

**POLITECNICO DI MILANO**

**Geoinformatics Engineering Master of Science**

**School of Civil, Environmental and Land Management Engineering**

**School of Industrial and Information Engineering**



**Implementing an algorithm based on soft computing for Burned  
Area mapping from Sentinel-2 data in a user-friendly QGIS  
workflow**

**Supervisor:** Assist. Prof. Giovanna Sona

**Co-Supervisor:** Dr. Daniela Stroppiana

**Thesis by:** Ruken Dilara Zaf

**Student ID:** 926368

**Academic year 2020/2021**

# Table of Contents

1. Introduction	1
2. Background Information	4
2.1. Remote sensing	4
2.2. Stages in GIS Development	11
2.3. R programming	13
3. Literature Review on BA mapping	15
4. Methodology	20
4.1. R and QGIS Integration	22
4.2. Theoretical Definitions and Code Modifications for Integration	27
4.2.1. Fuzzy Membership Function-Computation of Membership Degrees (MDs)	28
4.2.2. Computation of Ordered Weighted Averages and BA maps creation	31
4.2.3. Calculation of Accuracy Metrics	33
5. Test Area and Data	38
5.1. Description of Test Area and fire season	38
5.2. Geospatial Data Used for Burned Area Detection on Manavgat, Turkey fire event	40
6. Results and Discussion	51
6.1. MD and OWA results	51
6.1.1. Partial Evidence (MD) results	51
6.1.2. Global Evidence (OWA) results	56
6.2. Validation Results	59
7. Conclusion	65
8. BIBLIOGRAPHY	66

## Table of Figures

Figure 1. Burn severity (low-medium-high) detected in different forest types (California (CA) chaparral, Montana (MT) mixed-conifer forests, and Alaska (AK) black spruce forests) (Lentile et al., 2006) .....	2
Figure 2. Schematic description of remote sensing process .....	5
Figure 3. Active and Passive remote sensing .....	5
Figure 4. Scheme of three main types of RS vehicles: satellites, airplanes, and Unmanned Aerial Vehicles (UAVs) (Heisig, 2021).....	6
Figure 5. Electromagnetic spectrum regions and wavelengths .....	7
Figure 6. Spectral signatures as function of wavelengths for five typical surfaces (Zhuge & Zou, 2017).....	8
Figure 7. Major GIS use options (URL5).....	12
Figure 8. ESA average global BA product: MODIS FireCCI51 with subsequent extensions to include BA for the years 2018 to 2020. ( <a href="https://geogra.uah.es/fire_cci/firecci51.php">https://geogra.uah.es/fire_cci/firecci51.php</a> , last accessed on March 2022)....	16
Figure 9. Comparison of the MODIS BA products (FireCCI50, FireCCI51, and MCD64A1 c6) against the Sentinel-2 product (FireCCISFD11) for the year 2016. a) Northern Hemisphere Africa (NHAF) and b) Southern Hemisphere Africa (SHAF) (Lizundia-Loiola et al., 2020) .....	19
Figure 10. Fire areas in Mediterranean region used for : the parameter selection (b: Vesuvius National Park, Italy), integration phase (a: Pedrógão Grande, Portugal) and validation step (c: Manavgat, Antalya, Turkey) of the algorithm .....	21
Figure 11. Plugin panel-Processing R Provider window .....	23
Figure 12. (a) Options icon (b) Icon for R script creation .....	23
Figure 13. QGIS window to set up the important paths between QGIS and some main providers .....	24
Figure 14. QGIS Processing Toolbox interface .....	26
Figure 15. QGIS input/output interface example before the run .....	26
Figure 16. Process flow (from Sentinel-2 imagery to burned area mapping) .....	27
Figure 17. Step by step Input/Output for each part of the algorithm .....	28
Figure 18. Dzetsaka classification tool interface for QGIS .....	34
Figure 19. Forest Fires in Turkey between the years 1990-2018 (Republic of Turkey Ministry of Urban, Environment and Climate Change) .....	39
Figure 20. Detailed event map for Turkey Forest fires, 02/08/2021 (ECHO/European Commission) .....	39
Figure 21. Sentinel-2 images for pre and post fire dates in Manavgat, Turkey (prefire: 20/07/2021 and postfire: 29/08/2021).....	40
Figure 22. Turkey burnable land cover types .....	42
Figure 23. Distribution of percentage of polygon numbers with different land covers .....	46
Figure 24. Distribution of percentage of burnable areas w.r.t assigned codes .....	46
Figure 25. Classified CLC vector of Manavgat (as burnable or not) .....	47
Figure 26. Pre-and post-fire Planet images (ground truth) .....	48
Figure 27. Classified and filtered reference vector (from post-fire Planet image) .....	49
Figure 28. Sigmoid function results on Portugal site: MDs calculated for each selected band: PostRE2 (a), PostRE3 (b) PostNIR (c), $\Delta RE2$ (d), $\Delta RE3$ (e), $\Delta NIR$ (f) and $\Delta SWIR2$ (g). .....	52
Figure 29. Portugal: Density distribution of different input layers. MD1=PostRE2, MD2=PostRE3, MD3=PostNIR, MD4= $\Delta RE2$ , MD5= $\Delta RE3$ , MD6= $\Delta NIR$ and MD7= $\Delta SWIR2$ .....	53
Figure 30. Sigmoid function results on Turkey site: MDs calculated for each selected band: PostRE2 (a), PostRE3 (b) PostNIR (c), $\Delta RE2$ (d), $\Delta RE3$ (e), $\Delta NIR$ (f) and $\Delta SWIR2$ (g).....	54

Figure 31. Turkey: Density distribution of different input layers. MD1=PostRE2, MD2=PostRE3, MD3=PostNIR, MD4= $\Delta$ RE2, MD5= $\Delta$ RE3, MD6= $\Delta$ NIR and MD7= $\Delta$ SWIR2.....	55
Figure 32. Portugal: OWA maps with different weighting vectors: OWA_and(a), OWA_almost_and (b), OWA_average (c), OWA_almost_or (d), OWA_or (e) .....	57
Figure 33. Turkey: OWA maps with different weighting vectors: OWA_and(a), OWA_almost_and (b), OWA_average (c), OWA_almost_or (d), OWA_or (e) .....	58
Figure 34. Percentage of omission errors for five different OWA operators.....	61
Figure 35. Percentage of commission errors for five different OWA operators.....	61
Figure 36. OWA_or map (excluding the pixels with values smaller than 0.1) and Planet image of the area .....	62
Figure 37. Agreement maps by OWA maps and reference vector of Portugal: Agreement_and (a), Agreement_almost_and (b), Agreement_average (c), Agreement_almost_or (d), Agreement_or (e) ..	63
Figure 38. Agreement maps by OWA maps and reference vector of Turkey: Agreement_and (a), Agreement_almost_and (b), Agreement_average (c), Agreement_almost_or (d), Agreement_or (e) ..	64

## List of Tables

Table 1. Sentinel-2 band details .....	10
Table 2. R Syntax in Processing scripts (URL12) .....	25
Table 3. Sampled error/confusion matrix: $n_{ij}$ express the number of pixels of agreements (diagonal cells) or disagreements (off diagonal cells) between the BA product and the reference PLANET image. In the case of binary classification (burned areas and not burned areas) considering as target objective of the algorithm the identification of Burned areas we can make this equivalence: True Positives (TP= $n_{11}$ ), True Negatives (TN= $N_{22}$ ), False Positives (FP = $n_{12}$ ) and False Negatives ( FN = $n_{21}$ ). .....	35
Table 4. Metrics computed from the error/confusion matrix and range of variability .....	35
Table 5. Landcover types in the CLC(2012) product.....	44
Table 6. Accuracy Metrics of Portugal fire .....	59
Table 7. Accuracy Metrics of Turkey fire .....	59

## ABSTRACT

The idea of this thesis is implementing a Burned Area (BA) mapping algorithm into a QGIS workflow. The approach for burned area identification is based on soft computing techniques and on the integration of features (spectral bands and indices) derived from Sentinel-2 multispectral satellite images with ordered weighted average operators (OWAs). The algorithm was previously implemented only in R environment and the core objective of this thesis was to exploit the QGIS user-friendly environment for displaying and visualisation of the results. Further, the workflow created and implemented in QGIS is first tested on the previously analysed fire site Pedrógão Grande, Portugal (2017), and then run over a new Mediterranean fire site to assess the exportability of the algorithm output. The new test region is chosen among as one of the most recent fire events in Antalya, Turkey (2021). In both cases, BA maps are compared to fire reference perimeters derived from very high-resolution Planet images for accuracy assessment. In Portugal, for all applied ordered weighted averages (OWAs), the omission error (oe), ranging in [8,3-48,5%], is greater than the commission error (ce), [0.3-6.4%]. On the other hand, in Turkey, while the omission errors are in a similar range [3,6-60%] as Portugal, commission errors are significantly greater especially for OWA\_almost\_or (19,5%) and OWA\_or (23,2%). This increase leads to a decrease of the Dice coefficient with the lowest value of 0.57 for the OWA\_and operator. Despite that, the OWA\_average operator, which results with the best values in Turkey, still provides an acceptable oe (8.9%) and ce (11.1) with a high Dice coefficient (90%). Finally, results of the previous study on Mediterranean regions with the same algorithm (Sali et al., 2021) and our results together show that the algorithm is successfully implemented in QGIS and it is suitable to detect burned areas in Mediterranean region even without further calibration.

## ABSTRACT

L'obiettivo principale di questo lavoro di tesi è di implementare un algoritmo di mappatura delle aree bruciate in QGIS. L'approccio per il rilevamento delle aree percorse dal fuoco, che si basa sulla teoria dei sistemi fuzzy e sulle tecniche di soft computing, integra bande e indici spettrali derivati da immagini satellitari Sentinel-2, attraverso l'utilizzo di operatori OWA (ordered weighted average operators). L'algoritmo, già disponibile e codificato in R, è stato implementato come flusso di lavoro in QGIS per poter sfruttare le avanzate funzionalità di visualizzazione e processing di dati geospaziali che offre un ambiente GIS rispetto ad un linguaggio di programmazione. L'implementazione è stata verificata su un evento di incendio già precedentemente analizzato (Pedrógão Grande, Portogallo, 2017) e successivamente è stata testata su un nuovo caso di evento di incendio in Turchia (Antalya, 2021). Per entrambi i casi di test, le mappe prodotte di area bruciata sono state confrontate con perimetri di riferimento derivati da immagini ad alta risoluzione Planet per la validazione. In Portogallo, la validazione ha evidenziato errori di omissione e commissione variabili a seconda dell'operatore OWA considerato nel range [8,3-48,5%] e [0.3-6.4%], rispettivamente. In Turchia, gli errori di omissioni sono comparabili [3,6-60%] mentre la commissione è significativamente maggiore soprattutto per il caso degli operatori OWA\_almost\_or (19,5%) e OWA\_or (23,2%). L'operatore OWA\_and ha evidenziato valori più bassi del coefficiente di Dice tra tutti gli operatori testati (DC=0.57). Per il caso test in Turchia, l'operatore che ha fornito globalmente i risultati migliori è stato OWA\_average con oe=8,9%, ce=11,1% e Dice=90%. I risultati prodotti in questa tesi, combinati con evidenze già riportate in letteratura (Sali et al., 2021) hanno confermato che l'algoritmo implementato in QGIS è in grado di rilevare con soddisfacente accuratezza le aree percorse dal fuoco in ambiente Mediterraneo in modo semi-automatico.

## **ACKNOWLEDGEMENTS**

Foremost, I would like to express my sincere gratitude to my supervisor Assist.Prof. Giovanna Sona and co-supervisor Dr. Daniela Stroppiana for their consistent support, guidance, and the comments during the running of this study. Besides, I would like to thank to Assoc.Prof. Giovanna Venuti for her support and offering my name for this work.

I am very grateful to my lovely friends both in Turkey and in Milan. Their love, support and belief helped me not only in this period but also anytime I hit a stone wall.

Finally, I would like to express my deepest love and thanks to my family. They never stopped believing in me and encouraged me on the way of all my wishes.

## 1. Introduction

Fire is an integral and fundamental process of the Earth System as it affects the distribution of vegetation, the biogeochemical cycle, the landscape, and the climate. A fire regime, in general, describes intensity, frequency, size, severity (the degree of damage that fires cause to the landscape) and the seasonal patterns of fires in a local/regional area. The changes in the fire regimes are expected to continue during the coming years and decades, with potentially wide-ranging and severe impacts on different ecosystems (Rogers et al., 2020).

Despite fire being a natural phenomenon and a key disturbance for many ecosystems, the general increasing trend in the number of fires and affected surface areas in the Mediterranean region is spectacularly serious. The reasons for this increase are mainly the land-use (landcover) change and climate warming. The changes in the traditional land use by socio-economic transformations and lifestyles have led an increase in accumulated fuel. According to many regional studies, the rural depopulation and the increased cover of forest and shrublands (increased plant biomass and fuel load) is one of the main reasons of more frequent fires and increased hazard in Mediterranean areas (Chergui et al., 2018; Chuvieco, 1999; San-miguel-ayanz et al., 2012). Secondly, climate change increases drought frequency and extends the fire season. Obviously, climate change alone is insufficient to explain current fire-regime changes for wildfires as human behaviour (e.g., ignition patterns) is also important. However, the weather factors affecting the fire occurrence (wind, high temperatures, and low humidity) are potentially affected by climate change (Pausas & Keeley, 2021). These phenomena, together with anthropogenic impacts, place the Mediterranean forest ecosystem in an alarming situation.

The results of wildfires with thousands of hectares of forest areas burned every year and with serious economic damage and landscape changes also contribute to soil erosion, habitat degradation, and release of greenhouse gases (GHGs). By all its effects on the ecosystem balance, wildland fires are accepted as an Essential Climate Variable (ECV) by the Global Climate Observing System (GCOS). Thus, burned area (BA) detection is an important active research topic studied for many ecosystems. To evaluate the damages, point the post-fire management and implement medium to long-term restoration strategies, a timely and accurate detection of burned



areas is essential (Luca et al., 2021; Ramo et al., 2018). One of the most important parameters to quantify fire damage on an area is generally the fire severity.

Fire or burn severity (Figure 1) can be defined as the degree of the fire effect both on vegetation and soil (Rahman & Chang, 2017). The evaluation of fire severity can possibly be made by post-fire in-situ field surveys; however, because this approach requires enormous effort and budget, alternative solutions are gaining importance. Changes in spectral reflectance by vegetation removal acquired by remote sensing images are potential choices for this purpose both for local and global classification. Several studies showed that there is a high correlation between vegetation indices derived from post-fire data and burn severity (Santis & Chuvieco, 2007).

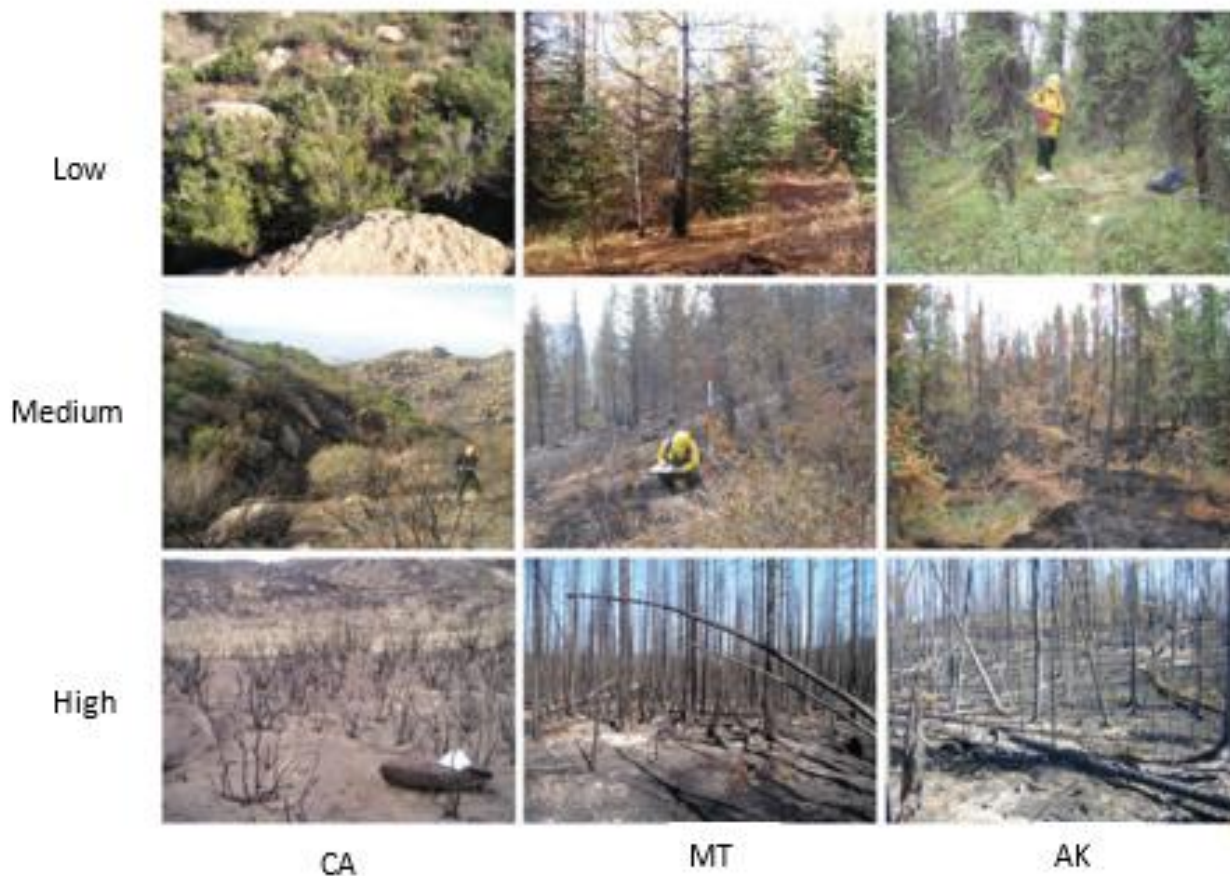


Figure 1. Burn severity (low-medium-high) detected in different forest types (California (CA) chaparral, Montana (MT) mixed-conifer forests, and Alaska (AK) black spruce forests) (Lentile et al., 2006)

On the other hand, using remote sensing images and data are very useful on fire mapping, there are many challenges to create algorithms and interpret the results of this approach. A local map can offer relatively high accuracy after the validation by a reference data, however using the same algorithm on another area may give completely ineffective results. In parallel with, creating extensive coverage maps, regional or global, can be extremely problematic. The independent variables may lead to completely different results in different areas because of the diverse conditions. This also arises a related problem which is lack of automation. Because the process generally includes determination of a trade-off between over-estimation (commission) and under-estimation (omission), when the variables affecting these rates have different impacts, an automatic procedure can fail without a recalibration between diverse locations. Furthermore, fire mapping being time sensitive also poses an obstacle for automation (Goodwin & Collett, 2014).

The idea of this study is integration of a BA algorithm which is proved that giving highly accurate results in various Southern European fire sites (Sali et al., 2021) into a new and a more user-friendly environment, and testing the algorithm's exportability.

## 2. Background Information

### 2.1. Remote sensing

The electro-magnetic (EM) energy, reflected or emitted by surfaces, can give meaningful information on the physical characteristics of specified objects or of an area. Remote Sensing (RS) (Figure 2) is the process of detecting and recording the EM energy reflected and/or emitted by a surface by using sensors on board platforms at a distance from the observed surface (e.g., satellites or aircraft).

Another definition of RS is that it is the science of collecting the spatial data about Earth's features from a distance by using the sensors in order to map, monitor and manage the information and create meaningful results (Coops & Tooke, 2017).

The information carrier in RS is the electromagnetic radiation which is measured either by passive or active sensors (Figure 3). Passive sensors (e.g., the sensors used in LANDSAT, SENTINEL missions) are those which sense natural radiations generated by the Sun, either reflected or emitted from the Earth. On the other hand, the sensors which produce their own electromagnetic radiation, are called active sensors (e.g. the sensors used in LIDAR, RADAR). The ensemble of data and techniques from RS is also called "Earth Observation" (EO) that helps researchers for decision making in many areas such as agriculture, natural resource management, forestry, urbanization, and transportation.

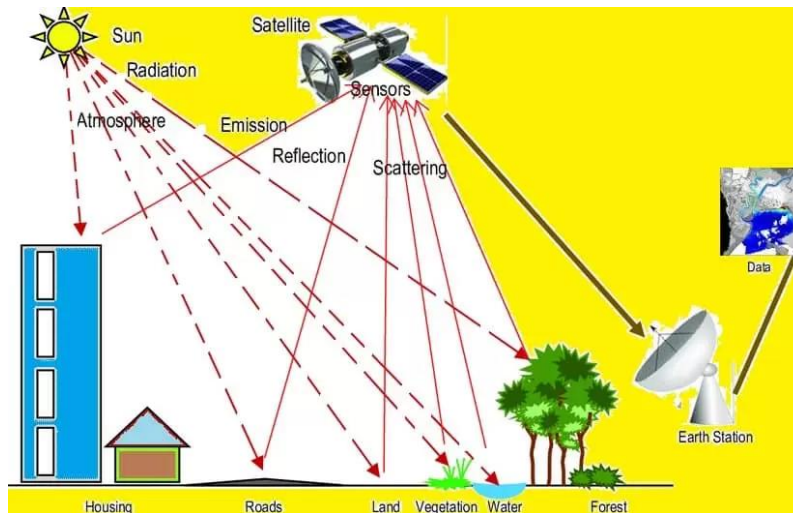


Figure 2. Schematic description of remote sensing process

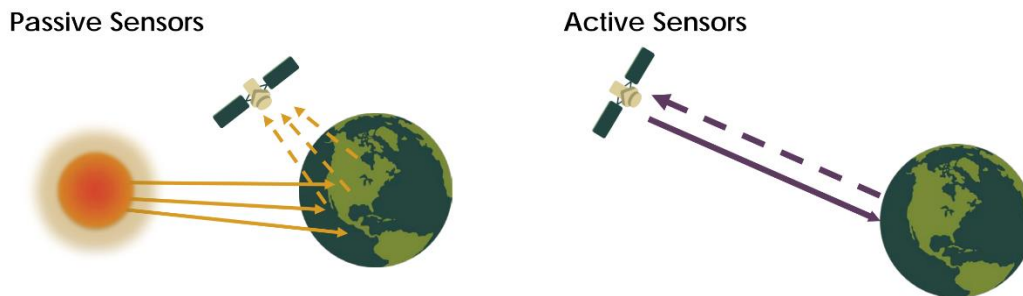


Figure 3. Active and Passive remote sensing

The sensors used in remote sensing can be classified also based on the vehicles that carry them. Airplanes, satellites, and unmanned aerial vehicles (UAVs) (Figure 4) have specialized platforms that carry sensors and each of them has its own advantages and disadvantages on their flight restrictions, image resolution and coverage. For example, UAVs have flexible availability, relatively low cost, and very high spatial resolution. They can be very effective for imaging a small area; however, with their short flight time, the surveyed area is limited, and they can easily be affected by weather conditions. Airplanes can serve again high spatial resolution (higher than satellites) while their cost can be very high, and their availability depends on weather conditions. Satellites can acquire clear and stable images and detect larger area within each image. Also, they are a great resource for historical data collection. Generally, the resolutions and revisit times

depend on the satellite and the specific sensor on it. The drawbacks of the satellite imagery are that the clouds may hide ground features, the cost can increase dramatically for high resolution images and data may not be collected at critical times because of the fixed schedule (Chang & Clay, 2016).

