



**Politecnico  
di Torino**

POLITECNICO DI TORINO

Master degree course in Computer Engineering (Data Science)

Master Degree Thesis

# Hotspot Disambiguation

Machine Learning techniques to improve existing approaches

**Supervisor**

prof. Paolo Garza

**Candidate**

Edoardo CALVI

matricola: 259044

**Internship Tutor**

dott. Edoardo Arnaudo

ACADEMIC YEAR 2020-2021

This work is subject to the Creative Commons Licence

# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
<b>2</b>	<b>Data Sources and Related works</b>	<b>9</b>
2.1	MODIS . . . . .	9
2.2	VIIRS . . . . .	11
2.3	Copernicus EMS and EFFIS . . . . .	11
2.4	Sentinel imagery usage . . . . .	13
2.5	Land Cover . . . . .	14
<b>3</b>	<b>Data Preparation</b>	<b>15</b>
3.1	Collections and Preparation . . . . .	15
3.1.1	Hotspot products . . . . .	15
3.1.2	Land Cover . . . . .	19
3.1.3	Sentinel-3 OLCI . . . . .	22
3.1.4	EFFIS burnt areas . . . . .	25
3.1.5	Copernicus EMS burnt areas . . . . .	26
3.2	Dataset Creation . . . . .	27
3.2.1	Ground Truth . . . . .	27
3.2.2	Dataset Balancing . . . . .	32
3.2.3	Incremental Data Integration . . . . .	37
3.3	Further Preprocessing and Encodings . . . . .	38
<b>4</b>	<b>Methodology</b>	<b>43</b>
4.1	Adopted Models . . . . .	44
4.2	Frameworks and Tools . . . . .	44
4.3	Experimental Design and Validation Procedure . . . . .	45
4.3.1	Experimental Design . . . . .	46
4.3.2	Validation Procedure . . . . .	47

<b>5</b>	<b>Results</b>	49
5.1	Baseline . . . . .	49
5.2	Land Cover experiments . . . . .	54
5.3	Sentinel-3 OLCI experiments . . . . .	54
5.4	Aggregation by hotspot collection and model . . . . .	56
<b>6</b>	<b>Conclusions</b>	61
	<b>References</b>	65

# Chapter 1

## Introduction

Climate change is an important challenge facing mankind [1], which threatens the lives of many with dangerous side effects that can lead to catastrophic natural events [2].

A warmer climate increases the likelihood of fires at a wide range of latitudes [3]. These can potentially expand to destroy entire forests, therefore releasing significant amounts of smoke, carbon dioxide and heat in the process. A few instruments have been developed that can provide satellite data useful to detect thermal anomalies, such as the Moderate-resolution Imaging Spectroradiometer (MODIS), and the Visible Infrared Imaging Radiometer Suite (VIIRS). These radiometers acquire data in multiple spectral bands, or groups of wavelengths, and are mounted on satellites so that the whole surface of Earth is covered twice a day.

Developing accurate automatic detection systems based on similar instruments could prove to be vital for detecting expanding forest fires in their early stages, especially when they spark in remote areas that are scarcely populated, both in order to reduce damage to natural heritage and minimise the efforts required in the intervention. For this purpose, a few algorithms have been developed to convert raw measurements into hotspot detections. It is worth mentioning the work of *Giglio et al. (2003)* [4] that provided a good starting point for this and other projects in the same field [5][6] by implementing and improving algorithms to refine Modis and Viirs data, which is then made available for free to the public.

The adopted algorithm changes depending on the instrument, given that both the spatial resolution and the observed spectral bands differ, but the

overall approach can be summarised as follows: starting from raw measurements, the areas covered by clouds and water bodies are masked. Then, potential thermal anomalies are detected based on threshold levels, which are computed contextually as reference value using the surroundings. Subsequently, any areas that could lead to false alarms, like deserts, are excluded, and an estimation of the power released by each remaining hotspot is made.

When a hotspot is identified, it cannot be located more precisely than what the spatial resolution of the instrument that detected it allows for. For this reason, these algorithms divide the projection space in cells in a grid-like pattern, where the smallest subdivision has approximately the same size as the spatial resolution of the given instrument. These subdivisions are called “firepixels”. Each hotspot is therefore assigned to the firepixel where it was detected, and its position is defined by the coordinates of said firepixel.

This aspect, combined with the relatively low spatial resolution of MODIS and the long intervals of time between each pass of the instruments, suggests that there is room for improvement both in terms of i) quality and resolution of the hotspot data, and ii) frequency of the measurements over a restricted geographic area. This led to the creation of other projects that aimed to expand on how satellite data can be used in this field.

One example of this is the adaptation of the aforementioned algorithm to VIIRS data, to exploit the improved resolution. Two products have been developed that use different spectral bands of the instrument.

After a hotspot is identified, one further disambiguation is performed to establish whether it is caused by a wildfire or by another source. One possible approach to improve on this specific task could be to integrate other data sources, in order to have a more complete perspective of the area affected by a hotspot. Possible inclusions that have been considered for this purpose are i) type of land cover of an area [7], to better evaluate when a hotspot could be located near vegetation, and ii) other satellite imagery like that obtainable with Sentinel [8], to capture the effects of a fire at higher resolution.

The goal of this thesis is to explore the results of integrating said data sources with existing hotspot detection systems, and to evaluate performances of common machine learning algorithms for the application of hotspot disambiguation.

As for similar works, validation is based on preexisting datasets of previous fire events. The ones that were chosen for this project are obtained from the Copernicus Emergency Management Service (EMS) [9] and the European Forest Fire Information System (EFFIS)[10], which roughly cover the entirety

of the European area, that is the focus of this research.

The other parts of this document are organised as follows. In Chapter 2, each adopted data source is introduced and motivated, showing examples of works that also use them to solve related tasks. In Chapter 3, the way this data is collected and preprocessed is presented (Section 3.1), leading to the creation of datasets (Section 3.2). In Chapter 4, a formalization of the problem is made, and the proposed experiments are described in conjunction with the validation process (Section 4.3). In Chapter 5, the obtained results are discussed. In Chapter 6, the author draws some conclusions for the work done and proposes some improvements for future developments.



# Chapter 2

## Data Sources and Related works

This chapter focuses on describing what data sources have been considered for inclusion and their motivation, together with examples of works that employ them for related tasks. In Section 2.1, a commonly used algorithm to process MODIS data is presented. This approach was the starting point to develop similar algorithms that are now often employed to process data from other satellites, including VIIRS, as presented in Section 2.2. In Section 2.3, two sources of historical data for past fire events are described, in particular Copernicus EMS and EFFIS. In Section 2.4, a few Sentinel satellite imagery products are introduced, and their differences presented. Finally, Section 2.5 provides a brief explanation on how land cover can be used for related tasks.

### 2.1 MODIS

MODIS was one of the first spectroradiometers that were sent in orbit around the Earth to monitor changes in large-scale global dynamics. One is mounted on the Terra satellite, which launched in 1999, and one on the Aqua satellite, launched in 2002. Fire observations are done four times per day, twice from each satellite. The instruments record data in 36 spectral bands, each with its respective spatial resolution. 29 out of the 36 have an approximate spatial resolution of 1 km, including the ones that are used for hotspot detection.

Each raw measurement consists of an unprojected portion of the MODIS

orbital swath that spans around 5 minutes, called “granule”. A granule measures approximately  $2340 \times 2030$  km in the along-scan and along-track directions, respectively. A good visualization of a granule is displayed in Figure 2 at page 11 of the MODIS Collection 6 Active Fire Product User’s Guide [11].

Each individual 1-km pixel is processed according to the algorithm described in *Giglio et al. (2003)* [4], which uses brightness temperatures derived from the 3.9  $\mu\text{m}$  and 11  $\mu\text{m}$  channels of MODIS to detect thermal anomalies, and those from the 12  $\mu\text{m}$  channel for cloud masking. A brightness temperature, or radiance temperature, is the temperature at which a black body would need to be in order to emit the same total flux of electromagnetic energy of the observed object at a given frequency, and is measured in degrees Kelvin.

Starting from raw measurements, the process follows these steps:

1. any pixel missing valid data is removed from further consideration;
2. an internal cloud mask and a water mask obtained from MODIS geolocation products are used to exclude other obvious non-fire pixels;
3. for remaining potential fire pixels, neighbouring pixels are used to establish a background value, if a minimum of 8 of those are still valid and not excluded by previous steps;
4. if the previous step was successful, a series of contextual threshold tests are applied to verify if the considered brightness temperatures differ from the estimated background value;
5. further tests are applied to exclude false detections caused by sun-glint, desert boundaries, and water reflections from areas not included in the water mask;
6. any remaining hotspot not excluded after previous steps is considered a fire;
7. Fire Radiative Power (FRP) is approximated for each of these from brightness temperatures by using a relationship described in *Wooster, Zhukov, and Oertel (2003)* [12] and *Wooster et al. (2012)* [13]. It is a measure of the rate at which energy is released in the considered firepixel.

As for the validation process, high-resolution fire maps obtained from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER)

were used [14]. This process showed how detection probability varies depending on fire size and time in the day-night cycle. Smaller fires are harder to detect also because of the moderate spatial resolution of the instrument.

Further processing, projections, and aggregation can be done to obtain more products, starting from the output of this algorithm. One of these, called *MCD14ML* is used for this thesis, and it will be described in detail in a later section.

## 2.2 VIIRS

Starting from the algorithm described for MODIS in the previous section, an equivalent has been implemented for the VIIRS instrument [15], although with a slightly higher spatial resolution of 750 m/pixel. The product obtained from this will be denoted from now on with the “*VNP14*” nomenclature used by NASA to refer to it. Performance wise, this product resembles what can be done with data retrieved from MODIS.

The 375 m/pixel channels of the instrument have also been repurposed for active fire detection following a similar approach. The product obtained this way is denoted as “*VNP14IMG*”. However, since these channels saturate more easily, data has to be handled in a special way to accommodate the differences with respect to other commonly used channels. The improved spatial resolution also allows to have better responses for smaller fires than what can be achieved with MODIS products, although this version of the algorithm is still experimental and under development. This disparity is shown in Figure 2.1, which shows the total count of detections over a reference period of 5 years for each collection, from 2015 to 2019, restricted over the European macro-area used for this work, which will be characterised in detail in later sections.

Both products are adopted for this thesis to be used in conjunction with MODIS hotspots, since none of the three collections completely includes all the activations detected by the others. Therefore, by having all three, it could be possible to have a wider coverage.

## 2.3 Copernicus EMS and EFFIS

In the wide field of fire management, an important role is played by historical data about previous fires, as it allows to identify their consequences and side effects. In particular, in order to build a ground truth about which hotspots

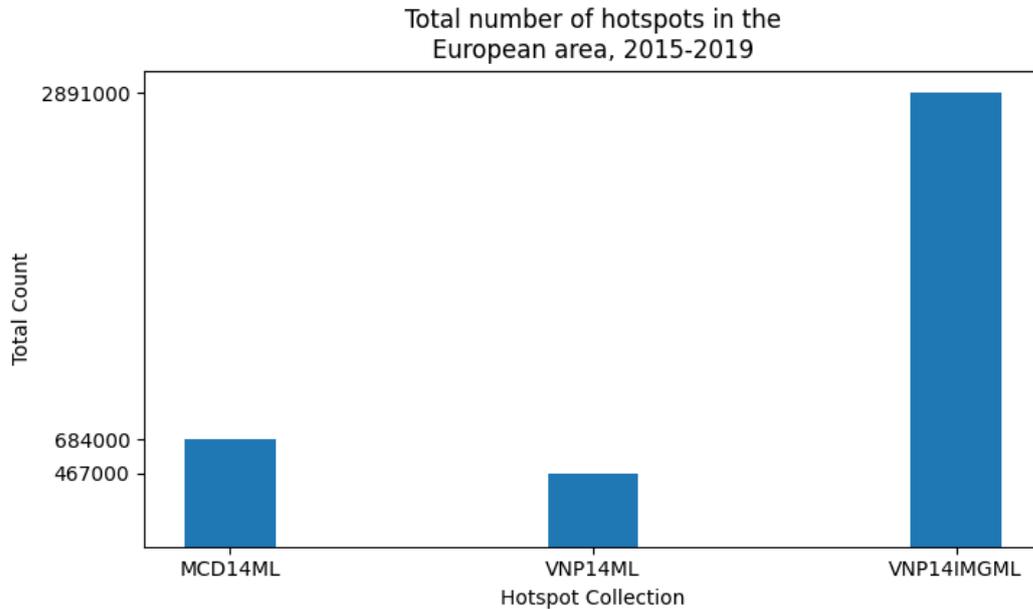


Figure 2.1. Total count of hotspot detections in the European area, over the years from 2015 to 2019, separately for the three considered products.

are related to a fire activation, two of the available data sources that cover Europe are Copernicus EMS and EFFIS.

The former is an on-demand service that provides detailed information about catastrophic events, primarily ones that take place in the European area. Albeit very detailed in terms of burned area delineation, the number of forest fire and wildfire activations is limited and the temporal information concerning the end of the event is often imprecise or even omitted.

The latter is a service open to all European countries based on cooperation. There are regulations in place which force Community members to monitor forests and collect a minimum set of data for each occurring fire. One of the products offered by this service is the burnt area database, which reportedly covers on average about 80% of the total burnt areas in the considered region, focusing on fires covering more than 30ha. This data source is characterised by a higher number of activations with respect to Copernicus EMS, although less detailed. Moreover, information about the end date of each event is not provided.

In *Nolde, Plank, Riedlinger (2020)* [16] it is shown how data from both of these services can be used in combination with MODIS data to solve the task

of burnt area monitoring. The authors also use the Sentinel-3 OLCI sensor that will be discussed in the next section.

