This is the peer reviewed version of the following article: De Luca, G., Silva, J. M. N., & Modica, G. (2021). **A workflow based on Sentinel-1 SAR data and open-source algorithms for unsupervised burned area detection in Mediterranean ecosystems**. GIScience and Remote Sensing, 58(4), 516–541., which has been published in final doi https://doi.org/10.1080/15481603.2021.1907896.

GISCIENCE & REMOTE SENSING
2021, VOL. 58, NO. 4, 516–541
https://doi.org/10.1080/15481603.2021.1907896

Taylor & Francis
Taylor & Francis Group

A workflow based on Sentinel-1 SAR data and open-source algorithms for unsupervised burned area detection in Mediterranean ecosystems

Giandomenico De Luca[a], João M.N. Silva[b] and Giuseppe Modica[a]

[a]Dipartimento di Agraria, Università degli Studi Mediterranea di Reggio Calabria, Reggio Calabria, Italy; [b]Centro de Estudos Florestais, Instituto Superior de Agronomia, Universidade de Lisboa, Lisbon, Portugal

ABSTRACT

This paper investigates the capability of the free synthetic aperture radar (SAR) Sentinel-1 (S-1) C-band data for burned area mapping through unsupervised machine learning open-source processing solutions in the Mediterranean forest ecosystems. The study was carried out in two Mediterranean sites located in Portugal (PO) and Italy (IT). The entire processing workflow was developed in Python-based scripts. We analyzed two time-series covering about one month before and after the fire events and using both VH and VV polarizations for each study site. The speckle noise effects were reduced by performing a multitemporal filter and the backscatter time averages of pre- and post-fire datasets. The spectral contrast between changed and unchanged areas was enhanced by calculating two single-polarization radar indices: the radar burn difference (RBD) and the logarithmic radar burn ratio (LogRBR); and two temporal differences of dual-polarimetric indices: the delta modified radar vegetation index (ΔRVI) and the delta dual-polarization SAR vegetation index (ΔDPSVI), all exhibiting greater sensitivity to the backscatter changes. The scene's contrast was enhanced by extracting the Gray Level Co-occurrence Matrix (GLCM) textures (dissimilarity, entropy, correlation, mean, and variance). A principal component analysis (PCA) was applied for reducing the number of the GLCM image layers. The burned area was delineated through unsupervised classification using the *k*-means clustering algorithm. A suitable number of clusters (*k* value) were set using a silhouette score analysis. To assess the accuracy of the resulting detected burned areas, an official burned area map based on multispectral Sentinel-2 (S-2) was used for PO, while for IT, a reference map was produced from S-2 data, based on the normalized burned ratio difference (ΔNBR) index. Recall (*r*), precision (*p*) and the *F-score* accuracy metrics were calculated. Our approach reached the values of 0.805 (*p*), 0.801 (*r*) and 0.803 (*F-score*) for PO, and 0.851 (*p*), 0.856 (*r*) and 0.853 (*F-score*) for IT. These results confirm the suitability of our approach, based on SAR S-1 data, for burned area mapping in heterogeneous Mediterranean ecosystems. Moreover, the implemented workflow, completely based on free and open-source software and data, offers high adaptation flexibility, repeatability, and custom improvement.

KEYWORDS

SNAP-python (snappy) interface; *k*-means clustering; scikit-learn libraries; radar vegetation index (RVI); dual-polarization sar vegetation index (DPSVI)

A workflow based on Sentinel-1 SAR data and open-source algorithms for unsupervised burned area detection in Mediterranean ecosystems

Giandomenico DE LUCA[1*], João M.N. SILVA[2], Giuseppe MODICA[1]

[1]Dipartimento di Agraria, Università degli Studi Mediterranea di Reggio Calabria,
Località Feo di Vito, I-89122 Reggio Calabria, Italy
[2]Forest Research Centre, School of Agriculture, University of Lisbon, Tapada da Ajuda, 1349-017 Lisbon, Portugal

giandomenico.deluca@unirc.it

joaosilva@isa.ulisboa.pt

giuseppe.modica@unirc.it

[*]Corresponding author

Abstract

This paper investigates the capability of the free synthetic aperture radar (SAR) Sentinel-1 (S-1) C-band data for burned area mapping through unsupervised machine learning open-source processing solutions in the Mediterranean forest ecosystems. The study was carried out in two Mediterranean sites located in Portugal (PO) and Italy (IT). The entire processing workflow was developed in Python-based scripts. We analysed two time-series covering about one month before and after the fire events and using both VH and VV polarisations for each study site.

The speckle noise effects were reduced by performing a multitemporal filter and the backscatter time-averages of pre- and post-fire datasets. The spectral contrast between changed and unchanged areas was enhanced by calculating two single-polarisation radar indices: the radar burn difference (RBD) and the logarithmic radar burn ratio (LogRBR); and two temporal differences of dual-polarimetric indices: the delta modified radar vegetation index (ΔRVI) and the delta dual-polarization SAR vegetation index (ΔDPSVI), all exhibiting greater sensitivity to the backscatter changes. The scene's contrast was enhanced by extracting the Gray Level Co-occurrence Matrix (GLCM) textures (dissimilarity, entropy, correlation, mean, and variance). A principal component analysis (PCA) was applied for reducing the number of the GLCM image layers. The burned area was delineated through unsupervised classification using the *k*-means clustering algorithm. A suitable number of clusters (*k* value) was set using a silhouette score analysis. To assess the accuracy of the resulting detected burned areas, an official burned area map based on multispectral Sentinel-2 (S-2) was used for PO, while for IT, a reference map was produced from S-2 data, based on the normalised burned ratio difference (ΔNBR) index. Recall (*r*), precision (*p*) and the *F-score* accuracy metrics were calculated. Our approach reached the values of 0.805 (*p*), 0.801 (*r*) and 0.803 (*F-score*) for PO, and 0.851 (*p*), 0.856 (*r*) and 0.853 (*F-score*) for IT. These results confirm the suitability of our approach, based on SAR S-1 data, for burned area mapping in heterogeneous Mediterranean ecosystems. Moreover, the implemented workflow, completely based on free and open-source software and data, offers high adaptation flexibility, repeatability, and custom improvement.

Keywords: SNAP-Python (snappy) interface, *k*-means clustering, scikit-learn libraries, radar burn difference(RBD), radar vegetation index (RVI), dual-polarization SAR vegetation index (DPSVI)

1. Introduction

In the Mediterranean basin, although wildfires are an integral part of natural ecosystems, their extent and impacts have increased in the last decades, with thousands of hectares of forest areas burned every year and with significant economic damages and landscape changes (Chuvieco, 2009; Gitas et al., 2012; Lanorte et al., 2012; Ruiz-Ramos et al., 2018; San-Miguel-Ayanza et al., 2019). Moreover, fires are a long-term threat, contributing to soil erosion and habitat degradation, releasing greenhouse gases (GHGs), affecting air quality and global climate (Chuvieco, 2009; Gitas et al., 2012; Rosa et al., 2011).

Timely and accurate detection and quantification of burned areas are necessary to assess the damages, address the post-fire management, and implement medium and long-term territorial and landscape restoration strategies (Chuvieco et al., 2019; Lasaponara and Tucci, 2019; Pepe et al., 2018). In this context, satellite remote sensing provides reliable tools and techniques for detecting and quantifying the extension of burned areas (Chu and Guo, 2013; Chuvieco et al., 2019; Filipponi, 2019; Lizundia-Loiola et al., 2020; Otón et al., 2019), permitting rapid, cost-effective, temporally constant coverage and monitoring of large and less accessible regions (Pepe et al., 2018). Several studies concerning the localisation and mapping of fires' effects on vegetation were based on multispectral satellite data (Chuvieco et al., 2019; Filipponi, 2019; Imperatore et al., 2017; Lizundia-Loiola et al., 2020; Mouillot et al., 2014; Otón et al., 2019). These sensors are very efficient for the purpose due to their sensitivity in the visible, near and short infrared (NIR and SWIR) bands to changes in the state of vegetation and soil (Pereira et al., 1999; Chuvieco et al., 2019; Meng et al., 2017; Tanase et al., 2020; Miller et al., 2007; De Santis et al., 2009; Fornacca et al., 2018; Filipponi et al., 2018; Fernández-Manso et al., 2016). The optical spectral signature of the burned vegetation is unique and distinguishable from other disturbance factors or phenological changes in the short-term period after a fire. This is mainly due to the combined effect of diverse factors: the reduction of vegetation amount, the presence of coal and ash, changes in the moisture content and temperature, and the reflectance of soil. However, some of these elements tend to be attenuated in a few weeks or months after the fire event, in particular where the fire severity was low (Pereira et al., 1999; Smith et al., 2005; Inoue et al., 2019), leading to a spectral confusion of burned areas with other disturbances or low unburned albedo surfaces (e.g., dark soils, water surfaces, shaded regions, ploughed fields, timber harvesting) (Imperatore et al., 2017; Kurum, 2015; Pepe et al., 2018; Fraser et al., 2000; Stroppiana et al., 2015; Dijk et al., 2021; Rodman et al., 2021). Moreover, optical signal data are influenced by different phenological and physiological vegetation phases (e.g., seasonal senescence, leaf-off conditions), especially in the case of burned vegetation detection and monitoring at larger time intervals after the event

(Gallagher et al., 2020; Verbila et al., 2008; Fraser et al., 2000). In this context, the synthetic aperture radar (SAR) sensors are active systems that avoid some of these problems, proving to be an alternative or complementary data source for burned area detection and fire effects monitoring (Lehmann et al., 2015; Lasko, 2019; Kurum, 2015; Stroppiana et al., 2015; Tanase et al., 2011; Martinis et al., 2017; Chuvieco et al., 2019 Lasaponara et al., 2019). The response of the radar signal is affected by the ensemble of environmental variables (e.g., land cover, vegetation cover structure, moisture content, dielectric property of objects, size/shape and orientation of the scatterers in the canopy) and variables related directly to the sensor (e.g., polarisation, wavelength, orbit) or the local surface properties (e.g., topography, orientation, surface roughness, local incident angle) (Gimeno and San-Miguel-Ayanz, 2004; Hachani et al., 2019; Imperatore et al., 2017; Lapini et al., 2020; Santi et al., 2019, 2017; Tanase et al., 2011, 2020, 2010). SAR data are more sensitive to canopy structure than optical-based products (Martins et al., 2016). In detecting burned areas, SAR technology uses the variations in microwave backscatter caused by vegetation cover and soil structure and moisture content modifications, which implies a dielectric permittivity variation, thus providing an efficient system for discriminating events that cause changes in objects on the Earth's surface (Chuvieco et al., 2019; Donezar et al., 2019; Imperatore et al., 2017; Kurum, 2015; Pepe et al., 2018; Santi et al., 2017; Tanase et al., 2011, 2020, 2015, 2010; Zhou et al., 2019). Ruiz-Ramos et al. (2018) noted that, in dry conditions, the backscatter signal tended to decrease even after several weeks after the fire, indicating how degraded conditions can persist significantly after the event. This highlights the efficiency of SAR data in monitoring burned areas and justifying the need for timely interventions to counteract the ecosystem degradation and avoid desertification phenomena (Hill et al., 2008; De Luis et al., 2001; Chuvieco, 2009).

The variation of the backscattering signal due to the fire's effect on reducing the crown structure can be of different evidence depending on the polarisation. Generally, cross-polarised signals (vertical-horizontal, VH, and horizontal-vertical, HV) show a decrease in the backscatter response due to the consequent reduced volumetric dispersion contribution. Conversely, the change in the co-polarised backscatter coefficients (vertical-vertical, VV or horizontal-horizontal, HH) can be attributed to higher soil exposure (Imperatore et al., 2017). Due to this different interaction with the various aspects of the effects of fire on the environment, both types of polarisation can be decisive in detecting burnt forest areas (Tanase et al., 2014). For other purposes, this aspect is already employed in vegetation monitoring through the use of radar-based polarimetric indices in which both types of polarisation are used depending on the type of product and the SAR sensor used (Gururaj et al., 2019; Mandal et al., 2020; Nasirzadehdizaji et al., 2019). The radar vegetation index (RVI) (Kim et al., 2009), full- or dual-polarimetric, is a well-established SAR index (Szigarski et al., 2018) and generally used in studies related to vegetation biomass growth (Kim et al., 2014), in the LAI (leaf area index) estimation (Pipia et al., 2019) or in the estimation of the water content of plants and soil (Kim et al., 2012; Trudel et al., 2015). Kim et al. (2012) demonstrated a high correlation between L-band RVI and other optical vegetation indices. The dual-polarisation SAR vegetation index (DPSVI) (Periasamy, 2018) also returned positive results for the study of plant biomass, demonstrating a good correlation with the normalised difference vegetation index (NDVI). However, single-polarisation indices were also used with excellent results to map the burnt areas or fire severity (Lasaponara and Tucci, 2019; Tanase et al., 2015). More generally, most of the studies explored the backscattering behaviour after a fire in the Mediterranean environment (Imperatore et al., 2017; Kurum, 2015; Minchella et al., 2009; Tanase et al., 2015), but few of these have focused on the ability of SAR data to map the burned areas by measuring their accuracy with analytical methods (Belenguer-Plomer et al., 2019; Gimeno et al., 2004; Gimeno and San-Miguel-Ayanz, 2004; Lasaponara and Tucci, 2019; Martinis et al., 2017; Zhang et al., 2019).

Several space missions provide satellite constellations operating SAR imaging dedicated to environment observation useful for fire monitoring purposes (Chuvieco, 2009; Chuvieco et al., 2019; Mouillot et al., 2014). Copernicus missions by the European Space Agency (ESA) provides free high spatial and temporal resolution SAR (S-1) and multispectral (S-2) data (ESA Sentinel Homepage, 2020). The S-1 constellation comprises two polar-orbiting satellites (S-1A and S-1B) performing C-band (from 3.75 cm to 7.5 cm wavelength) radar imaging. The good spatial and temporal resolutions added to the free distribution make the Sentinel mission particularly suitable for risk monitoring and rapid mapping (Martinis et al., 2017). Several studies have demonstrated the sensitivity of the C-band to changes in the vegetation and environment affected by fire (Imperatore et al., 2017; Kurum, 2015; Tanase et al., 2020, 2010).

One of the strengths of the S-1 and S-2 data is their high spatial and temporal resolution. The spatial resolution has a considerable effect on the detection of burnt areas and their subsequent monitoring, lowering the omission errors typical of the coarser resolution data in detecting the smallest areas and improving spectral discrimination (Verhegghen et al., 2016; Boschetti et al. 2015; Stroppiana et al., 2015; Belenguer-Plomer et al. 2019; Mouillot et al., 2014). The advantages become more evident when the acquisition revisit time of these products is a few days, allowing the monitoring of temporal trends at an appropriate temporal scale (Boschetti et al., 2015; Verhegghen et al., 2016; Gitas et al., al., 2012; Tanase et al., 2020).

Furthermore, ESA itself distributes the Sentinel application platform (SNAP) (ESA SNAP Homepage, 2020), a free and open-source software platform containing the toolboxes necessary for pre-processing and processing Sentinel data. The SNAP toolboxes, initially Java-based, can also be accessed from the Python programming language (The Python Language Reference, 2020), one of the most popular languages for remote sensing and scientific analysis, widely used in both operational and scientific domains (Hao and Ho, 2019), through the ESA SNAP-Python (snappy) interface (ESA SNAP Cookbook, 2020).

The present work aimed to develop a semi-automatic procedure for mapping burned areas in Mediterranean regions using SAR S-1 data and based on the *k*-means clustering algorithm for an unsupervised image classification approach. Therefore, supporting the state-of-the-art of SAR-based burned area mapping.

The *k*-means is one of the most straightforward iterative clustering algorithms, widely used in data mining and pattern recognition purposes (Dhanachandra et al., 2015; Nagpal et al., 2013, Jain 2010). One of the main difficulties for the *k*-means cluster analysis is to set the more suitable number of clusters (*k* value) in the initialisation phase. Among the different approaches proposed in the literature (Kodinariya and Makwana, 2013), in our approach, we used the silhouette score (Rousseeuw, 1987) to set the value of the *k* parameter, which statistically measures the average separation distance (dissimilarity) between points within neighbouring clusters. The entire processing workflow (Figure 2), excluding accuracy assessment, was developed in Python-based open-source libraries and scripts, implementing ESA-snappy for image pre-processing and Scikit-learn (Pedregosa et al., 2011) processing and classification. It consists of the following fundamental steps: 1) speckle-noise reduction by calculating the backscatter time average of pre- and post-fire datasets and then applying a multitemporal filter; 2) calculation of the radar burn difference (RBD) and the logarithmic radar burn ratio (LogRBR) single-polarisation indices and the dual-polarimetric S-1 indices (ΔRVI and ΔDPSVI) in order to emphasise the areas of change; 3) gray-level co-occurrence matrix (GLCM) texture features extraction; 4) data reduction using the principal components analysis (PCA) transformation; 5) silhouette score analysis in order to set the *k* parameter value; 6) unsupervised classification using the *k*-means clustering algorithm.

To confirm the method's applicability, it was tested on two scenes representing two Mediterranean forest environments located in two different countries (Italy and Portugal). The validation of the classification maps was performed by comparison with reference maps based on S-2 Multispectral images and calculating accuracy metrics (recall, *r*, precision, *p*, and the *F-score*).

2. Materials and Methods

2.1 Study sites

The implemented methodology was tested in two Mediterranean areas of Southern Europe (Figure 1). The first is located in Algarve, the southernmost region of Portugal (37° 18'N; 08° 30'W), a forest area in the Serra de Monchique mountain range (study site PO). The second is located in the central area of Sicily (South of Italy, 37° 43'N; 14° 39'E), the "Rossomanno-Grottascura-Bellia" regional nature reserve (study site IT). The extent of the two study sites was obtained manually based on the overlapping area of the tiles of the various orbits of S-1. The two study sites extend to approximately 2550 km^2 (IT) and 3600 km^2 (PO). The sites are located at the same latitude and present very similar and comparable typical Mediterranean vegetation contexts. Most parts of the two study areas were dominated by genus Eucalyptus species (Eucalyptus spp.) and typical Mediterranean conifers (Pinus spp.), deriving mainly from artificial planting during the end of the 19th century and the 20th century. However, both areas study sites are also covered by areas with dense typical Mediterranean forest vegetation of secondary broad-leaved (ex. *Quercus* spp.) and coniferous trees, interspersed with sclerophyllous shrublands (Camerano et al., 2011; San-Miguel-Ayanz et al., 2016; Sistema Nacional de Informação Geográfica (SNIG), 2020). The PO study site also includes agricultural areas and pastures.

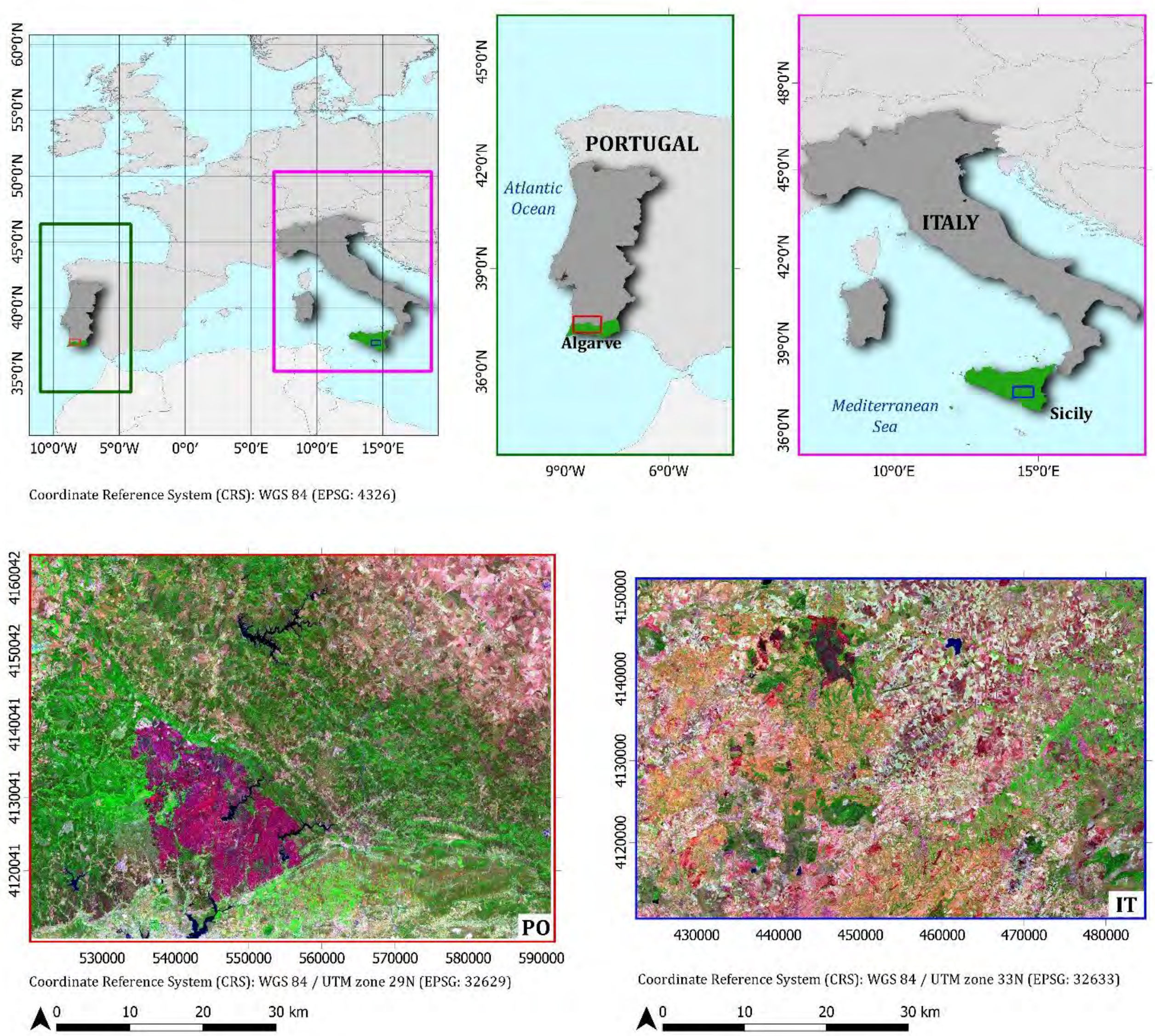


Figure 1. Study sites: in the top, location of the study sites in Europe and in the respective countries; in the bottom, the overviews of the two study sites (post-fire Sentinel-2 images, SWIR-NIR-Green false-colour composite) where the burned areas are clearly visible (the dark-purple area in PO; the darker area in IT).

The events occurred in August from the 3rd to the 10th, 2018, in the PO study site, covering 268.9 km^2, while, in the IT study site, the fire occurred on August 6th, 2017, covering an area of 38.51 km^2. Regarding the Sicilian natural reserve, the fire also affected neighbouring and similar forest areas outside its administrative boundaries. In the PO study site, fire affected the vegetation in a heterogeneous way at the spatial level, altering or removing the structure at various degrees, with a predominant crown fire occurrence, leaving residues of burns on the ground (ash and coal). In some places, where the severity was higher, the bare soil was exposed (Oom et al., 2018).

2.2 Dataset

2.2.1 Sentinel-1 dataset

The Sentinel-1A/B high-resolution ground range detected (GRDH) dual-polarised (VV and VH polarisations) time series, acquired in interferometric wide (IW) mode, was searched through the Copernicus Open Access Hub (2020). The spatial resolution of the product is 20 m x 22 m (ground range x azimuth), with a pixel spacing of 10 m x10 m (ground range x azimuth) on the image, corresponding to the mid-range value at mid-orbit altitude averaged over all sub-swaths (ESA Sentinel-1 User Guide, 2016). The bulk downloading process was carried out using the aria2 command-line downloader (aria2 download utility Homepage, 2020), allowing to automate and speed up the acquisition of huge datasets. In total, we acquired two S-1 image datasets, one for each of the two study sites, respectively. The images were acquired to cover a time frame of about a month before and after the event date during the summer fire season (July-September), taking into account the need for the absence of rain that could affect the backscatter signal. For the PO study site, the dataset was formed by eight images for the pre-fire period and five images for the post-fire period; for the IT study site, the pre-fire and the post-fire images were nine and five respectively (Table 1).

Table 1. Sentinel-1 dataset characteristics. The red line separates the images acquired before and after the fire occurrence.

PO Study site				**IT Study site**			
Mission	*Orbit*	*Product*	*Sensing Date - Hour*	*Mission*	*Orbit*	*Product*	*Sensing Date - Hour*
S-1B	Ascending	IW Level-1 GRDH	2018/07/01 – 18:34	S-1A	Ascending	IW Level-1 GRDH	2017/07/05 - 17:04
S-1A	Ascending		2018/07/07 – 18:35	S-1A	Descending		2017/07/06 - 05:04
S-1A	Descending		2018/07/08 – 06:35	S-1B	Descending		2017/07/12 – 05:04
S-1B	Ascending		2018/07/13 – 18:34	S-1A	Ascending		2017/07/17 - 17:04
S-1A	Ascending		2018/07/19 – 18:35	S-1B	Ascending		2017/07/18 - 16:55
S-1A	Descending		2018/07/20 – 06:35	S-1A	Ascending		2017/07/24 - 16:56
S-1B	Ascending		2018/07/25 – 18:34	S-1A	Ascending		2017/07/29 - 17:04
S-1A	Ascending		2018/07/31 – 18:35	S-1B	Ascending		2017/07/30 - 16:55
S-1A	Ascending		2018/08/12 – 18:35	S-1B	Descending		2017/08/05 – 05:04
S-1B	Ascending		2018/08/18 – 18:34	S-1A	Ascending		2017/08/17 - 16:56
S-1A	Ascending		2018/08/24 – 18:35	S-1A	Ascending		2017/08/22 - 17:04
S-1A	Descending		2018/08/25 – 06:35	S-1A	Descending		2017/08/23 - 05:04
S-1B	Ascending		2018/08/30 – 18:34	S-1B	Ascending		2017/08/23 - 16:55
				S-1A	Ascending		2017/08/29 - 16:56

2.2.2 Reference data

As reference data for the PO study site, the burned area perimeters provided by Instituto de Conservação da Natureza e das Florestas (ICNF) based on S-2 satellite imagery (SIG-ICNF, 2021) were adopted. The minimum extent of the mapped fires is 0.5 km^2. Due to the insufficient quality of the official data (see the Supplementary material), we downloaded two Sentinel-2B Level-1C images, acquired one before (sensing date: 2017/08/01, 09:50) and one after the fire (sensing date: 2017/08/11, 09:50), respectively, in order produce the reference map for the IT event.

