SAPIENZA
UNIVERSITÀ DI ROMA

A Scalable Fire Danger Index based on Sentinel Imagery

Department of Astronautical, Electrical and Energy Engineering
Faculty of Civil and Industrial Engineering
PhD School in Science and Technologies for Industrial Innovation
Energy and Environment (XXXV cycle)

Valerio Pampanoni
ID number 1480613

Advisor
Prof. Giovanni Laneve

Academic Year 2021/2022

Thesis defended on January 25 2023
in front of a Board of Examiners composed by:

Dr. Marco Sabatini (chairman)
Prof. Lucia Fontana
Prof. Fulvio Palmieri

A Scalable Fire Danger Index based on Sentinel Imagery
Sapienza University of Rome

© 2023 Valerio Pampanoni. All rights reserved

This thesis has been typeset by LaTeX and the Sapthesis class.

Version: January 17 2022

Author's email: valerio.pampanoni@uniroma1.it

Ab astris salus orbis
Carlo Pelanda

Abstract

The incidence of wildfires and megafires and their disastrous consequences is increasing all over the planet. According to the latest European Forest Fire Information System annual fire report [115], in 2021 alone wildfires burned a surface area more than twice the size of Luxembourg, including more than a thousand square kilometres of Natura 2000 protected areas. In addition, 2022 has registered the highest number of wildfires since 2006, and will also be recorded as one of the driest years on record. Assuming that the most efficient and cost-effective way limit the damage caused by wildfires consists in their prevention, building tools to allow the decision makers to allocate resources using state of the art technology and fresh data is of the utmost importance. To this end, the combined usage of data from weather and satellite platforms capable to provide data on a regional or national scale and at a high temporal frequency provides the optimal solution for assessing and monitoring the state of the vegetation. However, users of fire danger product users often complain about the resolution of the provided products. While moderate- or coarse-resolution products may be adequate to cover the regional or national scale, high-resolution products are required to properly describe the fire danger in relatively small-sized areas of high interest in fire danger modelling, such as wildland-urban interfaces, national parks or protected areas. Using a different fire danger product based on the spatial scale of the target may be impractical and increase the workload and training requirements for the personnel. For this reason, we propose a scalable fire danger index based on Sentinel imagery that is able to cover different spatial scales by exploiting the surface reflectances provided by different Sentinel products (i.e. Sentinel-2 and Sentinel-3).

This novel index, named Daily Fire Danger Index, exploits both weather and satellite data to estimate all the main variables of fire danger, such as the amount of dead fuel, moisture of the dead and live fuels, wind speed, evapotranspiration etc, and is calibrated using the historical records of wildfire occurrence in the target region. In particular, the live fuel moisture content is estimated using a state of the art procedure based on the inversion of radiative transfer models of the PROSAIL family. The index was tested in Sardinia, a region well-known for its proneness to wildfires and which is also regularly affected by megafires, and the performance comparison with the Canadian Fire Weather Index shows very significant improvements on the capability to discriminate fire danger even at a moderate resolution. Finally, the 2021 Planargia-Montiferru megafire was selected as a case study to showcase the added value of the high-resolution version of the index.

A photograph of a fire danger sign in the Gutturu Mannu forest in southern Sardinia, taken during a fuel sampling campaign in March 2022 by the PhD candidate.

Contents

Acronyms

ADAI Association for the Development of Industrial Aerodynamics. 61
ALA Average Leaf Angle. 47
AOI Area of Interest. 24
AVHRR Advanced Very High Resolution Radiometer. 5, 23

CFFWI Canadian Forest Fire Weather Index. 2, 7
CLC Corine Land Cover Map. 28
CNR National Research Council. 61

DBT Dry Bulb Temperature. 25
DEM Digital Elevation Model. 20
DFDI Daily Fire Danger Index. 2, 12
DFHI Daily Fire Hazard Index. 6
DMC Dry Matter Content. 9
DRMKC Disaster Risk Management Knowledge Center of the European Union. 3

EFFIS European Forest Fire Information System. 2, 29
EMC Equilibrium Moisture Content. 25
ERC Energy Release Component. 5
EWT Equivalent Water Thickness. 9

FAO Food and Agriculture Organization. 20, 94
FIRMS Fire Information for Resource Management System. 2
FPI Fire Potential Index. 5
FRP Fire Radiative Power. 73

GDAL Geospatial Data Abstraction Library. 27
GRASS Geographic Resources Analysis Support System. 27

GSA Global Sensitivity Analysis. 1, 45

GVMI Global Vegetation Moisture Index. 8

IBE Institute for BioEconomy. 61

IC Ignition Component. 5

ICONA Instituto para la Conservación de la Naturaleza. 7

INMG Indice de Risco Meteorologico. 7

JRC Joint Research Center. 29

LAI Leaf Area Index. 32, 38

LFMC Live Fuel Moisture Content. 1, 9

LIDF Leaf Inclination Distribution Function. 33, 38, 39

LUT Look-Up Table. 32

MDI Mean Decrease in Impurity. 73

MFPI Modified Fire Potential Index. 9

MSI MultiSpectral Instrument. 46

NDVI Normalized Difference Vegetation Index. 5, 23

NFDRS National Fire-Danger Rating System. 5, 25

NIR Near Infrared. 8

NOAA National Oceanic and Atmospheric Administration. 5

NRC Natural Resources Canada. 82

OLCI Ocean and Land Colour Instrument. 46

PAR Photosynthetic Active Region. 23

PREFER Space-based information support for the Prevention and Recovery of Forest Fires Emergency in the Mediterranean Area). 6

RG Relative Greenness. 5

RISICO RISchio Incendi e COordinamento. 7

RTM Radiative Transfer Model. 1, 8

SAC Specific absorption coefficients. 35

SAIL Scattering by Arbitrarily Inclined Leaves. 8, 32

SC Spread Component. 5

SFIDE	Satellite FIre DEtection. 85
SLSTR	Sea and Land Surface Temperature Radiometer. 46
SRF	Spectral Response Function. 57
SWIR	Shortwave Infrared. 8
TIR	Thermal Infrared. 38
USDA	United States Department of Agriculture. 5
WUI	Wildland-Urban Interface. 12

Introduction

This doctoral thesis proposes a novel, dynamic and scalable fire danger index based on weather data and different resolution satellite surface reflectance products. The problem of fire danger assessment is an interdisciplinary one, and many different approaches are extensively documented in the scientific literature. For these reasons, this thesis is divided into six chapters:

In Chapter 1 we give a brief summary of the history of fire danger modelling and an overview of the state of the art using weather and satellite data. This summary focuses mostly on the American fire danger literature, since our fire danger index descends from this specific methodological tradition, and aims to provide the reader with an understanding of the history of fire danger modelling with a particular focus on the satellite-based family of products. In addition, we mention a number of popular indices belonging to other methodological traditions that do not involve satellite data, some of which are still very widely used even outside of their countries of origin. Subsequently, we give an overview of the objectives of this thesis, namely on how we mean to improve the capability of assessing fire danger through an index based on both weather and satellite information, and in particular to exploit the latter to cover different spatial scales. Finally, we provide some basic information on the bioclimatic characteristics of the island of Sardinia, motivating its selection as study area for this work.

In Chapter 2 we dedicate a section to the enumeration and description of the variables involved in fire danger modelling which will be used in our algorithm, providing a justification for their inclusion in the process of fire danger modelling. Subsequently, we provide the source of the data when the variables are obtained from a third-party, and the description of the estimation process when the variables are produced by ourselves in order to ensure the full documentation of the calculation process and the reproducibility of the results.

In Chapter 3 we provide a detailed description of the Radiative Transfer Model (RTM)-based Live Fuel Moisture Content (LFMC) inversion process used to estimate this variable from satellite surface reflectances, which includes Global Sensitivity Analysis (GSA) of the PROSAIL family of models and a partial validation of the LFMC product using field data. Given the complexity of the issue and the importance of the variable in fire danger modelling, it was deemed necessary to dedicate an entire chapter to this matter.

In Chapter 4 we propose an additive fire danger model and define the formula of the novel fire danger index. Subsequently, we show how each component of the index is calculated, namely how each biophysical variable is transformed into a danger variable, normalized, weighted and integrated into a dimensionless index with values ranging from 0 to 100. Then we define the criteria for the calibration and performance evaluation of the index. The performance evaluation was carried out using fire occurrence data in the form of the aggregated MODIS Collection 6 [54] and VIIRS 375 m [118] active fire products and burned area perimeters provided by the Sardinian fire department. From now on, for the sake of brevity, we will refer to the aggregated MODIS and VIIRS active fire products as the Fire Information for Resource Management System (FIRMS) active fire products.

In Chapter 5 we show a comparison of performance between our index and the Canadian Forest Fire Weather Index (CFFWI), the main fire danger indicator of the European Forest Fire Information System (EFFIS) and in general one of the most widely used fire danger indices. We discuss differences between the fire danger assessment provided by the two indices using the fire seasons of 2019, 2020 and 2021 as our test framework.

In Chapter 6 we analyse the added value of the high-resolution index in the context of the Planargia-Montiferru megafire, a devastating wildfire that burned about 13000 hectares of Sardinian vegetation in the summer of 2021. First of all, we provide a description of the circumstances that caused the evolution of a relatively small fire into a megafire. Subsequently, we draw a comparison between the information provided by the moderate-resolution Daily Fire Danger Index (DFDI) maps and a small AOI of their high-resolution counterparts corresponding to the ignition zone of the original fire.

Chapter 1

State of the Art and Objectives of the Research

1.1 Terminology

Before even beginning any type of discussion on the topic of fire danger, it is absolutely mandatory to clarify the meaning of the term "danger", both *per se* and relatively to the context of integrated risk assessment, in order to avoid confusion with related terms such as "risk" and "hazard". This is necessary not only for readers not accustomed to the jargon of this specific branch of remote sensing and environmental management, but also for technicians and experts, since it is not rare to see the same terms used with different meanings in the scientific literature.

The international guidelines on integrated risk evaluation established that such assessments should include three aspects: hazard, exposure and vulnerability. To clarify what each of these terms means, the Science for Disaster Risk Management 2020 report by the Disaster Risk Management Knowledge Center of the European Union (DRMKC)[18] provides the following operational definition of the concept of risk:

> "*Risk is the potential for adverse consequences or impacts due to the interaction between one or more natural or human-induced hazards, exposure of humans, infrastructure and ecosystems, and systems' vulnerabilities*"

Therefore, modeling risk implies also modeling these three different variables and their interactions:

- "*Hazard* [...] *[is] the process, phenomenon or human activity that carries the potential to cause loss of life, injury or other health impacts, property damage, social and economic disruption or environmental degradation. Hazards can be natural (e.g. earthquakes, droughts, floods) or anthropogenic (e.g. oil spills, terrorist attacks) in origin and can be characterised by their location, likelihood of occurrence, intensity or magnitude, duration, and extent.*"
- "*Exposure* [...] *[is] the presence of people, infrastructure, housing, production capacities, species or ecosystems, and other tangible human assets in places and settings that could be adversely affected by one or multiple hazards.*"
- "*Vulnerability* [...] *[is] the propensity or predisposition of an individual, a community, infrastructure, assets or systems (incl. ecosystems) to be adversely*

> *affected. Vulnerability encompasses a variety of concepts and elements including sensitivity or susceptibility to harm and lack of capacity to cope and adapt."*

The translation of these concepts into the realm of fire management is not immediate. In particular, the term "hazard" in this context has often been used interchangeably with "danger" and with a different meaning than the one defined above: in fire management, fire hazard or danger is defined as the propensity of fire to ignite and propagate. The higher the fire danger, the higher the probability that vegetation ignites, and once ignited, the higher the chance that it propagates. In order to avoid confusion, we will adopt the "danger" terminology, and treat danger as one of the components of fire risk along with exposure and vulnerability.

After defining the different components of fire risk modelling, we should remark that the object of this thesis is fire danger alone, and therefore, we will deal exclusively with variables that concern the vegetation and its proneness to burn. While it is clear that, especially in Mediterranean climates and societies, the overwhelming majority of fires is caused either accidentally or voluntarily by humans, and more in general that the human factor is a fundamental variable of fire risk modelling, we will not deal with it. The human factor is usually incorporated in the exposure component of fire modelling, while we specifically deal with the assessment of the state of the vegetation in relation to its capability to ignite and support the spread of large fires.

1.2 A Brief History of Fire Danger

According to H. T. Gisborne's 1936 handbook titled "The Principles of Measuring Forest Fire Danger" [55], the history of fire danger as a subject of scientific research may be dated as back as 1922, when the U.S. Forest Service, with headquarters in Missoula, Montana, started to investigate which variables were involved in determining fire insurgence and spread and how should they be measured and used in the field. Gisborne's effort to translate these measurements into a numerical fire danger rating ranged across 17 million acres of Montana and Northern Idaho forests, and years of tests and successive experiments allowed to determine the most significant variables that concur to determining fire danger: 1) date or season, 2) fuel moisture or inflammability of specific materials, 3) wind speed, 4) visibility range and 5) activity of fire-starting agencies.

While the Northern Rocky Mountain method targeted on a specific area and was tailored for easy and standardized calculation by the fire control agents, most of these variables, such as fuel moisture and wind speed, are still at the base of today's most widely used and advanced fire danger indices. Furthermore, the description of the role of each of these variables play in determining fire danger helps us trace how their measurement or estimation evolved from direct or indirect field measurements to today's methodologies based on weather and satellite data. To this end, it is worthwhile citing at least part of Gisborne's descriptions:

- **Date, or Season**: *"even though temperature, humidity, wind and fuel moisture may be the same in mid-June as in mid-July or even mid-August, the green vegetation such as grass, weeds, and brush is maturing, curing and becoming less of a fire retardant and more of a fire accelerator as the season progresses.* [...] *Even more consistent with calendar date is the number of hours of dangerous burning weather, according to hours of sunshine each day."*

- **Fuel Moisture**: *"this variable, which determines forest inflammability, is the second most important one considered in the Northern Rocky Mountain method.* [...] *The drier these fuels, the greater the danger, and in determining current danger it does not matter whether this dryness is controlled by precipitation alone, humidity alone, or any combination of precipitation, temperature, humidity, wind and sunshine."*
- **Wind**: *"Many experienced men believe that wind is one of the most important variables of fire danger. Cases can be cited of crown fires occurring with snow on the ground, and of blow-ups during high humidity.* [...] *Wind velocity, therefore, cannot be omitted from any complete scheme of rating fire danger.*

The 1940 fire control conferences called by the United States Department of Agriculture (USDA) Forest Service marked the beginning of the effort towards a standardized, national fire-danger rating system, and by 1954 there were already eight different rating systems used across the United States [37]. These efforts culminated in 1972, with the first version of the National Fire-Danger Rating System (NFDRS) was published by Deeming et al. and would go on to see two major revisions in the following years [38, 13]. The NFDRS capitalized on the studies on the physics of fuel moisture that culminated in the Rothermel fire spread model [110]. This resulted in a method based on three indices derived from the three fire behaviour components: Spread Component (SC), Energy Release Component (ERC) and Ignition Component (IC) with values from 0 to 100. The NFDRS distinguished a total of five fuel classes, three dead and two live, and incorporated weather data such as temperature, relative humidity, precipitation and wind speed. In addition to the method itself, the philosophical framework developed in this context would turn out to be very influential in defining the applicability scope of fire danger indices: a fire danger rating system should deal with the containment capabilities attributable to fire behaviour, and not with extinguishment, which should be dealt with using other means; provide relative, not absolute measurements; evaluate the worst conditions in a given area (i.e., be evaluated at the time of day at which fire danger is higher or on extreme exposures [37]).

At the time, the assessment of the condition of the vegetation was conducted directly *in situ* and required a large amount of manpower to cover large areas with sufficient frequency. Furthermore, this often resulted in fire danger indices being representative of the areas where samples where gathered, while they would be more and more inaccurate the farther one would get from the sampling point. The rise of satellite communications offered the ideal means of gathering up-to-date measurements in a synoptic way on a continental scale with unprecedented frequency. In particular, the deployment of the polar orbiting satellite Advanced Very High Resolution Radiometer (AVHRR) of the National Oceanic and Atmospheric Administration (NOAA), allowed the scientific community to develop and test the Normalized Difference Vegetation Index (NDVI) [127]. Today very widely used, the index is defined in order to be sensitive to the photosynthetic activity in the biomass present in a given pixel. In his seminal 1993 work [14], Burgan made use of historical NDVI measurements to define the Relative Greenness (RG) with the aim to monitor vegetation greenness in a given pixel based on recorded maxima and minima. Five years later [15], Burgan himself would integrate the RG into a new fire danger index named Fire Potential Index (FPI), which was among the very first indices to integrate both weather information and satellite data in the fire danger modelling process. In 2002, the FPI was introduced to the European continent [86] with a significant update on the estimation of the dead fuel load: rather than

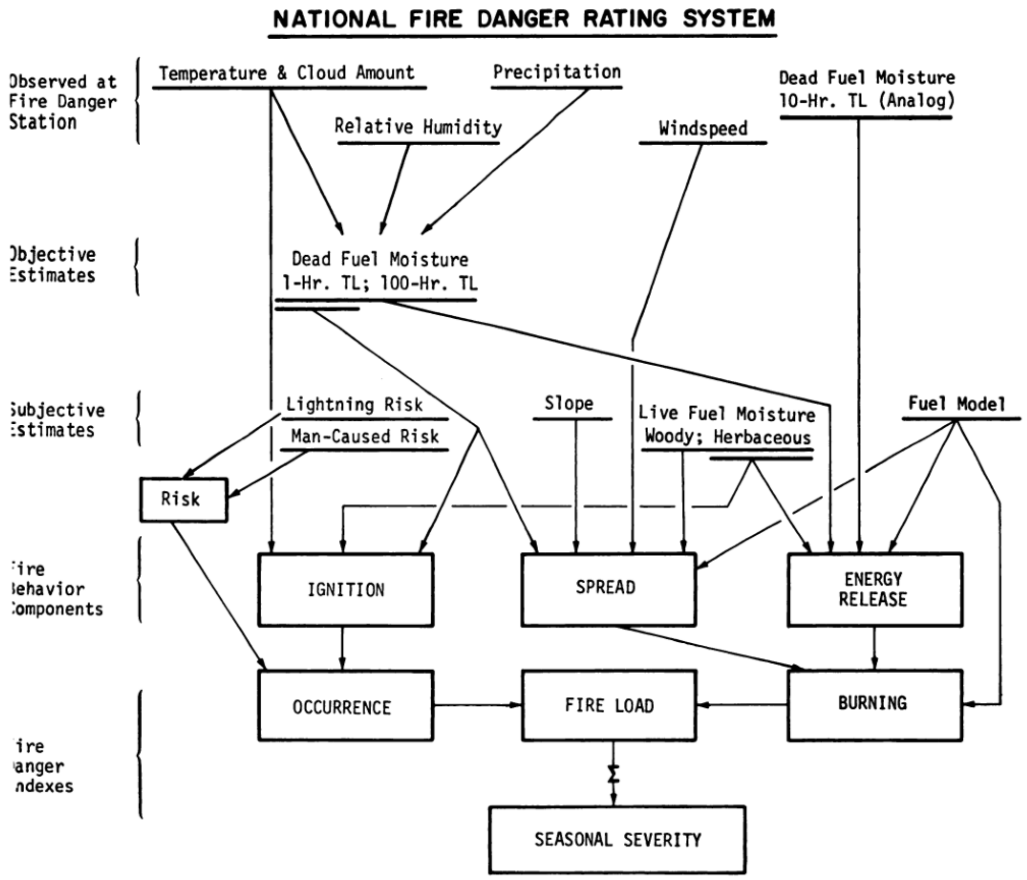

Figure 1.1. Structure of the National Fire Danger Rating System as shown in [37]. The scheme shows the three components of the Fire Behaviour model (ignition, spread and energy release) and all the weather variables involved in their calculation. The scheme also shows how the different components are integrated into a single seasonal severity indicator.

relying on the interpolation of the data gathered by a network of weather stations as in the US version, the European index used a complex algorithm to pick the best measurements from a 50 by 50 kilometre grid of weather stations.

Development of an FPI-derived index continued in the framework of the Space-based information support for the Prevention and Recovery of Forest Fires Emergency in the Mediterranean Area) (PREFER) project, which aimed to support the prevention, reaction and recovery phase of forest fire management through the use of space-based technology [82]. The FPI model was adapted to MODIS imagery, and corrections to the FPI variables were introduced using estimates of the vegetation water content through the equivalent water thickness and of the solar illumination conditions and wind speed through the evapotranspiration [81, 84]. Furthermore, weather forecast data was exploited in order to produce a fire danger forecast for the following days. This new index was called Daily Fire Hazard Index (DFHI) for its capability to provide updated fire danger forecasts on a daily basis.

The work described in this doctoral thesis has its roots in the experience of the development of the DFHI.

1.2.1 Weather-based Indices

In the 1940s, while the USDA was working on creating a national fire danger system, in the Soviet Union V.G. Nesterov was developing a simple algorithm to estimate fire danger using only weather data about air temperature, humidity and daily precipitation [95]. In the 1960s, A.G. McArthur developed the Australian Mk4 fire danger meter [91, 92], which consisted in a manually operated dial where the user would set the current values of the weather variables to obtain a fire danger rating. It would see an update to Mk5 in the following decade [61]. In the 1970s, the Canadian Forestry Service completed development of the Canadian Forest Fire Weather Index (CFFWI) [131], which simplified the calculation of the index by referring to a single fuel class, a standard pine species, and by requiring only weather data to be calculated. The ease of obtaining such data and the good performance of the index at large spatial scales would be at the base of the success of the CFFWI for the following decades. Many more indices were developed in the 1990s and in the early 2000s, such as the Indice de Risco Meteorologico (INMG) in Portugal [56], the Index of Risk [133] and the RISchio Incendi e COordinamento (RISICO) method [48] in Italy, the Instituto para la Conservación de la Naturaleza (ICONA) method [132] in Spain.

1.2.2 Probabilistic Approach

The increase in the availability of historical fire weather and fire occurrence data opened up the possibility to explore a more data-driven approach to the problem of fire danger. For this reason, a relatively recent line of research explores a probabilistic approach to the problem of fire danger assessment, and has already been applied to different spatial scales.

In some cases, the approach consists in building a logistic regression based on historical records of fire occurrence, weather variables, fire weather variables and fire danger indices [102, 2]. This approach has the advantage of providing a forecast of the number of wildfires and large wildfires, and also provides a lot of useful information on the predictive power of different variables on wildfire and large wildfire occurrence. For this reason, these analyses can also be used to assess the performance of different fire danger indices with respect to fire occurrence. Interestingly, [102] found that different variables were relevant to the probabilities of fire ignition and large wildfire occurrence, and in another instance, it was noted that different land cover types were more susceptible to fires in different stages of the fire season [2].

Another approach derives the fire occurrence probabilities from a high number of fire spread simulations [90, 47]: different fire behaviour models are used to run multiple simulations on the target area, allowing to identify fire activity hotspots. Compared to the earlier empirical approach, this simulation-based approach has the advantage of providing the possibility to investigate fire patterns and their causes under a wide range of different conditions. For instance, [90] analyses two different sets of weather conditions - "normal" and "extreme" - to find that the spatial distribution of higher danger areas at a fine scale is significantly different. In addition, [47] was able to incorporate the effects of fire suppression efforts into the model by introducing a probability of suppression. While the outputs of this approach are undoubtedly of great interest, their adoption in many areas of the world could be hindered by incomplete or scarce historical fire weather and fire occurrence records.

1.2.3 The Role of Satellite Remote Sensing in Fire Danger Assessment

An accurate estimation of fire danger is strongly dependent on the capacity to estimate the proneness to burn of live and dead vegetation. The relatively simple dynamics of dead fuel moisture content, which depend mostly on the vapour pressure deficit between the dead fuel and the environment, allow this estimation to be performed rather accurately using weather data exclusively.

Live fuels, on the other hand, pose a much greater modelling challenge due to their access to water resources that are not limited to the air humidity or surface soil moisture. Moreover, given that these fuels are in fact live plants, their mass changes along with the progression of their growth or curing state during the onset of different seasons. Furthermore, in order to properly assess fire danger, information on the amount of available fuel is also necessary, since if there is little fuel to for the fire to burn the danger will consequently be reduced.

Satellite data proved to offer a reliable means to assess both these issues since the early Nineties. The introduction of the NDVI and relative greenness in the FPI algorithm [14] allowed Burgan to provide an estimate of the amount of vegetation present in a given area and of the curing state of this vegetation relative to its history as recorded by the available data. Nevertheless, it is interesting to note that at the very beginning of the article he also stated "*we are not ready to provide live vegetation moisture content assessments*", implying that it should be included in the assessment of fire danger if it should become feasible at some point in the future. To this end, in the same years the development of Radiative Transfer Model (RTM)s specialized on leaf biochemistry such as PROSPECT [72] and canopy reflectance models such as Scattering by Arbitrarily Inclined Leaves (SAIL) [134] was starting to gain traction, even though despite the continuous improvement of the performance of computer systems, large scale application of such models to satellite images would remain inaccessible for years to come. For the next decade or so, satellite-based fire danger indices relied on the working assumption that an NDVI-based measured of greenness was an accurate enough representation of moist live vegetation [15]. However, the moisture content of live vegetation can undergo very significant variations even within the fire season, making any indirect estimate very unreliable.

The first attempts to estimate the moisture content of the live vegetation using satellite sensors usually relied on semi-empirical relationships between the observations, which in most cases were surface reflectances in visible channels and brightness temperatures or a combination of those into vegetation indices, and field data. However, the scarce amount of field data coupled with their availability on a restricted set of vegetation species, regions and climatic zones would impair the generalization of these methods on a continental or global scale. Other methods relied on the assumption that the curing state or the photosynthetic activity of the vegetation would be a reliable proxy for water content. However, this assumption holds true only for some species, since variations in chlorophyll content may also be related to other phenomena such as the phenological state of the plant, scarce availability of nutrients or poor air quality [21]. To this end, studies on leaf properties in the optical portion of the spectrum had already showed that the Shortwave Infrared (SWIR) region was the most sensitive to leaf water content [128, 53]. One of the first successful attempts to exploit satellite observations in the SWIR to directly estimate the moisture content of live fuels was Ceccato [22, 20], who developed the Global Vegetation Moisture Index (GVMI) for the SPOT satellite exploiting the sensitivity to vegetation water content of the SWIR portion of the spectrum. Using the PROSPECT model, Ceccato used the Near Infrared (NIR) and SWIR bands of SPOT-VEGETATION to construct a an index sensitive to the canopy water content per unit area as measured by the satellite. This index was introduced in Laneve's Modified Fire Potential Index (MFPI), which also accounted for the local solar illumination conditions using the vegetation evapotranspiration rate [83].

1.3 The Opportunity of Direct LFMC Estimation

The water content is just one of two variables that concur to determining the moisture content of the live vegetation. It is clear that the fuel moisture content of a leaf is determined not only by the amount of water it manages to retain in accordance with the local environmental conditions, but also by its maturing and curing stage, which is associated to changes in mass during its life cycle. Therefore, even if the water content does not change significantly over time, the simple progression of plant phenology may induce rather significant variations in live fuel moisture content, and consequently on fire danger. To this end, figure 1.2 contains some examples of timeseries where the measured Live Fuel Moisture Content (LFMC) is shown side by side with the corresponding Equivalent Water Thickness (EWT) and Dry Matter Content (DMC). The EWT represents the thickness that the water contained in the leaf would have if it were spread equally across the leaf area, and is therefore measured as a length, usually in centimetres. The DMC is defined as the leaf mass per unit area. This example is extremely clear in showing how a constant or even increasing EWT may sometimes counter-intuitively correspond to a decreasing LFMC or *vice versa* due to the fact that there is little correlation between the variation of leaf mass and its water content.

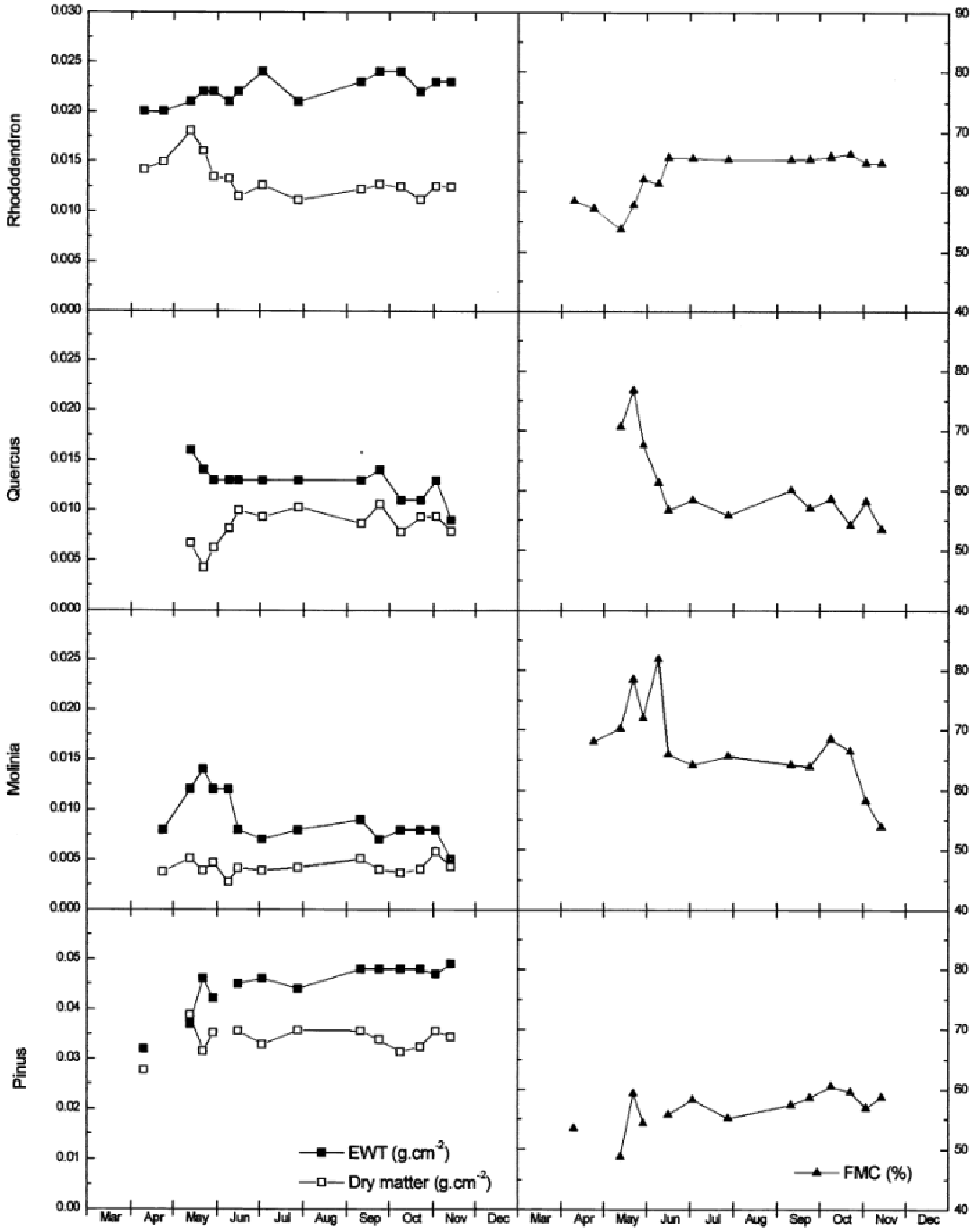

Figure 1.2. Variation of EWT (left graph, black squares), DMC (left graphs, white squares) and LFMC (right graphs, black triangles) as sampled by [57] and elaborated by [21] for the *Rhododendron ponticum*, *Quercus robur*, *Molina caerulea* and *Pinus sylvestris* species. The juxtaposition of these timeseries shows how knowledge of the EWT alone is insufficient for the purpose of inferring the LFMC progression for some species. This is particularly evident for the *Quercus robur*, where the initial drop in EWT is compensated by a sharper drop in DMC, resulting in an overall increase in LFMC. Subsequently, the EWT maintains a constant value, and the decrease in LFMC is only determined by an increase in DMC.

A complete estimation of the LFMC definitely requires the capability to estimate these variations in leaf mass using remotely sensed data. Providing such an estimate, which is also a very important variable of fire ignition and fire behaviour models, and integrating it directly in the process of fire danger assessment, is one of the main objectives of this thesis. Furthermore, classical FPI-derived fire danger indices do not directly include the impact of important variables such as the wind speed, the local slope and aspect and the number of sunlight hours into the fire danger modelling process. We aim to include these variables without altering drastically the FPI-like structure of the DFHI.

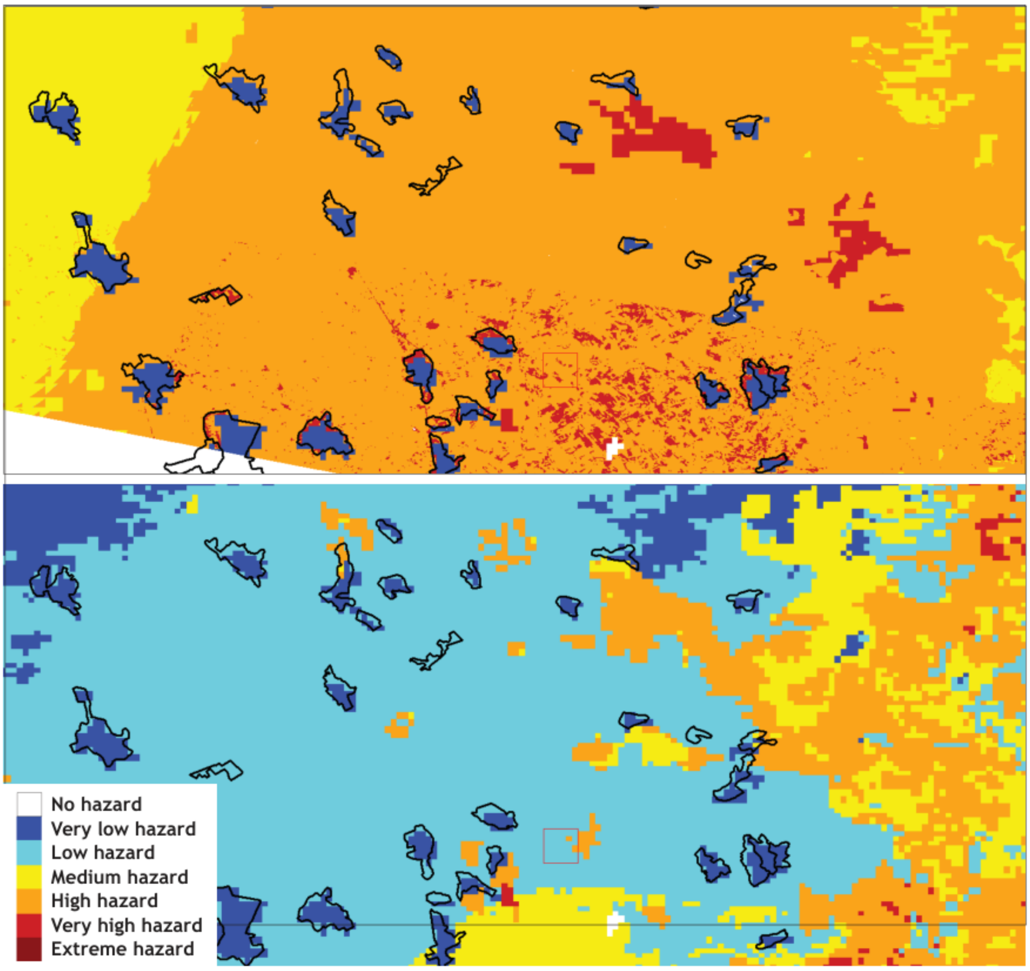

Figure 1.3. Detail of a DFHI map computed using a Landsat 8 image from 3 September 2016 (above) and a MODIS image acquired the same day (below). The MODIS based map was downscaled from 250 m to 30 m to match Landsat's resolution. Black polygons represent the area corresponding to Corine Land Cover class 1 (Artificial Surfaces). Re-elaborated from [81].