Similarly, in *Farasin et al. (2020)* [17], Copernicus EMS is used for the task of damage severity estimation. It is worth noting that the authors used data from Sentinel-2, which will be described in the next section as well.

In this thesis both are used in conjunction with hotspot detections in order to decide which anomalies are linked to forest fires.

## 2.4 Sentinel imagery usage

A different perspective on the wider context of wildfires management can be obtained by using higher resolution satellite imagery, or simply different portions of the electromagnetic spectrum. For this purpose, one of the available options is using Sentinel missions [8], which offer a variety of products. Among these, the considered ones are the following:

- Sentinel-2, which provides images with a spatial resolution of up to 10 m/pixel, 20 m/pixel, or 60 m/pixel, depending on the used bands. Despite the fact that two satellites of this type orbit the Earth to increase the revisit frequency, the time between each pass is 5 days, making this product only viable for non-time-sensitive applications.
- Sentinel-3, which has a significantly lower resolution but shorter time intervals between each pass, especially at latitudes farther from the equator. For instance, in the European macro-area, the expected revisit time goes from 0.5 to 1.5 days, depending on the latitude. Among the two products obtainable from this mission, the one with a higher spatial resolution of 300 m/pixel was chosen.
- Sentinel-5P, which aims to monitor our atmosphere and its composition performing measurements at a high spatio-temporal resolution, with 14 passes each day.

Sentinel products were used in other related works [16][17], as previously discussed.

## 2.5 Land Cover

In order to better characterise an area affected by a hotspot, a valuable piece of information is what type of terrain it is detected in. Often, instruments used for hotspot detection have bands that can be used dynamically to detect greenery, like in the cases of MODIS, VIIRS, Sentinel-2, Sentinel-3.. Nonetheless, having information about other types of terrain can prove just as useful, for the purpose of disambiguation - especially for those that are seldom associated with wildfires but can produce hotspots, like densely populated urban areas.

In the context of the Copernicus project, the first detailed thematic map of Land Coverage was introduced in 1985. Namely, Corine Land Cover (CLC), which periodically receives updates to this day, the last of which happened in 2018. For that last update, high resolution imagery from Sentinel-2 was the primary data source, which was visually interpreted by the respective countries to produce the final result [7].

As shown in *Sifakis et al. (2004)* [18], land cover was used to observe forest fires in conjunction with AVHRR imagery, which has a lower resolution than what is available now via MODIS with current technology.

# Chapter 3

## Data Preparation

This chapter presents how all the discussed data sources can be retrieved (Section 3.1) and combined to form a dataset (Section 3.2), before being loaded and preprocessed (Section 3.3) to be fed to models as presented in the next chapter.

### 3.1 Collections and Preparation

#### 3.1.1 Hotspot products

The products in this category provide high-level descriptions of detected hotspots, and are used as a starting point for the task of disambiguation. The selected ones are reported in Table 3.1 together with their spatial resolutions.

Product Name	Spatial Resolution	Angular Approximation
MCD14ML	1 km	0.01°
VNP14ML	750 m	0.007°
VNP14IMGML	375 m	0.0035°

Table 3.1. Spatial resolution of the hotspot products that are used in this thesis, together with their approximate equivalent in terms of latitude-longitude.

They are freely distributed by an FTP server hosted by the University of Maryland, as described in the MODIS and VIIRS user manuals [11][15]. At

the time of writing, the chosen collections are available following the directory tree presented below:

```
data
├── MODIS
│   └── C6
│       └── MCD14ML
├── VIIRS
│   └── C1
│       ├── VNP14ML
│       └── VNP14IMGML
```

The chosen products are available in the leaf nodes of the tree above, in aggregated form. Hotspot detections are grouped in textual files in CSV format on a monthly basis, and then compressed in archives. In order to have coverage from all the data sources, only data covering the time frame from January 2012 was considered. At the time of retrieval, the latest month for which data had already been made available was May 2020. MCD14ML is the name of the collection obtained from MODIS data. VNP14ML follows this heritage for VIIRS, while VNP14IMGML is produced using higher resolution channels, as previously discussed in Chapter 2. Tables 3.2, 3.3, 3.4, shown below, present an overview of what attributes are used to characterise hotspots in the textual files of the respective collections.

MCD14ML	
Attribute Name	Description
YYYYMMDD	Date of detection (UTC)
HHMM	Time of detection (UTC)
sat	Satellite which performed the detection
lat	Latitude of the center of the firepixel where the hotspot was detected
lon	Longitude of the center of the firepixel where the hotspot was detected
T21	Brightness temperature of the 3.9 $\mu\text{m}$ channel
T31	Brightness temperature of the 11 $\mu\text{m}$ channel
sample	Number of sample within the considered firepixel. In case of overflow, it starts counting from 0 again
FRP	Fire Radiative Power estimation in MW
conf	Confidence about the detection. Integer number from 0 to 100
type	Class of hotspot, for the disambiguation task
dn	Single char to represent if the detection was done during the day or during the night

Table 3.2. Description of MCD14ML attributes.

VNP14ML	
Attribute Name	Description
YYYYMMDD	Date of detection (UTC)
HHMM	Time of detection (UTC)
sat	Satellite which performed the detection
lat	Latitude of the center of the firepixel where the hotspot was detected
lon	Longitude of the center of the firepixel where the hotspot was detected
T_M13	Brightness temperature of the 4 $\mu\text{m}$ channel
T_M15	Brightness temperature of the 11 $\mu\text{m}$ channel
sample	Number of sample within the considered firepixel. In case of overflow, it starts counting from 0 again
pixarea	Area of the firepixel in $\text{km}^2$
FRP	Fire Radiative Power estimation in MW
conf	Confidence about the detection. Integer number from 0 to 100
type	Class of hotspot, for the disambiguation task

Table 3.3. Description of VNP14ML attributes.

VNP14IMGML	
Attribute Name	Description
YYYYMMDD	Date of detection (UTC)
HHMM	Time of detection (UTC)
sat	Satellite which performed the detection
lat	Latitude of the center of the firepixel where the hotspot was detected
lon	Longitude of the center of the firepixel where the hotspot was detected
T_I4	Brightness temperature of the I4 channel
T_I5	Brightness temperature of the I5 channel
sample	Number of sample within the considered firepixel. In case of overflow, it starts counting from 0 again
pixarea	Area of the firepixel in $km^2$
FRP	Fire Radiative Power estimation in MW introduced empyrical relationship
conf	Confidence about the detection. Categorical attribute, with values: low, nominal, high
type	Class of hotspot, for the disambiguation task

Table 3.4. Description of VNP14IMGML attributes.

Despite some differences, the three collections have several attributes in common. Among these, the one denoted as "*type*" is the class assigned by the current algorithms for the task of disambiguation. It is a numeric integer value from 0 to 3, extremes included. The meaning of each value is reported in Table 3.5 below. As will be discussed in detail in later chapters, this label is not always reliable, and it is in the scope of this thesis to improve this classification.

Concerning other attributes, the following operations are performed:

- information about date and time of detection are combined to form a timestamp
- latitude and longitude of the center of the firepixel are combined to form a so called "*POINT*", which is one of the proposed datatypes of the Well-Known Text representation of geometry objects (WKT), commonly used by GIS software.

Attribute Value	Description
0	Presumed vegetation fire
1	Active volcano
2	Other static land source
3	Offshore

Table 3.5. Explanation of the possible values for the *type* categorical attribute, shared among all hotspot data sources.

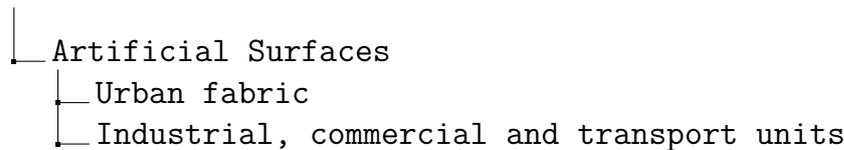
- any other field is left untouched for the time being

All of this information is then inserted in a spatial-enabled database, which is a database with GIS extensions to facilitate geo-spatial operations. Refer to Section 4.2 for what tools have been used. Having hotspot data in such a database allowed to intersect it and combine it with other types of geo-spatial data, like burnt areas from past fires.

### 3.1.2 Land Cover

As already discussed in previous chapters, the selected product for Land Cover is CLC, with its higher resolution version of 2018. It is a mosaic of multipolygons, each of which covers a certain region of space and has an associated type of land cover. "*MULTIPOLYGON*" is another datatype of the already mentioned WKT representation, which identifies a group of polygons with a finite number of intersections between each other. The categories that are used for this classification are organised in a hierarchical structure on 3 levels, that describe the type of land cover to increasing levels of detail. The exact structure can be seen in the documentation of CLC [19]. Only two of the three levels of depth will be used, so that more general categories are obtained. The resulting tree structure is displayed below. Leaf nodes are the categories that were adopted.

By adopting this aggregation, many of the small yearly changes in type of land cover can be accounted for. For this reason, and considering the higher resolution of this release of the product, the 2018 version was used as a time-independent indication of the type of terrain.



---

└─	Mine, dump and construction sites
└─	Artificial, non-agricultural vegetate areas
└─	Agricultural areas
└─	└─ Arable land
└─	└─ Permanent crops
└─	└─ Pastures
└─	└─ Heterogeneous agricultural areas
└─	Forest and seminatural areas
└─	└─ Forest
└─	└─ Shrub and/or herbaceous vegetation associations
└─	└─ Open spaces with little or no vegetation
└─	Wetlands
└─	└─ Inland wetlands
└─	└─ Coastal wetlands
└─	Water bodies
└─	└─ Inland waters
└─	└─ Marine waters
└─	Missing data
└─	└─ NODATA

The mosaic only covers all pan-European countries, therefore creating a wide Area of Interest (AoI). An AoI can be described as a squared region that envelopes a geometric element.

As an initial approach, the whole AoI was considered, except for the Northernmost portion of the African continent, due to the higher likelihood of false hotspot detections near desert boundaries. This restriction was particularly effective for the VNP14IMGML collection, which is more sensitive to smaller anomalies but is more prone to higher false detection rates. It allowed to filter out around 97% of the total available hotspots of the collection for the considered AoI, going from around 170M to about 5M. For the other hotspot products, MCD14ML went from 1.6M to 1.1M, and VNP14ML from 900k to 750k.

A delineation of the boundaries used for this process is shown in Figure 3.1. A further refinement of the considered bounding-box will be suggested in Chapter 6, together with other proposed improvements.

The product is made available for free directly on the CLC 2018 website, and is distributed in three forms which are: i) an RGB raster of the whole mosaic, ii) an ESRI geodatabase of the multipolygons, or iii) an SQLite

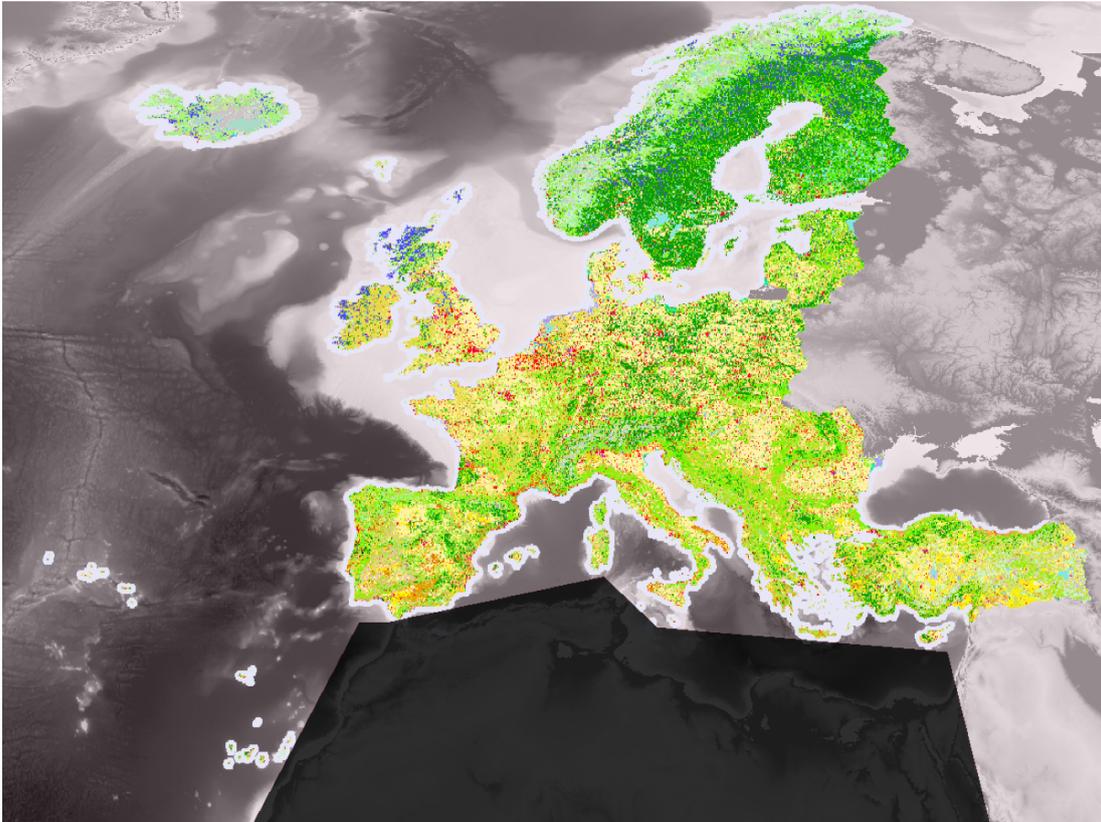


Figure 3.1. Considered European macro-area. The excluded portion of the African continent is shown in dark gray. Coloured areas are regions where CLC data is available. The remaining areas with no such information are represented in light gray.

database. The adopted form is the ESRI geodatabase, as this allows better flexibility in terms of encoding of the features, with respect to an already prepared raster. Multipolygons in this product are originally stored using the ESRI:3035 projection. Therefore, in order to be able to compare these geo-spatial features with hotspot data, they were reprojected to the same system, WGS84 EPSG:4326, based on latitude and longitude like hotspot data. All the multipolygons of the mosaic were then inserted in the geo-spatial database.