The two images were pre-processed (Section 2.4) and the normalised burn ratio (NBR) (Eq. 1) for the pre- and post-fire S-2 data and, consequently, their temporal difference represented by ΔNBR index (Eq. 2) (Key et al., 2006) was calculated:

$$NBR_{zj} = (NIR_{zj} - SWIR_{zy}) / (NIR_{zj} + SWIR_{zj}) = (B8A_{zj} - B12_{zj}) / (B8A_{zj} + B12_{zj}) \quad (1)$$

$$\Delta NBR = NBR_{prefire} - NBR_{postfire} \quad (2)$$

where *zj* represents a fire-related time period (pre- or post-fire); NIR is the near infra-red band that in this case was represented by the band B8A (865 nm) of S-2 data; SWIR is the short-wave infra-red band represented by the band B12 (2190 nm) of S-2 data. These two bands are very sensitive to burned vegetation (Lanorte et al., 2012). For this reason, this index is generally used as a reference layer since it allows to better identify the perimeter of the burned areas than other methods (Ban et al., 2020; Donezar et al., 2018; Tanase et al., 2015; Zhang et al., 2019; Kurum, 2015; Tanase et al., 2010), in the absence of good quality official data. The shapefile used as reference was obtained by converting the binary map composed of pixels with ΔNBR values greater than 0.1 (conventional burned / not-burned threshold (Keeley et al., 2009). Moreover, the interpretation was visually strengthened and guided by using the RGB false-colour combination (SWIR-NIR-Red). The IT reference shapefile was filtered, deleting all the polygons with an area less or equal to 0.05 km^2 to reduce redundancy and make the data consistent with the PO.

2.3 Processing libraries

The S-1 data pre-processing was carried out using the Sentinel-1 Toolbox implemented in ESA-SNAP v.7.0.4 (ESA SNAP Homepage, 2020) and executed through Snappy (ESA SNAP Cookbook, 2020), the SNAP-Python interface which enables accessing and managing the SNAP Java application programming interface (API) from Python. The application script was built on Python v.3.6.8 (The Python Language Reference, 2020), a version compatible with the Snappy interface. The image processing and classification were implemented in Scikit-learn v.0.23.1 (Pedregosa et al., 2011, Scikit-learn Homepage, 2020), an open-source Python-based library that provides a collection of different data-processing modules concerning machine learning analysis and modelling (Hao and Ho, 2019; Pedregosa et al., 2011). This library contains all the processing modules used in this study: the MinMaxScaler module (Section 2.5.1), the sklearn.decomposition.PCA module (Section 2.5.2), the sklearn.metrics.silhouette_score module (Section 2.5.3) and the sklearn.cluster.KMeans module (Section 2.5.4).Image pre-processing and layers creation

The S-1 data pre-processing steps (Figure 2), carried out for both the two time-series datasets, started by applying the auto-downloaded orbit file, followed by thermal noise removal. The implemented process code is available as a repository on the GitHub platform. The web-link is in the Websites Section (GitHub Code repository, 2021).

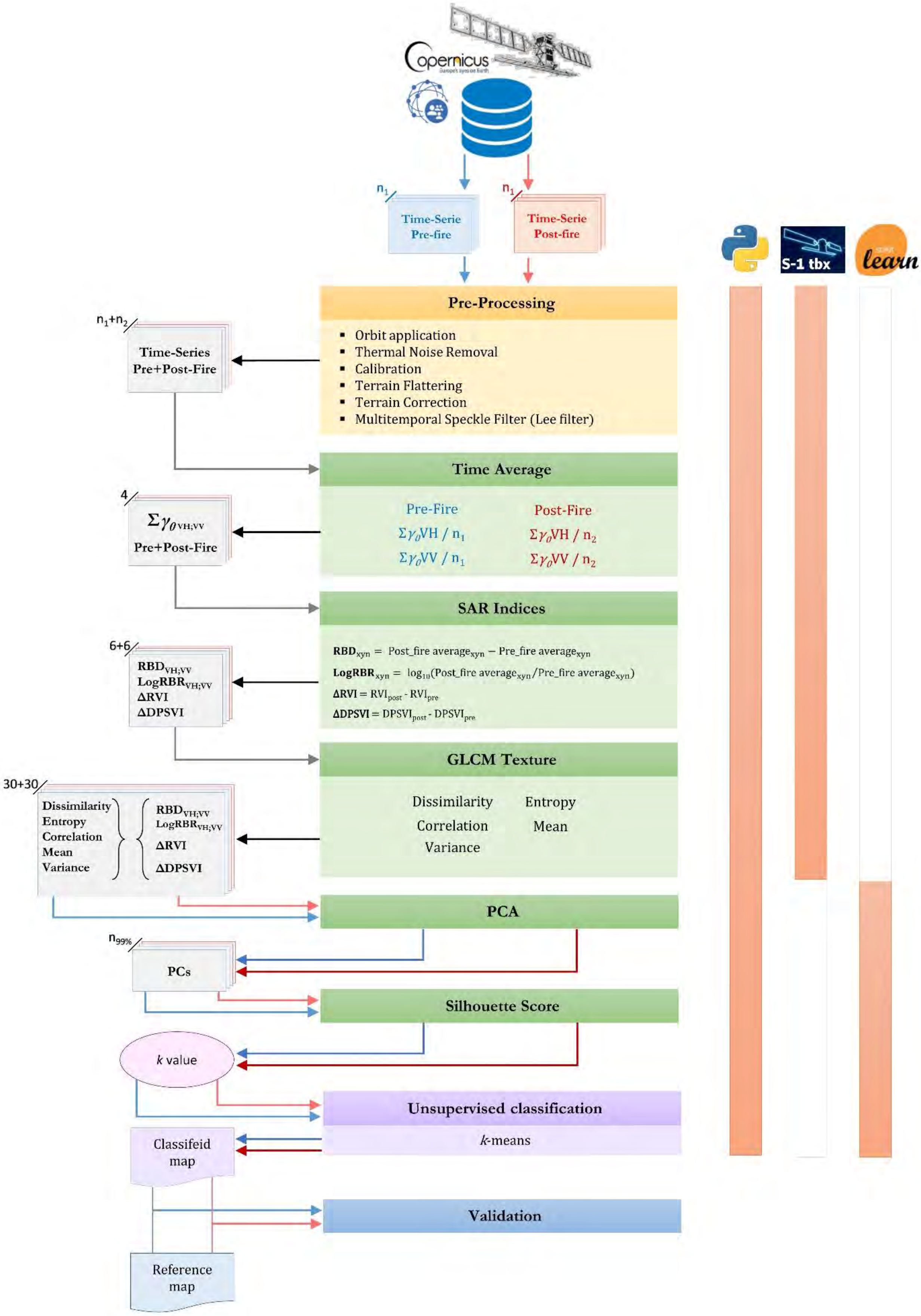


Figure 2. The workflow of the implemented approach.

The images were then radiometric calibrated and converted to beta (β_0) noughts backscatter standard conventions. Due to the rough terrain topography of both study areas and consequently the presence of geometric and radiometric distortions, a radiometric terrain flattening (RTC processing) and a terrain correction were performed using a digital elevation model (DEM) obtained from the shuttle radar topography mission (SRTM) (Farr et al., 2007; Small, 2011), presenting a spatial-sampling of 1 arc-second. The bilinear interpolation resampling method was used for both DEM and output image resampling. During the RTC processing, the images were converted from β_0 to gamma (γ_0) nought automatically. In contrast, in the terrain correction step, the images were projected to WGS84/UTM zone 29N and 33N for the PO study site and IT study site, respectively.

For each study site dataset, an image stack was made using the Create Stack Operator of Sentinel-1 Toolbox. The product geolocation was used as an initial offset method, and the extent of the master image was adopted on the slave images without resampling. The optimal master image for each dataset was chosen automatically by the tool.A multitemporal speckle Lee filter (Quegan et al., 2000; Santoso et al., 2015) of 15x15 pixel window size was carried out to apply a first reduction of the radar speckle noise. Subsequently, the speckle reduction was improved by calculating the backscatter time average (Lasaponara and Tucci, 2019), separately for the images before and after the fire, for each polarisation (VH and VV). Following the implemented pre-processing phase, four layers are obtained:

1. Pre-fire time average VH;
2. Pre-fire time average VV;
3. Post-fire time average VH;
4. Post-fire time average VV.

For both study sites, these individual layers were used to compute two single-polarisation radar indices for change detection: the RBD (Eq. 3) (the difference between pre- and post-fire backscattered time average for each polarisation) and the LogRBR (Eq. 4) (log-scaled ratio of the backscattering coefficients between pre- to post-fire for each polarisation). This latter index is derived from the radar burn ratio (RBR) (Tanase et al., 2015) used in change detection or fire severity detection (Lasaponara and Tucci, 2019; Tanase et al., 2015), scaled to logarithmic in order to optimise the noise distribution (Dekker, 1998).

The equations of the two indices are:

$$RBD_{xy} = \text{Post-fire TimeAverage}_{xy} - \text{Pre-fire TimeAverage}_{xy} \tag{3}$$

$$LogRBR_{xy} = \log_{10}(\text{Post-fire TimeAverage}_{xy} / \text{Pre-fire TimeAverage}_{xy}) \tag{4}$$

where xy represents a specific polarization (VV or VH).

Besides, two dual-polarimetric radar vegetation indices, the radar vegetation index (RVI) (Eq. 5) proposed by (Kim and Van Zyl, 2009) and modified for the S-1 dual-polarized SAR data (Gururaj et al., 2019; Mandal et al., 2020; Nasirzadehdizaji et al., 2019), and the dual-polarisation SAR vegetation index (DPSVI) proposed by (Periasamy, 2018) (Eq. 6) were computed for pre- and post-fire datasets, respectively:

$$RVI_{zj} = 4 \cdot {}_{TimeAverage}VH / ({}_{TimeAverage}VV + {}_{TimeAverage}VH) \quad (5)$$

$$DPSVI_{zj} = ({}_{TimeAverage}VV + {}_{TimeAverage}VH) / {}_{TimeAverage}VV) \quad (6)$$

where zj represents a fire-related time period (pre- or post-fire).

From these two vegetation indices, the relative temporal difference was calculated (ΔRVI and ΔDPSVI) (Eq. 7-8):

$$\Delta RVI = RVI_{post} - RVI_{pre} \quad (7)$$

$$\Delta DPSVI = DPSVI_{post} - DPSVI_{pre} \quad (8)$$

For the RBD_{VH}, $RBD_{VV,}$ $LogRBR_{VH}$, $LogRBR_{VV,}$ ΔRVI and ΔDPSVI index layers, five GLCM (Grey Level Co-occurrence Matrix) texture features (Haralick, 1979; Haralick et al., 1973) were computed for each of the two study sites (Tab 2) fixing the size of the search window to 11x11 pixels. The five GLCM textures were computed to exhibit a more marked contrast between changed and unchanged areas, adding extra spatial information to support classification accuracy performance (Hall-Beyer, 2017; Li et al., 2014).

The GLCM process originated a dataset consisting of 30 layers for each study-site, which constituted the input data for the next processing workflow step.

Table 2. Name, group, and equation of used GLCM (Grey Level Co-occurrence Matrix) texture measures. $P_{i,j}$ is the probability of values i and j occurring in adjacent pixels in the original image within the window defining the neighbourhood. i and j are the labels of the columns and rows (respectively) of the GLCM: i refers to the digital number value of a target pixel; j is the digital number value of its immediate neighbour. μ is mean and σ the standard deviation.

GLCM Features	Group	Equation
Dissimilarity	Contrast	$\sum_{i,j=0}^{N-1} P_{i,j}\,\lvert i-j\rvert$
Entropy	Orderliness	$\sum_{i,j=0}^{N-1} -\ln(P_{i,j})P_{i,j}$
Correlation	Statistics	$\sum_{i,j=0}^{N-1} P_{i,j}\left[\frac{(i-\mu_i)(i-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}}\right]$
Mean		$\mu_i = \sum_{i,j=0}^{N-1} i(P_{i,j})\,;\ \mu_j = \sum_{i,j=0}^{N-1} j(P_{i,j})$
Variance		$\sigma_i^2 = \sum_{i,j=0}^{N-1} P_{i,j}(i-\mu_i)^2\,;\ \sigma_j^2 = \sum_{i,j=0}^{N-1} P_{i,j}(j-\mu_j)^2$

The S-2 images downloaded to generate the IT reference data were pre-processed using the Sentinel-2 Toolbox. These were first resampled to 10 m × 10 m pixel size using the band B4 (Red; 664.6 nm) as reference source size and the bilinear interpolation as an upsampling method. Subsequently, the images were reprojected and clipped on the same area of the correspondent S-1 data. The Level-2A products (Bottom-of-Atmosphere) were generated using Sen2Cor v2.8 processor (ESA sen2cor Homepage, 2020).

2.4 Data preparation

2.4.1 Data normalisation

The data normalisation in the same continuous scale range [0-1] was carried out for all the S-1 single layers (Eq. 9). This operation converts the original values of the input data into the new range through rescaling. This step aimed to equalise the input features, reducing the influence of differences in their intervals, making them comparable in numerical values and optimising the learning algorithm process (Angelov and Gu, 2019, Subasi, 2020). The normalisation was carried out using the specific MinMaxScaler module contained in scikit-learn, given by:

$$x_{norm} = \frac{x-x_{min}}{x_{max}-x_{min}} \qquad (9)$$

where x_{norm} is the new normalised value, x is the value to be normalised, x_{min} and x_{max} are the smallest and the highest value of the data (feature range).

2.4.2 Data reduction: Principal Component Analysis (PCA) transformation

Considering the high number of input data layers, a principal component analysis (PCA) was performed to reduce the dimension of the dataset and select the optimum layer subset without losing the essential information (total variance) for image classification (Gimeno et al., 2004; Richards, 2013). The PCA module provides a linear dimensionality reduction based on singular value decomposition (SVD) in order to project the data in a lower-dimensional space (eigenspace) and derive the new principal components (PCs) representing the directions of maximum variance of the eigenspace (Subasi, 2020). In this study, the first transformed PCs that explained a high enough cumulative variance (greater than or equal to 99%) were considered the optimal reduced representation of the original dataset and used as input in the classification process.

2.5 Image Classification

2.5.1 Classification algorithm (*k*-means algorithm)

The burned areas' classification was carried out using the popular *k*-means algorithm, a data clustering method introduced by James MacQueen (1967). It is known as one of the simplest and fastest unsupervised machine learning algorithms (Dhanachandra et al., 2015; Nagpal et al., 2013; Soni and Patel, 2017), widely used in remote sensing applications (Celik, 2009; Dhanachandra et al., 2015; Li et al., 2014; Phiri and Morgenroth, 2017; Rehman et al., 2019; Senthilnath et al., 2017). Given a dataset, the algorithm is based on the grouping (clustering) of the pixels with homogeneous characteristics in a predefined number (k) of clusters. The homogeneity of the pixels is defined by the minimum distance between their value and the single cluster's centroid. The algorithm's initialisation starts with a first random definition of the k centroids, optimised by the *k-means++* method (Arthur and Vassilvitskii, 2007), and is based on the weighted distribution probability for the definition of the centroids. Then, it proceeds with the first assignment of each pixel to the nearest centroid, in terms of values' Euclidean distance, and therefore with the first k clusters' generation. After the first initialisation of the k centroids, each of them is recalculated many times over so that the dataset belonging to a cluster can be reassigned to the new cluster, obtaining the most appropriate assignment of each pixel to the clusters. This process is repeated iteratively until the centroids' arrangement ceases to change, the tolerance or error value is satisfied, or until the maximum number of defined iterations is reached (Dhanachandra et al., 2015; Soni and Patel, 2017). The centroid of a cluster is the point to which the sum of distances from all the pixels in that cluster is minimised. Therefore, the *k*-means could be defined as an iterative algorithm that minimises the value of the sum of squared errors (SSE) of distances from each object to its cluster centroid (Dhanachandra et al., 2015). The *k*-means algorithm used in this work was based on a combination with the expectation-maximisation (EM) model (Dempster et al., 1977).

2.5.2 Definition of a suitable number of clusters using the Silhouette Score

One of the main issues at initialising a clustering algorithm is setting the optimal number of clusters (*k* parameter) (Kodinariya and Makwana, 2013). To solve this issue, we used the silhouette score approach (Rousseeuw, 1987), which is based on the separation distance between clusters, according to the following formula (equation 10):

$$\text{Silhouette Score} = (b_i - a_i) / \max(a_i, b_i) \qquad (10)$$

where *i* is the value of a single-pixel contained in a cluster, *a* is the average distance (dissimilarity) between *i* and all other objects of the same cluster, and *b* is the average distance between *i* and the nearest cluster of which *i* is not a part (Rousseeuw, 1987). This coefficient measures how close each point in a cluster is to the neighbouring clusters' points for a given number of clusters. The computation of its average results is a simple method to address *k* value (Rousseeuw, 1987). We calculated the mean of the silhouette score for different *k* values (*k*-space, from 2 to 20) using the "relative" module provided in scikit-learn. To save on computation time, the calculation was performed on a sample of 100,000 points randomly distributed over the entire area of the two datasets. The score value can vary in a range from 1 (maximum separation: well clustered, best *k*-value) to -1 (minimum separation: misclassified, worst *k*-value).

2.5.3 Classification application and post-process enhancement

For each of the two transformed and reduced datasets, an unsupervised classification was carried out using the *k*-means algorithm. The number of clusters (*k* parameter) was set according to the silhouette score analysis result, while the default number of iteration (300) was left.

In order to identify the classes representing the real burned areas, the mean value of each radar index for each class was computed and plotted.

Despite the noise reduction operations, the SAR data still presents some outliers, which are persistent due to the signal's intrinsic characteristics. Moreover, since we used several images for each dataset covering a time-frame of about one month before and one month after the fire event, different surface-changes could have occurred (small fires, agronomic operations, etc.), leading to an erroneous assessment of commission errors. Therefore, following the raster data's vectorisation, pre and post-fire scenes were filtered, eliminating clusters covering an area less or equal to 0.05 km^2 (minimum mapping unit of reference data; see Section 2.2.2 Reference data).

2.6 Accuracy Assessment

The resulting classification maps were compared to the respective reference burned areas to assess their accuracy. The accuracy analysis regarded only those classes corresponding to the actual burned area, excluding the other classes. We chose these classes by observing the distribution of the average value of each of the six radar indices for each class. The classes that did not correspond to the burned area were aggregated together as "unburned class". Both the classified and the reference images were vectorised to facilitate their analytical comparison. Therefore, after their superimposing, each classified pixel was labelled into one of the following categories (pixel-based accuracy assessment) (Goutte and Gaussier, 2005; Modica et al., 2020; Shufelt, 1999; Sokolova et al., 2006):

- True Positive (TP): when a pixel classified as burned corresponded to burned class in the reference data (pixel correctly classified).
- False Negative (FN): when a pixel representing burned in the reference data was classified as not-burned (pixel not detected).
- False Positive (FP): when a pixel classified as burned did not correspond to burned class in the reference data (pixel erroneously detected).

After counting the number of pixels belonging to one of the three categories for each image, the recall (*r*), Precision (*p*) and *F-score* accuracy metrics were calculated (Equations. 11-13) (Goutte and Gaussier, 2005; Shufelt, 1999; Sokolova et al., 2006; Sokolova and Lapalme, 2009):

$$r = \frac{|TP|}{|TP+FN|} \qquad (11)$$

$$p = \frac{|TP|}{|TP+FP|} \qquad (12)$$

$$Fscore = 2 \cdot \frac{r \cdot p}{r+p} \qquad (13)$$

where *r* and *p* are functions of omission and commission errors. Their opposites, 1-*r* and 1-*p*, indicate the omission and commission errors, respectively. The *F*-score measures the overall accuracy using the harmonic mean of commission and omission errors. The *r*, *p*, and *F* can be in a range from 0 (total misclassification) to 1 (perfect classification) (Goutte and Gaussier, 2005; Modica et al., 2020; Sokolova and Lapalme, 2009).

3. Experimental Results

3.1 Data preparation

To detect burned areas, the radiometric changes that occurred after the fire had to be highlighted. For this reason, radar vegetation indices were calculated, of which two were single-polarisation (RBD and LogRBR) and two dual-polarimetric (RVI and DPSVI). Unlike the RBD and LogRBR indices that already express temporal differences, the respective ΔRVI and ΔDPSVI indices had to be derived from the original RVI and DPSVI. The RBD, LogRBR, ΔRVI and ΔDPSVI, used as the input layer for successive GLCM computation step, are shown in Figures 3 and 4 for the PO, and IT study sites, respectively.

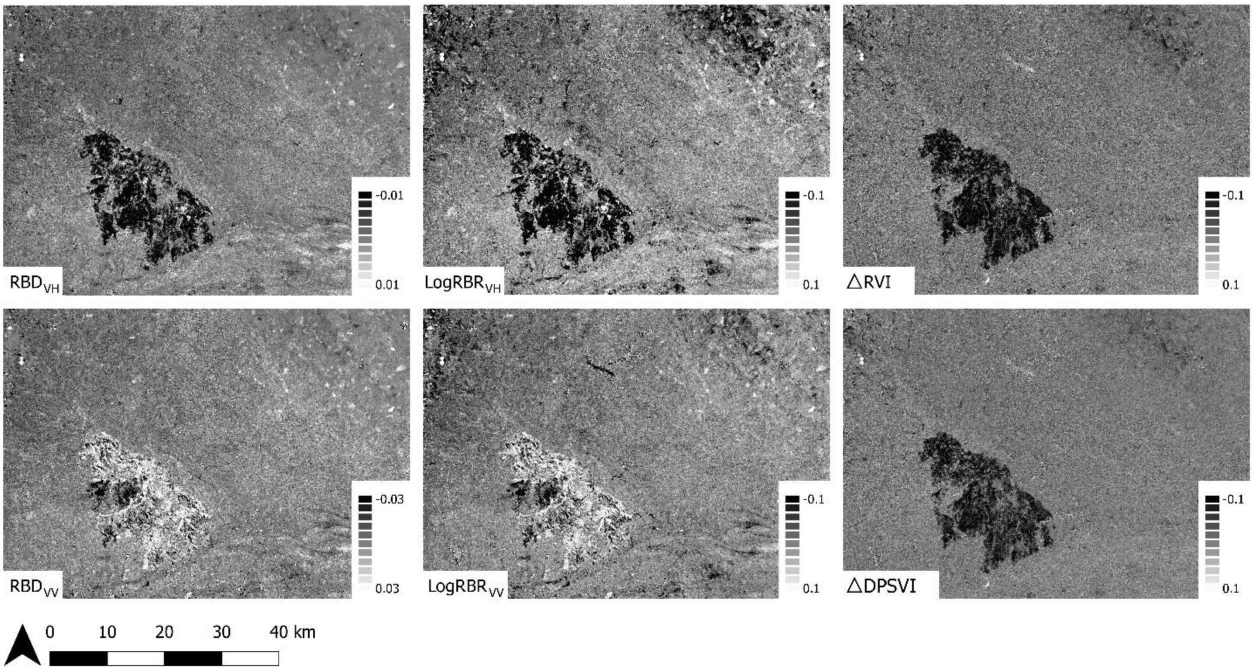


Figure 3. The S-1 indices (RBD$_{VH}$, LogRBR$_{VH}$, ΔRVI, RBD$_{VV}$, LogRBR$_{VV}$, and ΔDPSVI) were obtained in the PO dataset pre-processing steps. For each of these indices, the GLCM (Grey Level Co-occurrence Matrix) texture features were calculated.

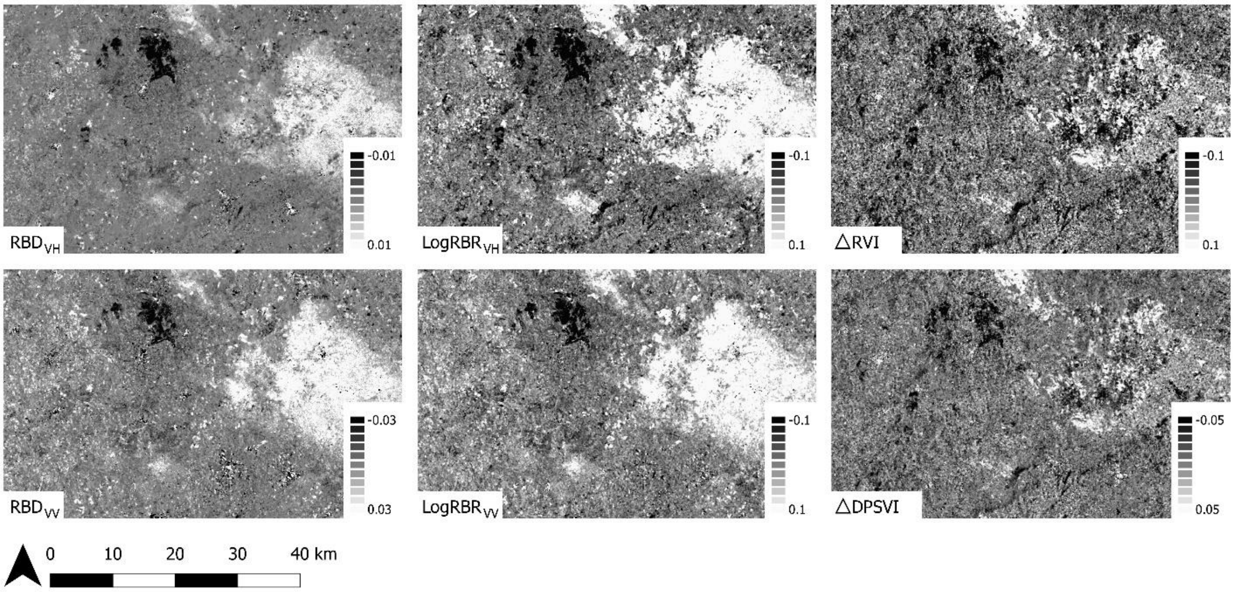


Figure 4. The S-1 indices (RBD$_{VH}$, LogRBR$_{VH}$, ΔRVI, RBD$_{VV}$, LogRBR$_{VV}$ and ΔDPSVI) were obtained in the IT dataset pre-processing steps. For each of these indices, the GLCM (Grey Level Co-occurrence Matrix) texture features were calculated.

3.2 PCA transformation

The PCA was performed on the entire dataset to reduce their dimension. The cumulative variance explained by the PCs is reported in Figures 5 (PO) and 6 (IT). As shown, the PO dataset reached the threshold (0.99) at the 9th PC, while the IT dataset expressed the same cumulative variance value at the 13th PC. These PCs, which for each dataset have reached the threshold and are represented by transformed images, have been chosen as input layers in the subsequent related processes.