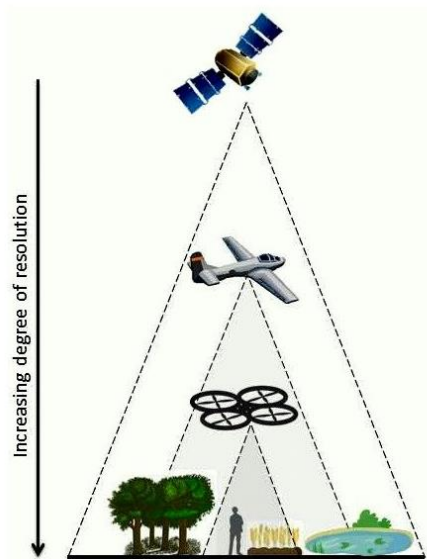


Figure 4. Scheme of three main types of RS vehicles: satellites, airplanes, and Unmanned Aerial Vehicles (UAVs) (Heisig, 2021)

The resolution of an image refers to the potential details provided by the imagery. In remote sensing, we refer to four types of resolution: spatial, spectral, radiometric, and temporal.

**Spatial Resolution** refers to the size of the ground area of each pixel that can be detected by a satellite sensor or displayed in a satellite image. For example, a spatial resolution of 250 m means that one pixel of the image represents an area 250 by 250 meters on the ground. The ability to "resolve," or separate, small details is one way of describing what we call spatial resolution.

**Spectral Resolution** refers to the ability of a satellite sensor to collect EM radiation separating specific wavelengths of the electromagnetic spectrum. For example, panchromatic indicates it accepts all colours, meaning the band has a very wide signal range. Multispectral indicates that the sensor has the capability to accept signal in various narrower bands separately.

**Radiometric resolution refers to sensor** sensitivity to the magnitude of the electromagnetic energy. The finer the radiometric resolution of a sensor, the more sensitive it is to detecting small differences in reflected or emitted energy. It determines how finely a system can represent or

distinguish differences of intensity and is usually expressed as several levels or a number of bits, for example 8 bits or 256 levels that is typical of digital image resolutions.

**Temporal resolution** refers to the time between two successive acquisitions of the same area with the same geometry of acquisition (revisiting time). The capability for satellites to provide images of the same geographical area more frequently has increased dramatically since the dawn of the space age, ranging from some days to few hours (URL1).

Remote sensing can also be broadly classified as optical and microwave.

In the electromagnetic spectrum, the wavelengths ranging between approximately 1 cm to 1 m belongs to the microwave portion. Microwaves have some special properties by their long wavelengths. This region of wavelengths is not susceptible to atmospheric scattering caused by cloud cover, haze, dust, and rainfall, while shorter optical wavelengths are highly affected by them. This property allows detection of microwave energy under almost all weather and environmental conditions so that data can be collected at any time (URL2). In addition, microwave sensors have the ability for day/night operation independent of sun or illumination conditions. As a result of its all-weather capability, the practice area of microwave RS is demonstrated for applications limited due to weather conditions like agriculture crop monitoring on a regular basis, flood mapping, disaster management (Mohan, 2015).

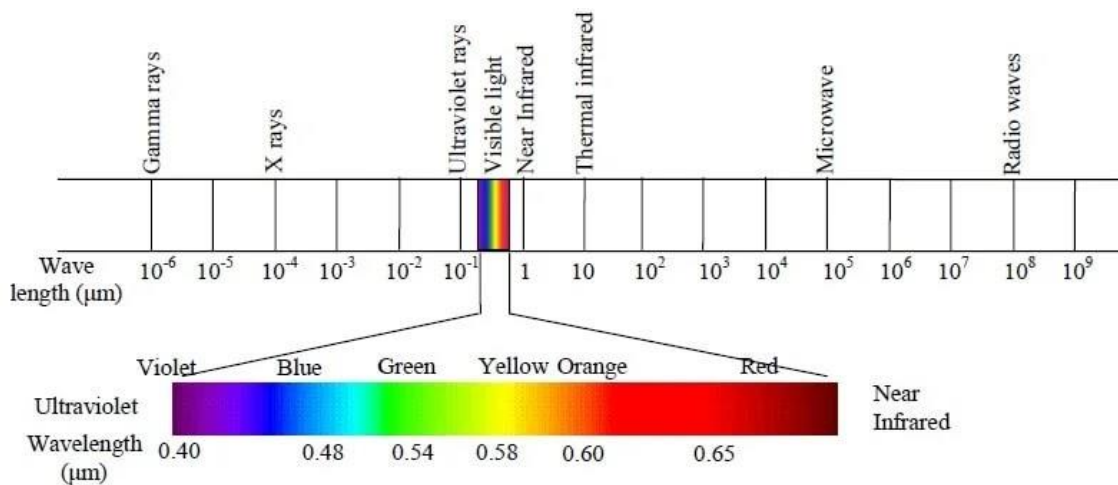


Figure 5. Electromagnetic spectrum regions and wavelengths

Optical remote sensing works with the sensors which detect the wavelengths from visible, near-infrared and middle-infrared up to thermal-infrared region (Figure 5) of the solar radiation. When solar radiation hits a target surface, it can be transmitted, absorbed, or reflected.

Spectral signature (Figure 6) as the basis of RS is a representation of the reflected radiation as a function of wavelength because different surfaces reflect radiation differently in various regions. Thus, a material can be identified from its spectral response (signature) at different wavelengths. For example, green vegetation has a very specific signature as the chlorophyll absorbs mainly the red light in the visible range and near-infrared light is reflected. Therefore, the difference between visible and near-infrared channels are great for vegetation while it is insignificant for bare soil.

There are several variables affecting spectral response even for the same target type. For example, seasonal changes on vegetation (e.g., green-ness of leaves) or the location of the same kind of surface can give different signatures (e.g. by the effect of different level of urbanisation). Knowing where to "look" spectrally and understanding the factors which influence the spectral response of the features of interest are critical to correctly interpreting the interaction of electromagnetic radiation with the surface (Mohammed et al., 2017).

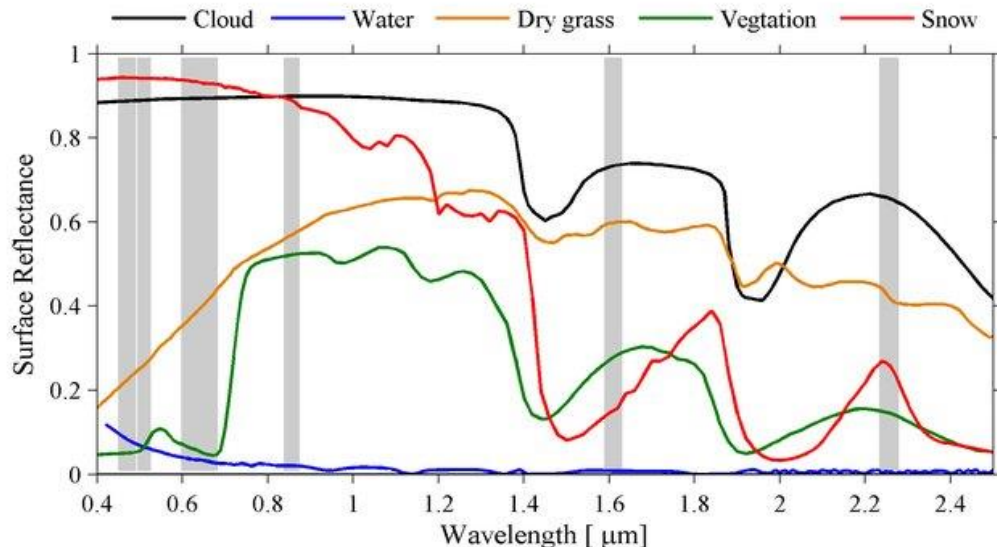


Figure 6. Spectral signatures as function of wavelengths for five typical surfaces (Zhuge & Zou, 2017)

Optical RS systems can be classified depending on the number of spectral bands used in the process such as panchromatic, multi-spectral, and hyper-spectral imaging systems. The main differences between panchromatic, multispectral and hyperspectral data acquisition are the width and number of recording bands. Whereas panchromatic sensors work with a single wide recording band, multispectral image consists of several relatively narrower bands of which two or three bands are in the visible range and there may also be some in near-infrared or middle infrared bands. A hyperspectral image consists of a large number of narrower bands spread across the electromagnetic spectrum. Its primary goal is to obtain a spectrum from each pixel in the image of a scene to find objects, detect processes, or identifying materials.

There are several multispectral imaging sensors used in different satellite missions. They have various resolutions from coarse to high and can have different numbers of bands. Medium-resolution satellite imagery is typically 10- to 30-meter resolution, gathered by satellites like the Landsat missions or Sentinel-2. They are widely used in Earth Observation because they are cost-effective, not data-intensive, regularly updated, and a good source providing data and information of crops, soil, and ground cover for agricultural application and production. Also, because they follow a consistent path and shorter revisit availability, they provide a high possibility to find a good image.

The aim of Sentinel-2 mission is monitoring variability in land surface conditions, and its wide swath width (290 km) and high revisit time (10 days at the equator with one satellite, and 5 days with 2 satellites under cloud-free conditions which results in 2-3 days at mid-latitudes), make its contribution to Copernicus themes such as climate change, land monitoring, emergency management, and security, highly significant (URL3). Sentinel-2 platform carries a multispectral sensor (MSI) covering 13 spectral bands (Table 1), from 443 to 2190 nm, visible and near-infrared bands are available at a 10 m spatial resolution, red-edge/shortwave-infrared bands at 20 m, while three atmospheric correction bands have 60 m resolution (Cogato et al., 2020). The 13 spectral bands guarantee consistent time series, showing variability in land surface conditions and minimising any defects introduced by atmospheric variability (URL4).

There are several advantages of multispectral Sentinel-2 data over other available monitoring satellites. By the access of Sentinel Hub, the Sentinel-2A imagery can be freely acquired (<https://scihub.copernicus.eu/>). Additionally, for image processing, the ESA created a tool named



Sentinel Application Platform (SNAP). It is a software specialized for Sentinel images and has a forum for discussions and consultations (Isbaex & Coelho, 2021).

*Table 1. Sentinel-2 band details*

<b>Spectral Band</b>	<b>Central Wavelength(nm)</b>	<b>Band width(nm)</b>	<b>Spatial Resolution(m)</b>
B1 (Coastal aerosol)	443	20	60
B2 (Blue)	490	65	10
B3 (Green)	560	35	10
B4 (Red)	665	30	10
B5 (Red-edge 1)	705	15	20
B6 (Red-edge 2)	740	15	20
B7 (Red-edge 3)	783	20	20
B8 (Near infrared)	842	115	10
B8a (Near infrared narrow)	865	20	20
B9 (Water vapor)	945	20	60
B10 (Shortwave infrared/cirrus)	1380	30	60
B11 (Shortwave infrared 1)	1910	90	20
B12 (Shortwave infrared 2)	2190	180	20

## 2.2. Stages in GIS Development

A Geographic Information System (GIS) is a system which links the concepts of “where” and “what”. It connects data to map, and helps users understand patterns, relationships, and geographic context (URL5). The first example of this linkage and idea of using spatial analysis takes us all the way back to 1854 and a cholera outbreak. An English Doctor, Dr. John Snow, did not accept the belief that the disease was being spread through the air. He discovered a pattern and demonstrated the potential of maps by identifying the link between an outbreak location of cholera and a contaminated water supply. John Snow’s cholera map was a major event connecting the what with the where and showed that GIS is a problem-solving tool (URL6).

Handling the data and coarse information was an issue so this problem accelerated the development of computer mapping. In the 1960s, Roger Tomlinson had created the Canadian Geographic Information System (CGIS) which was a unique system to handle maps by adopting layering approach (*One-Stop GIS , Stages in GIS Development, 2021*). The system provided accurate and relevant data which enabled Canada to begin its national land-use management program.

During the same period, one of the first automated computer mapping software programs, namely SYMAP, was created in Northwestern University by Howard fisher. It included spatial analytic capabilities applied to spatially distributed data. Then in 1965, Fisher established the Harvard Laboratory for Computer Graphics, and this Lab became a research centre for spatial analysis and visualization (URL5). After these seminal works, GIS was not anymore an experimental tool, it started to be an important part of the cartographic world (URL7).

GIS industry solidified its strength in the seventies. The foundation and products of Environmental System Research Institute (ESRI) have played a key role in the popularization of GIS. ESRI initially specialized in using GIS to give consultation services for land-use planning and resource management. In 1981, the firm developed the first commercial product, ARC/INFO. Then, in the late eighties, the first open-source GIS, GRASS (Geographic Resources Analysis Support System), appeared and it reached production status and supported both raster and vector data.

After the release of first commercial products, other companies also developed their own products and made GIS more powerful and user-friendly over the next ten years. With the development of internet and powerful computing tools, it became easier to gather and distribute data, so the functionality of GIS increased effectively.







 Identify problems	 Monitor change
 Manage & respond to events	 Perform forecasting
 Set priorities	 Understand trends

Figure 7. Major GIS use options (URL5)

In today's world, people can create their own digital maps to support solving real-world problems. Most of the organizations compare many different types of information by sharing their works, revealing the patterns and relationships. GIS can be counted both as a tool, technique or software for industry and a science for researchers by many tools for theory, evaluation and understanding. This variety of fields of usage induced the rise of open-source GIS software packages which are free, and the source code is openly available.

Quantum GIS (QGIS), as one of the most commonly used open-source GIS software, is an official project of the Open Source Geospatial Foundation (OSGeo) which is volunteer-driven. QGIS, with its nature, accepts contributions from users in the form of code, tool development, bug fixes, etc. This ability of all users to access and verify code promotes a high degree of quality control while the proprietary software packages are just available for the software developers and not to users (Flenniken et al., 2020).

The functionality of QGIS is accounted as the best one among the open source packages when compared with commercial ArcGIS by ESRI (WHO, 2018). QGIS design is built on plug-in architecture. This simplifies adding new functions or creating tools for specific applications and the archive is constantly expanding by contributions. Except for the plug-in archive, there are several standard tools for vector and raster analysis embedded in QGIS. The increasing number of plug-ins for QGIS suitable for processing and classification satellite images confirms the widespread use of this software by the RS community.

As an advantage of QGIS, the integration with the other open-source platforms is fully supported. As in our study, some symbiosis with other softwares can be very helpful to conduct sophisticated geospatial tasks. For example, while R is one of the best environments for processing and algorithm implementation, QGIS is the best for visualization of the results and preparing map layouts. Consequently, QGIS and R association has been counted as an important option (Muenchow et al., 2017).

### 2.3. R programming

As a programming language and software environment, R is a very useful tool for data analysis, and statistical computing. Also, it has been gaining high importance as a reference software tool for many different Research & Development areas.

R is a free software and can be used on various operating systems (UNIX, Windows and MacOS). It is available under the GNU General Public License. The major motivations for using R are its simplicity and availability for high level programming. The syntax is not complex, and it has a lot of powerful functions for doing mathematics and statistics. Further, the debugging is relatively easy because R is not a compiled language, that means you can execute one line at a time to verify the calculations (Wallach et al., 2019). Another main advantage of R is its flexibility to integrate with other languages. This allows to easy access for the user to many data sources and packages. By all these important properties, it is currently one of the most requested programming languages and it is growing day by day (URL8).

In the context of geographic data analysis, with the availability of large amounts of spatial data, the demand for high performance computers and upgradeable software has been widely increased. Handling this data requires flexible analyse, visualization and modelling. Neither a standard statistical package nor a standard GIS tool can achieve all these requirements by oneself (Brunsdon & Comber, 2019). R supports advanced geospatial statistics and being an interactive interface to other software, it can easily build a bridge to GIS software (URL9). This union can help to create more user-friendly and easier handling of geospatial data.

And in this work, R codes have been integrated in QGIS to implement the Burned Area (BA) mapping algorithm. This way we can exploit the user-friendly interface of QGIS and the full potentiality of the R programming language.

### 3. Literature Review on BA mapping

Remote sensing has been an important tool for BA mapping since the beginning of satellite observations. Because historical estimates done by fire management teams (field surveys) have been including differences in the methods and discontinuity in data collection, satellite imagery by reaching reliable information, permitting rapid, cost-effective, temporally constant coverage and monitoring of large and less accessible regions became a sound alternative (Chuvieco et al., 2019; Luca et al., 2021).

At the beginning, because of the impact of fire disturbance on climate, BA products were generally used by climate modelers. Then, by the increase in accuracy and continuous data collection, many different communities (e.g., civil protection services, environmental and forest protection services etc.) started to use these data. One of the main information provided by BA maps is the level of fire severity. Since the fire severity is an important factor that influences vegetation recovery, the primary post-fire management actions need this information to be taken. Also, it is used by researchers to study a variety of fire effects, including spatial patterns and trends in fire severity over time (Chuvieco et al., 2019; Lydersen et al., 2016).

The existing RS based BA products can be separated broadly as global and regional products. To develop global BA products, mainly two sensors have been chosen in the last years: VEGETATION (VGT), on board the SPOT satellites since 1998, and Moderate Resolution Imaging Spectroradiometer (MODIS), on the board the Terra and Aqua satellites since 2000. The first global BA product named Global Burned Area (GBA2000) was developed based on daily VGT images during the year 2000 and provided monthly estimates of BA with 1km spatial resolution. To detect the BA, seven regional algorithms were adapted for different fire conditions because methods of detection of burnt area vary for different ecosystem types. These algorithms were different by their region of developments and spectral indices used with possible different thresholds (Chuvieco et al., 2018; Tansey et al., 2004). After the first product, from VGT images, several following global products have been released: the L3JRC covering the period between the years 2000 and 2007, the Globcarbon for the period from 1998 to 2007 and several Copernicus products (the last one released in July 2020; with 300 m resolution by using Sentinel-3 data) (<https://land.copernicus.eu/global/products/ba> , last accessed March 2022).

Furthermore, NASA has also generated an archive of global BA products, starting with MCD45A1 using 500 m resolution MODIS imagery and it was standard NASA BA product from 2000 to 2016. In this product, in order to distinguish the significant differences between observed and predicted reflectance data, they used Bidirectional Reflectance Disturbance Function (BRDF). Then, by improving the algorithm by also using thermal anomalies, they created the recent set of products, namely MODIS MCD64A1 (Chuvieco et al., 2018, 2019). The current version of this algorithm (collection 6) provides considerably more sensitivity than the original and identifies 26% more global BA than the previous collection (Giglio et al., 2018). Recently, one of the main projects on global products namely Fire\_CCI, has been generated by the European Space Agency’s Climate Change Initiative. The first product was based on 300-m resolution MERIS images from the ENVISAT satellite, and the algorithm was also using MODIS hotspots as a supplement. The most recent product (FireCCI51 for the period 2001-2020) (Figure 8) were using the MODIS sensor’s red (R) and near-infrared (NIR) reflectances and thermal anomaly data, thus providing the highest spatial resolution (250 m) among the existing global BA datasets.

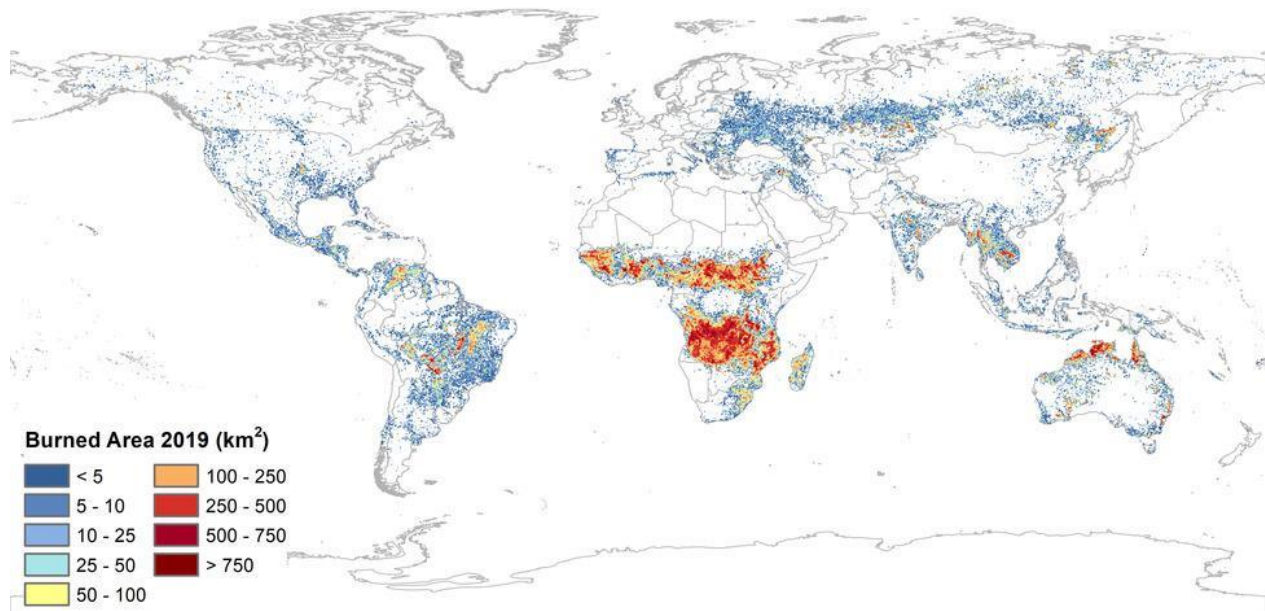


Figure 8. ESA average global BA product: MODIS FireCCI51 with subsequent extensions to include BA for the years 2018 to 2020. ([https://geogra.uah.es/fire\\_cci/firecci51.php](https://geogra.uah.es/fire_cci/firecci51.php), last accessed on March 2022)

On a regional basis, many countries have been producing BA maps for their fire monitoring systems. Several studies have been conducted by researchers to increase the accuracy on BA detection and automatization of the algorithms. While in global BA mapping, the coarse resolution sensors are employed; medium-to high resolution sensors are preferred in regional scale. A high percentage of recent studies are focused on the images derived from Landsat-8 and Sentinel-2 missions.