Finally, it has been noted [81] that the resolution of the satellite imagery used to estimate the state of the live vegetation may have a significant impact on the assessment of fire danger on a small scale. An example of the effect of satellite imagery on fire danger assessment is portrayed in figure 1.3. This may imply an overestimation of fire danger in some areas and an underestimation in some others.

To this end, we aim to exploit the Copernicus platform and Sentinel imagery to build an algorithm that is both image-agnostic (i.e., is not dependant on one specific satellite platform) and scalable in terms of resolution. We will make use of Sentinel-3 imagery and its 300-meter resolution to map fire danger at a regional, national or continental scale, and Sentinel-2 imagery and its 10-20-meter resolution for the local scale. Thus, we will have a way to provide danger indices at different scales while using the appropriate amount of resources for different purposes. While it may be attractive to use high-resolution imagery for the whole process, the amount of processing power required would increase drastically compared to the one needed to process moderate-resolution imagery, and at the same time the level of detail of such maps would make them extremely cumbersome to interpret on a national or even regional scale. On the other hand, only high-resolution imagery can provide a sufficient level of detail to properly describe fire danger in local high hazard or high value areas such as Wildland-Urban Interface (WUI)s or national parks and their surroundings.

Having the possibility to scale fire danger maps based on the size of the target could be a very powerful tool in the hand of the civil protection and fire management experts, especially if they could be exploited to plan prescribed burns and other prevention measures around high value protected areas, or to identify danger hotspots in WUIs and intervene appropriately.

1.4 Aims and Objectives

Summing up, the objective of this thesis is to build upon the foundation provided by the DFHI to define a novel fire danger index. This new fire danger index will be called Daily Fire Danger Index (DFDI), and will improve upon its predecessor and other danger indices by:

1. Exploiting satellite data and RTMs to estimate the fuel moisture content of live vegetation directly, rather than relying on water-sensitive vegetation indices or other proxies.

2. Incorporating the effect of wind speed on fire danger directly.

3. Incorporating the effect of evapotranspiration on fire danger directly. Since it is possible to model the effects of the local slope, aspect and daily hours of sunlight into evapotranspiration, we also aim to use evapotranspiration to indirectly incorporate their effects on fire danger.

4. Exploiting Sentinel imagery to provide a scalable-resolution product to fulfil the different needs of the decision makers.

In particular, the estimation of LFMC from satellite data requires multiple steps which will take up a significant portion of this work. For this reason, we will proceed as follows:

- Define the physical variables relevant to to fire danger assessment and identify those that will be included in the fire danger index; where necessary, describe in detail how each variable is calculated.

- Detail the entire procedure of LFMC estimation from satellite data: describe the radiative transfer models used for the purpose and how they are coupled; justify the use of different models for different land covers and describe the inversion scheme; perform a GSA of the models to identify the ideal satellite bands for LFMC inversion; attempt a validation of the inversion procedure on field data.

- Define the fire danger index formula and describe the normalization procedure for each involved variable; evaluate the performance of the index on the test area and compare it with the CFFWI.

1.5 Test Area

Sardinia is one of the 20 regions of Italy and is second largest island of the Mediterranean basin after Sicily. Its orography is mostly characterized by hills and low mountains, with the highest mountain peaking at 1834 meters above sea level, but at the same time it contains one of the most extended agricultural lands in the country [89]. While the climate is a typical Mediterranean temperate climate characterized by relatively mild winters and by dry and hot summers, the Sardinian case is particularly interesting for an intense water deficit that lasts from May to September [25]. Furthermore, Sardinia's wind profile tends to be skewed towards moderate-strong winds both in the Winter and in the Summer season. Almost half of the island is covered by pastures and cultivated lands, while shrubs add up to almost a third of the entire surface. Forests, with a prevalence of oaks, cover about 16% of the terrain [111].

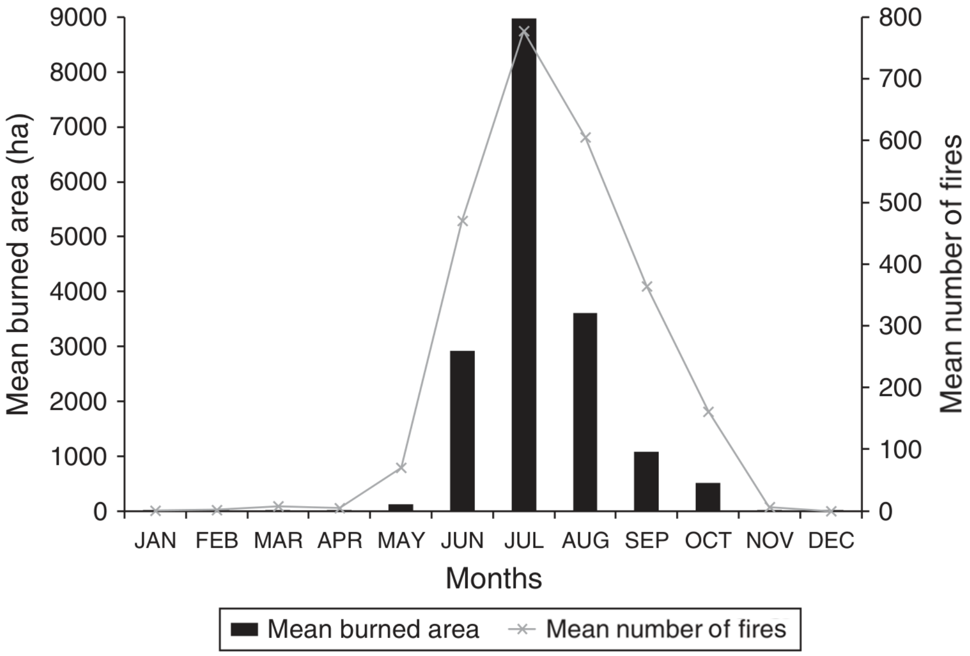

Figure 1.4. Mean wildfire occurrences by month (grey line with "x" shaped markers) and related burned surface area (black bars) for the 1995-2009 period in Sardinia according to the Sardinian Forest Service, from [111].

Sardinia is one of the Italian regions most affected by wildfires [98], and as shown in figure 1.4, most of the wildfires occur in the June-September window. In particular, Salis [111] notes that the worst fire seasons in terms of total burned surface were characterized by a concurrence of severe droughts, high temperatures, strong winds and large amounts of accumulated dead fuels. Additionally, Casu [19] states that the most damaging wildfires in this region are associated to mistral, libeccio and the sirocco winds, with the former being characterized by the highest average speeds. Indeed, Salis [112] reports that in the period 1998-2015, mistral winds were linked to 79% of the fires that burned areas larger than 100 hectares. As in other Mediterranean regions, the vast majority (90%) of fires in Sardinia are anthropogenic, roughly equally divided between accidental and intentional ignitions [111].

Sardinia was chosen as a test area not only for its climatic characteristics and for its diverse vegetation: the rich scientific literature in the framework of wildfire management and the abundance of historical wildfire occurrence and burned area data also contribute to make it an ideal candidate for the purpose.

Chapter 2

Methods

2.1 The Variables of Fire Danger

This chapter acts as a summary of the variables involved in fire danger modelling, some of which have been cursorily named during the literature review, in order to clarify their role and significance and to justify their inclusion in or exclusion from the danger modelling process. Subsequently, each variable will be discussed separately in order to provide information on the methods used to measure and estimate the values of these variables. In addition, we will provide information on the temporal and spatial scale used to sample these variables.

As done in [84], we will divide the variables in two categories based on the timescale of their variability: *dynamic variables*, whose values may change significantly on a daily, weekly of monthly basis, and *static variables*, which change on a timescale that is irrelevant to short-term fire danger prediction.

2.1.1 Dynamic Variables

This section is dedicated to dynamic variables relevant to fire danger estimation. Their values change on a relatively short timescale, and therefore a reliable fire danger assessment must be based on an accurate and recent estimation of these variables.

Weather Variables

Temperature, relative humidity, precipitation and wind speed all play a role in determining fire danger. It is clear that prolonged periods of high temperature, low relative humidity and scarce precipitation will increase the fire danger in the affected area, and it is also clear that, once ignited, the fire will spread faster and farther if the wind speed is high. Nevertheless, the vegetation, and especially the live vegetation, reacts to changes in the local environment on a timescale that varies from fuel type to fuel type and from species to species, and the inertia of the vegetation to these changes makes it so that the instantaneous values of these variables by themselves are not a reliable indicator of fire danger. The wind direction is worth a separate discussion: most fire danger experts will agree that wind direction also plays an important role in fire danger, especially at a local scale where historically very large or destructive fires may be associated to specific winds [112]. However, even at a regional level, the characteristics of a certain wind may change drastically depending on the origin of the wind itself[1], and therefore considerations about the effect of a specific wind direction may only apply to a very specific area. Precipitation is also sometimes included as a factor in fire danger modelling. Nevertheless, given the fact that precipitation is usually accompanied by a decrease in temperature and by an increase in relative humidity, its effect on the dead vegetation can be inferred from the changes in these two variables or can be included in the process of dead fuel moisture modelling itself, while its effect on the moisture of live vegetation can be estimated using remotely sensed reflectances.

Vegetation Greenness

Vegetation greenness is at the same time a measure of the amount of available vegetation and of its health. The role this variable plays in fire danger is dual, and strictly tied to that of the moisture content and curing state of the vegetation. A large amount of healthy, green vegetation will act as a fire retardant, even if dry dead fuel is present and may be prone to ignite fires. On the other hand, a large amount of stressed, dry vegetation may pose a great danger by acting as fuel and allowing the propagation of very large wildfires.

Dead Fuel Moisture

Dead fuels derive from dead trees, tree branches or leaves, shrubs or grass, and their moisture content can be estimated rather reliably knowing only the local environmental conditions. The moisture content of dead leaves depends only on the ever-changing equilibrium reached with the environmental conditions set by their surroundings, and differently than live leaves, they do not have any residual capacity to gather and store moisture. These fuels play a crucial role in determining the proneness to ignition of a given area, because they are much easier to ignite than live fuels, and their accumulation in large amounts is associated to faster ignition and propagation of wildfires.

Live Fuel Moisture

Differently than their dead counterparts, live fuels have access to a variety of mechanisms to procure and store water. Consequently, the estimation of their moisture content is strongly species-dependent, and more complex models are required to obtain such estimates through remotely sensed data. The entire Chapter 3 is dedicated to the description of this task.

Evapotranspiration

With the term evapotranspiration we refer to the combination of the processes of evaporation and transpiration, in which water is lost by the soil surface and by the vegetation respectively. Since these two processes happen simultaneously and have the same effect, there is no practical way of distinguishing them. For this reason, they are treated as a single phenomenon.

In particular, the reference evapotranspiration is related to a standard crop with pre-defined characteristics [3]. Intuitively, evapotranspiration depends on the topsoil water availability, but is mostly determined by the amount of solar radiation that

---

[1] For instance, in Sardinia the mistral wind is cool and moist in the Southern part of the region, where it comes from the sea, while it is hot and dry in the Northern part, where it has already crossed the mainland.

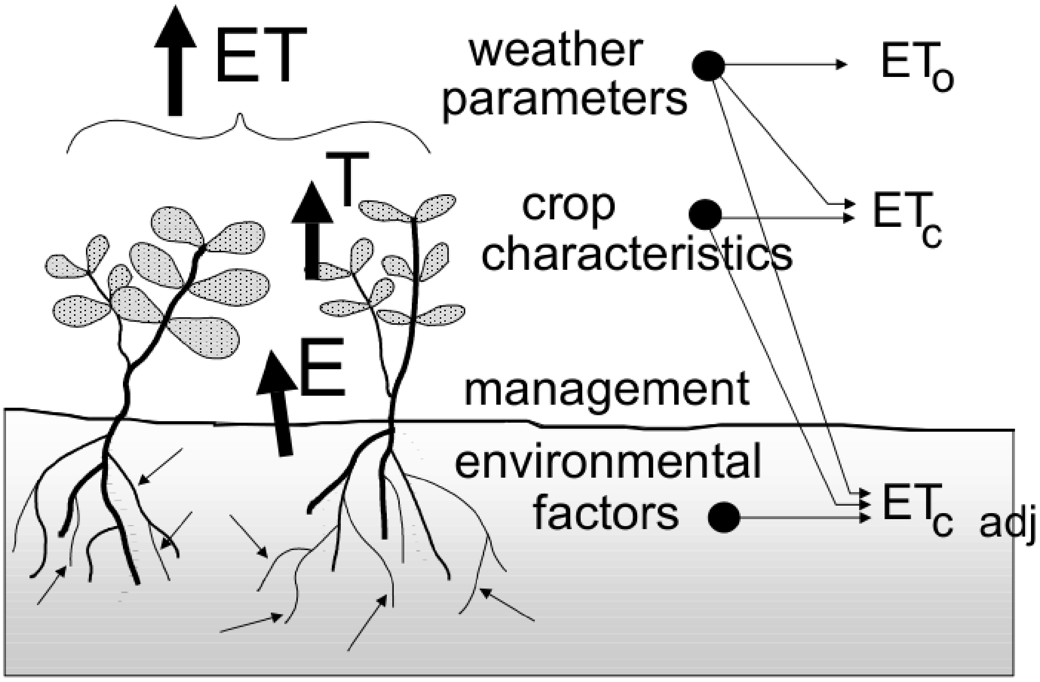

Figure 2.1. Factors that affect evapotranspiration according to the FAO Irrigation and Drainage Paper No. 56 [3]

reaches the vegetation. Therefore, a low vegetation cover will result in a higher evapotranspiration, because more radiation will reach the soil and more water will evaporate, while a lush vegetation will create more shade and induce less evaporation. The denser the vegetation, the more relevant the transpiration in comparison with the evaporation. Another factor that affects the amount of solar radiation that reaches the soil is the elevation and morphology of the terrain, which will create different illumination conditions. For this reason, the latest version of the DFHI [84] made use of the reference evapotranspiration to include, if indirectly, the effects of the local terrain morphology on fire danger.

Day of the Year, Date or Season

Since the beginning of the development of fire danger modelling, it was clear that, if indirectly, the date was a primary factor [55]. While on one hand, the date is strictly related to the progression of plant phenology, and therefore to its development stage, on the other it is intuitive to link the increase of the amount of hours of scorching weather in the beginning of summer to drops in fuel moisture content. While it would be easy to include this value in any type of index, it is difficult to generalize its meaning in different climatic areas and in different ecosystems. The capability to directly estimate live fuel moisture, and therefore to track the evolution of leaf dry matter in addition to leaf water content, should allow us to monitor these effects using remotely sensed data rather than pre-determined tables such as phenological calendars.

2.1.2 Static Variables

These variables change on a timescale that is not relevant to the problem of short-term fire danger assessment. Therefore, we will obtain such variables at the adequate resolution for our purposes and re-use them for each daily or hourly calculation, and update them whenever new versions are available.

Slope and aspect angle

The local slopes and aspect angle determine different microclimatic conditions for the vegetation by allowing different amounts of solar radiation and different winds to reach the surface, in turn controlling the amount of water available to the vegetation, and effectively allowing different types of vegetation to grow and prosper on either side of the slope. For instance, in the Northern hemisphere, it is well-known that south-facing slopes will receive more sunlight, making them more suitable to host sparse, sun-resistant vegetation, while north-facing slopes will be more humid and capable of growing trees and denser vegetation in general. The steeper the slope, the more intense these effects.

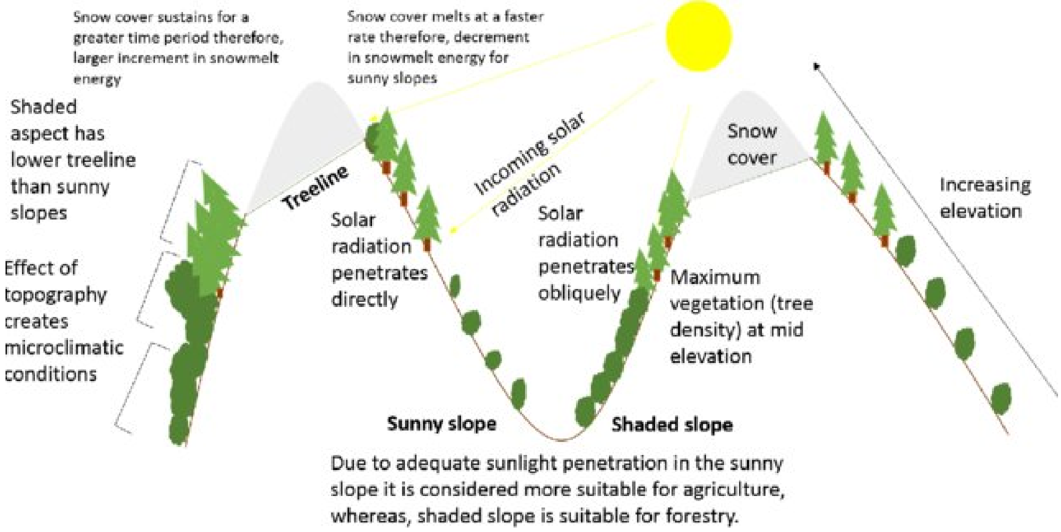

Figure 2.2. Schematic of how different slopes and slope orientations affect the illumination conditions and therefore concur to determining the vegetation type and density, from [121].

Fuel type map

Different fuels will behave differently in terms of their capability to retain moisture under different weather conditions, in terms of their seasonal growth, maturing and aging trends, and in terms of their proneness to burn. Naturally, fire-resistant tree species such as the cork oak will be associated to a lower fire danger than ones that produce highly flammable oils such as the blue gum eucalyptus. Furthermore, knowledge of the extinction moisture of different species can be useful to determine whether or not the local weather conditions allow a fire to ignite and spread at all.

2.2 Estimation of the Variables of Fire Danger

In this section we will describe in detail how each variable used in the fire danger modelling process is obtained or estimated.

2.2.1 Weather Variables and Evapotranspiration

For the purpose of compiling the historical archive of the fire danger index, the highest possible resolution historical weather data was necessary. Furthermore, the capability of providing an instantaneous snapshot of fire danger at different times of day requires hourly weather data. To this end, we decided to use ERA-5 Land hourly data reanalysis [93], which provides global weather datasets with a 0.10 degrees resolution from 1950 to today. This allowed us to obtain the following variables:

- 2m air temperature
- 2m dewpoint temperature
- 10m u-component of wind
- 10m v-component of wind

While the 2-meter air temperature can be used straight away for our purposes, minimal processing is required to obtain some of the other variables we need. For instance, ERA-5 Land does not contain the relative humidity directly, but allows us to obtain by including the dewpoint temperature. Furthermore, the wind speed is sampled at 10 meters height, while we will use wind speed at 2 meters height, i.e. the same height as the air temperature, which is more representative of the wind speed that affects the vegetation. We will now describe the two processes in detail.

Relative Humidity

The relative humidity is defined as the ratio between the amount of water vapour present in the air to the amount of water vapour that the air could contain if it saturated in the same conditions of pressure and temperature. The relative humidity is strictly related to the values of the actual air temperature and of the dewpoint temperature, which is the temperature at which water vapour starts to condensate into dew. For this reason, the relative humidity can be calculated as a ratio between the actual water vapour pressure e_a, i.e., the one determined by the dewpoint temperature, and the saturation water vapour pressure e_s at the current air temperature:

$$rh_{\%} = \frac{e_a}{e_s} \times 100 \tag{2.1}$$

An accurate formula that allows us to obtain the vapour pressure of water in function of its temperature is given by [124], who improved the formulations derived by [139]:

$$\ln e_w = -6096.9385T^{-1} + 21.2409642 - 2.711193 * 10^{-2}T + 1.67395210^{-5}T^2 + 2.433502\ln(T) \tag{2.2}$$

If T is the air temperature at 2 meters height and T_{dew} is the corresponding dewpoint temperature, applying the exponential function to both sides of the equation and substituting into equation 2.1:

$$rh_{\%} = \frac{\exp -6096.9385T_{dew}^{-1} + 21.2409642 - 2.711193 * 10^{-2}T_{dew} + 1.67395210^{-5}T_{dew}^2 + 2.433502\ln(T_{dew})}{\exp -6096.9385T^{-1} + 21.2409642 - 2.711193 * 10^{-2}T + 1.67395210^{-5}T^2 + 2.433502\ln(T)} \times 100 \tag{2.3}$$

Downscaling of Wind Speed

The value of wind speed changes with height, and in normal conditions it tends to decrease the closer to the surface it gets. For this reason, in meteorology wind speed is measured using instruments set at the standard height of 10 meters, which is also the height at which this parameter is provided by ERA-5 Land. Nevertheless, in some disciplines such as agrometeorology, experts are interested in wind speed values at a height of 2 or 3 meters, rather then meteorology's canonical 10. Therefore, in order to exploit the weather data for such purposes, a correction must be applied to the 10-meter wind speed value. To this end, we refer to the formula proposed in [3] to obtain the wind speed value used in the procedure for the calculation of the reference evapotranspiration:

$$u_2 = u_z \frac{4.87}{\ln 67.8z - 5.42} \tag{2.4}$$

where u_2 is the wind speed at 2 meters height and u_z is the input wind speed at a generic z meters height.

Evapotranspiration

In order to calculate daily values of the reference evapotranspiration, we follow the procedure described in the Food and Agriculture Organization (FAO) irrigation and drainage paper 56 [3], to which we refer the reader for an extensive and detailed explanation of each single variable and calculation step. We are only interested in the calculation of the reference evapotranspiration, denoted as ET_0, for hourly timescales using weather data. However, the aforementioned paper contains a number of different ways in which the variables involved in the process depending on the available data type, that may be either weather forecast data or measurements directly obtained from weather stations, and on the different types of variables that may or may not be available. Furthermore, in order to model the effect of the local topography on the illumination conditions, we used topographic information to calculate the actual hours of sunlight. Writing down the entire procedure in this chapter would be too long of a digression from the main topic of this thesis. However, in order to properly document the entire procedure and to guarantee the repeatability of the results, we will move the complete description to the Appendix, and limit the current discussion the calculation of the actual sunlight hours.

The FAO procedure [3] relies on the usage of a Campbell-Stokes sunshine recorder for direct measurement of the actual duration of sunshine. In our case, we neglect the effect of the clouds, but we can exploit knowledge of the topography based on a Digital Elevation Model (DEM) to determine whether a given pixel, at a given time of a given day of the year is illuminated or in shade. To do so, we follow the process shown in [42], which is based on equations developed by [52].

Let us consider a surface located at a latitude ϕ, when the local declination is δ. The angle between the normal to the surface and the horizontal plane (slope) is β, and the displacement between the horizontal projection of the normal from the local meridian (azimuth) is γ. If ω is the solar hour angle at the target time, the cosine of the angle between the incoming beam radiation and the normal to the surface θ is:

$$\begin{aligned} \cos\theta = {} & \sin\delta[\sin\phi\cos\beta - \cos\phi\sin\beta\cos\gamma] && (2.5)\\ & + \cos\delta[\cos\phi\cos\beta\cos\omega + \sin\phi\sin\beta\cos\gamma\cos\omega && (2.6)\\ & + \sin\beta\sin\gamma\sin\omega] && (2.7) \end{aligned}$$

While the slope angle β can be directly obtained using GDAL's `gdaldem` function, some work must be done to obtain the γ angle. `gdaldem` outputs the aspect angle, which has the same significance as γ, but is 0 when the surface is facing North and is defined between 0 and 360 degrees. To obtain γ from the aspect angle we need only subtract 180 degrees:

$$\gamma = aspect - 180^\circ \tag{2.8}$$

A scheme of the illumination geometry of a south-facing surface (i.e. $\gamma = 0$) with a β sloped is represented in 2.3.

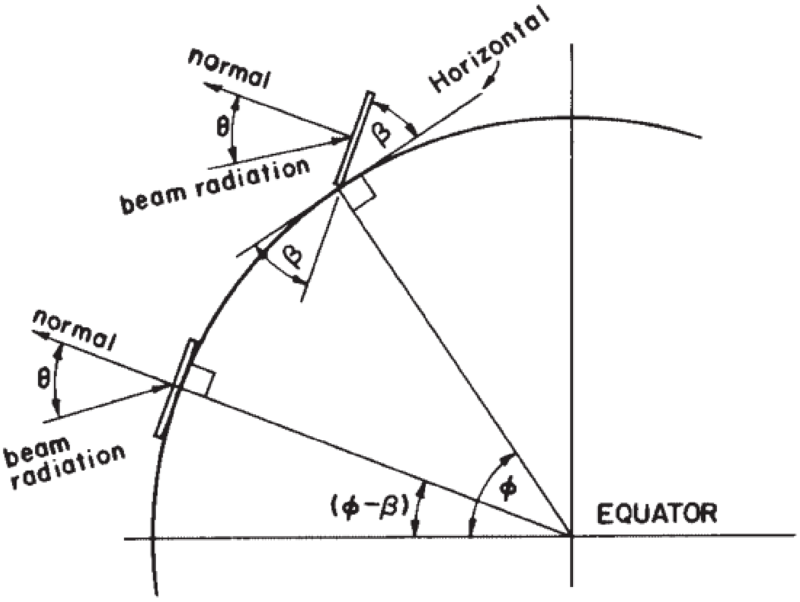

Figure 2.3. Scheme of the geometry of the incoming radiation on an inclined surface for a south-facing slope (i.e. $\gamma = 0$), adapted from [42]. β represents the slope, ϕ represents the latitude and θ is the angle between the normal to the local surface and the incoming beam radiation.

The complementary angle to the beam incidence angle θ represents the solar elevation with respect to the surface. We are more interested in this angle than in θ itself, because when it is positive and sufficiently large we can assume that the surface is directly illuminated by the Sun rays. Let the solar elevation angle be ϵ. We can calculate it as:

$$\begin{aligned}\epsilon = 180^\circ - \theta = 180^\circ - \arccos(\sin\delta[\sin\phi\cos\beta - \cos\phi\sin\beta\cos\gamma] \quad (2.9)\\ + \cos\delta[\cos\phi\cos\beta\cos\omega + \sin\phi\sin\beta\cos\gamma\cos\omega \quad (2.10)\\ + \sin\beta\sin\gamma\sin\omega]) \quad (2.11)\end{aligned}$$

For each hour of potential sunlight N, and therefore for each solar hour angle ω between the sunset hour angle ω_n and the sunrise sunset angle $\omega_n - \frac{\pi}{N}$, we will check if this value is higher than 5° for each individual pixel and its topography. If it is, we will add 1 hour to n, otherwise we will not. The resulting value will give us the approximate total amount of sunlight hours accounted for shade caused by the local topography.

As specified in the description of each parameter, all meteorological variables should be either measured or converted at 2 meters height to maintain uniformity, and in general the obtained values refer to an extensive surface of green grass that covers the soil which is not suffering from water scarcity conditions.

In our case, the reference evapotranspiration is calculated for all the pixels in the area of interest, using the latest available meteorological data at 2 m height and in particular the temperature and wind speed. Even though the vegetation cover of the individual pixel may not correspond at all to the description of the reference plot defined in [3], the evapotranspiration value is still considered representative of the different evaporation and transpiration conditions associated to that area: a pixel hit by an higher amount of radiation, higher temperatures and higher wind speeds will necessarily be more stressed than a pixel characterized by lower mean temperatures and by a topography that causes it to be more hours in shade. An example of the evapotranspiration obtained using this process is shown in figure 2.4.

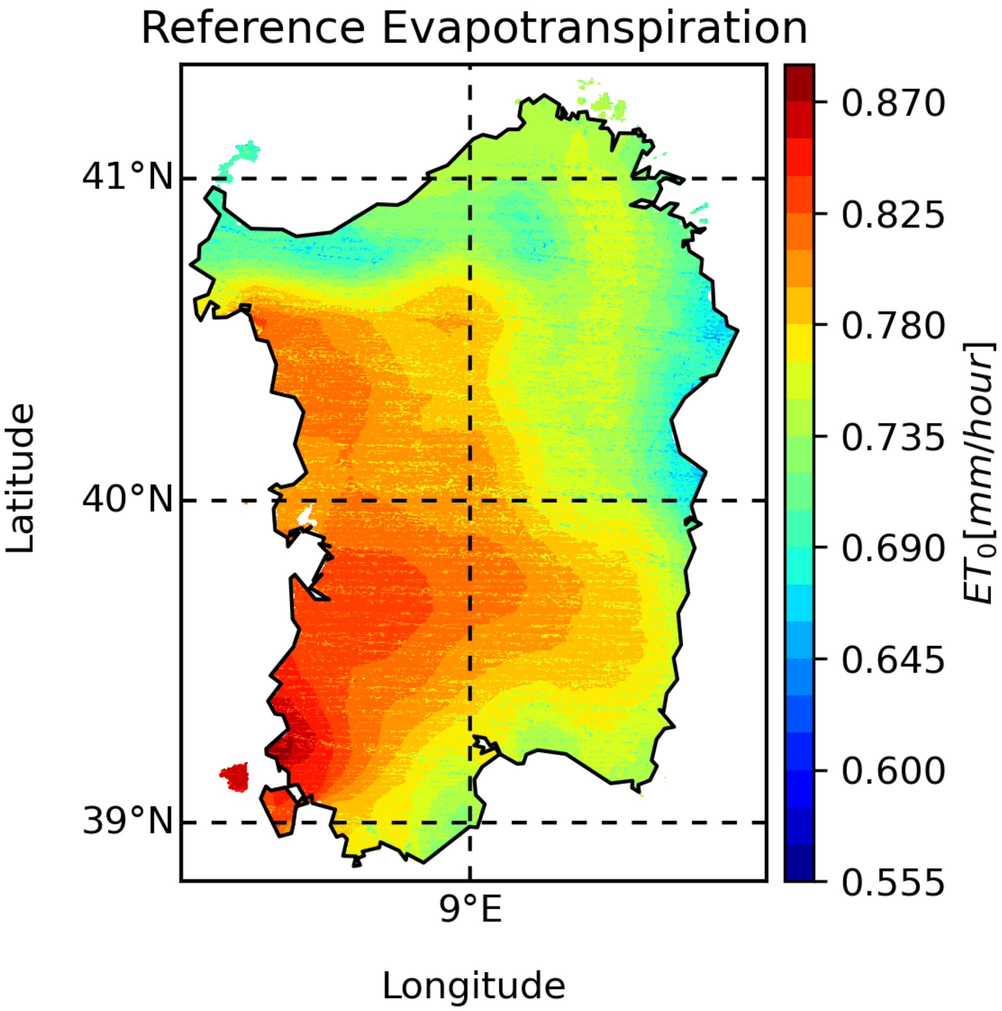

Figure 2.4. Reference evapotranspiration in Sardinia at 12:00 of the 21st of June 2022.

2.2.2 Fraction of Live Vegetation

Satellite imagery has been used to estimate the greenness of the vegetation for the purposes of fire danger modelling at least since [14]. Exploitation of long-term (i.e., years-long) records of indices sensitive to vegetation greenness allows to compare the latest acquisitions with the accumulated historical knowledge, and therefore to provide a relative measure of the health or stress level of a specific pixel or area. To this end, the most widespread vegetation greenness index is undoubtedly the Normalized Difference Vegetation Index (NDVI), which was first calculated on a large spatial and temporal scale for the Advanced Very High Resolution Radiometer (AVHRR) in [127]. As indicated by its name, the index is calculated by normalizing the difference of the red and near infrared surface reflectances captured by an optical sensor:

$$NDVI = \frac{NIR - RED}{NIR + RED} \tag{2.12}$$

It is well-known that the interaction of leaves with the incoming optical radiation is mostly regulated by the process of photosynthesis [129]. Healthy leaves that are actively photosynthesizing carbohydrates by absorbing incoming light and using it to combine the oxygen in the atmosphere and the water in the mesophyll will consequently have a low reflectance in the so-called Photosynthetic Active Region (PAR), and at the same time they will reflect more radiation in the near infrared to avoid overheating. On the other hand, stressed or unhealthy vegetation with low or non-existent photosynthetic activity will absorb less radiation in the PAR (hence, reflect more) and reflect less radiation in the near infrared. For this reason, positive and high NDVI values are associated to healthy and dense vegetation, while lower positive values are associated to unhealthy vegetation or even bare soil.

In our case, we used Sentinel-2 L2A and Sentinel-3 SYNERGY imagery to obtain the timeseries of the index on the area of interest. Table 2.1 contains the list of the bands we used to calculate the NDVI from the two Sentinel products. The historical Sentinel-3 NDVI maxima and minima in Sardinia are shown in figure 2.6 and 2.5 respectively.