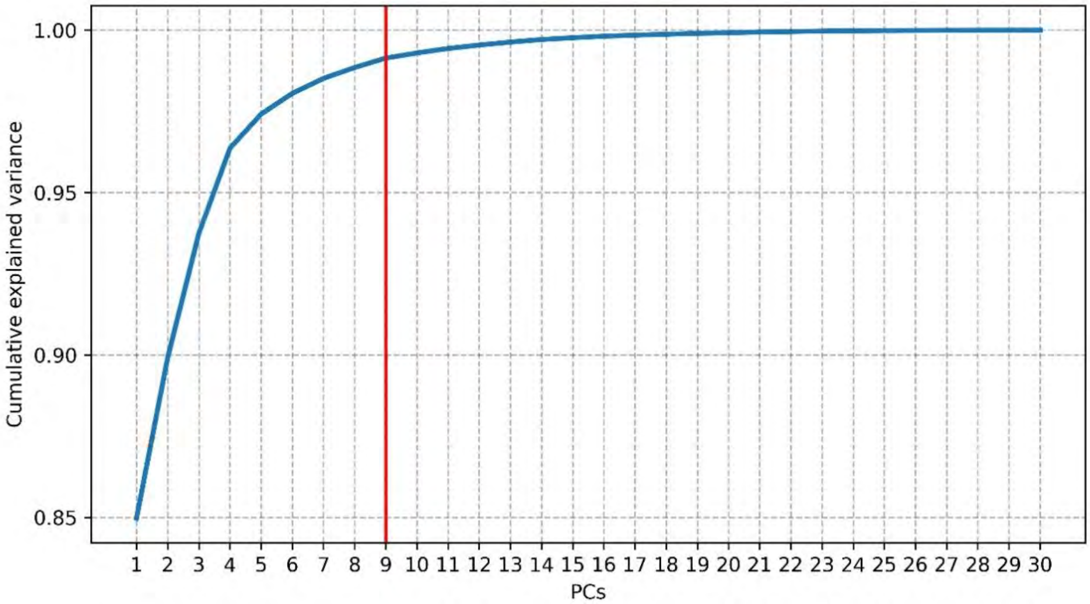


Figure 5. The cumulative variance explained by the principal components (PCs) for the PO study site. The red line identifies the first PCs that reached a cumulative variance of 0.99.

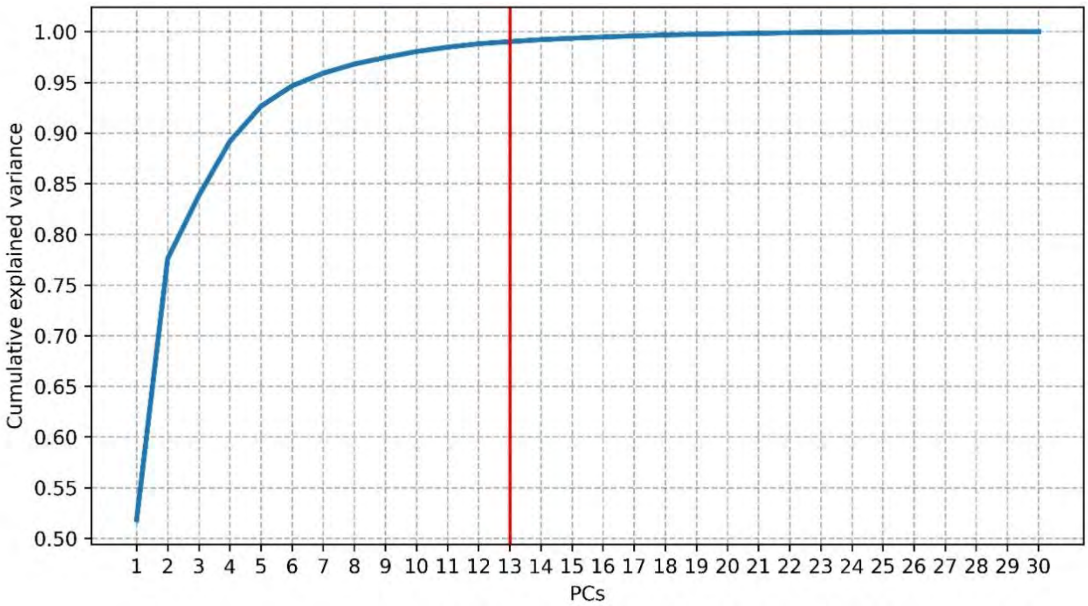


Figure 6. The cumulative variance explained by the principal components (PCs) for the IT study site. The red line identifies the first PCs that reached a cumulative variance of 0.99.

3.3 Silhouette score

Figures 7 (PO) and 8 (IT) show the trend of the averaged silhouette score calculated on relative PCA outputs for a k-space ranging from 2 to 20 clusters and for a sample of 100,000 random points. The highest values resulted from lower k-values, with the maximum value described by $k = 2$ for both datasets. The next highest value was found when $k = 7$ in both datasets with a Silhouette score of 0.166 and 0.191 for PO and IT, respectively.

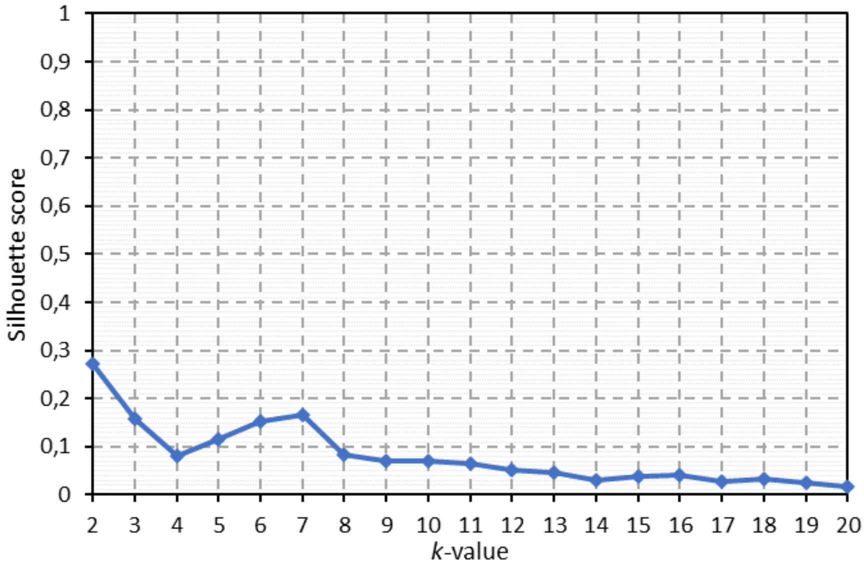


Figure 7. Silhouette score values, for the PO dataset, for a *k*-space range (*k* values) between 2 and 20.

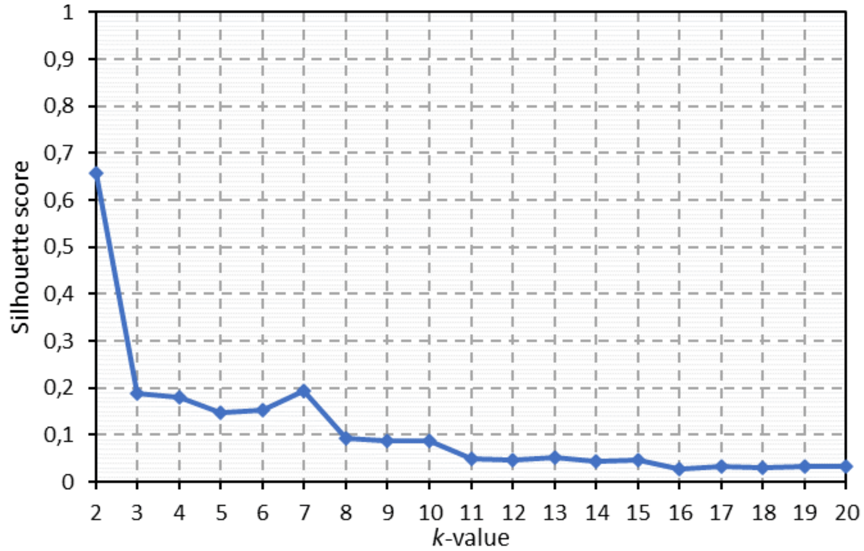


Figure 8. Silhouette score values, for the IT dataset, for a *k*-space range (*k* values) between 2 and 20.

3.4 Image Classification and accuracy assessment

The clusters resulting from the two datasets are shown in Figure 9. The number of classes resulting from the classification was equal to seven for both study sites, resulting from the silhouette analysis, which defined k parameter setting.

From a first visual interpretation of the entire classified maps, the association between the resulting classes and the burned areas is evident if these are visually compared with the radar indices of Figures 3 and 4.

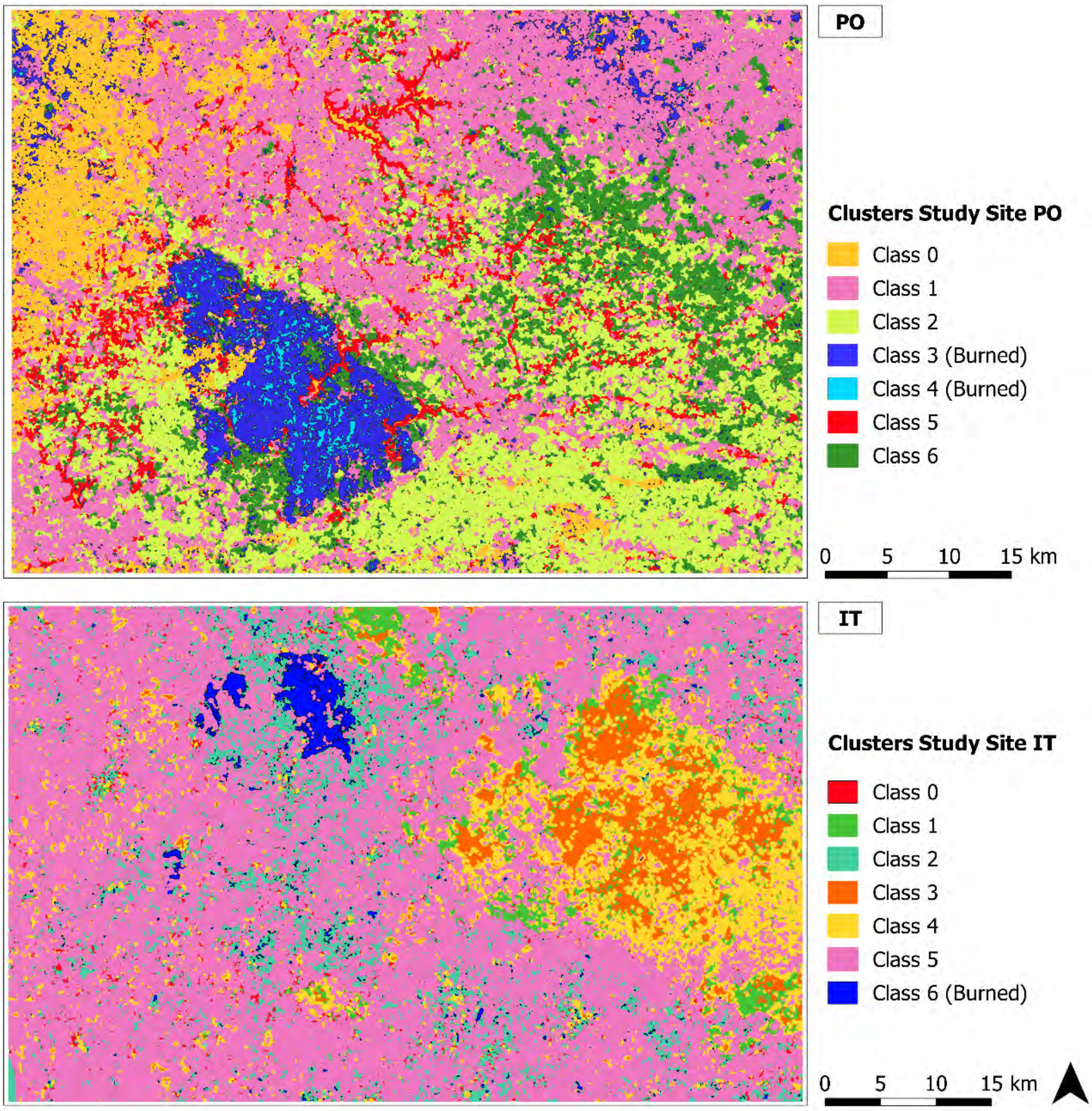


Figure 9. Classification results, showing the seven classes for both study areas. The blue clusters (classes 3-4 in the PO, and 6 in IT) represent the burned areas' classes.

Figure 10 shows the distribution of the mean value of each of the six radar indices for each class (at the top). At the bottom is showed the mean of all the six indices for each class. For the PO study site, the RBD (both polarisations) and the LogRBR_VV maintain stable behaviour for all classes and do not allow class discrimination. The LogRBR_VV shows a slight increase in classes 3, 4 and 6, while the RBD_VV in classes 3 and 4. The other three indices clearly show different behaviour in classes 3 and 4 with lower values, especially observing the RVI and observing the IT plots, the LogRBR (both polarisation), and the RVI lower value in class 6. Classes 1 and 3 are characterised by a positive peak given by some indices: DPSVI, RVI, LogRBR_VH-VV and RBD_VV, respectively. Also, in this case, the RBD_VH had stable behaviour between the classes.

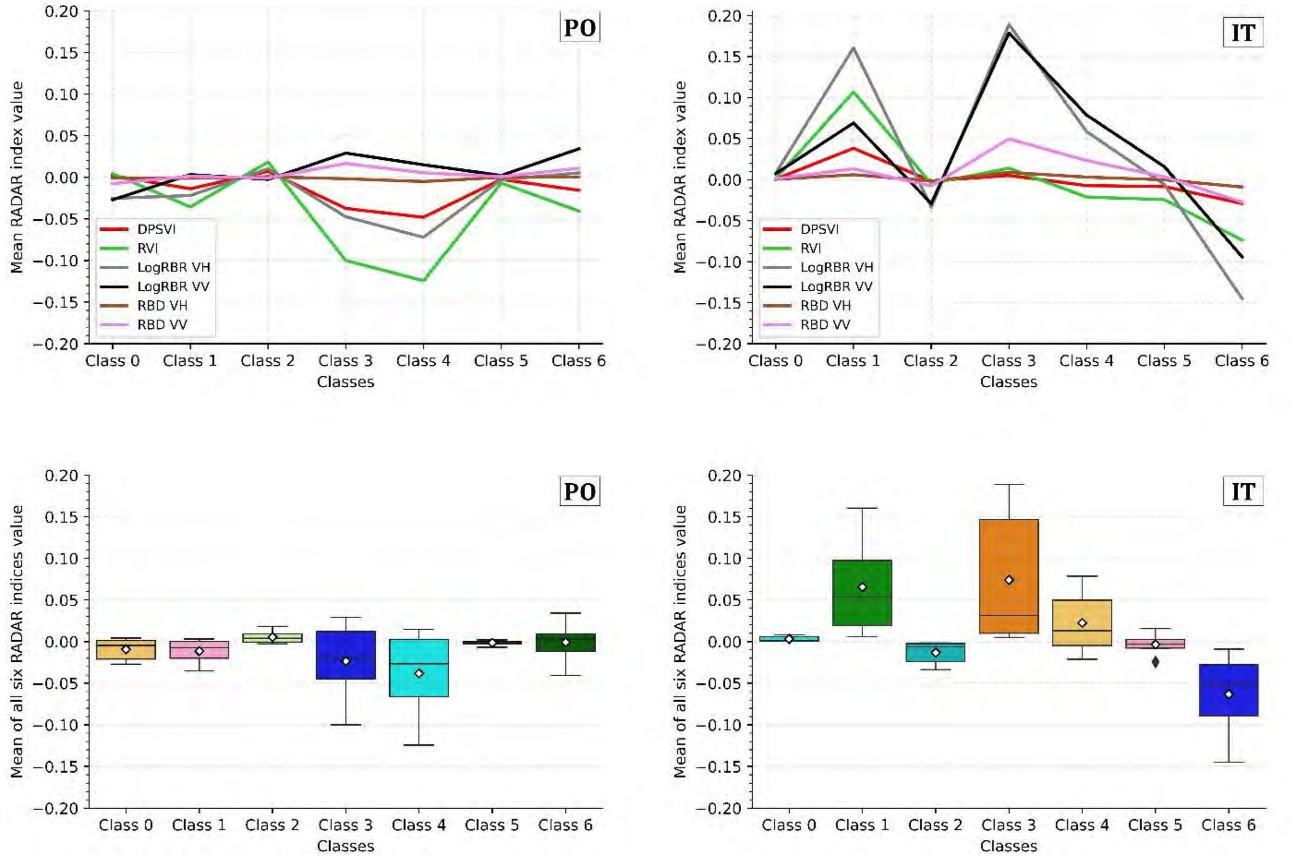


Figure 10. The figure shows the distribution of the mean value of each radar index across all six classes for both study sites (PO and IT) (at the top). At the bottom, boxplots of indices values for each class are reported (the white rhombus marker indicates the mean values).

In Figure 10, it is possible to clearly distinguish the classes that have lower and negative values and a mean below -0.02 for both study sites. Since we are using temporal difference indices, we assume that classes 3, 4 (PO) and 6 (IT) represent the burned areas. In total, considering only these fire-related classes, they covered an area equal to 300.10 km^2 in PO (classes 3 and 4 together) and 51.59 km^2 in IT (class 6). However, we noted the presence of several small areas distributed over all the scenes. For this reason, all the single clusters with a size less or equal than 0.05 km^2 belonging to the fire-related classes were excluded. This threshold corresponds to the minimum mapping unit of the reference data used in the accuracy assessment. The remaining filtered burned clusters covered an area of 269.67 km^2 in PO and 43.28 km^2 in IT.

We filtered the classification results and compared the pixels belonging to the fire-related classes with the reference burned area, according to the three accuracy categories (TP, FN, and FP) to analyse the classification's accuracy. Table 3 shows the distribution of the number of pixels in each of the three accuracy categories.

Table 3. Distribution of each dataset's pixels and the three accuracy categories (true positive, TP; false negative, FN; false positive, FP) for both study sites (PO and IT).

GLCM Features	PO	IT
TP	80.47 %	85.57 %
FP	19.95 %	14.94 %
FN	19.53 %	14.43 %

A visual overview showing the spatial distribution of the accuracy assessment categories (TP, green; FP, yellow; FN, red) is presented in Figure 11 for both study areas. In the same figure, the perimeter of the S-2 based reference burned area, used for accuracy assessment, has been superimposed (blue border).

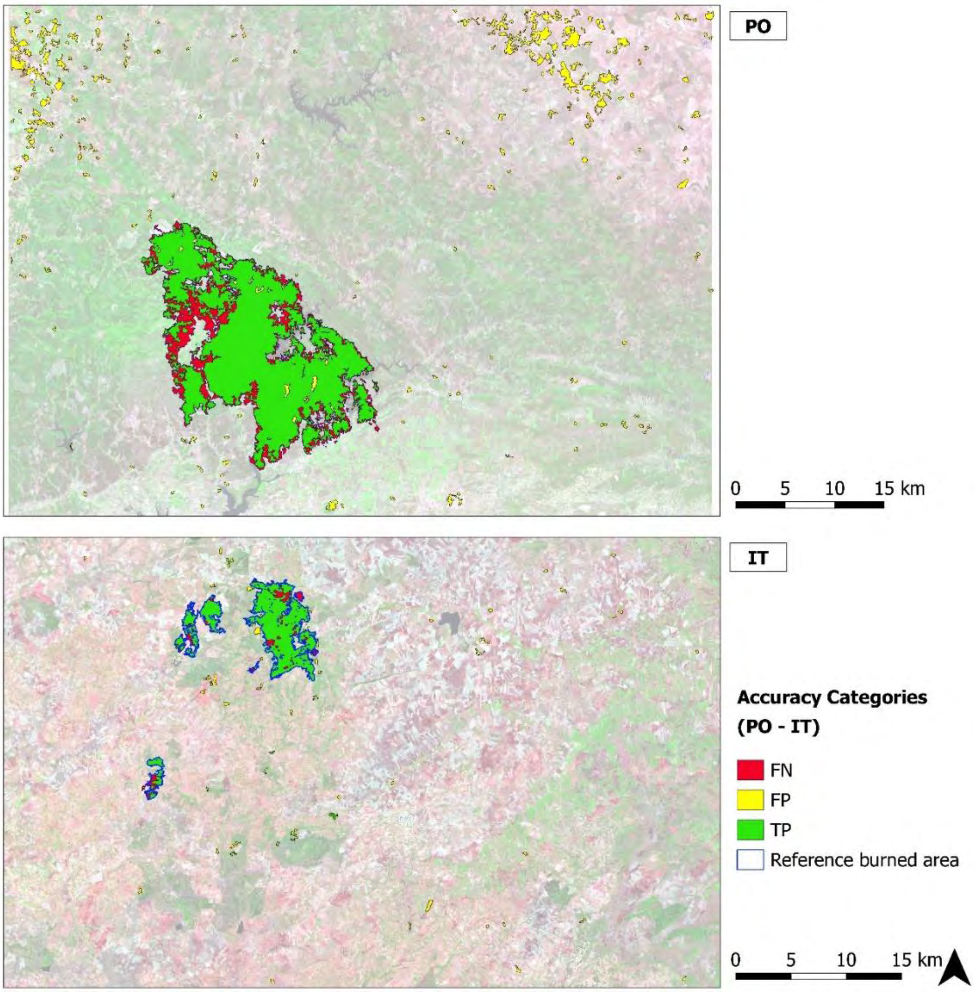


Figure 11. The maps show the spatial distribution of the three accuracy categories, true positive (TP, green), false positive (FP, yellow), false negative (FN), for IT and PO study sites, using the reference layer (blue) derived from S-2 data.

The *r*, *p* and *F-score* accuracy metrics were calculated. The results show that the highest values for *p* and *r* and the *F-score* were reached by the IT classification, with 0.0.851, 0.0.856 and 0.853, respectively, compared to those produced by the PO dataset, which are 0.805, 0.801 and 0.803, respectively.

4. Discussion

4.1 SAR dataset and indices

SAR data entails a more complicated extraction, management and understanding of the extracted information. Compared to the generally more stable accuracy performance of optical data, under optimal time conditions, it must be considered that the research on these is much more consolidated over time, and numerous methodologies of analysis and optimisations have been developed (Chuvieco et al., 2019; Pereira et al., 1999; Meng et al., 2017; Tanase et al., 2020; Miller et al., 2007; De Santis et al., 2009; Fornacca et al., 2018; Filipponi et al., 2018; Fernández-Manso et al., 2016). Tanase et al. (2020) also stated that the development of methodologies for detecting burned areas from SAR sensors is infancy compared to the optical sensors. Further contributions in this field could improve the results. Some studies using deep learning with SAR data, have already shown that accuracy can be high (Ban et al., 2020). We consider that the two types of data should be used as complementary to fill each other's gaps and optimise their usage potential (Lehmann et al., 2015; Stroppiana et al., 2015; Lasko, 2019).

Concerning the number and dates of images used, we have decided to include approximately one month before and one month after the event, represented, in this case, by the more drastic months of the summer fire season (July and August). Since the events under study did not occur precisely on August 1st, this resulted in a different number of pre- and post-fire images. We have not included additional images (i.e., from September) to avoid rain interference, which would have involved further analyses in interpreting the noise. The imbalance in the number of pre- and post-fire images may affect their time average, an issue not explored in the present study. Nevertheless, even with a small number of post-fire images, the aim of reducing speckle noise has been fulfilled.

This study aimed to test and establish the workflow's functionality, focusing mainly on extracting valid and useful information from the SAR data. The approach has been applied only to two regions of the Mediterranean, presenting similar vegetation, climate, and latitude. If further tested and optimised, this method could be easily applicable and with good results, at least in the Mediterranean environments.

The S-1 radar indices (Equations 3, 4, 7, and 8), calculated from the time-averaged data layers and for both study sites (PO and IT), present a well-defined area of low backscatter (darker area), indicative of the fire occurrence. This is in agreement with several research works (e.g., Belenguer-Plomer et al., 2019; Carreiras et al., 2020; Imperatore et al., 2017; Tanase et al., 2015, 2010; Zhang et al., 2019) that show how a progressive fall in the cross-polarised intensity of the radar backscatter is always observed after a forest fire. This is related to the forest structure's loss, leading to a less reflection of the C-band signal (Carreiras et al., 2020; Chuvieco et al., 2019; Donezar et al., 2019; Imperatore et al., 2017; Kurum, 2015; Pepe et al., 2018; Santi et al., 2017; Tanase et al., 2010, 2011, 2020, 2015; Zhang et al., 2019), and the soil changes following the fire occurrence (Hachani et al., 2019; Kurum, 2015; Martinis et al., 2017; Ruiz-Ramos et al., 2018; Tanase et al., 2010).

A clear difference is observed in the PO study site in the co-polarised indices (RBD and LogRBR) obtained from VV polarisation. The corresponding burned areas are represented by lighter pixels (higher backscatter), but in any case, always distinguishable from the rest of the scene. This is partly observable in the upper plots of Figure 10. However, it must be taken into account that they represent the classes deriving from the classification and therefore affected by commission and omission errors. This particularity is not observed in the IT study site, demonstrating a different property of the signal from each polarisation and the possibility of having a different result, depending on a multitude of local conditions as stated in several studies (Belenguer-Plomer et al., 2019; Donezar et al., 2019; Imperatore et al., 2017; Tanase et al., 2010). This is because polarisations have a different interaction with vegetation scattering components based on their size and space orientation. Standing vertical tree trunks depolarise the incoming waves with different strengths than branches or leaves (Flores et al., 2019). The total backscatter coefficient from vegetation surface is the combination of the scattering components given by the volume of the stand, by the volume of the soil, and the combination of these two (Richards J.A., 2009; Flores et al., 2019). The backscatter from co-polarisation is typically stronger for rough surface scattering components (e.g., bare ground). The cross-polarised backscatter form vegetation is associated with the distribution of volume scatterers from leaves and small branches (Flores et al., 2019; Carreiras et al., 2020). So, the cross-polarised backscatter coefficient has higher sensitivity for volume changes, decreasing with the increase of burn severity at all frequencies due to the destruction of the canopy volume component (Tanase et al., 2010; Imperatore et al., 2017; Carreiras et al., 2020). The co-polarised signal VV is not so markedly affected by the loss of the canopy components but is affected by greater exposure of the underlying soil after the destruction of the canopy. As hypothesised by other studies (Tanase et al., 2010; Imperatore et al., 2017), this can result in a different and opposite behaviour compared to the cross-polarised signal, with an increase in backscattering. The sensitivity of the signal to the vegetation structure also depends on the wavelength. It determines the signal's penetration capacity (the longer the band, the lower the frequency, the more the radar waves can penetrate the canopy of trees) and diffusion from the smaller or larger woody components of the forest. Therefore, it affects the degree of interaction of the signal with the underlying components such as the soil, whose contribution increases after disastrous events such as a fire (Saatchi et al., 2016;

Hosseini et al., 2017; Flores et al., 2019). The combined use of both polarisations, using dual-polarimetric difference indices (ΔRVI, ΔDPSVI), represents an effective tool for integrating the information. In general, the use of both polarisations (VV, VH) allows capturing the volume and structure variability of different sizes and orientations of the vegetation (Flores et al., 2019). Polarisation impacts differently how each element of the surface affects the backscatter. Therefore, the use of combined polarisation can help improve the retrieval of more information (Santi et al., 2019; Tanase et al., 2014), and it has already been shown how polarimetric data have high sensitivity towards changes in vegetation conditions (Engelbrecht et al., 2017; Chang et al., 2018; Mandal et al., 2020). Chen et al. (2018) show how indices that combine cross- and co-polarised bands had better performance than single-polarisation when used to map post-fire regrowth in different recovery intervention conditions. Plank et al. (2019) investigated the different behaviours of the quad-polarimetric L-band SAR backscatter properties during active fire and post-fire conditions. Moreover, a series of polarimetric decomposition procedures, including the RVI index, were computed to map the burned scar with an overall accuracy similar to the one we obtained in this research. Martinins et al. (2016) used several dual- and quad polarimetric L-band indices for monitoring forest degradation after the fire, demonstrating that these are very sensitive to forest structure and its modifications. However, none of them was able to discriminate between the intermediate levels of degradation. Dos Santos et al. (2013) show that L-band polarimetric indices can be applied to quantify and monitor the carbon stocks in the tropical forest affected by the fire. Other studies investigated the capability of dominant scattering mechanisms in fully-polarimetric data to detect burned areas using polarimetric decompositions models (Engelbrecht et al., 2017; Goodnough et al., 2011; Czuchlewski et al., 2005; Martins et al., 2016; Tanase et., 2014). All these researches obtained high accuracy values, demonstrating that polarimetric data increase SAR measurement sensitivity for scar detection and classification.