At a later stage, during the creation of the dataset as described in Section 3.2, these geometric objects will be rasterized into images, to characterise the surroundings of a hotspot. Images of this type have 16 channels, one for each used land cover category, and are generated at a low resolution of 16x16 to

have comparable size with those obtainable from Sentinel-3 imagery. Each pixel of a channel has a value of 1 if the corresponding area is covered by the category of terrain represented by the channel, and 0 otherwise.

Even though having one channel for each category allows for a clearer distinction, it can increase dimensionality. Therefore, by adopting fewer more general categories, it is possible to fight against the curse of dimensionality problem, while also improving on generalization.

### 3.1.3 Sentinel-3 OLCI

The selected product for Sentinel-3 data is based on the Ocean and Land Colour Instrument (OLCI), which has a spatial resolution of 300-m, higher than other options.

Data can be freely retrieved from a service named Sentinel-hub [20], which provides a wide range of utilities and helper functions for the satellite image retrieval process. Data is available starting from the 26th of April 2016, and it comes in the form of numpy arrays or geo-tiff images.

The obtained images have 21 channels, one for each band detected by the instrument. They are reported in Table 3.6 together with common use cases for each one.

For each request, it is mandatory to specify a few parameters that define the image, such as:

- **bbox**: bounding box of the requested area, which also includes information about the used reference system. In other words, it is a rectangular region in the specified projected space.
- **size**: of the image, in pixels. It can be determined if both the bounding box and the spatial resolution of the instrument are defined. In case a resolution different from the one of the instrument is chosen, data will be interpolated to accommodate the requested size.
- **time\_interval**: it is a time frame, used to restrict the selection of images to ones that were taken within this interval. It is expressed as the interval comprised between a start date and an end date, both included.

The service imposes some restrictions on the number and frequency of requests that can be made to it. For this reason, all the requests were made

Band	$\lambda$ of the center of the band (nm)	Example use cases
Oa01	400	Aerosol correction
Oa02	412.5	Yellow substance and detrital
Oa03	442.5	Chlorophyll absorption maximum, biogeochemistry, vegetation
Oa04	490	High Chlorophyll
Oa05	510	Chlorophyll
Oa06	560	Chlorophyll reference (Chlorophyll minimum)
Oa07	620	Sediment loading
Oa08	665	Chlorophyll (2nd Chlorophyll absorption maximum), sediment, yellow substance/vegetation
Oa09	673.75	Improved fluorescence retrieval
Oa10	681.25	Chlorophyll fluorescence peak
Oa11	708.75	Chlorophyll fluorescence baseline
Oa12	753.75	O <sub>2</sub> absorption/clouds, vegetation
Oa13	761.25	O <sub>2</sub> absorption band/aerosol correction
Oa14	764.375	Atmospheric correction
Oa15	767.6	Fluorescence over land
Oa16	778.75	Atmos. corr./aerosol corr.
Oa17	865	Atmos. corr./aerosol corr., clouds, pixel co-registration
Oa18	885	Water vapour absorption reference band, Vegetation monitoring
Oa19	900	Water vapour absorption reference band, Vegetation monitoring (maximum reflectance)
Oa20	940	Water vapour absorption, Atmos. corr./aerosol corr.
Oa21	1020	Atmos. corr./aerosol corr.

Table 3.6. Description of the bands available in Sentinel-3 OLCI images, together with some reported example use cases.

after the selection of what data points will be included in a dataset. This process will be described in the following section, 3.2. The adopted approach for collection was to request an image for each selected hotspot, centered around the center of its firepixel, restricted to the date of its detection, using a squared bounding box in the latitude-longitude projected space with side equal to double the spatial resolution of its detecting instrument. Resolution is expressed in degrees. In case no data is available for the selected date, the time interval is expanded to include the previous day and the one after the detection date.

However, this approach doesn't consider that distances behave differently in a projected space. In other words, the equivalent in meters of  $1^\circ$  of longitude is not uniform at all latitudes. Therefore, despite requesting a squared bounding box defined using latitude and longitude, images are not squared in terms of pixels and their sizes vary depending on the latitude. A few examples of this issue are displayed in Figure 3.2.

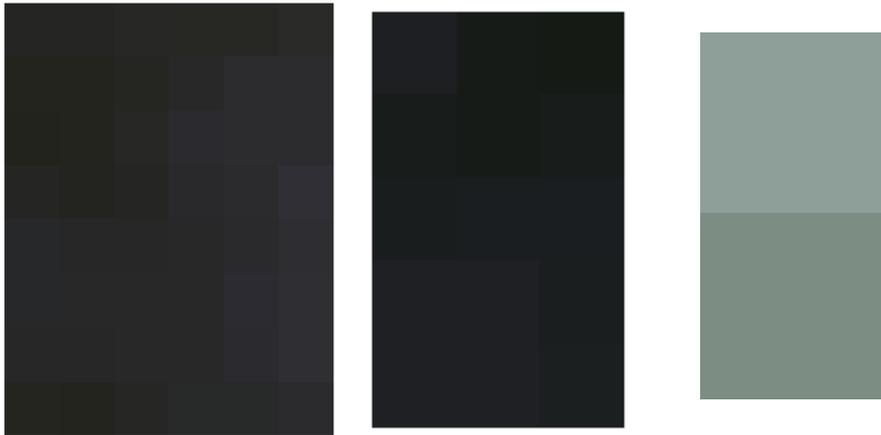


Figure 3.2. From left to right: a 6x8 pixels image obtained for a MCD14ML hotspot at latitude  $39.033^\circ$ ; a 3x5 pixels image for a VNP14ML hotspot at latitude  $57.074^\circ$ ; a 1x2 pixels image for a VNP14IMGML hotspot at latitude  $60.888^\circ$ . The displayed images use bands Oa06, Oa04, Oa08 as RGB channels for visualization purposes only.

Due to the aforementioned restrictions, it was not possible to retrieve new copies of the images with an adjusted procedure within the time frame for this thesis. However, after it was verified that none of the selected hotspots would be so close to the poles to cause degenerate images with width equal to 0 pixels, it was decided that images would be uniformed in size following the procedure described in 3.3 and used anyway for some experiments, since they

contain at least some information. The adjusted method will be discussed, together with other proposed improvements, in Chapter 6.

### 3.1.4 EFFIS burnt areas

The EFFIS burnt areas product reportedly covers around 80% of the total burnt area in the covered regions. The remaining 20% is composed of fires that burn less than 30ha, which are not considered. The product was obtained via a data request form that can be accessed directly from the EFFIS website [10]. The collection comes packaged as a compressed geo-json file, in which the geometry of the burnt areas is provided in WKT format, as well as other attributes as presented in Table 3.7. All geometries in this product share the same WKT representation, in the *MULTIPOLYGON* format. The coordinate reference system adopted by the creators is once again ESRI:3035. Therefore all points were converted to the same reference system of the hotspots, WGS84 EPSG:4326, based on latitude and longitude, before being inserted in the geo-spatial database.

At the time of retrieval, this product included a total of 14914 activations, which locations can be visualised as displayed in Figure 3.3

Attribute Name	Description
id	Internal ID of the activation
FIREDATE	Date of the activation (UTC)
COUNTRY	2-letter abbreviation of country names
placename	Integer number to identify areas
PROVINCE	Name of the affected location was detected
YEARSEASON	Legacy attribute, containing info about year and season of the fire
geometry	WKT representation of the burnt area

Table 3.7. Description of EFFIS burnt areas attributes.

It should be noted how the only temporal information provided for each fire is one date, without any further specifications on whether it represents the starting date or the ending date of the activation, or anything about the duration of the fire. However, since this work requires to be able to intersect hotspots and past fire events both spatially and temporally, this attribute plays an important role. For this reason, an extra analysis has

been performed to better evaluate the ending date of each activation, and it will be discussed in [3.2.1](#).



Figure 3.3. Location of EFFIS activations. Each coloured point represents the center of a burned area.

### 3.1.5 Copernicus EMS burnt areas

Some data about previous forest fires and wildfires can be freely downloaded from the Copernicus EMS rapid mapping service [9]. For each activation, some information is provided about location, affected countries, and starting date of the fire, optionally seldom including an indication of the time of day. Since little to no indication about the end dates is provided, data from this source went through an extra analysis step similarly to what was mentioned for EFFIS in a previous subsection. The details of this process will be discussed in [3.2.1](#).

The service offers four types of products:

- **reference** products establish the situation before the event;
- **first estimates** roughly identify and assess the most affected locations;
- **delineation** products assess the extent of the event;

- **gradings** evaluate the intensity and scope of the damage.

Both delineation and grading types happen after the event, with gradings being more definitive. If data is only of medium quality, multiple releases are produced. Moreover, more versions are published if errors are found. However, since products are created on-demand, not every activation contains the same types of products, with some activations not including any.

For the purpose of this thesis, reference products have been excluded, as they don't bring direct information about a fire event on their own. For each activation, grading products were given the highest priority, followed by delineation, and finally first estimates.

Products are distributed in the form of compressed archives, each containing shapefiles about a variety of related aspects to the fire, like the AOI for the event, or information about the local hydrography.

Unfortunately, there is no uniformity in terms of what aspects are covered by a product. Some of these even lack the delineation of the burnt area, only providing the AOI. Any activation for which neither the delineation or the AOI were provided was excluded from further consideration. For the remaining ones, the most detailed available shapefile was considered, and every *POLYGON* and *MULTIPOLYGON* therein contained was considered as a portion of the burnt area for the activation. They were then inserted in the geo-spatial database, where they were subsequently combined together in order to form a unique geometric representation of each activation, similarly to how EFFIS data is provided, instead of a fragmented one.

This filtering process left only around 80 activations out of the 117 that were reported at the time of the data retrieval. Their locations can be visualised as shown in Figure 3.4.

## 3.2 Dataset Creation

### 3.2.1 Ground Truth

In order to build a ground truth for the hotspot disambiguation task, which establishes whether a hotspot is connected to a fire event or not, the adopted approach was to perform an intersection between detected hotspots and available previous fire events. More in detail, both of the following conditions need to be met in order to have a match between the two:



Figure 3.4. Location of Copernicus EMS fire activations. Each point represents the center of a burned area.

- Temporal intersection: the considered hotspot has to be detected at a time comprised between the starting date and the ending date of the given fire event.
- Spatial intersection: the center of the hotspot firepixel has to be detected in the area affected by the given fire event, barring the spatial resolution of the hotspot.

If both conditions are met, the considered hotspot is assigned to the positive class. Otherwise, it will be considered as part of the negative one. This technique has at least two critical points that need to be considered.

Firstly, it is required that most of the fire events in the considered time interval are detected, if not all. This implies that hotspots that cannot be linked to a fire event can only be in the negative class. Unfortunately, the available data sources cover for only about 80% of the burnt areas, mostly caused by bigger fires, since tracking damage of smaller ones is a harder task for satellite data. This also means that some noise is inevitable with the provided context, as hotspots related to those fires causing the remaining 20% will be assigned a negative label, therefore being false negatives. However,

it should be noted that there is no information about what is the percentage of hotspots that is linkable to those fires.

Secondly, if criteria for intersection are not correctly defined, those hotspots that fall near the decision boundary could be classified with the wrong label, adding to the noise problem that was mentioned.

In terms of spatial comparisons, this issue is accounted for by including a buffer region surrounding the burnt area of a fire, with minimum size comparable to the uncertainty of measurements. This way, if a hotspot is detected in a firepixel that is centered outside of a burnt area, but near to it and in the right time frame, it will be included in the positive class.

Concerning the temporal aspect, the duration of fire events from the adopted products is not clearly defined. As already discussed in previous sections, both Copernicus EMS and EFFIS report only a single date for any fire event. Considering that big fire events such as those covered by these products can last more than a single day, each hotspot detected after the reported date would be assigned to the negative class and act as noise.

For this reason, one further analysis was performed for both of the EFFIS and Copernicus EMS burnt areas products, with the goal of inferring a likely end date for each activation. The chosen strategy can be described as follows.

For each activation, starting from the reported date of the fire, the number of hotspots intersecting the burnt area in the given date is counted, including all three hotspot collections. If the total for that date is higher than or equal to a minimum threshold level, then the considered day is included as part of the fire event. The process stops when it finds a date with a count lower than the threshold, which is meant to be seen as a hyperparameter.

In order to find a reasonable value, a few alternatives have been considered: {1, 2, 3, 5, 10}. In Figure 3.5 it is shown how using different values as threshold influences the inferred fire duration for EFFIS activations.