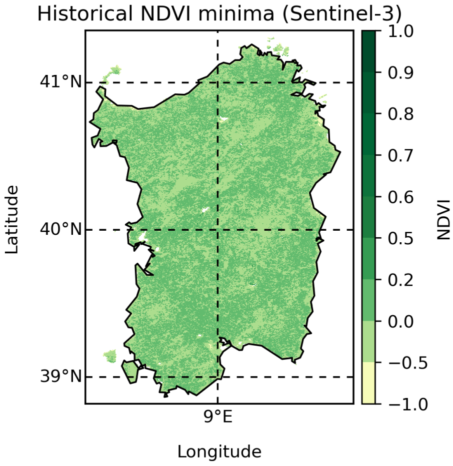

Figure 2.5. Sentinel-3 historical NDVI minima in Sardinia

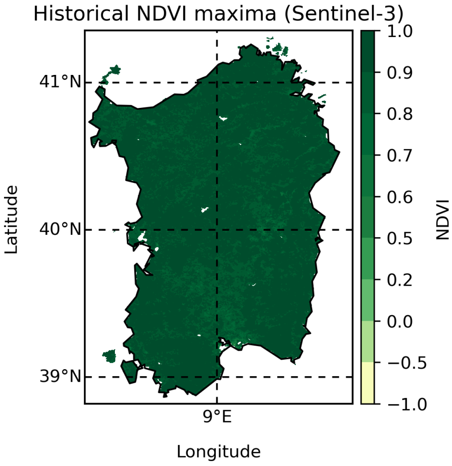

Figure 2.6. Sentinel-3 historical NDVI maxima in Sardinia

Product	Red Band	NIR Band
Sentinel 2 L2A	Band 04	Band 08
Sentinel 3 SYNERGY	SYN18/S2	SYN19/S3

Table 2.1. Satellite image bands used to compute the NDVI.

In particular, the NDVI values were stored and used to build archives for the entire Area of Interest (AOI). Building such archives may require a different orders of magnitude in terms of processing loads depending on the image resolution and the size of the AOI, but platforms such as Google Earth Engine allow to drastically reduce this workload even on high-resolution imagery such as Sentinel-2. The NDVI archives are then processed to obtain the historical minima and maxima for each pixel, and these values are stored in two dedicated rasters. For each new observation (i.e., new satellite image), the current NDVI values are used to calculate the Relative Grenness in the original formulation described by [14]:

$$RG = \frac{NDVI - minNDVI}{maxNDVI - minNDVI} \cdot 100 \tag{2.13}$$

where $minNDVI$ and $maxNDVI$ are the minimum and maximum NDVI values registered in a given pixel in the historical archive. For the purpose of the calculation of a fire danger index, we prefer to use metrics scaled between 0 and 1, and thus we define the relative greenness fraction RG_f as follows:

$$RG_f = \frac{NDVI - minNDVI}{maxNDVI - minNDVI} \tag{2.14}$$

Needless to say, with each new observation the historical maxima and minima NDVI archives are updated, so that the resulting RG_f values are always between 0 and 1. Low RG values will correspond to an higher amount of stressed or aging vegetation compared to the history of that specific pixel. Conversely, high RG values will be associated to healthy and lush vegetation, again, relative to that pixel's historical record. It must be stressed that the RG is a relative measurement of grenness, and therefore looking at this quantity by itself may bring the user to draw wrong conclusions by comparing the values of different pixels. A high RG value in a pixel may correspond to very little vegetation presence, but to a good health of such vegetation. On the other hand, low RG values in another pixel may correspond to a dense forest suffering from diseases or other heat or moisture induced stresses. Therefore, in the framework of fire danger assessment, low RG values are associated to a relatively higher fire danger, because the vegetation in the affected pixels may be stressed or unhealthy and consequently more prone to catching and spreading fire. Nevertheless, fire danger is also higher if the amount of stressed vegetation is larger. For this reason, in order to scale this quantity with an absolute measure of the vegetation presence, we use the historical $maxNDVI$: the higher its value, the stronger the vegetation presence. This value is used to rescale RG_f as follows:

$$L_f = RG_f \cdot \left(\frac{1 + maxNDVI}{2} \right) \tag{2.15}$$

where L_f is the Live Vegetation Fraction [84], which has all the characteristics we need from a vegetation greenness index: high L_f values are associated to healthy and dense vegetation, while low values are associated either with a very low relative greenness, i.e. with a high amount of stressed vegetation compared to the historical records, with a low NDVI maximum value, i.e. with a low historical vegetation presence, or with a combination of the two.

2.2.3 Dead Fuel Moisture

In order to obtain realistic values for the moisture content of dead fuels, we refer to the procedure developed in the context of the National Fire-Danger Rating System (NFDRS) by Deeming et al. [38]. These handbooks contain detailed procedures to derive the fire danger indices of the NDFRS, but we are only interested in the parts that concern the estimation of dead fuel moisture content based on the weather variables of temperature and relative humidity. The description of the moisture exchange mechanics between dead fuels and the surrounding environment is based on the research work documented in [50]. According to this approach, the behaviour of dead fuels can be quantified in terms of their Equilibrium Moisture Content (EMC).

Equilibrium Moisture Content

The EMC represents the amount of moisture content held by dead woody material that has reached a steady state. EMC values can be determined using the so called Dry Bulb Temperature (DBT)[2] and relative humidity using the equations developed by Simard [120] on the basis of the tables that can be found in the Wood Handbook [130]. All temperatures are hereby expressed in degrees Fahrenheit, relative humidity is expressed as a percentage, and EMC is expressed as percent moisture content.

$$\begin{cases} EMC = 0.03229 + 0.2981073 * rh - 0.000578 * T_f * rh, & \text{if } rh \leq 10\% \\ EMC = 2.22749 + 0.160107 * rh - 0.014784 * T_f, & \text{if } 10 \leq rh < 50\% \\ EMC = 21.0606 + 0.005565 * rh^2 - 0.00035 * rh * T_f - 0.483199 * rh, & \text{if } rh \geq 50\% \end{cases} \tag{2.16}$$

We can use these equations to evaluate the EMC for the observation time or to compute the minimum and maximum daily EMC, that we will refer to with the following variable names:

$$\begin{cases} EMCOBS = f(TMPOBS, RHOBS) \\ EMCMIN = f(TMPMAX, RHMIN) \\ EMCMAX = f(TMPMIN, RHMAX) \end{cases} \tag{2.17}$$

Where:

- $\_OBS$ refers to the afternoon observation time values
- $\_MIN$ refers to the 24-hour minimum observed value
- $\_MAX$ refers to the 24-hour maximum observed value

Since the discovery that the moisture content behaviour of dead fuels follows an exponential function, i.e. their moisture content adapts to the vapour pressure gradient that exists with the surrounding environment following an exponential function, dead fuels are usually categorized into timelag classes. The concept of timelag is related to the time necessary to lose $1 - 1/e$ (i.e. about two thirds) of their original moisture content [10]. Thus, the NFDRS defines 1-, 10-, 100-, and 1000-hour timelag classes based on the diameter or characteristic dimension of the fuels as shown in table 2.2.

---

[2] The DBT is the temperature of the air measured by a thermometer freely exposed to the air, but shielded from radiation and moisture. DBT is the temperature that is usually thought as air temperature.

Timelag Class	Diameter Range
1-h	up to 1/4 in
10-h	1/4 to 1 in
100-h	1 to 3 in
1000-h	3 to 8 in

Table 2.2. Timelag classes based on fuel size according to the NFDRS

An example of a 1000-hour timelag fuel moisture map generated by the US Rocky Mountain Reasearch Station with the procedures detailed in [10] is portrayed in figure 2.7. 1000-hour timelag fuels are the most complex to model, and since their behaviour is more similar to that of live fuels, they are often used as a proxy for LFMC in fire danger modelling processes. Since in our case we rely on a direct estimation of LFMC, we will only focus on 10-hour timelag classes, which are the most important in the processes of fire ignition and fire spread.

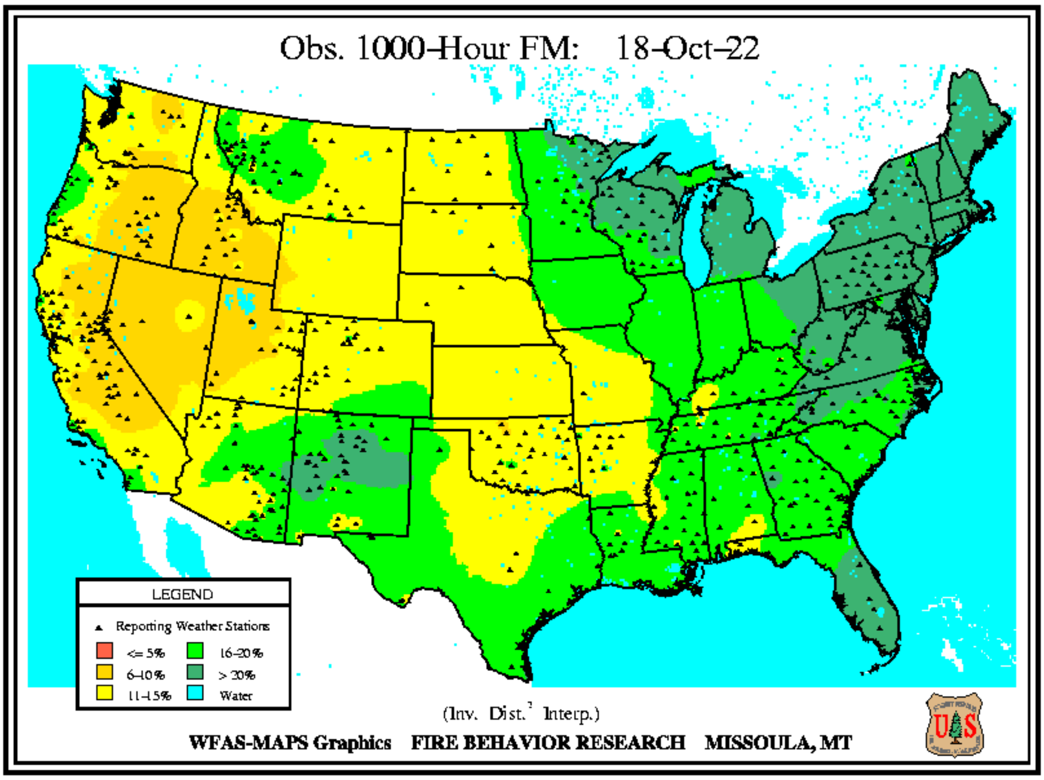

Figure 2.7. Daily NFDRS 1000-hour timelag fuel moisture for the United States as downloaded from the NFDRS website on October 18th 2022.

1-Hour Timelag Fuel Moisture

1-hour timelag fuels respond so rapidly to changes in the environmental conditions that the only variable required to estimate their moisture content is the potential moisture content. The field approach relies on fuel stick data or estimates of the relative humidity and dry bulb temperature of the air in immediate contact with the fuel elements. These values are then corrected at instrument height according to the insolation conditions [16] using tabled values. In our case, we can simply use the air temperature and relative humidity at two meters height to compute the EMC using equation 2.16, and then apply the following correction to derive the 1-hour timelag fuel moisture:

$$MC_1 = 1.03 * EMCPRM \tag{2.18}$$

If it is raining at the observation time we simply set the following fixed value for the MC1:

$$MC_1 = 35.0 \tag{2.19}$$

10-Hour Timelag Fuel Moisture

Field methods that rely on fuel sticks use an empirical method to account for the weight loss that sticks undergo as they weather. In our case, we can simply calculate the EMCPRM as in 2.16 and correct as follows to obtain the 10-Hour Timelag fuel moisture:

$$MC_{10} = 1.28 * EMCPRM \tag{2.20}$$

Fraction of 10-Hour Timelag Fuel Moisture

10-Hour Timelag Fuel Moisture can be normalized using a fuel map and the corresponding moisture of extinction of a given fuel type. The moisture of extinction, expressed as a mass percentage, represents the minimum amount of moisture that the associated fuel must contain in order not to ignite. Therefore, fuels with a higher extinction moisture will be more prone to burning than fuel with a lower extinction moisture, because the latter will be able to prevent ignition even in drier conditions. This normalized quantity is called Fraction of 10-Hour Timelag Fuel Moisture, and is referred to as TN_f:

$$TN_f = \frac{MC_{10} - 2}{MX - 2} \tag{2.21}$$

In order to avoid obtaining values higher than unity, when the 10-Hour Timelag fuel moisture is higher than the corresponding extinction moisture value we reset the fractional value to one, assuming that the contribution to fire danger is minimal whenever this value is reached and does not change significantly for higher values.

2.2.4 Live Fuel Moisture

As previously mentioned, the estimation of the fuel moisture content of live fuels is one of the major objectives of this thesis. Given the extensive modelling effort required to obtain such estimates, the entire Chapter 3 is dedicated to this issue.

2.2.5 Static Variables

The high-resolution digital elevation model used in the algorithm was obtained from the Sardinian administration web portal for geographic information. In order to compute the slope and aspect values from the DEM, we used the Geospatial Data Abstraction Library (GDAL) [137] function `gdaldem`. The entire GDAL library is free software and open source, and in particular, the `gdaldem` function is documented in the Geographic Resources Analysis Support System (GRASS) manual [23]. In particular, to obtain both the slope and the aspect rasters, we used the Horn function, which according to the manual is most suited to rough terrain. The DEM used for the AOI is shown in figure 2.8, while the derived slope and aspect are shown in figures 2.9 and 2.10 respectively.

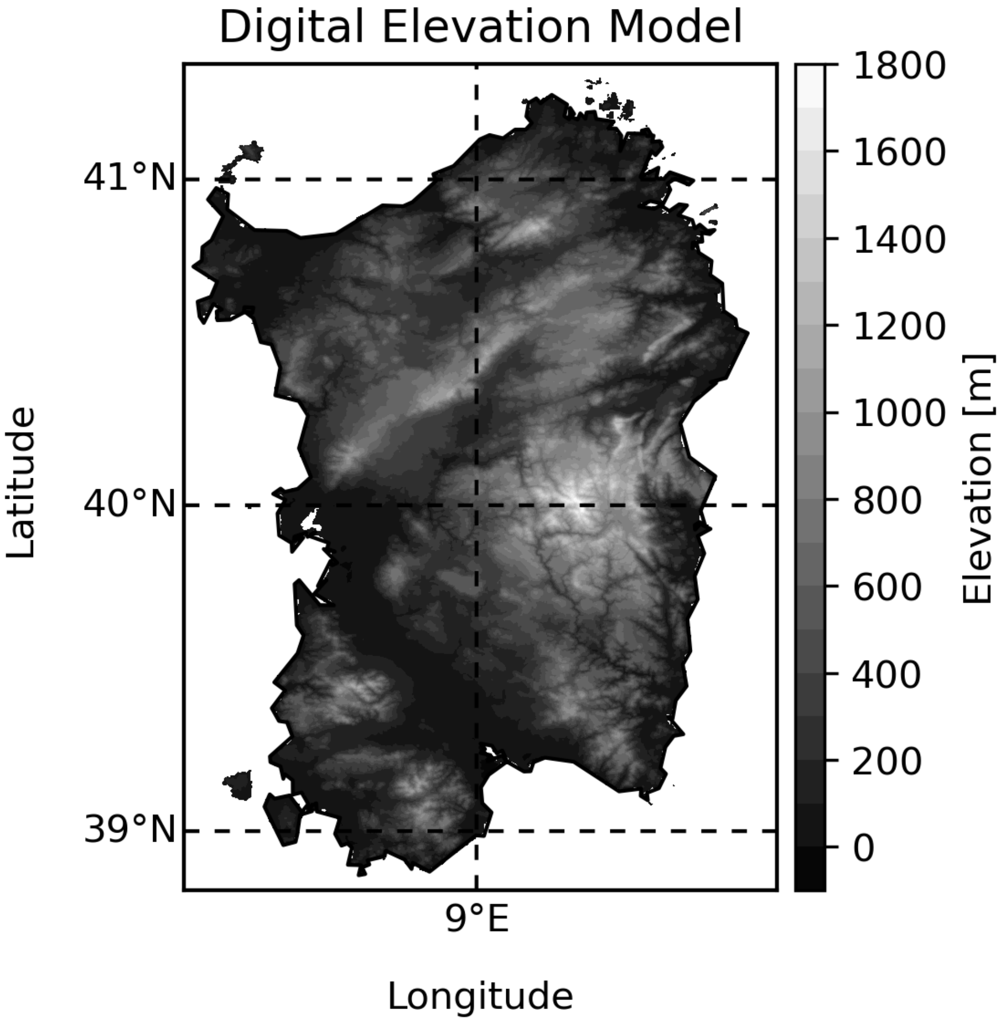

Figure 2.8. High-resolution Digital Elevation Model of Sardinia, provided by the Sardinian administration. Lighter pixels represent higher elevations.

The 100-meter resolution Corine Land Cover Map (CLC) 2018 [32] and the 10-meter resolution ESA Worldcover [146] were used for the moderate-resolution product based on Sentinel-3 images and for the high-resolution product based on Sentinel-2 images respectively. In both cases, the original land cover classes were remapped into three basic classes: grassland, shrubland and forest. Similarly as [105], different classes corresponds to different LFMC inversion methods, which will be detailed in Chapter 3. The land cover reclassification schemes are shown in table 2.3.

Finally, in addition to a land cover map, a fuel type map was necessary to obtain information about the characteristics of the vegetation with respect to fire behaviour. In particular, the extinction moisture of each fuel is used to evaluate the fire danger of the dead fuel classes by comparison with their fuel moisture content. To this end,

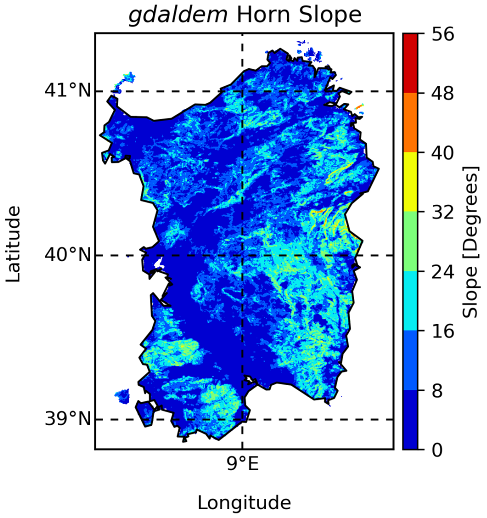

Figure 2.9. Slope inclination map obtained using `gdaldem` and the Horn algorithm. The raster was upscaled to match the Sentinel-3 resolution using a "mean" algorithm.

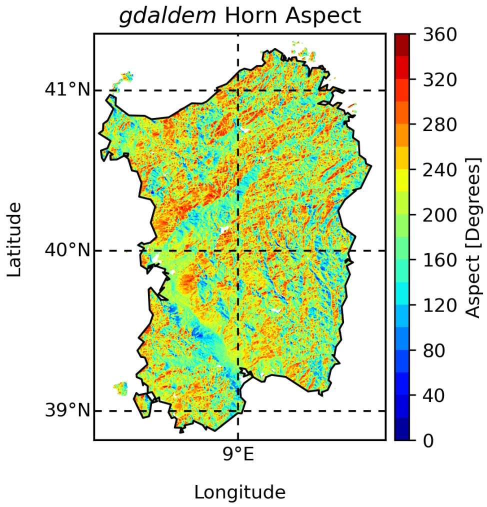

Figure 2.10. Slope orientation or aspect map obtained using `gdaldem` and the Horn algorithm. The raster was upscaled to match the Sentinel-3 resolution using a "mean" algorithm.

CLC Map Code	CLC Class Description	Target Class
211	Non-irrigated arable land	Grassland
212	Permanently irrigated land	Grassland
213	Rice fields	Grassland
221	Vineyards	Forest
222	Fruit trees and berry plantations	Forest
223	Olive Groves	Forest
231	Pastures	Grassland
241	Annual crops associated with permanent crops	Grassland
242	Complex cultivation patterns	Grassland
243	Land principally occupied by agriculture...	Forest
244	Agro-forestry areas	Forest
311	Broad-leaved forest	Forest
312	Coniferous forest	Forest
313	Mixed forest	Forest
321	Natural grassland	Grassland
322	Moors and heathland	Shrubland
323	Sclerophyllous vegetation	Shrubland
324	Transitional woodland scrub	Shrubland
ESA WC Map Code	**ESA WC Class Description**	**Target Class**
10	Tree Cover	Forest
20	Shrubland	Shrubland
30	Grassland	Grassland
40	Cropland	Grassland

Table 2.3. Remapping scheme of CLC and ESA Worldcover classes into our three target classes of grassland, shrubland and forest. If a map code is not included in the table it is considered non-vegetated and ignored.

we exploited the European fuel map provided by the Joint Research Center (JRC) in the European Forest Fire Information System (EFFIS) project [43].

2.2.6 Summary of the Data and its Providers

In this section we provide a summary of the variables used to generate the variables and fire danger maps shown throughout the thesis. The first table 2.4 describes the weather variables and the static fuel map and digital elevation model, while the second table 2.5 contains the references to the specific products of the exploited satellite imagery.

Variable	Provider	Data Type	Spatial Resolution	Timeliness
Air Temperature @2m	ECMWF	Raster	0.10°	Daily
Dewpoint Temperature @2m	ECMWF	Raster	0.10°	Daily
Wind Speed u-component @10m	ECMWF	Raster	0.10°	Daily
Wind Speed v-component @10m	ECMWF	Raster	0.10°	Daily
Corine Land Cover 2018	Copernicus	Raster	$100m$	-
ESA WorldCover 2021 v200	ESA	Raster	$10m$	-
Fuel Map	JRC	Shapefile	-	-
Copernicus EU DEM 1.1	Copernicus	Raster	$25m$	-

Table 2.4. Summary table of the providers of each variable used in the fire danger modelling process. The resolution and format of the data are also included. If a link to the dataset is available, it is provided as a hyperlink on the provider name.

Product	Provider	Processing Level	Spatial Resolution	Revisit Time
Sentinel-3 SYNERGY	ESA	Level-2	$300m$	$<$ 2 days
Sentinel-2 L2A	ESA	Level-2A	$10 - 20 - 60m$	5 days

Table 2.5. Summary table of the satellite image products used in the fire danger modelling process. The product guides are provided as an hyperlink on the provider name.

Chapter 3

Live Fuel Moisture Content Estimation using Sentinel Imagery

LFMC directly affects the flammability of the vegetation, and for this reason it is one of the most important physical variables in fire behaviour and fire ignition models [107, 12]. The interaction of LFMC with fire determines the time to ignition, because the fire has to spend energy evaporating the water before igniting the fuel. Thus, ignition of fuels with a higher moisture content will require stronger fires, and in general moister live fuels will behave more like fire retardants than like fuels with respect to fire spread. Laboratory studies of fire ignition and behaviour have shown that the importance of LFMC in fire behaviour may change depending on conditions such as the flame temperature [49] or the considered vegetation species [149]. However, [105] notes that when shifting the analysis from the laboratory to a large spatial and temporal scale, the LFMC undoubtedly becomes a primary factor, citing a number of studies dedicated to historical fire analysis in the United States [117, 40] and Mediterranean areas [27, 77]. Consequently, the capability of estimating it reliably is fundamental for fire danger assessment [26].

In the scientific literature, LFMC is generally defined using the terminology associated to field sampling: if FW is the leaf sample fresh weight and DW is the weight of the leaf matter after oven-drying, the LFMC can be calculated as follows:

$$LFMC_{\%} = \frac{FW - DW}{DW} \times 100 \tag{3.1}$$

Usually, in field sampling campaigns, freshly cut leaf samples are weighted before being placed in special ovens at a temperature higher than 100 degrees for a number of hours. After drying in the oven, the samples are then re-weighed, and the weight difference between the two measurements corresponds to the mass of the water contained in the sample. This value is finally divided by the dry mass, providing the Live Fuel Moisture Content value. Unfortunately, only few institutions have the resources and know-how that allows them to provide LFMC field measurements with daily or weekly frequency, and in general the availability is of such measurements is concentrated in the Mediterranean countries, United States, and Australia [145]. Even where these measurements are available, it remains difficult to extrapolate the samplings away from the field site with the objective to provide fire danger maps, as noted by [55]. Therefore, the capability to exploit satellite remote sensing for this purpose is extremely attractive for any scale larger than the local. In particular, satellite platforms such as MODIS, VIIRS and Sentinel-3 provide the attractive possibility to map fuel moisture content with a high temporal resolution (daily) and at a moderate spatial resolution ($250 - 500\ m$).

However, obtaining a reliable estimate of LFMC requires a much stronger scientific and computational effort than its dead counterpart, which can be estimated using weather-based variables and indices [105]. The approaches towards LFMC estimation from remotely sensed data can be grouped in two main classes: empirical-statistical models, which rely on fitting timeseries of LFMC field measurements with remotely sensed data from a satellite or aerial platform, and RTM-inversion models [144, 97]. Empirical-statistical methods exploit the availability of field samples to link the measured values to concurrent satellite acquisitions of surface reflectances and the associated vegetation indices [125, 17, 39], sometimes also including weather data [51, 26]. The resulting regression functions are then used to estimate LFMC values in other areas and times. However, their applicability is usually limited to the targeted satellite platform and their accuracy in areas other than the field sampling sites is significantly decreased: using field measurements from a limited number of field sites makes the dataset inherently incapable of reproducing the vast amount of different viewing and illumination conditions, vegetation species and soil characteristics [144].

RTM-inversion methods, on the other hand, have the advantage of a physical foundation that allows their application on a global scale [144] and the knowledge of the satellite observation angles allows their inclusion in the inversion method. In optical remote sensing of vegetation, RTMs are used to model the phenomena of light interception by plant canopies and to interpret vegetation reflectance in terms of biophysical characteristics [75]. RTMs attempt to describe the processes of scattering and absorption, the two main physical processes involved in such a medium. For this reason, they have been widely used to design and exploit vegetation indices [22, 20, 28, 101, 24] and to develop inversion procedures to estimate vegetation parameters from remotely sensed data [63, 147, 64, 143, 105, 97].

In the past 50 years, a number of different RTMs based on different modelling approaches have been developed. Nevertheless, the PROSPECT leaf optical properties and the Scattering by Arbitrarily Inclined Leaves (SAIL) canopy bidirectional reflectance model are still the most popular and accurate [71]. In the nineties, these models have been coupled into PROSAIL, which allows to model both the spectral and the directional variation of canopy reflectance as a function of leaf biochemical properties such us chlorophyll content, water content, pigment concentration on the leaf side, and in terms of Leaf Area Index (LAI), leaf orientation on the canopy side. Furthermore, in 2001 SAIL was coupled with the Jasinski geometric model [76] to create GeoSail [69], allowing to properly describe radiation reflected by discontinuous vegetation. In the following paragraphs, we will:

- describe these models and how they are coupled, and provide a complete description of their input variables to motivate their use to model different types of vegetation and their capability to provide estimates of the LFMC
- perform GSAs of the coupled models, both across the entire supported spectrum and in the Sentinel-2 and Sentinel-3 channels, in order to identify the ideal bands for LFMC inversion
- describe the vegetation modelling scheme
- describe the model inversion procedure based on Look-Up Table (LUT)s and how the optimal inversion parameters were obtained
- perform a partial validation using field data

3.1 Description of the Radiative Transfer Models and their Variables

SAIL is one of the most widely used canopy reflectance models [72], and was developed as an extension of the one-dimensional non-lambertian directional reflectance model developed by Suits in 1972 [126]. Suits himself had extended the functionality of previously existing models by implementing the capability to reproduce the effect of the viewing angle on the directional reflectance, and to trace back these effects to a specific class of canopy components. Suit's canopy model corresponded to a finite number of infinitely extended horizontal layers, each layer containing homogeneously mixed, but randomly distributed components. In 1984, Verhoef [134], proposed a new solution of Suit's equations in order to improve the model performance related to changing viewing angles. According to Verhoef, the simplifications adopted by the previous model resulted in unrealistic spectra for certain view angle distributions. By introducing the LAI and the Leaf Inclination Distribution Function (LIDF) in the calculation of the extinction coefficients, SAIL was able to model canopies with leaves of arbitrary inclination, hence the name.

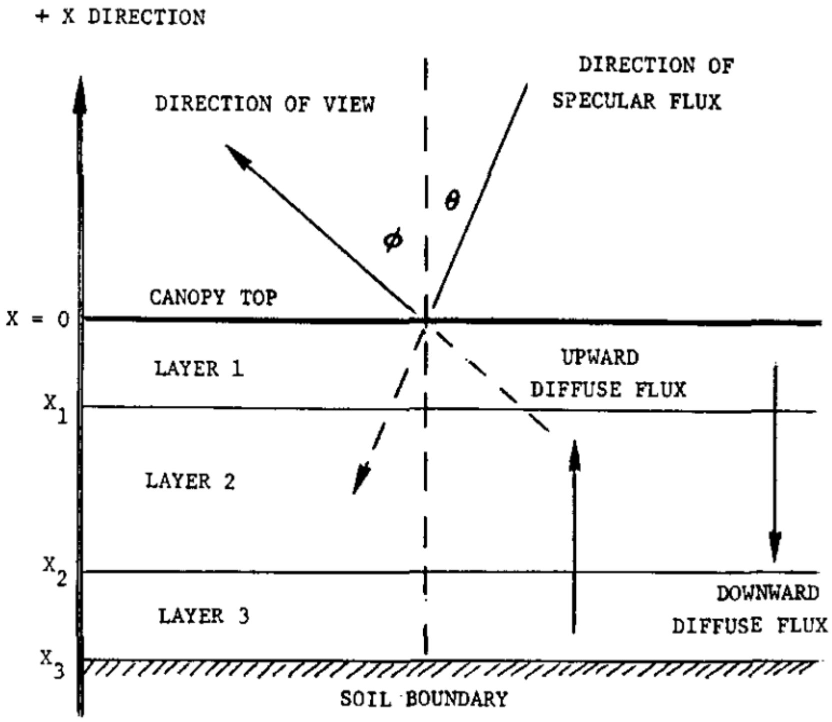

Figure 3.1. A three layer example of Suit's canopy model, adapted from [126]. Each canopy layer (e.g. grain, stalk, leaves) is modeled as a horizontal, infinitely extended layer. Each layer is composed of randomly distributed and homogeneously mixed components. The incident solar flux is divided into two components: the specular flux, which reaches canopy with an incidence θ with respect to the normal to layer surface, and founds no obstacle in its path, and the diffuse flux, which is intercepted by the canopy at least once. When the solar flux interacts with the canopy layers, it leaves the specular component and is either absorbed or contributes to the diffuse flux. The bottom-most canopy layer is always bounded by the soil, which reflects all the downward diffuse flux in the upward direction.

PROSPECT is a radiative transfer model based on Allen's generalized plate model, which was capable of describing the reflectance and transmittance of monocotyledon and dicotyledon leaves by introducing a void area index to model the different internal structure of different leaf species [4]. In 1990, Jacquemoud and Beret introduced PROSPECT to improve upon Allen's foundation, but at the same time to keep the number of model parameters to a minimum in order to facilitate model inversion [72]. This first version of PROSPECT was based on three parameters representing the leaf internal structure, pigment concentration parameter and water content. The modern version employs two main categories of parameters: the leaf structure parameter N, and the leaf biochemical content parameters.

The leaf structure parameter is related to the cellular structure of the leaf, and therefore assumes different values based on the plant species and curing state. The leaf biochemical content parameters describe the physiological state of the leaf through its mass, water content, and pigment composition. In particular, leaf pigments play a fundamental role in the ability of plants to adapt to their environment, and their parametrization has evolved the most over time compared to the other model variables. In the first model version, the authors used acetone 80% to manually extract chlorophyll a and b and carotenoids, but were unable to isolate other pigments such as tannins, brown pigments which appear in senescent leaves, because no extraction methods were available at the time [72]. In general, leaves contain three main families of pigments: chlorophylls, carotenoids and anthocyanins. Chlorophylls are responsible for the photosynthetic process, and for vascular plants they can be divided in chlorophyll-a and b. Carotenoids also contribute to light absorption, but are also well-known for their photo-protective capabilities. Finally, not very well-understood until the past few decades, anthocyanins play a wide variety of roles that range from shielding the leaf from UV-B light, to increasing resistance to drought and attacks by herbivores [60].

Subsequent updates allowed the inclusion of new types of pigments and increased the complexity of the model: in 2008, Feret was able to achieve the separation of the contributions of carotenoids and chlorophylls [45], creating PROSPECT-5. For this project, we used the 2017 version PROSPECT-D, where the D stands for "Dynamic" to refer to its capability to reproduce leaf phenology dynamics thanks to the addition of anthocyanins [46].

The inversion of PROSPECT is relatively easy, but from a remote sensing point of view, inverting the reflectance spectra of a singular leaf has limited applicability. On the other hand, SAIL provides a description of a leaf canopy, but its inversion from satellite data can be feasible only when several measurements from different viewing angles are available, which is almost never the case. To solve this issue, the two models were coupled into PROSAIL [73] since the early nineties, and their parallel evolution allowed different versions to be coupled together to create different joint models [75].

In particular, for this project we used the combination of 4SAIL and PROSPECT-D to describe continuous vegetation such as grasslands and shrublands, and the combination of PROSPECT-D and GeoSail [69], not to be confused with GeoSAIL [136], to describe horizontally discontinuous vegetation such as forest canopies. GeoSail was created in 2001 by Huemmrich using the Jasinski geometric model [76] to weigh SAIL's output based on a simple geometric description of forest structure: trees are assumed to be all the same size and shape (cylindrical or conical), to not cast shadow on each other and to be small compared to the pixel size. Since at the time of this research no public code was available that allowed to combine GeoSAIL with modern implementation of PROSPECT, we forked a freely available, open-source Python implementation [62] that allowed to combine PROSPECT-D

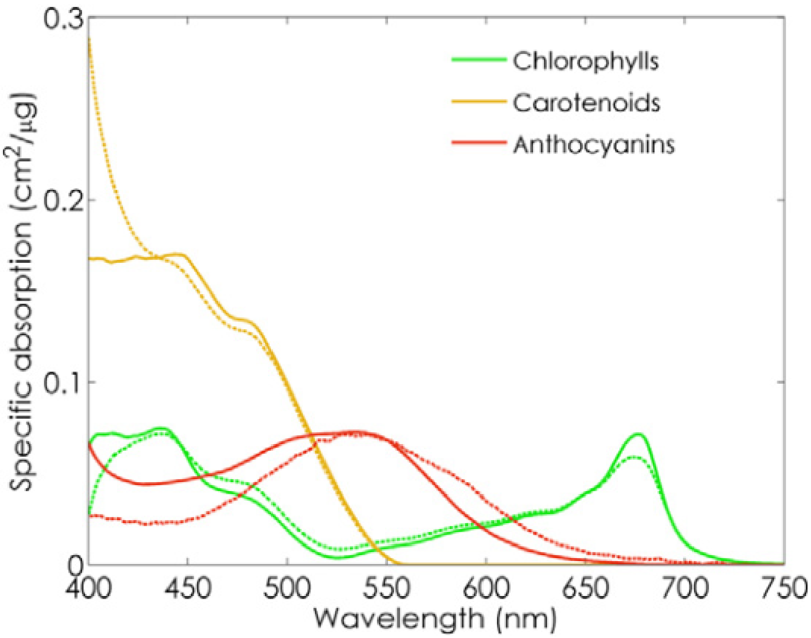

Figure 3.2. Specific absorption coefficients (SAC) of the three pigment families of PROSPECT-D, adapted from [46]. The solid lines represent the SAC curves used by the PROSPECT-D model, while the corresponding dashed lines refer to the curves used by PROSPECT-5 for the chlorophylls and carotenoids, and to the measurements carried out by [99] for the anthocyanins.