Although the potential of polarimetric indices and backscatter decomposition models has been proven in these mentioned research, some of these dealt with the L-band use (Chen et al., 2018; Plank et al., 2019; Martins et al., 2016; Dos Santos et al., 2013;). Our research is the first to deal with ΔRVI and DPSVI in mapping burned areas using S-1 C band data to our best knowledge. Therefore more research should be carried out to investigate this issue deeply.

4.2 GLCM texture extraction and PCA transformation

For GLCM texture calculation, the square processing window size is crucial since it defines the number of neighbour pixels used for texture calculations (Coburn et al. 2004). GLCM analysis results largely depend on the relationship between the processing window's size and the objects' size and variability across the image (Coburn et al. 2004). Several studies confirmed that small sizes could miss important information for texture characterisation, failing to capture enough surface patterns, while too large windows could introduce systematic errors (Dorigo et al. 2012; Hall-Beyer et al., 2017; Coburn et al. 2004; Franklin et al., 2020; Murray et al. 2010; Caridade et al. 2008). This last hypothesis occurs when the window is too large, overlapping more land-use class edges (Franklin et al., 2000; Dorigo et al., 2012). Coburn et al. (2004) and Murray et al. (2010) demonstrated that using medium-high window size (between 7x7 and 15x15 pixels), there are improvements in the overall accuracy. In our case, small fires (i.e., less than 0.5 km2) were not considered. Moreover, given our research's purpose (i.e., a binary detection of burned/not burned areas), delta indices are useful, considering that they highlight only those areas where changes occurred. Indeed, these indices do not provide any information on the actual land use cover. We fixed the window size to 11x11 pixels following these considerations and based on Muthukumarasamy et al. (2019) research aimed at land cover classification using S-1 and S-2 data. However, if small and scattered burned areas have to be mapped, smaller window sizes should be considered. Similar consideration could be addressed about the window size used for the spatial averaging in each image of the time-series in multitemporal speckle filtering (Quegan et al., 2000).

The datasets transformed and reduced by the PCA can be considered an optimal representation subset of the original ones. On the one hand, it maintains the most useful information in a few layers, speeding up the calculation process. On the other hand, the linear transformation performed on the original images, as a function of the maximum variance expressed, created new, improved imagery, able to discriminate better the changes caused by the fire, and therefore, optimising the unsupervised classification, as already pointed out by Gimeno et al. (2004).

The first PC represents the maximum proportion of the original dataset variance (Fung and Ledrew, 1987). In our case, we used the first PCs obtained that explained a cumulative variance larger than 99%, which revealed with high contrast the area affected by the fire. This is evident in the first PC, as shown in Figures S2 and S3 (Supplementary material). This aspect is important so that the various characteristics of the scene can be circumscribed and classified within the various classes, directly influencing the values obtained in subsequent analyses.

4.3 k-means classification and accuracy assessment

The silhouette score in the preliminary choice of the most suitable number of clusters has solved the well-known problem of parameter setting that allowed reducing the algorithm's implementation time, i.e., avoiding a series of trial-and-error tests. It is evident from the graphs shown in Figures 7 and 8 that for lower k values (<10), the silhouette score and, therefore, the clusters' separation is more significant. A value of 7 seemed to be optimal to discriminate the various areas that characterised the scene, which was an expression of the different surface change levels.

The *k*-means unsupervised classification was applied to the transformed dataset (PCs) to discriminate the burned areas without having prior knowledge of the characteristics and the number of classes characterising the surface background. Although the easy to use and speed execution time characterising the standard *k*-means algorithm has been widely recognised (Nagpal et al., 2013), extensions like the *k-means*++ (Arthur and Vassilvitskii, 2007) improved the reliability of the obtained classifications. Indeed, the standard *k*-means algorithm is very prone to the different numerical distribution of the individual layers' values, making up the datasets, to the so-called outliers with extreme values. The choice of a centroid is generally random in this algorithm, leading to the definition of always different centroids, even in identical and repeated conditions, limiting the results' repeatability. Therefore, all data must be reported on the same scale. In our case, a normalisation (Eq. 9) of all layers values in the range [0, 1] has been carried out. Normalisation is a crucial step when the different input data have different values range. However, although MinMax normalisation is one of the most common ways to rescale the data, it keeps all the data values, including any outliers that can influence the result (Kandanaarachchi et al. 2020). These are very different values from the rest of the other data values, and the k-means algorithm is sensitive to them, affecting its performance (Gan et al., 2017; Hautamäki et al., 2005). These arise from common noise or errors in remotely sensed data (Liu et al., 2017) with anomalous values concerning the surrounding pixels (Alvera-Azcárate et al., 2012). Several methods of outliers detection and correction are present in the literature for general data analysis (Kandanaarachchi et al. 2020, Campos et al. 2016, Angelov et al. 2019, Gan et al. 2017, Hautamäki et al. 2005) and specific remote sensing contexts (Liu et al. 2017, Alvera-Azcárate et al. 2012). Gan et al. (2017) reported a series of related work concerning outliers detection, dedicated to cluster analysis and specific to the k-means algorithm. Given the good results of the first test of the classification, this topic has not been addressed in this study case, but it could be further investigated in future work developments.

Since the quality of the final clustering results depends on the arbitrary selection of initial centroid (Dhanachandra et al., 2015), the *k-means++* (Arthur and Vassilvitskii, 2007), and implemented in the scikit-learn module, optimise the standard *k*-means algorithm by choosing the initial cluster centroids basing on the weighted distribution probability metric and only the first centroid is randomly selected. This seeding method yields a better performing algorithm and consistently finds a better clustering with lower resources than the standard *k*-means (Arthur and Vassilvitskii, 2007).

To estimate SAR S-1 data accuracy in detecting burned areas, the classified maps were compared to the relative reference burned area obtained from S-2 images. From a first visual assessment of the classified maps (Figure 11), the 3, 4 (in IT) and 6 (in PO) classes seem to have detected a large part of the relative affected area, a condition confirmed by observing TPs' distribution in Figure 10. Nevertheless, the *F-score*, *p* and *r* accuracy metrics are those that give an analytic and objective picture of the classification algorithm performance (Modica et al., 2020; Shufelt, 1999). The results indicated a satisfying global accuracy, represented by the *F-score*, for both the study sites, similar to other works using only the SAR data (Belenguer-Plomer et al., 2019; Carreiras et al., 2020; Donezar et al., 2019; Gimeno et al., 2004; Gimeno and San-Miguel-Ayanz, 2004; Lasaponara and Tucci, 2019; Zhang et al., 2019; Goodnough et al., 2011).

However, some commission and omission errors occurred. It should be noted that the omission and commission errors, represented by the opposite of *r* and *p,* respectively, presented similar values in both study sites. Figure 10 shows how most FPs are located in scattered areas throughout the scene and probably represented by local surface changing conditions (i.e., topography, roughness, humidity, local incidence angle) affecting the backscatter signal (Belenguer-Plomer et al., 2019; Donezar et al., 2019; Gimeno et al., 2004; Gimeno and San-Miguel-Ayanz, 2004; Kurum, 2015). Concerning the effects of the terrain conformation and the sensor geometry, these were attenuated by using images deriving from both ascending and descending orbits (Tab. 1), allowing to observe the burned surfaces from multiple angles of incidence of radar beams. This is due to the reliefs' topographic characteristics that determine the radar beam's local incident angle, which plays a fundamental role in the radiometric radar response of the surface (Gimeno and San-Miguel-Ayanz, 2004; Kurum, 2015; Tanase et al., 2010). Also, Donezar et al. (2019) observed how the low detection of some burned areas could be since orography overshadowed these areas facing the side opposite the radar beam, while this problem did not occur when using images of both orbits. This increases the chance that a burned surface that was shadowed in one image would be illuminated on another. The same behaviour was observed in Sayedain et al. (2020), where the use of both ascending and descending orbit directions improved the accuracy of land use classification with S-1 data.

Still, regarding the commission errors, it is necessary to consider the variations inherent in the observed scenario within the time considered from the first pre-fire acquisition date to the last post-fire image date. During this time-frame, other environmental and agricultural changes could also occur. More investigations should be carried out in these contexts. Taking these critical aspects into account, the time-series on which the backscatter was averaged has probably contributed to producing a better result, reducing the intrinsic noises of the radar data (Lasaponara and Tucci, 2019). Obviously, previous knowledge of the meteorological conditions present at the date of image acquisition must be taken into account to select an optimal time series or possibly consider the effects of rains (Gimeno et al., 2004). The multitemporal Lee filter's use allowed further reduction of the noise and amalgamated pixels with different intensities to be similar to their neighbours, thus eliminating small isolated regions (Imperatore et al., 2017).

4.4 Advantages and shortcomings of the implemented workflow

The use of specific Python-based libraries allowed us to build a complete workflow and enclose it in a single script. Furthermore, the use of Python scripts offers the repeatability of the proposed model with high flexibility, allowing any further improvement (e.g., more reliable classification algorithm) with only small script changes. The process is not entirely automatic. Many steps require the user's intervention, such as the imagery selection and the analysis of the results for clusters related to the burned areas. However, the availability of free and open-source software dedicated to remote sensing image processing such as ESA snappy allow connecting the first pre-processing steps to a large number of free toolkits and libraries for exploration, in-depth analysis, data processing, implementing advanced algorithms and graphics (Hao and Ho, 2019; Pedregosa et al., 2011).

The main advantages of the approach developed here were related to (i) self-adaptation to local scattering conditions without the need for a priori information of the observed area; (ii) total free and open-source based workflow, from satellite data to the libraries used in the processing; (iii) possibility of adaptation and interchangeability of parts of the Python-based script (essential for custom improvements); (iv) ability to detect burnt areas during the summer period in territories with heterogeneous vegetation and topographical characteristics, typical in the Mediterranean environment. On the other hand, the main limitations concerned: (i) the misclassification of non-fire related modifications; (ii) dependence of accuracy on variables influencing radar scattering processes (e.g., type of ecosystem, topography). Therefore, there is a need for further improvements to reduce these limitations.

5. Conclusions and Recommendations

Our study showed the potential of the implemented approach, based on Sentinel-1 SAR data, for semi-automated and accurate detection of burned areas in Mediterranean contexts, which is the first and necessary operational step for any subsequent investigations the disturbing effects on vegetation and the environment. This sensor showed to be sensitive to fire-induced changes, and this has been enhanced through the use of radar difference indices. In particular, the dual-polarimetric vegetation indices, RVI and DPSVI, used as differences between pre- and post-event (Δ), have never been used to the best of our knowledge for this purpose with S-1 data. Therefore more investigation will have to be done to find out more about their behaviour. It could be interesting to study these two for the medium and long-term monitoring of post-fire effects and vegetative dynamics.

The pre-processing approaches adopted have made it possible to reduce the adverse geometric and radiometric effects of sensor characteristics and local surface conditions (topography, roughness, humidity, local incidence angle, etc.). These factors mentioned above are those that most affect the backscatter signal. Meanwhile, the combination of using a time-average of the pre- and post-fire time series with a multitemporal speckle-filter can reduce the intrinsic speckle noise of the SAR data. The PCA analysis, reducing the amount of data deriving from pre-processing steps, allowing to decrease the time and computational resources requesting.

Our findings confirm the reliability of open-source and Python-based processing solutions. On the one hand, they allow building an almost complete processing and analysis workflow, with a high degree of interchangeability and flexibility in the choice of components. On the other hand, they offer full repeatability when similar conditions arise or partially repeatability, in this case, using some parts of a process even if some steps requires user intervention.

The research was conducted in two Mediterranean areas with similar environmental characteristics, located in different countries, to test the operability of the methodological workflow and its various components. Future developments may involve testing our approach over larger study areas affected by large and small fires in order to assess the impact of the spatial pattern of burned areas on the classification accuracy. It is also planned to improve some workflow components, such as the use of other radar indices or the use of more robust machine learning techniques, to minimise the presence of commission errors, resulting from signal confusion between burned areas and other land cover types.

Acknowledgements

6. References

Alvera-Azcárate, A., et al. 2012. Outlier detection in satellite data using spatial coherence. Remote Sens. Environ. 119, 84–91. https://doi.org/10.1016/j.rse.2011.12.009

Angelov, P., Gu, X., 2019. Empirical Approach to Machine Learning, IEEE Transactions on Cybernetics. Springer International Publishing. https://doi.org/10.1109/TCYB.2017.2753880

Arthur, D., Vassilvitskii, S., 2007. k-means++: The Advantages of Careful Seeding, in: SODA '07: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms. pp. 1027–1035.

Ban, Y., Zhang, P., Nascetti, A., Bevington, A. R., & Wulder, M. A. (2020). Near Real-Time Wildfire Progression Monitoring with Sentinel-1 SAR Time Series and Deep Learning. *Scientific Reports*, *10*(1), 1–15. https://doi.org/10.1038/s41598-019-56967-x

Belenguer-Plomer, M.A., Tanase, M.A., Fernandez-Carrillo, A., Chuvieco, E., 2019. Burned area detection and mapping using Sentinel-1 backscatter coefficient and thermal anomalies. Remote Sens. Environ. 233, 111345. https://doi.org/10.1016/j.rse.2019.111345

Boschetti, L.; Roy, D.P.; Justice, C.O.; Humber, M.L. MODIS–Landsat fusion for large area 30m burned area mapping. *Remote Sens. Environ.* **2015**, *161*, 27–42. https://doi.org/10.1016/j.rse.2015.01.022

Camerano, P., Cullotta, S., Varese, P., 2011. Strumenti conoscitivi per la gestione delle risorse forestali della Sicilia. Tipi Forestali. Compagnia delle Foreste, Arezzo.

Campos, G.O., Zimek, A., Sander, J. et al. On the evaluation of unsupervised outlier detection: measures, datasets, and an empirical study. Data Min Knowl Disc 30, 891–927 (2016). https://doi.org/10.1007/s10618-015-0444-8

Caridade, C. M. R., A. R. S. Marçal, and T. Mendonça. 2008. "The Use of Texture for Image Classification of Black & White Air Photographs." International Journal of Remote Sensing 29 (2): 593–607. doi:10.1080/01431160701281015.

Carreiras, J.M.B., Quegan, S., Tansey, K., Page, S., 2020. Sentinel-1 observation frequency significantly increases burnt area detectability in tropical SE Asia. Environ. Res. Lett. 15. https://doi.org/10.1088/1748-9326/ab7765

Celik, T., 2009. Unsupervised Change Detection in Satellite Images Using Principal Component Analysis and k-Means Clustering. IEEE Geosci. Remote Sens. Lett. 6, 1457–1457. https://doi.org/10.1109/LGRS.2009.2025059

Chang, J.G., Shoshany, M., Oh, Y., 2018. Polarimetric radar vegetation index for biomass estimation in desert fringe ecosystems. IEEE Trans. Geosci. Remote Sens. 56 (12), 7102–7108.

Chen, W., Jiang, H., & Moriya, K. (2018). Monitoring of post-fire forest regeneration under different restoration treatments based on ALOS / PALSAR data. New Forests, 49(1), 105–121. https://doi.org/10.1007/s11056-017-9608-2

Chu, T., Guo, X., 2013. Remote sensing techniques in monitoring post-fire effects and patterns of forest recovery in boreal forest regions: A review. Remote Sens. 6, 470–520. https://doi.org/10.3390/rs6010470

Chuvieco, E., 2009. Earth observation of wildland fires in mediterranean ecosystems, Earth Observation of Wildland Fires in Mediterranean Ecosystems. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-01754-4

Chuvieco, E., Mouillot, F., van der Werf, G.R., San Miguel, J., Tanasse, M., Koutsias, N., García, M., Yebra, M., Padilla, M., Gitas, I., Heil, A., Hawbaker, T.J., Giglio, L., 2019. Historical background and current developments for mapping burned area from satellite Earth observation. Remote Sens. Environ. 225, 45–64. https://doi.org/10.1016/j.rse.2019.02.013

Coburn, C. A., and A. C. B. Roberts. 2004. A Multiscale Texture Analysis Procedure for Improved Forest Stand Classification. International Journal of Remote Sensing 25: 4287–4308. doi:10.1080/ 0143116042000192367.

Czuchlewski, K.R.; Weissel, J.K., 2005. Synthetic Aperture Radar (SAR)-based mapping of wildfire burn severity and recovery. Int. Geosci. Remote Sens. Symp, 1, 1–4. https://do/10.1109/igarss.2005.1526102

De Luí s, M., Francisca Garcí a-Cano, M., Cortina, J., Raventós, J., Carlos González-Hidalgo, J., Rafael Sánchez, J., 2001. Climatic trends, disturbances and short-term vegetation dynamics in a Mediterranean shrubland. Forest Ecology and Management 147 (1), 25-37. https://doi.org/10.1016/s0378-1127(00)00438-2 , INRMMMiD:13642243

Dekker, R.J., 1998. Speckle filtering in satellite SAR change detection imagery. Int. J. Remote Sens. 19, 1133–1146. https://doi.org/10.1080/014311698215649

Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum Likelihood from Incomplete Data Via the EM Algorithm . J. R. Stat. Soc. Ser. B. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x

Dhanachandra, N., Manglem, K., Chanu, Y.J., 2015. Image Segmentation Using K-means Clustering Algorithm and Subtractive Clustering Algorithm. Procedia Comput. Sci. 54, 764–771. https://doi.org/10.1016/j.procs.2015.06.090

Dijk, D. Van, Shoaie, S., Leeuwen, T. Van, & Veraverbeke, S. (2021). Spectral signature analysis of false positive burned area detection from agricultural harvests using Sentinel-2 data. International Journal of Applied Earth Observations and Geoinformation, 97(September 2020), 102296. https://doi.org/10.1016/j.jag.2021.102296

Donezar, U., De Blas, T., Larrañaga, A., Ros, F., Albizua, L., Steel, A., Broglia, M., 2019. Applicability of the multitemporal coherence approach to sentinel-1 for the detection and delineation of burnt areas in the context of the copernicus emergency management service. Remote Sens. 11. https://doi.org/10.3390/rs11222607

Dorigo, W., A. Lucieer, T. Podobnikara, and A. Carnid. 2012. "Mapping Invasive Fallopia Japonica by Combined Spectral, Spatial, and Temporal Analysis of Digital Orthophotos." International Journal of Applied Earth Observation and Geoinformation 19: 185–195. doi:10.1016/j.jag.2012.05.004.

dos Santos, J. R., F. Martins, L. S. Galvão, and H. A. M. Xaud. 2013. "Contribution of Polarimetric SAR Attributes for Modeling of the Tropical Forest Biomass Affected by Fire".In Proc. of 33rd EARSeL Symposium: Towards Horizon 2020: Earth Observation and Social Perspectives. edited by R. Lasaponara, N. Masini, and M. Biscione, 219–226. Matera, Italy: EARSeL and CNR.

Engelbrecht, J., Theron, A., Vhengani, L., & Kemp, J. (2017). A Simple Normalized Difference Approach to Burnt Area Mapping Using Multi-Polarisation C-Band SAR. Remote Sensing, 9, 764(1), 9–11. https://doi.org/10.3390/rs9080764

Farr, T. G., et al. (2007), The Shuttle Radar Topography Mission, Rev. Geophys., 45, RG2004, doi:10.1029/2005RG000183.

Filipponi, F., 2019. Exploitation of sentinel-2 time series to map burned areas at the national level: A case study on the 2017 Italy wildfires. Remote Sens. 11. https://doi.org/10.3390/rs11060622

Flores, A., Herndon, K., Thapa, R., & Cherrington, E. (2019). SAR Handbook: Comprehensive Methodologies for Forest Monitoring and Biomass Estimation. In *THE SAR HANDBOOK Comprehensive Methodologies for Forest Monitoring and Biomass Estimation*. https://doi.org/10.25966/nr2c-s697

Franklin, S. E., R. J. Hall, L. M. Moskal, and M. B. Lavigne. 2000. "Incorporating Texture into Classification of Forest Species Composition from Airborne Multispectral Images." International Journal of Remote Sensing 21: 61–79. doi:10.1080/014311600210993.

Fraser R., Li Z., Cihlar J., 2000. Hotspot and NDVI differencing synergy (HANDS): A new technique for burned area mapping over boreal forest. Remote Sensing of Environment, 74 (2000), pp. 362-376 https://doi.org/10.1016/S0034-4257(00)00078-X

Fung, T., Ledrew, E., 1987. Application of principal components analysis to change detection. Photogramm. Eng. Remote Sens. 53, 1649–1658.

Gallagher, M. R., Skowronski, N. S., Lathrop, R. G., Green, E. J., Gallagher, M. R., Skowronski, N. S., & Lathrop, R. G. (2020). An Improved Approach for Selecting and Validating Burn Severity Indices in Forested Landscapes An Improved Approach for Selecting and Validating Burn Severity Indices in feux dans des milieux forestiers. Canadian Journal of Remote Sensing, 46(1), 100–111. https://doi.org/10.1080/07038992.2020.1735931

Gan G., Ng M.K.P. (2017) K-means clustering with outlier removal. Pattern Recognit. Lett., 90, pp. 8-14.

https://doi.org/10.1016/j.patrec.2017.03.008

Gimeno, M., San-Miguel-Ayanz, J., 2004. Evaluation of RADARSAT-1 data for identification of burnt areas in Southern Europe. Remote Sens. Environ. 92, 370–375. https://doi.org/10.1016/j.rse.2004.03.018

Gimeno, M., San-Miguel-Ayanz, J., Schmuck, G., 2004. Identification of burnt areas in Mediterranean forest environments from ERS-2 SAR time series. Int. J. Remote Sens. 25, 4873–4888. https://doi.org/10.1080/01431160412331269715

Gitas, I., Mitri, G., Veraverbeke, S., Polychronaki, A., 2012. Advances in Remote Sensing of Post-Fire Vegetation Recovery Monitoring - A Review. Remote Sens. Biomass - Princ. Appl. https://doi.org/10.5772/20571

Goodenough, D.G.; Chen, H.; Richardson, A.; Cloude, S.; Hong,W.; Li, Y., 2011. Mapping fire scars using Radarsat-2 polarimetric SAR data. Can. J. Remote Sens., 37, 500–509. https://do/10.5589/m11-060

Goutte, C., Gaussier, E., 2005. A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation. Lect. Notes Comput. Sci. 3408, 345–359. https://doi.org/10.1007/978-3-540-31865-1_25

Gururaj, P., Umesh, P., Shetty, A., 2019. Assessment of spatial variation of soil moisture during maise growth cycle using SAR observations. https://doi.org/10.1117/12.2532953

Hachani, A., Ouessar, M., Paloscia, S., Santi, E., Pettinato, S., 2019. Soil moisture retrieval from Sentinel-1 acquisitions in an arid environment in Tunisia: application of Artificial Neural Networks techniques. Int. J. Remote Sens. 40, 9159–9180. https://doi.org/10.1080/01431161.2019.1629503

Hall-Beyer, M., 2017. Practical guidelines for choosing GLCM textures to use in landscape classification tasks over a range of moderate spatial scales. Int. J. Remote Sens. 38, 1312–1338. https://doi.org/10.1080/01431161.2016.1278314

Hao, J., Ho, T.K., 2019. Machine Learning Made Easy: A Review of Scikit-learn Package in Python Programming Language. J. Educ. Behav. Stat. 44, 348–361. https://doi.org/10.3102/1076998619832248

Haralick, R.M., 1979. Statistical and structural approaches to texture. Proc. IEEE. https://doi.org/10.1109/PROC.1979.11328

Haralick, R.M., Dinstein, I., Shanmugam, K., 1973. Textural Features for Image Classification. IEEE Trans. Syst. Man Cybern. https://doi.org/10.1109/TSMC.1973.4309314

Hautamäki V., Cherednichenko S., Kärkkäinen I., Kinnunen T., Fränti P. (2005) Improving K-Means by Outlier Removal. In: Kalviainen H., Parkkinen J., Kaarna A. (eds) Image Analysis. SCIA 2005. Lecture Notes in Computer Science, vol 3540. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11499145_99

Hill, J., Stellmes, M., Udelhoven, T., Röder, A., & Sommer, S. (2008). Mediterranean desertification and land degradation. Mapping related land use change syndromes based on satellite observations. Global and Planetary Change, 64(3–4), 146–157. https://doi.org/10.1016/j.gloplacha.2008.10.005

Imperatore, P., Azar, R., Calo, F., Stroppiana, D., Brivio, P.A., Lanari, R., Pepe, A., 2017. Effect of the Vegetation Fire on Backscattering: An Investigation Based on Sentinel-1 Observations. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 4478–4492. https://doi.org/10.1109/JSTARS.2017.2717039

Inoue, Y., Qi, J., Olioso, A., Kiyono, Y., Horie, T., Asai, H., Saito, K., Ochiai, Y., Douangsavanh, L., Qi, J., Olioso, A., Kiyono, Y., Horie, T., Asai, H., Saito, K., & Ochiai, Y. (2019). *Reflectance characteristics of major land surfaces in slash - and - burn ecosystems in Laos*. *1161*. https://doi.org/10.1080/01431160701442039

Jain A., (2010) .Data clustering: 50 years beyond k-means. Pattern Recognit. Lett., 31 (8), pp. 651-666. https://doi.org/10.1016/j.patrec.2009.09.011

Kandanaarachchi, S., Muñoz, M.A., Hyndman, R.J. et al. On normalization and algorithm selection for unsupervised outlier detection. Data Min Knowl Disc 34, 309–354 (2020). https://doi.org/10.1007/s10618-019-00661-z

Key, C., Benson, N., 2006. Ground measure of severity, the Composite Burn Index; and Remote sensing of severity, the Normalized Burn Ratio. chapter Landscape assess- ment (LA): Sampling and analysis methods In: RMRS-GTR-164, G.T.R. (Ed.), FIREMON: Fire Effects Monitoring and Inventory System. USDA Forest Service, Rocky Mountain Research Station, Ogden, pp. 1–51.