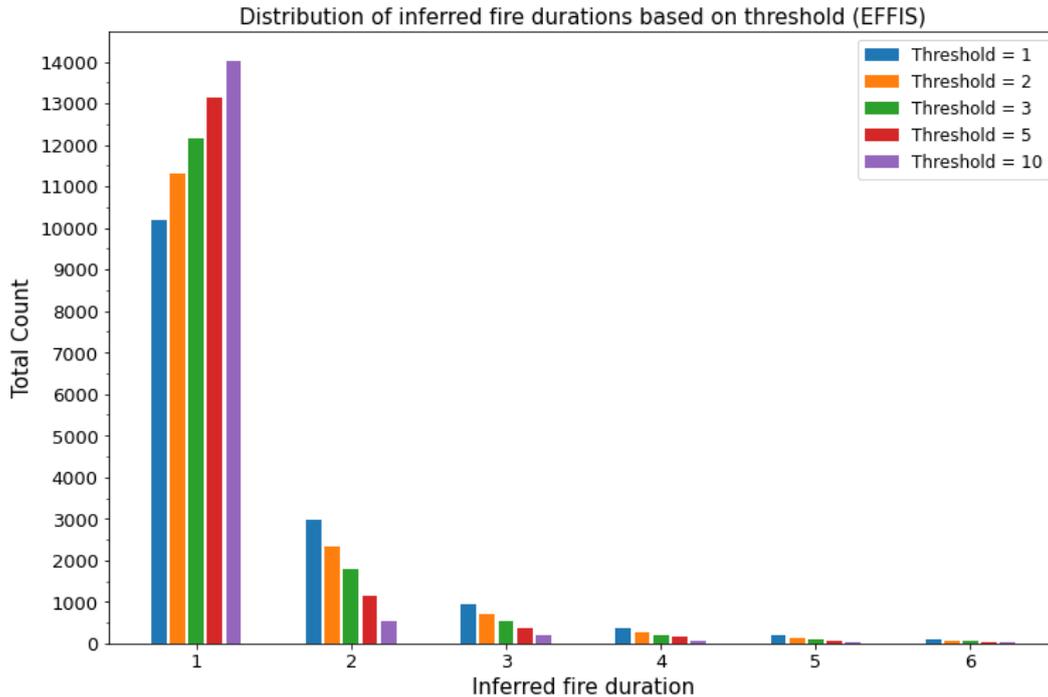


Figure 3.5. Effect of threshold (number of hotspots) on the inferred fire duration, focusing on the more numerous EFFIS activations. A higher threshold leads to reduced inferred durations.

Even by employing this technique, the majority of the fire activations has an inferred duration of one day. Nonetheless, it can be seen that low thresholds allow to reveal the longer duration of a significant portion of the recorded fire events. The adopted threshold was 2, as a trade-off between sensitivity and resilience to false detections.

Following this step, the intersection between hotspots and burnt area products was made once again in order to assign each hotspot to a class, therefore obtaining what will be used as ground truth in later sections.

The location of all positive hotspots is shown in Figure 3.6. In order to adjust for the sheer number of detections, they are visualised in an aggregated form via a hex-bin plot, color-coding each region based on the number of hotspots it contains.

While the same approach can be used for the negative class, it is necessary to perform a random sampling for visualization purposes due to the significantly larger number of entries. The results with a reduced sample size of 1/10 are shown in Figure 3.7. The disparity between the two classes will be

addressed in Subsection 3.2.2.

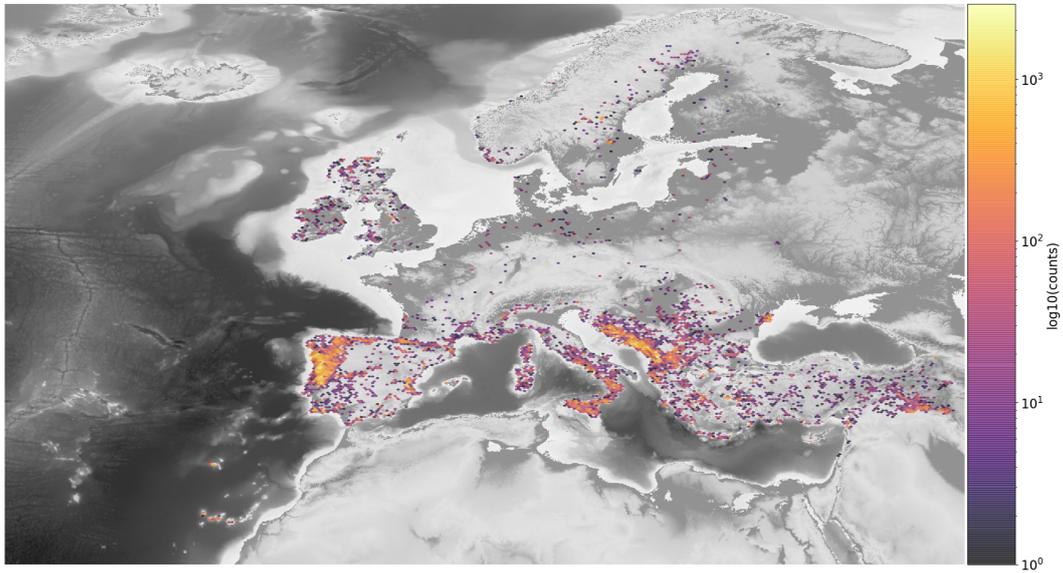


Figure 3.6. Hexbin plot of positive hotspot locations. Each bin is color-coded to represent the number of contained hotspots, with a logarithmic scale.

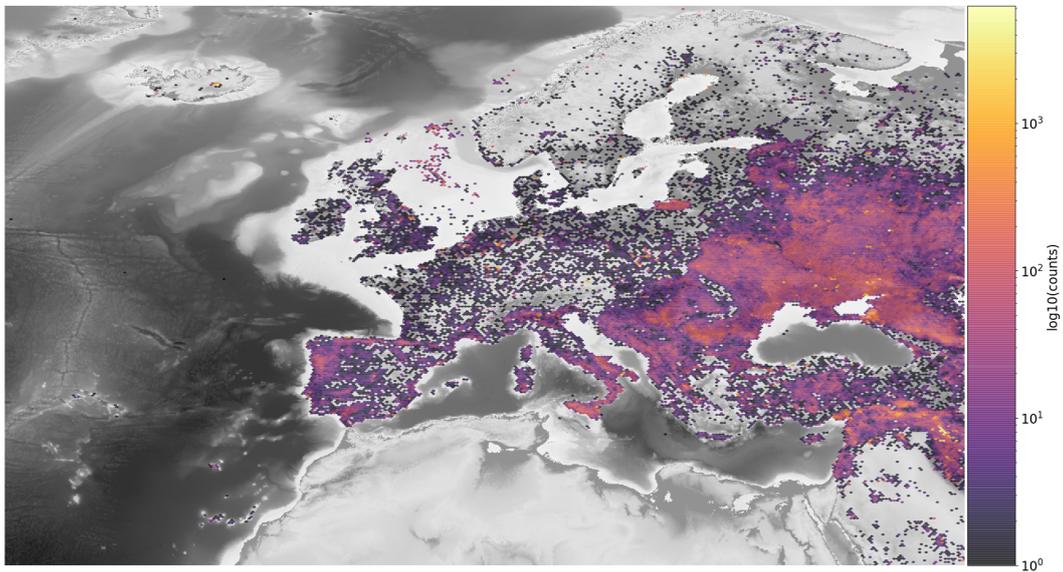


Figure 3.7. Hexbin plot of negative hotspot locations, sampled to a 1/10 of the available amount. Each bin is color-coded to represent the number of contained hotspots, with a logarithmic scale.

### 3.2.2 Dataset Balancing

Data collected in the way described in previous sections leads to a severely unbalanced distribution between positive and negative class, as shown in Figure 3.8. In order to prevent a situation where a model can classify each sample as negative and score over 99% accuracy, a balancing operation has been performed to get closer to a 50%-50% distribution.

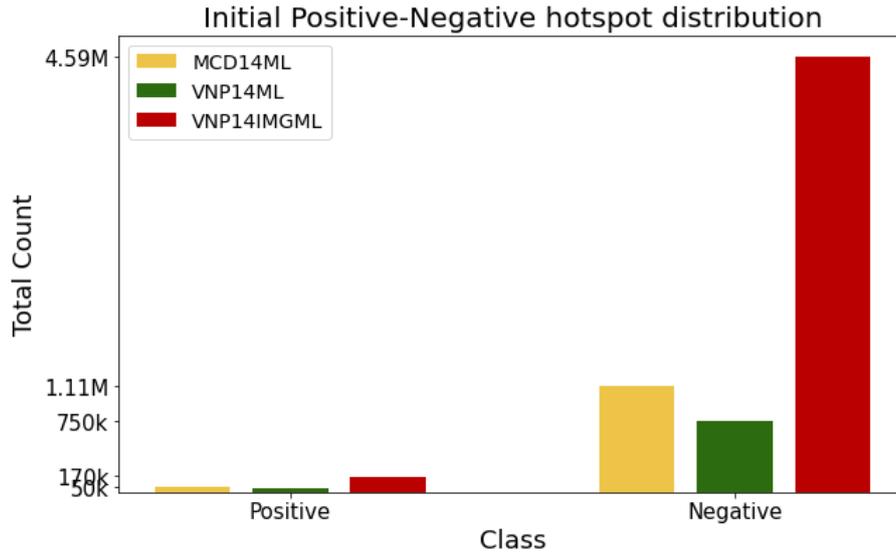


Figure 3.8. Initial balancing of the positive and negative hotspot classes. The negative class is over-represented for every hotspot collection.

In particular, a sampling of the negative class was performed to reduce its size to be comparable with the positive one. While doing so, a secondary goal that was kept in consideration was to maintain a similar geographic distribution of hotspots in the negative class. To achieve this objective, a stratification based on latitude and longitude coordinates of detection was done.

Since they are continuous attributes, it was required to first perform a discretization of these variables. The procedure can be described with the following steps.

Starting from the bounding box of the considered pan-European area, projected in the WGS:84 space, its width (longitude) is subdivided in 2684 splits, so that each one measures approximately  $0.03^\circ$ . The number of height subdivisions (latitude) is computed by using  $0.03^\circ$  as measure of each split, rounding up, obtaining a grid with squared cells in the projected space. Since the number of intervals along the latitude direction was rounded up, the size of this grid is greater than the size of the bounding box. Therefore, an offset is applied in the latitude direction to adjust for the excessive height. Finally, each hotspot is assigned to the cell of the grid that contains its location.

A visual representation of how the created stratification grid compares to the starting bounding box is shown in Figure 3.9.

The value of  $0.03^\circ$  was chosen to be comparable with the minimum approximate spatial resolution out of the three hotspot collections, which is  $0.01^\circ$  for the MCD14ML product obtained from MODIS, while being slightly greater in order to reduce the number of cells in the grid to be lower than the number of hotspots. This approach allows to transition to coarser aggregations, if required by future studies.

The selection process is performed separately for each hotspot product, so that a dataset built after this selection can be balanced on all three collections and can consider each of them independently. This allows to adjust for the different quantities of data of the three collections, as previously discussed.

The steps are as follows. For each cell containing positive hotspots, include all of them. After that, create a ranking of the negative hotspots in the same cell, based on their temporal distance from any of the positives. Then, pick as many negative hotspots as the positive ones, prioritising those detected not close to positives.

For other cells that only contain negative hotspots, if their number of hotspots is below a threshold, exclude them from further consideration. Otherwise, perform a random sampling without repetition, to reduce sample size to a factor of  $1/threshold$ .

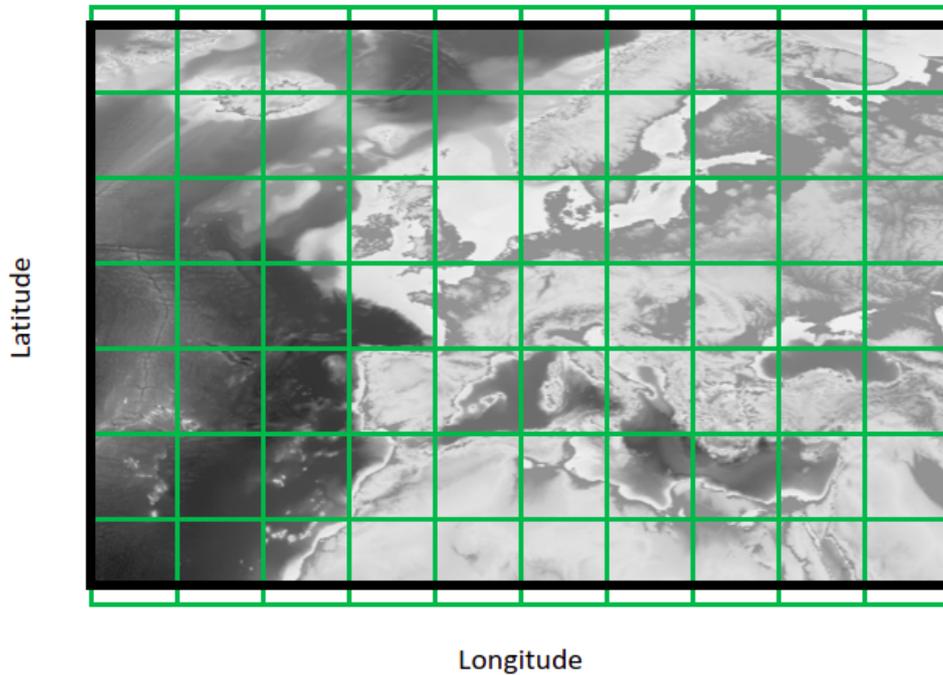


Figure 3.9. Representation of the stratification grid. The bounding box of Europe is shown in black, while the stratification grid is in green.

The ranking described above aims to exclude those hotspots that were left out of the positive class during the inferral of end dates for fire events: by prioritizing those that are detected farther from the decision boundary, there is a higher chance they are true negatives. This approach would allow to characterise an area both during a fire event and farther from it. However, if only this step was adopted, not enough negative hotspots would be available to balance the positive class, as displayed in Figure 3.10.

Therefore, with the adjustments adopted for other cells, it is possible to include extra negative hotspots, while adding a threshold to exclude areas with false detections, and reducing their count.