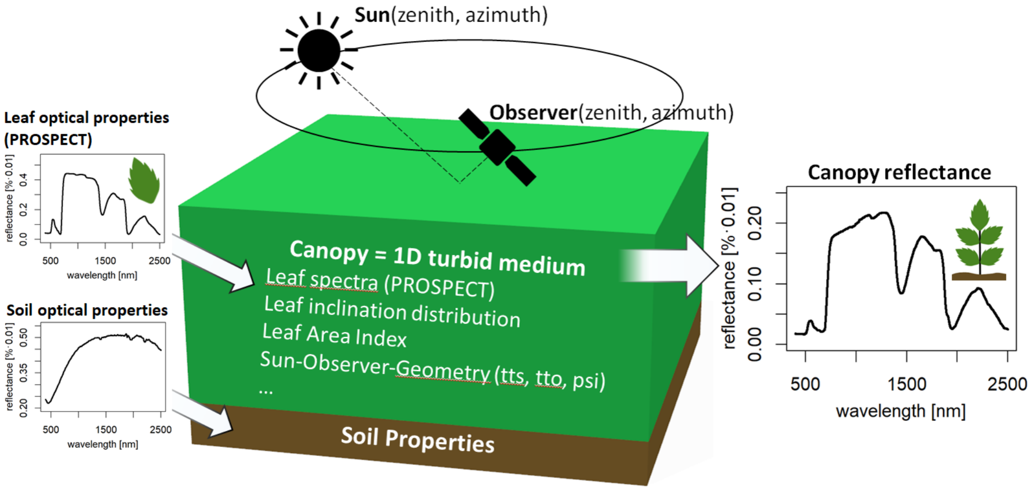

Figure 3.3. PROSPECT + SAIL coupling scheme, from [78]: PROSPECT uses a description of the leaf in terms of its structure (N) and biochemical composition ($C_{ab}, C_{anth}, C_w, \dots$) to output the leaf transmittance and reflectance. These values are used by SAIL to build the canopy model in conjunction with the leaf inclination and density ($LIDF, LAI$) and a soil spectrum (ρ_{soil}, p_{soil}). The solar illumination parameters (SZA, SAA) and the viewing angles (VZA, VAA) allow to determine the canopy reflectance that reaches the viewer.

and 4SAIL, and ported Huemmrich's Fortran code [68] to Python and interfaced it with 4SAIL.

The symbols, names and units of the model parameters of PROSPECT-D, 4SAIL and Geo are listed in tables 3.1, 3.2 and 3.5 respectively. In the following paragraphs, we will provide a detailed description of each variable of the three involved models, and then proceed to document the model implementation in Python, the sensitivity analysis and the inversion technique.

3.1.1 PROSPECT-D Model Variables

PROSPECT-D models leaf reflectance and transmittance between 400 and 2400 nm as a function of their dry mass, water content and pigment concentration.

Symbol	Quantity	Unit
N	Leaf structure parameter	-
C_w	Equivalent Water Thickness	cm
C_{dm}	Dry matter content	$g\ cm^{-2}$
C_{ab}	Clorophyll $a+b$ content	$\mu g\ cm^{-2}$
C_{ar}	Carotenoid content	$\mu g\ cm^{-2}$
C_{bp}	Brown pigments content	-

Table 3.1. Parameters of the PROSPECT-D model

Leaf Structure Parameter N

The leaf structure parameter N allows to represent the internal structure of the leaf, modelling discontinuities in matter and therefore in refractive index. As is well-known, vascular plants can be divided into two groups based on their internal structure: monocotyledons, which have only one type of linear, densely packed mesophyll, and dicotyledons, which contain a mixture of more compact palisade mesophyll and a less dense, net-like spongy mesophyll.

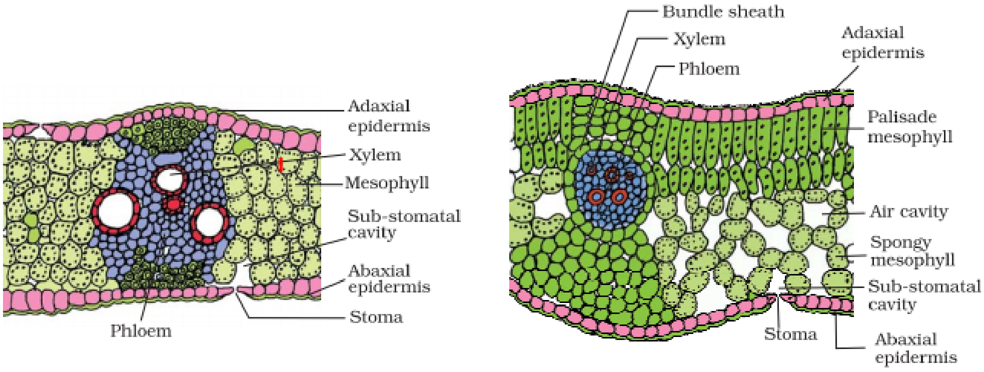

Figure 3.4. Difference between the linear, densely packed mesophyll of a monocotyledon leaf (left) and the mix between palisade and spongy mesophyll of a dicotyledon (right) leaf. In particular, the dicotyledon diagram clearly shows the cavities that permeate the spongy mesophyll of dicotyledons. Adapted from [100]

The first version of Allen's model [5] was only able to represent compact monocotyledon leaves, and was later generalized to the non-compact case [4], providing the base for PROSPECT [72]. $N = 1$ corresponds to the albino maize leaf, and values up to 1.5 correspond generally to monocotyledon species that are characterized by very compact mesophyll in the form of long, parallel vein-like structures; values higher than 1.5 and up to 2.5 can be used to represent dicotyledon leaves, where the presence of the less densely packed spongy mesophyll allows the presence of more air inside the leaf. Finally, values higher than 2.5 normally correspond to senescent leaves, with a decaying or collapsing internal structure. Jacquemoud showed that a variation of the leaf structure parameter affects the transmittance more than the reflectance [72], remarking the importance of modelling both.

Water concentration C_w and Equivalent Water Thickness EWT

The water concentration C_w represents the amount of water contained in the leaf expressed in g/cm^2. This quantity has the same value, but not the same dimensions, of the Equivalent Water Thickness EWT which is defined as the hypothetical thickness that the water contained in the leaf would have if it were spread over the surface of the leaf itself [36]. For this reason, the name EWT is often used interchangeably to refer to the concentration or to the actual thickness, creating some confusion. The water concentration can be obtained by multiplication of the Equivalent Water Thickness and the water density ρ_w:

$$C_w\ [g/cm^2] = EWT\ [cm]\ \rho_w\ [g/cm^3] \tag{3.2}$$

However, given the fact that $\rho_w \approx 1.0\ g/cm^3$, the two quantities generally have the same exact value. The strong water absorption bands in the SWIR and NIR make this quantity a good candidate for estimation through remote sensing.

Dry Matter Content C_{dm}

The dry matter content represents the dry mass per unit area of the leaf. Its value depends on the species and on the aging and metabolic state of the leaf.

Chlorophyll Content C_{ab}

This parameter represents the concentration of the chlorophyll a and b pigments, which play the main role in photosynthetic processes. Chlorophyll a absorbs the most light in the blue and red wavebands, and reflects more light in the green. Chlorophyll b absorbs most of the radiation in the blue wavelengths.

Carotenoid Content C_{ar}

Carotenoids, are a family of pigments which can be divided into oxygen-free carotenes and oxygen-containing xanthophylls. Among those that play a role in photosynthesis, Lightenthaler [85] names $\beta-$carotene, luten, violaxanthins and neoxanthins. Their absorption peak is in the blue wavelengths, around $450nm$.

Brown pigment C_{bp}

Differently than the other pigments, the brown pigment is not represented as a concentration, but as a dimensionless value that ranges from the value 0, associated to young or healthy leaves with no presence of senescent material, to 1, which corresponds to cured vegetation with the maximum content of brown pigment. Brown pigments are usually associated to the family of tannins, but their role as a PROSPECT input is rather marginalized in the scientific literature, where its value is often set to zero [33, 105] when the authors specifically aim to model green vegetation or are able to remove brown leaves in a laboratory environment [72]. As noted by Danner [35], brown pigments can be synthesized by the leaf without altering the leaf colour, which is usually triggered by a decrease of photosynthetic activity associated to chlorophyll breakdown.

3.1.2 4SAIL Model Variables

The original version of SAIL, in addition to the leaf reflectance and transmittance and to the viewing angles, only required the value of the Leaf Area Index (LAI) and the specification of a Leaf Inclination Distribution Function (LIDF) to run [134], as the model evolved new parameters began being included. In 1991, Kuusk [80] modeled the hotspot effect in canopy reflectance, and this work was integrated into SAIL creating SAILH [135] in the late nineties. In 2007, a further update allowed to take into account for the different temperatures of illuminated and shaded leaves, and further extended the applicability of the model to the Thermal Infrared (TIR) domain. Being based on a four-stream radiative transfer model, this version was named 4SAIL.

Symbol	**Quantity**	**Unit**
ρ	Leaf reflectance	-
τ	Leaf transmittance	-
LAI	Leaf area index	m^2/m^2
LIDF	Leaf inclination distribution function	-
Hot	Hotspot parameter	m/m
ρ_s	Soil reflectance assumed Lambertian or not	-
sza or θ_s	Solar zenith angle	deg
vza or θ_v	Viewing zenith angle	deg
raa or φ_{sv}	Relative azimuth angle	deg

Table 3.2. Parameters of the 4SAIL model

Leaf Area Index

As defined by Verhoef in [135], the Leaf Area Index of a canopy layer is the ratio between the total, one-sided leaf canopy area to the associated ground surface area. Canopies with a higher LAI will be more dense, occupying more of the available area around the tree, and shadowing more of the ground.

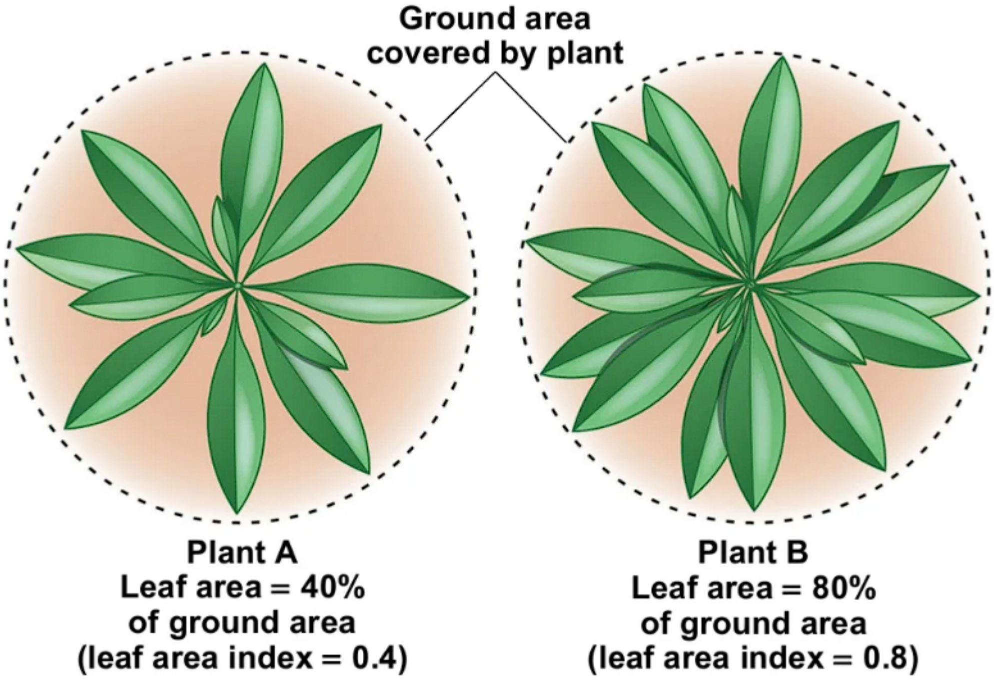

Figure 3.5. Graphical comparison between two different LAI values: on the left, plant A with fewer leaves only manages to cover 40% of the ground area ($LAI = 0.4$), while on the right the much lusher plant B covers 80% of the ground area ($LAI = 0.8$). Sourced from [1]

Leaf Inclination Distribution Function

The Leaf Inclination Distribution Function (LIDF) is used to represent in a simple way a number of different leaf orientations. Verhoef [135] describes a method to define a number of different LIDFs using two parameters: average leaf slope $LIDF_a$, and distribution bi-modality $LIDF_b$. The implementation used for our purposes supports the functions listed in table 3.3.

LIDF type	$LIDF_a$	$LIDF_b$
Planophile	1	0
Erectophile	-1	0
Plagiophile	0	- 1
Extremophile	0	1
Spherical	-0.35	-0.15
Uniform	0	0

Table 3.3. Two Parameter LAD parametrization, where a and b represent the average leaf slope and distribution bimodality respectively.

Hotspot Factor and Hotspot Effect

It is well-known that the backscatter direction corresponds to the maximum value of the reflected diffuse radiation. The same effect with different names can be found across different disciplines: *heiligenschein* in meteorology, opposition effect in astronomy, and hotspot effect in optical remote sensing [80, 65]. The explanation of this effect is that in the backscatter direction, the shadow cast by the particles of the medium the radiation is impacting will not be visible, while in any other direction a portion of the shadow will necessarily be visible, lowering the value of the reflected radiance. In order to model this effect, SAILH introduced the hotspot parameter Hot, which represents the ratio of leaf width to canopy height. For this reason, the values of this parameter may vary a lot depending on the species considered. Some examples of these values for different crops as presented by [8] are summarized in table 3.4.

Crop Type	Hot $[m/m]$
Maize	0.01 - 0.20
Wheat	0.01 - 0.50
Rice	0.01 - 0.10
Soybean	0.20
Sugar Beet	0.20 - 0.40

Table 3.4. Values of the hotspot parameter for different types of crops, from [8]

Soil Reflectance and Brightness Factor

In order to model the contribution of the upper soil layer to the reflectance, and the changes in this reflectance that may be caused by varying levels of soil moisture, two soil-related parameters are defined: a soil reflectance factor ρ_{soil}, which depends on the wavelength, and a dimensionless soil brightness factor α_{soil} that can be used to scale the soil reflectance value [70, 108]. The Python bindings used for this project offer two soil spectra, a dry soil spectrum named $soil\_spectrum\_1$ and a wet soil spectrum $soil\_spectrum\_2$, which are represented in figure 3.6, and offer the possibility to merge them using the brightness factor and a dryness factor p_{soil}. The resulting ρ_{soil} is easily computed as a weighted sum of the two reflectance spectra:

$$\rho_{soil} = \alpha_{soil} \cdot [p_{soil} \cdot soil\_spectrum\_1 + (1 - p_{soil}) * soil\_spectrum\_2] \tag{3.3}$$

Setting the soil dryness factor p_{soil} to zero implies using the wet soil only, while setting it to 1 implies using the dry spectrum only. The spectra can also be specified by the user using the $soil\_spectrum\_1$ and $soil\_spectrum\_2$ keywords.

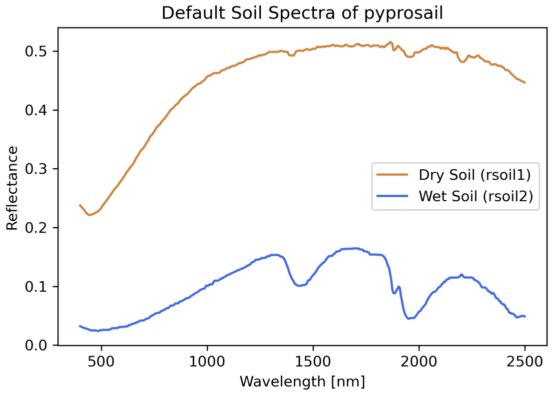

Figure 3.6. Dry (brown line) and wet (blue line) soil spectra included in the `pyprosail` Python library. The library also allows the users to provide their own soil spectra.

Illumination and Sensor Viewing Geometry

Finally, in order to define the illumination geometry and the amount of radiation that reaches the observer, the following angles must be set:

- **Sun Zenith Angle** sza or θ_s, the angle between the local zenith and the incident Sun rays. It is zero when the Sun is at the zenith, and is 90 degrees when the Sun is on the horizon.
- **View Zenith Angle** vza or θ_v, the angle between the local zenith and the rays that impact the viewer.
- **Relative Azimuth Angle** raa or φ_{sv}, the angular displacement between the azimuth of the Sun and the viewer.

A graphical scheme of the coordinate system that represents the scene geometry can be found in figure 3.7.

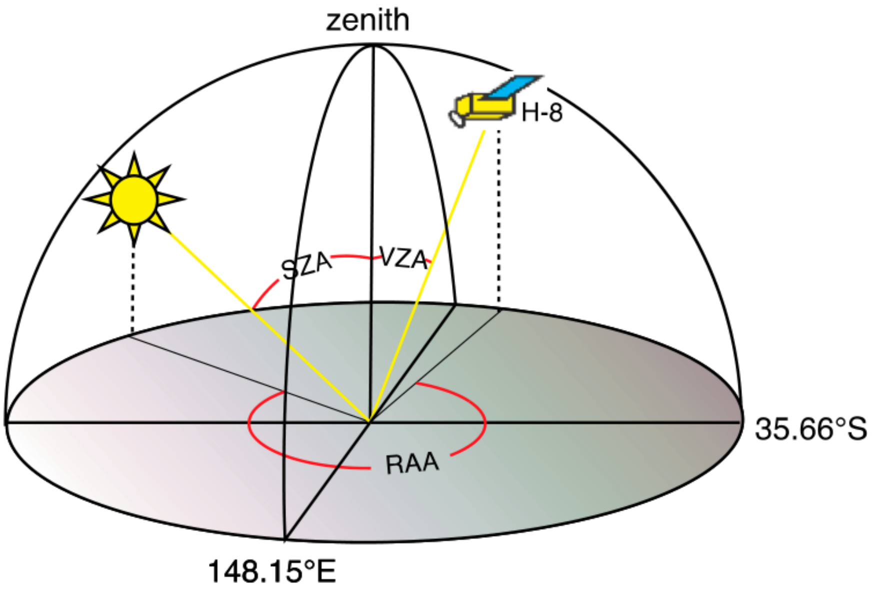

Figure 3.7. Sun-Sensor geometry scheme adapted from [87]

3.1.3 Geo4Sail

Since there was no open-source publicly available library that coupled the Jasinski geometric model with a modern version of SAIL, we forked the `pyprosail` library and ported Huemmrich's Fortran code to it. Since GeoSail was a simple addition to SAIL, we could exploit the 4SAIL functions of `pyprosail`, reducing the code porting workload only to the Geo functions. Since at the time of writing only the functions for cylinders and cones were implemented, this meant rewriting in Python the GeoSail `GEOCONE` and `GEOCYLI` functions [68]. However, following Jasinski's instructions [76] it will be possible to implement other shapes in the future. Each function requires the inputs summarized in table 3.5, the first three of which are outputs of 4SAIL:

Symbol	Quantity	Unit
tdo	Canopy hemispherical-directional transmittance	-
rdo	Canopy hemispherical-directional reflectance	-
rso	Canopy bi-directional reflectance	-
ρ_s	Soil reflectance assumed Lambertian or not	-
chw	Crown height-to-width ratio	m/m
$ccov$	Crown coverage	-
$cshp$	Crown shape	-

Table 3.5. Main variables of the Geo model

In the cylinder case, the scene reflectance is calculated as the sum of the component reflectances of the illuminated and shadow portions of the background and of the canopy. In the cylinder case, the contribution of the shadowed crown is neglected, because it is assumed to be quantitatively negligible with respect to the other components.

Geocyli

For square cylinders, the ratio of canopy area to shadowed area for the individual crown η can easily be calculated as:

$$\eta = chw \ \tan(sza) \tag{3.4}$$

Using η we can compute the fraction of shadowed soil as follows:

$$sfrac = 1 - ccov - (1 - ccov)^{\eta+1} \tag{3.5}$$

The fraction of illuminated soil can then be calculated by subtracting from one the fraction of soil covered by the canopy crown $ccov$ and the fraction of shadowed soil $sfrac$:

$$ilsoil = 1 - ccov - sfrac \tag{3.6}$$

The reflectance of the shadowed background $rssh$ can be calculated using the hemispherical-directional transmittance through the crown tdo outputted by 4SAIL and the soil reflectance spectrum $rsoil$:

$$rssh = tc \ \rho soil \tag{3.7}$$

Finally, assuming that the contribution of the shadowed canopy is negligible, the scene reflectance can be calculated as follows:

$$rsc = ccov \ rc + sfrac \ rssh + ilsoil \ \rho soil \tag{3.8}$$

where the first summand is the reflectance of the illuminated crown, the second summand is the reflectance of the shadowed background and the third component is the reflectance of the illuminated background.

Geocone

The cone aspect angle $caspa$ depends only on the crown height to width ratio chw:

$$caspa = \arctan\left(\frac{1}{2\ chw}\right) \tag{3.9}$$

Knowing the cone aspect angle, we can use it and the sun zenith angle to calculate the angle describing the location of the edge of the shadowed area on the cone β:

$$\beta = \arccos\frac{\tan(caspa)}{\tan(sza)} \tag{3.10}$$

This allows us to compute the ratio of canopy area to shadowed area η for an individual conical crown:

$$\eta = \frac{\tan\beta - \beta}{\pi} \tag{3.11}$$

In the conical case, the fraction of shadowed crown is simply the ratio between the β angle and π:

$$fcsh = \frac{\beta}{\pi} \tag{3.12}$$

The fraction of shadowed background can be computed with the same formula as the cylindrical case thanks to the usage of the similarity factor η:

$$sfrac = 1 - ccov - (1 - ccov)^{\eta+1} \tag{3.13}$$

The same goes for the fraction of illuminated soil:

$$ilsoil = 1 - ccov - sfrac \tag{3.14}$$

The reflectance of the shadowed crown is by definition the product of the hemispherical-directional transmittance tc and the hemispherical-directional reflectance rc outputted by 4SAIL:

$$rchs = tc\ rc \tag{3.15}$$

The reflectance of the shadowed background can be computed as before:

$$rssh = tc\ \rho soil \tag{3.16}$$

Finally, the scene reflectance is given by:

$$rsc = ccov(1 - fchs)\ rc + (ccov\ fcsh)\ rcsh + sfrac\ rssh + ilsoil\ \rho soil \tag{3.17}$$

where the first summand represents the contribution of the illuminated crown, the second summand the contribution of the shadowed crown, the third summand the contribution of the shadowed background and the fourth summand the contribution of the illuminated background.

Addition of Other Shape Functions

In principle, in order to add trees with different shapes, one need only define the η factor as in [76] in relation to the target shape. The article already provides the similarity parameters of circular cylinders and spheres, as summarized in table 3.6.

Shape	$\eta = A_s/A_t$
Circular Cylinder	$\frac{4\ tan(sza)\ chw}{\pi}$
Square Cylinder	$\frac{tan(sza)\ chw}{\pi}$
Cone	$\frac{\cot\psi - \frac{\pi}{2} + \psi}{\pi}$
Sphere	$\tan(sza) \sin(sza)$

Table 3.6. Similarity parameter of canopy geometries, adapted from [76]. The parameter ψ is defined as arcsin $\frac{\frac{1}{chw}}{\tan(sza)}$. $\eta = A_s/A_t$ refers to the ratio between shadowed and total area.

3.2 Vegetation modelling and RTM Inversion Strategy

Our vegetation modelling strategy is heavily based on [105], which can be considered the state of the art regarding LFMC inversion from satellite images based on the inversion of PROSAIL-adjacent RTMs. As anticipated in the previous chapter, we distinguish three basic vegetation types: grassland, shrubland and forest. The combination of PROSPECT-D and 4SAIL is used to model grasslands and shrublands, which are assumed to be homogeneous vegetated pixels. The combination of PROSPECT-D, 4SAIL and GEO allows us to model potentially discontinuous vegetation, and therefore we use it to model forested pixels. Differently than [105], we do not use a grassland canopy reflectance to model the forest understory, but we combine the `pyprosail` dry and wet soil spectra in different proportions using the *psoil* parameter.

The RTM inversion strategy applied to these two model combinations is based on the use of LUTs, again similarly as [105], but with small improvements on the LUT creation side. Before describing in detail the methodology adopted to create and invert the LUTs, we will perform a sensitivity analysis of the two model combinations in order to verify that the model is sensitive to the variables that allow us to calculate the LFMC, and in particular to identify the ideal Sentinel-2 and Sentinel-3 channels to use to successfully perform the inversion. While many PROSAIL sensitivity analyses can be found in the scientific literature, almost none of them are specifically tailored to the problem of live fuel moisture content estimation, and since many versions of PROSPECT and SAIL exist, it was deemed useful to perform a GSA on the specific model versions and combinations we are going to use.

3.3 Sensitivity Analyses of the Combined Models

Before attempting to retrieve the biophysical parameters from satellite observations using a model inversion technique, it is mandatory to verify the feasibility of the retrieval on a variable by variable basis. Indeed, there may be large differences in the quantitative contribution to the model output of the individual variables, i.e. the variation of one particular variable or of a group of variables may determine most of the model output, while the contribution of the remaining ones may not determine any significant quantitative change. If a variable does not have any significant effect on the variation of the model output, it will be very difficult to invert it, because the model will barely react even to large variations of the variable.

Furthermore, given the fact that we are interested in inverting the model from satellite data, only a limited number of wavebands will be at our disposal. Therefore, in addition to verifying that the variation of our target variables has a significant impact on the model output, we will also need to verify that this applies to the portions of the spectrum sample by the satellite platform providing the observations.

To this end, Global Sensitivity Analysis (GSA) are a family of mathematical procedures that allow to quantify the relative contribution of each variable of the model to the model output [122, 123], and are often used to quantify the effect of the uncertainty on a variable on the uncertainty of the model output [114]. In remote sensing, they are often used to assess the feasibility of the retrieval of the parameters of a certain model from remotely sensed data [21, 9, 103, 97]. In particular, the Sobol method allows to obtain both first-order sensitivity indices, which quantify the contribution of each variable *per se*, i.e. while all the other variables are kept constant, and higher-order sensitivity indices that account for the interactions between variables. The possibility to take into account the interactions between variables is fundamental, because some variables that may appear to be not significant at the first-order, may become significant through interactions [103].

Let $y = f(x_0, x_1, ..., x_m)$ be our m-parameter model, and let $V(x_i)$ be the variance of the i-th model parameter x_i. Sobol's method allows to decompose the total unconditioned variance $v(y)$ into the conditional variances of each individual parameter $v(f_i(x_i))$ and into the conditional variances associated to the interactions between parameters $v(f_{ij}(x_i, x_j))$. Since each parameter can interact with all the others, the order of the interactions will be up to m. The total unconditioned variance can therefore be decomposed as the sum of the first-order variances and the sum of the higher-order variances:

$$v(y) = \sum_{i=1}^{m} v_i + \sum_{i}^{m} \sum_{j=i+1}^{m} v_{ij} + v_{ijk} + ... + v_{ijk...m} \tag{3.18}$$

The first-order variance associated to the i-th variable x_i can be calculated as the variance of the expected value of f when x_i is fixed and all the other parameters $x_{\neq i}$:

$$v_i = v[f_i(x_i)] = v_{x_i}[E_{x_{\neq i}}(y|x_i)] \tag{3.19}$$

Finally, the sensitivity indices for each individual variable can be obtained using their variance and dividing it by the total unconditioned variance. The first-order sensitivity indices S_i of the i-th variable is therefore defined as the portion of total unconditioned variance associated to the variation of the parameter x_i without taking into account the interactions with the other parameters $x_{\neq i}$:

$$S_i = \frac{v_i}{v(y)} = \frac{v_{x_i}[E_{x_{i \neq j}}(y|x_i)]}{v(y)} \tag{3.20}$$

The k-th order sensitivity index associated to the interaction of the i-th model parameter x_i with other $n-1$ variables can intuitively be obtained by subtracting their n first-order variances to the expected value of the model when those n parameters are fixed and all the other $m-n$ parameters are allowed to vary:

$$S_{ij...n} = \frac{v_{ij..n}}{v(y)} = \frac{v_{x_i x_j ... x_n}[E_{x_{i \neq ij...n}}(y|x_i, x_j, ..., x_n)]}{v(y)} \tag{3.21}$$

Finally, the total-order sensitivity index of the i-th variable is the sum of the first-order sensitivity index of that variable and of all the higher-order sensitivity indices associated to its interactions with the other variables:

$$S_{Ti} = S_i + \sum S_{ij} + ... \sum S_{ij...m} = \frac{E_{x_{\neq i}}[v_{x_i}(y|x_{\neq i}]}{v(y)} \tag{3.22}$$

These calculations were carried out using the SALib Python library [66], which makes it straightforward to construct the necessary sequences for each model parameter through the Saltelli module, requiring only the definition of the variation interval of each variable as an input. The module returns quasi-random, low-discrepancy sequences calculated in order to sample the parameter space as uniformly as possible by exploiting Saltelli's sampling scheme, which is an improvement of Sobol's [113, 66]. The sequences are then used as inputs for our fork of the `pyprosail` Python library, and the resulting outputs are ran through the `sobol.analyze` function which performs the Sobol Global Sensitivity Analysis returning both first-order indices and total-order indices.

In the following paragraphs, we will present the results of the sensitivity analysis for the PROSPECT-D + 4SAIL model, which will be applied to grasslands and shrublands, and for the PROSPECT-D + 4SAIL + GEO model, which will be applied to forests. Each paragraph will contain a table detailing the boundaries of the model variables fed as inputs to SALib Saltelli module, graphs of the results of the GSA for the entire portion of the spectrum covered by PROSPECT, and finally the integration of the results in the Sentinel channels. For the sake of brevity, the results of the GSA are shown only for the Sentinel-3A and for the Sentinel-2A channels. Due to the similarity of the spectral response functions of the "A" sensors and their "B" counterparts, the results are graphically indistinguishable, and will therefore be omitted for the sake of brevity.

3.3.1 PROSPECT-D + 4SAIL

The boundaries of the parameters used to run the GSA of the combined PROSPECT-D and 4SAIL are listed in table 3.7. The values of the variables are sourced from the most recent literature [103, 105], and refer to the range these parameters usually occupy in grasslands and shrublands. Furthermore, to simplify the management of the different leaf inclination distribution functions, we exploited the `pyprosail` feature that allows to set the LIDF using an ellipsoidal function defined only by an average leaf slope angle, and letting it vary between 0 and 90 degrees.

The results of this GSA over the entire spectrum supported by PROSPECT-D are shown in figure 3.8, while the results in the Sentinel-3 SYNERGY S3A channels are shown in figure 3.9 and 3.10 for Ocean and Land Colour Instrument (OLCI) and Sea and Land Surface Temperature Radiometer (SLSTR) respectively. Finally, figure 3.11 shows the results in the Sentinel-2A MultiSpectral Instrument (MSI) channels. All these graphs show the values of the total-order sensitivity indices of all the variables of PROSPECT-D and 4SAIL, except for the angular variables, which have not been included in the GSA due to the fact that they are always known when using satellite data. The results for the Sentinel-3 case were anticipated in the conference paper [97].

First of all, if we assume a significance threshold of 5% as done in [103], we can state that the hotspot parameter Hot and the carotenoid concentration C_{ar} are not significant across the entire spectrum supported by PROSPECT-D, because their

Symbol	Quantity	Unit	Min-Max	Type
PROSPECT-D				
N	Leaf structure parameter	-	$[1 - 3.0]$	float
C_w	Water concentration	$g \cdot cm^{-2}$	$[0.0001 - 0.0360]$	float
C_{dm}	Dry matter content	$g \cdot cm^{-2}$	$[0.0017 - 0.096]$	float
C_{ab}	Clorophyll $a + b$ concentration	$\mu g \cdot cm^{-2}$	$[1 - 100]$	float
C_{ar}	Carotenoid concentration	$\mu g \cdot cm^{-2}$	$[1 - 10]$	float
C_{ant}	Anthocyanin concentration	$\mu g \cdot cm^{-2}$	$[0 - 40]$	float
C_{bp}	Brown pigment	-	$[0 - 1]$	float
4SAIL				
LAI	Leaf area index	-	$[0.5 - 7]$	float
lidfa	Average leaf angle	deg	$[0 - 90]$	float
p_{soil}	Dry-wet soil factor	-	$[0 - 1]$	float
r_{soil}	Soil brightness factor	-	$[0 - 1]$	float
Hot	Hotspot parameter	(m/m)	$[0.01 - 0.40]$	float

Table 3.7. Parameter bounds used for the GSA of the combination of PROSPECT-D and 4SAIL.

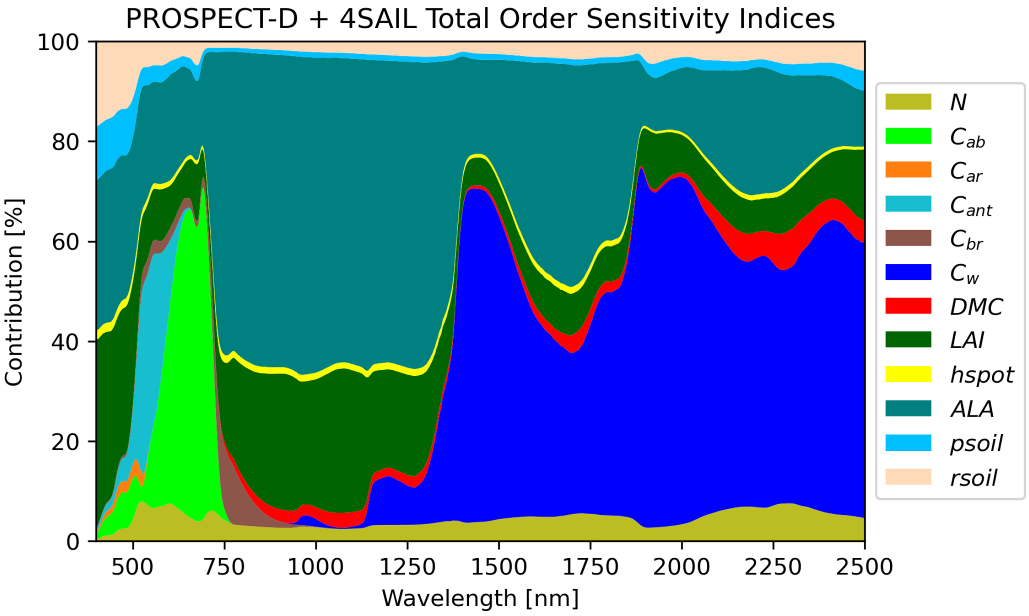

Figure 3.8. GSA of the PROSPECT-D + 4SAIL radiative transfer models. The stacked plot shows the progression of the total order sensitivity indices of each parameter of the combined model in the portion of the spectrum supported by PROSPECT. Viewing angles have not been included in the GSA, since their value is always known.

total order sensitivity index never reaches the significance threshold. In general, the most significant biophysical parameters for PROSPECT-D + 4SAIL in the visible spectrum are the LAI, Average Leaf Angle (ALA), anthocyanin concentration C_{ant} and chlorophyll a and b concentration C_{ab}.