According to a study on exploring the utility of Sentinel-2 MSI and Landsat-8 OLI in BA mapping of a heterogenous savanna landscape, the highest contribution to burned area detection is gained by NIR, Red, Red-edge and Blue spectral bands in the studied ecosystem. The classification is done by Random Forest (RF) classifier. The study carried on two different sites and while Sentinel-2 gives higher accuracy in one site, Landsat-8 had a higher accuracy on the other site. However, according to their results, although Landsat 8 OLI has a low spatial and temporal resolution as compared to Sentinel-2, it can still be utilized in burn area mapping (Ngadze et al., 2020). Another study (Fitriana et al., 2018) showed that in mountain regions (Java Island in Indonesia), Landsat-8 imageries give the highest sensitivity in the NIR band and Normalized Burn Ratio long (NBR<sub>L</sub>) to detect the burned area. According to another research on Landsat-8 dataset in Indonesia, using TIRS, SWIR and NBR combination can ease the differentiation of burnt and unburnt area in rural while highlighting that this combination does not give good results in urban areas and needs to be combined with land use data (Indratmoko & Rizqihandari, 2019). In a different study conducted in a temperate region (Southern Europe), with single-date Landsat TM/ETM+ images for burned area mapping, an overall accuracy of 99.72%, commission error of about 21% and omission error less than 3% have been reached by a semi-automated OWA and region growing algorithm based on Spectral Indices and soft computing techniques (Stroppiana et al., 2012).

There is an increase in employment of Sentinel-2 imagery for BA mapping purposes. Filipponi (2019) offers a novel approach for medium-high resolution BA mapping by using Sentinel-2 MSI time series data. The study includes a four-step supervised method to develop an automated and transferable procedure that relies on a change detection approach, based on the identification of structural changes in vegetation cover, and exploits complementary information through a region-growing algorithm for 2017 wildfires in Italy. The proposed procedure resulted in generating a commission error of around 25% and an omission error of around 40% and the best performance



gained by NBR spectral index (Filipponi, 2019). In development of FireCCISFD11 product, on sub-Saharan Africa, a multitemporal two-phase algorithm for BA detection has been developed (Roteta et al., 2019). They have used short and near-infrared bands from Sentinel-2 MSI measurements and hotspots by MODIS. The aim of selecting such a large area was comparing this local BA area map with global products. The results showed that the total BA estimated from this product was 4.9 M km<sup>2</sup>, around 80% larger area than what the NASA BA product (MCD64A1 c6) detected in the same period (2.7 M km<sup>2</sup>) with higher dice coefficient (DC) and much lower omission error. A comparison of different global BA products and FireCCISFD11 is demonstrated in Figure 9 (Lizundia-Loiola et al., 2020). In the end, it is proved that Sentinel-2 imagery can give a great contribution to better understanding the impacts of small fires in global fire regimes even if it has less temporal reporting accuracy than global BA products (Roteta et al., 2019). In addition to its positive contribution to detect small burned areas, Sentinel-2 has a potential to effectively used to map the wetland burned areas which is a challenging ecosystem as a result of spectral confusion between open water and burned area, rapid post-fire vegetation regrowth, and high annual precipitation limiting clear-sky satellite observations (Vanderhoof et al., 2021).

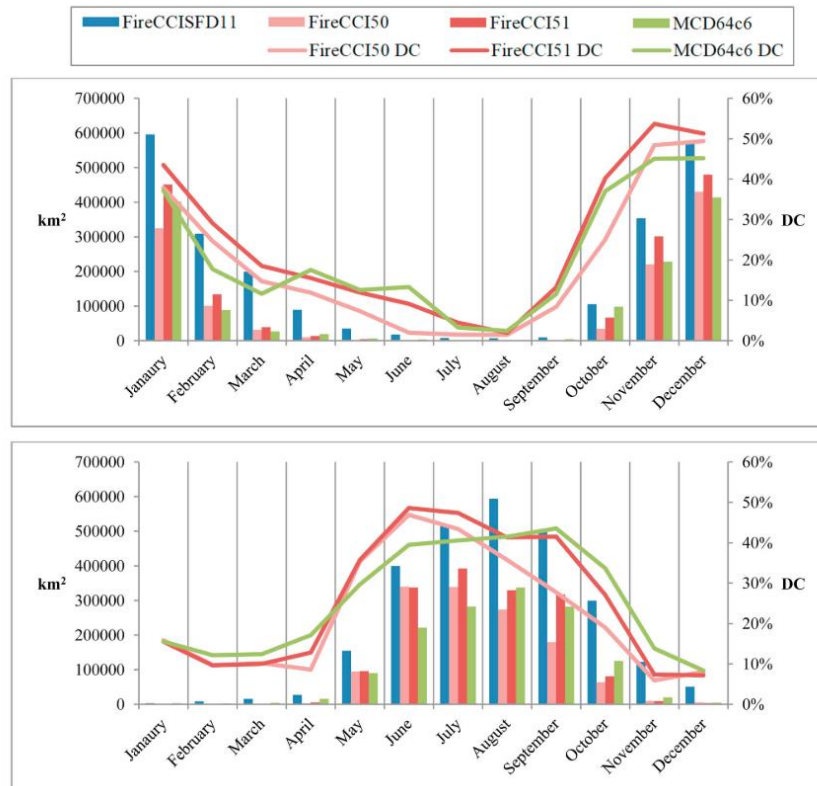


Figure 9. Comparison of the MODIS BA products (FireCCI50, FireCCI51, and MCD64A1 c6) against the Sentinel-2 product (FireCCISFD11) for the year 2016. a) Northern Hemisphere Africa (NHAf) and b) Southern Hemisphere Africa (SHAf) (Lizundia-Loiola et al., 2020)

And lastly, in the pioneer work of this study, an algorithm initially developed for post-fire Landsat data was assessed to see if it is adoptable to multi-temporal Sentinel-2 data with some improvements. Sali et al. (2021) already proposed a new version of the fuzzy BA algorithm applied to S2 data but in this thesis the objective is twofold: i) to implement the algorithm in a QGIS workflow to have a suitable environment for displaying and visualisation of the results and ii) to test the performance of the algorithm over a new site in Turkey. The proposed algorithm combines the partial evidence of burn which is obtained from the S2 bands reflectance data through the membership functions, into a synthetic score of global evidence by means of an ordered weighted averaging (OWA) operator. This semi-automatic robust and self-adaptive algorithm is offered to maximize mapping accuracy delivered by Sentinel-2 imagery. The BA algorithm was trained over the Vesuvius National Park, Italy, which was highly affected by fires in 2017 and contained different degrees of severity. After the training phase, accuracy assessment was carried out over a wide range of conditions/locations in Southern Europe for the 2017 summer fire season (Sali et al., 2021). In this thesis the algorithm is further tested to assess its exportability to a new region.

## 4. Methodology

In this study, we can divide the process in two separate parts.

The first one is the reorganization of the previously developed codes written in R as a sequence of separated steps and the integration of these codes for a direct use in QGIS in a more user-friendly way. The second part is testing the algorithm in QGIS over a new test site in another Mediterranean region (Turkey). This step allowed both testing the code itself as integrated in the new GIS environment and validating the output burned area maps in a different geographic area.

The first phase implied a deep analysis of the codes used to create burned area maps from Sentinel-2 images. The QGIS and R integration phase was at first tested on the previously analysed fire site Pedrógão Grande, Portugal (2017), in order to check step by step the correctness of the code's modifications. And lastly, further test and validation area was chosen as one of the most recent fire events in Antalya, Turkey (2021).

The core of this study is the analysis of the R codes step by step, then modification and implementation of them in QGIS. Before going deeper on what I did, a theoretical background of the algorithm could be helpful for the reader. The process workflow consists of three main steps which are:

- a. Computation of Membership Degrees (MDs) for each pixel
- b. Computation of Ordered Weighted Averages (OWAs) for each pixel
- c. Validation of the output Burned Area (BA) maps by comparing to reference data

The above listed steps address two main parts of the work: burned area mapping (a, b) and validation (c). The inputs for the first part of the algorithm (part a,b) are the Sentinel-2 images acquired for the dates before and after of the fire event. In order to select the useful bands or the band combinations for the discrimination of burned and unburned surfaces, a separability analysis was carried out on the burned area over Vesuvius National Park, Italy (2017) in a previous study (Sali et al., 2021). We had used the same algorithm parameters which had been tuned in the same research. To choose the bands to be used, they have calculated the separability metric  $M$  by using frequency distributions of training pixels on their study.

$$M = \frac{\mu_u - \mu_b}{\sigma_u + \sigma_b}$$

where:  $\mu_u$  and  $\mu_b$  represent the mean values of unburned (u) and burned (b) classes, while  $\sigma_u$  and  $\sigma_b$  represent the standard deviation for the same two classes in the training data.

The result of the formula is a value varying between 0 and  $+\infty$ . If  $M > 1$ , the two classes are separable. Furthermore, the more  $M$  is greater the more the classes are separable. If  $M$  takes values below of the unit, the two classes cannot be considered separable.

In our study, we have used the same seven input bands which were selected from the separability analysis for the implementation of the algorithm and to be used in the membership function. These bands are: post-fire NIR (B8), post-fire RE2 (B6) and RE3 (B7) and temporal difference (difference between post-pre fire reflectances) of the same three bands and additionally of SWIR2 (B12) (Sali et al., 2021).

In the following sections, first, the basis of R and QGIS integration phase and then, the theoretical definitions of the three main algorithm steps, explanations of the separated R codes and the modifications between R and QGIS will be explained.



Figure 10. Fire areas in Mediterranean region used for : the parameter selection (b: Vesuvius National Park, Italy), integration phase (a: Pedrógão Grande, Portugal) and validation step (c: Manavgat, Antalya, Turkey) of the algorithm

#### 4.1. R and QGIS Integration

This step is the main objective of the thesis because it makes the whole process of creation of burned area maps run in a more user friendly and semi-automatic way.

The purpose in fact is to speed up the use of the code, and to allow a much easier way to visualize and analyse the results through QGIS interface.

Furthermore, the creation of a link between QGIS and R can be especially essential for sharing your applications or algorithms with people who have limited skills in programming or spatial data analysis, that would be required by using the algorithm implemented, for example, in R.

QGIS includes a number of plugins that you can activate in order to extend its functionality. The scripting language for QGIS is Python and the creation of a new plugin is only possible by writing the algorithms in Python to make it an open source. Python is fully integrated and cooperates with QGIS.

On the other hand, there are many spatial extensions of R that support a wide range of geospatial capabilities. According to the research of GoGeomatics, on the top programming languages in GIS world, R will play an important role in the years to come (URL10).

Since the algorithm had been developed and implemented only in R, the plugin “*Processing R Provider*” (Figure 11) is used, which allows the execution of R scripts directly on data loaded in QGIS; however, these codes could also have been run in QGIS with some modifications.

In the first place, to use the scripts on QGIS, R has to be installed separately in the computer because the plugin searches for the installation automatically. Then, after the installation of *Processing R Provider*, by clicking the “options” icon (Figure 12 (a)) in the processing toolbox, you can configure the links between R and QGIS. This will bring up a window (Figure 13) where you can find the important paths to step up. Finally, you can find the icon for creation of new R scripts on the processing toolbox (Figure 12(b)).

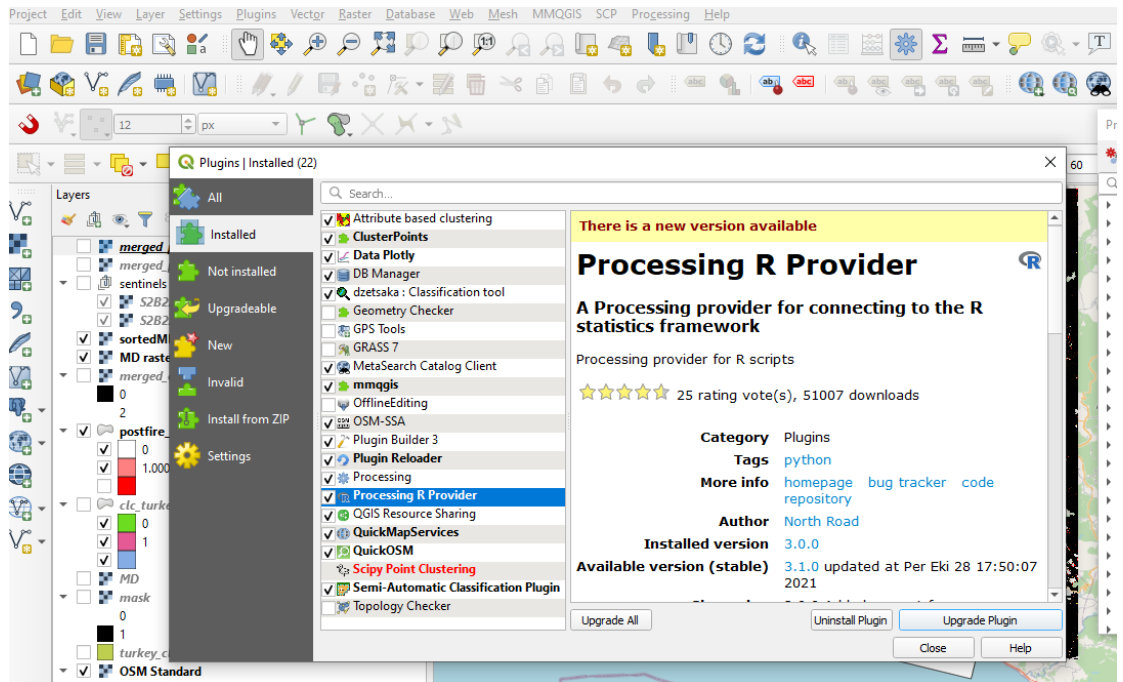


Figure 11. Plugin panel-Processing R Provider window

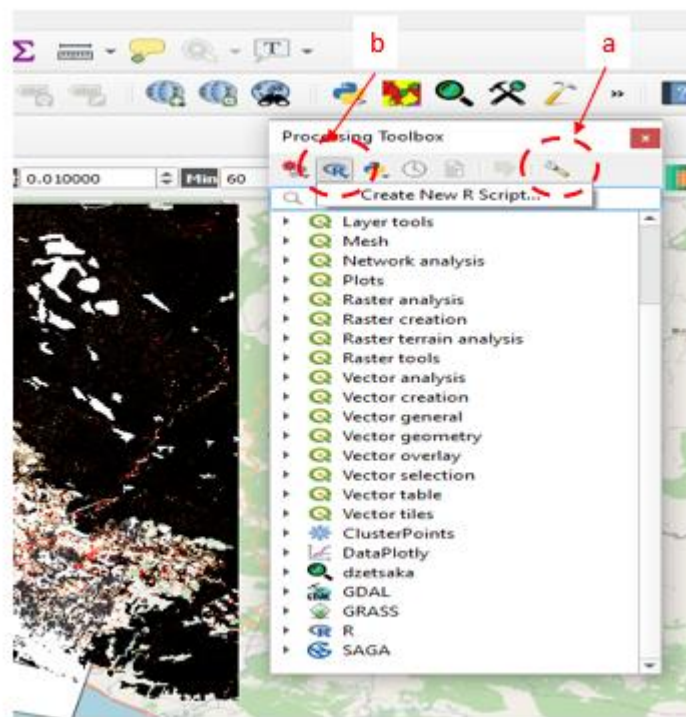


Figure 12. (a) Options icon (b) Icon for R script creation

After the installation of R and setting up the directories for linkage, you can either create a new R code or copy-paste your existing codes on QGIS. In my case, I directly pasted our codes part by part. The main issue with R codes running in QGIS is the correct definition of inputs and outputs in QGIS syntax (Table 2 Table 2. R Syntax in Processing scripts (URL12)). There are some specific code lines which must be specified at the beginning of the code. They, with several metadata, define the interaction between QGIS and R. These lines that start with hashtag # should be placed in the code so that QGIS can understand the R code. By using the correct statements, you can create a group for your models under the R segment (Figure 14) in the processing toolbox. Furthermore, these specified lines determine how the user interface will look like (URL11).

However, QGIS doesn't run the code immediately when you click the name of the code, because it needs to understand what the inputs and outputs of this run are. Indeed, in running from the processing toolbox list an R code, QGIS at first just searches for it then, independently from the code is working or not, it opens a panel (Figure 15) to require the inputs and define the repository for your output.

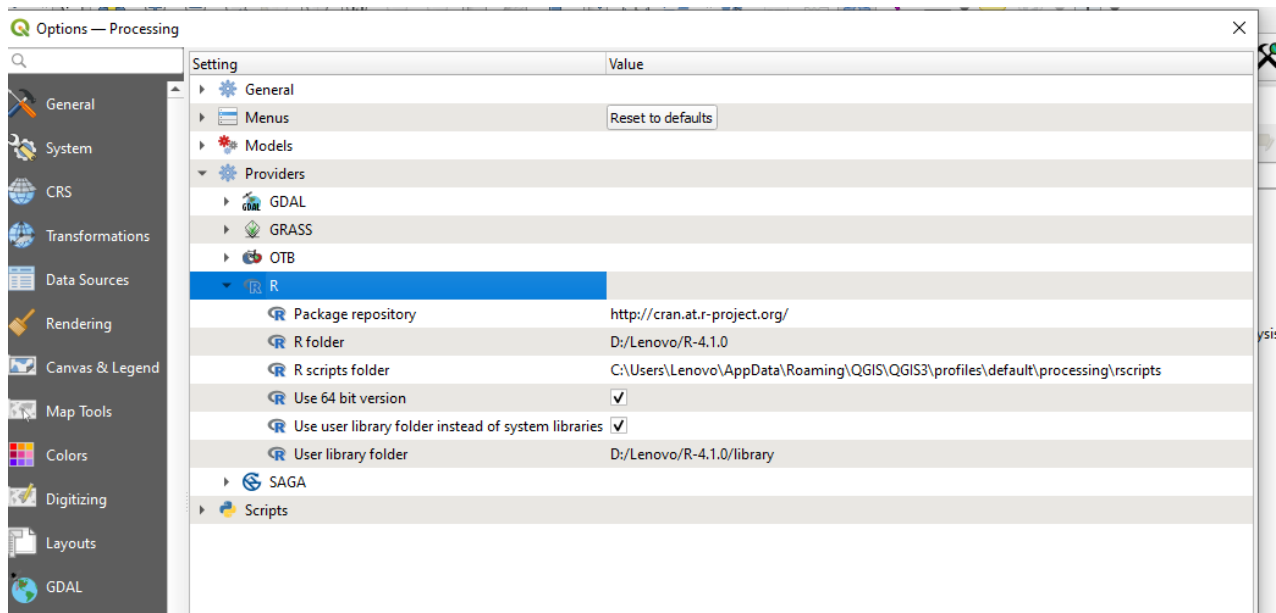


Figure 13. QGIS window to set up the important paths between QGIS and some main providers

Table 2. R Syntax in Processing scripts (URL12)

Inputs	Input Explanation	Outputs
##My Group=group	Group creation	##output = output vector
##Name = name	Naming the algorithm	##output = output raster
##Layer = vector or raster	Specifying the input layer is vector or raster	##output = output table
##Field= Field Layer	Where layer is the name of the input layer	##showplots
##Bol = boolean	To control the selection on the input/output interface	
##Str = string		
##Num = number		

The necessary input files should be added to QGIS layers panel because the input/output interface requires input selection from the already opened files.

Until this point, properties or functionality of the R code is not checked by QGIS, it only deals with the metadata written at the beginning of the code (#). This makes the selectivity of this integration fail on a certain level. For example, in your code, you may have an algorithm that optionally creates more than one output, however either you need only one, or some, or all of them, QGIS asks you either to create a temporary folder or a repository for each of them, before starting the run. This may be enormously time consuming depending on the complexity of the run. In order to deal with this issue, some parts of the algorithms had to be recoded.



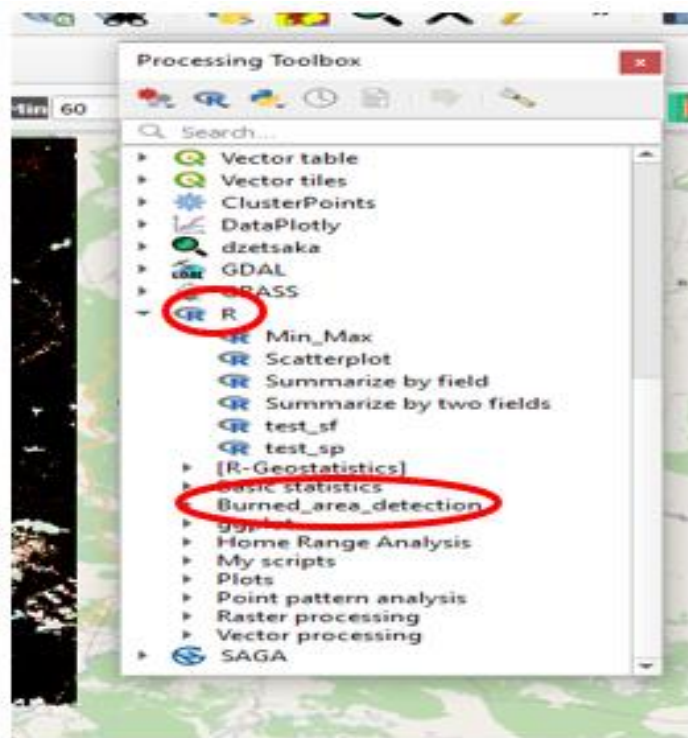


Figure 14. QGIS Processing Toolbox interface

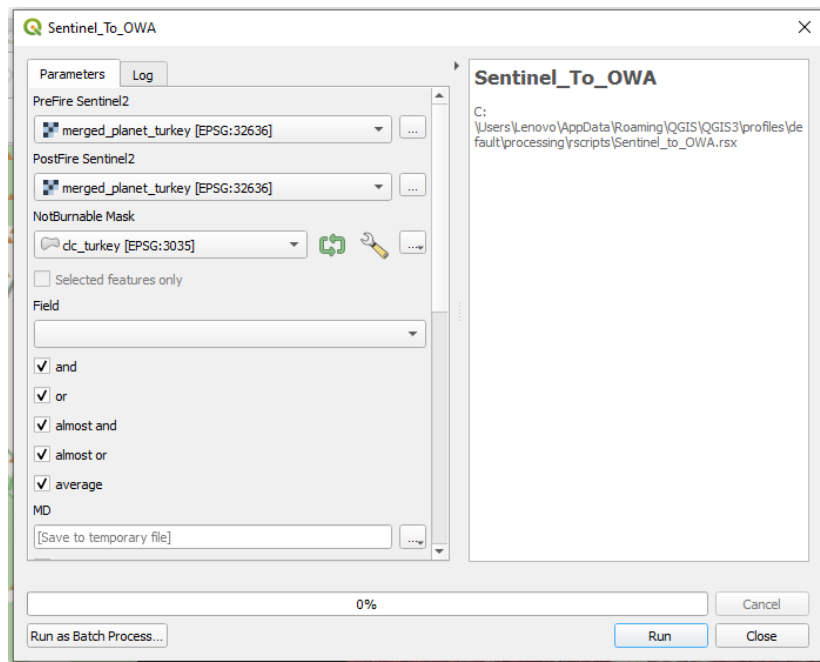


Figure 15. QGIS input/output interface example before the run

In the following sections, explanation of each code and the required modifications to run the algorithm on QGIS will be explained in detail.