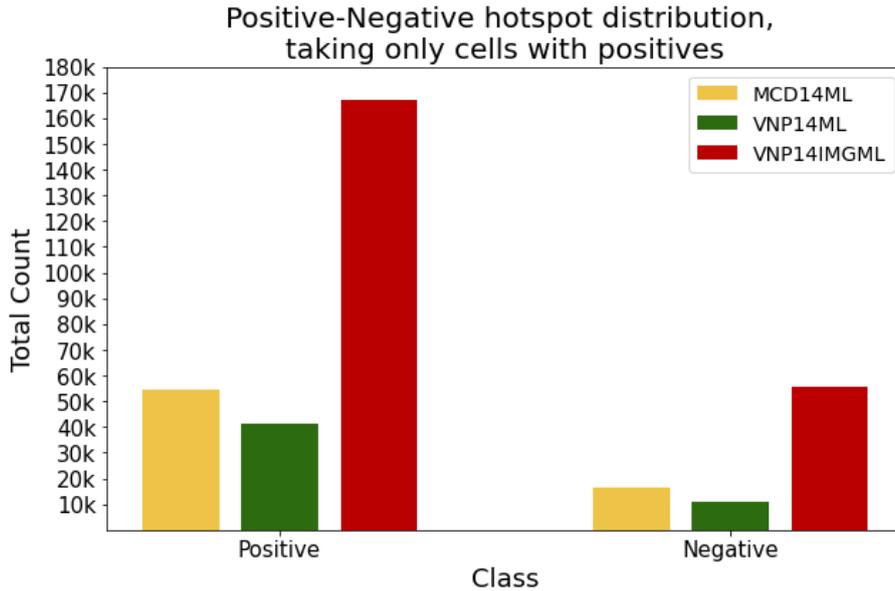


Figure 3.10. Resulting balancing of only taking cells which contain at least one positive hotspot.

As will be described in Subsection 3.2.3, two versions of the dataset were created, denoted as " $v2$ " and " $v3$ ". Each version has its specific threshold values, because different groups of positive hotspots are considered during stratification, leading to a different selection. The values used for each are reported in Table 3.8. They were chosen empirically with the two goals of i) having the two classes represented as close as possible to a 50%-50% split for each hotspot collection, and ii) not having a selection with more positive hotspots than negative ones, since this would hint an underlying data distribution that doesn't match the whole collection. Considering that these threshold levels are integer numbers, the chosen values are the closest ones to achieve these objectives.

The results of the balancing process are reported in Figure 3.11 and Figure 3.12 respectively.

Dataset version	MCD14ML	VNP14ML	VNP14IMGML
$v2$	9	7	20
$v3$	6	5	11

Table 3.8. Threshold levels used to evaluate cells in the stratification grid.

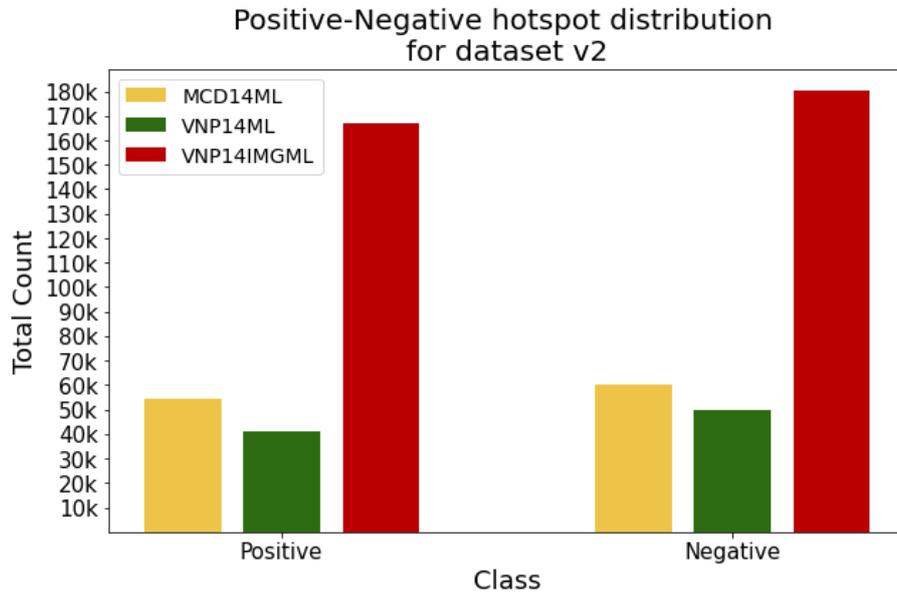


Figure 3.11. Resulting balancing of the  $v2$  version of the dataset.

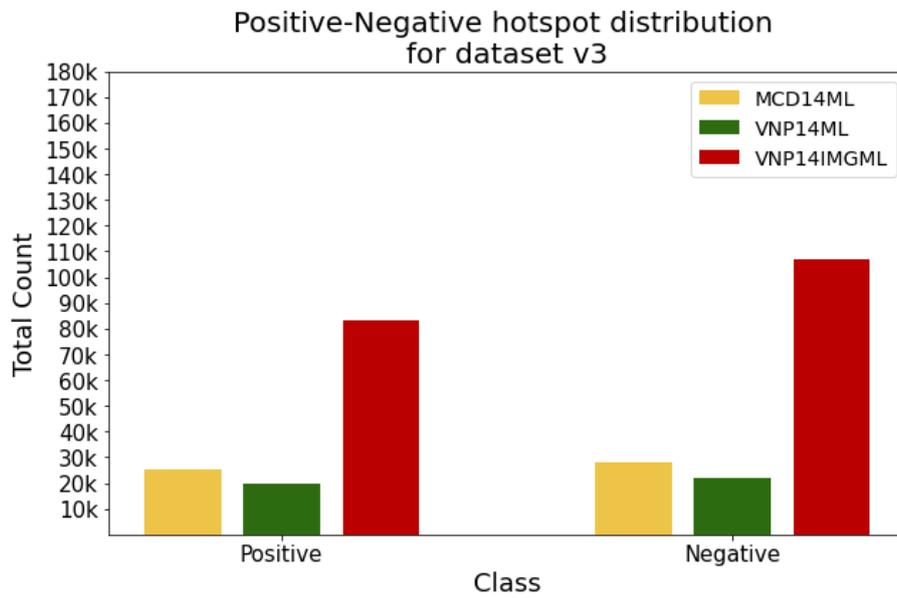


Figure 3.12. Resulting balancing of the  $v3$  version of the dataset.

### 3.2.3 Incremental Data Integration

In order to assess the effects spawned from integrating land cover and Sentinel imagery together with hotspot data, an incremental approach was adopted. Considering the time availability of each collected data source, reported in Table 3.9, it can be seen that there is no data for Sentinel-3 before April 2016.

Data Source	Time Coverage
MCD14ML	2012 - May 2020
VNP14ML	2012 - May 2020
VNP14IMGML	2012 - May 2020
CLC2018	Considered time-invariant
Sentinel-3 OLCI	April 2016 - Present

Table 3.9. Time availability of the adopted data sources.

For this reason, two versions of the dataset were created, named "*v2*" and "*v3*" after one first failed attempt. Each version underwent a separate process of stratification, as presented in previous sections, due to the differing selection of available hotspots.

*v2* covers from year 2012 to May 2020, and it includes hotspot data and land cover, but no data from Sentinel sources. The reasoning behind this version is to have a wider sample size in terms of hotspot data, while still being able to assess the impact of land cover for the disambiguation task.

*v3* covers from year 2017 to May 2020, in order to facilitate future integration with other Sentinel products that may not be available before. It includes all sources that were used for *v2*, while adding Sentinel-3 OLCI imagery. Images were retrieved after the stratification process for hotspot data had taken place, as a way to accommodate for the restrictions on requests for Sentinel imagery, as already presented in Subsection 3.1.3.

Similarly, images for land cover were rasterized at this point in time, starting from the geometry objects obtained in Subsection 3.1.2, to represent the surroundings of the area impacted by a hotspot.

The important differences between the two versions are described in Table 3.10.

	<i>v2</i>	<i>v3</i>
Total number of hotspots	553157	284636
MCD14ML positive-negative count	54415-60444	25078-28197
VNP14ML positive-negative count	41313-49708	19629-21762
VNP14IMGML positive-negative count	167074-180203	82879-106155
Overall class balancing	47.5%-52.5%	44.9%-55.1%
Time coverage	2012 - May 2020	2017 - May 2020
Includes land cover	yes	yes
Includes Sentinel-3 OLCI	no	yes

Table 3.10. Main differences between the two versions of the dataset.

### 3.3 Further Preprocessing and Encodings

As a last step before being used for experiments, data is loaded from the dataset and encoded. Each data source is first addressed separately, before being combined with the rest and normalized to zero mean and unit variance.

Although the three hotspot collections can be treated separately, one of the considered approaches is to combine them in one, exploiting the fact that they share the majority of attributes. In particular, since each hotspot collection uses brightness temperatures based on different spectral bands, these represent physical measurements with different meanings and they cannot be considered as shared. Therefore, a conventional value of 0 was used as a way to represent missing brightness temperature values. This is visualized in Figure 3.13. The next paragraphs contain a description of the preprocessing performed for each source.

#### Hotspot Data

All three hotspot collections go through the same processing, which can be summarised as follows:

- Any attribute that was used as an identifier is removed, namely *id* and *sample*;
- Latitude and longitude attributes are removed;
- Any remaining categorical attribute is one-hot encoded;
- Information about the year of detection is removed;

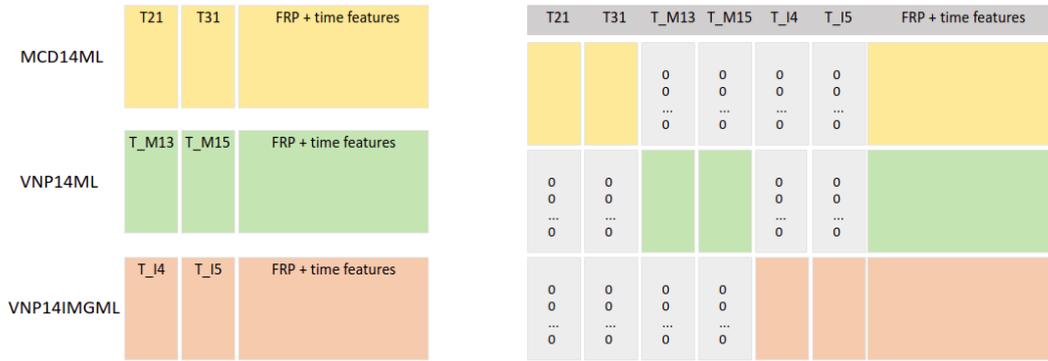


Figure 3.13. Adopted approach to combine all hotspot collections for a unified model. MCD14ML features are represented in yellow, VNP14ML is in green, VNP14IMGML is in orange. On the left: attributes of the three hotspot collections, with FRP and time features in common. On the right: 0-padding is added to represent missing data, and the three collections are combined.

- Remaining information about date and time of detection is encoded with a periodical sinusoidal representation;
- Other numerical attributes are left untouched.

The year of detection was removed in order to facilitate the learning of year-invariant features and to avoid potential shortcuts. Similarly, latitude and longitude were removed with a similar concern in mind, considering how most available fire activations are clustered in the Mediterranean area.

Regarding the sinusoidal representation, it is used as a way to express periodicity: for each point in time, its progression over a period can be represented with a polar parametrization to encapsulate the cyclic nature of the attribute, in the same way as sine and cosine functions can be used to draw circles.

The considered periodicities are: i) hour progression over 24-hour cycle, ii) day progression over 366-day cycle, iii) week number over 53-week cycle. For each, 2 features are extracted starting from the detection timestamp, for a total of 6 features.

Figure 3.14 shows how this encoding can be applied for the week number, as an example.

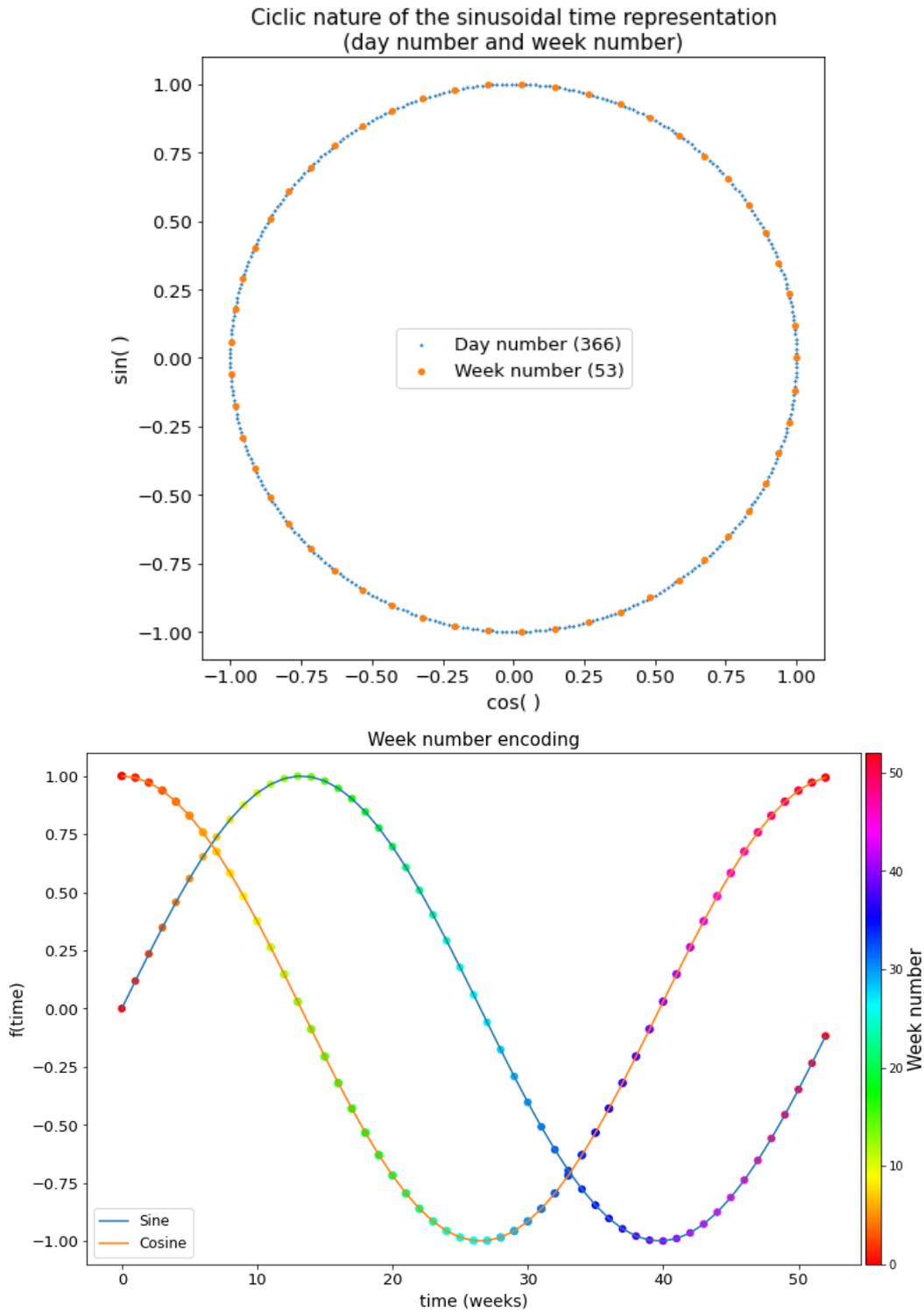


Figure 3.14. Encoding of time attributes in a sinusoidal representation. Above: the cyclic nature of this encoding can be visualized as a circle. Below: two features are obtained starting from the week number.