In the violet-blue wavelengths, the soil brightness and dryness parameters reach their highest significance values, with the LAI and ALA still making up most of the reflectance. In the visible spectrum, the anthocyanin concentration C_{ant} and the chlorophyll a and b concentration C_{ab} take the lead, with the former being the highest significance parameter in the blue wavelengths and the latter becoming prominent in the green and red wavelengths. The LAI and ALA make up most of the output reflectance in the NIR, until we reach the SWIR and the water thickness C_w takes on the majority of the significance. Finally, in the longer SWIR wavelengths, the dry matter content C_{dm} reaches its highest significance values, even though most of the reflectance is still traceable to C_w and ALA variations. The brown pigment C_b is most significant in the shorter NIR wavelenghts, while the leaf structure parameter N contributes relatively consistently across the whole spectrum.

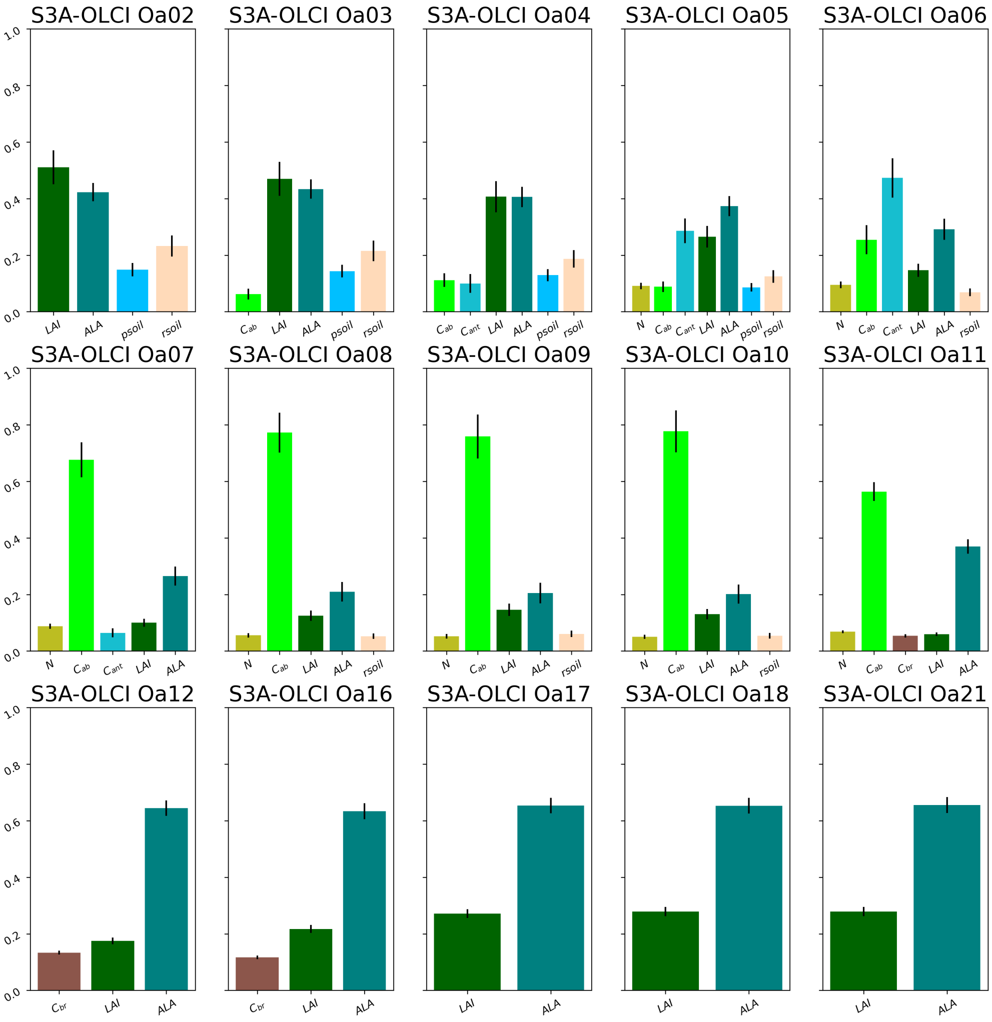

Figure 3.9. Total order sensitivity indices of the PROSPECT-D + 4SAIL parameters in the Sentinel-3A OLCI channels.

The graph in figure 3.9 shows 15 out of the 16 OLCI bands included in the SYNERGY product. The OLCI Oa01 channel was disregarded since it does not completely fall into the portion of the electromagnetic spectrum covered by PROSAIL. The LAI and the ALA are the most significant variables in the blue channels, while the anthocyanin content plays a prominent role as we transition into the green channels Oa05 and Oa06, remarking the importance of its addition in the PROSPECT-D model. The chlorophyll content takes over from channel Oa07 onwards. The OLCI NIR channels (Oa11 and following) are mostly dominated by the ALA, but the brown pigment and the LAI also contribute to a significant portion of the reflectance. The other model parameters such as dry matter content, leaf structure parameter, and soil parameters are generally of low significance.

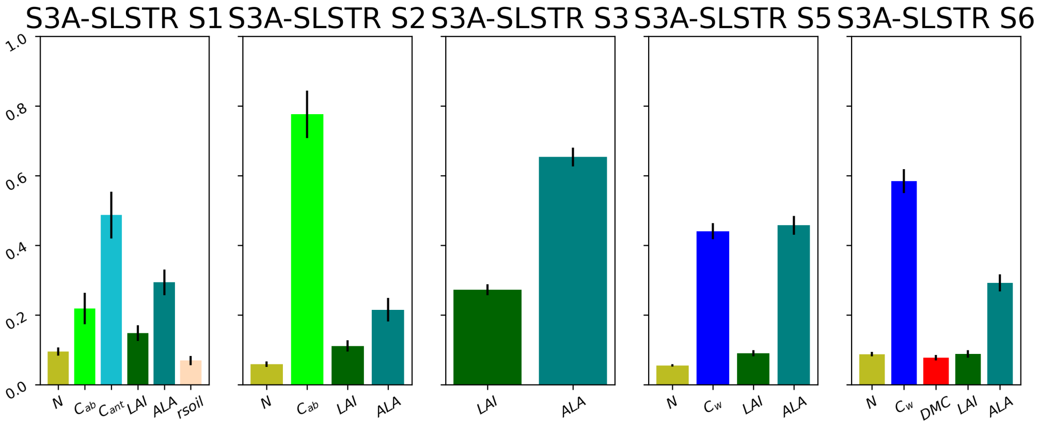

Figure 3.10. Total order sensitivity indices of the PROSPECT-D + 4SAIL parameters in the Sentinel-3A SLSTR channels.

The SYNERGY product contains all the SLSTR solar-reflective channels except for S4. Channels S1 and S2 fall in the green and red respectively, and show very similar results to the overlapping OLCI bands Oa06 and Oa08. The same can be said for channel S3, which falls in the NIR and encompasses OLCI Oa17. Unsurprisingly, SWIR channels S5 and S6 show very significant water thickness contribution. Furthermore, S6 is also the SYNERGY channel where the dry matter content contribution is most significant. Consequently, successful inversion of the LFMC cannot disregard these two channels.

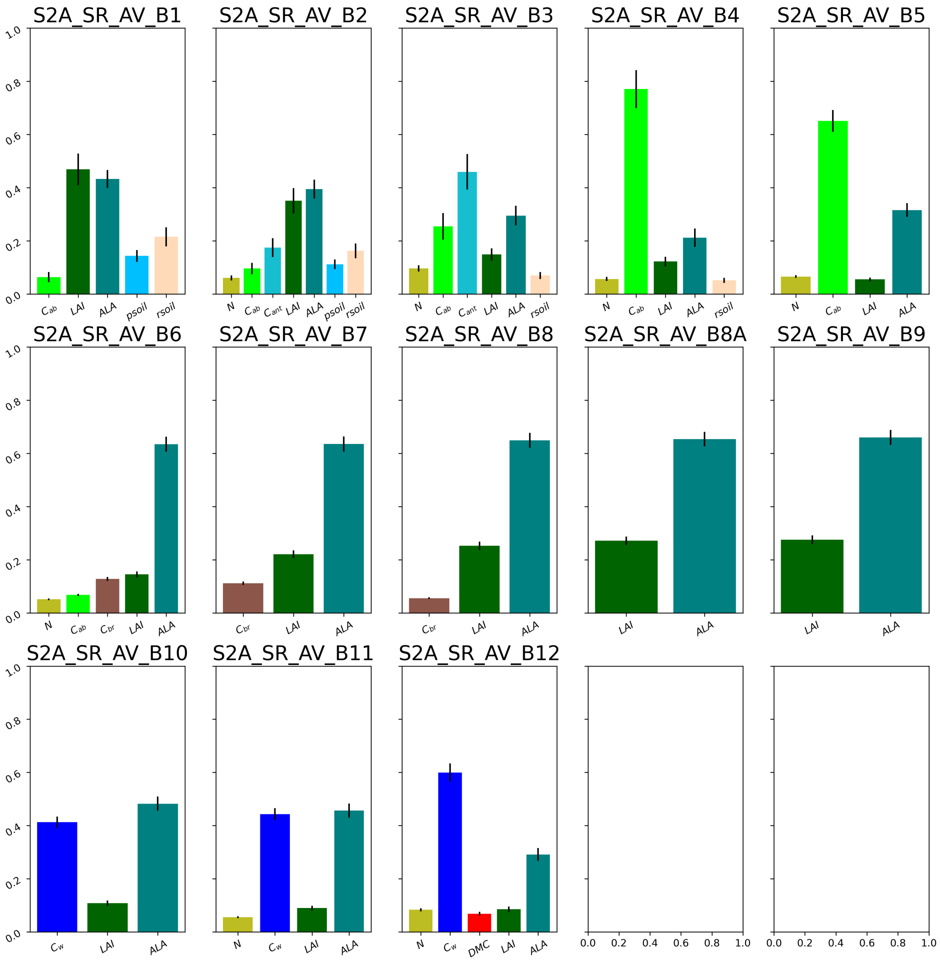

Figure 3.11. Total order sensitivity indices of the PROSPECT-D + 4SAIL parameters in the Sentinel-2A S2MSI2A channels.

Of the twelve Sentinel-2 S2MSI2A channels, only eleven are distributed as actual surface reflectance products, since channel B10 is used for atmospheric correction purposes and does not contain surface information. In this case, channels B11 and B12 are the most important with respect to successful LFMC inversion, similarly as SLSTR's S5 and S6 with which they share very similar central wavelengths.

3.3.2 PROSPECT-D + 4SAIL + GEO

The boundaries of the parameters used to run the total order sensitivity indices of the combined PROSPECT-D, 4SAIL and GEO are listed in table 3.8. The values of the variables are sourced from the most recent literature [105], and refer to the range these parameters usually occupy in forest environments. The same considerations made in the previous paragraph about the LIDF apply here.

The results of the GSA in terms of total order sensitivity indices across the entire PROSPECT-D spectrum are shown in figure 3.12. While the SZA was not included in the sensitivity analysis for the same reasons stated in the PROSAIL case, the VZA and RAA have been set to zero, since GEO only supports a nadir viewing geometry. Furthermore, the GSA was conducted for a cylinder shaped crown. For this reason, the only additional variables compared to the PROSAIL case will be the

Symbol	Quantity	Unit	Min-Max	Type
PROSPECT-D				
N	Leaf structure parameter	-	$[1.05-2.74]$	float
C_w	Water concentration	$g \cdot cm^{-2}$	$[0.0001-0.0290]$	float
C_{dm}	Dry matter content	$g \cdot cm^{-2}$	$[0.0018-0.0189]$	float
C_{ab}	Clorophyll $a+b$ concentration	$\mu g \cdot cm^{-2}$	$[1-107]$	float
C_{ar}	Carotenoid concentration	$\mu g \cdot cm^{-2}$	$[1-10]$	float
C_{ant}	Anthocyanin concentration	$\mu g \cdot cm^{-2}$	$[0-40]$	float
C_{bp}	Brown pigment	-	$[0-1]$	float
4SAIL				
LAI	Leaf area index	-	$[0.5-5]$	float
lidfa	Average leaf angle	deg	$[0-90]$	float
p_{soil}	Dry-wet soil factor	-	$[0-1]$	float
r_{soil}	Soil brightness factor	-	[0.3]	float
Hot	Hotspot parameter	(m/m)	$[0.01-0.40]$	float
sza or θ_s	Solar zenith angle	deg	$[27-80]$	float
GEO				
chw	Crown height to width ratio	m/m	$[1-3]$	float
$ccov$	Crown coverage	m^2/m^2	$[0.2-1]$	float
$cshp$	Crown shape	-	Cone, cylinder	-

Table 3.8. Parameter bounds used for the GSA of the combination of PROSPECT-D, 4SAIL and GEO. Note the absence of the viewer azimuth angle and the relative azimuth angle, due to the fact that GEO only supports nadir viewing geometry.

crown height-to-width ratio chw and the crown coverage $ccov$. In addition, given the important role of the soil reflectance in this model, we set the soil brightness parameter $rsoil$ to a constant value of 0.3. This allowed us to obtain more useful results from the GSA, since otherwise the soil reflectance would dominate the output, especially in the visible portion of the spectrum.

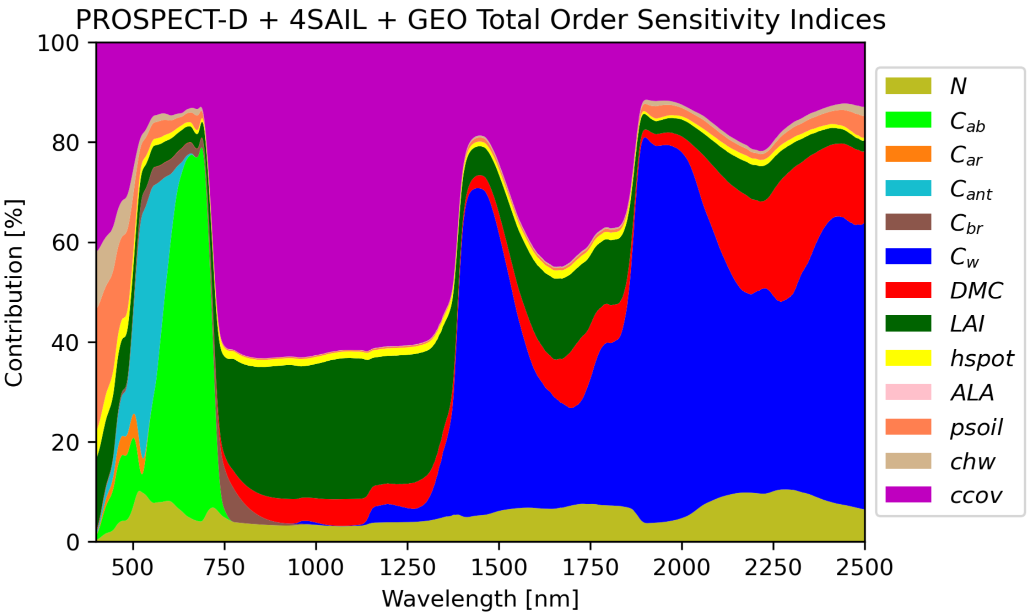

Figure 3.12. GSA of the PROSPECT-D + 4SAIL + GEO radiative transfer models. The stacked plot shows the progression of the total order sensitivity indices of each parameter of the combined model in the portion of the spectrum supported by PROSPECT. Viewing angles have not been included in the GSA, since their value is always known.

As expected, the crown coverage *ccov* has a prominent role in determining the output reflectance, because it directly controls the proportion of radiation that will be associated to the canopy and to the soil. The other additional variable, i.e. the crown height-to-width ratio, is not significant across most of the NIR and SWIR spectrum, and reaches its highest significance values in the visible spectrum in general and in the blue wavelengths in particular. Qualitatively, the considerations made in the previous section regarding the importance of the leaf biochemical parameters still apply here with no significant change, but quantitatively their role is significantly reduced by that of the crown coverage. Another major difference with the previous model combination is the role of the ALA, previously very important in the NIR-SWIR, now insignificant across the entire spectrum due to the prominence of the the crown coverage.

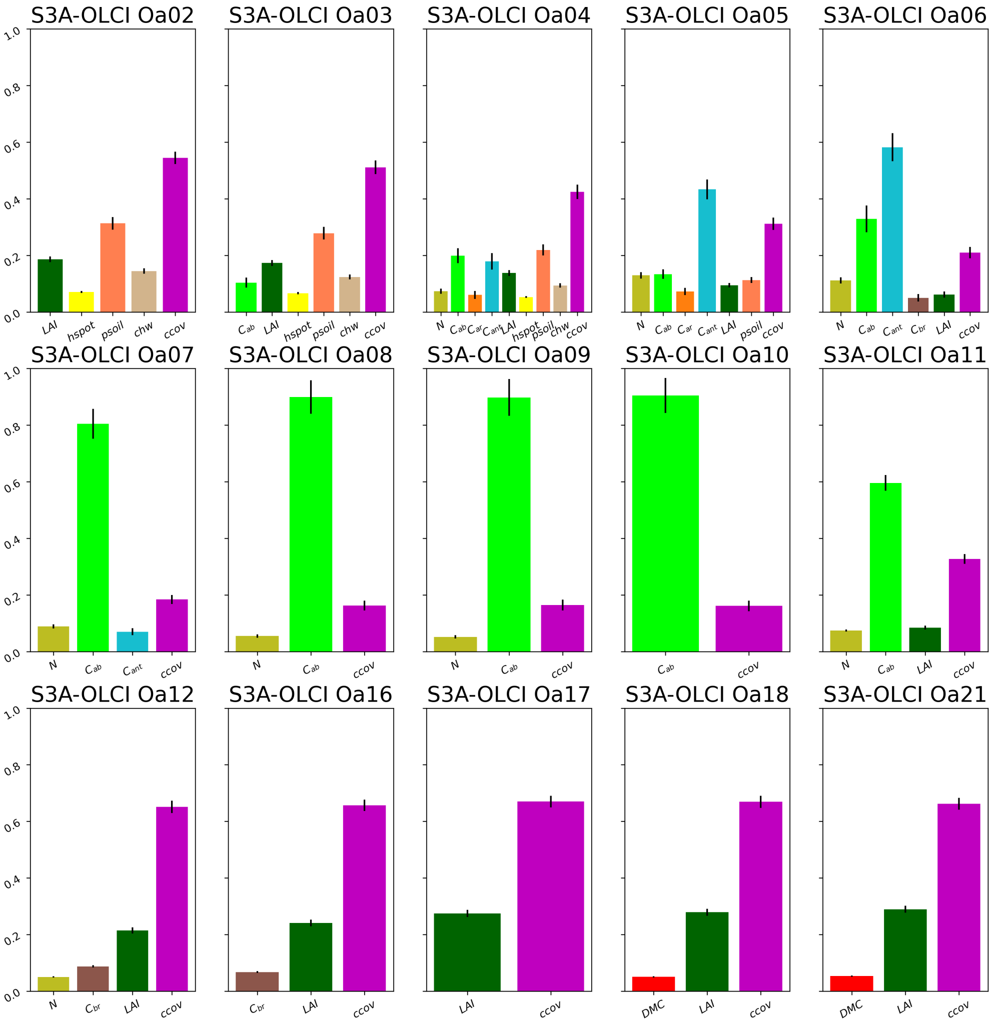

Figure 3.13. Total order sensitivity indices of the PROSPECT-D + 4SAIL + GEO parameters in the S3A OLCI channels.

With respect to the PROSAIL case, the LAI has much less influence in the output of the OLCI visible channels, with most of the reflectance being associated to variations in the soil dryness parameter and in the crown coverage. The crown height-to-width ratio is also significant from Oa02 to Oa04. The anthocyanins and chlorophyll concentration take up their usual roles in the blue and green-red wavelengths respectively, while the brown pigment and LAI rise up in importance in the NIR channels.

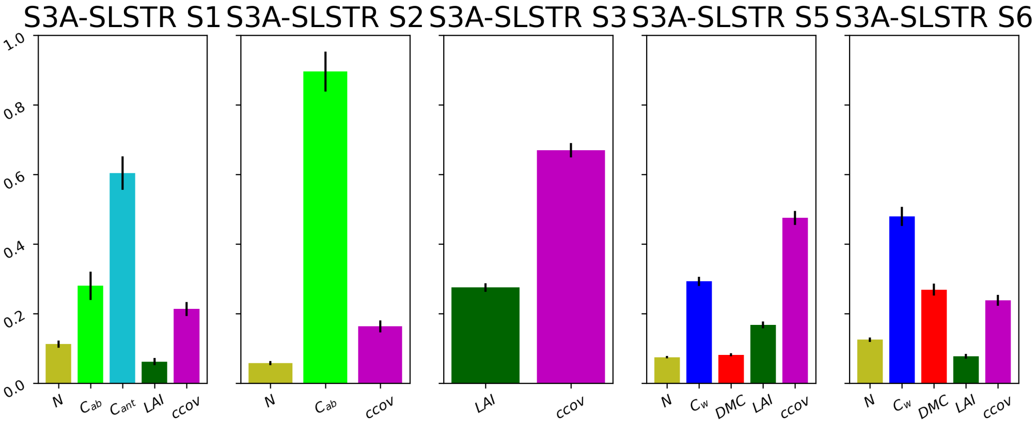

Figure 3.14. Total order sensitivity indices of the PROSPECT-D + 4SAIL + GEO parameters in the S3A SLSTR channels.

The SLSTR channels show the usual water thickness prominence in the SWIR channels S5 and S6, but the crown coverage has a strong effect in both channels. In S5 it reaches a higher total sensitivity index than the water thickness itself. Compared to the PROSAIL case, the dry matter content maintains generally higher sensitivity index values.

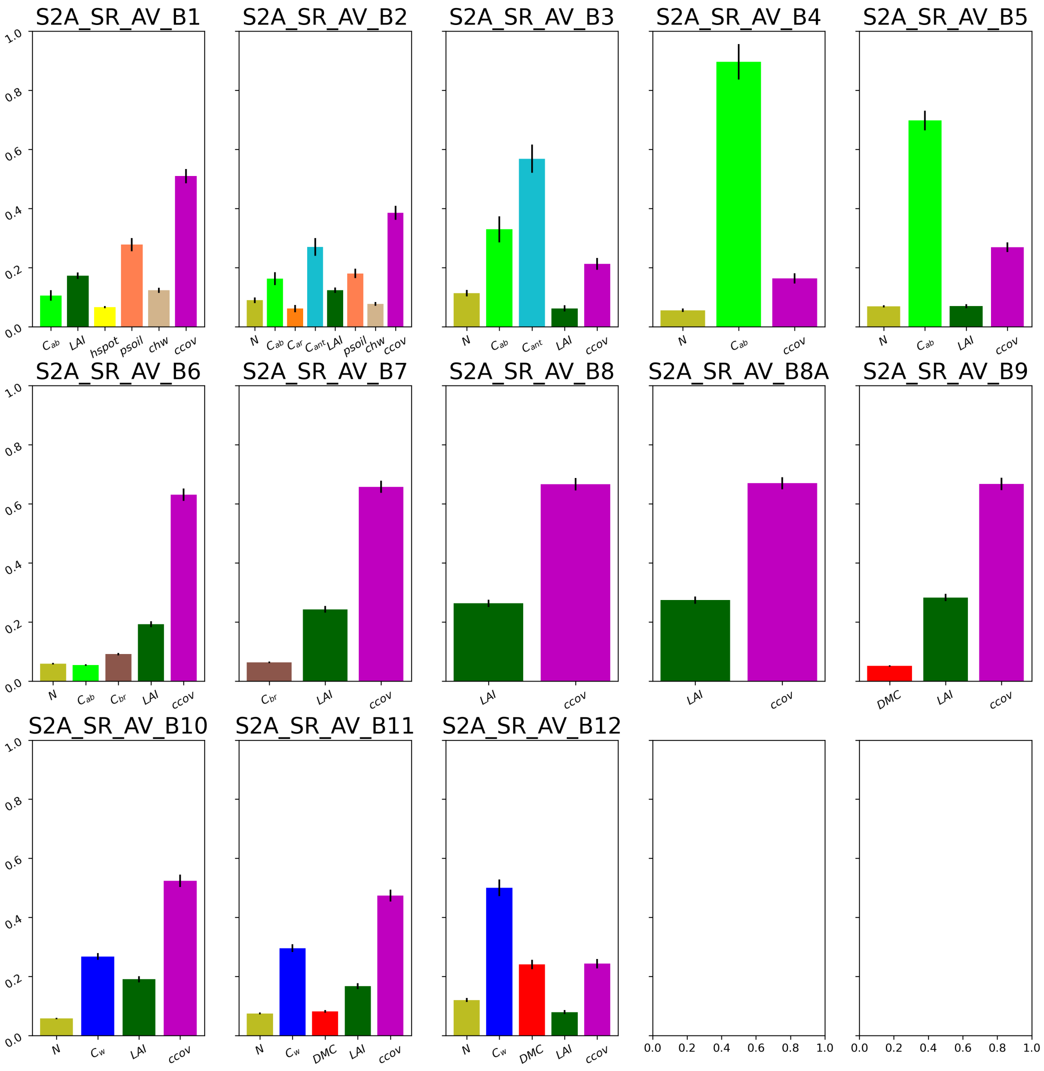

Figure 3.15. Total order sensitivity indices of the PROSPECT-D + SAIL + GEO parameters in the S2A S2MSI2A channels.

Similar results can be observed in the Sentinel-2A S2MSI2A channels, where the dry matter content shows up among the significant parameters even in the NIR band B9.

3.4 LUT-based Inversion

Since our objective is to reconstruct the combination of the model biophysical parameters from the measured reflectances, we face a typical example of remote sensing inverse problem. In this particular domain, the PROSAIL family of radiative transfer models has been the subject of an intense scientific effort thanks to its relatively small number of input variables and the high interest they raise for vegetation monitoring purposes [75].

Let us consider a generic physical model m of an arbitrary number of variables $x_i, i = 1, 2, ..., k$. The dependent variable, which in our case is a reflectance, can be represented as follows:

$$R = M(x) + \varepsilon \tag{3.23}$$

Where ε represents the residual error between the simulated and the observed variables. Equation 3.23 is the compact form of the set of n equations:

$$\Big\{ R_i = m_i(x_1, x_2, ..., x_k) + \varepsilon_i \quad i = 1, 2, ..., n \tag{3.24}$$

According to [75], there are two categories of inversion methods applied to the inverse problem of a nonlinear model such as PROSAIL. Both are fundamentally based on the minimization of a cost function δ^2 that can be written as the sum of the summands:

$$\delta^2 = \sum_{i=1}^{n} \left[\frac{R_i - m_i(x_1, x_2, ..., x_k)}{\sigma_{Ri}} \right]^2 + \sum_{j=1}^{k} \left[\frac{x_j - x_j^{prior}}{\sigma_{xj}} \right]^2 \tag{3.25}$$

The first summand represents the difference between the observed and simulated output, while the second represents the difference between the input variables to estimate and the *a priori* knowledge available at the time of the inversion. It is clear that in order to obtain the optimal inversion result one should attempt to minimize both summands. The divisors σ_{Ri} and σ_{xj} in equation 3.25 represent the uncertainties on the observations and on the model, and the uncertainty on the *a priori* information respectively. Since their values are usually unknown, minimization of the cost function must rely on some assumptions, with the drawback of possibly returning solutions associated to local minima.

The two approaches to the inversion problem distinguish themselves on the prioritization of the first or second summand of equation 3.25:

- Approaches that focus on the minimization of the first summand, i.e. on the observation space, define the solution searching for the minimum difference between the observed and the simulated values. Look-up tables, iterative optimization and Monte Carlo methods are typical examples of this approach.

- The approaches that focus on the second term are based on the definition of a parametric function or model which was previously trained to reproduce a set of outputs from a set of inputs. An example of these methods are artificial neural networks and vegetation index-based methods.

Even confining the research to the PROSAIL family of RTMs, there exist many examples of implementations based on both approaches. Machine learning methods saw a surge of interest in the past decade for their ease of implementation and their computational speed after the training phase, but their applicability is limited on one hand from the relative scarcity of LFMC measurements and their uneven distribution around the globe [145], and on the other, for the specific case of Sentinel-3 SYNERGY, by the lack of publicly available datasets from 2019 to today.

For this reason, we focus on the approaches based on the observation space, and in particular to those based on LUTs. Mathematically, LUTs are among the simplest inversion methods [7], and consist in sampling the input parameters of the model according to their sensitivity and then using them to create a table of simulated observations. The actual observations are then compared to this pre-computed set. The assumption is that the inputs that generated the simulated observations which are more similar to the observed ones are the best guess of the actual input parameters, and therefore the optimal solution of the inverse problem. The advantage of using LUTs, in addition to their relative simplicity compared to other inversion methods, is their capability to explore the global solution space if the parameter space is sufficiently sampled, avoiding being trapped in local minima [7]. Examples of successful utilization of LUTs for the inversion of biophysical variables from satellite data are plenty [29, 136, 79, 138, 58], and despite a recent surge in the interest towards machine learning methods [148, 140, 34], their popularity remains strong even for the specific case of LFMC and PROSAIL models [143, 141, 142, 104, 105].

3.4.1 LUT Parametrization Criteria and Calculation Procedure

It is well-known that the inversion problem is under-determined, due to the fact that there are more unknowns than measured variables, and ill-posed, due to the fact that very similar spectra may result from different combination of input variables jeopardizing the uniqueness of the solution [7]. For these reasons, the LUT parametrization process, i.e. the process of selecting the combinations of input variables for the target RTM, is of paramount importance for the success of the inversion. To this end, [105] notes that biophysical variables are not independent of each other, and that unlikely combinations may aggravate the ill-posedness of the inversion problem by producing spectra similar to others that result from realistic combinations. To solve this issue, the authors define a set of "ecological rules" that allow to discard these unlikely combinations before calculating the LUT.

The ecological rules are based on measured relationships between two biophysical parameters. The authors exploited the globeLFMC database [145] for LFMC samples, the LOPEX1993 [67] and ANGERS2003 [74] database for EWT, DMC and C_{ab} samples, and the MODIS product MCD15A3H [94] for LAI estimates to verify the correlation of different combinations of these variables on different land cover types. The strongest correlations, quantified in terms of correlation coefficient r, were found for:

- LAI and LFMC, $r = 0.71$ on grasslands
- C_{ab} and C_{dm}, $r = 0.58$ on grasslands
- C_w and C_{dm}, $r = 0.66$ on forests
- C_{ab} and C_w, $r = 0.52$ on forests

Weaker correlations were found for LAI and FMC on forests and C_{ab} and C_w on shrublands.

Consequently, we used only the strongest relationships in the compilation process of our LUTs, which we will now describe step-by-step. It should be noted that, while [105] contains a description of the definition and application of the ecological rules, it does not contain the LUT calculation algorithm, and in general, finding a complete description of such an algorithm is quite rare in the adjacent literature. For this reason, we will now describe our LUT calculation algorithm citing the sources and noting our additions to the literature where necessary.

A LUT is generated at each passage of the target sensor, in order to use the SZA interval recorded by the sensor as *a priori* information for the LUT to reduce the number of stored variables, allowing us to better sample the observation space. This is the main improvement of our LUT calculation algorithm compared to the literature, since usually a single LUT is generated for each land cover allowing the SZA to vary in a the entire interval observed by the target sensor. After getting the SZA values from the satellite image, the LUT calculation process takes place as follows:

1. **Definition of the LUT target size**: we tested LUT sizes ranging from 10000 to 200000 samples, and found that the benefit of increasing the LUT size resulted in drastically diminishing returns above the 100000 samples mark. For this reason, we ultimately used a target sample size of 100000 samples.

2. **Definition of the LFMC minimum and maximum values**: we used the same minimum value of 20% and maximum value of 200% for all three land covers.

3. **Creation of first-guess distributions for each variable**: after setting a seed for the random number generator, we independently create distributions of values for each model parameters. The number of samples generated in this phase is 15 times the number of target LUT samples, in order to keep a wide margin of surviving samples after enforcing the ecological rules. Some model parameters are generated using Gaussian distributions and the mean and standard deviation values shown in [105]. The other parameter values are either set using simple uniform distributions between the minima and maxima shown in [105], or set to fixed values due to their low significance as measured by the GSA. The distribution type, minima and maxima for each parameter of the model and for each of the three target land cover types are shown in tables 3.9, 3.10, 3.11.

4. **Elimination of unlikely combinations**: for each couple of strongly correlated model parameters, we applied the selection procedure as shown in [105]. After removing all the unlikely combinations, if the remaining number of samples is lower than the target sample size, we return to step 1).

5. **Sample selection**: once we have an adequate number of samples, in order to match the LUT size to the target sample size, we choose among the available profiles those that generate an equal number of LFMC samples in a number of 10%-wide LFMC bins between the minimum and maximum target LFMC values. This ensures that we have the same amount of profiles in each of this bins, and therefore that we sample equally profiles that generate different ranges of LFMC. If the number of remaining profiles is insufficient to satisfy this requirement, we go back to step 1). Once the requirement is satisfied, we proceed to calculate the LUT using the appropriate model based on the land cover.

6. **LUT calculation**: the reflectance spectra outputted by the model are convolved using the Spectral Response Function (SRF) of the target sensor. The

Sentinel-3 S3A SRFs for the sensors OLCI and SLSTR are shown in figures 3.16 and 3.17 respectively, while the Sentinel-2 S2A SRFs for the sensor MSI are shown in figure 3.18.

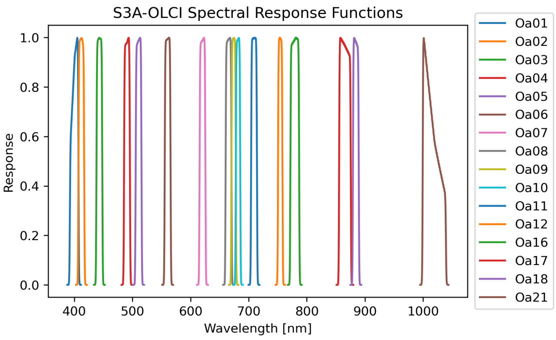

Figure 3.16. Sentinel-3A OLCI spectral response functions, available on the Copernicus website.