#### 4.2. Theoretical Definitions and Code Modifications for Integration

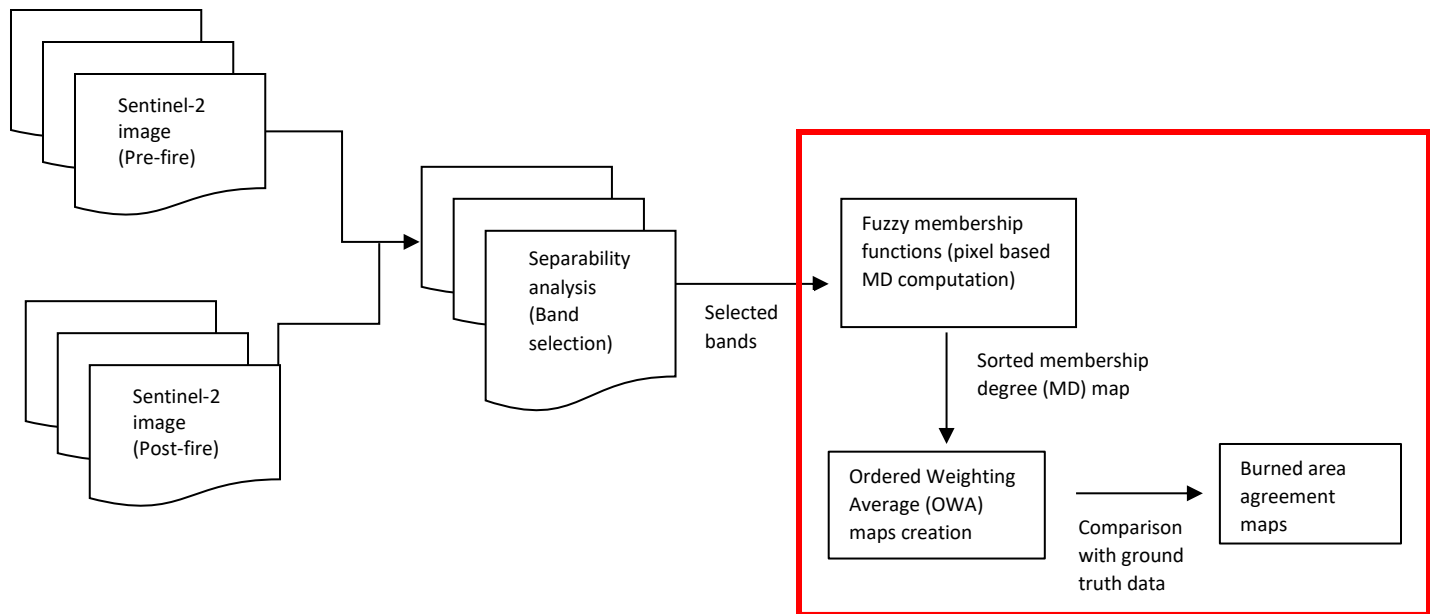


Figure 16 . Process flow (from Sentinel-2 imagery to burned area mapping)

Figure 16 shows the flowchart of the algorithm: a pair of Sentinel-2 images (pre- and post-fire dates) is the input of the whole system and by adding information from some other systems (e.g. land cover data) to Sentinel data, we end up with burned area maps. In our case, the main issue was not running the codes because it was proven that the codes were working properly on R. At the beginning, we had four separated R codes (1 is intermediate) and they were not connected to each other. The first step for our study was having a deep understanding of the codes and testing them during each step of integration in QGIS. The need for an interoperable code between R and QGIS asked for modifications of some parts of the code. Here, we will explain firstly the theory behind each separate module of the algorithm, secondly the code owning the module and lastly the modifications and tests applied during the integration.

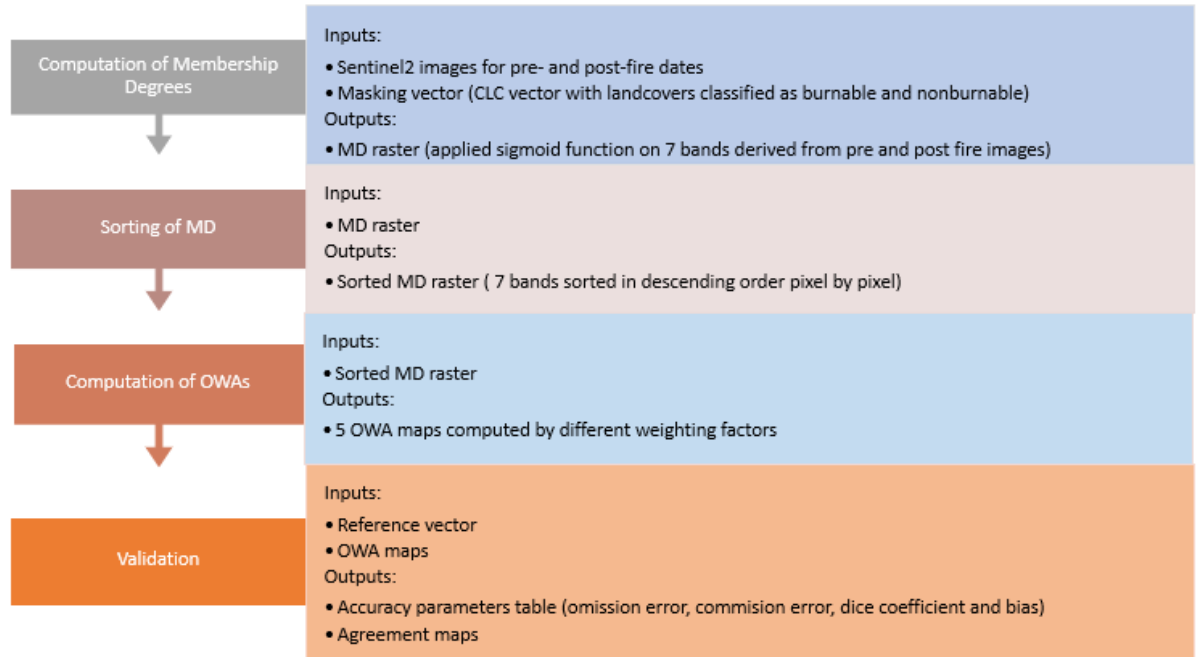


Figure 17. Step by step Input/Output for each part of the algorithm

#### 4.2.1. Fuzzy Membership Function-Computation of Membership Degrees (MDs)

A Membership Function (MF), in definition, is a function used to specify the degree for a given input belonging to a class. The output of this kind of a function is called a membership degree that ranges between 0 and 1. For example, in our system, a degree closer to 1 represents an area which belongs more likely to the “burned area” class, while a degree closer to 0 can be accepted in the “unburned area” class. In order to derive a burned area map (i.e. a binary classification of burned/unburned areas), a threshold value ( $\in [0,1]$ ) is chosen and the pixels having values greater than this threshold are considered as “burned”.

In the algorithm developed by Sali et al. (2020), logistic regression, which is a classification model that uses an S-shaped curve to separate values of different classes, has been used to calculate the Membership Degree (MD) for each pixel.

$$MD = f(x) = \frac{L}{1 + e^{-k(x-x_0)}}$$

where:

x = a real number

x<sub>0</sub> = the x value of the sigmoid midpoint

k = logistic growth rate

L = the curve's maximum value (here 1, MD should be between 0 and 1)

The spectral response (x) from Sentinel-2 images for the selected seven bands are the input values for the sigmoid functions. That means, we run the MF for each of the seven bands separately for each pixel. The other parameters (k and x<sub>0</sub>) were also tuned during the study on the fire event at Vesuvius National Park, Italy (2017). We did not change these parameters for test areas because one of our research goals is the automatization of the process for whole Mediterranean region. So, this thesis work contributed to testing the exportability of the BA algorithm as developed by Sali et al. (2020) to a new geographic area. Thus, the only area specific parameter is the spectral response acquired from Sentinel-2 image of the AOI.

In the R code, the process flow of the computing MD module can be summarized as below:

- a. Loading of the necessary R packages
- b. Definition of the sigmoid membership function
- c. Setting band specific parameters (k and x<sub>0</sub>)
- d. Reading the input dataset (pre-fire and post-fire Sentinel images and, binary classified Corine Land Cover (CLC) vector)
- e. Conversion (rasterization) of the CLC shapefile to raster (based on the attribute field of the shapefile that provides information on burnable and not burnable polygons)
- f. Setting output directory for storing/saving MD output raster
- g. Calculating layers used in the processing from S2 spectral bands
- h. Application of sigmoid function by inserting the values on steps c and h
- i. Load of available data into a single raster (stacking)

- j. Masking the stack with burnable/not burnable mask derived from the CLC
- k. Writing the output file in the directory at step g

These above steps were modified in some points for the integration with QGIS. The required packages are much less when working on QGIS because it already has most of the tools needed for geospatial processes. And also, especially the input/output requirements are quite different for R and QGIS interfaces as mentioned in Section R and QGIS Integration4.1. The added and removed lines for code integration are as below:

**Added Lines:**

*For Definition of Inputs/Outputs:*

```
##Burned_area_detection=group    >> Group name for every code of the algorithm
##a_Compute Membership Degrees=name >> Algorithm name
##PreFire_Sentinel2=raster      >> Input Sentinel-2 image for pre-fire
##PostFire_Sentinel2=raster    >> Input Sentinel-2 image for post-fire
##NotBurnable_Mask=vector      >> Input Masking vector
##MD=output raster            >> Output masked MD map
##mask=output raster          >> Output mask (to check)
##Field=Field NotBurnable_Mask >> Masking vector required layer
```

*For setting the input rasters and vector at the same CRS:*

```
if (!compareCRS(crs(NotBurnable_Mask),crs(PreFire_Sentinel2) )){
crs = crs(PreFire_Sentinel2)
NotBurnable_Mask <- st_transform(NotBurnable_Mask, crs)}
```

*For showing the outputs in the QGIS Result Viewer:*

```
mask <- mask_forest
MD <- bande_md
```

**Removed Lines:**

*For definition of input/output directories:*

```
in1 <- "directory of Sentinel-2 pre-fire image"
in2 <- " directory of Sentinel-2 post-fire image "
mask <- st_read("directory of Masking vector (CLC shapefile)")
tiff_fname_md <- "output directory for MD map"
```

*For writing the created map into the output file:*

```
writeRaster(bande_md, tiff_fname_md, overwrite=TRUE)
```

After the calculation of MDs for each band, the intermediate step was sorting them in descending order pixel by pixel. This output was used in the next step, the OWA map computation. The added lines for I/O definition on QGIS are:

```
##Burned_area_detection=group
##b_Sort Membership Degrees=name
##MD=raster
##MD_raster_sort=output raster
```

#### 4.2.2. Computation of Ordered Weighted Averages and BA maps creation

The sorted MD multiband raster is then used in Ordered Weighted Averaging (OWA) as the second step of the algorithm.

OWA family is a tool which is first introduced by Yager (Yager, 1988) to deal with the problem of aggregating multicriteria functions to form an overall decision function. It provides an aggregation lying between two extremes (“and” and “or”). At one extreme (“and”), we desire that an alternative satisfy “all” the criteria. While, at the other extreme (“or”), the alternative should satisfy at least one of the criteria. In many cases the type of aggregation operator desired lies somewhere between these two extremes. Following this conceptualization, the decision maker’s attitude is highly important to form an overall decision function, the decision maker should state the number or the percent of the criteria that must be satisfied.

**Definition:** An OWA of dimension  $N$  and weighting vector  $W$ , with  $i=1, \dots, N$   $\sum w_i = 1$ , aggregates  $N$  input values  $[d_1, \dots, d_N]$  and computes an aggregated value  $a$  in  $[0, 1]$  as follows:

$$OWA: [0, 1]^N \rightarrow [0, 1] \quad a = OWA([d_1, \dots, d_N]) = \sum_{i=1}^N w_i * g_i$$

in which  $g_i$  is the  $i$ -th largest value of the  $d_1, \dots, d_N$ . In this case study,  $([d_1, \dots, d_N])$  are the MDs of the seven input factors ( $N=7$ ): RE2, RE3, NIR,  $\Delta$ RE2,  $\Delta$ RE3,  $\Delta$ NIR,  $\Delta$ SWIR2.

As a multicriteria evaluation, unlike classical logic which is either true or false, the OWA operations allow the user to change the form of attributes contribution from a minimum (Logical AND, “optimistic”) to a maximum (Logical OR, “pessimistic”) combination on the maps created. All intermediate combinations, obtained by changing the order and magnitude of weights, may also create meaningful results for decision makers (Chapman et al., 2003).

In our algorithm, five different combinations of weights were used as OWA:

$w_{and} = [0, \dots, 0, 1]$  □ most optimistic

$w_{almost\_and} = [0, \dots, 0.5, 0.5]$

$w_{almost\_or} = [0.5, 0.5, \dots, 0, 0]$

$w_{average} = [1/n, \dots, 1/n]$

$w_{or} = [1, \dots, 0, 0]$  □ most pessimistic

The added lines for OWA code integration are as following:

**Added Lines:**

*For I/O definition:*

```
##Burned_area_detection=group    >> Group name
##c_Compute OWA=name            >> Algorithm name
##MD_raster_sort=raster        >> Input Sorted MD map
```

```
##and=Boolean
##or=Boolean
##almost_and=Boolean
##almost_or=Boolean
##average=Boolean
```

} >> User interface selection panel creation

```
##OWA_and=output raster
##OWA_almost_and=output raster
##OWA_or=output raster
##OWA_almost_or=output raster
##OWA_average=output raster
```

} >> Output BA maps for different weight combinations

*For time efficiency:*

```
OWA_and<- raster(nrow=1,ncol=1)
values(OWA_and)<-0
OWA_almost_and<- raster(nrow=1,ncol=1)
values(OWA_almost_and)<-0
OWA_or<- raster(nrow=1,ncol=1)
values(OWA_or)<-0
OWA_almost_or<- raster(nrow=1,ncol=1)
values(OWA_almost_or)<-0
OWA_average<- raster(nrow=1,ncol=1)
values(OWA_average)<-0
```

>> Primary setting OWA  
rasters as “0” not to lose time  
on the unselected computations

All the lines to load packages are removed in this part of the code, because all necessary packages were loaded in the primary MD computation step.

Lastly, before computation of each OWA combination, we inserted an “*if*” statement, so that if the user does not select one of the combinations in the selection panel, that BA map will not be created.

#### 4.2.3. Calculation of Accuracy Metrics

The third part of the implementation steps is computation of the accuracy metrics, i.e. validation (step c). Validation provides a quantitative assessment of the accuracy of geoproducts, burned area maps in this case; accuracy is relevant information for end users, while facilitating critical information for end users (Congalton and Green 1999). To assess accuracy, the training dataset must be compared by a reference. In our case, this spatial dataset should get as close as to the ground truth to have a high accuracy. Since the in-situ fire data is hardly feasible to be achieved over large areas and we need a representative and reliable data as reference, the use of high resolution remotely sensed data is widely accepted. The first rule for a reference dataset to be considered as ground truth is that this dataset must have a high accuracy and be generated independently (Morissette et al., 2006). In fact, in this work we used reference burned perimeters (as ground-truth data) derived from very high-resolution Planet images which is a high frequency satellite monitoring



system with a spatial resolution of 3.7 m in four multispectral bands (Red, Green, Blue and NIR). We decided to use only the post-fire Planet image as reference raster.

The selected reference image was classified by Dzetsaka Classification Tool (Figure 18), a semi-automatic classification plugin. This plugin is very fast, easy and supports different classification models such as Gaussian Mixture Model, Random Forest, K-nearest neighbours, Support Vector Machines. We have used Random Forest classification algorithm, whose working principle depends on building decision trees on different samples and the final output is based on average or majority ranking (URL13).

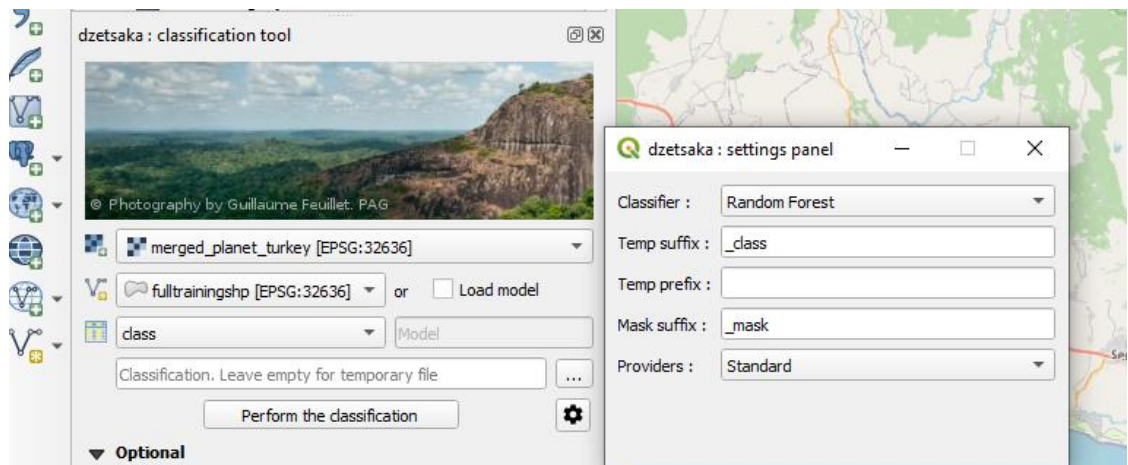


Figure 18. Dzetsaka classification tool interface for QGIS

The next step after the classification was applying majority filter (SAGA) on the output raster to reduce the noise. And finally, vectorization was applied to the classified and filtered reference raster. By using this reference vector, we calculated the accuracy metrics for each of the OWA maps, to decide which weighting function gives a closer result to the ground truth.

There are many different measures for validation in the literature. In our study, validation step was carried out by calculating the omission /commission errors, Dice coefficients and bias between the reference data and different OWA maps. All of these metrics are calculated by using the True Positive, True Negative, False Positive and False Negative pixels (Table 3).

Table 3. Sampled error/confusion matrix:  $n_{ij}$  express the number of pixels of agreements (diagonal cells) or disagreements (off diagonal cells) between the BA product and the reference PLANET image. In the case of binary classification (burned areas and not burned areas) considering as target objective of the algorithm the identification of Burned areas we can make this equivalence: True Positives ( $TP = n_{11}$ ), True Negatives ( $TN = N_{22}$ ), False Positives ( $FP = n_{12}$ ) and False Negatives ( $FN = n_{21}$ ).

		Reference		
		Burned	Unburned	Total
BA algorithm	Burned	$n_{11}$	$n_{12}$	$n_{1+}$
	Unburned	$n_{21}$	$n_{22}$	$n_{2+}$
Total		$n_{+1}$	$n_{+2}$	

Errors of commission occur when a pixel is incorrectly included in burned class while the pixel is not burned in the reference, while errors of omission occur when a burned pixel is left out and classified as unburned (Table 4).

Table 4. Metrics computed from the error/confusion matrix and range of variability

Accuracy metric name	Formula	Range
Commission error	$Ce = \frac{n_{12}}{n_{1+}}$	$[0,1]$
Omission Error	$Oe = \frac{n_{21}}{n_{+1}}$	$[0,1]$
Dice Coefficient	$DC = \frac{2n_{11}}{2n_{11} + n_{12} + n_{21}}$	$[0,1]$
Relative Bias	$relB = \frac{n_{21} - n_{12}}{n_{+1}}$	$[-1,+1]$

Deciding on which accuracy measure between BA omission and commission errors is preferred is almost impossible because they are equally important. They show the proportion of area of agreement and disagreement but comparison on them may not yield a meaningful burned accuracy. Another reason for this is that we deal with a two-class problem. Thus, another measure, Dice Coefficient (DC), summarizes both omission and commission errors, is reported, and it shows the degree of how similar the pixels in the two maps compared (URL14).

In the accuracy computation code, the aim is not only to tabulate the accuracy metrics as a measure of the agreement between reference and classified maps, but also to depict the spatial distribution of the agreement and disagreement pixels through the *agreement maps*; these maps show whether pixels in the selected OWA map and reference map fall in one of the classes: correctly burned or unburned, omission or commission.

The added and removed lines for the integration of this last portion of the codes are as follows:

**Added Lines:**

*For I/O definition:*

```
##Burned_area_detection=group          >> Group name
##Compare_with_reference_vector=name   >> Algorithm name
##reference=vector                    >> Input ground truth (classified reference vector)
##raster_to_validate=raster           >> Input OWA raster
##threshold=number                   >> threshold for each pixel to be assumed as burned
##Field=Field reference               >> the classification layer on reference vector
##tab=output table                   >> Output tabulated accuracy metrics
##agreement_map=output raster        >> Output agreement map
```

Since we have different input sources and they may not be set in the same CRS, we added the following lines to the code, as in MD computation module, to manage different reference systems:

```
if (!compareCRS(crs(reference),crs(raster_to_validate) )){
crs = crs(raster_to_validate)
reference <- st_transform(reference, crs)}
```

In QGIS, since we have an interactive user panel, we do not need to define the threshold (that we have used to convert the continuous map to binary) in the code, we can change it in the user interface. So, the threshold definition line was removed from the code.

Finally, we have added some lines to create a table and show the outputs in the layer panel:

```
rownames <- c("value")
colnames <- c("omission error", "commission error", "Dice Coefficient", "Bias", "Relative
Bias")
tab <- matrix(c(oe, ce, dc, bias, relB),ncol=5,byrow=TRUE)
rownames(tab) <- c("value")
colnames(tab) <- c("omission error", "commission error", "Dice Coefficient", "Bias",
"Relative Bias")
tab <- as.table(tab)
agreement_map <- accuracy
```

In the following chapter, the description of the test area (Manavgat, Turkey) and the data used in the algorithms explained in this section will be displayed.

## 5. Test Area and Data

### 5.1. Description of Test Area and fire season

Manavgat is the second largest district of Antalya, Turkey, with a 2.283 km<sup>2</sup> surface area. The area extending from the coastline to the Taurus Mountains is mainly agricultural area. It is covered with various plants, fruits, and trees. Next to the agricultural lands, as you climb to the Taurus Mountains, maquis and forest areas begin. Red pine is located in the lower part, South of the Taurus Mountains. As you go higher, black pine, spruce, cedar, and juniper take the place of red pine. Willow and plane tree are common in river valleys. Summers are quite hot and can exceed 45-degree Celsius.

In July 2021, Turkey had faced one of the most severe fires in its history. Especially the fire that broke out at four different points in Manavgat caused high damage in forest areas while many animals lost their life. Manavgat is among the most fire-affected places in Turkey, and while seasonal fires are normal and even healthy for the local ecosystem, the scale of 2021 fires was threatening. With the landscape parched and strong winds – in particular, one that blows from the northeast– the authorities were struggling to move fast enough to control things (URL15).

According to Global Disaster Alert and Coordination System (GDACS), this fire in Manavgat lasted nine days and the area affected was 35438 ha (URL16) and seven people have lost their lives. When we look at Figure 19, we can see that the size of the area affected by the fire in 2021 in the Manavgat region is greater than the annual totals in Turkey between 1990 and 2018. And, as it can be seen in Figure 20, MODIS data showed that there were lots of active fire zones in the area even in the fifth day from the first event.