### **Sentinel-3 OLCI**

Sentinel-3 OLCI images obtained during collection (Section 3.1) are characterised by varying image size, going from 1 to 7 pixels on the horizontal direction, and from 2 to 8 pixels on the vertical one. The number of channels is fixed and is equal to 21.

In order to obtain a consistent number of features from these images, while maintaining a low dimensionality, the used approach was to average all the pixels in each band, obtaining a total of 21 features.

### **Land Cover**

Images obtained in previous sections have a resolution of 16x16, with 16 bands - one for each used land cover category. Each channel was then down-sampled to 4x4x16, to further reduce dimensionality, in analogy to what was performed for Sentinel imagery.

Finally they were flattened to a one-dimensional array.



# Chapter 4

## Methodology

This chapter focuses on formalizing the addressed problem, as well as defining a proposed solution and the adopted methodology. In particular: Section 4.1 describes the used models; Section 4.2 details what external frameworks and tools were used in the context of this thesis; Section 4.3 presents the performed experiments and the validation process.

The focus of this work is to disambiguate hotspot detections that are related to wildfires from ones that are not. More formally, given a hotspot detection, characterised with:

- timestamp of detection;
- brightness temperature measurements obtained from two spectral bands;
- the estimation of its Fire Radiative Power;
- (optionally) a raster of the surrounding land cover (an image with 16 channels);
- (optionally) a Sentinel-3 OLCI acquisition of the same area (an image with 21 channels);

the goal is to assign a numeric label in  $\{0, 1\} \subset \mathbf{N}$ , in order to establish if it belongs to the class of hotspots related to wildfires or not. The problem is therefore structured as a binary classification task. A set of hotspots, for which the class is known, is used to train a classification model in a supervised way, which is then evaluated against a different set of hotspots.

## 4.1 Adopted Models

For the purpose of this thesis, two models were adopted. For both, different sets of hyperparameters were considered. The best for each experiment set was chosen following the validation approach described in Section 4.3.

The first model is a Random Forest Classifier, an ensemble of Decision Trees where each tree is built from a subset of the total available features. Individual trees look for boundaries in the feature space that would allow to reduce impurity. In order to contain overfitting, it is possible to limit the growth of a tree. One of the ways to do so is to impose a minimum number of samples required for a leaf, therefore stopping any further split for that branch.

This type of model is known for its interpretability: a ranking of feature importance for classification can be obtained after training, and individual decision trees of the ensemble can be visualized. An example of application of this model in a related field can be found in *Ramo and Chuvieco (2017)* [21], which addresses the burnt area monitoring task.

The second model is a Multi Layer Perceptron (MLP), defined as an ensemble of neurons, organised in feed-forward fully-connected layers, where each node in a layer is connected to all the nodes in the following. The minimum number of layers is equal to 3, considering that one is an input layer, and one is the output one; every other one in-between is denoted as "hidden". Unlike a linear perceptron, each node introduces non-linearities by adopting a non-linear activation function. The higher is the layer count, the more non-linearities can be represented.

This model was chosen in place of convolutional neural networks due to the limited amount of features that are available with respect to other common applications in the field of image processing, for both Sentinel-3 OLCI and Copernicus EMS.

## 4.2 Frameworks and Tools

In this section, the software components used for this work are introduced, as well as details on the hardware of the machine that ran the experiments.

The whole project was developed in Python with the aid of the following external libraries. Geo-spatial data was processed with *shapely*, *Fiona* and

*GeoPandas*, and then inserted in a PostgreSQL database with GIS extension enabled (referred to as PostGIS), where it was handled via *SQLAlchemy* and *GeoAlchemy2*. Rasters produced starting from geometry objects contained therein were obtained via *rasterio*. Sentinel imagery was obtained via *Sentinel-Hub*. TIF images were loaded in memory via *scikit-image* and *tifffile*. Pre-processing and visualization were performed with *Numpy*, *Pandas*, *Matplotlib*. Training of models and validation were performed with *scikit-learn*. All the versions are specified in Table 4.1.

Software Package	Version
Python	3.6.9
shapely	1.7.1
Fiona	1.8.18
GeoPandas	0.8.1
PostGIS	13.0
SQLAlchemy	1.3.20
GeoALchemy2	0.8.4
rasterio	1.2.2
Sentinel-Hub	3.2.1
scikit-image	0.17.2
tifffile	2020.9.3
numpy	1.19.4
Pandas	1.1.4
Matplotlib	3.3.4
scikit-learn	0.24.2

Table 4.1. Versions of adopted software.

The experiments were run on a workstation with an Intel Core i9-7940X @ 3.10GHz with 128GB of RAM.

## 4.3 Experimental Design and Validation Procedure

This section presents and motivates all the considered experiments (Subsection 4.3.1) as well as the validation process, and provides the used values for hyperparameters (Subsection 4.3.2).

### 4.3.1 Experimental Design

The first experiments are focused on determining a baseline - a reference value to be used for comparison with all later experiments. In particular, it was obtained starting exclusively from hotspot data, with different combinations of attributes. The following scenarios have been considered for both versions of the dataset:

- Comparing the classification label provided by the existing hotspot detection algorithms (attribute *type*) with the constructed ground truth. No models were trained for this;
- Using exclusively FRP and time features, to assess the performance of using only attributes that are not collection-specific, separately for each hotspot collection;
- Using FRP and time features combined with collection-specific brightness temperatures, separately for each hotspot collection, to check if they could bring any improvements;
- Using FRP, time features, brightness temperatures and including the *type* attribute, together with its reported *confidence*. The goal of this experiment is to verify that this label negatively influences generality of the learned features for classification by being used as a shortcut, and it was therefore excluded from further considerations. More on this in Chapter 5. This was done by training a random forest with the selected features and plotting the resulting feature importance. No MLP was trained for this purpose;
- Using FRP, time features and brightness temperatures of all three collections combined in one as previously described, in order to evaluate the option of having a single unified model.

The second batch of experiments focuses on the introduction of land cover data as further characterisation of a hotspot, available for both versions of the dataset. Therefore the following scenarios were considered for both:

- Using FRP, time features, brightness temperatures of hotspots and adding the land cover of the considered area, separately for each hotspot collection;
- Using FRP, time features, brightness temperatures and land cover of all three hotspot collections combined in one.

The third batch focuses on the introduction of Sentinel-3 OLCI data, only available for version *v3*, which is therefore the only version considered for the following experiments:

- Using FRP, time features, brightness temperatures of hotspots, combined with land cover of the surroundings, and adding Sentinel-3 OLCI imagery of the same area, separately for each hotspot collection;
- Using FRP, time features, brightness temperatures of hotspots, with land cover and Sentinel-3 OLCI data of the area, combining all three hotspot collections in one.

All experiments were performed with both presented models, except for the two for which it was specified otherwise. The results are presented and discussed in Chapter 5.

### 4.3.2 Validation Procedure

For each described experiment that required training a model, the dataset was randomly divided in two splits, stratified for class: one for training and hyperparameter tuning, and another for testing purposes. The proportions of the two are respectively 90%-10% of the total.

Then, for the training and tuning of each model, a grid search with K-fold cross-validation was performed using the f1-score as guiding metric for the process. The number of folds K is equal to 5. Finally, the best performing model is evaluated against the testing split in terms of accuracy, precision, recall, and f1-score.

The considered values for hyperparameters are shared among all experiments that use the same model. They are reported in Table 4.2 for the Random Forest Classifier, and in Table 4.3 for the Multi-Layer Perceptron.

An early-stopping criterion is adopted for the training of MLPs: a small fraction of the available data is set aside, and the model is evaluated against it at each epoch. If several epochs bring no improvements, the training process is stopped.

Hyperparameter	Considered values
Impurity criterion	{GINI, entropy}
Number of trees	100
Samples per leaf (minimum)	{2, 5, 10, 20}
Number of considered features	$\sqrt{feature\_count}$
random_state	42
Number of parallel jobs	4

Table 4.2. Considered values for the hyperparameters of Random Forests.

Hyperparameter	Considered values
Hidden Layer sizes	{(100), (128, 24), (100, 32, 8)}
Activation Function	ReLU
Solver	SGD
Batch size	200
Initial learning rate	$1 \times 10^{-3}$
Learning rate	1/5
Momentum	0.9
Max number of epochs	200
Early stopping	enabled
Number of unsuccessful epochs before early stopping	10
Validation fraction for early stopping	0.05
random_state	42

Table 4.3. Considered values for the hyperparameters of MLPs.

# Chapter 5

## Results

This chapter focuses on reporting the obtained results from the experiments presented in Chapter 4 and discussing them. In particular: Section 5.1 defines what will be considered as baseline; Section 5.2 presents the performances after integrating land cover data; Section 5.3 describes the results obtained after adding Sentinel-3 OLCI imagery; Section 5.4 reports the results aggregated by hotspot collection and model, instead of by experiment, to better visualize improvements.

### 5.1 Baseline

In Table 5.1 the results of evaluating the labels provided by existing algorithms against the proposed ground truth are shown.

The constructed ground truth is obtained based on the hypothesis that burnt areas as provided by EFFIS cover on average about 80% of the total, as reported in their statistics. If this holds true, then the vast majority of hotspots related to those fire events will be assigned to the positive class. Although admittedly a part of positive hotspots are left out and assigned to the negative class in this work, the results above show how existing algorithms lie heavily on the opposite side, with a very high recall on the positive class at the expenses of accuracy and recall for the negative class.

All the following experiments are to be interpreted while bearing in mind that models trained in this thesis use f1-score as guiding metric, therefore aiming to balance the two classes. More on this point will be discussed in Chapter 6.

In Table 5.2 the results of only using time features and FRP are shown. This experiment represents the starting point for all the following ones, which

Dataset version	Hotspot Collection	Acc.	Precision		Recall		f1
			(pos)	(neg)	(pos)	(neg)	
v2	MCD14ML	.6334	.5646	.9692	.9889	.3134	.7188
	VNP14ML	.4548	.4543	.9578	.9999	.0018	.6247
	VNP14IMGML	.6914	.6117	.9610	.9815	.4224	.7537
v3	MCD14ML	.6554	.5788	.9672	.9863	.3605	.7295
	VNP14ML	.4753	.4749	.9411	.9999	.0015	.6440
	VNP14IMGML	.7101	.6047	.9642	.9760	.5029	.7467

Table 5.1. Comparison between pre-existing labels and the proposed ground truth.

incrementally add more features and are expected to improve performances.

Dataset version	Model	Hotspot Collection	Acc.	Precision		Recall		f1
				(pos)	(neg)	(pos)	(neg)	
v2	<b>RF</b>	MCD14ML	.7912	.7736	.8077	.7907	.7917	.7821
		VNP14ML	.7983	.7777	.8155	.7781	.8151	.7779
		VNP14IMGML	.7973	.7772	.8176	.8113	.7843	.7939
	MLP	MCD14ML	.7642	.7566	.7709	.7407	.7854	.7486
		VNP14ML	.7777	.7559	.7956	.7533	.7978	.7547
		VNP14IMGML	.7498	.7296	.7703	.7626	.7380	.7458
v3	<b>RF</b>	MCD14ML	.8019	.7866	.8160	.7956	.8076	.7911
		VNP14ML	.8074	.7899	.8239	.8095	.8054	.7996
		VNP14IMGML	.8147	.7791	.8441	.8051	.8222	.7919
	MLP	MCD14ML	.7570	.7442	.7682	.7380	.7739	.7411
		VNP14ML	.7652	.7545	.7747	.7491	.7797	.7518
		VNP14IMGML	.7606	.7290	.7848	.7214	.7911	.7252

Table 5.2. Performances by using only FRP and time features.

As expected, recall for the positive class dropped, while accuracy increased together with performance on the negative class.

Table 5.3 shows the scores obtained by adding collection-specific brightness temperatures as features for the respective models.