Keeley, J.E.: Fire intensity, fire severity and burn severity: a brief review and suggested usage. Int. J. Wildl. Fire 18, 116–126 (2009). https://doi.org/10.1071/WF07049

Kim, Y., Jackson, T., Bindlish, R., Lee, H., Hong, S., 2012. Radar vegetation index for estimating the vegetation water content of rice and soybean. IEEE Geosci. Remote Sens. Lett. 9 (4), 564–568. https://doi.org/10.1109/LGRS.2011.2174772

Kim, Y.; Jackson, T.; Bindlish, R.; Hong, S.; Jung, G.; Lee, K., 2014. Retrieval of Wheat Growth Parameters with Radar Vegetation Indices. IEEE Geosci. Remote Sens. Lett., 11, 4259–4272. https://10.1109/LGRS.2013.2279255

Kim, Y., Van Zyl, J.J., 2009. A time-series approach to estimate soil moisture using polarimetric radar data. IEEE Trans. Geosci. Remote Sens. https://doi.org/10.1109/TGRS.2009.2014944

Kodinariya, T.M., Makwana, P.R., 2013. Review on determining number of Cluster in K-Means Clustering. Int. J. Adv. Res. Comput. Sci. Manag. Stud. 1, 2321–7782.

Kurum, M., 2015. C-Band SAR Backscatter Evaluation of 2008 Gallipoli Forest Fire 12, 1091–1095.

Lanorte, A., Danese, M., Lasaponara, R., Murgante, B., 2012. Multiscale mapping of burn area and severity using multisensor satellite data and spatial autocorrelation analysis. Int. J. Appl. Earth Obs. Geoinf. 20, 42–51.

https://doi.org/10.1016/j.jag.2011.09.005

Lapini, Pettinato, Santi, Paloscia, Fontanelli, Garzelli, 2020. Comparison of Machine Learning Methods Applied to SAR Images for Forest Classification in Mediterranean Areas. Remote Sens. 12, 369. https://doi.org/10.3390/rs12030369

Lasaponara, R., Tucci, B., 2019. Identification of Burned Areas and Severity. IEEE Geosci. Remote Sens. Lett. 16, 917–921. https://doi.org/10.1109/LGRS.2018.2888641

Lasko, K. (2019). Incorporating Sentinel-1 SAR imagery with the MODIS MCD64A1 burned area product to improve burn date estimates and reduce burn date uncertainty in wildland fire mapping Incorporating Sentinel-1 SAR imagery with the MODIS MCD64A1 burned area product to improve burn date. Geocarto International, 0(0), 1–21. https://doi.org/10.1080/10106049.2019.1608592

Lehmann, E. A., Caccetta, P., Lowell, K., Mitchell, A., Zhou, Z. S., Held, A., Milne, T., & Tapley, I. (2015). SAR and optical remote sensing: Assessment of complementarity and interoperability in the context of a large-scale operational forest monitoring system. Remote Sensing of Environment, 156, 335–348. https://doi.org/10.1016/j.rse.2014.09.034

Li, M., Zang, S., Zhang, B., Li, S., Wu, C., 2014. A review of remote sensing image classification techniques: The role of Spatio-contextual information. Eur. J. Remote Sens. 47, 389–411. https://doi.org/10.5721/EuJRS20144723

Liu Q., Klucik R., Chen C., Grant G., Gallaher D., Lv D., Shang L. (2017). Unsupervised detection of contextual anomaly in remotely sensed data. Remote Sensing of Environment, Volume 202, Pages 75-87, ISSN 0034-4257. https://doi.org/10.1016/j.rse.2017.01.034.

Lizundia-Loiola, J., Otón, G., Ramo, R., Chuvieco, E., 2020. A spatio-temporal active-fire clustering approach for global burned area mapping at 250 m from MODIS data. Remote Sens. Environ. 236, 111493. https://doi.org/10.1016/j.rse.2019.111493

MacQueen, J., 1967. Some methods for classification and analysis of multivariate observations, in: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability.

Mandal, D., Kumar, V., Ratha, D., Dey, S., Bhattacharya, A., Lopez-sanchez, J.M., Mcnairn, H., Rao, Y.S., 2020. Remote Sensing of Environment Dual polarimetric radar vegetation index for crop growth monitoring using sentinel-1 SAR data. Remote Sens. Environ. 247, 111954. https://doi.org/10.1016/j.rse.2020.111954

Martinis, S., Caspard, M., Plank, S., Clandillon, S., Haouet, S., 2017. Mapping burn scars, fire severity and soil erosion susceptibility in Southern France using multisensoral satellite data. Int. Geosci. Remote Sens. Symp. 2017-July, 1099–1102. https://doi.org/10.1109/IGARSS.2017.8127148

Martins, F. D. S. R. V., J. R. dos Santos, L. S. Galvão, and H. A. M. Xaud. 2016. “Sensitivity of ALOS/ PALSAR Imagery to Forest Degradation by Fire in Northern Amazon.” International Journal of Applied Earth Observation and Geoinformation 49: 163–174. doi:10.1016/j.jag.2016.02.009.

Meng, R.; Zhao, F. Remote sensing of fire effects. A review for recent advances in burned area and burnseverity mapping. InRemote Sensing of Hydrometeorological Hazards; Petropoulos, G.P., Islam, T., Eds.;CRC Press: Boca Raton, FL, USA, 2017; pp. 261–276Minchella, A., Del Frate, F., Capogna, F., Anselmi, S., Manes, F., 2009. Use of multitemporal SAR data for monitoring vegetation recovery of Mediterranean burned areas. Remote Sens. Environ. 113, 588–597. https://doi.org/10.1016/j.rse.2008.11.004

Modica, G., Messina, G., De Luca, G., Fiozzo, V., Praticò, S., 2020. Monitoring the vegetation vigor in heterogeneous citrus and olive orchards. A multiscale object-based approach to extract trees’ crowns from UAV multispectral imagery. Comput. Electron. Agric. 175, 105500. https://doi.org/105500

Mouillot, F., Schultz, M.G., Yue, C., Cadule, P., Tansey, K., Ciais, P., Chuvieco, E., 2014. Ten years of global burned area products from spaceborne remote sensing-A review: Analysis of user needs and recommendations for future developments. Int. J. Appl. Earth Obs. Geoinf. 26, 64–79. https://doi.org/10.1016/j.jag.2013.05.014

Murray H., Lucieer A., Williams R. 2010 Texture-based classification of sub-Antarctic vegetation communities on Heard Island Int. J. Appl. Earth Obs. Geoinf., 12 (2010), pp. 138-149

Nagpal, A., Jatain, A., Gaur, D., 2013. Review based on data clustering algorithms. 2013 IEEE Conf. Inf. Commun. Technol. ICT 2013 298–303. https://doi.org/10.1109/CICT.2013.6558109

Nasirzadehdizaji, R., Sanli, F.B., Abdikan, S., Cakir, Z., Sekertekin, A., Ustuner, M., 2019. Sensitivity analysis of multi-temporal Sentinel-1 SAR parameters to crop height and canopy coverage. Appl. Sci. https://doi.org/10.3390/app9040655

Oom D., Silva J., Aguiar F., Petit N., (Instituto Superior de Agronomia – University of Lisbon, Lisbon, Portugal). Personal communication, 2018

Otón, G., Ramo, R., Lizundia-Loiola, J., Chuvieco, E., 2019. Global detection of long-term (1982-2017) burned area with AVHRR-LTDR data. Remote Sens. 11. https://doi.org/10.3390/rs11182079

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: Machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Pepe, A., Stroppiana, D., Calo, F., Imperatore, P., Boschetti, L., Bignami, C., Brivio, P.A., 2018. Exploitation of Copernicus Sentinels data for sensing fire-disturbed vegetated areas. Institute for the Electromagnetic Sensing of the Environment (IREA). 7593–7596.Periasamy, S., 2018. Significance of dual polarimetric synthetic aperture radar in biomass retrieval: An attempt on Sentinel-1. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2018.09.003

Phiri, D., Morgenroth, J., 2017. Developments in Landsat Land Cover Classification Methods: A Review. Remote Sens. 9, 967. https://doi.org/10.3390/rs9090967

Pipia, L., Muñoz-Marí, J., Amin, E., Belda, S., Camps-Valls, G., & Verrelst, J. (2019). Fusing optical and SAR time series for LAI gap fillingwith multioutput Gaussian processes. Remote Sensing of Environment, 235(September), 111452. https://doi.org/10.1016/j.rse.2019.111452

Pereira J.M.C., Sa A.C.L., Sousa A.M.O, Silva J.M.N., Santos T.N., Carreiras J.M.B., 1999. Spectral characterisation and discrimination of burnt areas E. Chuvieco (Ed.), Remote Sensing of Large Wildfires in the European Mediterranean Basin, Springer-Verlag, Berlin (1999), pp. 123-138

Quegan, S., Toan, T. Le, Yu, J.J., Ribbes, F., Floury, N., 2000. Multitemporal ERS SAR analysis applied to forest mapping. IEEE Trans. Geosci. Remote Sens. https://doi.org/10.1109/36.842003

Rehman, T.U., Mahmud, M.S., Chang, Y.K., Jin, J., Shin, J., 2019. Current and future applications of statistical machine learning algorithms for agricultural machine vision systems. Comput. Electron. Agric. 156, 585–605. https://doi.org/10.1016/j.compag.2018.12.006

Richards, J.A., 2013. Remote sensing digital image analysis: An introduction.5th ed., Remote Sensing Digital Image Analysis: An Introduction. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-30062-2

Richards, J. A. (2009). Remote Sensing with Imaging Radar. 1st ed. In Springer (1st ed.). Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-642-02020-9

Rodman, K. C., Andrus, R. A., Veblen, T. T., & Hart, S. J. (2021). Remote Sensing of Environment Disturbance detection in landsat time series is influenced by tree mortality agent and severity , not by prior disturbance. Remote Sensing of Environment, 254(December 2020), 112244. https://doi.org/10.1016/j.rse.2020.112244

Rosa, I.M.D., Pereira, J.M.C., Tarantola, S., 2011. Atmospheric emissions from vegetation fires in Portugal (1990-2008): Estimates, uncertainty analysis, and sensitivity analysis. Atmos. Chem. Phys. 11, 2625–2640. https://doi.org/10.5194/acp-11-2625-2011

Rousseeuw, P.J., 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 20, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Ruiz-Ramos, J., Marino, A., Boardman, C.P., 2018. Using sentinel 1-SAR for monitoring long term variation in burnt forest areas. Int. Geosci. Remote Sens. Symp. 2018-July, 4901–4904. https://doi.org/10.1109/IGARSS.2018.8518960ù

Sayedain S. A., Maghsoudi Y., Eini-Zinab S. (2020) Assessing the use of cross-orbit Sentinel-1 images in land cover classification, International Journal of Remote Sensing, 41:20, 7801-7819, DOI: 10.1080/01431161.2020.1763512

San-Miguel-Ayanz, J., de Rigo, D., Caudullo, G., Houston Durrant, T., Mauri, A., 2016. European Forest Tree Species. Publication Office of the European Union, Luxemburg.

San-Miguel-Ayanza, J., Durrant, T., Boca, R., Libertà, G., Branco, A., de Rigo, D., Ferrari, D., Maianti, P., Vivancons, T.A., Costa H., Lana, F., 2018. Forest Fires in Europe, Middle East and North Africa 2017. JRC Technical Report. Publications Office of the European Union. https://doi.org/10.2760/476964

San-Miguel-Ayanza, J., Durrant, T., Boca, R., Libertà, G., Branco, A., de Rigo, D., Ferrari, D., Maianti, P., Vivancons, T.A., Oom, D., Pfeiffer, H., 2019. Forest Fires in Europe, Middle East and North Africa 2018. JRC Technical Report. Publications Office of the European Union. https://doi.org/10.2760/1128

Santi, E., Dabboor, M., Pettinato, S., Paloscia, S., Notarnicola, C., Greifeneder, F., Cuozzo, G., 2019. On The Use of Machine Learning and Polarimetry For Estimating Soil Moisture From Radarsat Imagery Over Italian And Canadian Test Sites. IGARSS 2019 - 2019 IEEE Int. Geosci. Remote Sens. Symp. 5760–5763. https://doi.org/10.1109/igarss.2019.8900269

Santi, E., Paloscia, S., Pettinato, S., Fontanelli, G., Mura, M., Zolli, C., Maselli, F., Chiesi, M., Bottai, L., Chirici, G., 2017. The potential of multifrequency SAR images for estimating forest biomass in Mediterranean areas. Remote Sens. Environ. 200, 63–73. https://doi.org/10.1016/j.rse.2017.07.038

Santoso, A.W., Pebrianti, D., Bayuaji, L., Zain, J.M., 2015. Performance of various speckle reduction filters on Synthetic Aperture Radar image. 2015 4th Int. Conf. Softw. Eng. Comput. Syst. ICSECS 2015 Virtuous Softw. Solut. Big Data 11–14. https://doi.org/10.1109/ICSECS.2015.7333103

Senthilnath, J., Kandukuri, M., Dokania, A., Ramesh, K.N., 2017. Application of UAV imaging platform for vegetation analysis based on spectral-spatial methods. Comput. Electron. Agric. 140, 8–24. https://doi.org/10.1016/j.compag.2017.05.027

Shufelt, J.A., 1999. Performance evaluation and analysis of monocular building extraction from aerial imagery. IEEE Trans. Pattern Anal. Mach. Intell. 21, 311–326. https://doi.org/10.1109/34.761262

Small, D., 2011. Flattening gamma: Radiometric terrain correction for SAR imagery. IEEE Trans. Geosci. Remote Sens. 49, 3081–3093. https://doi.org/10.1109/TGRS.2011.2120616

Smith A.M.S., Wooster M.J., Drake N.A., Dipotso F.M., Falkowski M.J., Hudak A.T., 2005. Testing the potential of multispectral remote sensing for retrospectively estimating fire severity in African Savannahs. Remote Sens. Environ., 97 , pp. 92-115, 10.1016/j.rse.2005.04.014

Sokolova, M., Japkowicz, N., Szpakowicz, S., 2006. Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation. pp. 1015–1021. https://doi.org/10.1007/11941439_114

Sokolova, M., Lapalme, G., 2009. A systematic analysis of performance measures for classification tasks. Inf. Process. Manag. 45, 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

Soni, K.G., Patel, D.A., 2017. Comparative analysis of K-means and K-medoids algorithm on Iris data. Int. J. Comput. Intell. Res. 5, 33–37. https://doi.org/10.5120/888-1261

Stroppiana, D., Azar, R., Calò, F., Pepe, A., Imperatore, P., Boschetti, M., Silva, J. M. N., Brivio, P. A., & Lanari, R. (2015). Integration of optical and SAR data for burned area mapping in Mediterranean regions. Remote Sensing, 7(2), 1320–1345. https://doi.org/10.3390/rs70201320

Subasi, A., 2020. Chapter 2 - Data prepocessing, in: Practical Machine Learning for Data Analysis Using Python. pp. 27–89. https://doi.org/10.1016/B978-0-12-821379-7.00002-3

Tanase, M., de la Riva, J., Santoro, M., Pérez-Cabello, F., Kasischke, E., 2011. Sensitivity of SAR data to post-fire forest regrowth in Mediterranean and boreal forests. Remote Sens. Environ. 115, 2075–2085. https://doi.org/10.1016/j.rse.2011.04.009

Tanase, M.A., Belenguer-Plomer, M.A., Roteta, E., Bastarrika, A., Wheeler, J., Fernández-Carrillo, Á., Tansey, K., Wiedemann, W., Navratil, P., Lohberger, S., Siegert, F., Chuvieco, E., 2020. Burned Area Detection and Mapping: Intercomparison of Sentinel-1 and Sentinel-2 Based Algorithms over Tropical Africa. Remote Sens. 12, 334. https://doi.org/10.3390/rs12020334

Tanase, M.A., Kennedy, R., Aponte, C., 2015. Radar Burn Ratio for fire severity estimation at canopy level: An example for temperate forests. Remote Sens. Environ. 170, 14–31. https://doi.org/10.1016/j.rse.2015.08.025

Tanase, M. A., Santoro, M., Aponte, C., & de la Riva, J. (2014). Polarimetric Properties of Burned Forest Areas at C- and L-Band. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(1), 267–276. doi:10.1109/jstars.2013.2261053

Tanase, M.A., Santoro, M., De La Riva, J., Pérez-Cabello, F., Le Toan, T., 2010. Sensitivity of X-, C-, and L-band SAR backscatter to burn severity in Mediterranean pine forests. IEEE Trans. Geosci. Remote Sens. 48, 3663–3675. https://doi.org/10.1109/TGRS.2010.2049653

Trudel, M., Charbonneau, F., & Leconte, R. (2012). Using RADARSAT-2 polarimetric and ENVISAT-ASAR dual-polarisation data for estimating soil moisture over agricultural fields. Canadian Journal of Remote Sensing, 38(4), 514–527. https://doi.org/10.5589/m12-043

U.S. Geological Survey, 2013. Landsat 8: U.S. Geological Survey Fact Sheet. https://doi.org/10.3133/fs20133060

Verbyla D. L., Kasischke E. S., Hoy E. E., 2008. “Seasonal and topographic effects on estimating fire severity from Landsat TM/ETM + data,” Int. J. Wildland Fire, vol. 17, no. 4, pp. 527–534. https://doi.org/10.1071/WF08038

Zhang, P., Nascetti, A., Ban, Y., Gong, M., 2019. An implicit radar convolutional burn index for burnt area mapping with Sentinel-1 C-band SAR data. ISPRS J. Photogramm. Remote Sens. 158, 50–62. https://doi.org/10.1016/j.isprsjprs.2019.09.013

Zhou, Z., Liu, L., Jiang, L., Feng, W., & Samsonov, S. V. (2019). Using long-term SAR backscatter data to monitor post-fire vegetation recovery in tundra environment. Remote Sensing, 11(19). https://doi.org/10.3390/rs11192230

6.1 Websites

aria2 download utility Homeage (2020): https://aria2.github.io/. Last access 05/10/2020

Copernicus Open Access Hub (2020): https://scihub.copernicus.eu/. Last access 05/10/2020

ESA sen2cor Homepage (2020): https://step.esa.int/main/snap-supported-plugins/sen2cor/. Last access 30/12/2020

ESA Sentinel-1 User Guide (2016): https://sentinel.esa.int/web/sentinel/user-guides/sentinel-1-sar/resolutions/level-1-ground-range-detected. Last access 02/01/2021

ESA Sentinel Homepage (2020): https://sentinel.esa.int/web/sentinel/home. Last access 08/09/2020

ESA SNAP Cookbook (2020): https://senbox.atlassian.net/wiki/spaces/SNAP/pages/24051769/Cookbook. Last access 05/10/2020

ESA SNAP Homepage (2020): http://step.esa.int/main/toolboxes/snap/. Last access 08/09/2020

GitHub Code repository (2021): https://github.com/Jando93/Sentinel-1-unsupervised-burned-area-detection. (DOI: 10.5281/zenodo.4556927). Uploaded 15/01/2021.

Scikit-learn Homepage (2020): https://scikit-learn.org/. Last access 15/09/2020

Sistema Nacional de Informação Geogrãfica (SNIG) (2021) : https://snig.dgterritorio.gov.pt/. Last access 15/01/2021

SNIG - Cartografia nacional de áreas ardidas, area ardita 2018 (2021). : https://sig.icnf.pt/portal/home/item.html?id=983c4e6c4d5b4666b258a3ad5f3ea5af. Last access 20/01/2021.

The Python Language Reference (2020): https://docs.python.org/3/reference/. Last access 15/09/2020