Selecting Manavgat as our study area is meaningful in two perspectives. Firstly, its being in Mediterranean area gives us the opportunity to test if the algorithm and the parameters tuned in Southern Italy and Portugal, can be extended without changing any parameters and creating a fully automatic process for regional mapping. Secondly, this event is very important in size and damage, while there are not many studies done on the area yet.

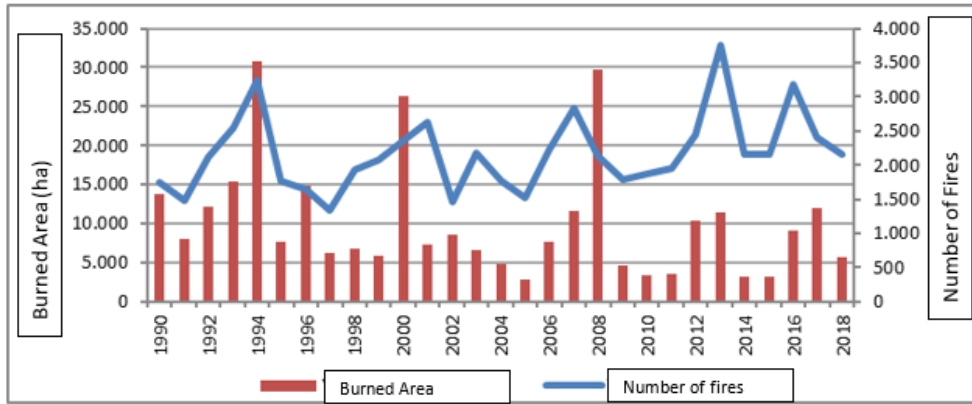


Figure 19. Forest Fires in Turkey between the years 1990-2018 (Republic of Turkey Ministry of Urban, Environment and Climate Change)

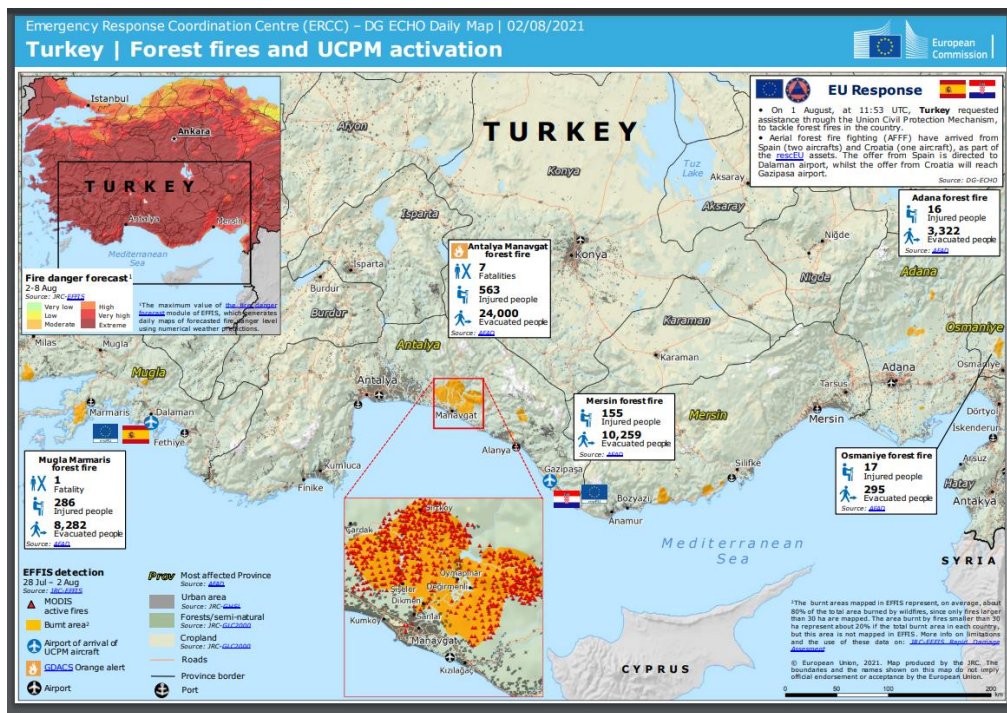


Figure 20. Detailed event map for Turkey Forest fires, 02/08/2021 (ECHO/European Commission)

## 5.2. Geospatial Data Used for Burned Area Detection on Manavgat, Turkey fire event

As explained in process flow of the algorithm, we needed to obtain the pre- and post-fire Sentinel-2 images on our study area (Figure 21). We used a R toolbox called Sen2r which is developed to download and pre-process Sentinel 2 images over a user-specified Area Of Interest (AOI) and time window, and exporting them as multi-band geotiff format (Ranghetti et al., 2020).

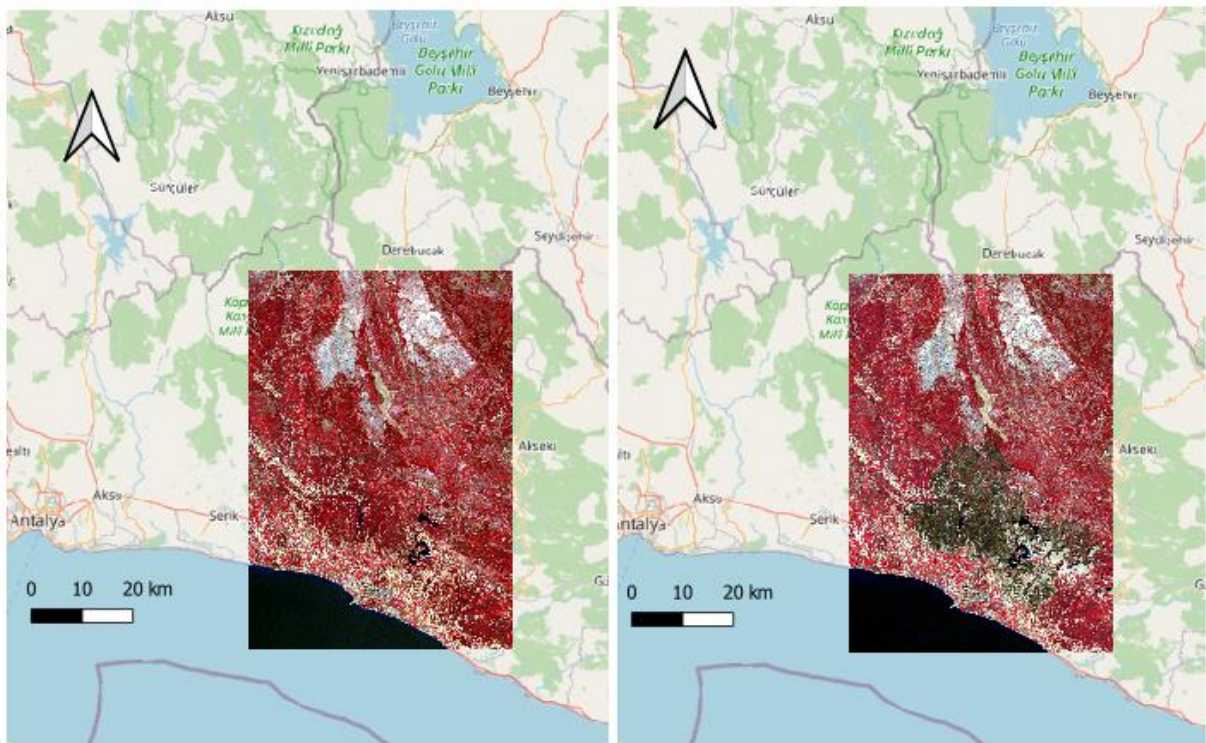


Figure 21. Sentinel-2 images for pre and post fire dates in Manavgat, Turkey (prefire: 20/07/2021 and postfire: 29/08/2021)

In the module of the algorithm for the computation of MD, burnable/unburnable raster mask is foreseen to mask out pixels that cannot burn. This mask was produced by using the Corine Land Cover (CLC) product from Copernicus Land Monitoring Service (Figure 22). This file includes information about the biophysical characteristics of the Earth and we needed to add a field to the original file to define and merge the classes that can be counted as burnable or not (Figure 25).

As we knew that Manavgat is a forestry district, an important amount of the surface area is classified as burnable being covered by woodland and forest vegetation.



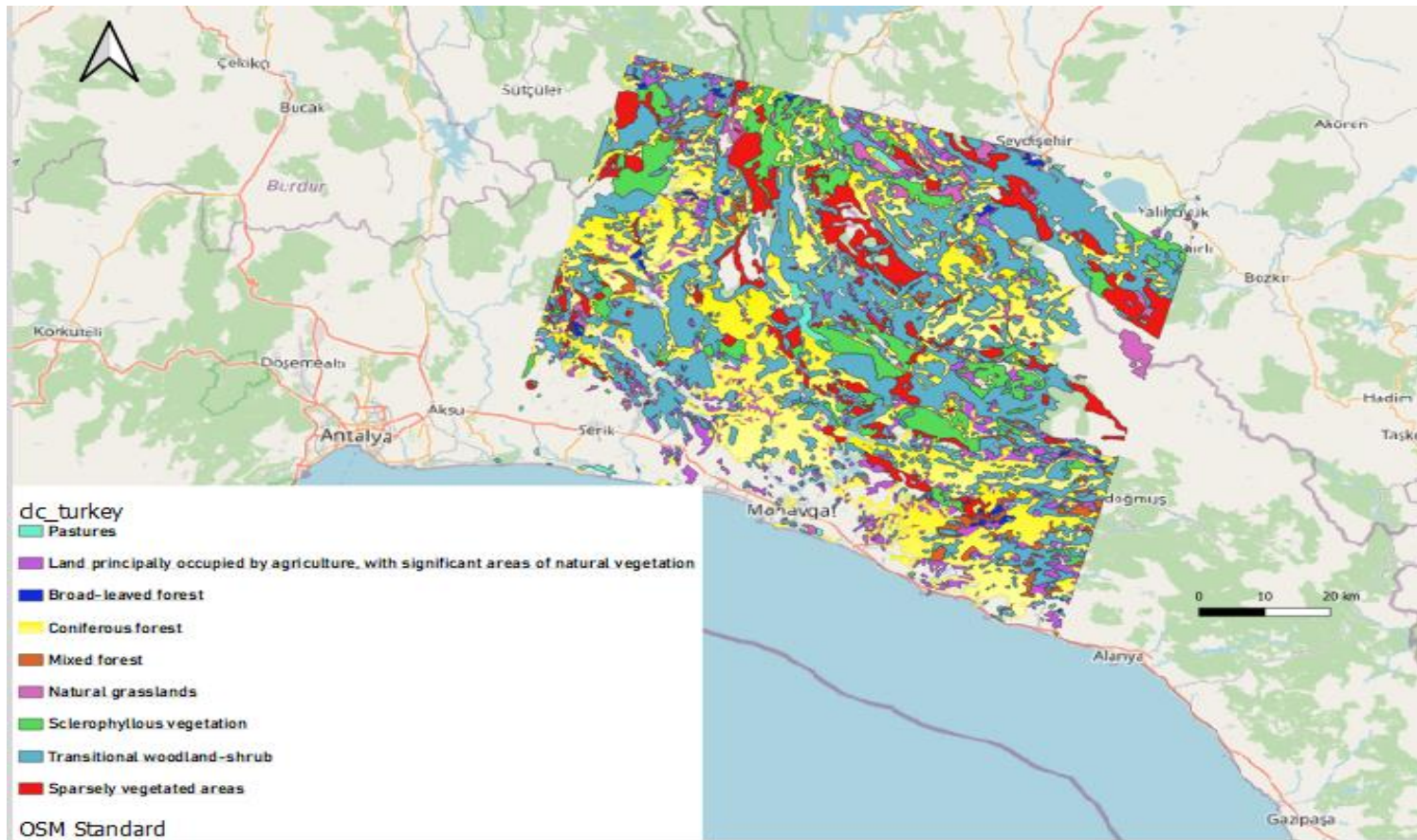


Figure 22. Turkey burnable land cover types

In the CLC metadata file, each polygon is identified by a label and a code number to be associate to a land cover class. In our AOI, there are five main codes and 44 sub codes. For example, code=5 identifies water bodies, but subcodes gives more details about the type of water bodies (e.g., maritime wetlands, inland waters). The list of the main codes and the land covers for each is as following:

Table 5. Landcover types in the CLC(2012) product

Code	Label	Sub-labels
1	Artificial surfaces	<ul style="list-style-type: none"> <li>Continuous urban fabric</li> <li>Discontinuous urban fabric</li> <li>Industrial or commercial units</li> <li>Road and rail networks and associated land</li> <li>Port areas</li> <li>Airports</li> <li>Mineral extraction sites</li> <li>Dump sites</li> <li>Construction sites</li> <li>Green urban areas</li> <li>Sport and leisure facilities</li> </ul>
2	Agricultural areas	<ul style="list-style-type: none"> <li>Non-irrigated arable land</li> <li>Permanently irrigated land</li> <li>Rice fields</li> <li>Vineyards</li> <li>Fruit trees and berry plantations</li> <li>Olive groves</li> <li>Pastures</li> <li>Annual crops associated with permanent crops</li> <li>Complex cultivation patterns</li> <li>Land principally occupied by agriculture, with significant areas of natural vegetation</li> <li>Agro-forestry areas</li> </ul>
3	Forest and semi-natural areas	<ul style="list-style-type: none"> <li>Broad-leaved forest</li> <li>Coniferous forest</li> <li>Mixed forest</li> </ul>

		<p>Natural grasslands</p> <p>Moors and heathland</p> <p>Sclerophyllous vegetation</p> <p>Transitional woodland-shrub</p> <p>Beaches, dunes, sands</p> <p>Bare rocks</p> <p>Sparsely vegetated areas</p> <p>Burnt areas</p> <p>Glaciers and perpetual snow</p>
4	Wetlands	<p>Inland marshes</p> <p>Peat bogs</p> <p>Salt marshes</p> <p>Salines</p> <p>Intertidal flats</p>
5	Water bodies	<p>Water courses</p> <p>Water bodies</p> <p>Coastal lagoons</p> <p>Estuaries</p> <p>Sea and ocean</p>

In Figure 23, the distribution of the overall land cover types is presented. As mentioned before, most of the land (58%) is covered by forested and semi-natural areas. The second greatest percentage (37%) belongs to the agricultural area. This distribution is the direct result of Copernicus Land Monitoring Service. When we work on it and classify the subsets manually as burnable and unburnable, we reached out Figure 24. The burnable class is formed of the CLC codes with 2 (agricultural) and 3 (forested areas).

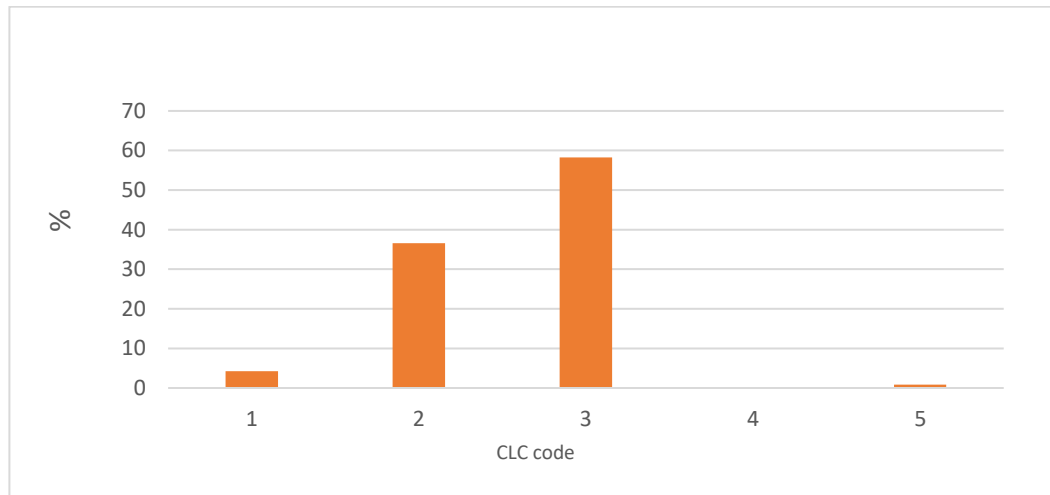


Figure 23. Distribution of percentage of polygon numbers with different land covers

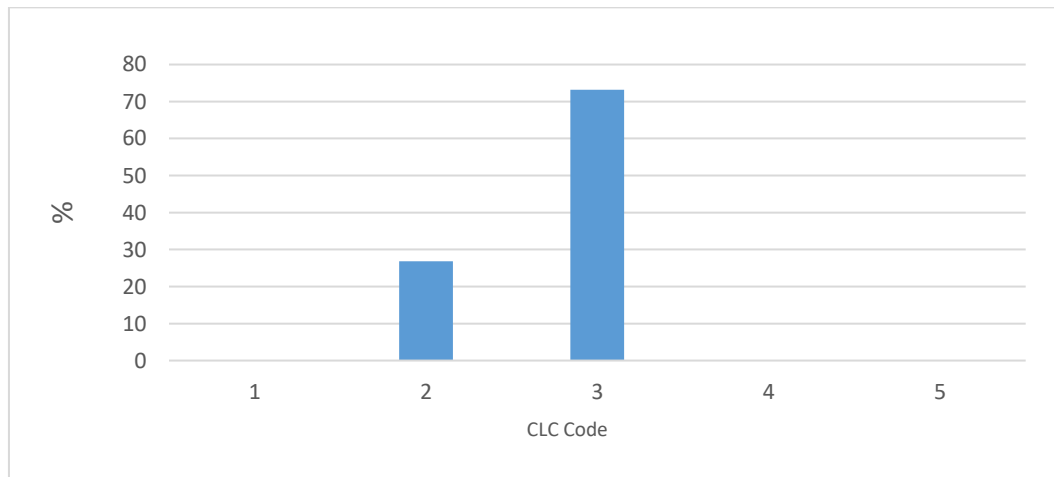


Figure 24. Distribution of percentage of burnable areas w.r.t assigned codes

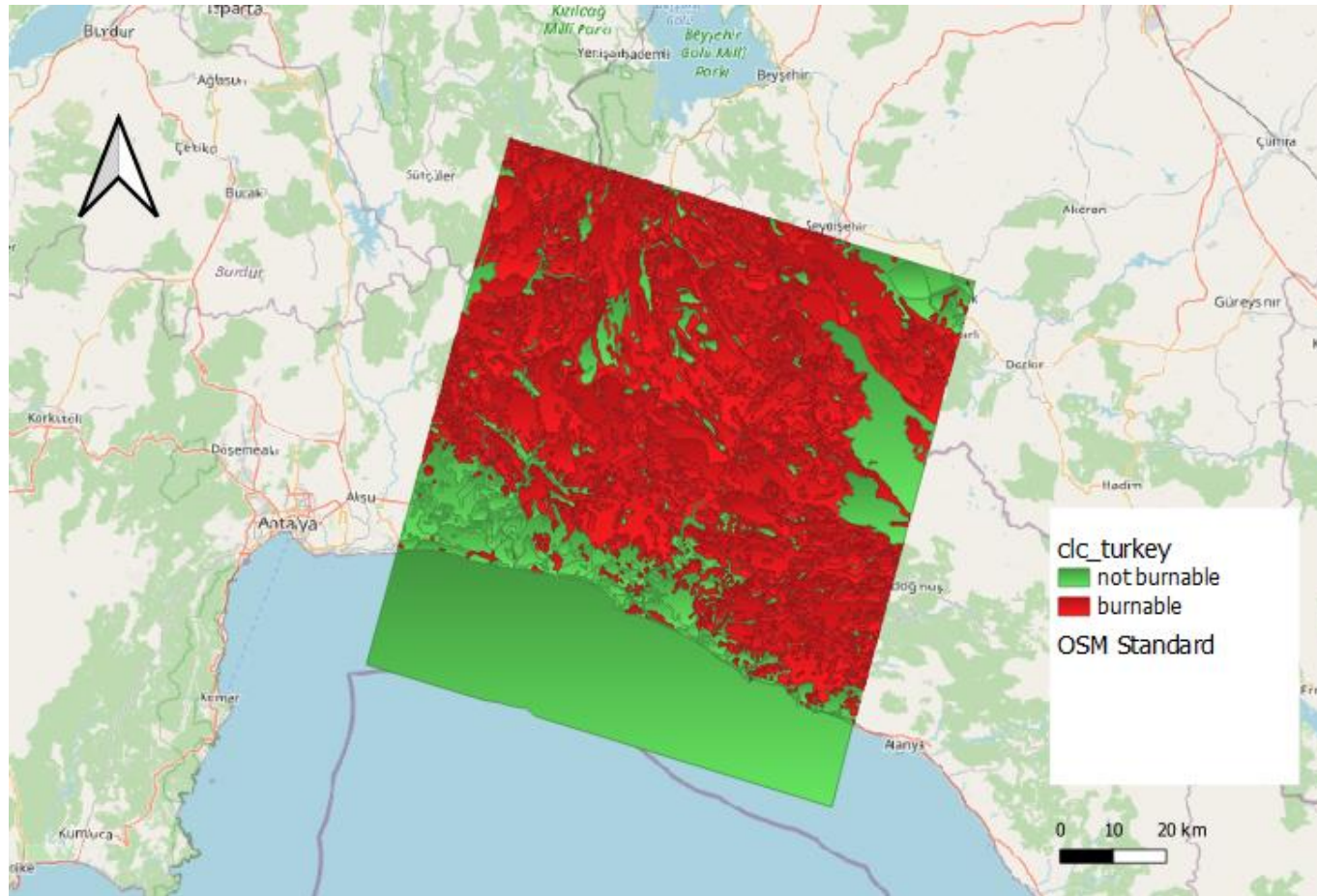


Figure 25. Classified CLC vector of Manavgat (as burnable or not)