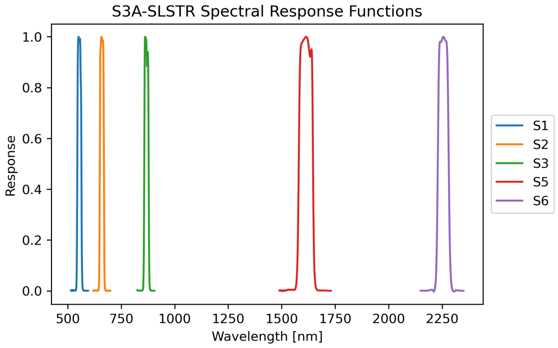

Figure 3.17. Sentinel-3A SLSTR spectral response functions, available on the Copernicus website.

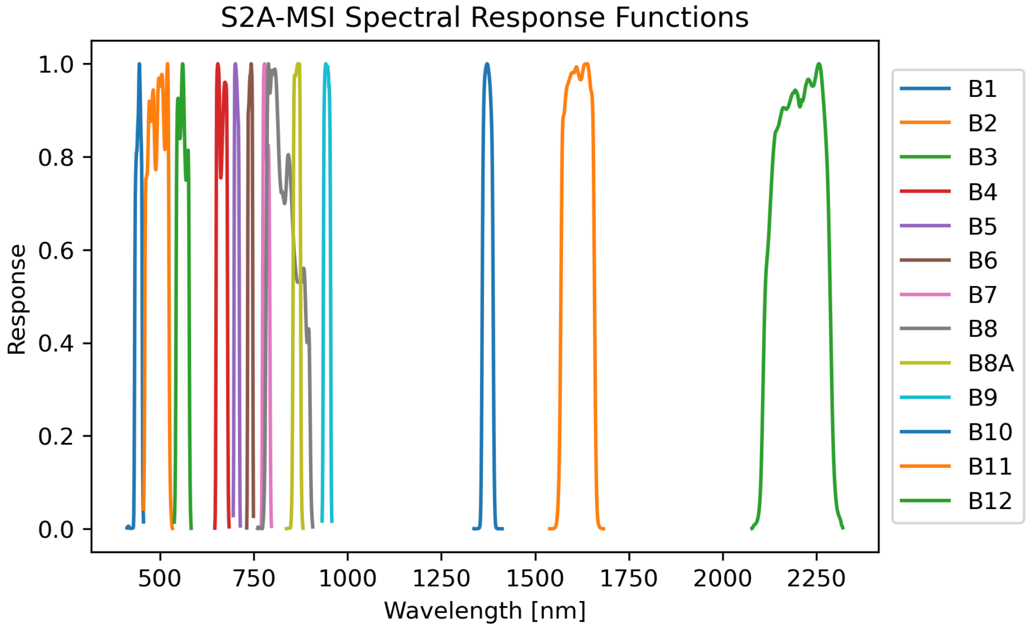

Figure 3.18. Sentinel-2A S2MSI spectral response functions, available on the Copernicus website.

Parameter	Unit	Distribution	Min-Max	μ, σ
PROSPECT-D				
N	-	Gaussian	$[1 - 3]$	$1.7, 0.32$
C_w	$g \cdot cm^{-2}$	Gaussian	$[0.0001 - 0.0360]$	$0.0131, 0.0071$
C_{dm}	$g \cdot cm^{-2}$	Gaussian	$[0.0017 - 0.0096]$	$0.0042, 0.0018$
C_{ab}	$\mu g \cdot cm^{-2}$	Gaussian	$[1 - 110]$	$43.5, 19.29$
C_{ar}	$\mu g \cdot cm^{-2}$	Constant	8	-
C_{ant}	$\mu g \cdot cm^{-2}$	Constant	20	-
C_{bp}	-	Constant	0	-
4SAIL				
LAI	m^2/m^2	Gaussian	$[0.5 - 7]$	$1.12, 1.21$
lidfa	-	-	Erect., spher., plan.	-
p_{soil}	-	Uniform	$[0 - 1]$	-
r_{soil}	-	Uniform	$[0 - 1]$	-
Hot	m/m	Constant	$[0.20]$	-
vza	deg	Constant	5	-
raa	deg	Constant	-30	-

Table 3.9. Description of the functions used to sample to input parameters of PROSPECT-D and 4SAIL to generate LUTs for the grassland land cover type.

Parameter	Unit	Distribution	Min-Max	μ, σ
PROSPECT-D				
N	-	Gaussian	$[1.27 - 3]$	$1.79, 0.36$
C_w	cm	Gaussian	$[0.0001 - 0.0520]$	$0.0110, 0.0061$
C_{dm}	$g \cdot cm^{-2}$	Gaussian	$[0.0017 - 0.0330]$	$0.0053, 0.0033$
C_{ab}	$\mu g \cdot cm^{-2}$	Gaussian	$[0.78 - 77.53]$	$35.37, 22.02$
C_{ar}	$\mu g \cdot cm^{-2}$	Constant	10	-
C_{ant}	$\mu g \cdot cm^{-2}$	Constant	20	-
C_{bp}	-	Constant	0	-
4SAIL				
LAI	m^2/m^2	Gaussian	$[0.5 - 7]$	$1.76, 1.56$
lidfa	-	Uniform	$[50 - 90]$	-
p_{soil}	-	Uniform	$[0 - 1]$	-
r_{soil}	-	Uniform	$[0 - 1]$	-
Hot	m/m	Constant	$[0.20]$	-
vza	deg	Constant	5	-
raa	deg	Constant	-30	-

Table 3.10. Description of the functions used to sample to input parameters of PROSPECT-D and 4SAIL to generate LUTs for the shrubland land cover type.

Parameter	**Unit**	**Distribution**	**Min-Max**	μ, σ
PROSPECT-D				
N	-	Gaussian	$[1.05 - 2.74]$	$1.54, 0.27$
C_w	$g \cdot cm^{-2}$	Gaussian	$[0.0001 - 0.0290]$	$0.0098, 0.0037$
C_{dm}	$g \cdot cm^{-2}$	Gaussian	$[0.0018 - 0.0189]$	$0.0052, 0.0027$
C_{ab}	$\mu g \cdot cm^{-2}$	Gaussian	$[1 - 107]$	$41.13, 20.63$
C_{ar}	$\mu g \cdot cm^{-2}$	Constant	10	-
C_{ant}	$\mu g \cdot cm^{-2}$	Constant	20	-
C_{bp}	-	Constant	0	-
4SAIL				
LAI	m^2/m^2	Gaussian	$[0.5 - 5]$	$1.67, 1.23$
lidfa	-	-	Plag., erect., spher.	-
p_{soil}	-	Uniform	$[0 - 1]$	-
r_{soil}	-	Uniform	$[0 - 1]$	-
Hot	m/m	Constant	0.20	-
vza	deg	Constant	5	-
raa	deg	Uniform	-30	-
GEO				
chw	-	Uniform	$[1 - 3]$	-
$ccov$	-	Uniform	$[0.2 - 1]$	-
$cshp$	-	Uniform	Cone, cylinder	-

Table 3.11. Description of the functions used to sample to input parameters of PROSPECT-D, 4SAIL and GEO to generate LUTs for the forest land cover type.

3.4.2 LUT Calculations and Small Improvements on the Established Literature

The same input LUTs are generated for a specific land cover type, but the reflectance values outputted by the model are then convolved using the spectral response functions of the target satellite sensors. This means that we will have the same input LUTs for grasslands, but different output LUTs for Sentinel-3A and Sentinel-3B, even though given the similarity of the SRFs the difference between the output spectra will be quantitatively small. The same can be said for the Sentinel-2A and Sentinel-2B bands.

Compared to the established literature, we exploited the fast LUT calculation times to generate different LUTs for different SZA intervals, rather than generating a single LUT for each land cover type and allowing the SZA to vary in the interval experienced by the sensor. In particular, during a Sentinel-3 acquisition the SZA varies by about ±8 degrees around a mean value, while for Sentinel-2 imagery the SZA only varies by small amounts, usually a little over 1 degree. Nevertheless, in both cases we calculate or read[1] the mean passage SZA, round it to the nearest integer and create a LUT associated to that SZA value, and then proceed slightly differently for the two satellite platforms:

- For each Sentinel-3 observation, we allow the SZA value to vary in an interval of ±8 degrees around the rounded mean passage value.
- For each Sentinel-2 observation, we simply set the SZA to the mean passage value.

This allowed us to significantly improve the inversion results, both on simulated data and on the field measurements, compared to using a single LUT as suggested

---

[1] In the Sentinel-2 S2MSI2A case, the mean passage SZA can be read directly using the provided `MTD_TL.xml` file in the `Mean_Sun_Angle` section. In the Sentinel-3 SYNERGY case, the value must be calculated from the one-dimensional netcdf field in the OLCI tiepoints file.

in the related literature. Nevertheless, rather than having three LUTs (i.e., one for each land cover type) for each satellite, we need to calculate new LUTs for each new mean passage SZA, significantly increasing the amount of storage space required for the LUTs.

3.4.3 Extraction of the Optimal Inversion Parameters

[105] contains an extensive analysis of the impact on the LFMC inversion performance based on a number of different parameters:

- Amount of spectral information used (satellite channels, vegetation indices)
- Amount of closest matching spectra among which to search the solution
- Cost Functions used to rank the potential solutions
- Measures of Central Tendency used to harmonize the potential solutions

Building on this foundation, we used field data provided by the Department of Mechanical Engineering, Association for the Development of Industrial Aerodynamics (ADAI) of the Coimbra University and the Institute for BioEconomy (IBE) of the Italian National Research Council (CNR) to test the impact of different inversion parameters on the performance of our inversion procedure. The optimal Sentinel-2 and Sentinel-3 bands to use for the inversion, were identified using simulated observations and the results of the GSA, differently than [105]. Subsequently, for each of the remaining inversion parameters, we varied only the target parameter and kept the others fixed and used the field data to test the effect on the accuracy of the retrieval. The results we obtained apply to the Sentinel-2 observations for the ADAI dataset and to both the Sentinel-2 and the Sentinel-3 observations for the CNR dataset. In the following paragraphs, we will give a brief description of the field sites, and then proceed to describe the results of the tests for each inversion parameter.

3.4.4 ADAI Dataset: Lousa Site

The ADAI dataset contains daily LFMC values for two tree species (*Pinus pinaster* and *Eucalyptus globulus*), and two shrub species (*Calluna vulgaris* and *Chamaespartium tridentatum*). The measurement campaign started in 1996 and continued uninterrupted until today. The measurements are gathered with different frequency depending on the time of the year: less frequently in the winter, and more frequently - even daily - during the fire season. In our case, we used the portion of this dataset that starts in October 2018 and ends in April 2022.

The sampling area shwon in figure 3.19 corresponds to a singular Sentinel-3 pixel, which is cut roughly in the middle by a road, which is surrounded by a mix of trees and shrubs. This implies that the Sentinel-3 pixel is not perfectly homogeneous, and indeed, we were unable to obtain acceptable LFMC inversions using this satellite platform. At the moment our model is not capable of mixing the spectra of trees and shrubs. Nevertheless, since both the Corine Land Cover and the ESA WorldCover classified this area as forested, we decided to focus on the two tree species, and tested the inversion results on the Sentinel-2 images.

After comparing a few estimated LFMC timeseries with the measured timeseries of the various species, it became apparent that our estimated progression matched more closely that of the *Pinus pinaster* than that of the *Eucalyptus globulus*. Even

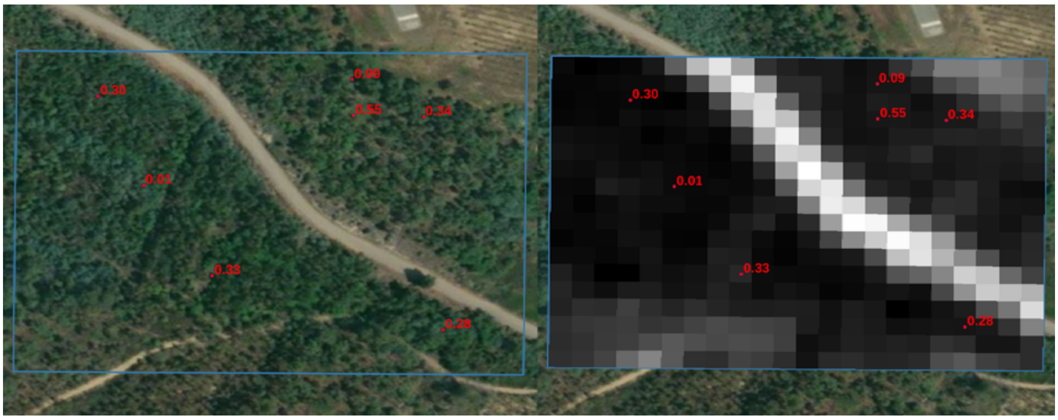

Figure 3.19. Bing Satellite map of the Lousa sampling site, roughly delimited by the blue rectangle. The red dots mark the Sentinel-2 pixels used for the validation, and the associated value represents the coefficient of determination (R^2) of the estimated timeseries with respect to the average dataset.

averaging the two measured timeseries, the correlation proved to be significantly worse than that with the sole *Pinus pinaster*. Therefore, we will only show the results pertaining to this species. The right side of figure 3.19 shows the Sentinel-2 pixels used for this validation and the associated coefficient of determination R^2, calculated by comparing the "predicted" LFMC values estimated with our procedure and the "true" values measured with by ADAI. For the sake of brevity, only the final results will be reported hereafter.

3.4.5 CNR Dataset: Capo Caccia Site

Capo Caccia is located in the northwestern part of Sardinia, and the sampling site shown in Figures 3.20 and 3.21 has been used by CNR since the early 2010s, with weekly or bi-weekly measurement campaigns. The local vegetation is dominated by shrubs, namely *Juniperus phoenicea*, with a smaller but significant presence of *Pistacia lentiscus* and *Phillyrea angustifolia*. Other species worth mentioning are *Cistus monspeliensis* and *Rosmarinus officinalis*, which are more sensitive to seasonal changes, especially during the summer.

The dataset provided by CNR contains the raw measurements of the dominant species, i.e. the *Juniperus phoenicea*, and a timeseries of the average of all the aforementioned species, *Juniperus phoenicea* included.

Amount of Spectral Information

The results of the simulations showed that the highest number of variables could be retrieved with an adequate coefficient of determination R^2 if all the bands were used for the retrieval, but, as expected, this approach was not optimal for the retrieval of LFMC. This result was shown in [97] using an early version of the LUTs. Rather than progressively reducing the number of bands, we started from the two most sensitive bands to C_w, ($SLSTR\ S_5$ and $SLSTR\ S_6$ for Sentinel-3, B_{11} and B_{12} for Sentinel-2) and from the most sensitive band to C_{dm} ($SLSTR\ S_6$ for Sentinel-3, B_{12} for Sentinel-2) and progressively added other bands to verify the effect on the LFMC retrieval using the coefficient of determination and the RMSE as figures of merit. The LFMC retrieval accuracy increases if bands containing

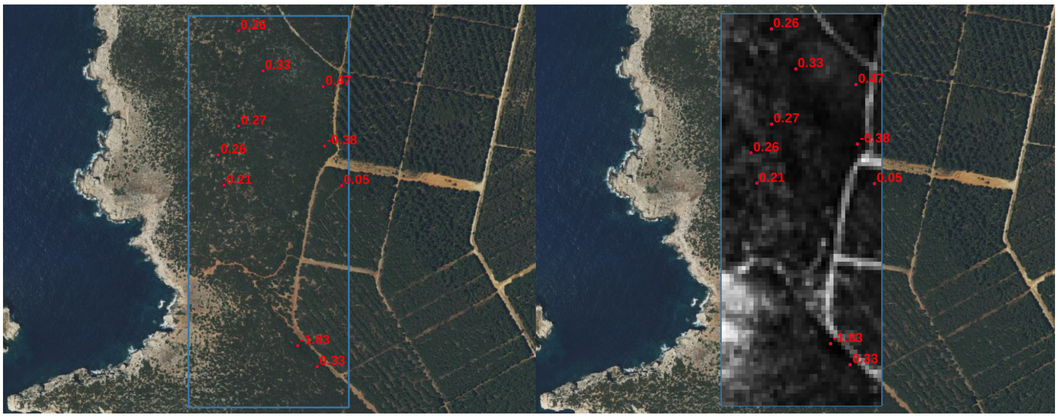

Figure 3.20. Bing Satellite map of the Capo Caccia sampling site, roughly delimited by the blue rectangle. The red dots mark the Sentinel-2 pixels used for the validation, and the associated value represents the coefficient of determination (R^2) of the estimated timeseries with respect to the average dataset.

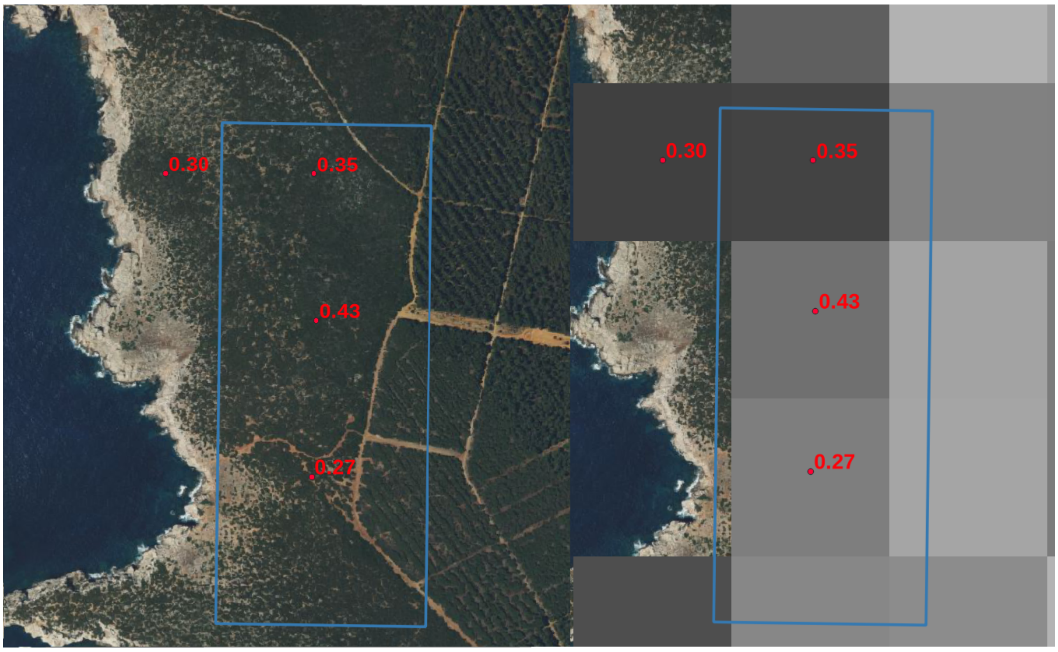

Figure 3.21. Bing Satellite map of the Lousa sampling site, roughly delimited by the blue rectangle. The red dots mark the Sentinel-3 pixels used for the validation, and the associated value represents the coefficient of determination (R^2) of the estimated timeseries with respect to the average dataset.

information about the chlorophyll content, anthocyanin content and LAI are added, such as $OLCI\ Oa_{06}$ or $SLSTR\ S_3$ for Sentinel-3, and B_2 and B_8 for Sentinel-2. The addition of further bands, either brought no significant improvement or significantly worsened the performance.

Nevertheless, the LFMC retrieval for Sentinel-3 was only slightly improved using $OLCI\ Oa_{06}$, and given the small performance difference an argument could be made for using only SLSTR to invert the LFMC, which may pave the way for near-real-time LFMC products based on the Sentinel-3 Level 1 product family.

Table 3.12 contains the list of channels used for the inversion of the LFMC for each satellite platform.

Platform	Product	Channels
Sentinel-2	$S2MSI2A$	2, 8, 11, 12
Sentinel-3	SY_2_SYN	Oa06/SYN06, S3N/SYN19, S5N/SYN20, S6N/SYN21

Table 3.12. List of the best performing channels of the Sentinel-2 S2MSI2A and Sentinel-3 SY_2_SYN for the inversion of the LFMC.

3.4.6 Cost Function

Among the many cost functions tested in [105], we only compared the performance of the Root Mean Square Error function ($RMSE_f$), the exponential function, and the Bhattacharya Divergence. Table 3.13, where we denote the measured reflectances as v_i, the reflectances stored in the LUT as w_i and the number of involved channels as n contains the results of our tests.

Cost Function	Formula	ADAI R^2	CNR R^2
RMSE	$\sqrt{\sum_{i=1}^{n} \frac{(v_i - x_i)^2}{n}}$	0.47	0.76
Bhattacharya Divergence	$-\ln 1 + \sum_{i=1}^{n} \left(\sqrt{v_i w_i} - \frac{v_i + w_i}{2} \right)$	0.19	0.74
Exponential	$\sum_{i=1}^{n} \lvert w_i \left(\exp - \frac{v_i - w_i}{v_i} \right) \rvert$	0.53	0.74

Table 3.13. LFMC retrieval performance with different cost functions on the ADAI and CNR dataset.

The RMSE and the Bhattacharya divergence show very similar performance, while the exponential function performs slightly worse in Capo Caccia and slightly better in Lousa. Even on a point-by-point basis, the exponential function and RMSE tend to give rather similar results, making it difficult to identify a best performer. This is not surprising, since even in [105] the two functions show comparable performance. In addition, we should remark that our analysis is limited to two sites, and while the variation of the inversion performance tends to agree in almost all instances, more data is needed to draw definitive conclusions.

3.4.7 Number of Selected Solutions

Let us denote the number of selected solutions with N: for each N, we will calculate the cost function value for each LUT entry, and sort these values from smallest to largest. The spectra associated to the smallest N values will be the only ones considered during the extraction of the solution. In [105], the LFMC retrieval performance worsens gradually after the 1% mark, implying that the bottom 1% cost function values were used for the inversion. The absolute number of LUT samples is not specified. In our case, the performance of the inversion worsens after the 30-40 mark ($0.03 - 0.04\%$), while the optimum seems to always be located between 15 and 25 ($0.015 - 0.025\%$) closest spectra. While both cases show that using a relatively small amount of solutions is optimal, it is difficult to explain the large difference between the percentages since we are unable to compare the LUT compilation algorithms directly.

N	ADAI R^2	CNR R^2
20	0.55	0.76
50	0.55	0.76
100	0.48	0.75

Table 3.14. LFMC retrieval performance with an increasing number of selected solutions. The N selected solutions are those that minimize the chosen cost function.

3.4.8 Measures of Central Tendency

Once a certain number of solutions are selected from the LUT, we need to harmonize the results in order to associate a unique solution to the measured spectrum. Usually, this is done by applying a certain measure of central tendency such as the arithmetic or geometric mean, the median or the mode, to the selected solutions. Compared to [105], we only tested the arithmetic mean and median. In the referenced article, the performance differences between the six tested functions are marginal for grasslands, and more significant for shrublands and forests. The median seems to work best overall, and in particular for forests and shrublands. Our tests, summarized in table 3.15 confirm this result.

Function	ADAI R^2	CNR R^2
mean	0.27	0.55
median	0.48	0.76

Table 3.15. LFMC retrieval performance using different measures of central tendency to harmonize the selected solutions.

3.4.9 Final Inversion Parameters

Following the tests in paragraphs 3.4.6, 3.4.7 and 3.4.8, we selected the inversion parameters summarized in table 3.16 to generate the LFMC datasets which will be used to calculate the fire danger index.

Inversion Parameter	Value
Number of Selected Solutions	20
Cost Function	RMSE
Measure of Central Tendency	median

Table 3.16. Inversion parameters used to generate the LFMC estimates.

3.5 Early Validation of the Inversion Procedure

In this paragraph we will show the results of the validation using the optimal inversion parameters listed in paragraph 3.4.9. Before comparing them with the field data, the estimated timeseries extracted using Sentinel-2 or Sentinel-3 data are smoothed using a i algorithm with a 7-sample window.

3.5.1 Sentinel-2 Product

ADAI Dataset

As mentioned in 3.4.4, the estimated timeseries follow rather closely the *Pinus pinaster* measurements, while a significant bias persists if we compare the estimated timeseries to the *Eucalyptus globulus* measurements and to a lesser extent to the average of *Pinus pinaster* and *Eucalyptus globulus* measurements. Therefore, we used the *Pinus pinaster* both for identifying the optimal inversion parameters and the validation. The 7 Sentinel-2 points used for the validation are shown in Figure 3.19, and were selected to avoid road pixels and mixed pixels and privileging more densely tree-covered areas in order to minimize spurious contributions to the reflectance.

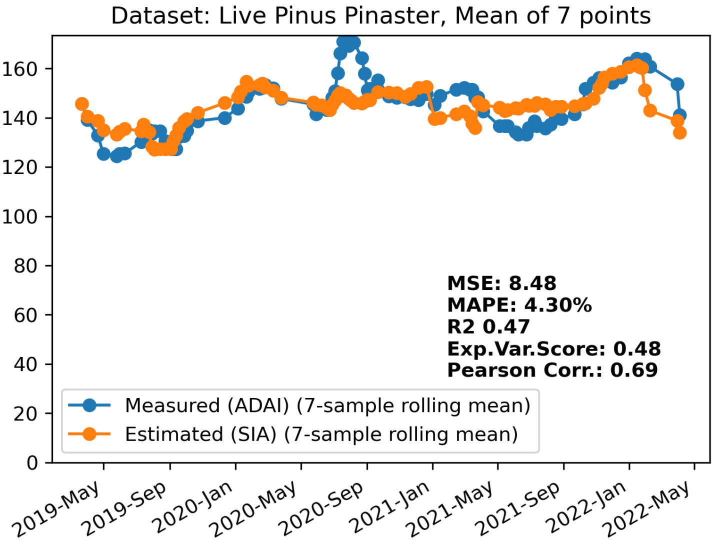

Figure 3.22. Average of the 7-day rolling mean timeseries of the Sentinel-2 LFMC product shown in figure 3.19 (orange) and ADAI *Pinus pinaster* measurements (blue).

If we compare the individual pixel timeseries with the *Pinus pinaster* measurements the best performing pixel reaches a coefficient of determination R^2 of 0.55, while the worse performing shows an R^2 of 0.01. If we average the timeseries of the seven pixels, i.e. if we generate another timeseries by aggregating the daily mean of the LFMC estimates, we obtain the result shown in figure 3.19, with a satisfying R^2 of 0.47: this value, while not higher than the 0.55 of the individual best performer, is significantly higher than all the other six measurement points.

CNR Dataset

In this case, the LFMC estimated timeseries follows more closely the measured dataset of the average of all the sampled species. The higher homogeneity and larger size of the sample site compared to the Lousa one allowed us to obtain an extremely high coefficient of determination with respect to the values usually found in the literature related to satellite-based inversions [105].

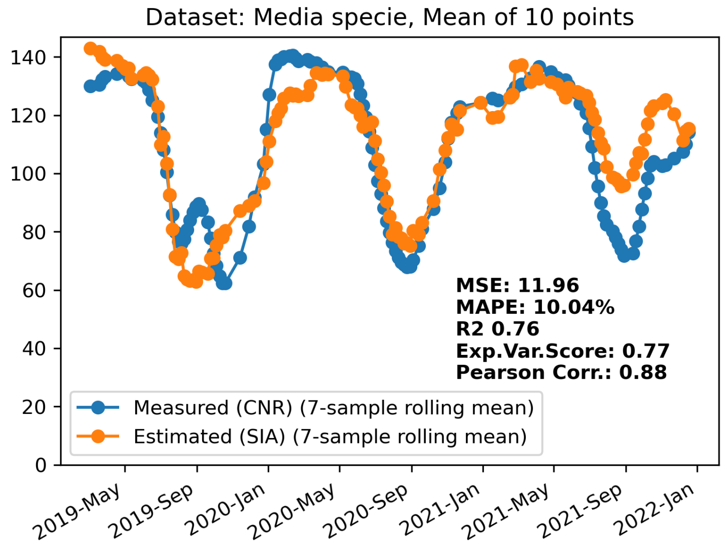

Figure 3.23. Average of the 7-day rolling mean timeseries of the Sentinel-2 LFMC product shown in figure 3.20 (orange) and CNR average species measurements (blue).

Figure 3.23 shows the progression of the average of the estimated timeseries of the ten Sentinel-2 points shown in figure 3.20. The coefficient of determination is much higher than that of the individual point timeseries, reaching a value of 0.76.

3.5.2 Sentinel-3 Product

Successful inversion of the LFMC product based on Sentinel-3 imagery was only achieved on the CNR Capo Caccia dataset. The homogeneity and large size of this site even with respect to the Sentinel-3 300-meter pixel size allowed us to obtain good correlation: if we compare the average timeseries extracted from the four Sentinel-3 pixels that overlap the sampling area, shown in figure 3.21 we obtain an R^2 of 0.42.

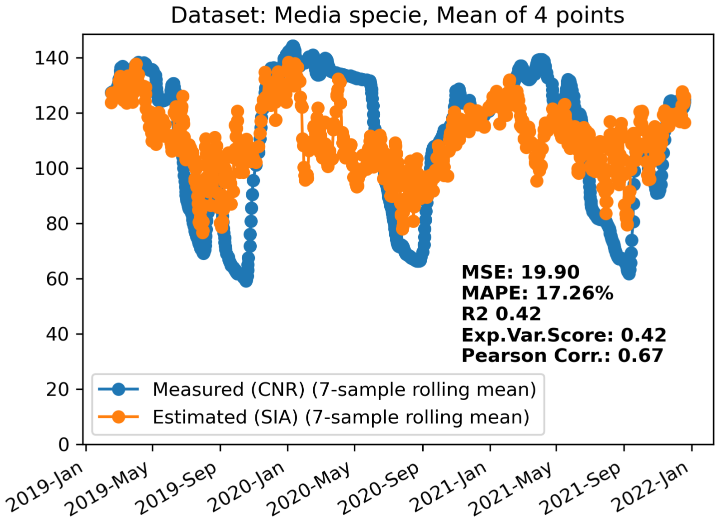

Figure 3.24. Average of the 7-day rolling mean timeseries of the Sentinel-2 LFMC product shown in figure 3.20 (orange) and CNR average species measurements (blue).

Individually, the four sampled pixels return coefficients of determination that range from 0.27 to 0.43. In general, the R^2 and RMSE values shown in figure 3.24 are comparable with the ones found in [105] for the shrubland land cover.

Chapter 4

The Fire Danger Index

In the previous chapters we have identified the biophysical and meteorological variables relevant to fire danger, and we have described the estimation procedure of each of them. In order to provide an easy to interpret fire danger index, these variables must be properly normalized and combined together, and finally translated into categorical danger ratings.

4.1 State of the Art of Satellite-based Fire Danger Indices

The scientific literature that constitutes the base of this index goes from the FPI to the DFHI. The FPI model was rather simple: it was defined as the product of the normalized versions of two main variables: the relative greenness, which provided an estimate of the amount of live vegetation, and at the same time was used as a proxy for live fuel moisture content, and the timelag fuel moisture, which provided an estimate on the dryness level of the dead vegetation layer [15].

$$FPI_u = 100 - (RG_f\ L_f + TN_f\ D_f) * 100 \tag{4.1}$$

Equation 4.1 shows the "uncorrected" FPI equation, where RG_f is the relative greenness fraction we mentioned in equation 2.14, TN_f is the fractional ten-hour timelag fuel moisture calculated in equation 2.21 and D_f represents the dead fuel load. The FPI used a fuel model both to calculate the fractional ten-hour timelag fuel moisture and to derive the live and dead fuel loads. As mentioned in the previous chapters, we remark that Burgan himself, in the introduction to his 1993 article where he introduced the relative greenness, stated that at the time the technology to assess live fuel moisture content from satellite observations was not ready. For this reason, the FPI relies on the relative greenness as a proxy for vegetation moisture.

The DFHI built upon the foundations of the FPI by correcting the basic model variables to include the effects of the topography, wind speed and equivalent water thickness in the fire danger assessment [84]. In particular, the DFHI exploited Ceccato's Global Vegetation Moisture Index to estimate the equivalent water thickness from two MODIS surface reflectance bands. While this was not yet a proper live fuel moisture content assessment for all the reasons stated in paragraph 1.3, it proved to be a very simple way to monitor the vegetation water content without significantly altering the computational load required to calculate the index.

These additions to the FPI allowed to improve the performance of the fire danger index, but did not deviate significantly from the original two-variable model.

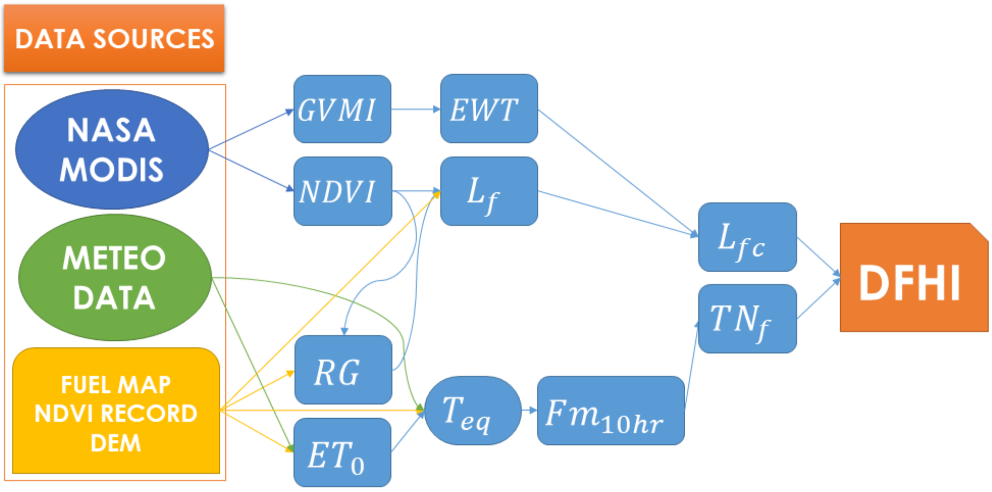

Figure 4.1. Conceptual map of the DFHI algorithm, sourced from [84]. The scheme shows how the basic satellite surface reflectances, meteorological variables and static data are used to calculate the two final FPI variables. In particular, the EWT and the ET_0 are used to correct the live vegetation fraction L_f and the temperature T_{eq} used to calculate the ten-hour timelag fuel moisture FM_{10hr} respectively.

4.2 Objectives of the New Index

In this thesis, we propose a novel fire danger index based on satellite imagery. We aim to build upon the foundations provided by the FPI and DFHI by:

- Proposing an additive fire danger model, in order to overcome the limitations of the two-variable model established by the FPI. This will allow us to take into account more directly the role of variables such as the wind speed and live fuel moisture content.
- For the first time, exploit the surface reflectances provided by the target satellite platforms to estimate both the vegetation greenness and the live fuel moisture content.