Dataset version	Model	Hotspot Collection	Acc.	Precision		Recall		f1
				(pos)	(neg)	(pos)	(neg)	
v2	<b>RF</b>	MCD14ML	.8117	.7930	.8294	.8153	.8084	.8040
		VNP14ML	.8142	.7912	.8339	.8025	.8240	.7968
		VNP14IMGML	.8052	.7843	.8263	.8208	.7907	.8021
	MLP	MCD14ML	.7727	.7709	.7742	.7402	.8020	.7552
		VNP14ML	.7880	.7692	.8033	.7614	.8101	.7653
		VNP14IMGML	.7589	.7536	.7637	.7413	.7752	.7474
v3	<b>RF</b>	MCD14ML	.8254	.8103	.8392	.8219	.8285	.8160
		VNP14ML	.8199	.8046	.8343	.8197	.8201	.8121
		VNP14IMGML	.8261	.7937	.8525	.8146	.8350	.8040
	MLP	MCD14ML	.7693	.7644	.7735	.7380	.7973	.7510
		VNP14ML	.7727	.7500	.7948	.7816	.7646	.7655
		VNP14IMGML	.7675	.7437	.7848	.7157	.8079	.7294

Table 5.3. Performances by using only FRP, time features and brightness temperatures.

There are slight improvements in terms of accuracy and recall for the negative class, more pronounced for the MODIS collection than for the others. It is also important to note how performances are comparatively slightly higher on  $v3$  than on  $v2$ , presumably due the reduced sample size.

Dataset version	Model	Hotspot Collection	Acc.	Precision		Recall		f1
				(pos)	(neg)	(pos)	(neg)	
v2	<b>RF</b>	MCD14ML	.8314	.8038	.8593	.8523	.8127	.8273
		VNP14ML	.8185	.7938	.8399	.8107	.8250	.8022
		VNP14IMGML	.8374	.8024	.8763	.8783	.7994	.8386
	MLP	MCD14ML	.7989	.7739	.8237	.8131	.7861	.7930
		VNP14ML	.7886	.7689	.8049	.7640	.8091	.7664
		VNP14IMGML	.8037	.7655	.8479	.8535	.7575	.8071
v3	<b>RF</b>	MCD14ML	.8480	.8244	.8709	.8608	.8366	.8422
		VNP14ML	.8230	.8058	.8394	.8263	.8201	.8159
		VNP14IMGML	.8617	.8182	.9005	.8798	.8477	.8479
	MLP	MCD14ML	.8108	.7839	.8373	.8262	.7970	.8045
		VNP14ML	.7616	.7467	.7752	.7532	.7692	.7499
		VNP14IMGML	.8219	.7679	.8728	.8503	.7998	.8070

Table 5.4. Performances obtained by using FRP, time features, brightness temperatures and pre-existing labels.

Table 5.4 shows the effects of adding the pre-existing label and its confidence to the set of features. Some extra minor improvements appear. In order to explore the impact these features have in terms of decisions for the models, feature importance as learned by the Random Forest during training is plotted. Figure 5.1 focuses on MODIS, and also shows this ranking as obtained for the previous experiments for reference.

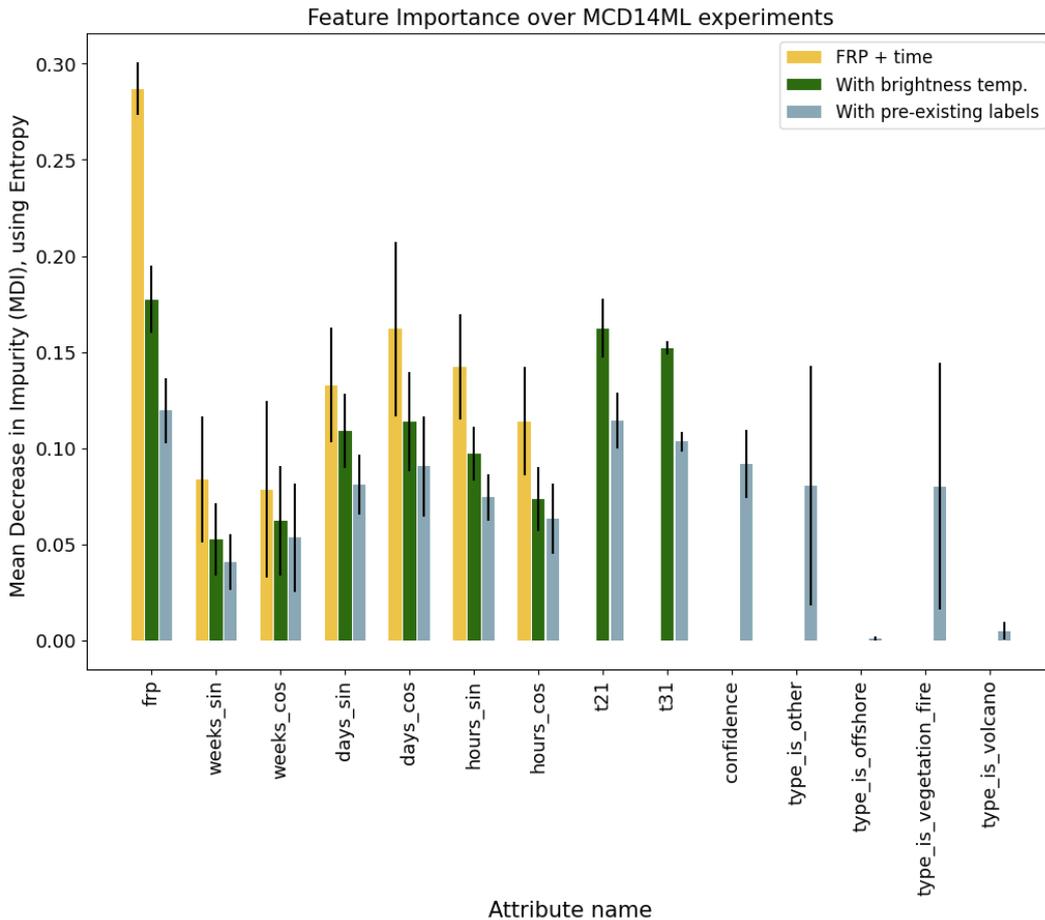


Figure 5.1. Comparison of feature importance using different sets of attributes. It is computed as Mean Decrease in Impurity (MDI) from the trained Random Forest models. The adopted impurity measure for these experiments was Entropy.

When included, the provided label obtains a significant importance, and in some trees it is even higher than any other feature. Since previous experiments have shown how other features can be sufficient on their own to train a rudimentary model, and considering how they can get obscured when the label is included, it will be excluded from further considerations.

The last experiment of this batch tries to combine all hotspot data sources to create a unified model. The results are displayed in Table 5.5.

Dataset version	Model	Acc.	Precision		Recall		f1
			(pos)	(neg)	(pos)	(neg)	
v2	<b>RF</b>	.8109	.7924	.8286	.8157	.8066	.8039
	MLP	.7572	.7391	.7742	.7556	.7586	.7473
v3	<b>RF</b>	.8229	.8021	.8400	.8044	.8380	.8033
	MLP	.7593	.7351	.7787	.7264	.7863	.7307

Table 5.5. Performances obtained by using FRP, time features and brightness temperatures, while combining all hotspot collections.

## 5.2 Land Cover experiments

The second batch of experiments assesses the effects of the addition of land cover data for the surroundings of the area affected by a hotspot. Results are shown in Table 5.6.

For both versions of the dataset, there are improvements with respect to the baseline results. While these are more marked for the recall of the positives, this metric increased for both classes, leading to an improvement on accuracy as well. In particular, the drop in recall of the positive class is now limited, when compared to the baseline. The obtained results even show better performances than what was achieved with models including pre-existing labels.

## 5.3 Sentinel-3 OLCI experiments

The last batch of experiments incorporates Sentinel-3 OLCI imagery to the previous ones. It is only available for the *v3* dataset version, since *v2* doesn't include this data source. The results are reported in Table 5.7.

Dataset version	Model	Hotspot Collection	Acc.	Precision		Recall		f1
				(pos)	(neg)	(pos)	(neg)	
v2	<b>RF</b>	MCD14ML	.8714	.8137	.9419	.9448	.8052	.8744
		VNP14ML	.8836	.8236	.9490	.9462	.8316	.8807
		VNP14IMGML	.8608	.7985	.9442	.9505	.7776	.8679
		Combined	.8634	.8020	.9416	.9459	.7887	.8681
	MLP	MCD14ML	.8624	.8094	.9256	.9283	.8031	.8648
		VNP14ML	.8693	.8120	.9307	.9264	.8218	.8655
		VNP14IMGML	.8460	.7859	.9264	.9346	.7639	.8538
		Combined	.8515	.7910	.9289	.9344	.7766	.8567
v3	<b>RF</b>	MCD14ML	.8928	.8376	.9573	.9582	.8345	.8939
		VNP14ML	.9081	.8516	.9756	.9766	.8462	.9098
		VNP14IMGML	.8948	.8248	.9682	.9646	.8404	.8892
		Combined	.8999	.8351	.9706	.9687	.8438	.8969
	MLP	MCD14ML	.8898	.8396	.9468	.9471	.8388	.8901
		VNP14ML	.9043	.8521	.9650	.9660	.8486	.9055
		VNP14IMGML	.8869	.8178	.9592	.9544	.8344	.8808
		Combined	.8907	.8233	.9654	.9635	.8312	.8879

Table 5.6. Performances obtained by adding land cover information to hotspot data.

Dataset version	Model	Hotspot Collection	Acc.	Precision		Recall		f1
				(pos)	(neg)	(pos)	(neg)	
v3	<b>RF</b>	MCD14ML	.8949	.8351	.9665	.9677	.8301	.8965
		VNP14ML	.9082	.8506	.9771	.9780	.8451	.9099
		VNP14IMGML	.9031	.8374	.9705	.9668	.8535	.8975
		Combined	.9088	.8490	.9720	.9697	.8590	.9053
	MLP	MCD14ML	.8855	.8321	.9474	.9482	.8298	.8863
		VNP14ML	.8964	.8449	.9562	.9572	.8415	.8975
		VNP14IMGML	.8907	.8330	.9471	.9389	.8531	.8828
		Combined	.8962	.8407	.9535	.9491	.8530	.8916

Table 5.7. Performances obtained by <sup>55</sup>integrating Sentinel-3 OLCI imagery with hotspot data and land cover.

The values show very minor improvements for Random Forest in terms of positive class recall, although consistently on all hotspot products. This could possibly be caused by the limited quantity of information due to the spatial resolution of the instrument and the adopted collection process for this work. On the other side, performances for MLP show a minor decline for the MCD14ML and VNP14ML products, but a minor improvement on VNP14IMGML and the combination of all hotspot products. Further investigation focused on the learning process for this model would be needed, and it will be suggested in Chapter 6.

## 5.4 Aggregation by hotspot collection and model

This section presents the obtained results aggregated by collection and model, so that the incremental improvements brought by the integration of the adopted data sources are visible, and adds some considerations to summarize them.

In particular: Tables 5.8 and 5.9 focus on the MCD14ML product; Tables 5.10 and 5.11 are for VNP14ML; Tables 5.12 and 5.13 are for VNP14IMGML; Tables 5.14 and 5.15 are for combined hotspots.

MCD14ML, Random Forest							
Dataset v.	Experiment	Accuracy	Precision		Recall		f1-score
			(pos)	(neg)	(pos)	(neg)	
v2	Baseline	.6334	.5646	<b>.9692</b>	<b>.9889</b>	.3134	.7188
	FRP + time	.7912	.7736	.8077	.7907	.7917	.7821
	Brightness temp.	.8117	.7930	.8294	.8153	<b>.8084</b>	.8040
	Landcover	<b>.8714</b>	<b>.8137</b>	.9419	.9448	.8052	<b>.8744</b>
v3	Baseline	.6554	.5788	<b>.9672</b>	<b>.9863</b>	.3605	.7295
	FRP + time	.8019	.7866	.8160	.7956	.8076	.7911
	Brightness temp.	.8254	.8103	.8392	.8219	.8285	.8160
	Landcover	.8928	.8376	.9573	.9582	<b>.8345</b>	.8939
	Sentinel	<b>.8949</b>	<b>.8351</b>	.9665	.9677	.8301	<b>.8965</b>

Table 5.8. Progressive changes in performances for Random Forest models on the MCD14ML product in the presented experiments.

MCD14ML, Multi-Layer Perceptron							
Dataset v.	Experiment	Accuracy	Precision		Recall		f1-score
			(pos)	(neg)	(pos)	(neg)	
v2	Baseline	.6334	.5646	<b>.9692</b>	<b>.9889</b>	.3134	.7188
	FRP + time	.7642	.7566	.7709	.7407	.7854	.7486
	Brightness temp.	.7727	.7709	.7742	.7402	.8020	.7552
	Landcover	<b>.8624</b>	<b>.8094</b>	.9256	.9283	<b>.8031</b>	<b>.8648</b>
v3	Baseline	.6554	.5788	<b>.9672</b>	<b>.9863</b>	.3605	.7295
	FRP + time	.7570	.7442	.7682	.7380	.7739	.7411
	Brightness temp.	.7693	.7644	.7735	.7380	.7973	.7510
	Landcover	<b>.8898</b>	<b>.8396</b>	.9468	.9471	<b>.8388</b>	<b>.8901</b>
	Sentinel	.8855	.8321	.9474	.9482	.8298	.8863

Table 5.9. Progressive changes in performances for MLP models on the MCD14ML product in the presented experiments.

VNP14ML, Random Forest							
Dataset v.	Experiment	Accuracy	Precision		Recall		f1-score
			(pos)	(neg)	(pos)	(neg)	
v2	Baseline	.4548	.4543	<b>.9578</b>	<b>.9999</b>	.0018	.6247
	FRP + time	.7983	.7777	.8155	.7781	.8151	.7779
	Brightness temp.	.8142	.7912	.8339	.8025	.8240	.7968
	Landcover	<b>.8836</b>	<b>.8236</b>	.9490	.9462	<b>.8316</b>	<b>.8807</b>
v3	Baseline	.4753	.4749	.9411	<b>.9999</b>	.0015	.6440
	FRP + time	.8074	.7899	.8239	.8095	.8054	.7996
	Brightness temp.	.8199	.8046	.8343	.8197	.8201	.8121
	Landcover	.9081	<b>.8516</b>	.9756	.9766	<b>.8462</b>	.9098
	Sentinel	<b>.9082</b>	.8506	<b>.9771</b>	.9780	.8451	<b>.9099</b>

Table 5.10. Progressive changes in performances for Random Forest models on the VNP14ML product in the presented experiments.