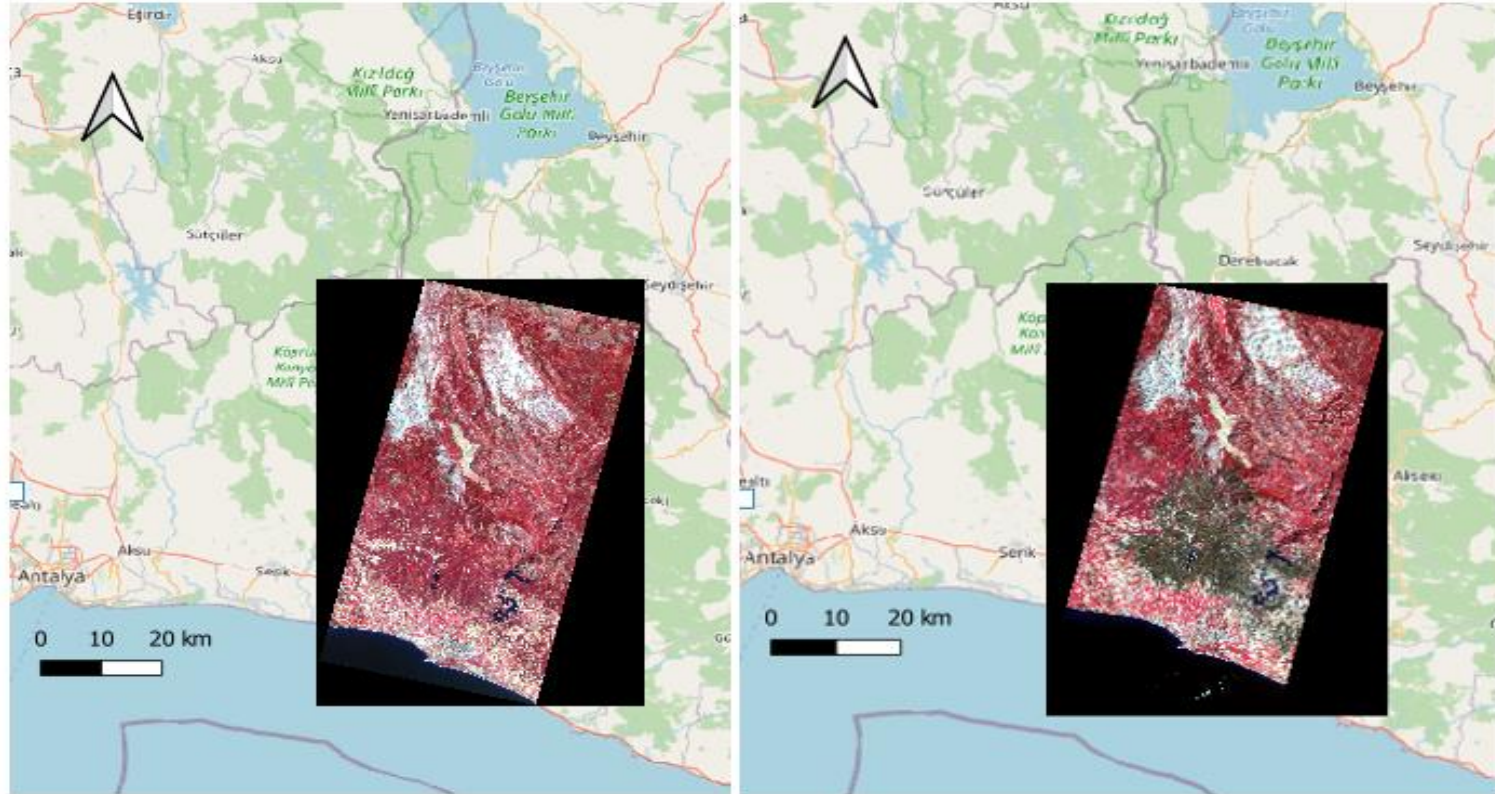


Figure 26. Pre-and post-fire Planet images (ground truth)

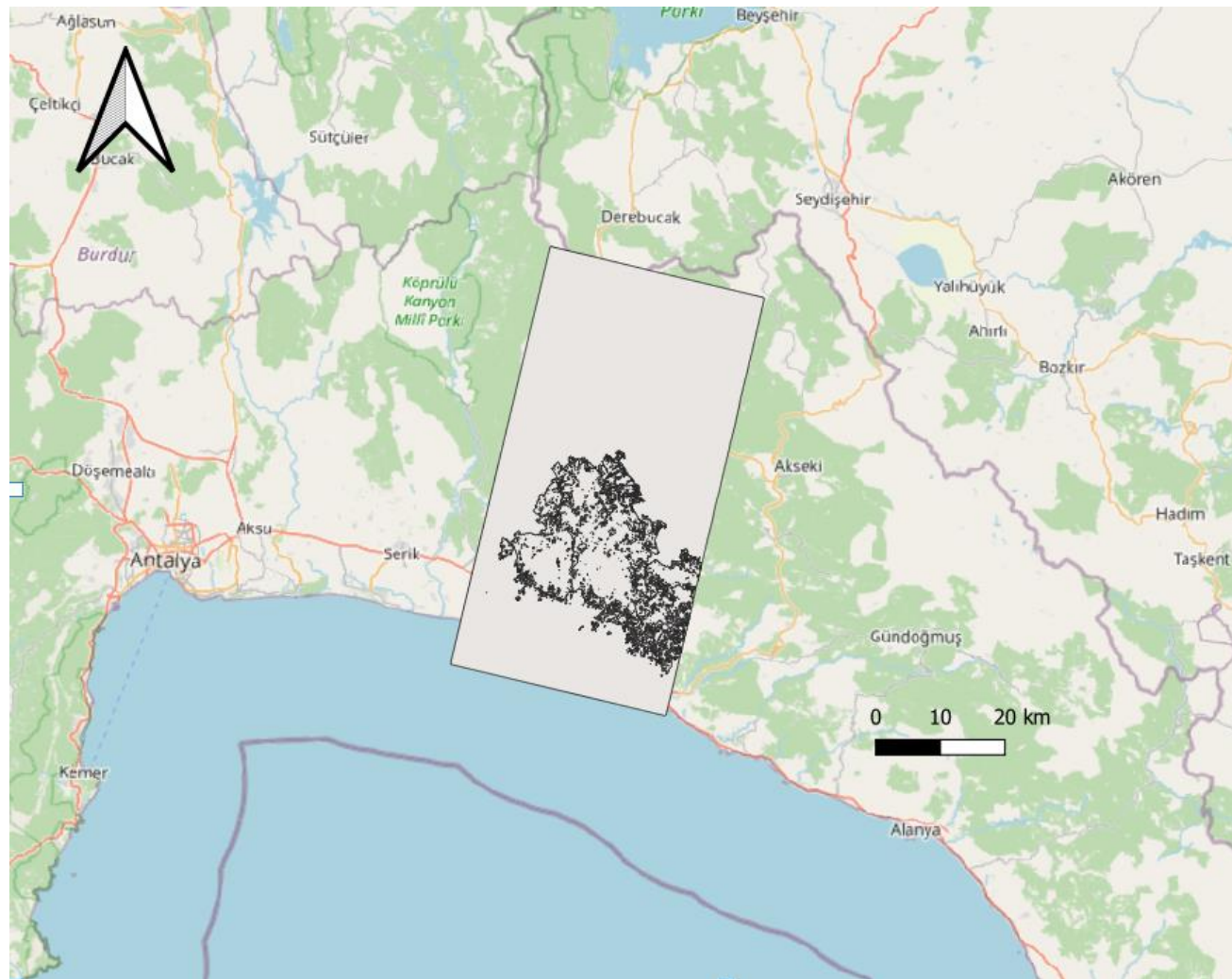


Figure 27. Classified and filtered reference vector (from post-fire Planet image)



After the computation of MD raster and OWA maps, the next step was the calculation of the accuracy metrics. The inputs for this process are the created OWA maps and a reference vector (Figure 27). To generate a reference (ground truth) vector, pre and post fire Planet images (Figure 26) were downloaded for the same dates of Sentinel-2 images.

## 6. Results and Discussion

As mentioned at the beginning of the methodology chapter, the first results were obtained as part of the tests for the codes developed for the implementation of the algorithm in QGIS. In this part, we have used the data on the fire site Pedrógão Grande, Portugal (2017). Working on this site supplied us a cross check because similar results were available from a previous study (Sali et al., 2021) but without GIS involved.

The testing section for the assessment of the algorithm performance was carried in the study area in Turkey. In this chapter, the results of the three modules of the algorithm (MD, OWA and accuracy metrics calculations) will be presented on these two Mediterranean areas.

### 6.1. MD and OWA results

#### 6.1.1. Partial Evidence (MD) results

The first step from theory to visualization of burn is carried out by applying MFs. The results are the maps of partial evidence of (Membership Degree, MD) showing the ratings of each pixel (i.e., the membership to the burned class). The MD values on the maps ( Figure 28 and Figure 30) show the burn likelihood of each pixel: higher MD (closer to 1, yellow in the figure) represents higher burn probability.

The histograms (Figure 29 and Figure 31) show the density distribution of MDs for each input band: histograms are clearly bimodal representing degrees for the unburned (peak at around 0) and burned (peak around 1) categories. For Portugal (Figure 29), the distributions of the values are better separated than the Turkey case. It means that the two classes are spectrally better separated, and it is likely that a better classification can be achieved. Less pixels with intermediate values are present and most of the pixels are concentrated toward 0 and 1. For the Turkey site we can observe that in the MD maps of the input bands  $\Delta RE2$ ,  $\Delta RE3$  and  $\Delta NIR$  there are several pixel values in the range [0, 0.2] mainly concentrated in the Northern portion of the area affected by the fire (blue regions in Figure 30 (d,e,f)). The MD maps for the Portugal site are apparently more homogeneous over the burned area with most of the values that tend to 1. However, in the Turkey case (Figure 31), density of MD=1 is relatively much lower when compared with Portugal.

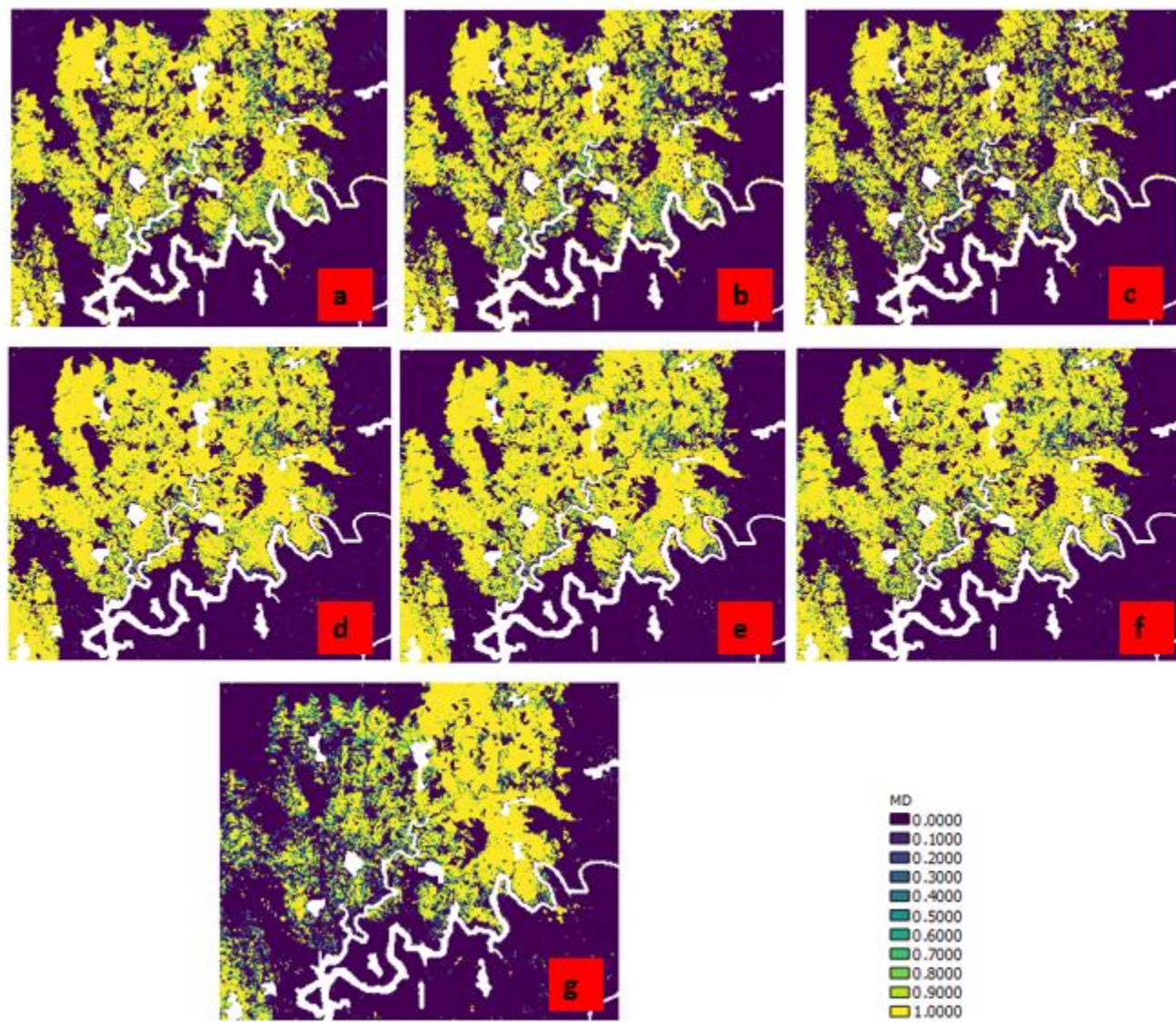


Figure 28. Sigmoid function results on Portugal site: MDs calculated for each selected band: PostRE2 (a), PostRE3 (b) PostNIR (c),  $\Delta RE2$  (d),  $\Delta RE3$  (e),  $\Delta NIR$  (f) and  $\Delta SWIR2$  (g).

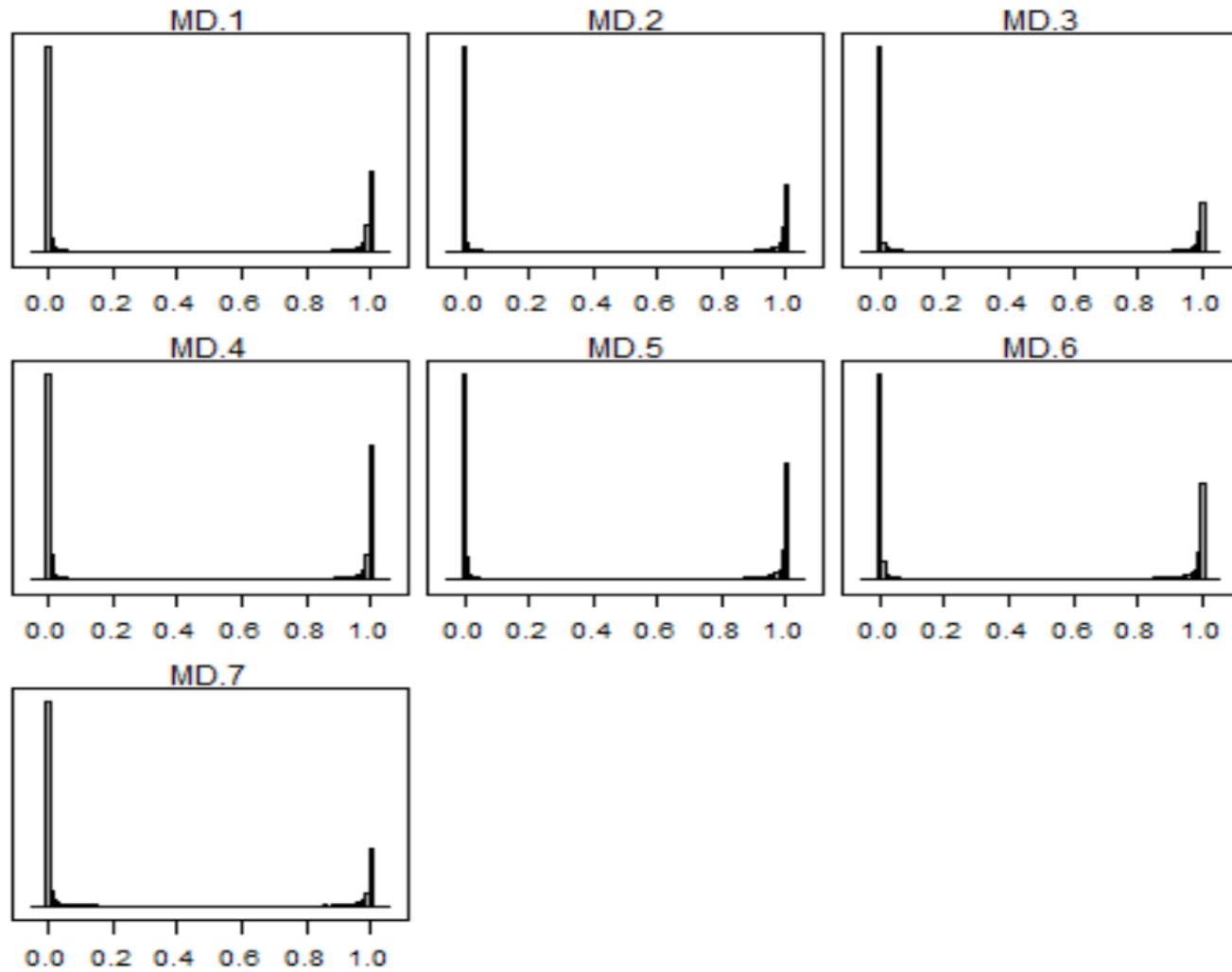


Figure 29. Portugal: Density distribution of different input layers. MD1=PostRE2, MD2=PostRE3, MD3=PostNIR, MD4= $\Delta$ RE2, MD5= $\Delta$ RE3, MD6= $\Delta$ NIR and MD7= $\Delta$ SWIR2

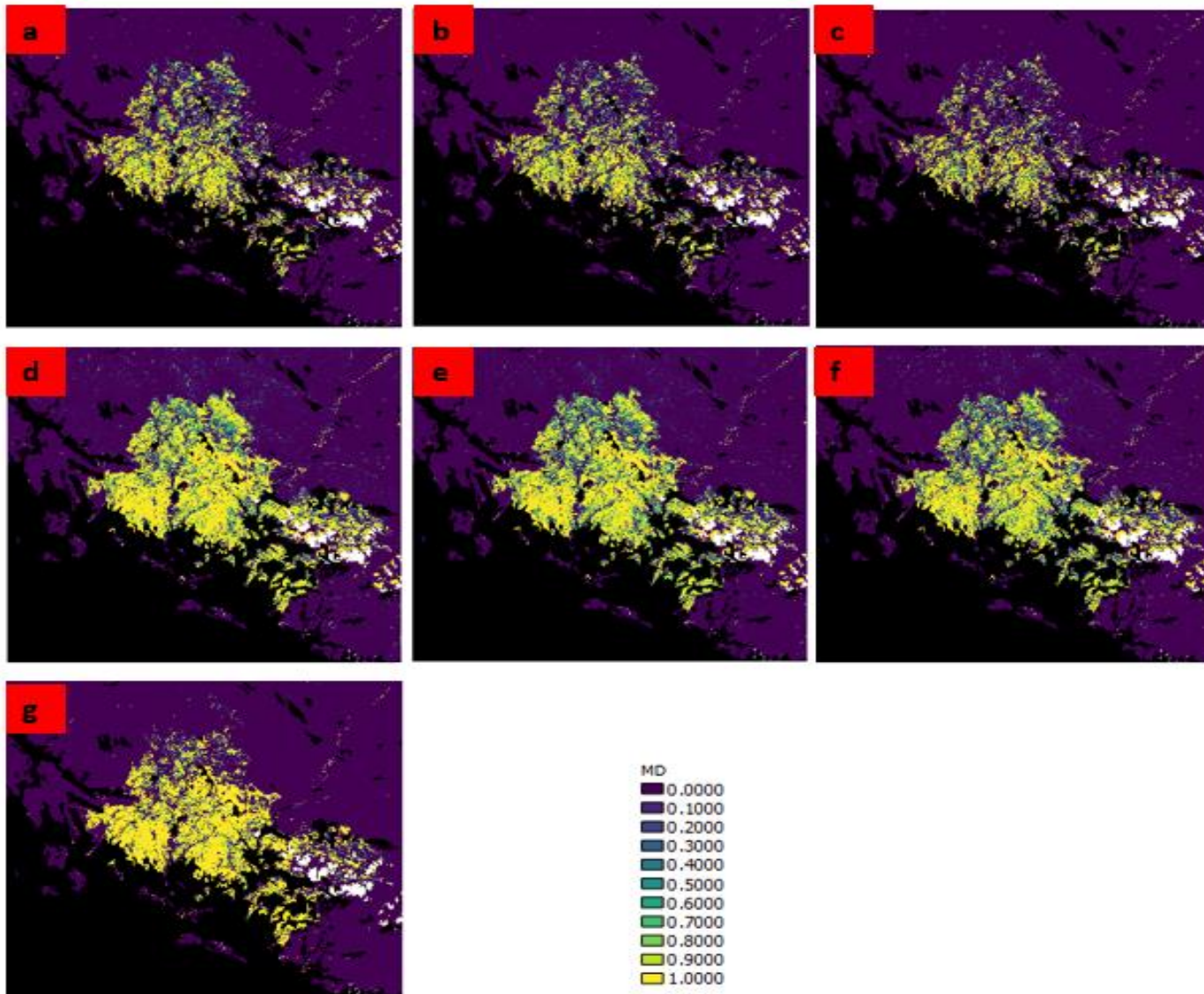


Figure 30. Sigmoid function results on Turkey site: MDs calculated for each selected band: PostRE2 (a), PostRE3 (b) PostNIR (c),  $\Delta RE2$  (d),  $\Delta RE3$  (e),  $\Delta NIR$  (f) and  $\Delta SWIR2$  (g).