In the following paragraphs we will describe the procedure that allows us to obtain the DFDI values from the biophysical variables that determine fire danger, and how the values are transformed in a fire danger scale. Finally, we will validate the index using historical wildfire occurrence data in the study area.

4.3 The DFDI Algorithm

Since the DFDI is a fire danger index, each member of the DFDI equation must represent a danger variable. This implies that, while some variables require not only normalization, but also their rearrangement into an equivalent danger variable. For instance, high LFMC values are associated to a lower fire danger. For this reason, we will now use the $\overline{VAR}$ notation to refer to the normalized version of the original biophysical variables. Those variables that require only normalization will therefore appear as $\overline{VAR}$ in the DFDI equation, while those which require both normalization and "inversion" will appear as the one's complement of their normalized version $(1-\overline{VAR})$.

DFDI values on a scale from 0 to 100 can be obtained using the following formula:

$$DFDI = f_1\ (1-\overline{LFMC}) + f_2\ (1-L_f) + f_3\ (1-TN_f) + f_4\ \overline{w} + f_5\ \overline{Et_0} \tag{4.2}$$

where:

- $\overline{LFMC}$ is the historical minmax-normalized LFMC
- L_f is the fraction of live vegetation as calculated in 2.15
- TN_f is the fractional moisture content of the dead vegetation as calculated in 2.21
- $\overline{w}$ is the historical minmax-normalized wind speed
- $\overline{Et_0}$ is the historical minmax-normalized reference evapotraspiration
- f_i are the weights of each parameter

In the following paragraphs we will now proceed to describe the normalization process of each variable and how the weights were obtained.

4.3.1 Variable Normalization

Unsurprisingly, given the fact that they were the two factors used to obtain the DFHI, the fraction of live vegetation L_f and the fraction of 10-hour timelag fuel moisture TN_f are already normalized. However, compared with [84], the definition of the L_f is slightly different in order to obtain values between 0 and 1. Furthermore, the DFHI used to correct these two variables using the equivalent water thickness and the evapotranspiration, respectively. This allowed to indirectly include the effects of changes in vegetation moisture and evapotranspiration through simple means, and without altering the FPI model.

In our case, estimating the live fuel moisture content from the satellite data allowed us to include its effect directly in the index. Furthermore, since it is well-known that wildfires can ignite and propagate even when only some of the danger variables overcome danger thresholds, we decided to model their contribution separately and then aggregate them using a simple sum. This allows us to avoid the issue of previous indices that were calculated as a product of different variables, where only one factor returning very low values would necessarily translate into very low final danger ratings. In our case, even if one or two variables are associated to low danger, the final danger rating will not necessarily be very low thanks to the contribution of the other variables. As usual, proper normalization of the danger variables, so that their values range from 0 when their values pose the relative minimum danger, and 1 when danger is the highest, is necessary. In addition, we take into account that not all variables contribute to danger the same way, also depending on the geographical location of the area of interest, and therefore proper weighting of the individual danger factors is needed.

Live Fuel Moisture Content Normalization

Higher fuel moisture content values are associated to lower fire danger. However, different vegetation types may vary their LFMC in significantly narrower or larger ranges [145]. Therefore, live fuel moisture content values are normalized using the historical recorded minima and maxima of each pixel of the AOI:

$$\overline{LFMC} = \frac{LFMC - LFMC_{min}}{LFMC_{max} - LFMC_{min}} \tag{4.3}$$

This returns values between 0 and 1, where 0 corresponds to the historical minimum (driest recorded day) and 1 to the historical maximum (moistest recorded day). The live fuel moisture content related danger rating should be higher when the vegetation is relatively dry, and therefore we use its one's complement $(1 - \overline{LFMC})$ in the index equation.

Fraction of Live Vegetation

The live vegetation fraction is already a normalized version of the relative greenness. Since the associated danger rating should be higher when a lower amount of live vegetation than usual is present, we use its one's complement $(1 - L_f)$ in the index equation. This parameter is sometimes called fraction of dead vegetation.

Fraction of 10-hour timelag Fuel Moisture

Higher values of fraction of 10-hour timelag fuel moisture are associated to a lower fire danger, given the fact that the vegetation contains more water and is consequently harder to ignite. For this reason, we use its one's complement $(1 - TN_f)$.

Wind Speed

Horizontal and vertical wind speed values u and v are used to calculate the absolute wind speed value w. This value is then normalized using the historical records for the given area of interest:

$$\overline{w} = \frac{w - w_{min}}{w_{max} - w_{min}} \tag{4.4}$$

Since higher wind speeds are associated to higher fire danger, this value is directly used in the DFDI equation.

Evapotranspiration

Hourly evapotranspiration values are normalized using the historical recorded values for the area of interest:

$$\overline{Et_0} = \frac{Et_0 - Et_{0min}}{Et_{0max} - Et_{0min}} \tag{4.5}$$

Since higher evapotranspiration values are associated to a greater loss of water, and therefore with a higher fire danger, this value can directly be used in the DFDI equation.

4.3.2 Weighting the DFDI Parameters

Even though from a qualitative point of view all the considered variables contribute to determining fire danger, their quantitative contribution will not necessarily be the same. But how to quantify the relative contribution to fire danger of each variable without having another measure of fire danger? In order to solve this issue without relying on another existing fire danger index, we decided to use the Fire Radiative Power (FRP) estimates provided by FIRMS in the entire Sentinel-3 availability timeframe. The rationale behind this choice relies on the assumption that a well performing fire danger index should be able to indicate where fires with a higher FRP are more likely to occur. Therefore, we may use historical FRP estimates and the associated danger conditions to assess which variables contribute the most on average.

After collecting the fire danger variables associated to each FIRMS record's date and geographical position, we used `Scikit-learn`'s [11] Random Forest regressor to fit them to the FRP values. Random forest methods are routinely used in classification and regression tasks. They are based on a random aggregation of a high number of decision trees, and the algorithm output is obtained through majority voting for random forest classifiers, and as the average output of the decision trees for the random forest regressors. The randomization allows to overcome the proneness to overfitting associated to simple decision trees, and `Scikit-learn`'s tool offers a number of options to further reduce this possibility. Furthermore, this tool also allows us to quickly obtain the feature importance of each feature fed into it as part of the training set choosing from a number of importance measures. Among these, the Gini importance, also known as Mean Decrease in Impurity (MDI) quantifies variable importance based on the number of node splits that include a certain feature and the number of split samples, averaging over all the forest trees. Compared to other importance metrics, it is known for favouring categorical variables over numeric ones, due to the fact that the latter can be split over more trees, and for its lower capability to distinguish relevant variables when they are strongly correlated. While no categorical variables are involved in the definition of our index, avoiding the inclusion of strongly correlated variables is an utmost priority.

This method of weighting the fire danger variables also allows us to calibrate the index on a specific region: in some areas, fire occurrence may be more frequently linked to strong winds than others where fires are more driven by the accumulation of dead fuel or by intense moisture deficit in the live vegetation. The results for the AOI and for the Portugal Coimbra region are detailed in table 4.1.

Features **Region**	LFMC (f_1)	L_f (f_2)	TN_f (f_3)	w (f_4)	ET_0 (f_5)	R^2
Sardinia	18	33	15	20	14	0.66
Coimbra	17	20	16	27	20	0.87

Table 4.1. Gini importance [%] of the fire danger variables and coefficient of determination R^2 of the random forest regressor in the Sardinia region in Italy and in the Coimbra region in Portugal.

Running the same algorithm in Sardinia and Portugal returns significant differences: while in Sardinia the most discriminating feature is the fraction of live vegetation by a wide margin, in Portugal the wind speed is the most important, with the fraction of live vegetation and evapotranspiration tied as the second most important. Concerning the Sardinian case, the results match the description of the wildfire history in the region given by Salis [111]: the largest fires in the region are usually associated to strong winds and with an accumulation of dead fuels, explaining the supremacy of the two aforementioned factors.

4.4 Calibration and Performance of the DFDI

The definition of the fire danger index given in 4.2 is based on a weighted sum of fire weather variables, normalized using the historical records of a given area. For this reason, it will be very unlikely to find very low index values, which would be associated to the variables concurrently assuming values close to the historical low, and very high values, which would be associated to the variables concurrently assuming values close to the historical high. Therefore, the index should be calibrated accordingly, i.e. in order to provide fire danger ratings consistent with the seasonal distribution of its values. To this end, the histogram of the DFDI values calculated in the entire Sentinel-3 timeframe shown in figure 4.2 confirms our previous assumption: an insignificant amount of pixels are associated to danger values between 0 and 10 and above 80. As a consequence, the thresholds defined to transform the index values to easily interpretable fire danger categories will be differently sized, and the associated thresholds will necessarily be concentrated in the $20-70$ interval.

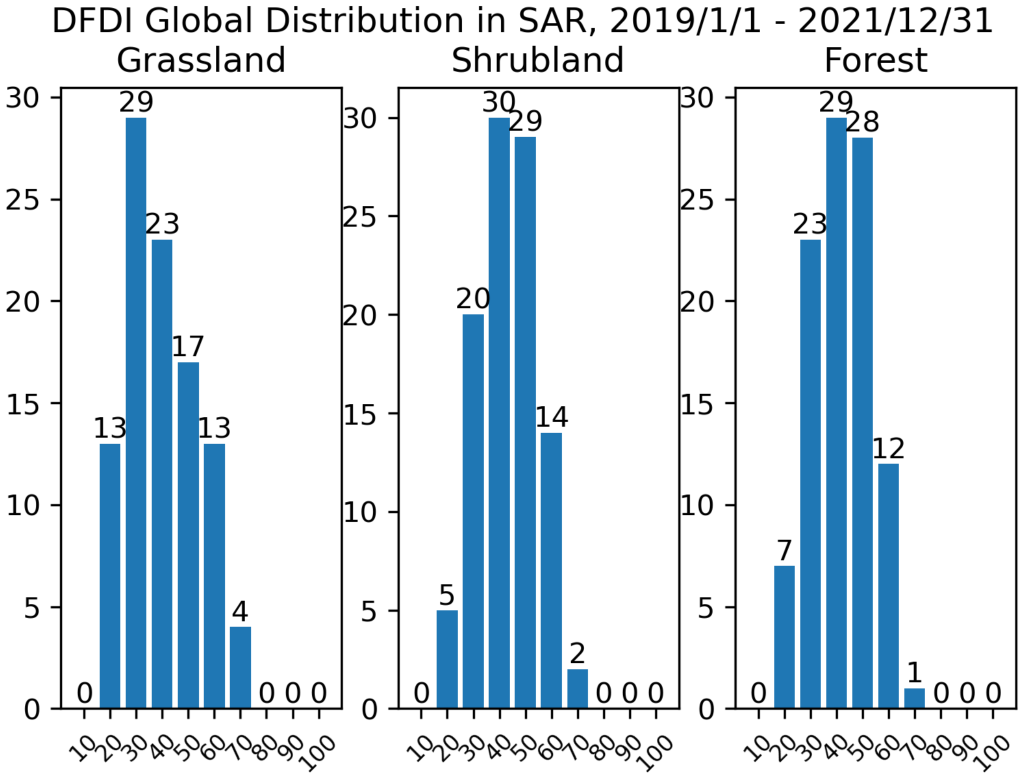

Figure 4.2. Frequency histogram of the DFDI values in Sardinia by land cover in the timeframe 2019/1/1 - 2021/12/31. Values indicate the percentage of pixels falling in the interval that ranges from the previous tick value to the tick value centered on the bar.

Ideally, a useful fire danger index should have the following characteristics:

A) The index should be able to recognize high fire danger areas in a given region without indiscriminately returning high danger values in an excessively large portion of the territory: if on a given day the fire danger is high in too large a portion of the managed territory, the index will provide no useful information for allocating firefighting resources. If this happens regularly during the fire season, the index will be of no use to the firefighting community.

B) The index should be capable to recognize the areas where fires are likely not only to ignite, but to spread and burn large patches of territory: therefore, large fires and burnt areas should be regularly associated to high fire danger values. Even if the index satisfies the previous requirement, i.e. if it is calibrated in such a way that it routinely recognizes a manageable amount of high danger areas, its ability to predict fire danger will be reduced if burnt patches and large fires tend to occur often in areas classified as low danger.

Therefore, a fire danger index should be calibrated in order to satisfy requirement A), while its performance will be evaluated using requirement B). Difficulties in satisfying both requirements, or insufficient performance are indicative of fundamental issues in the definition of the index. For instance, a fire danger distribution that could be deemed acceptable according to criterion A), may perform badly according to criterion B). Therefore, the calibration of the index can be performed iteratively as follows:

1. Define index thresholds for five danger classes: very low, low, medium, high, extreme
2. Test the distribution of the index on the fire season and off the fire season
3. If the distribution is not satisfactory according to criterion A), (i.e., too many pixels classified as high or extreme danger at once) return to step 1)
4. Test the performance of the index using FIRMS FRP and burnt area data
5. If the performance is not satisfactory according to criterion B), (i.e., too many pixels classified as low or very low danger on hotspots and burnt areas) return to step 1)

The index was tested in the period 2019-2021, which marks the availability timeframe of Sentinel-3 imagery and at the same time allows us to use validated FRP and burnt area datasets for our own validation. The occurrences of FIRMS hotspots with FRP higher than 10 MW by day of the year were counted in order to assess the beginning and ending of the fire season in the test area of Sardinia. This value was chosen in order to exclude from the analysis smaller fires, which may still be detected by the FIRMS sensors, but may be associated to controlled burns or self-extinguishing small fires. The definition of fire danger entails both the ignition and the propagation of the fire, and dangerous propagation requires an environment that favours the spread of fire. Ignition, on the other hand, is often of anthropogenic nature, especially in Mediterranean countries. For all these reasons, we deemed necessary to focus on the higher-power, more dangerous wildfires to calibrate the index.

The results, shown in figure 4.3, confirm the trends discussed by [111] and portrayed in figure 1.4, with an up-tick in the fires occurring in October, which are even higher than those in September. This trend is explained by the rainfall

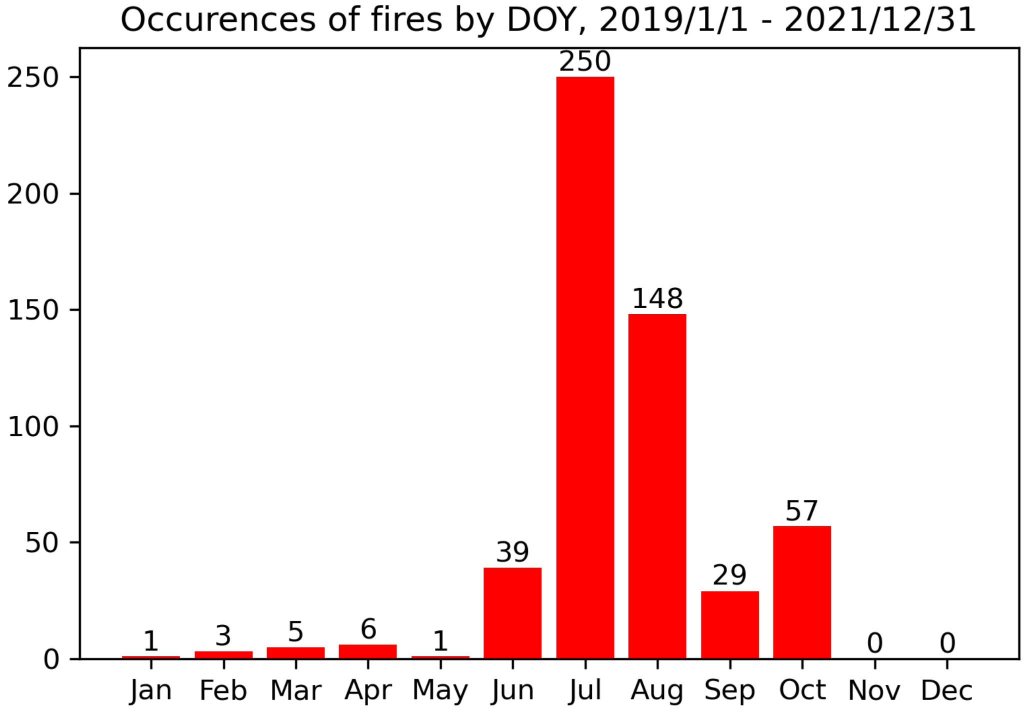

Figure 4.3. Occurrence of hotspots with FRP higher than 10 MW using FIRMS records in the timeframe 2019/1/1 - 2021/12/31.

patterns of the last few years, which saw relatively rainy Septembers followed by relatively dry and hot Octobers.

For this reason, we set the beginning of the fire season on the June 1st, and the ending on October 15th. The remaining days will be considered off fire season. The index thresholds identified using the iterative calibration process are listed in table 4.2, and the distribution of the index in the study area on and off the fire season is shown in figures 4.4 and 4.7, 4.4 and 4.8 and 4.6 and 4.9 respectively for years 2019, 2020 and 2021.

Danger Class	Index Interval
Very Low	0 – 30
Low	30 – 40
Medium	40 – 55
High	55 – 65
Extreme	65 – 100

Table 4.2. Classes of the DFDI and associated thresholds.

Looking at the three analysed years, we can observe the same general fire danger behaviour for each land cover type, with a progressive year-by-bear increase of the danger values during the fire seasons, and with shrublands returning higher danger ratings outside of the fire season and grasslands returning higher danger ratings during the fire season. Unsurprisingly, the analysis returns 2021 as the worst of the three processed years, since this year's fire season was considered the worst fire season in Sardinia in 24 years in terms of burned area [109], mostly because of the

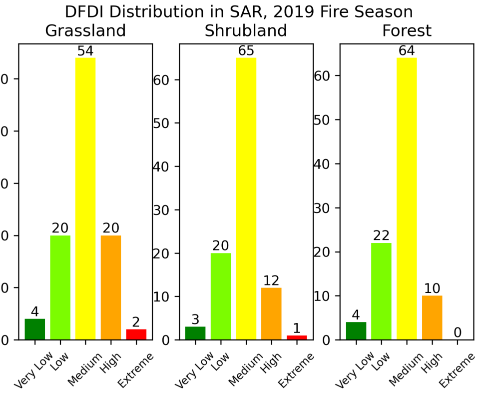

Figure 4.4. DFDI distribution by land cover in Sardinia, Fire season of 2019

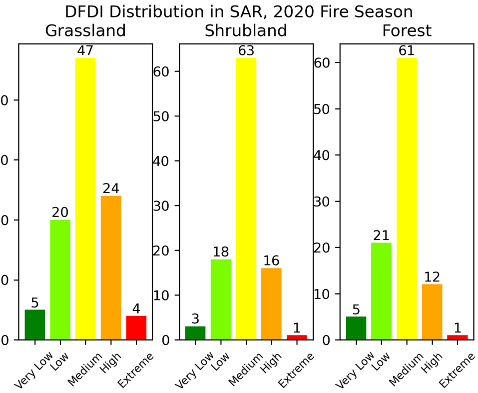

Figure 4.5. DFDI distribution in Sardinia, Fire season of 2020

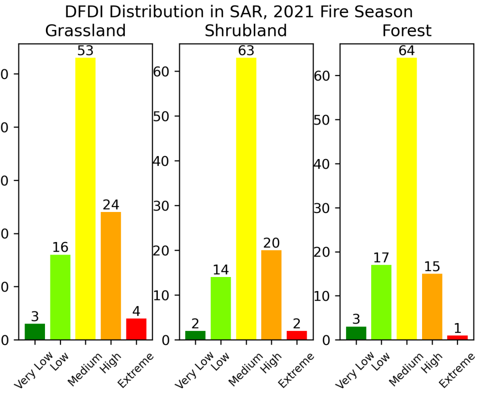

Figure 4.6. DFDI distribution in Sardinia, Fire season of 2021

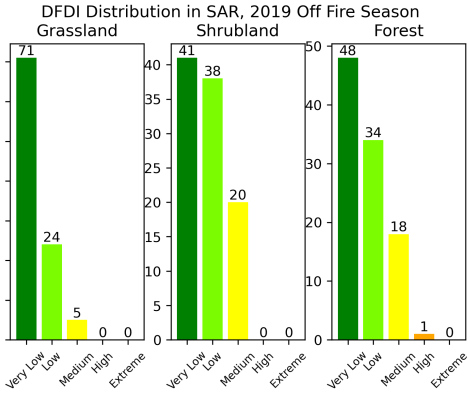

Figure 4.7. DFDI distribution by land cover in Sardinia, Off Fire season of 2019

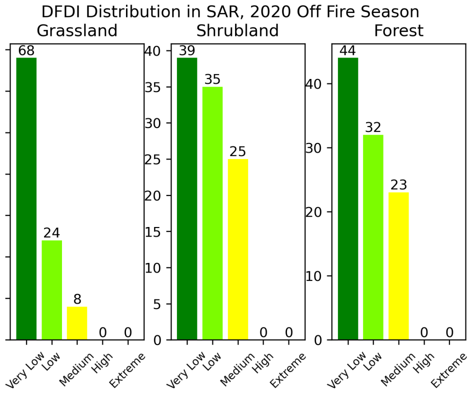

Figure 4.8. DFDI distribution in Sardinia, Off Fire season of 2020

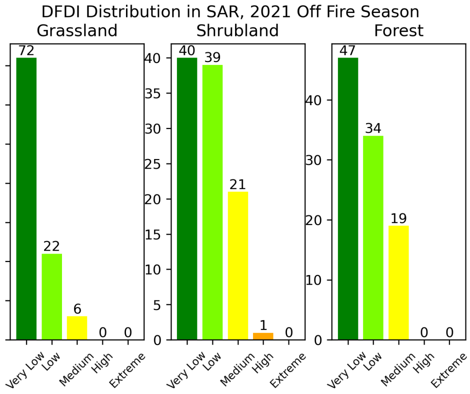

Figure 4.9. DFDI distribution in Sardinia, Off Fire season of 2021

devastating Planargia-Montiferru megafire, which destroyed over 13000 hectares of shrubs, cultivated areas and forests over a period of eight days.

Outside the fire season, the DFDI returns mostly very low or low danger values. On average, shrublands tend to show the highest fire danger values, but virtually no pixels are assigned an high or extreme danger rating outside of the fire season: only the forests in 2019 and the shrublands in 2021 round to 1% high danger rated pixels. Lower fire danger ratings are assigned to forests and grasslands, with 75-95% of the pixels being rated as very low or low danger.

During the fire season, grasslands tend to show the highest fire danger values, being regularly rated as the highest fire danger class, peaking in 2021 with 28% of the pixels rated as high or extreme danger, while forests tend to show the lowest fire danger values, again peaking in 2021 with 16% of pixels rated as high or extreme danger. Nevertheless, the amount of pixels rated as high or extreme danger remains well within the limit of manageability by fire danger experts: in forests, high or extreme danger rated pixels range from the 10% of 2019 to the 16% of 2021; in 2019, 13% of shrubland pixels are rated high or extreme danger, reaching a 22% peak in 2021. In the worst case scenario, which is associated to grasslands in the 2020 and 2021 fire seasons, we reach a total of 28% pixels rated high or extreme fire danger. Even in these two cases, only 4% of the pixels are assigned an extreme danger rating, allowing the experts to discriminate the most endangered areas even within a generalized context of high fire danger. Furthermore, it should be noted that fire management experts are generally more interested in managing shrublands and forests, for which the worst case scenarios consist in 22% and 16% of pixels respectively being rated high or extreme.

The performance of the index against FIRMS hotspot FRP and burnt area perimeters provided by the Sardinian fire department are shown in figure 4.10 and 4.11 respectively. In order to exclude small fires, only hotspots with an FRP higher than 10 MW and burnt areas larger than 10 Ha have been included in the analysis. While each hotspot corresponds to a single pixel, burnt areas were large enough to span multiple pixels. For this reason, each burnt area was rasterized to the Sentinel-3 grid, and all corresponding DFDI values were included in the analysis.

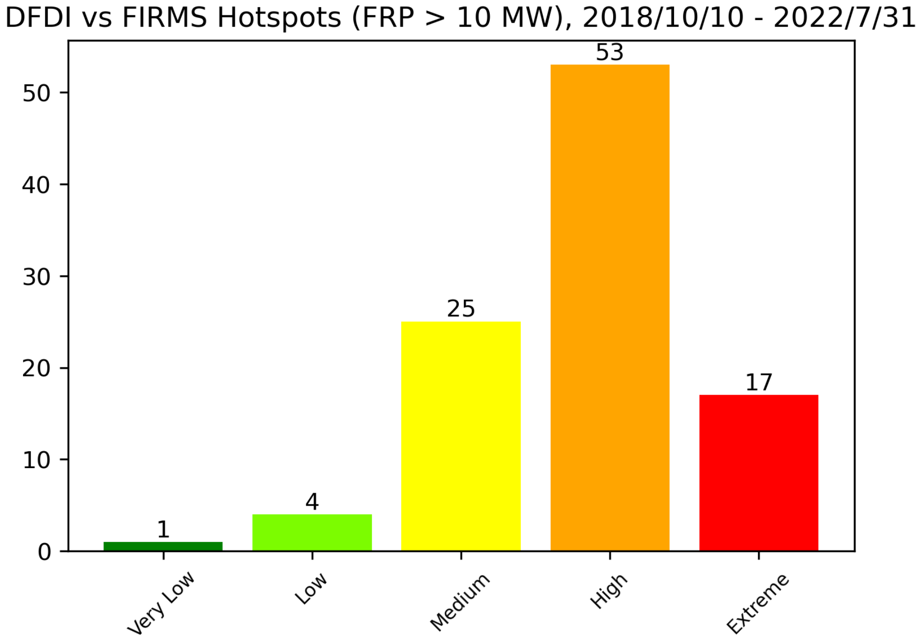

Figure 4.10. Distribution of the DFDI in Sardinia over FIRMS hotspots in the timeframe 2019/1/1 - 2022/12/31 represented as a percentage of total pixels assigned to each class.

The performance of the index against FIRMS hotspot FRP is very satisfactory:

only 5% of the sampled pixels were rated as very low or low danger, while 70% were rated as being high or extreme danger. In particular, the amount of pixels rated as extreme amounts to 17%, which is significantly larger then the average amount of pixels rated as such during the fire season for any land cover type: even in the worst of the three analysed fire seasons, this amount peaks at 4%, while it maintains values between 0 and 2% in all the other cases.

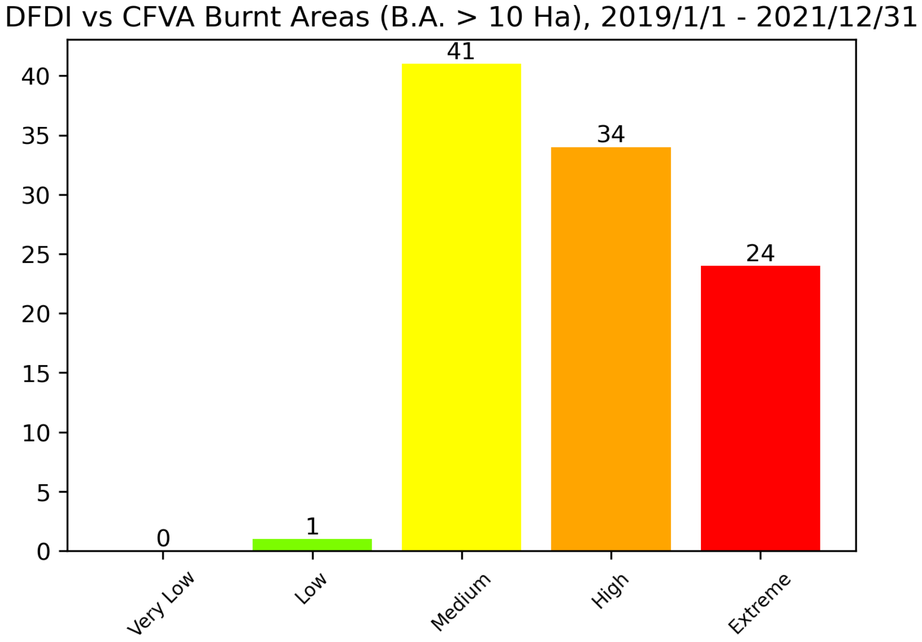

Figure 4.11. Distribution of the DFDI in Sardinia over burnt areas in the timeframe 2019/1/1 - 2021/12/31 represented as a percentage of total pixels assigned to each class.

The results over burnt areas fully confirm these impressions: an even smaller percentage of affected pixels were classified as low danger, and less than 1% were rated very low danger. Since we sampled the entire burnt areas and not only their geometric center or presumed ignition point, we would have expected that even lower danger pixels would be caught as collateral damage in large fires ignited in higher danger areas, and this consideration indeed applies to the medium danger pixels. Nevertheless, in this case almost 60% of the burnt area pixels were classified as high or extreme danger, with an even higher amount of extreme danger pixels than the previous case. Furthermore, compared to the results obtained in [84] on the DFHI (i.e. the DFDI's predecessor) using the same validation technique, these results show a generalized improvement, and in particular with respect to the almost complete absence of low danger pixels associated to burnt areas. This specific improvement is probably one of the advantages of the new additive model compared to the two-variable FPI model, which could return low danger values if only one of the two danger variables returned very low values for any kind of reason due to the fact that the two variables were multiplied to obtain the danger rating. With the additive model, even if one or more danger variables return low values, the final danger rating can still be high if the other danger variables return high values.

In general, these results confirm the capability of the index to discriminate fire danger in a way that is useful to the firefighting community, and at the same time that high fire danger areas are indeed more likely to suffer wildfires, and finally, that wildfires ignited in high fire danger areas are more likely to spread and cause massive damage.

Chapter 5

Comparison with the Canadian Fire Weather Index

In 2007, the EFFIS network adopted the Canadian Forest Fire Weather Index as the main method to assess fire danger at the European level after testing the performance of a number of different fire danger indices [84]. The CFFWI algorithm is made of six individual components that contribute to determining the fire danger. The inputs needed to calculate these components are temperature, relative humidity, wind speed and precipitation, i.e. only weather variables. The possibility of obtaining a fire danger rating using only relatively easy to obtain and process weather data is at the base of the CFFWI widespread adoption and success.

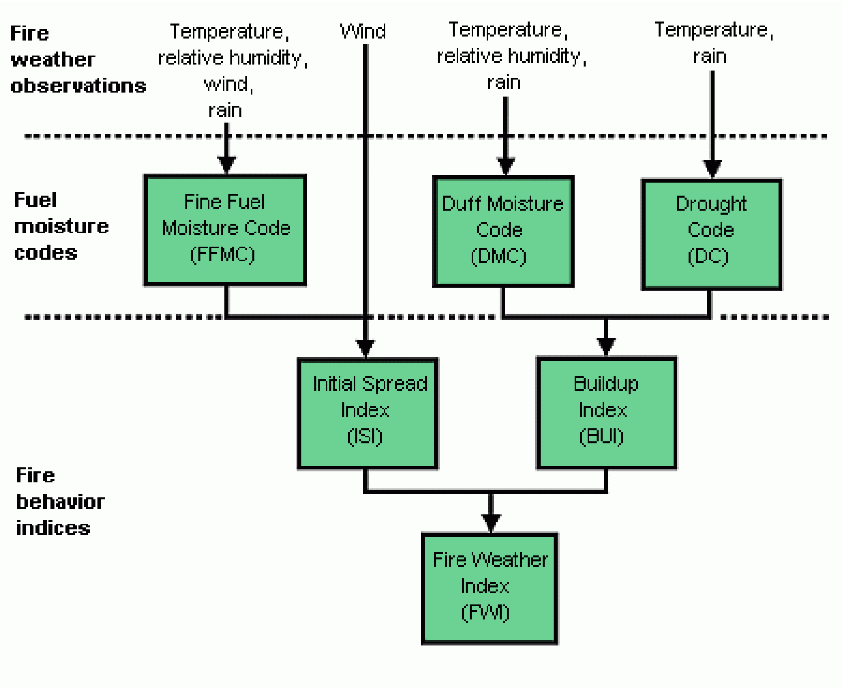

Figure 5.1. CFFWI calculation scheme, from the NRC website.

The six components of the CFFWI are:

1. **Fine Fuel Moisture Code**: related to the moisture content of litter and fine

dead fuels.

2. **Duff Moisture Code**: related to the moisture content of loosely compacted organic layers of moderate depth.
3. **Drought Code**: related to the moisture content of deep, compact organic layers.
4. **Initial Spread Index**: related to the expected fire spread rate.
5. **Buildup Index**: related to the amount of fuel available for the fire.
6. **Fire Weather Index**: a measure of fire intensity, normally used as a rating of fire danger.

For more information about the CFFWI, the reader may refer to [131] and to the Natural Resources Canada (NRC) website. The CFFWI data for the study area in the analysed time period were obtained using the Copernicus Climate API [106], and correspond to the consolidated dataset. The original data at 0.25° resolution was downscaled to the 300-meter Sentinel-3 resolution using a simple bilinear algorithm. The index thresholds used to turn the index values into fire danger ratings are shown in table 5.1, where, compared to the classification used by EFFIS, the "very high" and "extreme" danger classes were grouped in a single "extreme" class to simplify the comparison with the DFDI. The distribution of the CFFWI in the study area on and off the fire season is shown in figures 5.2 and 5.5, 5.2 and 5.6 and 5.4 and 5.7 respectively for years 2019, 2020 and 2021.

Danger Class	Index Interval
Very Low	0 – 5.2
Low	5.2 – 11.2
Moderate	11.2 – 21.3
High	21.3 – 38.0
Extreme	> 38.0

Table 5.1. Classes of the CFFWI and associated thresholds used in this study.

Outside of the fire season, the behaviour of the CFFWI is rather similar to that of the DFDI, but the differences between the three land cover types are much less pronounced. This was to be expected given the fact that the CFFWI is only based on weather data, and makes no distinction based on the underlying land cover. However, during the fire season the CFFWI tends to return a much higher percentage of high and extreme fire danger ratings compared to the DFDI. Even in the least severe fire season of 2019, the CFFWI classifies more than 60% of the pixels as high or extreme danger. In the worst case scenario, this value reaches 71%. Needless to say, a fire danger index with such a distribution can hardly provide useful information for the allocation of firefighting resources at the relatively large regional scale. Limiting the scope to pixels classified as extreme danger would drastically reduce the area under examination, but the affected surface area value still ranges from a fifth to a quarter of the total regional area. The CFFWI agrees with the DFDI in rating the 2021 fire season as the worst among the three analysed in this instance. However, the CFFWI rates the 2020 fire season as the least dangerous overall in terms of high and extreme rated pixels, even though this is partially compensated by the higher amount of extreme rated pixels in the 2021 season.