Figure 1

Click here to access/download;Figure;Figure 1.jpg

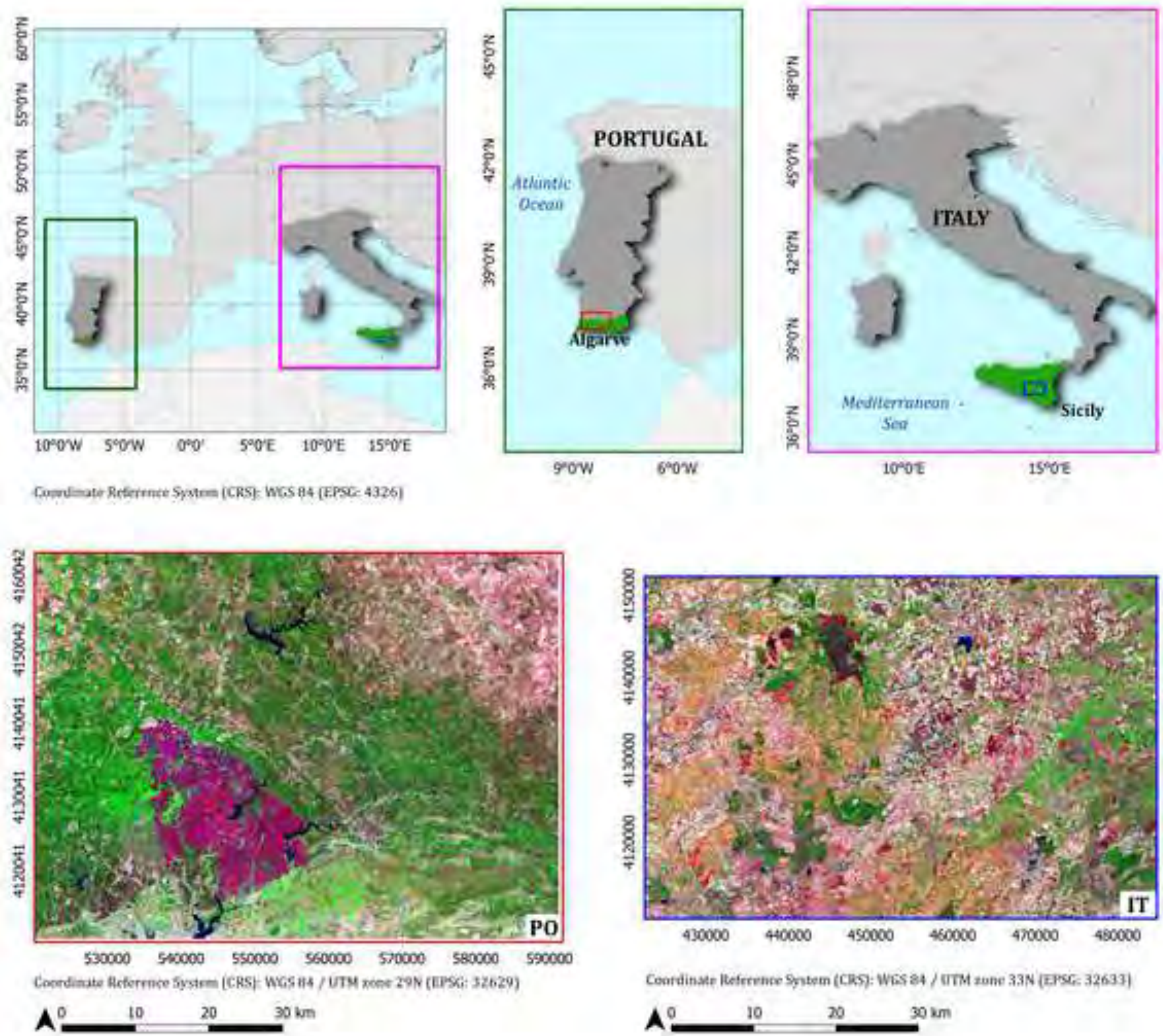


Figure 2 Click here to access/download;Figure;Figure 2.jpg

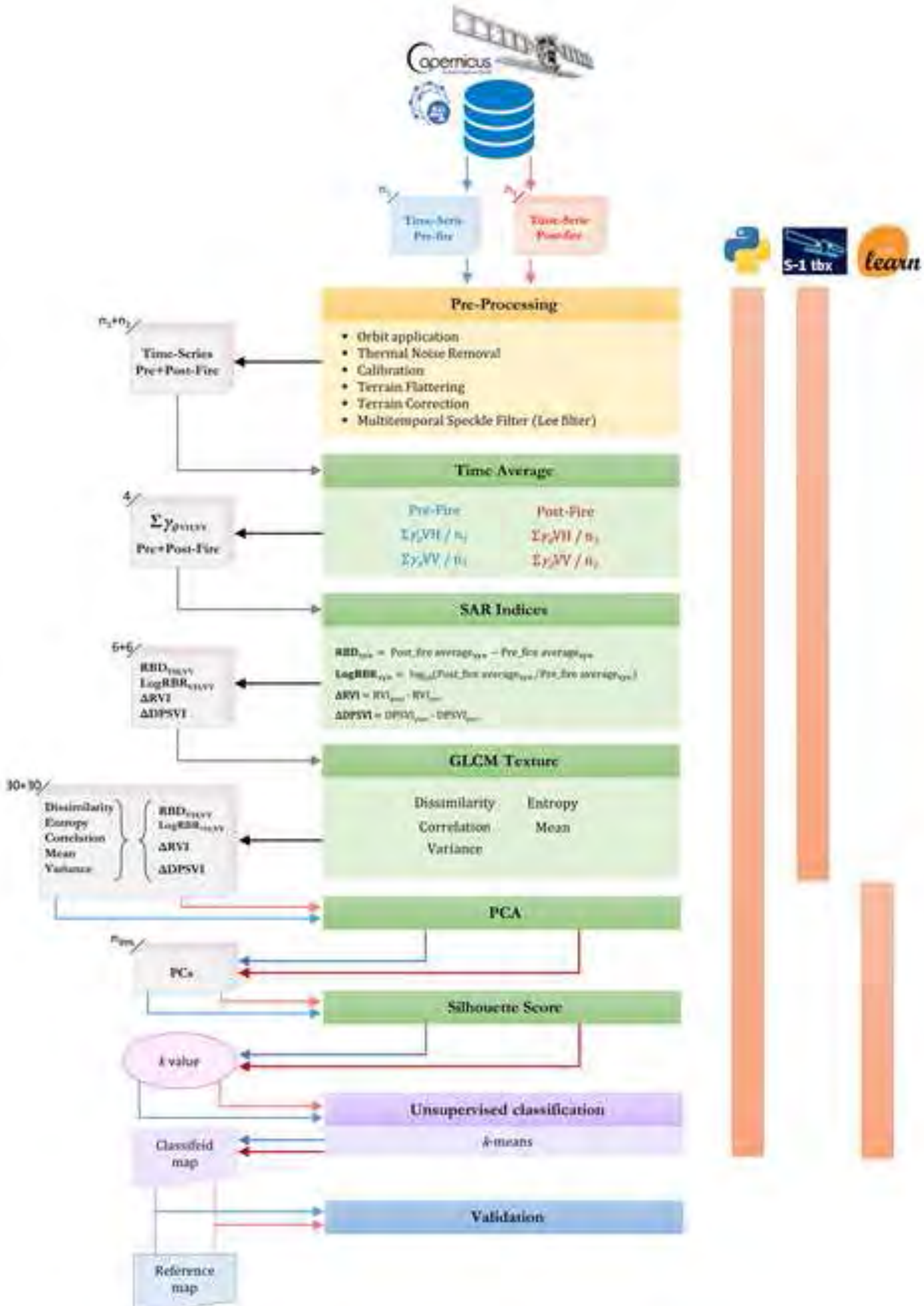
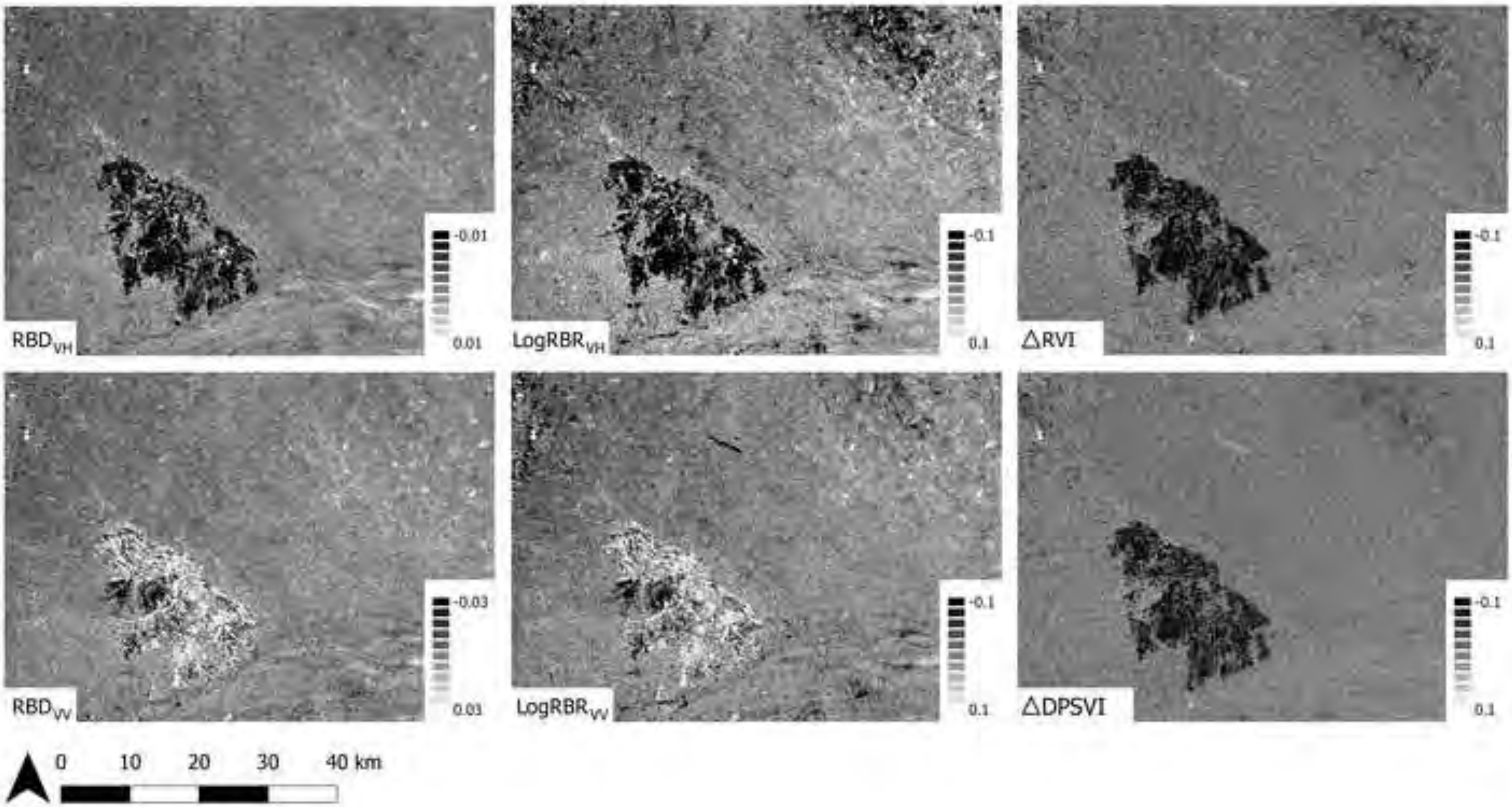


Figure 3



RBD VH
-0.01
0.01
LogRBR VH
-0.1
0.1
ΔRVI
-0.1
0.1
RBD VV
-0.03
0.03
LogRBR VV
-0.1
0.1
ΔDPSVI
-0.1
0.1
0 10 20 30 40 km

Figure 4

Click here to access/download;Figure;Figure 4.jpg

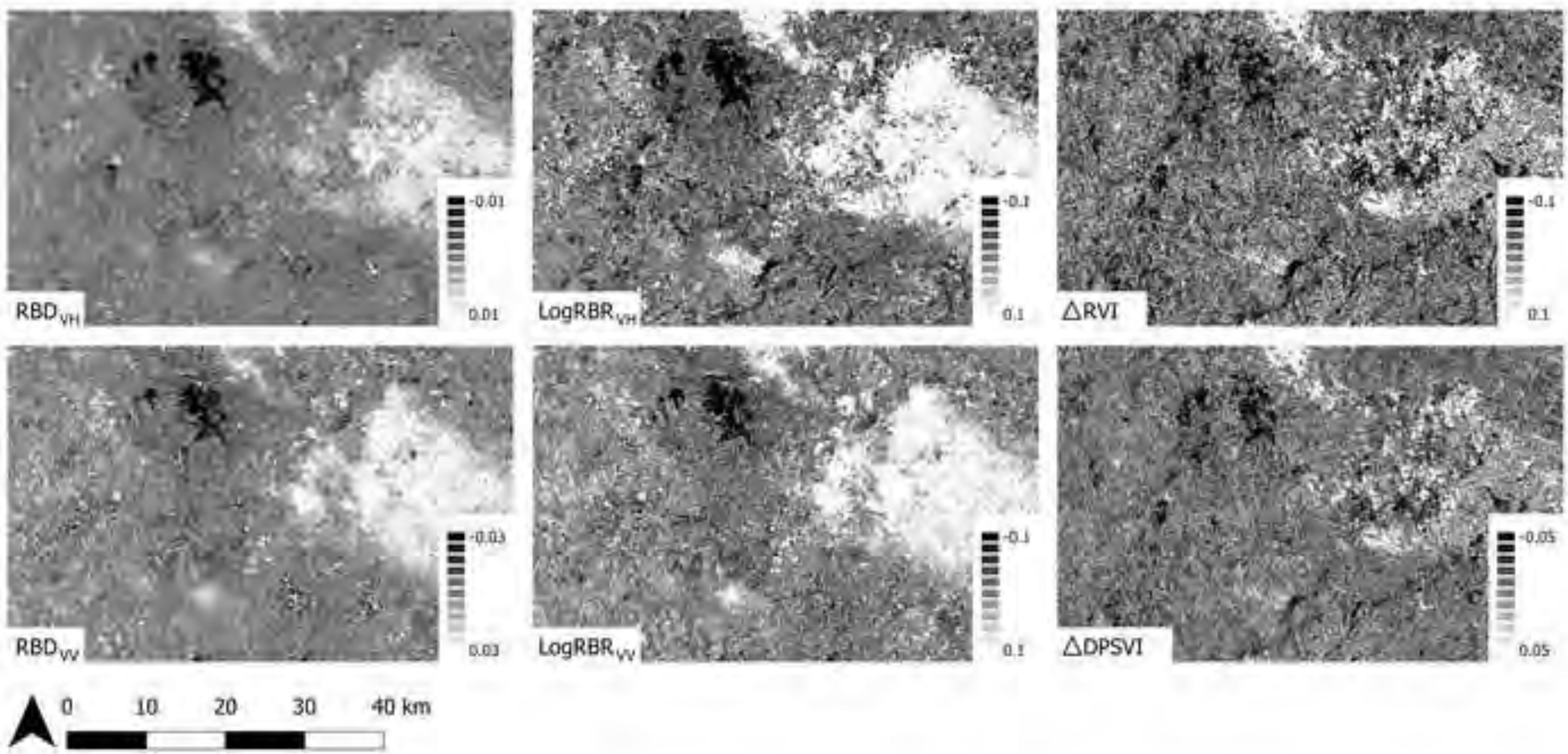
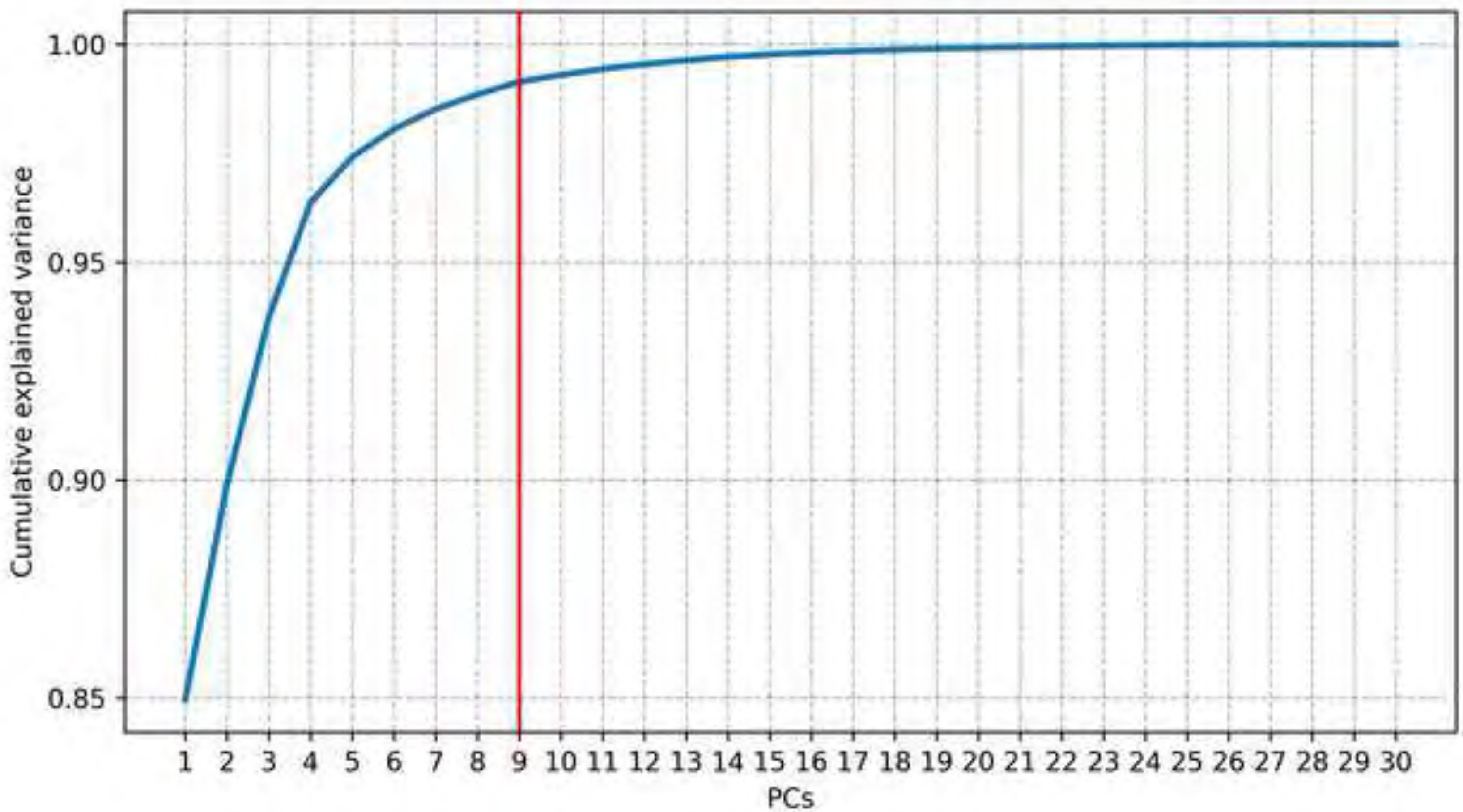


Figure 5



Cumulative explained variance
1.00
0.95
0.90
0.85
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
PCs

Figure 6

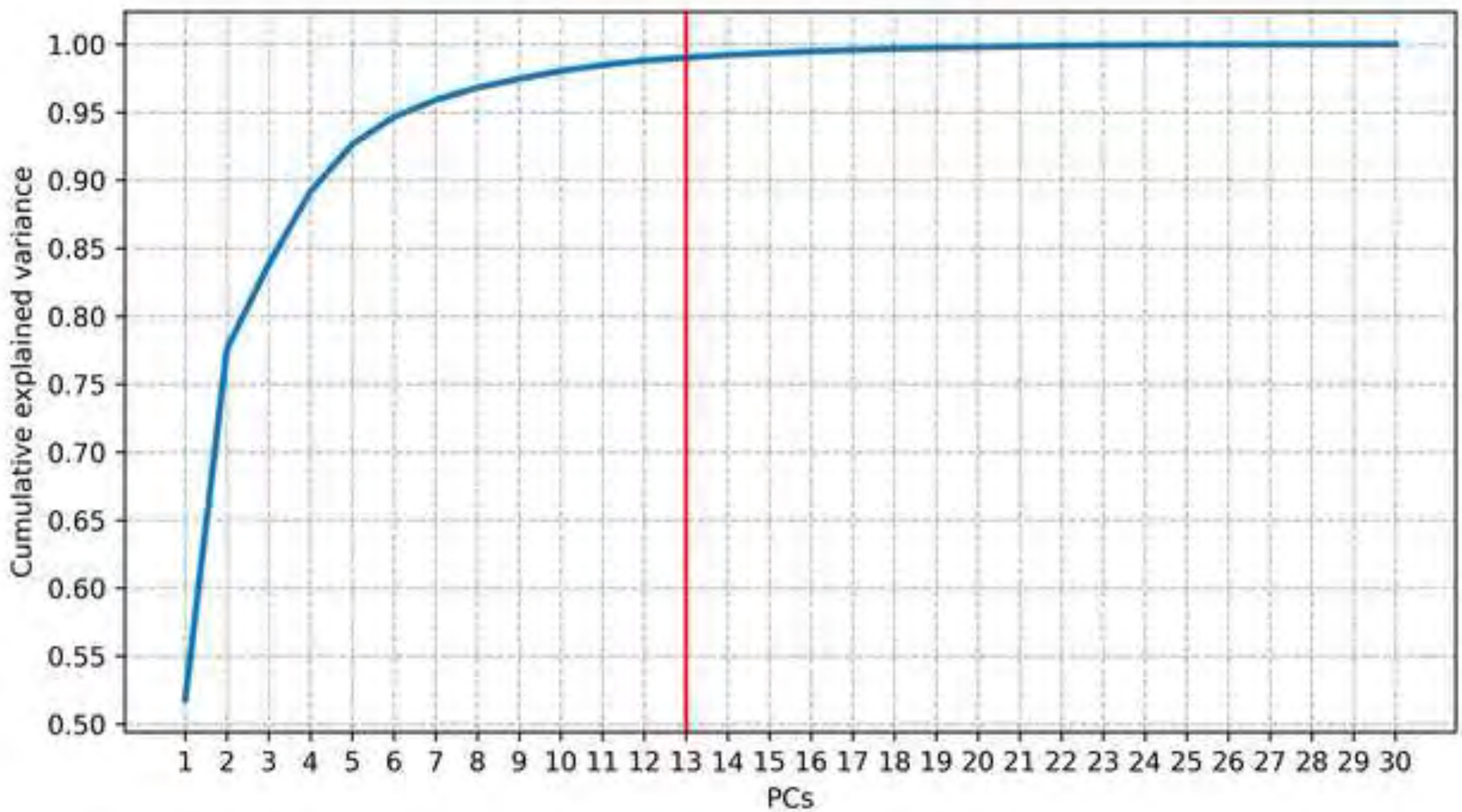

1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
Cumulative explained variance
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
PCs

Figure 7

Click here to access/download;Figure;Figure 7.png

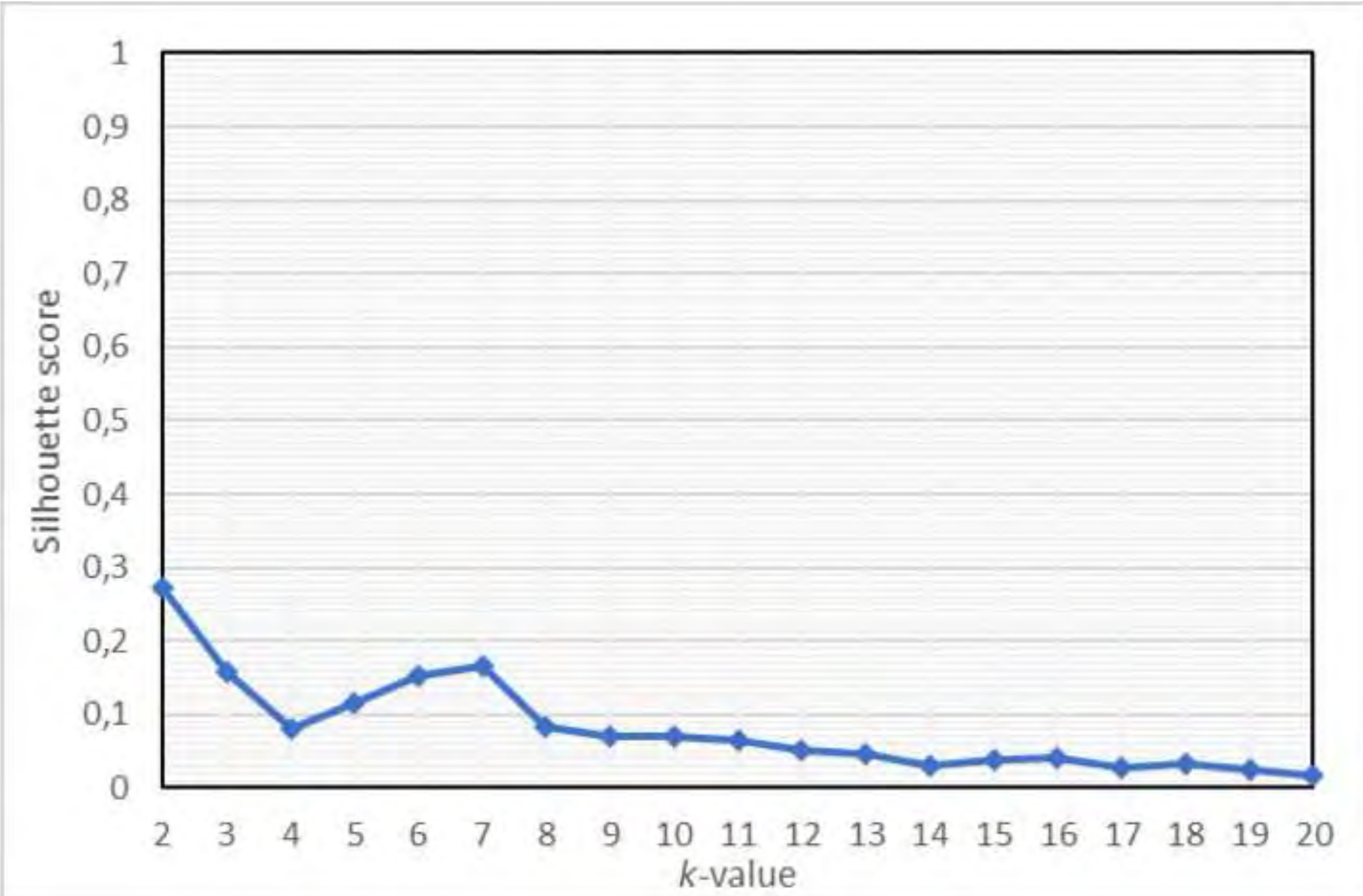
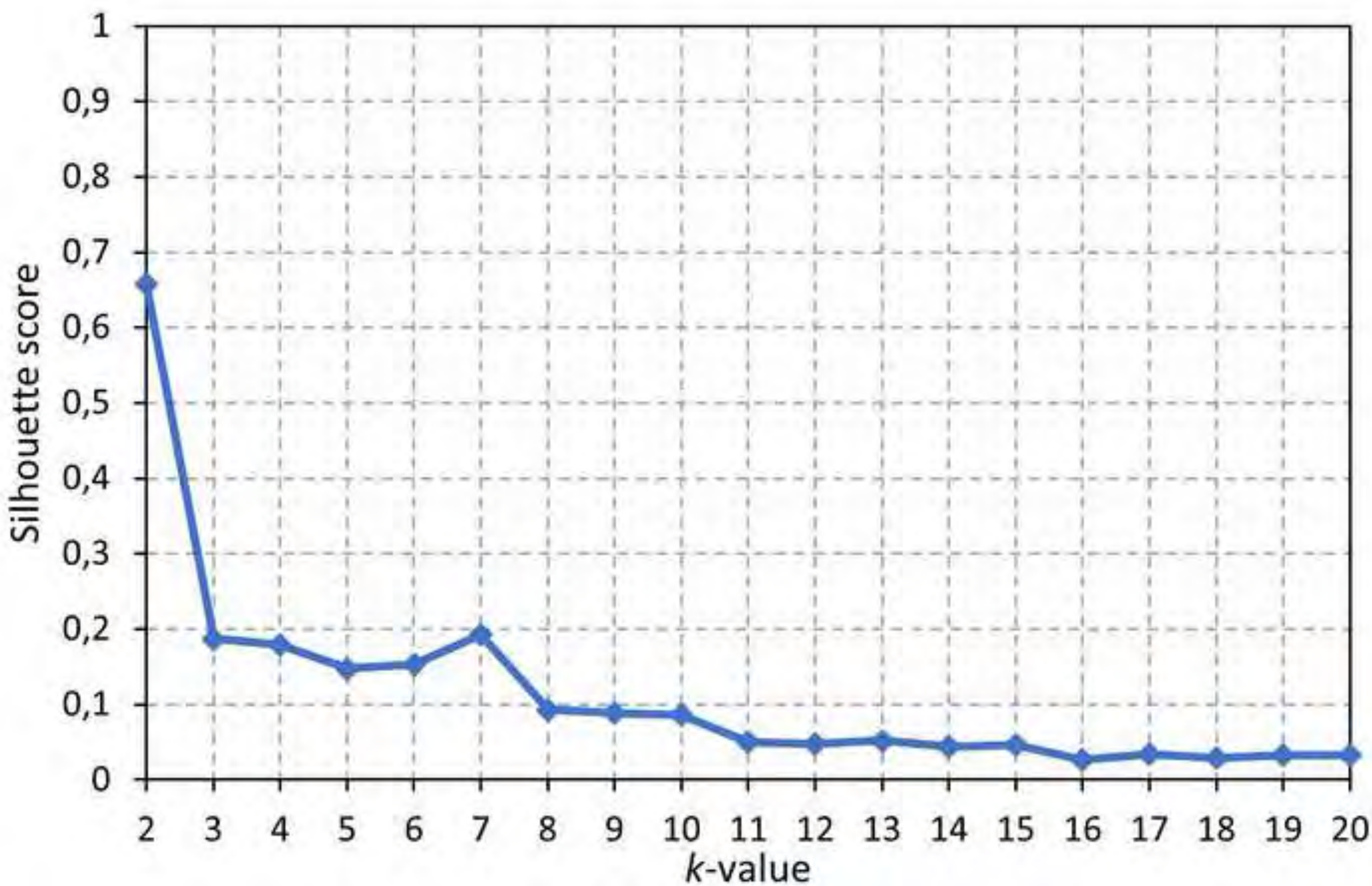