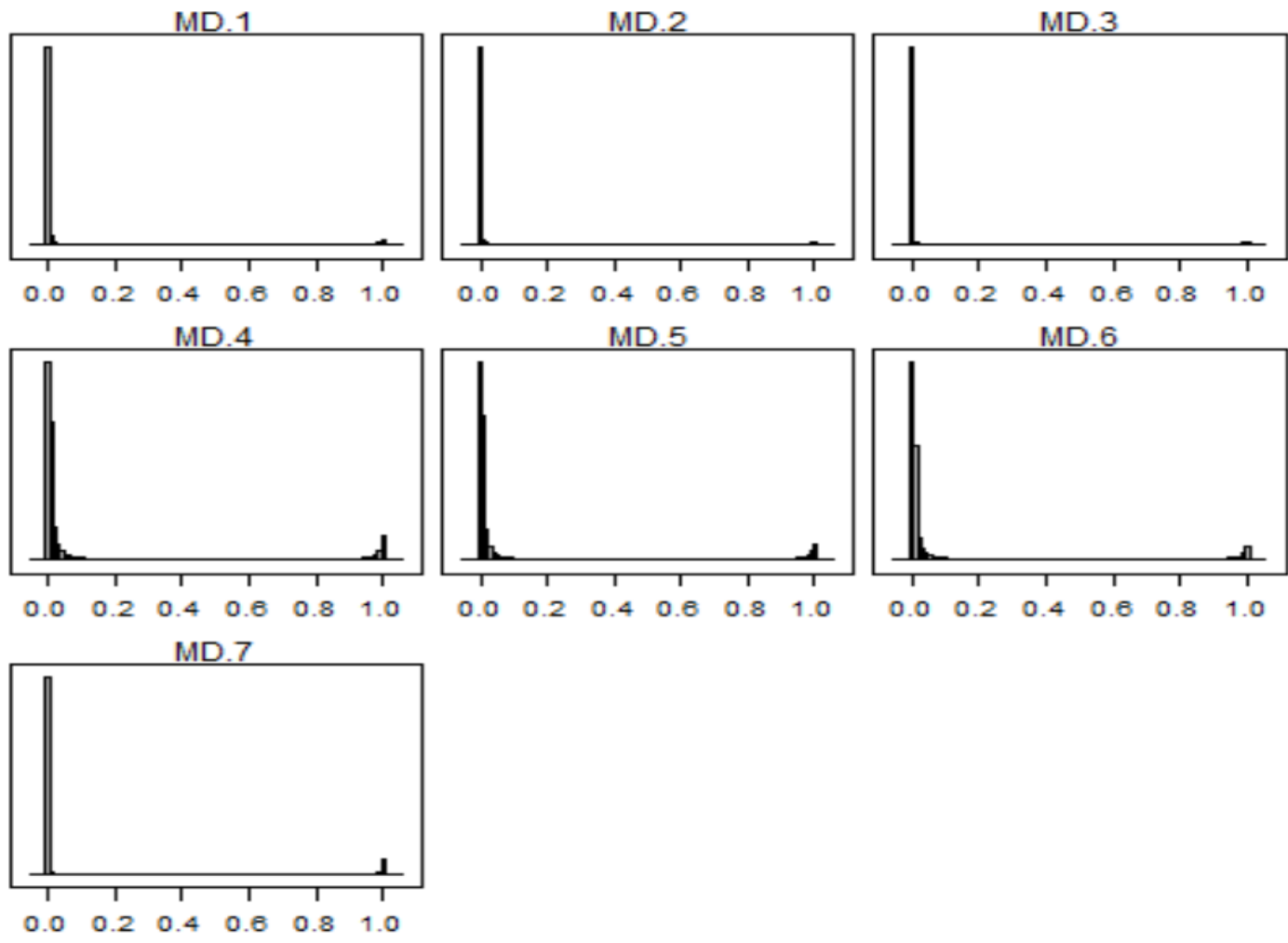


Figure 31. Turkey: Density distribution of different input layers. MD1=PostRE2, MD2=PostRE3, MD3=PostNIR, MD4= $\Delta$ RE2, MD5= $\Delta$ RE3, MD6= $\Delta$ NIR and MD7= $\Delta$ SWIR2

### 6.1.2. Global Evidence (OWA) results

By integrating the partial evidence with OWAs and by using different weighting vectors (see 4.2.2), we obtained the global evidence, the OWA maps (Figure 32 and Figure 33). OWA scores range within [0,1], with the greatest values showing points with the highest possibility of burn. As expected, the greatest proportion of burn area is obtained by the “or” ( $W=[1, 0, 0, \dots]$ ) and “almost\_or” ( $W=[0.5, 0.5, 0.0, \dots]$ ) operators that assign the highest weights to the input features with the highest partial evidence; the opposite for “and” ( $W=[0, 0, \dots, 1]$ ) and “almost\_and” ( $W=[0, 0, \dots, 0.5, 0.5]$ ) operators. These two groups of OWAs implement different attitudes/semantics being more (“and”) or less (“or”) conservative thus depicting a less or more optimistic/pessimistic picture of the fire event, respectively (Yager, 1988). “average” weighting comes with all equal participation of the seven input layers. As it can be seen from Figure 32 and Figure 33, using equal weights for each MD value yields greater spatial variability than other cases. The reason can be either different levels of severity or different land cover types across the area.

After a fire, the first priority is emergency stabilization in order to prevent further damage to life, property or natural resources. The longer-term rehabilitation effort to repair damage caused by the fire begins after the fire is extinguished and continues for several years. Rehabilitation focuses on the lands unlikely to recover naturally from wildland fire damage (URL17). OWA maps can be very helpful to land managers and planners when determining the priority areas for rehabilitation and restoration because they can be an indicator of the burn severity and of the level of damage that have occurred to the vegetation/ecosystem.

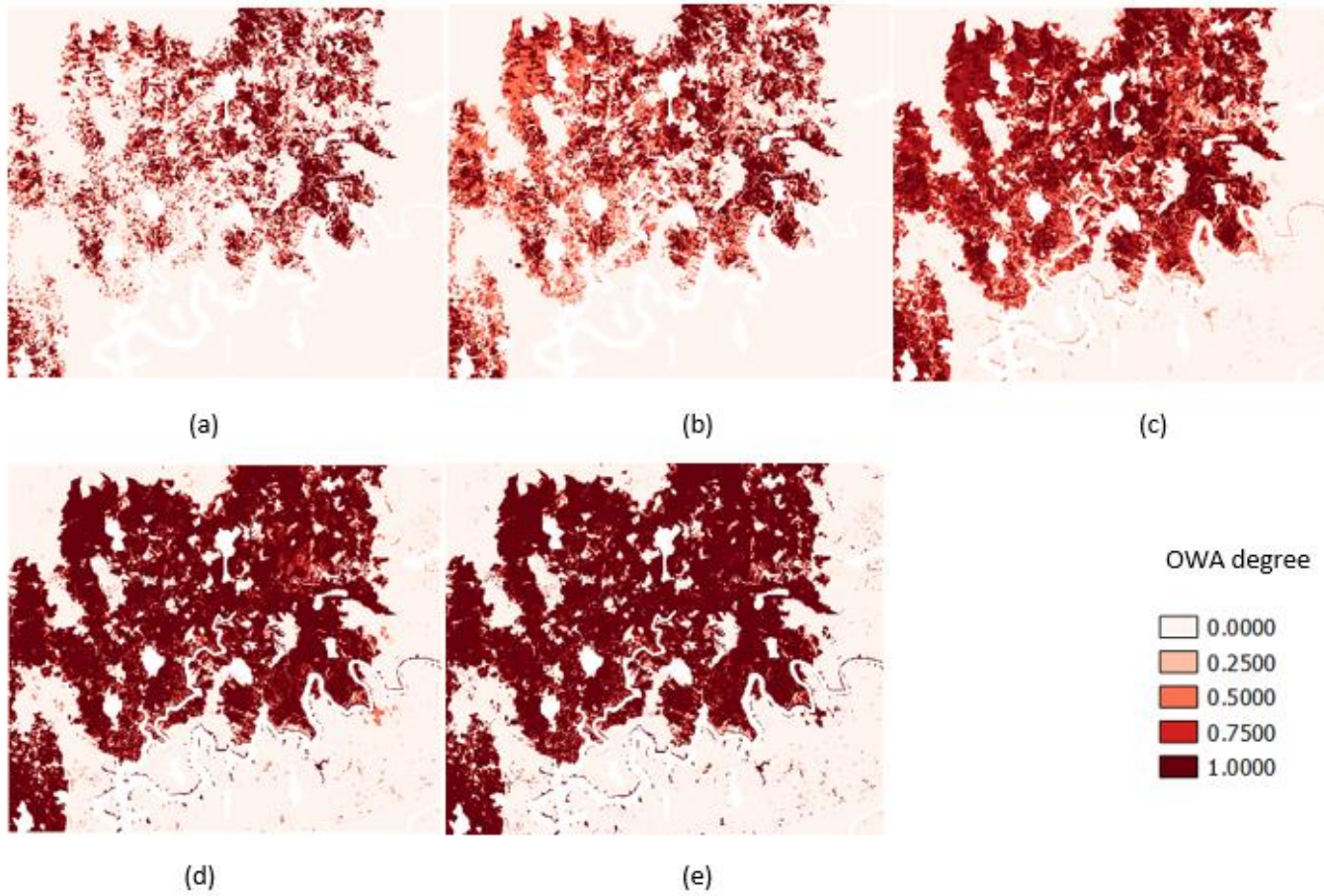


Figure 32. Portugal: OWA maps with different weighting vectors: OWA\_and(a), OWA\_almost\_and (b), OWA\_average (c), OWA\_almost\_or (d), OWA\_or (e)



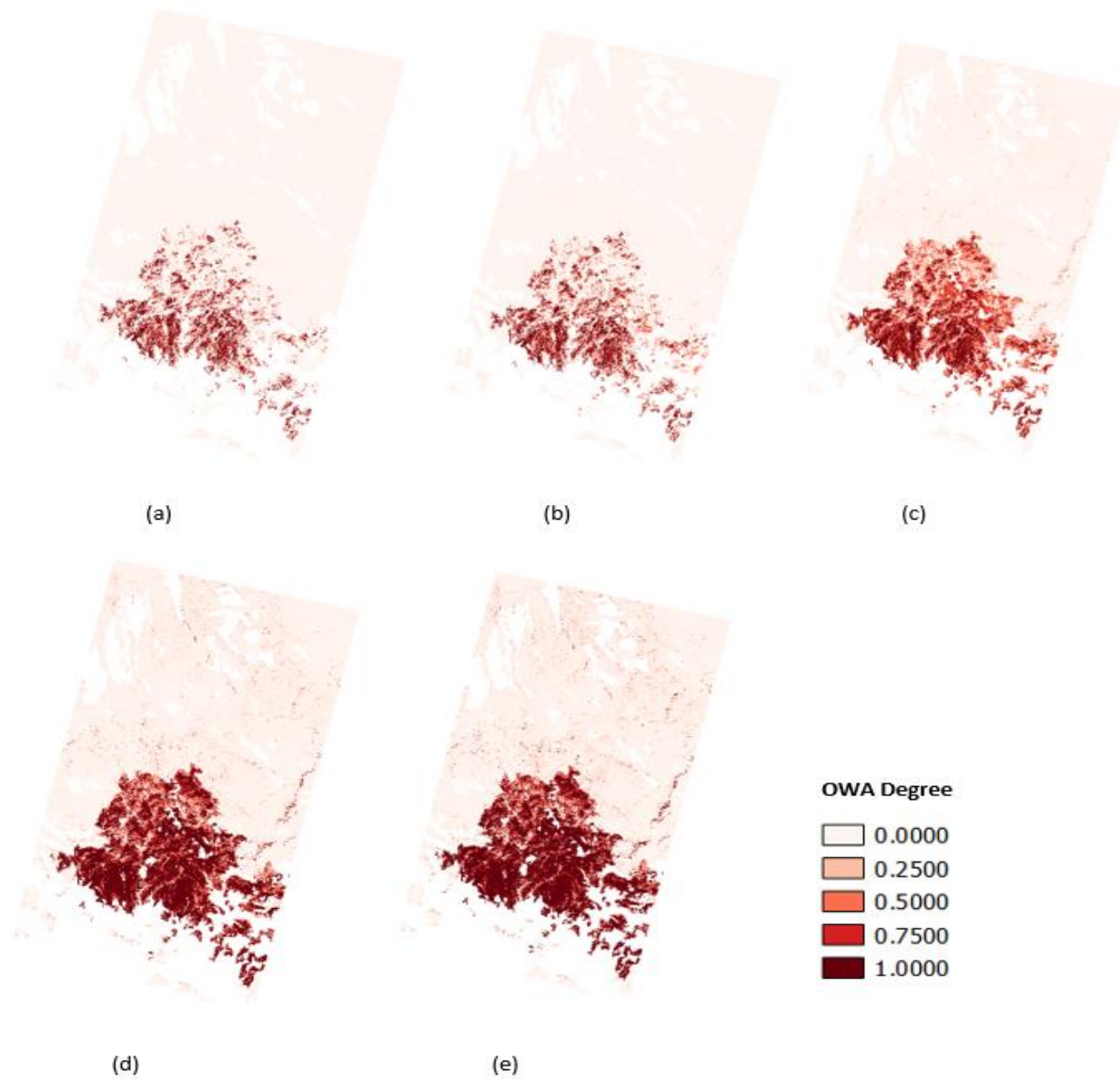


Figure 33. Turkey: OWA maps with different weighting vectors: OWA\_and(a), OWA\_almost\_and (b), OWA\_average (c), OWA\_almost\_or (d), OWA\_or (e)

## 6.2. Validation Results

The OWAs maps can be converted to burned area maps by applying a threshold and converting to a binary information burned/unburned (1/0). In this work we assumed that all pixels with OWA value greater than zero are burned. Once the Burned Area (BA) maps were generated, accuracy assessment was carried out by comparison with fire reference perimeters derived from Planet images. The comparison between classified (from OWA) and reference BA maps generates the confusion matrix and the accuracy metrics. accuracy metrics for the Portugal and Turkey sites are summarized in Table 6 and Table 7, respectively, for the five OWAs. Figure 37 and Figure 38, show the agreement maps where areas in green are presenting the agreement on unburned while the blue pixels show the commission errors, the yellow pixels show the omission errors, and the red ones represent agreement on burned with both OWA and reference (Planet) dataset.

As we discussed in the partial evidence section, it is proved that the algorithm gives a higher accuracy in Portugal. Looking at the Dice coefficient, that provides a synthetic measure of omission and commission errors, in Portugal site  $DC > 0.67$  with best values for the “or”, “average” and “almost-or” operators.

Table 6. Accuracy Metrics of Portugal fire

OWA weights	Omission	Commission	Dice Coefficient	Bias	Relative Bias
and	0.485	0.003	0.679	-586609	-0.484
almost_and	0.296	0.003	0.825	-356360	-0.294
<b>average</b>	<b>0.107</b>	<b>0.047</b>	<b>0.922</b>	<b>-75713</b>	<b>-0.062</b>
<b>almost_or</b>	<b>0.087</b>	<b>0.059</b>	<b>0.927</b>	<b>-36793</b>	<b>-0.030</b>
<b>or</b>	<b>0.083</b>	<b>0.064</b>	<b>0.926</b>	<b>-23918</b>	<b>-0.020</b>

Table 7. Accuracy Metrics of Turkey fire

OWA weights	Omission	Commission	Dice Coefficient	Bias	Relative Bias
and	0.600	0.004	0.572	-2139872	-0.597
almost_and	0.506	0.020	0.657	-1777218	-0.496
<b>average</b>	<b>0.089</b>	<b>0.111</b>	<b>0.900</b>	<b>90725</b>	<b>0.025</b>
almost_or	0.043	0.195	0.874	679092	0.189
or	0.036	0.232	0.855	911920	0.254

The results show that in Portugal, for all applied OWAs, the omission error, ranging in [8,3-48,5%], is greater than the commission error, [0.3-6.4%]. On the other hand, in Turkey, while the omission errors are in a similar range [3,6-60%] with Portugal, commission errors are significantly greater especially for OWA\_almost\_or (19,5%) and OWA\_or (23,2%). This increase leads to a decrease in the Dice coefficient with the lowest value of 0.57 for the “and” operator.

There can be possible reasons for the observed difference in the performance of the algorithm in the two sites; for example, the Portugal fire was very severe, in this case the classification can be much clearer than a moderately severe case. Alternatively, the land cover classes which are burned can be deeply green before the fire and the change in the signals can be easily noticeable in that case. Hence, several reasons concur to the greater accuracy of the BA maps for the Portugal site: fire severity, vegetation characteristics and conditions.

As it can be seen from Figure 38, most of the commission error in the Turkey site comes from the northern part of the area, which is outside of the fire perimeter. This can be because during the algorithm run, we defined a threshold for OWA maps for binary classification (burned or not). Both for Portugal and Turkey, we set the threshold value as 0.1. That means for any of the OWA maps, if there is a pixel with a value higher than 0.1, it is assumed as positive evidence of burn. So, even if the northern area has small values (around 0.2-0.3) but higher than 0.1, they are also assumed as burned. We can see the same result of thresholding on Portugal agreement maps (Figure 37), however the effect on the accuracy is less than the effect in Turkey. The reason for this difference is most likely that the parameters are not tuned for Turkey at the beginning of MD computation algorithm. When we look at Figure 36, the commission error mainly comes from salt and pepper effect. The topography in the northern region contains large hills and rocks. This nature results in shadows and can rise the commission error because the reflectance in this region is close to but higher than our threshold. Also, the masking vector that we used to define burnable areas is the Corine Land Cover product of Copernicus programme which is produced for European countries. Using a local land cover map for Turkey can improve these kinds of errors because they may supply a better classification.

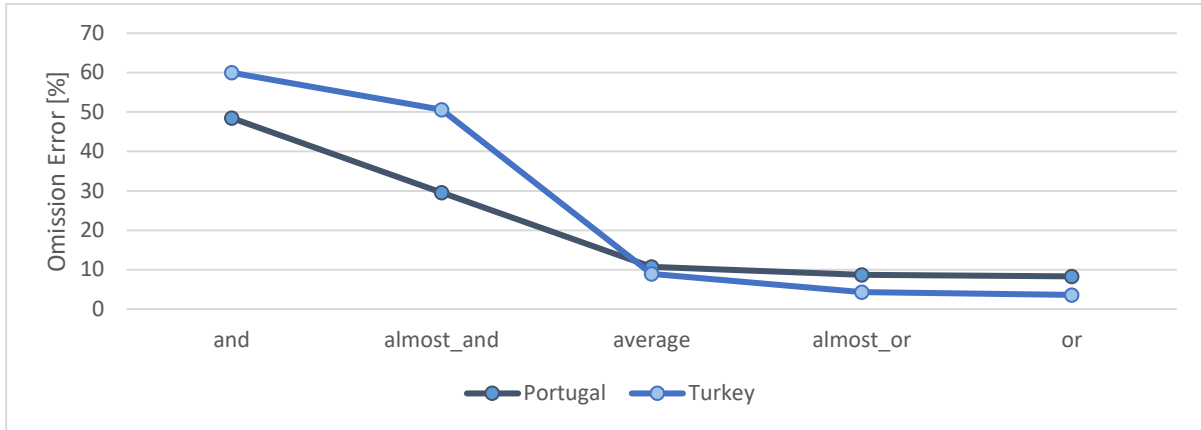


Figure 34. Percentage of omission errors for five different OWA operators

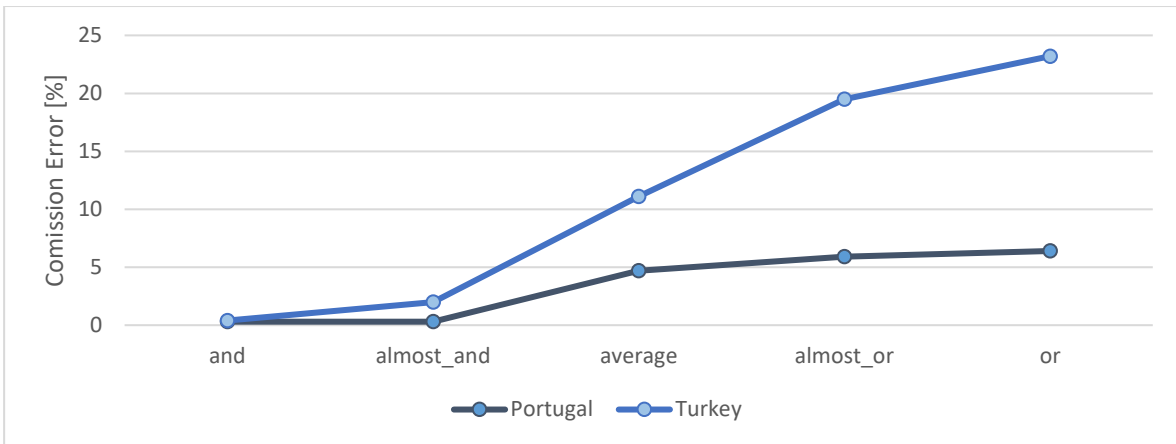


Figure 35. Percentage of commission errors for five different OWA operators

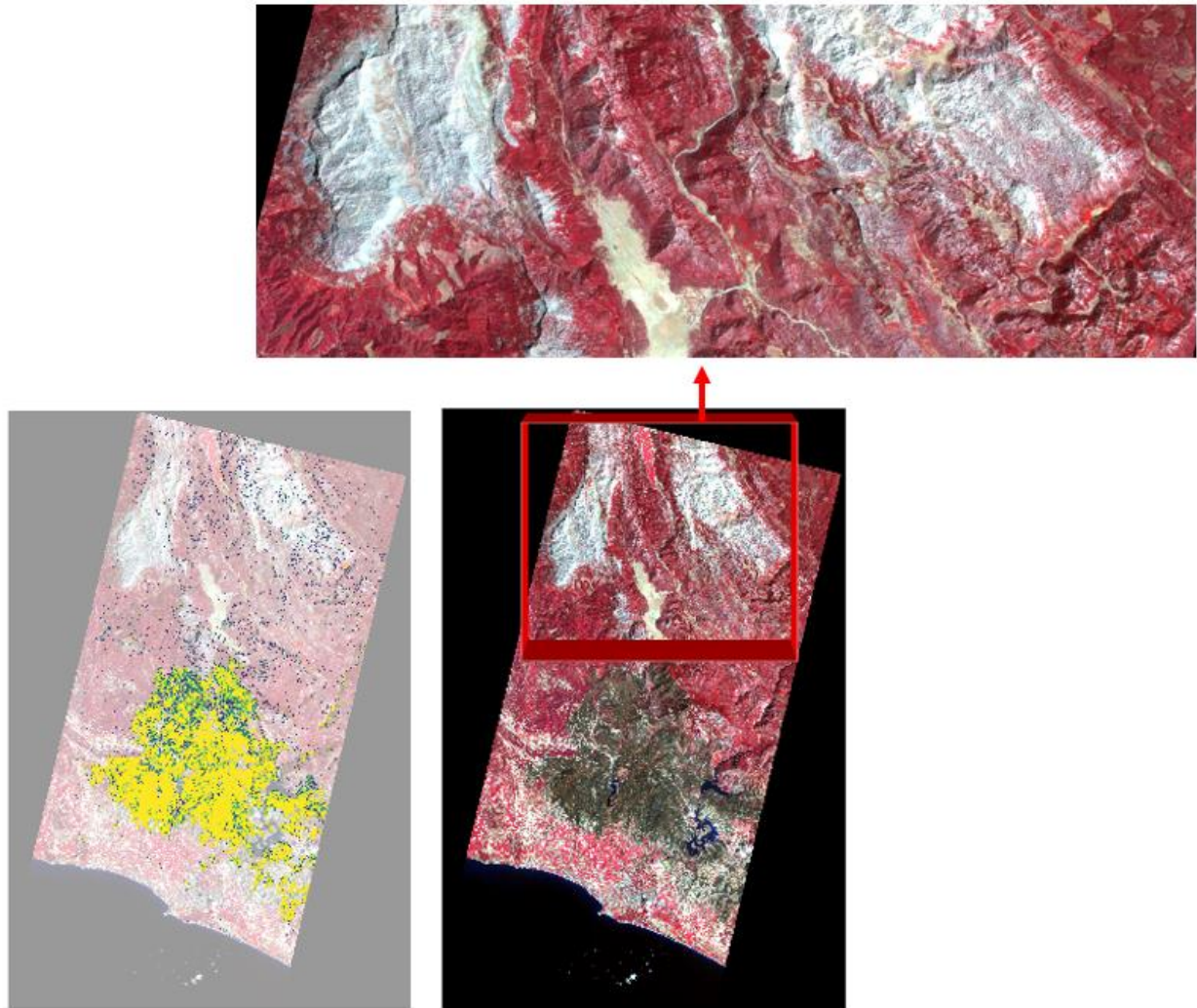


Figure 36. OWA\_or map (excluding the pixels with values smaller than 0.1) and Planet image of the area

Ultimately, the aim of this study is the automatization of the algorithms for regional scale in a more user-friendly way. Therefore, as long as the overall accuracy falls in an acceptable range, tuning is not necessary for each different site, and this proves the exportability of the algorithm. When we look at the both results of the previous study on Mediterranean regions (Sali et al., 2021) and our results, we can say that the algorithm is successfully integrated to QGIS and the algorithm itself is suitable to detect burned area in Mediterranean region even without further calibration.