VNP14ML, Multi-Layer Perceptron							
Dataset v.	Experiment	Accuracy	Precision		Recall		f1-score
			(pos)	(neg)	(pos)	(neg)	
v2	Baseline	.4548	.4543	<b>.9578</b>	<b>.9999</b>	.0018	.6247
	FRP + time	.7777	.7559	.7956	.7533	.7978	.7547
	Brightness temp.	.7880	.7692	.8033	.7614	.8101	.7653
	Landcover	<b>.8693</b>	<b>.8120</b>	.9307	.9264	<b>.8218</b>	<b>.8655</b>
v3	Baseline	.4753	.4749	.9411	<b>.9999</b>	.0015	.6440
	FRP + time	.7652	.7545	.7747	.7491	.7797	.7518
	Brightness temp.	.7727	.7500	.7948	.7816	.7646	.7655
	Landcover	<b>.9043</b>	<b>.8521</b>	<b>.9650</b>	.9660	<b>.8486</b>	<b>.9055</b>
	Sentinel	.8964	.8449	.9562	.9572	.8415	.8975

Table 5.11. Progressive changes in performances for MLP models on the VNP14ML product in the presented experiments.

VNP14IMGML, Random Forest							
Dataset v.	Experiment	Accuracy	Precision		Recall		f1-score
			(pos)	(neg)	(pos)	(neg)	
v2	Baseline	.6914	.6117	<b>.9610</b>	<b>.9815</b>	.4224	.7537
	FRP + time	.7973	.7772	.8176	.8113	.7843	.7939
	Brightness temp.	.8052	.7843	.8263	.8208	<b>.7907</b>	.8021
	Landcover	<b>.8608</b>	<b>.7985</b>	.9442	.9505	.7776	<b>.8679</b>
v3	Baseline	.7101	.6047	.9642	<b>.9760</b>	.5029	.7467
	FRP + time	.8147	.7791	.8441	.8051	.8222	.7919
	Brightness temp.	.8261	.7937	.8525	.8146	.8350	.8040
	Landcover	.8948	.8248	.9682	.9646	.8404	.8892
	Sentinel	<b>.9031</b>	<b>.8374</b>	.9705	.9668	<b>.8535</b>	<b>.8975</b>

Table 5.12. Progressive changes in performances for Random Forest models on the VNP14IMGML product in the presented experiments.

VNP14IMGML, Multi-Layer Perceptron							
Dataset v.	Experiment	Accuracy	Precision		Recall		f1-score
			(pos)	(neg)	(pos)	(neg)	
v2	Baseline	.6914	.6117	<b>.9610</b>	<b>.9815</b>	.4224	.7537
	FRP + time	.7498	.7296	.7703	.7626	.7380	.7458
	Brightness temp.	.7589	.7536	.7637	.7413	<b>.7752</b>	.7474
	Landcover	<b>.8460</b>	<b>.7859</b>	.9264	.9346	.7639	<b>.8538</b>
v3	Baseline	.7101	.6047	<b>.9642</b>	<b>.9760</b>	.5029	.7467
	FRP + time	.7606	.7290	.7848	.7214	.7911	.7252
	Brightness temp.	.7675	.7437	.7848	.7157	.8079	.7294
	Landcover	.8869	.8178	.9592	.9544	.8344	.8808
	Sentinel	<b>.8907</b>	<b>.8330</b>	.9471	.9389	<b>.8531</b>	<b>.8828</b>

Table 5.13. Progressive changes in performances for MLP models on the VNP14IMGML product in the presented experiments.

Combined hotspots, Random Forest							
Dataset v.	Experiment	Accuracy	Precision		Recall		f1-score
			(pos)	(neg)	(pos)	(neg)	
v2	Brightness temp.	.8109	.7924	.8286	.8157	.8066	.8039
	Landcover	<b>.8634</b>	<b>.8020</b>	<b>.9416</b>	<b>.9459</b>	<b>.7887</b>	<b>.8681</b>
v3	Brightness temp.	.8229	.8021	.8400	.8044	.8380	.8033
	Landcover	.8999	.8351	.9706	.9687	.8438	.8969
	Sentinel	<b>.9088</b>	<b>.8490</b>	<b>.9720</b>	<b>.9697</b>	<b>.8590</b>	<b>.9053</b>

Table 5.14. Progressive changes in performances for Random Forest models on combined hotspots in the presented experiments.

Combined hotspots, Multi-Layer Perceptron							
Dataset v.	Experiment	Accuracy	Precision		Recall		f1-score
			(pos)	(neg)	(pos)	(neg)	
v2	Brightness temp.	.7572	.7391	.7742	.7556	.7586	.7473
	Landcover	<b>.8515</b>	<b>.7910</b>	<b>.9289</b>	<b>.9344</b>	<b>.7766</b>	<b>.8567</b>
v3	Brightness temp.	.7593	.7351	.7787	.7264	.7863	.7307
	Landcover	.8907	.8233	<b>.9654</b>	<b>.9635</b>	.8312	.8879
	Sentinel	<b>.8962</b>	<b>.8407</b>	.9535	.9491	<b>.8530</b>	<b>.8916</b>

Table 5.15. Progressive changes in performances for MLP models on combined hotspots in the presented experiments.

In general, better performances were obtained on the  $v3$  version of the dataset, presumably both due to the reduced sample size and better data quality.

The results show an overall improvement with each additional characterization of the affected area, hinting that further refinements in terms of data cleaning and adopted data sources could be beneficial to better solve the task.

In terms of models, Random Forests performed better than their MLP counterparts in any given experiment, although they all followed the same general trend. This suggests that a finer tuning could be needed to fully harness its effectiveness.

# Chapter 6

## Conclusions

This work explored a different perspective on the task of hotspot disambiguation, albeit incomplete on some aspects and improvable on many. Firstly, it asserted that pre-existing techniques are unbalanced on this specific task, although very good at detection. While this means that most fire events will be observed, it also implies that further inspections are needed before calling for an intervention.

Secondly, it showed how integrating more and more data sources for better characterisation of an area can prove to be beneficial for the task. The fact that performances for both classes increase by adding more characterising data sources suggests that they bring useful information for the addressed problem. Overall, this study was successful in its attempt, but several aspects could be improved upon in its future developments.

Starting from data sources, a valid idea would be to use hotspot detections performed by satellites that can provide data for the considered area of interest with a higher frequency. An example of this would be to adopt data from geo-stationary satellites, assuming that the whole area could be covered. This would allow quicker response times in case of a fire, and reduce the cost of misclassifying a single data point.

Considering those sources that were adopted in this work, one further piece of information that could prove useful is raw imagery of the affected area, with measurements from all bands. As discussed in related works, a fire event alters the affected area with respect to the surroundings, and *Giglio et al. (2003)* [4] use this to define contextual thresholds for the detection. Instead, the current implementation only has access to absolute features, which allow for quick identification but do not provide a complete characterization of the

context.

One further refinement would be to more precisely define the boundaries of the considered European macro-area, in order to exclude any hotspot detection for which no land cover information is available, and therefore further reducing noise.

Another interesting idea would be to use satellite data with higher spatial resolution, both for hotspot data and other characterizations. As an example, Sentinel-5P imagery could be evaluated for integration.

Last but not least, having verified data for the entirety of the fire activations in the considered area would make the definition of a ground truth much easier and more sound, even if made available only for a short time interval.

Considering data preprocessing and preparation, geo-spatial features could be reprojected to a different reference system that better represents the Pan-European area at all latitudes, like ESRI:3035.

Related to this, and independently from the adopted reference system, it would be appropriate to always have a consistent number of features at all coordinates. In particular, for land cover rasters and sentinel imagery, an improvement could be made by ensuring that the considered region has the same required linear size independently from its latitude.

In terms of models and methodology, an improvement could be made in terms of performances by further fine-tuning each model. A possible experiment could be to use different metrics to guide training, assigning different weights to the positive and negative classes, in order to have more fair comparisons with the baseline and put more emphasis on the positives.

In regards to combining data from multiple hotspot collections for a unified model, the presented experiments couldn't find any significant changes in terms of performances with respect to the split counterparts. More detailed investigation would be required.

One further experiment could be to adopt neural networks combined with all the previous suggestions as a possible model, as a way to address this non-linearly-separable problem.

# Acknowledgements

This thesis is meant to be viewed in the context of the SAFERS project [22], funded by the European Union's Horizon 2020 research, and coordinated by LINKS Foundation, that allowed me to perform this study.

Thanks to my internship tutor, Edoardo Arnaudo, who guided me all the way.

Special thanks to Professor Paolo Garza for supervising my work.

And last but not least, thanks to my fiancé, my family and my friends, who supported me emotionally and financially through the struggles of a process of trial and error, which proved to be hard to cope with on my own.



# References

- [1] Houghton, J T, Jenkins, G J, and Ephraums, J J. Climate change, 1990.
- [2] Knutson, T., McBride, J., Chan, J. et al. Tropical cyclones and climate change. *Nature Geoscience*, 3:157–163, 2010.
- [3] San-Miguel-Ayanz, J., Durrant, T., Boca, R., Liberta', G., Branco, A., De Rigo, D., Ferrari, D., Maianti, P., Artes Vivancos, T., Pfeiffer, H., Loffler, P., Nuijten, D., Leray, T. and Jacome Felix Oom, D.,. Forest Fires in Europe, Middle East and North Africa 2018. Technical report, European Union, 2019.
- [4] Louis Giglio, Jacques Descloitres, Christopher O. Justice, and Yoram J. Kaufman. An enhanced contextual fire detection algorithm for modis. *Remote Sensing of Environment*, 87(2):273–282, 2003.
- [5] Valeria Di Biase and Giovanni Laneve. Geostationary sensor based forest fire detection and monitoring: An improved version of the sfide algorithm. *Remote Sensing*, 10(5), 2018.
- [6] Wilfrid Schroeder, Patricia Oliva, Louis Giglio, and Ivan A. Csiszar. The new viirs 375m active fire detection data product: Algorithm description and initial assessment. *Remote Sensing of Environment*, 143:85–96, 2014.
- [7] European Union. Corine land cover. Available online: <https://land.copernicus.eu/pan-european/corine-land-cover>.
- [8] European Union. Sentinel missions. Available online: <https://sentinels.copernicus.eu/web/sentinel/home>.
- [9] European Union. Copernicus EMS. Available online: <https://emergency.copernicus.eu/>.
- [10] European Union. EFFIS. Available online: <https://effis.jrc.ec.europa.eu/>.
- [11] Giglio et al. *MODIS Collection 6 Active Fire Product User's Guide*. Available online: [https://modis-fire.umd.edu/files/MODIS\\_C6\\_Fire\\_User\\_Guide\\_B.pdf](https://modis-fire.umd.edu/files/MODIS_C6_Fire_User_Guide_B.pdf).

- [12] M.J. Wooster, B. Zhukov, and D. Oertel. Fire radiative energy for quantitative study of biomass burning: derivation from the bird experimental satellite and comparison to modis fire products. *Remote Sensing of Environment*, 86(1):83–107, 2003.
- [13] M.J. Wooster, W. Xu, and T. Nightingale. Sentinel-3 slstr active fire detection and frp product: Pre-launch algorithm development and performance evaluation using modis and aster datasets. *Remote Sensing of Environment*, 120:236–254, 2012. The Sentinel Missions - New Opportunities for Science.
- [14] Jeffrey T. Morissette, Louis Giglio, Ivan Csiszar, and Christopher O. Justice. Validation of the modis active fire product over southern africa with aster data. *International Journal of Remote Sensing*, 26(19):4239–4264, 2005.
- [15] Giglio et al. *Visible Infrared Imaging Radiometer Suite (VIIRS). 375 m & 750 m. Active Fire Detection Data Sets Based on NASA VIIRS Land Science Investigator Processing System (SIPS). Reprocessed Data - Version 1*. Available online: [https://cdn.earthdata.nasa.gov/conduit/upload/9019/VIIRS\\_activefire\\_User\\_Guide\\_v1\\_3.pdf](https://cdn.earthdata.nasa.gov/conduit/upload/9019/VIIRS_activefire_User_Guide_v1_3.pdf).
- [16] Michael Nolde, Simon Plank, and Torsten Riedlinger. An adaptive and extensible system for satellite-based, large scale burnt area monitoring in near-real time. *Remote Sensing*, 12(13), 2020.
- [17] Alessandro Farasin, Luca Colomba, and Paolo Garza. Double-step u-net: A deep learning-based approach for the estimation of wildfire damage severity through sentinel-2 satellite data. *Applied Sciences*, 10(12), 2020.
- [18] N. Sifakis, D. Paronis, and I. Keramitsoglou. Combining avhrr imagery with corine land cover data to observe forest fires and to assess their consequences. *International Journal of Applied Earth Observation and Geoinformation*, 5(4):263–274, 2004.
- [19] European Union. Classes for corine land cover. Available online: <https://land.copernicus.eu/user-corner/technical-library/corine-land-cover-nomenclature-guidelines/html>.
- [20] Sentinel-Hub. Available online: <https://www.sentinel-hub.com/>.
- [21] Rubén Ramo and Emilio Chuvieco. Developing a random forest algorithm for modis global burned area classification. *Remote Sensing*, 9(11), 2017.
- [22] European Union. SAFERS. Available online: <https://safers-project.eu/>.