Figure 8



Silhouette score
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
k-value

Figure 9

Click here to access/download;Figure;Figure 9.png

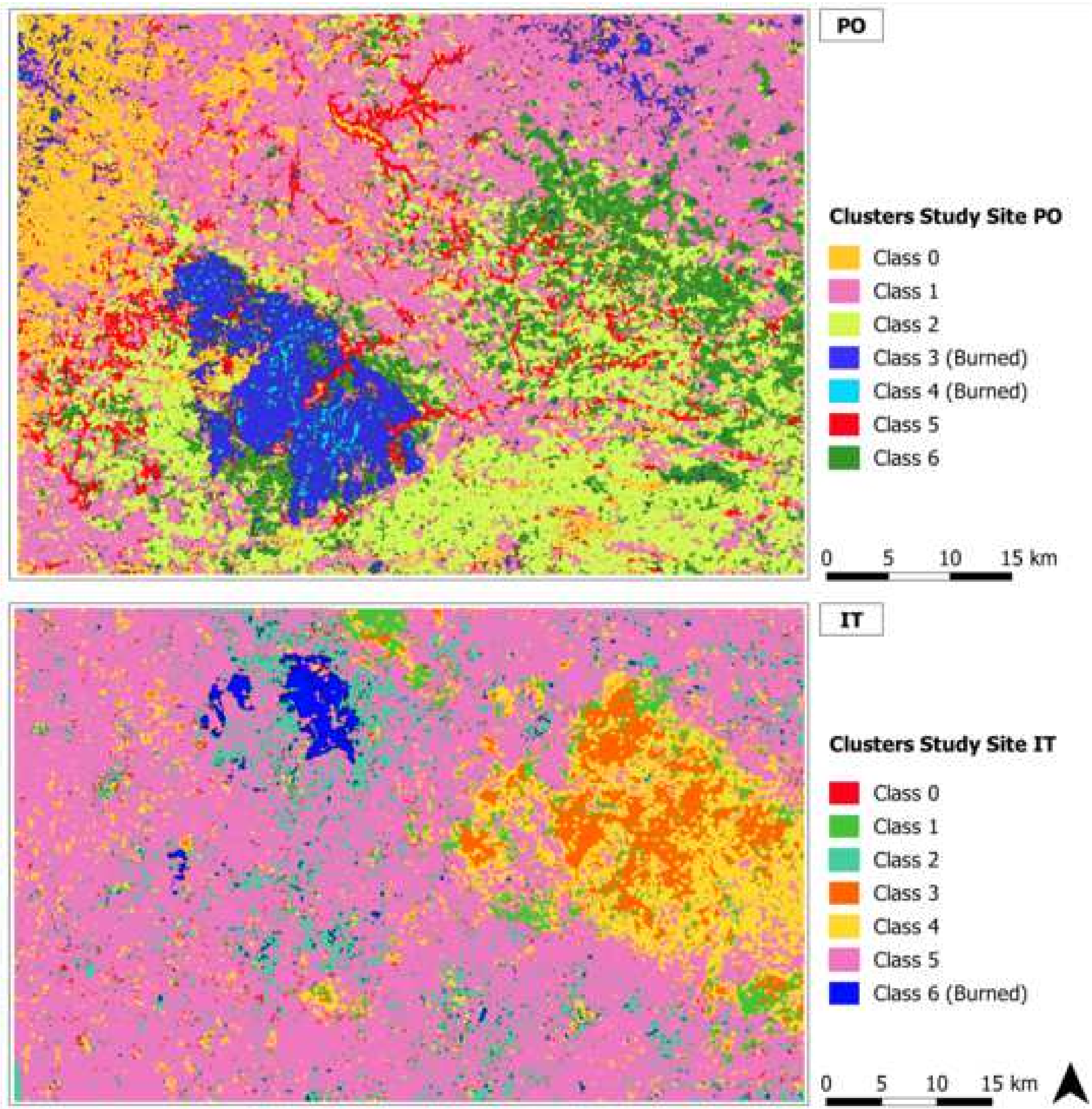


Figure 10 Click here to access/download;Figure;Figure 10.jpg

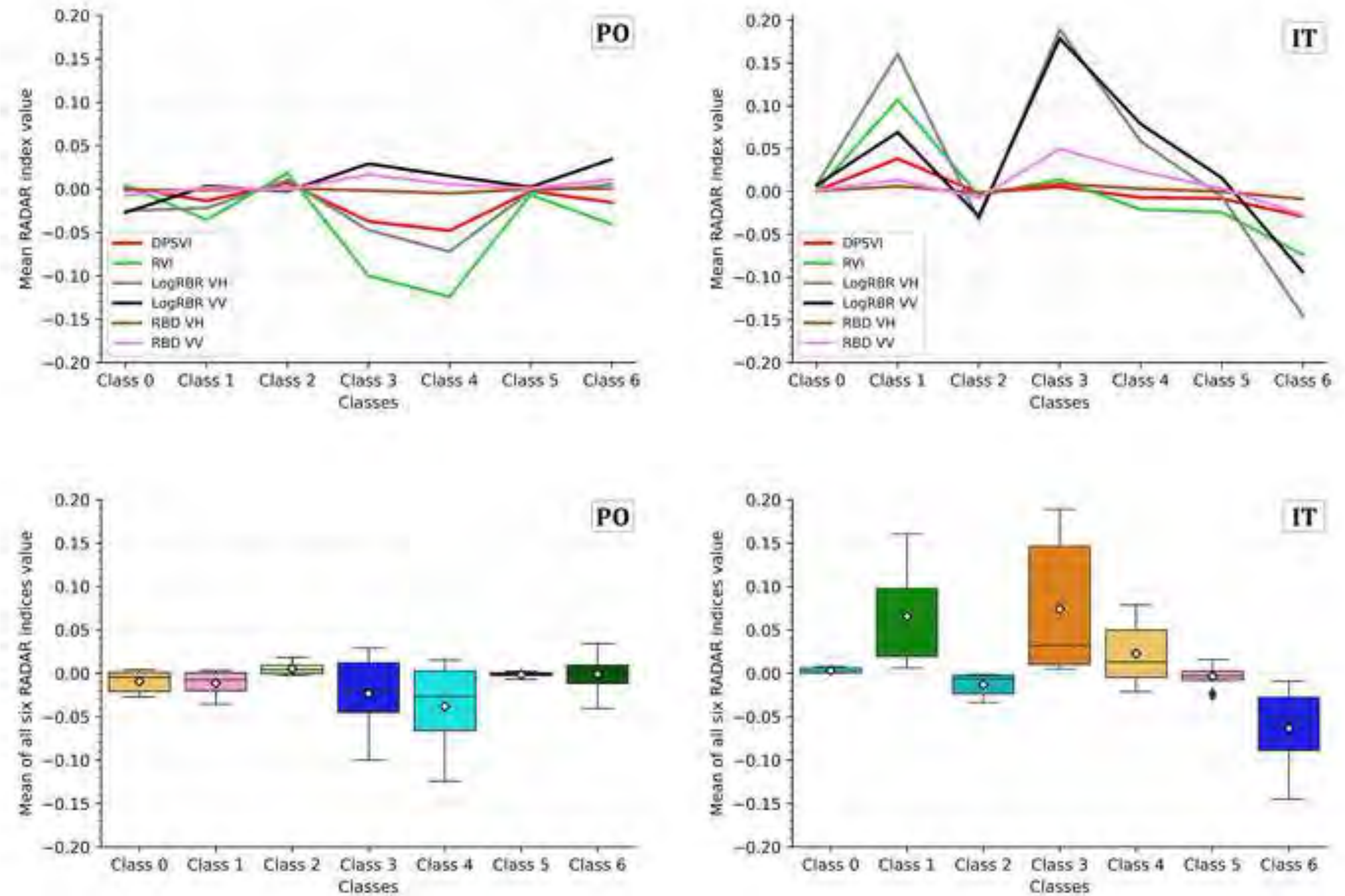
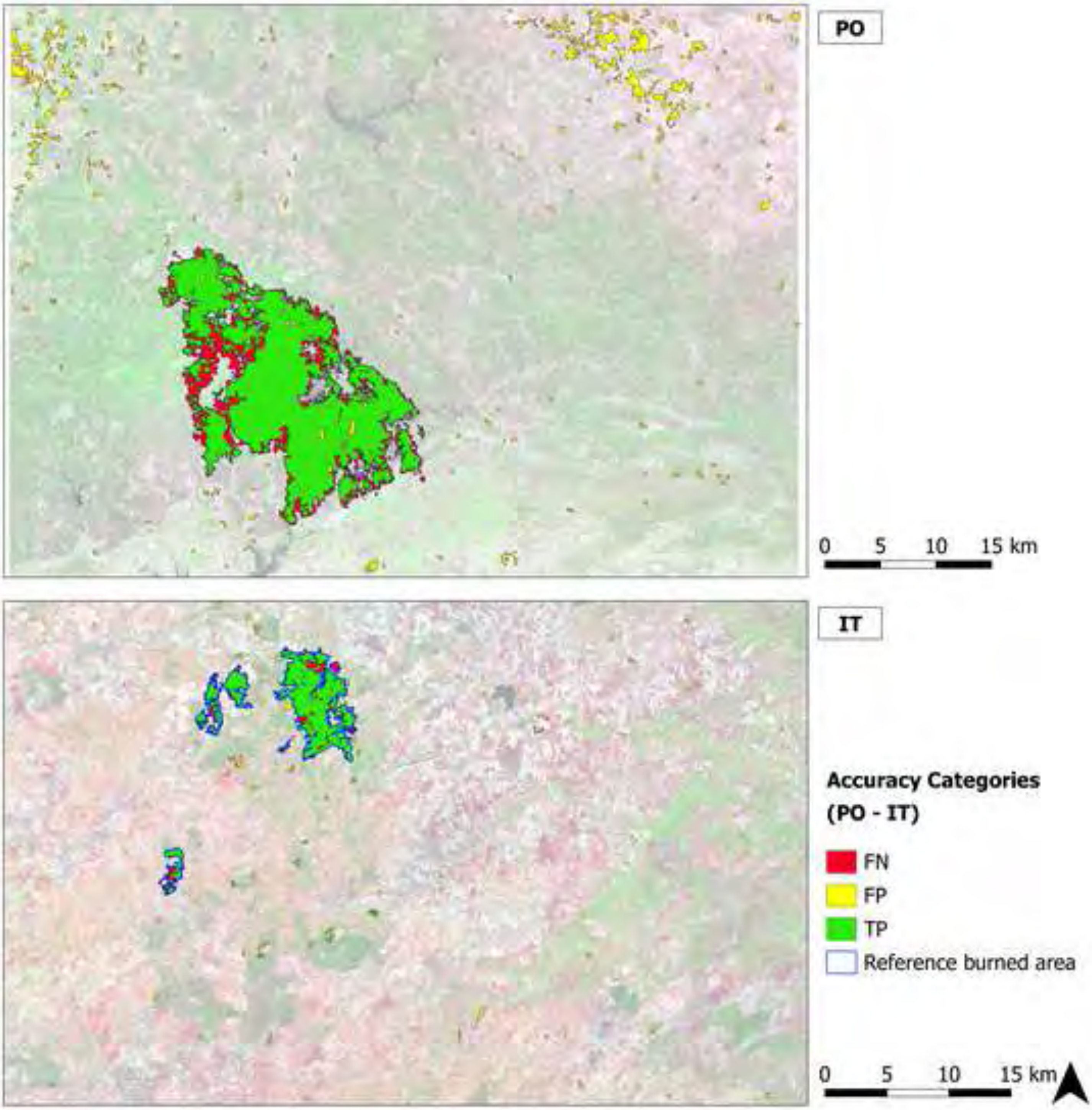


Figure 11 Click here to access/download;Figure;Figure 11.jpg



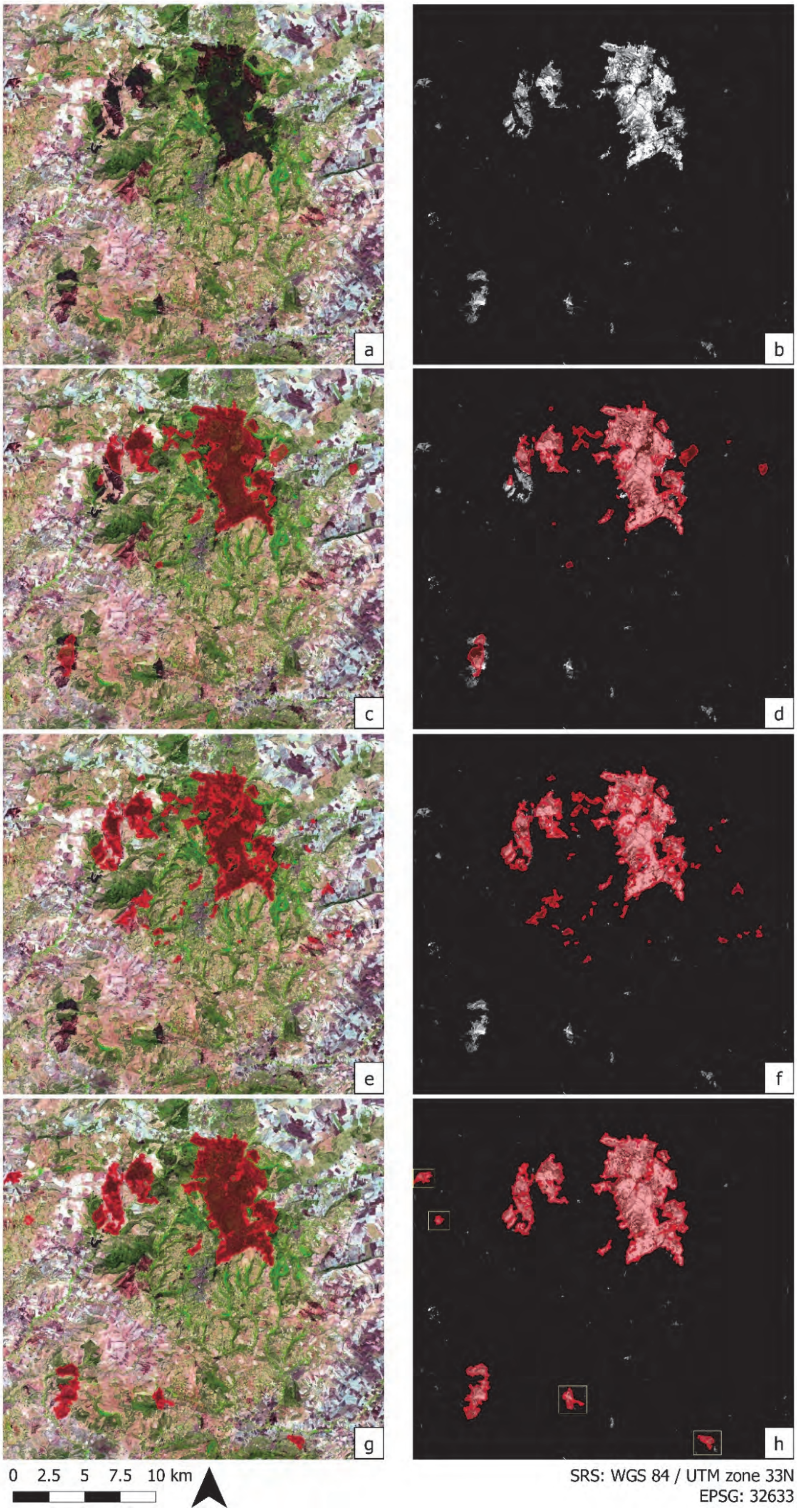


Fig. S1. The figure shows a subset of the IT study site focused on the burned area. The base maps are the Sentinel-2 false-color composite (SWIR-NIR-RED) (a; on the left side), the ΔNBR map where the light pixels have values greater than 0.1 (b; on the right side). The maps c and d show, superimposed, the official data of the Forest Information System (SIF) of Regione Sicilia, based on a visual interpretation of aerial and satellite data. Maps e and f show, superimposed, the official data "EMSR213 -Forest Fire in Southern Italy-Piazza Armerina-v1", provided by COPERNICUS Emergency Management Service as Rapid Mapping Service activated for Southern Italy in the fire season of 2017 (https://emergency.copernicus.eu/mapping/ems-product-component/EMSR213_16piazzaarmerina_02grading_monit01/1). Maps g and h show, superimposed, the reference data derived from ΔNBR and used in this study (§ 2.2.2). The four polygons in yellow (h) are not related to the main event (confirmation received from local authorities). For this reason, they are not considered in the reference map. As can be seen, the event's official data are not entirely consistent with the actual situation; therefore, they would lead to incorrect error detections. It can be seen, especially from the comparison with the ΔNBR map, that many areas burned have not been intercepted by official data (false negatives), particularly near the burned area's boundaries (Fig. 1 d-f). In contrast, others have been mistakenly considered to be part of the event (false positives). The latter are mainly related to fire events that occurred previously in the same fire season.

	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8	PC 9	PC 10	PC 11	PC 12	PC 13	PC 14	PC 15	PC 16	PC 17	PC 18	PC 19	PC 20	PC 21	PC 22	PC 23	PC 24	PC 25	PC 26	PC 27	PC 28	PC 29	PC 30
ΔDPSVI Dissimilarity	0,011	0,025	-0,117	-0,169	0,036	0,118	-0,088	0,005	0,251	0,164	-0,071	0,025	-0,067	0,152	-0,037	-0,023	-0,105	0,269	0,108	0,014	-0,316	-0,279	-0,225	0,333	-0,272	0,285	0,101	-0,452	-0,050	-0,057
ΔDPSVI Entropy	0,118	0,028	-0,215	-0,393	0,129	0,248	-0,061	-0,376	0,130	-0,401	-0,323	0,311	-0,104	-0,099	-0,045	-0,251	0,021	-0,066	0,113	-0,059	0,075	0,171	0,007	-0,046	0,072	0,123	0,062	0,120	-0,020	-0,007
ΔDPSVI Correlation	0,366	-0,092	0,135	0,045	0,067	0,190	0,090	0,017	0,044	-0,095	0,280	0,247	0,149	-0,030	0,157	0,253	-0,063	0,059	-0,193	0,222	0,432	0,279	0,066	0,247	-0,123	0,141	0,100	-0,237	-0,035	0,024
ΔDPSVI Mean	0,097	-0,077	-0,165	0,137	-0,011	0,023	-0,027	0,023	-0,073	0,115	-0,017	-0,062	0,088	0,093	0,050	0,036	0,164	0,062	-0,030	0,002	0,020	-0,032	-0,019	0,073	-0,184	0,369	0,309	0,557	-0,046	-0,528
ΔDPSVI Variance	0,024	-0,042	-0,134	0,091	-0,018	0,044	-0,004	0,053	-0,002	0,050	-0,089	-0,017	0,053	0,079	0,019	0,023	0,136	0,052	0,003	0,016	0,027	-0,082	-0,023	0,075	-0,193	0,189	0,200	0,312	0,167	0,820
ΔRVI Dissimilarity	0,023	0,019	-0,255	-0,235	0,033	0,085	-0,230	-0,079	0,240	0,314	0,015	-0,077	0,039	0,257	0,040	0,211	-0,270	0,424	-0,144	0,025	0,067	0,117	0,337	-0,236	0,133	-0,163	-0,077	0,147	0,028	0,017
ΔRVI Entropy	0,247	-0,077	-0,268	-0,125	0,095	-0,149	-0,295	-0,529	-0,334	0,080	0,149	-0,231	0,163	-0,063	-0,033	0,163	0,209	-0,192	-0,124	0,069	-0,033	-0,180	-0,092	0,082	-0,063	-0,154	-0,043	-0,130	0,016	0,027
ΔRVI Correlation	0,351	-0,112	0,066	0,121	0,071	0,096	0,051	-0,011	-0,097	0,115	0,281	0,143	-0,115	0,275	0,007	-0,659	0,234	0,186	-0,169	0,004	-0,156	-0,043	0,097	-0,120	0,044	-0,099	-0,044	-0,033	0,005	0,024
ΔRVI Mean	0,041	-0,118	-0,381	0,298	-0,083	0,111	0,015	0,154	-0,092	0,001	-0,283	0,035	0,174	0,041	0,014	0,111	0,182	-0,024	-0,201	-0,043	-0,120	0,110	-0,027	-0,148	0,497	0,309	-0,042	-0,317	-0,051	0,019
ΔRVI Variance	0,013	-0,072	-0,290	0,173	-0,046	0,226	0,071	0,198	0,092	-0,076	-0,342	0,123	0,075	0,074	0,064	0,029	0,231	0,014	0,000	0,053	0,223	-0,242	0,007	0,086	-0,315	-0,562	-0,137	-0,030	0,005	-0,121
LogRBR$_{VH}$ Dissimilarity	0,015	0,065	-0,203	-0,172	0,133	-0,114	-0,153	0,285	0,406	0,135	0,255	0,045	0,035	0,225	-0,110	-0,038	0,146	-0,569	-0,063	-0,249	0,108	0,138	-0,091	-0,097	-0,089	0,029	0,027	-0,053	0,037	-0,008
LogRBR$_{VH}$ Entropy	0,224	0,071	-0,251	-0,091	0,229	-0,489	-0,253	0,459	-0,236	-0,201	-0,002	0,234	0,045	-0,150	0,010	-0,084	-0,211	0,146	0,132	0,146	0,005	-0,124	0,063	0,021	0,027	0,009	-0,007	0,026	-0,006	0,003
LogRBR$_{VH}$ Correlation	0,357	-0,034	0,073	0,103	0,035	0,087	0,017	0,053	-0,210	0,198	-0,303	-0,084	-0,450	0,166	-0,507	0,105	-0,283	-0,198	-0,007	0,083	0,143	0,033	-0,049	-0,065	-0,074	0,045	-0,044	-0,001	-0,010	0,001
LogRBR$_{VH}$ Mean	0,030	0,215	-0,338	0,411	-0,121	-0,009	-0,048	-0,125	0,021	-0,128	0,188	0,050	-0,137	-0,108	-0,058	-0,014	-0,251	0,086	-0,075	-0,251	-0,256	0,431	-0,143	0,234	-0,165	-0,159	-0,088	0,076	0,127	-0,003
LogRBR$_{VH}$ Variance	0,013	0,179	-0,256	0,290	-0,069	0,151	0,070	-0,134	0,236	-0,103	0,287	-0,065	-0,145	-0,142	0,012	-0,089	-0,270	-0,203	0,059	0,159	0,152	-0,517	0,234	-0,040	0,193	0,083	0,124	0,012	-0,125	0,021
LogRBR$_{VV}$ Dissimilarity	0,012	0,076	-0,092	-0,231	-0,199	0,275	-0,156	0,229	-0,048	0,229	0,054	-0,091	-0,235	-0,317	0,279	-0,121	-0,017	-0,060	-0,166	0,357	-0,006	0,091	-0,460	-0,034	0,133	-0,110	0,055	0,112	-0,012	0,015
LogRBR$_{VV}$ Entropy	0,223	0,028	-0,024	-0,265	-0,397	0,189	-0,028	0,267	-0,078	-0,498	0,154	-0,498	0,036	0,124	-0,127	0,022	0,050	0,110	0,009	-0,168	-0,003	-0,026	0,082	0,008	-0,031	0,039	-0,038	-0,023	0,008	-0,004
LogRBR$_{VV}$ Correlation	0,365	-0,051	0,147	0,025	0,093	0,139	0,105	0,023	0,207	-0,103	0,157	0,251	0,098	-0,110	-0,100	0,413	0,046	0,055	0,097	-0,057	-0,383	-0,208	-0,287	-0,328	0,069	-0,101	-0,083	0,200	0,011	0,010
LogRBR$_{VV}$ Mean	0,035	0,412	0,094	0,082	0,004	-0,139	-0,143	-0,021	0,035	-0,062	-0,040	0,025	-0,345	-0,006	-0,030	0,188	0,392	0,205	-0,029	-0,083	0,069	0,040	0,040	-0,125	0,076	-0,173	0,557	-0,178	-0,081	-0,003
LogRBR$_{VV}$ Variance	0,016	0,394	0,090	0,042	0,135	0,080	-0,157	0,033	0,008	-0,056	-0,032	0,009	-0,276	0,055	0,217	0,161	0,302	-0,074	-0,060	0,171	-0,121	-0,029	0,247	0,097	-0,087	0,270	-0,564	0,087	0,075	-0,014
RBD$_{VH}$ Dissimilarity	0,005	0,040	-0,080	-0,067	0,072	0,023	0,132	0,050	0,100	0,121	0,125	-0,041	0,059	0,044	-0,368	0,027	0,257	0,068	0,340	0,221	-0,021	0,144	0,017	0,513	0,409	-0,153	-0,063	0,170	-0,184	0,035
RBD$_{VH}$ Entropy	0,060	0,141	-0,293	-0,232	0,308	-0,149	0,761	0,011	-0,062	0,044	0,023	-0,218	-0,147	-0,060	0,131	0,053	0,022	0,074	-0,141	-0,051	-0,001	0,020	-0,041	-0,066	-0,062	0,017	-0,014	-0,031	0,025	-0,001
RBD$_{VH}$ Correlation	0,364	-0,060	0,073	0,058	0,071	0,068	-0,046	0,024	-0,022	0,192	-0,136	-0,160	-0,083	0,005	0,536	0,014	-0,091	-0,121	0,454	-0,402	0,019	0,022	0,076	0,164	0,182	-0,081	0,048	-0,016	0,026	0,019
RBD$_{VH}$ Mean	0,024	0,073	-0,129	0,156	-0,022	-0,007	-0,016	-0,067	0,017	0,067	0,072	-0,161	0,110	-0,046	0,006	-0,088	0,097	0,109	0,500	0,194	0,122	0,260	-0,103	-0,433	-0,299	0,120	-0,143	-0,123	-0,395	0,063
RBD$_{VH}$ Variance	0,004	0,035	-0,051	0,054	-0,001	0,038	0,031	-0,029	0,041	0,027	0,058	-0,063	0,052	-0,024	-0,092	-0,043	0,078	0,054	0,319	0,228	0,100	0,043	-0,020	-0,134	0,053	0,050	0,058	-0,138	0,850	-0,150
RBD$_{VV}$ Dissimilarity	0,016	0,201	-0,066	-0,186	-0,245	0,217	0,069	0,094	-0,261	0,383	0,091	0,284	0,126	-0,433	-0,223	-0,004	0,095	0,010	0,060	-0,323	0,022	-0,070	0,305	-0,021	-0,128	0,079	-0,034	-0,061	0,016	0,000
RBD$_{VV}$ Entropy	0,226	0,250	-0,007	-0,113	-0,627	-0,362	0,217	-0,154	0,063	0,081	-0,141	0,241	0,116	0,287	0,141	0,000	-0,052	-0,151	0,033	0,203	-0,074	0,003	-0,011	0,016	0,018	-0,013	-0,010	0,025	-0,002	-0,003
RBD$_{VV}$ Correlation	0,323	0,069	0,171	0,099	0,023	-0,199	-0,038	-0,010	0,462	0,087	-0,302	-0,315	0,222	-0,451	-0,127	-0,215	0,049	0,050	-0,238	0,035	-0,021	0,045	0,103	0,055	-0,047	0,017	-0,041	0,011	0,020	-0,004
RBD$_{VV}$ Mean	0,012	0,495	0,110	0,060	0,125	0,071	-0,039	-0,057	-0,097	0,031	-0,039	-0,010	0,302	0,177	-0,058	-0,122	-0,099	0,177	-0,072	-0,244	0,397	-0,179	-0,447	0,021	0,167	0,060	-0,169	0,074	0,014	0,001
RBD$_{VV}$ Variance	0,007	0,370	0,084	0,004	0,253	0,307	0,020	0,092	-0,183	-0,026	-0,110	-0,095	0,388	0,148	0,019	-0,052	-0,186	-0,256	-0,012	0,233	-0,369	0,127	0,191	-0,027	-0,080	-0,149	0,288	-0,028	-0,030	0,002

Tab. S1. The table represents the eigenvector matrix resulting from principal component analysis (PCA) of the PO dataset. The columns represent the thirty principal components (PCs) (or eigenvectors), and the rows represent the thirty input layers. The eigenvectors matrix reports the statistical correlation between the input layers and the eigenvectors, indicating each input layer's proportion to each PC.