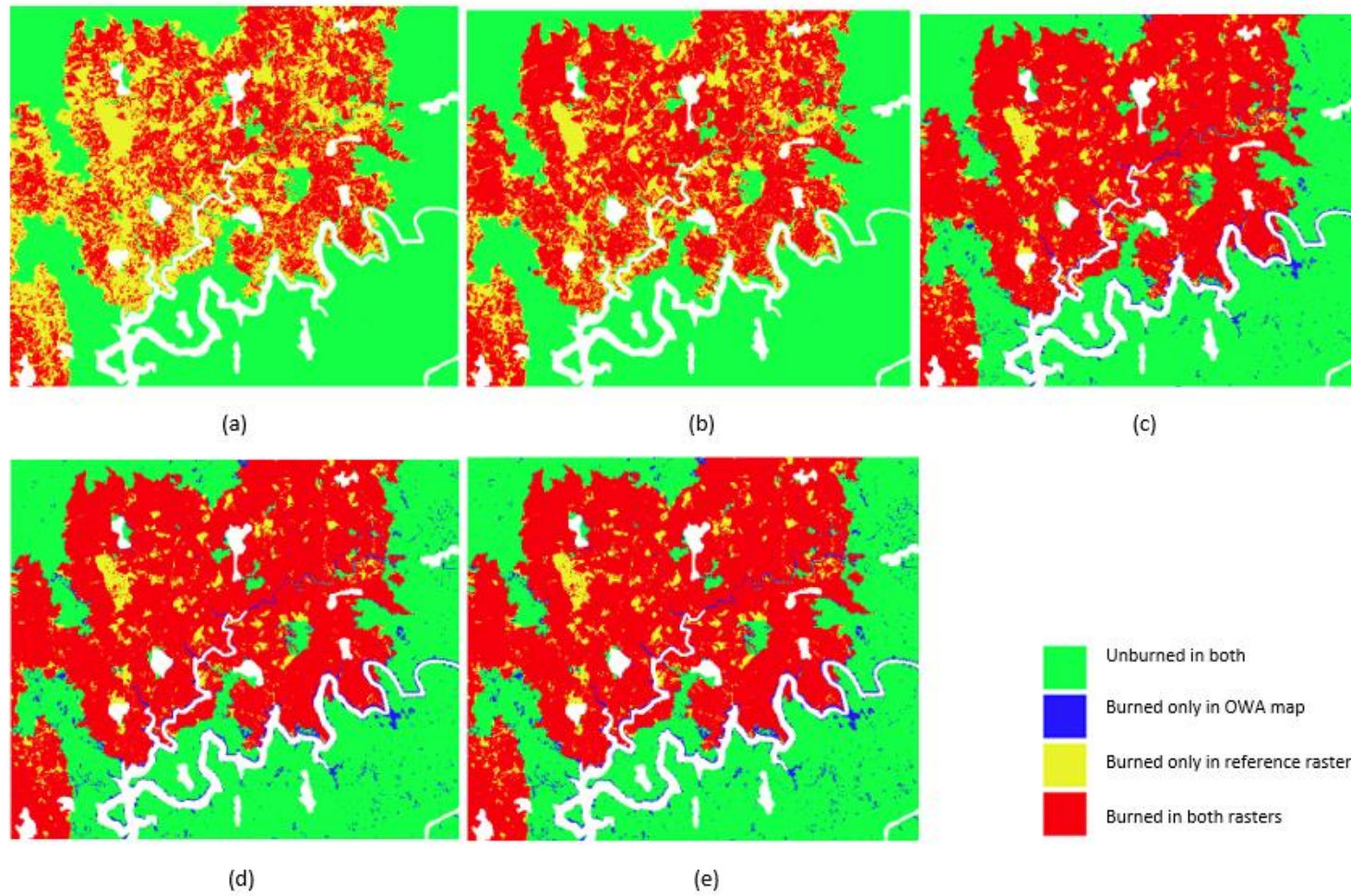


Figure 37. Agreement maps by OWA maps and reference vector of Portugal: Agreement\_and (a), Agreement\_almost\_and (b), Agreement\_average (c), Agreement\_almost\_or (d), Agreement\_or (e)

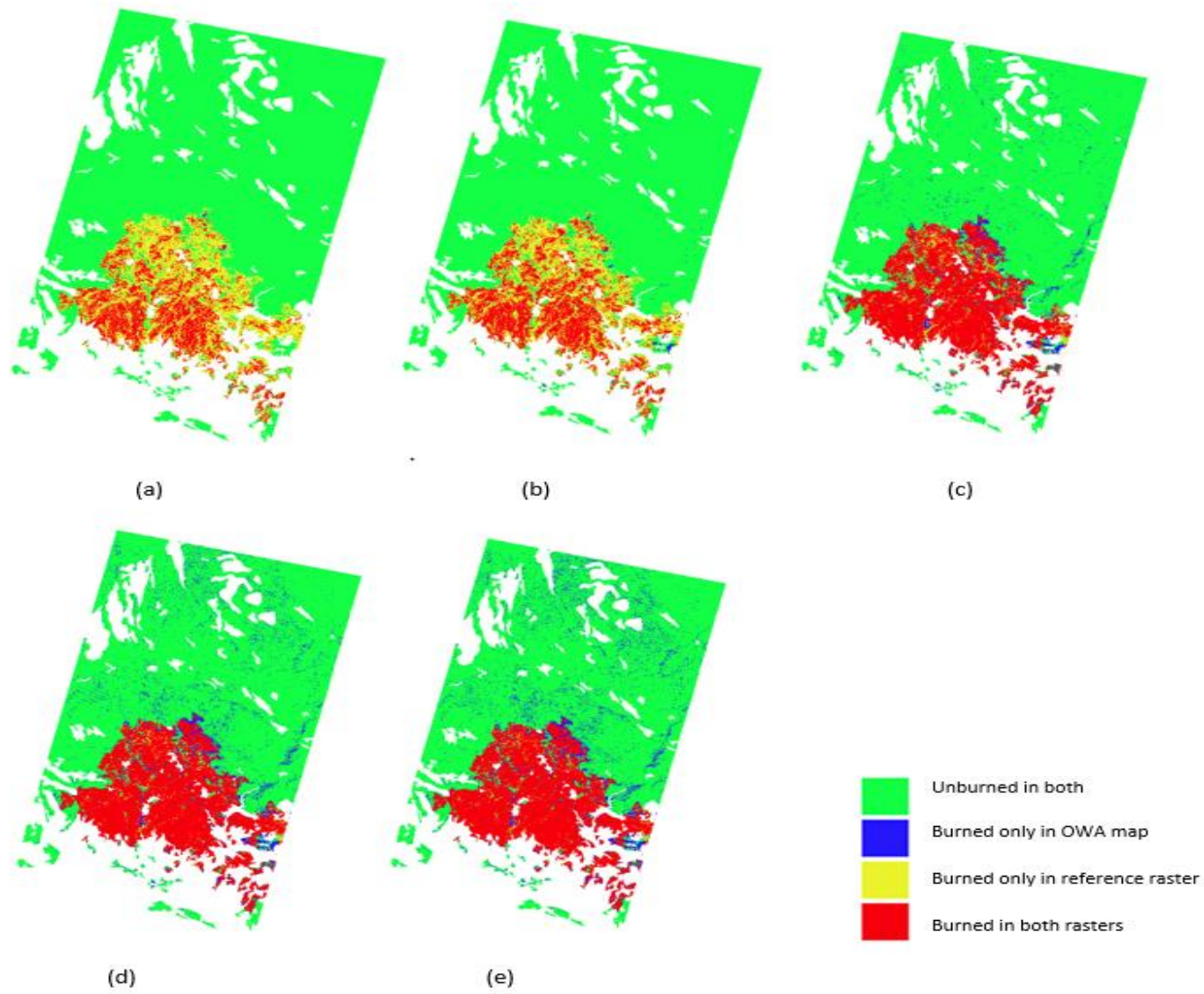


Figure 38. Agreement maps by OWA maps and reference vector of Turkey: Agreement\_and (a), Agreement\_almost\_and (b), Agreement\_average (c), Agreement\_almost\_or (d), Agreement\_or (e)

## 7. Conclusion

In this thesis the objective is implementing a soft computing algorithm in a QGIS workflow which is a user-friendly environment for displaying and visualisation of the results and testing the performance of the algorithm over a new site in Turkey to assess the exportability of the BA mapping process over the Mediterranean region. The algorithm has three main steps i) membership degree (MD, partial evidence) computation ii) ordered weighted averages (OWA, global evidence) computation and iii) validation. The results show that the algorithm which was coded as separate parts in R environment is successfully integrated into QGIS and now the algorithm can be run on an easier user interface. In Portugal, for all applied OWAs, the omission error (oe), ranging in [8,3-48,5%], is greater than the commission error (ce), [0.3-6.4%]. On the other hand, in Turkey, while the omission errors are in a similar range [3,6-60%] with Portugal, commission errors are significantly greater especially for OWA\_almost\_or (19,5%) and OWA\_or (23,2%). This increase leads to a decrease in the Dice coefficient with the lowest value of 0.57 for the “and” operator. Most of the commission error in the Turkey site comes from the northern part of the area, which is outside of the fire perimeter. Two of the possible reasons for the lower accuracy are i) the topography in the northern region of Turkey site contains large hills and rocks. This nature results in shadows and can rise the commission error because the reflectance in this region is close to but higher than our threshold (0.1) ii) the parameters are not tuned for Turkey at the beginning of MD computation algorithm. Despite that, the OWA\_average operator, which results with the best values in Turkey, still provides an acceptable oe (8.9%) and ce (11.1) with a high Dice coefficient (90%). Finally, results of the previous study on Mediterranean regions with the same algorithm (Sali et al., 2021) and our results together show that the algorithm is successfully integrated to QGIS and the algorithm itself is suitable to detect burned area in the Mediterranean region even without further calibration.



## 8. BIBLIOGRAPHY

- Brunsdon, C., & Comber, L. (2019). *An Introduction to R for Spatial Analysis and Mapping*.
- Chang, J., & Clay, D. (2016). *Matching Remote Sensing to Problems*. November.
- Chapman, T., Flegel, C., Walters, D., & Shrubsole, D. (2003). *GIS – Multicriteria Evaluation with Ordered Weighted Averaging (OWA): Case Study of Developing Watershed Management Strategies* *GIS ^ multicriteria evaluation with ordered weighted averaging (OWA): case study of developing watershed management strategi*. October. <https://doi.org/10.1068/a35156>
- Chergui, B., Fahd, S., Santos, X., & Pausas, J. G. (2018). Socioeconomic Factors Drive Fire-Regime Variability in the Mediterranean Basin. *Ecosystems*, 21(4), 619–628. <https://doi.org/10.1007/s10021-017-0172-6>
- Chuvieco, E. (1999). *Remote Sensing of Large Wildfires in the European Mediterranean Basin*.
- Chuvieco, E., Lizundia-loiola, J., Pettinari, M. L., Ramo, R., Padilla, M., Tansey, K., Mouillot, F., Laurent, P., Storm, T., & Heil, A. (2018). *Generation and analysis of a new global burned area product based on MODIS 250 m reflectance bands and thermal anomalies*. 2015–2031.
- 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 Sensing of Environment*, 225(February), 45–64. <https://doi.org/10.1016/j.rse.2019.02.013>
- Cogato, A., Meggio, F., Collins, C., & Marinello, F. (2020). *Medium-Resolution Multispectral Data from Sentinel-2 to Assess the Damage and the Recovery Time of Late Frost on Vineyards*.
- Coops, N. C., & Tooke, T. R. (2017). *Introduction to Remote Sensing*. <https://doi.org/10.1007/978-1-4939-6374-4>
- 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 Sensing*, 11(6). <https://doi.org/10.3390/rs11060622>
- Fitriana, H. L., Prasasti, I., & Khomarudin, M. R. (2018). *Mapping burned areas from landsat-8 imageries on mountainous region using reflectance changes*. 12, 1–7.
- Flenniken, J. M., Stuglik, S., & Iii, B. V. I. (2020). *Quantum GIS (QGIS): An Introduction to a Free Alternative to More Costly GIS Platforms 1*. 1–7.
- Giglio, L., Boschetti, L., Roy, D. P., Humber, M. L., & Justice, C. O. (2018). Remote Sensing of Environment The Collection 6 MODIS burned area mapping algorithm and product. *Remote Sensing of Environment*, 217(July), 72–85. <https://doi.org/10.1016/j.rse.2018.08.005>
- Goodwin, N. R., & Collett, L. J. (2014). Remote Sensing of Environment Development of an automated method for mapping fire history captured in Landsat TM and ETM + time series across Queensland, Australia. *Remote Sensing of Environment*, 148, 206–221. <https://doi.org/10.1016/j.rse.2014.03.021>
- Heisig, J. (2021). *Detecting Drought Effects on Tree Mortality in Forests of Franconia with Detecting Drought Effects on Tree Mortality in Forests of Franconia with Remote Sensing Master 's Thesis submitted Chair of Climatology*. January. <https://doi.org/10.13140/RG.2.2.27763.35364>

- Indratmoko, S., & Rizqihandari, N. (2019). Burn Area Detection Using Landsat 8 OLI TIRS Burn Area Detection Using Landsat 8 OLI TIRS. *IOP Conference Series: Earth and Environmental Science*. <https://doi.org/10.1088/1755-1315/338/1/012035>
- Isbaex, C., & Coelho, A. M. (2021). *The Potential of Sentinel-2 Satellite Images for Land-Cover/ Land-Use and Forest Biomass Estimation: A Review*.
- Lentile, L. B., Holden, Z. A., Smith, A. M. S., Falkowski, M. J., Hudak, A. T., Morgan, P., Lewis, S. A., Gessler, P. E., & Benson, N. C. (2006). Remote sensing techniques to assess active fire characteristics and post-fire effects. *International Journal of Wildland Fire*, 319–345.
- 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 Sensing of Environment*, 236. <https://doi.org/10.1016/j.rse.2019.111493>
- Luca, G. De, Silva, J. M. N., Modica, G., Luca, G. De, Silva, J. M. N., A, G. M., & Luca, G. De. (2021). A workflow based on Sentinel-1 SAR data and open-source algorithms for unsupervised burned area detection in Mediterranean ecosystems. *GIScience & Remote Sensing*, 58(4), 516–541. <https://doi.org/10.1080/15481603.2021.1907896>
- Lydersen, J. M., Collins, B. M., Miller, J. D., Fry, D. L., & Stephens, S. L. (2016). RELATING FIRE-CAUSED CHANGE IN FOREST STRUCTURE TO REMOTELY SENSED ESTIMATES OF FIRE SEVERITY. *Fire Ecology*, 12(3), 99–116. <https://doi.org/10.4996/fireecology.1203099>
- Mohammed, Z., Deshmukh, R. ., & Syed, S. B. (2017). An Overview of Different Optical Remote Sensing Techniques. *International Journal of Engineering and Advanced Technology (IJEAT)*, 6(17), 2249–8958. [https://www.researchgate.net/publication/320947556\\_An\\_Overview\\_of\\_Different\\_Optical\\_Remote\\_Sensing\\_Techniques](https://www.researchgate.net/publication/320947556_An_Overview_of_Different_Optical_Remote_Sensing_Techniques)
- Mohan, S. (2015). *Potentials and applications of Microwave remote sensing*. July.
- Morissette, J. T., Baret, F., Privette, J. L., Myneni, R. B., Nickeson, J. E., Garrigues, S., Shabanov, N. V., Weiss, M., Fernandes, R. A., Leblanc, S. G., Kalacska, M., Sánchez-azofeifa, G. A., Chubey, M., Rivard, B., Stenberg, P., Rautiainen, M., Voipio, P., Manninen, T., Pilant, A. N., ... Seufert, G. (2006). *Validation of Global Moderate-Resolution LAI Products : A Framework Proposed Within the CEOS Land Product Validation Subgroup*. 44(7), 1804–1817.
- Muenchow, J., Schratz, P., & Brenning, A. (2017). RQGIS: Integrating R with QGIS. *The R Journal*, 9(2), 409–428. <https://cran.r-project.org/package=RQGIS>
- Ngadze, F., Shaun, K., Id, M., Kavhu, B., Ndaimani, H., & Maremba, S. (2020). *Exploring the utility of Sentinel-2 MSI and Landsat 8 OLI in burned area mapping for a heterogenous savannah landscape*. 1–13. <https://doi.org/10.1371/journal.pone.0232962>
- One-Stop GIS , Stages in GIS Development*. (2021).
- Pausas, J. G., & Keeley, J. E. (2021). *Wildfires and global change*. 387–395. <https://doi.org/10.1002/fee.2359>
- Rahman, S., & Chang, H.-C. (2017). Assessment of Fire Severity and Vegetation Response using Moderate-Resolution Imaging Spectroradiometer. 11. *International Conference on Sensing Technology (ICST)*.

- Ramo, R., Garcia, M., Rodríguez, D., & Chuvieco, E. (2018). A data mining approach for global burned area mapping. *International Journal of Applied Earth Observation and Geoinformation*, 73(April), 39–51. <https://doi.org/10.1016/j.jag.2018.05.027>
- Ranghetti, L., National, I., Nutini, F., National, I., Busetto, L., & National, I. (2020). “sen2r”: An R toolbox for automatically downloading and preprocessing Sentinel-2 satellite data “sen2r”: an R toolbox for automatically downloading and preprocessing Sentinel-2. June. <https://doi.org/10.1016/j.cageo.2020.104473>
- Rogers, B. M., Balch, J. K., Goetz, S. J., Lehmann, C. E. R., & Turetsky, M. (2020). *Focus on changing fire regimes : interactions with climate , ecosystems , and society Focus on changing fire regimes : interactions with climate , ecosystems , and society*.
- Roteta, E., Bastarrika, A., Padilla, M., Storm, T., & Chuvieco, E. (2019). Development of a Sentinel-2 burned area algorithm : Generation of a small fire database for sub-Saharan Africa Remote Sensing of Environment Development of a Sentinel-2 burned area algorithm : Generation of a small fire database for sub-Saharan Africa. *Remote Sensing of Environment*, 222(January), 1–17. <https://doi.org/10.1016/j.rse.2018.12.011>
- Sali, M., Piasek, E., Boschetti, M., Brivio, P. A., Sona, G., Bordogna, G., Stroppiana, D., & Marcos, E. (2021). *A Burned Area Mapping Algorithm for Sentinel-2 Data Based on Approximate Reasoning and Region Growing*. 1–28.
- San-miguel-ayanz, J., Rodrigues, M., Oliveira, S. S. De, Pacheco, C. K., Moreira, F., & Duguay, B. (2012). *Land Cover Change and Fire Regime in the European Mediterranean Region Land Cover Change and Fire Regime*. October 2017. <https://doi.org/10.1007/978-94-007-2208-8>
- Santis, A. De, & Chuvieco, E. (2007). *Burn severity estimation from remotely sensed data : Performance of simulation versus empirical models Burn severity estimation from remotely sensed data : Performance of simulation versus empirical models*. June. <https://doi.org/10.1016/j.rse.2006.11.022>
- Stroppiana, D., Bordogna, G., Carrara, P., Boschetti, M., Boschetti, L., & Brivio, P. A. (2012). A method for extracting burned areas from Landsat TM/ETM+ images by soft aggregation of multiple Spectral Indices and a region growing algorithm. *ISPRS Journal of Photogrammetry and Remote Sensing*, 69, 88–102. <https://doi.org/10.1016/j.isprsjprs.2012.03.001>
- Tansey, K., Gre, J., Stroppiana, D., Pereira, M. C., Boschetti, L., Maggi, M., Brivio, P. A., & Fraser, R. (2004). *Vegetation burning in the year 2000 : Global burned area estimates from SPOT VEGETATION data*. 109, 1–22. <https://doi.org/10.1029/2003JD003598>
- Vanderhoof, M. K., Hawbaker, T. J., Teske, C., Ku, A., Noble, J., & Picotte, J. (2021). *Mapping Wetland Burned Area from Sentinel-2 across the Southeastern United States and Its Contributions Relative to Landsat-8 ( 2016 – 2019 )*. 8.
- Wallach, D., Makowski, D., W.Jones, J., & Brun, F. (2019). *Working with Dynamic Crop Models*. <https://doi.org/10.1016/B978-0-12-811756-9.00002-2>
- WHO. (2018). *Comparison of Geographic Information Systems (GIS) software* (Issue January).
- Yager, R. R. (1988). *On Ordered Weighted Averaging Aggregation Operators in Multicriteria Decisionmaking*. 1, 183–190.
- Zhuge, X., & Zou, X. (2017). *A Fast Cloud Detection Algorithm Applicable to Monitoring and Nowcasting*

*of Daytime Cloud Systems. October.* <https://doi.org/10.1109/TGRS.2017.2720664>

- URL1:** [https://cimss.ssec.wisc.edu/sage/remote\\_sensing/lesson3/concepts.html](https://cimss.ssec.wisc.edu/sage/remote_sensing/lesson3/concepts.html), last accessed December 2021
- URL2:** <https://www.nrcan.gc.ca/maps-tools-and-publications/satellite-imagery-and-air-photos/tutorial-fundamentals-remote-sensing/microwave-remote-sensing/9371>, last accessed December 2021
- URL3:** <https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-2/mission-objectives>, last accessed December 2021
- URL4:** [https://www.esa.int/Applications/Observing\\_the\\_Earth/Copernicus/Sentinel-2\\_overview](https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2_overview), last accessed February 2022
- URL5:** <https://www.esri.com/en-us/what-is-gis/overview>, last accessed January 2022
- URL6:** <https://www.bcs.org/articles-opinion-and-research/the-history-of-geographic-information-systems-gis/>, last accessed January 2022
- URL7:** <https://volaya.github.io/gis-book/en/History.html>, last accessed March 2022
- URL8:** <https://www.geeksforgeeks.org/r-programming-language-introduction/>, last accessed January 2022
- URL9:** <https://geocompr.robinlovelace.net/intro.html>, last accessed March 2022
- URL10:** <https://gogeomatics.ca/what-are-the-top-programming-languages-in-the-gis-world>, last accessed December 2021
- URL11:** <https://north-road.github.io/qgis-processing-r/script-syntax/>, last accessed January 2022
- URL12:** [https://docs.qgis.org/3.4/en/docs/user\\_manual/appendices/qgis\\_r\\_syntax.html](https://docs.qgis.org/3.4/en/docs/user_manual/appendices/qgis_r_syntax.html), last accessed December 2021
- URL13:** <https://www.analyticsvidhya.com/blog/2021/06/understanding-random-forest/>, last accessed February 2022
- URL14:** <https://medium.datadriveninvestor.com/deep-learning-in-medical-imaging-3c1008431aaf>, last accessed March 2022
- URL15:** <https://www.aljazeera.com/news/2021/8/3/turkey-wildfires-despair-and-questions-as-southern-coast-burns>, last accessed February 2022
- URL16:** <https://www.gdacs.org/report.aspx?eventid=1001229&episodeid=7&eventtype=WF>, last accessed March 2022
- URL17:** <https://www.fs.usda.gov/science-technology/fire/after-fire>, last accessed February 2022