	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8	PC 9	PC 10	PC 11	PC 12	PC 13	PC 14	PC 15	PC 16	PC 17	PC 18	PC 19	PC 20	PC 21	PC 22	PC 23	PC 24	PC 25	PC 26	PC 27	PC 28	PC 29	PC 30
ΔDPSVI Dissimilarity	-0,003	0,006	-0,065	0,049	0,098	0,034	-0,061	-0,009	0,084	-0,045	0,175	0,026	0,080	0,072	-0,036	0,224	0,159	0,023	0,149	0,172	-0,139	0,269	-0,047	-0,038	-0,025	0,091	0,746	-0,338	-0,033	-0,149
ΔDPSVI Entropy	-0,028	0,022	-0,257	0,214	0,417	0,177	-0,379	-0,164	0,177	-0,094	0,183	0,066	-0,428	0,013	0,141	-0,123	0,235	0,099	0,048	-0,162	-0,127	-0,173	-0,225	0,071	-0,087	-0,039	-0,083	0,059	-0,014	0,001
ΔDPSVI Correlation	-0,407	-0,135	0,098	0,014	0,114	0,109	-0,023	0,101	-0,090	-0,289	0,130	0,285	-0,047	-0,245	0,080	-0,182	0,061	0,000	-0,012	0,005	0,213	0,030	0,608	0,096	0,018	-0,121	0,125	0,090	-0,081	-0,064
ΔDPSVI Mean	0,023	0,038	-0,011	0,164	-0,086	-0,024	0,042	0,000	-0,037	0,150	0,074	-0,130	0,055	-0,070	-0,014	0,020	-0,022	-0,043	0,031	-0,046	-0,037	0,062	-0,108	0,136	-0,032	-0,064	0,154	0,652	-0,275	-0,579
ΔDPSVI Variance	0,001	0,020	-0,001	0,099	-0,040	0,005	0,027	0,005	0,011	0,056	0,102	-0,076	0,011	-0,072	0,010	0,012	-0,007	-0,012	0,058	0,022	-0,013	0,012	-0,038	-0,007	0,074	0,019	0,324	0,475	-0,129	0,777
ΔRVI Dissimilarity	-0,009	0,013	-0,128	0,106	0,152	0,033	-0,149	-0,023	0,064	-0,010	0,188	0,018	0,237	0,120	-0,130	0,286	0,155	-0,005	0,098	0,356	-0,080	0,501	0,135	0,110	0,153	-0,110	-0,456	0,131	0,051	0,069
ΔRVI Entropy	-0,053	0,014	-0,262	0,251	0,256	-0,014	-0,425	-0,211	-0,249	0,287	-0,392	-0,078	0,221	-0,100	-0,022	0,007	-0,314	-0,128	-0,109	0,032	0,142	-0,027	0,174	-0,108	0,037	0,059	0,112	-0,051	0,004	0,004
ΔRVI Correlation	-0,383	-0,121	0,095	0,080	0,004	0,007	0,014	0,063	-0,347	-0,019	-0,111	0,103	0,329	-0,402	0,157	-0,041	0,245	0,147	0,077	-0,031	-0,212	0,062	-0,474	-0,051	-0,009	0,005	-0,091	-0,056	0,037	0,030
ΔRVI Mean	-0,016	0,076	0,004	0,442	-0,190	-0,007	0,100	0,034	-0,033	0,231	0,282	-0,175	-0,018	-0,158	-0,020	-0,061	-0,107	-0,043	-0,097	-0,202	0,035	0,176	0,060	0,399	-0,412	-0,159	-0,030	-0,258	0,163	0,100
ΔRVI Variance	-0,018	0,060	0,016	0,303	-0,105	0,084	0,070	0,012	0,132	0,067	0,403	-0,265	-0,051	-0,294	0,067	-0,140	-0,122	0,114	-0,047	0,263	0,055	-0,126	0,033	-0,470	0,341	0,165	-0,093	-0,110	-0,003	-0,093
LogRBR$_{VH}$ Dissimilarity	-0,010	-0,013	-0,192	0,029	0,106	-0,217	0,031	0,011	0,133	-0,250	0,236	0,101	0,316	-0,181	-0,366	0,392	-0,056	0,015	-0,320	-0,235	-0,099	-0,399	-0,001	0,058	0,053	-0,008	0,002	0,002	-0,010	0,001
LogRBR$_{VH}$ Entropy	-0,034	-0,112	-0,486	-0,037	0,012	-0,745	0,124	0,073	-0,132	-0,061	0,077	0,001	-0,151	-0,006	0,062	-0,216	-0,040	0,081	0,199	0,131	0,049	0,067	-0,009	0,001	-0,021	-0,007	-0,002	-0,010	0,002	-0,001
LogRBR$_{VH}$ Correlation	-0,398	-0,093	0,008	0,027	-0,022	-0,031	0,031	0,089	-0,157	0,229	-0,016	-0,368	-0,090	0,386	0,134	0,346	0,186	0,410	-0,051	0,001	0,046	-0,263	0,168	0,085	0,051	0,018	0,000	-0,011	-0,015	0,003
LogRBR$_{VH}$ Mean	-0,129	0,395	-0,022	0,313	-0,254	-0,191	-0,027	0,080	-0,076	-0,057	-0,089	0,297	-0,156	0,210	0,149	0,133	0,233	-0,218	-0,288	-0,261	-0,022	0,195	0,035	-0,271	0,146	0,123	-0,005	0,045	0,056	-0,008
LogRBR$_{VH}$ Variance	-0,105	0,444	0,074	0,330	-0,148	0,001	-0,025	0,034	0,086	-0,233	-0,238	0,193	0,063	0,132	-0,225	-0,107	-0,161	0,292	0,280	0,342	-0,045	-0,262	-0,045	0,136	-0,094	-0,083	0,008	-0,009	-0,040	0,002
LogRBR$_{VV}$ Dissimilarity	-0,006	-0,010	-0,169	0,011	0,034	0,141	-0,072	0,317	0,064	-0,024	0,150	0,048	0,333	0,257	-0,011	-0,186	0,145	-0,010	-0,084	0,015	0,681	-0,027	-0,277	-0,046	-0,144	0,101	0,016	0,028	-0,002	0,000
LogRBR$_{VV}$ Entropy	-0,039	-0,102	-0,433	0,006	-0,168	0,242	0,001	0,593	0,354	0,034	-0,352	-0,082	-0,071	-0,207	0,034	0,059	-0,018	0,007	0,006	-0,042	-0,195	0,064	0,080	0,023	0,038	-0,005	-0,002	-0,006	0,002	0,000
LogRBR$_{VV}$ Correlation	-0,402	-0,113	0,061	-0,021	0,050	0,044	0,036	0,011	0,056	-0,167	0,036	0,051	-0,324	0,028	-0,019	0,330	-0,591	-0,102	0,022	0,011	0,198	0,211	-0,339	0,005	-0,005	0,075	-0,034	0,017	-0,006	-0,003
LogRBR$_{VV}$ Mean	-0,124	0,342	-0,052	-0,254	-0,066	-0,062	-0,148	-0,055	0,078	-0,211	0,000	-0,352	0,068	-0,095	0,299	0,068	0,030	-0,219	-0,134	0,150	0,065	-0,054	-0,082	-0,078	-0,059	-0,606	0,063	-0,043	-0,013	0,009
LogRBR$_{VV}$ Variance	-0,097	0,395	0,040	-0,219	0,078	-0,039	-0,183	0,051	-0,023	-0,258	0,034	-0,423	0,105	-0,071	-0,115	-0,141	-0,052	0,130	0,206	-0,387	-0,017	0,223	0,095	0,087	0,054	0,402	-0,043	0,011	0,012	-0,005
RBD$_{VH}$ Dissimilarity	-0,007	0,028	-0,027	0,029	0,085	0,019	0,107	0,005	0,046	0,010	0,058	0,030	0,124	-0,002	0,153	0,256	-0,023	-0,107	0,389	-0,094	-0,058	-0,099	0,193	-0,483	-0,611	0,080	-0,140	0,077	-0,078	0,014
RBD$_{VH}$ Entropy	-0,060	0,161	-0,181	0,146	0,438	0,151	0,721	-0,213	0,075	-0,103	-0,229	-0,129	0,065	0,038	0,067	-0,073	0,059	-0,023	-0,111	-0,021	0,054	0,093	-0,003	0,033	0,070	0,003	0,018	-0,002	0,016	0,000
RBD$_{VH}$ Correlation	-0,406	-0,117	0,018	0,005	0,009	0,013	0,008	0,034	0,008	0,067	0,024	-0,205	-0,057	0,213	-0,529	-0,298	0,151	-0,487	0,009	0,088	-0,229	-0,094	-0,029	-0,134	-0,075	-0,016	0,000	0,001	0,036	0,005
RBD$_{VH}$ Mean	-0,026	0,091	-0,002	0,055	-0,015	-0,025	0,056	0,013	0,026	0,176	0,025	0,054	0,013	-0,071	0,076	0,141	0,109	-0,374	0,420	-0,130	0,165	-0,186	-0,024	0,245	0,325	0,009	-0,116	-0,275	-0,503	0,043
RBD$_{VH}$ Variance	-0,009	0,035	0,008	0,022	0,017	0,005	0,032	0,016	0,027	0,076	0,029	0,033	0,025	-0,059	0,075	0,123	0,035	-0,200	0,389	-0,067	0,118	-0,200	0,009	0,093	0,228	-0,033	0,107	0,175	0,775	-0,086
RBD$_{VV}$ Dissimilarity	-0,025	0,082	-0,170	-0,075	0,065	0,195	0,047	0,187	-0,239	0,020	0,318	0,151	0,246	0,373	0,305	-0,199	-0,397	-0,078	-0,006	-0,074	-0,403	-0,093	0,020	0,060	0,134	-0,051	0,008	-0,019	-0,012	0,000
RBD$_{VV}$ Entropy	-0,040	0,149	-0,492	-0,246	-0,378	0,381	0,094	-0,319	-0,350	0,030	0,114	0,107	-0,146	-0,103	-0,240	0,080	0,071	0,093	0,068	-0,005	0,102	0,002	0,006	-0,046	-0,023	-0,011	-0,004	0,002	0,005	-0,001
RBD$_{VV}$ Correlation	-0,354	-0,022	-0,077	-0,160	-0,219	-0,091	-0,030	-0,417	0,588	0,275	0,011	0,183	0,254	0,048	0,151	-0,167	-0,047	0,119	-0,012	-0,102	-0,032	0,059	0,020	0,038	0,003	0,075	0,002	0,017	0,012	-0,002
RBD$_{VV}$ Mean	-0,086	0,347	0,015	-0,268	0,164	-0,047	0,022	0,134	-0,018	0,263	0,084	0,120	-0,114	-0,214	0,117	0,072	0,062	-0,165	-0,243	0,425	-0,054	-0,142	0,005	0,245	-0,209	0,426	-0,027	0,054	0,042	-0,004
RBD$_{VV}$ Variance	-0,054	0,292	0,104	-0,190	0,311	-0,063	0,055	0,242	0,017	0,473	0,076	0,204	-0,101	-0,085	-0,293	-0,031	-0,060	0,258	0,050	-0,191	0,053	0,121	-0,028	-0,229	0,081	-0,373	0,034	-0,022	-0,012	0,003

Tab. S2. The table represents the eigenvector matrix resulting from principal component analysis (PCA) of the IT dataset. The columns represent the thirty principal components (PCs) (or eigenvectors), and the rows represent the thirty input layers. The eigenvectors matrix reports the statistical correlation between the input layers and the eigenvectors, indicating each input layer's proportion to each PC.

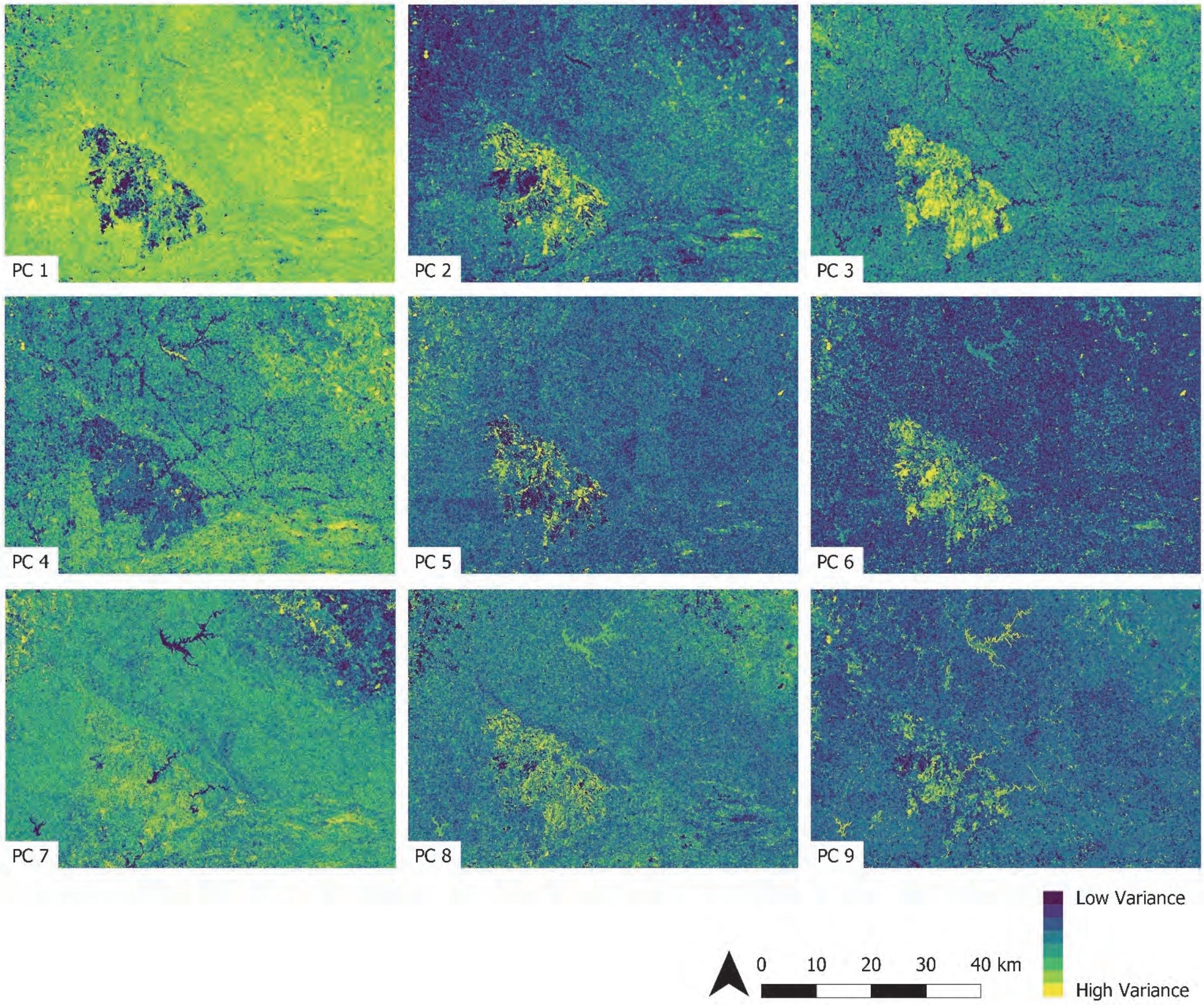


Fig. S2. The image shows all the first principal components (PCs), deriving from the principal component analysis (PCA) of dataset PO, that reached a cumulative variance ≥ 99%. The variance values are expressed by the color palette in each image, with the darkest color expressing the lowest variance and the lightest color expressing the highest variance for each respective image. The first PC, which represents the maximum proportion of the entire original dataset's variance, differentiates with great contrast the area affected by the fire with lower variance values than the rest of the scene (lowest values of the maximum proportion of the variance).

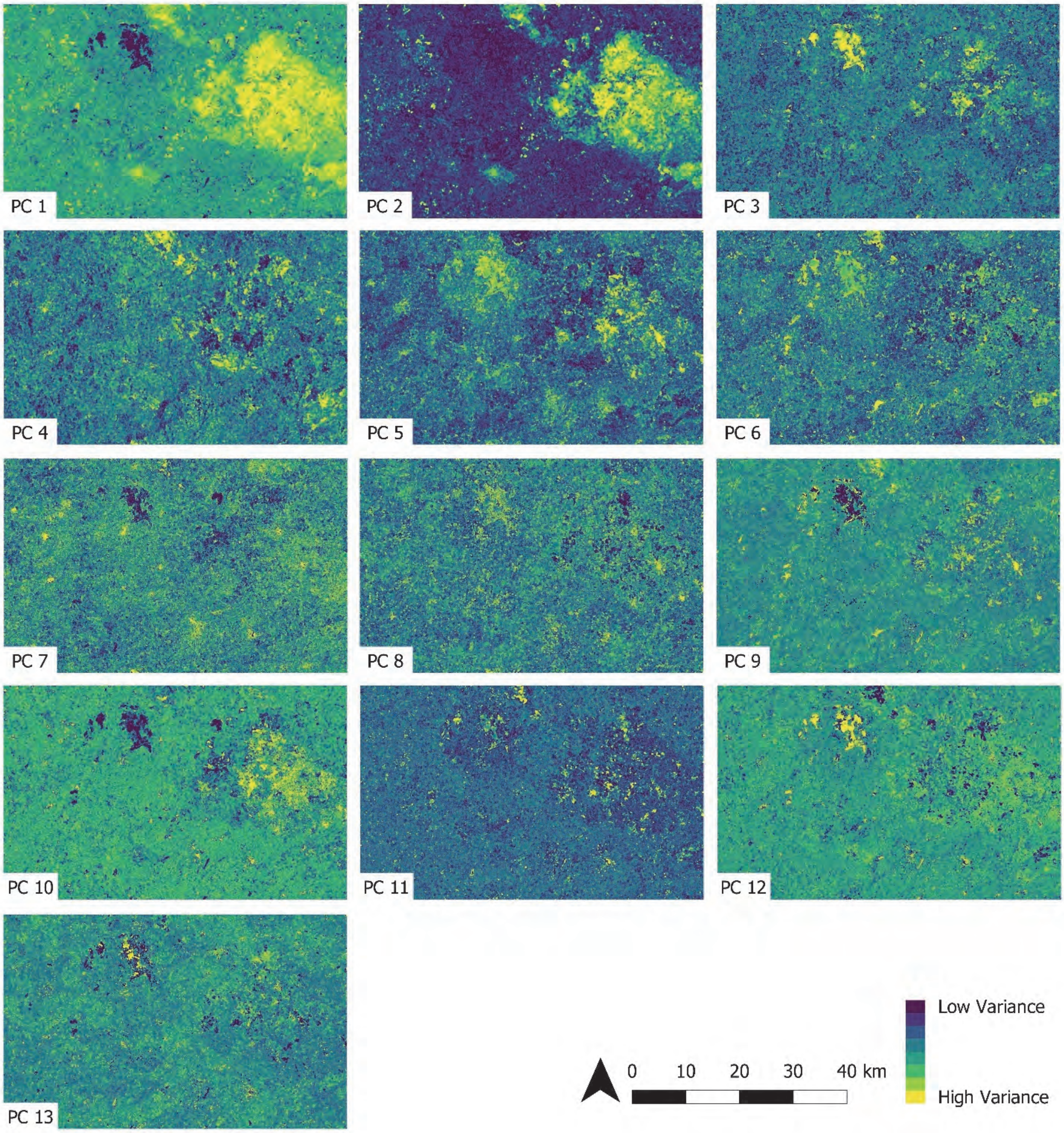


Fig. S3. The image shows all the first principal components (PCs), deriving from the principal component analysis (PCA) of dataset IT, that reached a cumulative variance ≥ 99%. The variance values are expressed by the color palette in each image, with the darkest color expressing the lowest variance and the lightest color expressing the highest variance for each respective image. The first PC, which represents the maximum proportion of the entire original dataset's variance, differentiates with great contrast the area affected by the fire with lower variance values than the rest of the scene (lowest values of the maximum proportion of the variance).

Classes	PO	IT
Class 0	7,843,107	1,573,744
Class 1	1,472,788	5,144,170
Class 2	1,187,595	451,379
Class 3	4,873,240	8,872,372
Class 4	10,488,656	2,793,112
Class 5	4,187,654	1,888,440
Class 6	5,866,831	4,712,221

Tab. S3. The table shows the number of pixels that fell into each of the seven classes (i.e., clusters).

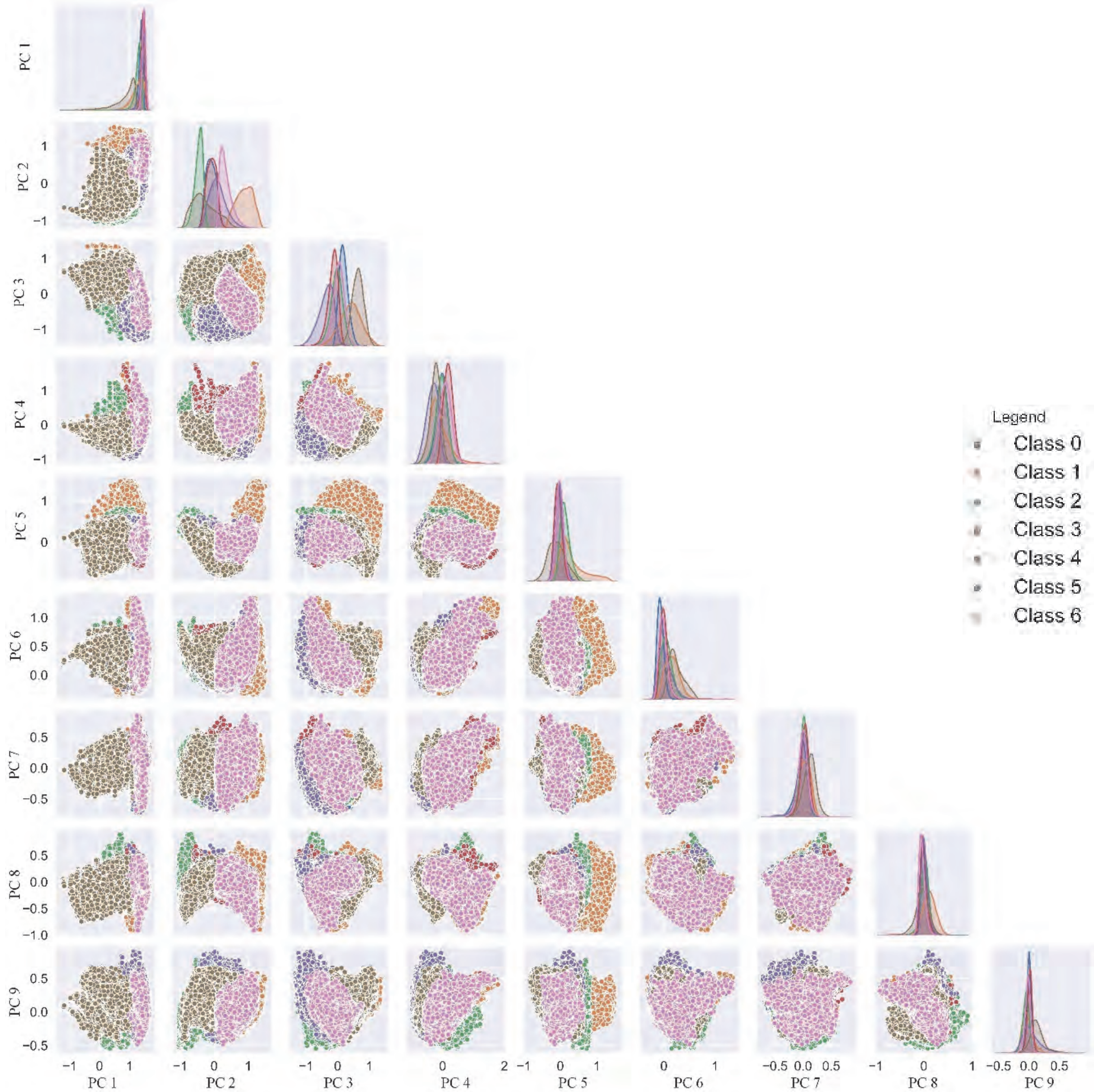


Fig. S4. Pair plot showing the clusters resulted from the k-means classification of the PO dataset.

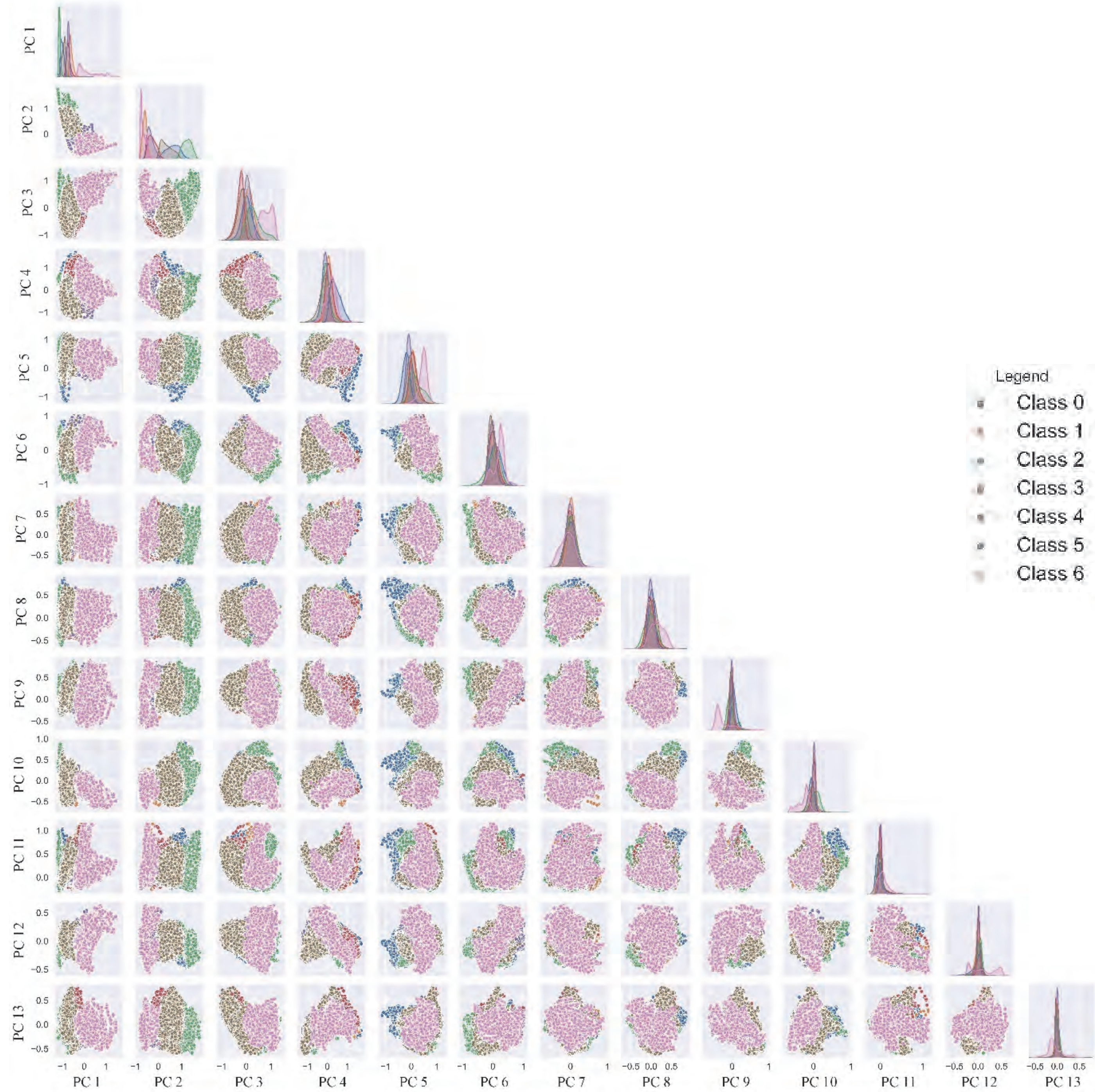


Fig. S5. Pair plot showing the clusters resulted from the k-means classification of the IT dataset.