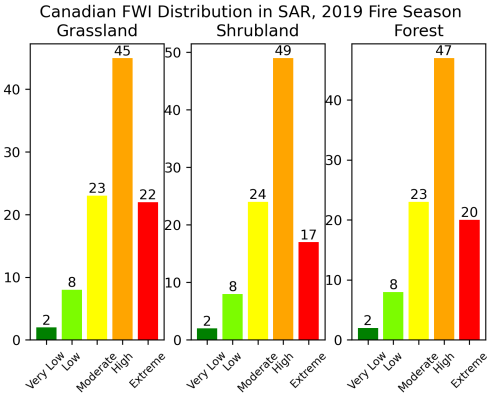

Figure 5.2. CFFWI distribution by land cover in Sardinia, Fire season of 2019

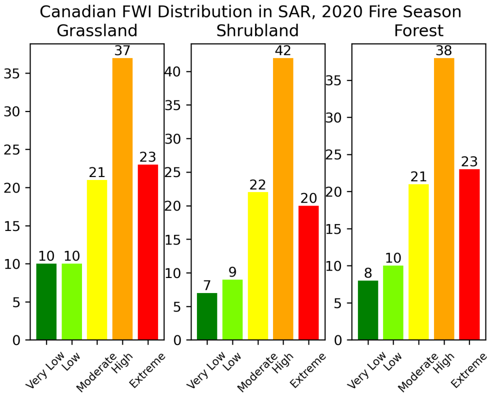

Figure 5.3. CFFWI distribution in Sardinia, Fire season of 2020

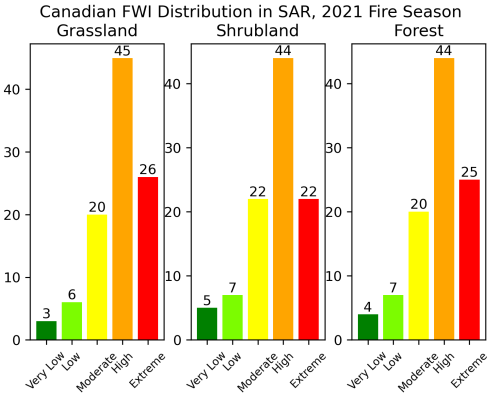

Figure 5.4. CFFWI distribution in Sardinia, Fire season of 2021

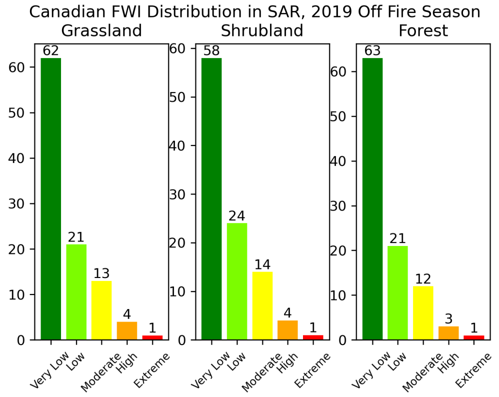

Figure 5.5. CFFWI distribution by land cover in Sardinia, Off Fire season of 2019

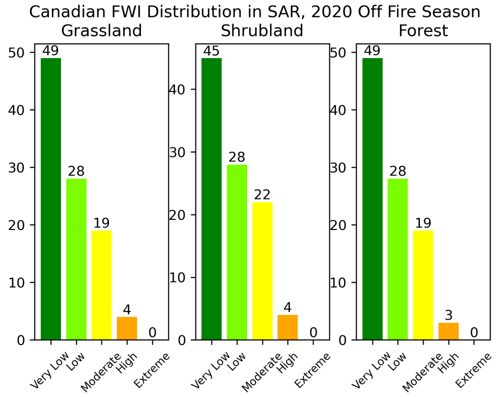

Figure 5.6. CFFWI distribution in Sardinia, Off Fire season of 2020

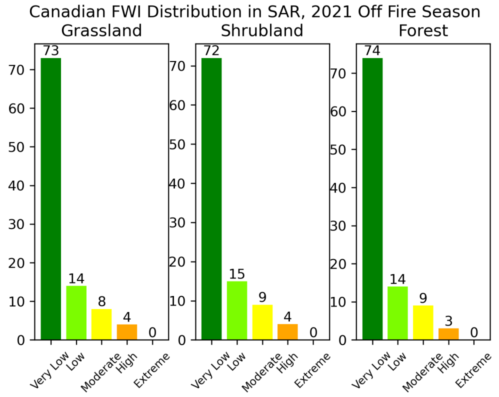

Figure 5.7. CFFWI distribution in Sardinia, Off Fire season of 2021

Chapter 6

The Planargia-Montiferru Megafire: A High-Resolution Case Study

The objective of this chapter is to compare the information provided by the moderate- and high-resolution maps in a context of high interest to the decision makers and to the scientific firefighting community in general, i.e. that of a megafire, and to evaluate how they complement each other and how they compare on the local scale. To this end, we selected the Planargia-Montiferru megafire, which occurred in Western Sardinia, more specifically in the Oristano province, in the middle of the summer of 2021. This wildfire burned almost 13000 hectares of vegetation, including 4000 hectares of forest [44]. These estimates made it the most destructive wildfire that occurred in Sardinia in 24 years. In particular, we will focus our analysis on a small area between the towns of Santu Lussurgiu and Bonarcado, where the original fire started. First of all, we will describe how the fire started and developed on the 23rd of July, and how on the next day the smouldering ashes re-ignited starting the megafire. We will then discuss the fire danger scenario on a regional scale using the moderate-resolution product, and use the high-resolution product to target the local scale around the original ignition point. Finally, we will compare the information provided by the moderate- and high-resolution products on the local scale, in order to assess if the moderate-resolution product would have provided enough information by itself, or if the added value of the high-resolution product justifies the increased complexity and computational cost necessary to generate it.

6.1 Development of the Fire

According to the local authorities, on July 23 of 2021 a fire erupted in the area between Santu Lussurgiu and Bonarcado. The wildfire was likely ignited by a vehicle that caught fire on the Provincial Road 15, and not without difficulties, was put out by the local fire department after burning around 20 hectares of land [116]. At the time of writing, the 2022 Google Street View capture still shows the residue of the car fire along the Provincial Road, which allowed us to pinpoint the exact ignition point of the wildfire. The Fire Brigade picture and the corresponding Google Street View capture are shown in figures 6.1 and 6.2 respectively.

Unfortunately, on the next day the weather conditions were characterized by very high temperatures and by a strong, hot sirocco wind. This allowed the smouldering ashes to re-ignite and to start secondary fires in multiple spots North-West from the

Figure 6.1. Sardinian Fire Brigade intervening on the scene of the accident of the 23rd of July 2021. The picture credits belong to the Fire Brigade, and the picture was downloaded from an online article of the local newspaper La Nuova Sardegna [116] published on the same day of the fire.

Figure 6.2. Google Street View 2022 capture [59] showing a patch of residue from the car fire still present on the Provincial Road 15. Using the picture on the left, it was possible to find the precise coordinates of the car, and therefore the likely ignition point. Note the electrical pole on the left.

original location. The images in figure 6.3 show the first hotspot detections of the Satellite FIre DEtection (SFIDE) system [41], based on the Geostationary satellite Meteosat 2nd Generation, and of the FIRMS system, based on the MODIS satellite, which was the first of the FIRMS satellite to overpass the area during the fire on that day.

This made interventions very difficult, and the wildfire would burn through the forests of Santu Lussurgiu and Cuglieri, the hills of Scano di Montiferro and the valleys of Tresnuraghes and Sennariolo before running out of fuel in close proximity of the sea in Porto Alabe [44]. The authorities ruled out the possibility of the fire being intentional, and therefore no arrests were made in conjunction with these events.

6.2 Fire Danger Maps

As specified in the previous chapters, the moderate-resolution DFDI maps are based on 300-meter resolution Sentinel-3 images, while the high-resolution maps are based on 10-20-meter resolution Sentinel-2 images. The variables based on weather data were calculated using the same 0.10 degrees resolution ERA-5 Land product, and then downscaled to the resolution of the target imagery using a simple bilinear algorithm. In addition, while the Sentinel-3 satellites can provide new images almost every day, the closest Sentinel-2 image that captured the AOI was acquired on July 22, one day before the fire. However, it is reasonable assume that the state of the vegetation in terms of greenness and moisture does not change significantly in a single day. For this reason, in general we can calculate the components of the DFDI based on satellite observations by simply using the latest available satellite image, as long as the latest observation is not older than seven days.

6.2.1 Moderate-Resolution DFDI Maps: A Regional Fire Danger Perspective

Figure 6.4 contains the regional DFDI maps based on Sentinel-3 imagery covering the timeframe that starts on the 17th of July 2021 and ends eight days later on the 25th. The blue 'x' marks the position of the original ignition point, assuming that it

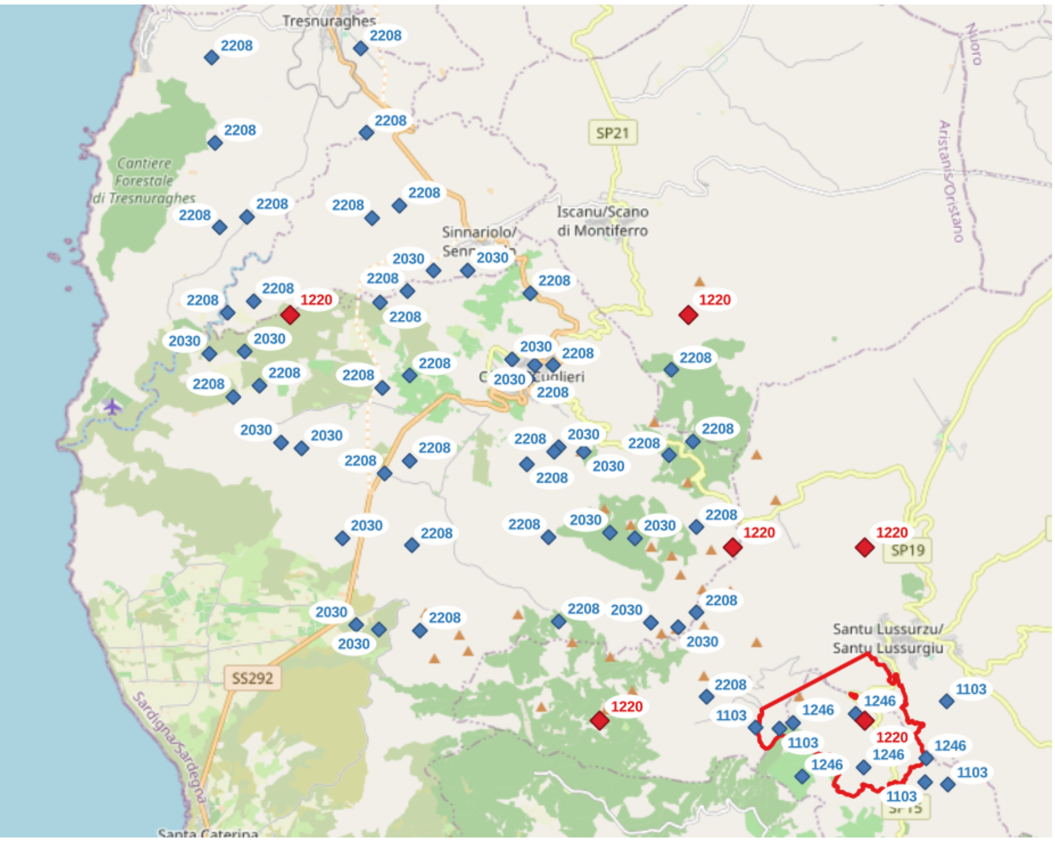

Figure 6.3. First detection of the re-ignitions of the 24th of July from the Meteosat 2nd Generation geostationary satellite using the SFIDE algorithm (4̃km resolution, red diamonds), and MODIS detections (1km resolution, blue diamonds). The labels represent the acquisition time in HHMM format. The red outlined polygon on the bottom right encompasses the area between Santu Lussurgiu and Bonarcado where the fire first erupted on the 23rd of July. Basemap provided by OpenStreetMap [96].

coincides with the car fire residue patch shown in figure 6.2. The analysed timeframe opens with a relatively mild danger scenario: on the 17th, most of the region is rated as low or medium danger, with the northeastern area corresponding to the Sassari province and the Oristano Gulf being rated low or even very low danger. The largest patches of high danger pixels can be found in the northern continental area around Ozieri, and along the southern coastline. The situation progressively worsens in the following days: on the 18th, the eastern half of the region is rated high danger, with the appearance of an extreme danger patch in the south-east around Lanusei. The fire danger map of the 19th shows a very similar scenario, while on the 20th a major portion of continental Sardinia is rated as high danger. On the 21st, the situation slightly improves, but a patch of extreme danger pixels appears in the central area between Macomer and Nuoro. On the 22nd, the situation worsens further, and another large patch of extreme danger pixels appears in the southern part of the region extending south-west to north-east from the Sulcis-Iglesiente to the Campidano areas. By the 23rd, the day of the original fire between Santu Lussurgiu and Bonarcado, barely any low danger pixels can be spotted in the entire region, but the fire danger scenario becomes truly dramatic on the 24th, the day of the Planargia-Montiferru megafire: due to the extremely hot weather and strong, hot sirocco wind, most of the southwestern part of the region is rated as extreme danger. The situation improves significantly on the 25th, when our timeframe ends.

Observing this fire danger map timeseries, and the grave scenario of the 24h in particular, it is not surprising that a megafire would occur on this very day: the amount of pixels rated as extreme danger, which in the entire 2021 fire season add up to 4% of the total grassland pixels, 2% of the total shrubland pixels and 1% of the total forest pixels, is higher than all the other days combined. Exploiting such fire danger maps, especially if sustained by accurate weather forecasts, could greatly help the decision makers to take countermeasures in such extreme situations, and possibly to help prevent megafires not only by deploying resources on the territory and putting the fire fighters on high alert, but also by properly informing the local population.

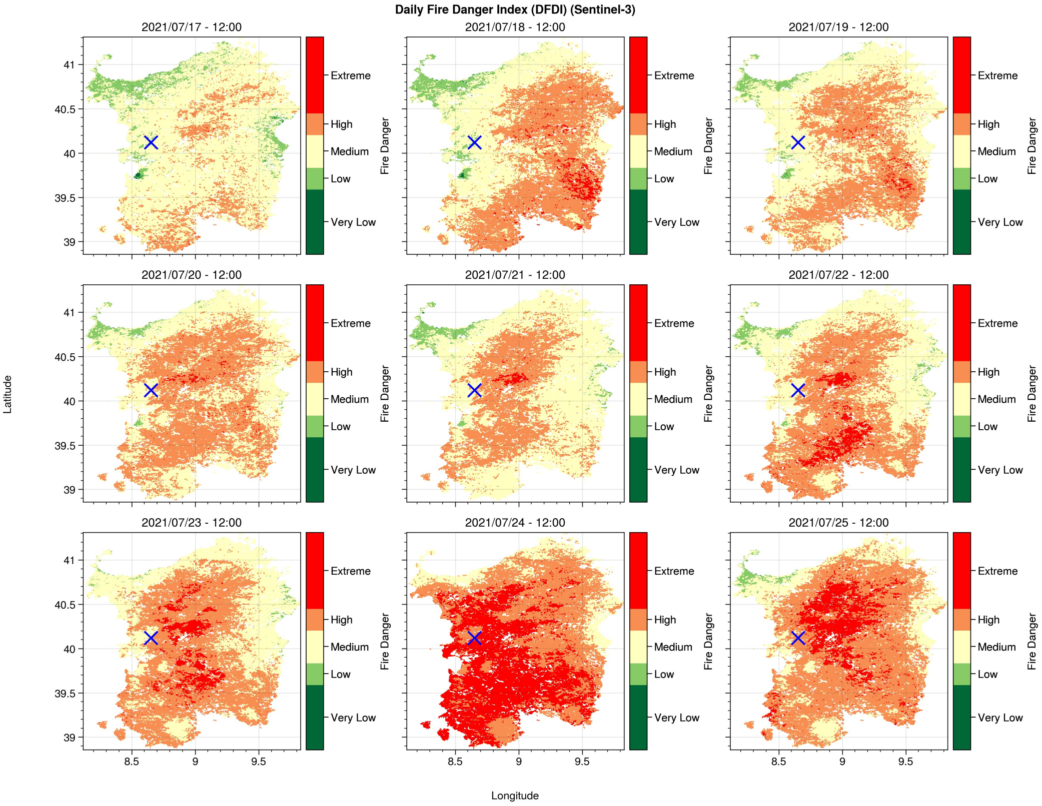

Figure 6.4. Sentinel-3 based (300-meter resolution) Daily Fire Danger Index maps in the timeframe 2021/07/17-2021/07/25 for the entire region of Sardinia. All indices are calculated using the weather data at 12:00 UTC. The location of the original wildfire that occurred between Santu Lussurgiu and Bonarcado on the 23rd is marked by a blue 'x'. The fire was suppressed by the Sardinian Fire Brigade, but the next day it re-ignited and sustained by a hot and strong sirocco would go on to devastate the Planargia and Montiferru areas.

Figure 6.5 shows the values of the variables that concur to the calculation of the DFDI on the 24th, the day of the megafire, and the DFDI map itself. Looking at the individual variables, the high evapotranspiration and low equivalent moisture content clearly show how the torrid weather affected the vegetation. In addition, the wind map shows that strong sirocco winds transported the smouldering ashes from the Santu Lussurgiu-Bonarcado towards North-West, igniting the entire Planargia-

Montiferru.

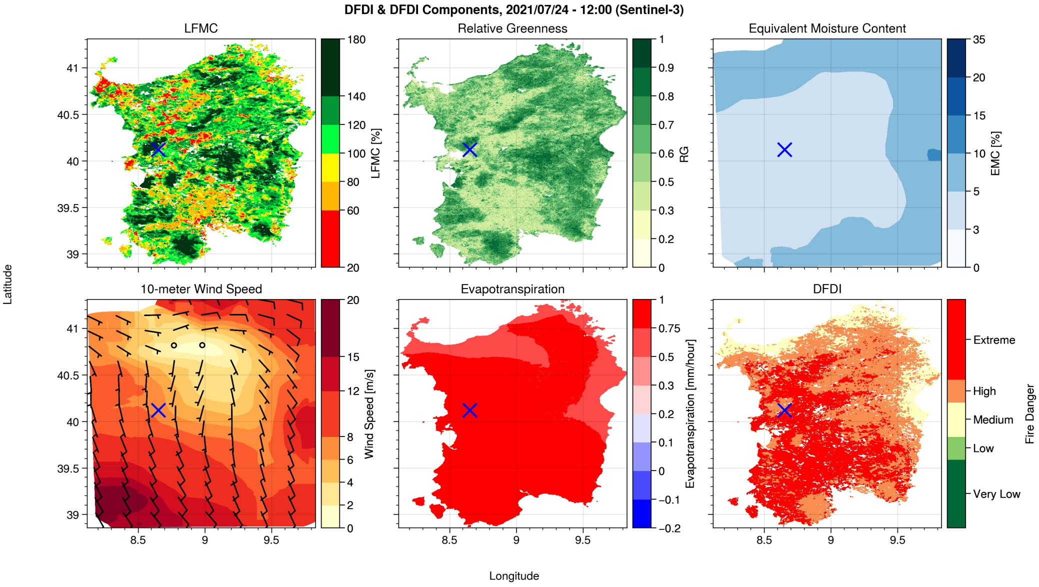

Figure 6.5. Variables that concur to the calculation of the DFDI and DFDI map of the 24th of July, date of the Planargia-Montiferru megafire. All weather data refer to 12:00 UTC. The location of the original wildfire that occurred between Santu Lussurgiu and Bonarcado on the 23rd is marked by a blue 'x'.

6.2.2 High-Resolution DFDI Maps: Fire Danger in the Sardinian WUI

Figure 6.6 shows high-resolution DFDI maps generated using Sentinel-2 imagery in the area between Santu Lussurgiu and Bonarcado in the same timeframe as the previous paragraph. The blue 'x' marks the position of the car fire residue patch shown in figure 6.2, which corresponds to a point on the Provincial Road 15 which we assume to be the ignition point of the fire that occurred on the 23rd.

In the first three days of the analysed timeframe, the fire danger scenario is relatively mild, with most of the area being rated as very low, low or medium danger. A few patches of high danger pixels show up, mainly in the northern part of the area, but most of the Provincial Road 15, which travels in the middle of the eastern half, is surrounded by medium or low danger pixels. The situation worsens significantly from the 20th onwards, when barely any low danger pixels remain and most of the AOI is covered by high danger pixels. A few patches of extreme danger pixels show up in the northern part. After a small improvement on the 22nd, we reach the 23rd, which is the day of the car fire that ignited the wildfire that burnt most of the vegetation in this area. On this day, the southern half of the AOI is almost entirely rated as high danger, and in particular, the area around the 'x' that marks the position of the car along the Provincial Road 15 is almost entirely encompassed by high danger pixels. By consulting this fire danger map, it is therefore not surprising that a fire could start and propagate easily in this WUI, explaining also the difficulty that the firefighters experienced in suppressing it. On the 24th, the situation is as dramatic as the regional level, and on this day the smouldering ashes would re-ignite to start the Planargia-Montiferru megafire.

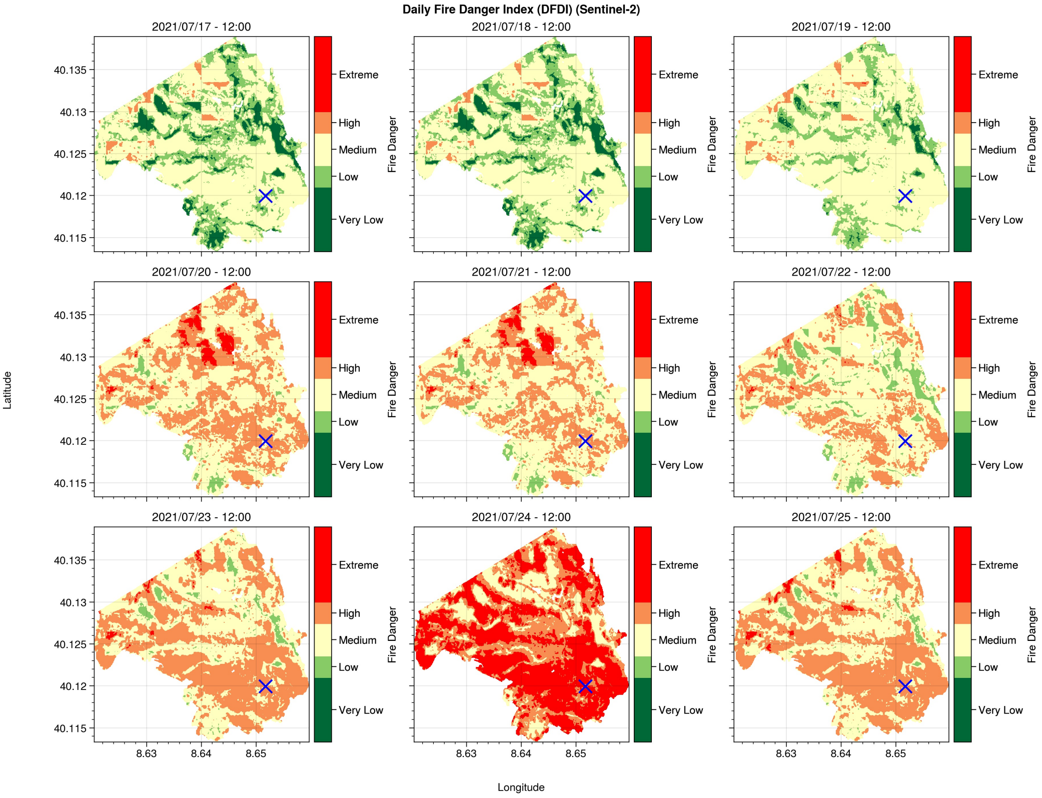

Figure 6.6. Sentinel-2 based (10-20-meter resolution) Daily Fire Danger Index maps in the timeframe 2021/07/17-2021/07/25 for the Santu Lussurgiu-Bonarcado area. All indices are calculated using the weather data at 12:00 UTC. The blue 'x' marks the position of the car that caught fire on the Provincial Road 15 on the 23rd.

6.2.3 Comparison of the Moderate- and High-Resolution Fire Danger Maps

Figure 6.7 shows a direct comparison between the moderate- and high-resolution fire danger maps in the area between Santu Lussurgiu and Bonarcado on the 23rd of July, when a car caught fire along the Provincial Road 15 in the position marked by a blue triangle in the south-east of the area of interest.

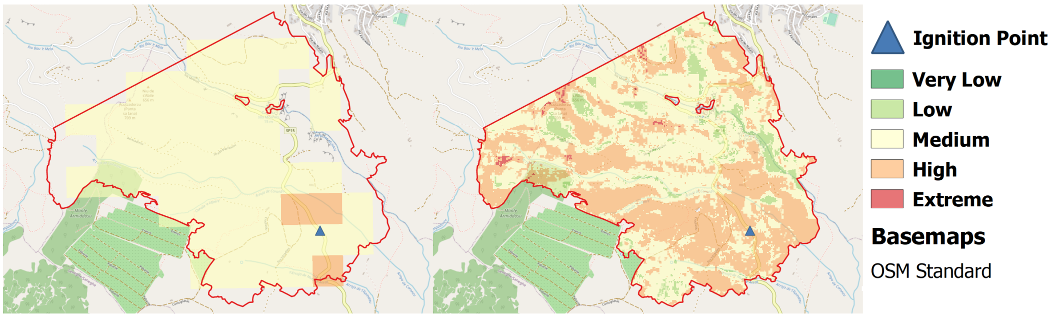

Figure 6.7. Sentinel-3 based (left) and Sentinel-2 based (right) Daily Fire Danger Index maps on the 23rd of July 2021. The blue triangle marks the position of the car that caught fire along the Provincial Road 15. Basemap provided by OpenStreetMap [96]

The moderate-resolution product is dominated by medium danger pixels, with only three high danger pixels. However, these high danger pixels are localized around the ignition point, implying that even at a 300-meter resolution this area was recognized as being the most dangerous. Nevertheless, it is evident how the level of detail provided by the 10-20-meter resolution Sentinel-2 images is much higher: the area around the ignition point, and in general along the Provincial Road 15 in the southern part of the AOI is dominated by high danger pixels. The situation improves slightly moving northwards along the Provincial Road, especially in close proximity to the small Sos Molinos stream, before returning to high danger levels entering the town of Santu Lussurgiu. Focusing on the WUI along the Provincial Road, it is clear how the DFDI maps were able to identify two major areas characterized by high fire danger on both sides of the road: one on the southern side, where the car caught fire starting the wildfire on the 23rd, and one on the northeastern side towards the entrance of Santu Lussurgiu. However, while on the northeastern side the fire danger distribution around the WUI was less worrisome, on the southern side it was clear that not only a fire could easily start, but that it could easily spread to the surrounding area dominated by high fire danger pixels.

In general, this comparison shows how useful a high-resolution fire danger product could be to provide accurate information on the local scale to the local administrators and decision makers: having access to a high-resolution fire danger forecast could have allowed the municipalities to plan and optimize fuel reduction campaigns on a very short term, greatly improving the chance to prevent large fires altogether. At the same time, the fire danger maps could be used to inform the citizens, especially those who live in the WUI and would be put at great risk by such fires, and possibly even encourage them to enact prevention measures independently in order to improve the security of their family and of their property.

Conclusions and Future Directions

After describing and selecting the fire weather variables most relevant to fire danger, this research work produced a procedure for Live Fuel Moisture Content retrieval from satellite images based on the latest literature, and applied it to Sentinel-2 and Sentinel-3 images with promising results according to an early validation performed on field data. In addition, the LFMC Sentinel-3 product is - to our knowledge - the very first LFMC product based on this satellite platform, and the Sentinel-2 product is among the very first of its kind [119, 88, 31].

Subsequently, a novel dynamic Fire Danger Index based entirely on European weather and satellite products was proposed, introducing an additive model to combine the fire danger variables. The index was calibrated using historical fire occurrence records, and its performance was tested against the CFFWI, which is EFFIS' fire danger index of choice. The performance of the new index on the test area proved to be superior to the CFFWI in all the analysed years according to the selected test metric. Even at the moderate resolution provided by Sentinel-3, firefighting operators using the DFDI would be in a position to discriminate the fire danger in a given region to a much finer degree than they would if they were using the CFFWI for the same purpose, where they would often find themselves in a situation where fire danger is high or extreme in all the region, crippling their ability to properly allocate the resources devoted to fire fighting and prevention.

Finally, the moderate- and high-resolution maps were compared using the Planargia-Montiferru megafire as a case study. The analysis of the regional scale maps confirmed how they were able to detect an anomalous, extreme fire danger scenario on the day of the megafire. On the local level, the added value of the high-resolution maps allowed to identify fire danger hotspots at the WUI level, and in particular in the area where the fire started along the Provincial Road 15. The high-resolution product could provide valuable information to enact prevention measures at a local level, optimizing the logistic and economic effort of small municipalities and private citizens.

This work also pointed out a number of possible future research lines that may further increase its value for the firefighting community:

- First of all, the LFMC vegetation model could be improved by 1) adding the capability to mix different vegetation spectral profiles pixels (e.g. shrub and forest) in order to model mixed pixels, 2) adding more tree shapes to the GEO module, and 3) by generalizing the geometric model to non-nadir views.

- Sentinel-3 SYNERGY data is provided with non-time-critical timeliness, implying that usually the latest products are delivered one or two days after acquisition. Even though, most of the time, the state of the vegetation does not change so quickly as to make a two-day delay a deal breaker for the usage of the data for operational purposes, it would be interesting to explore the impact of using near-real-time data on the accuracy of the fire danger assessment. At the same time, it should be noted that handling low processing level near-real-time imagery would significantly increase the workload of the fire danger data provider.

- Furthermore, it would be interesting to explore the impact of using high-resolution weather data and a scalable fuel map such as [6] in place of the static one currently used in the fire danger assessment. This would provide adaptability to different spatial scales also at the level of the fuel maps and properties, and would be especially interesting for the high-resolution product.
- Finally, the analysis of the Planargia-Montiferru megafire showed that not only it happened on a day of extreme fire danger, but at the same time when a sharp, generalized increase of the fire danger occurred at the regional scale. It would be interesting to repeat the analysis on other megafires to verify if a similar pattern can be recognized.

Acknowledgements and Thanks

I wish to thank Professor Giovanni Laneve for his invaluable guidance during this PhD programme, Professor Domingos Xavier Viegas and Luis Mario Ribeiro of the Forest Fire Research Center of ADAI of the University of Coimbra for providing the Live Fuel Moisture Content measurements for Lousa, Grazia Pellizzaro and Valentina Bacciu of CNR for providing Live Fuel Moisture Content data for the Capo Caccia site, and the FirEUrisk project [30] in general. Special thanks to my friend Marco Mastrofini for spotting the residue patch along the Provincial Road 15 in Sardinia that allowed me to pinpoint the exact ignition spot of the Santu Lussurgiu-Bonarcado wildfire. I also send my sincerest thanks to the reviewers Prof. Luigi Boschetti and Dr. Marìa Lucrecia Pettinari, both for their compliments and for their comments and advice, which helped shaping this thesis into something I am very proud of.

Appendix

Software used for this project

Only free and open source software was used in the development of this research project. All the existing software components mentioned in the thesis are freely available, and all the code that was created specifically for the innovative parts of the project was written in Python in an Anaconda environment. In particular, our fork of `jgomezdans/pyprosail` will soon be made publicly available on GitHub.

- Anaconda
- Cartopy
- GDAL
- Geopandas
- LaTeX
- Matplotlib
- Numpy
- Pandas
- Pyprosail
- Python
- Proplot
- SaLib
- Sentinelsat
- Scikit-image

Procedure used to Calculate the Reference Evapotranspiration

In order to calculate daily values of the reference evapotranspiration, we follow the procedure described in the Food and Agriculture Organization (FAO) irrigation and drainage paper 56 [3], to which we refer the reader for an extensive and detailed explanation of each single variable and calculation step. We are only interested in the calculation of the reference evapotranspiration, denoted as ET_0, for hourly timescales using weather data. However, the aforementioned paper contains a number of different ways in which the variables involved in the process depending on the available data type, that may be weather forecast data or directly obtained from weather stations, and on the different types of measurements that may or may not be available. In order to ensure the repeatability of our process, we will now describe how we obtained each individual variable.

The first half of the procedure relies on atmospheric data to quantify the vapour pressure deficit. The first step consists in using a digital elevation model to calculate the atmospheric pressure at the mean elevation of each of our pixels. If z is the elevation above sea level in meters, equation 6.1 returns the atmospheric pressure P in kPa:

$$P = 101.3 \left(\frac{293 - 0.0065z}{293} \right) \tag{6.1}$$

Since DEMs are generally calculated at a resolution that is much higher than that of the weather data, we need to determine at which resolution to calculate this variable. Since the fire danger index maps will be calculated at the same resolution of the satellite images, it makes sense to upscale the DEM and downscale the weather data at the same resolution as the satellite images. The simplest way to achieve this is to upscale the DEM using an averaging function and to downscale the weather data using a bilinear algorithm.

The atmospheric pressure is then used to compute the psychrometric constant γ:

$$\gamma = \frac{c_p P}{\varepsilon \lambda} \tag{6.2}$$

where:

- γ psychrometric constant in [kPa $^\circ$C^{-1}].
- $c_p = 1.013 \cdot 10 - 3 [MJ\ kg - 1^\circ C^{-1}]$ specific heat at constant pressure
- P atmospheric pressure in [kPa] as calculated in 6.1
- $\varepsilon = 0.622$ ratio between the molecular weight of water vapour and dry air
- $\lambda = 2.45 [MJ\ kg^{-1}]$ latent heat of vaporization

In particular, the value of the latent heat of vaporization λ is referred to the average atmospheric conditions. Thus, the psychrometric constant varies only with the atmospheric pressure, and consequently only with elevation.

Then, we use the air temperature and dewpoint temperature to calculate the actual vapour pressure and the saturation vapour pressure as done previously to compute the relative humidity:

$$e_a = e^0(T_{dew}) = 0.6108 * \exp \frac{17.27\ T_{dew}}{T_{dew} + 237.3} \tag{6.3}$$

$$e_s = e^0(T) = 0.6108 * \exp \frac{17.27\ T}{T + 237.3} \tag{6.4}$$

The vapour pressure deficit $e_s - e_a$ is then defined as the difference between saturation vapour pressure and the actual vapour pressure. The air temperature is also used to calculate the slope of the saturation vapour pressure curve Δ:

$$\Delta = \frac{4098\left[0.6108\exp\left(\frac{17.27\,T}{T+237.3}\right)\right]}{(T+237.3)^2} \tag{6.5}$$

where:

- $\Delta[kPa°C^{-1}]$ slope of saturation vapour pressure curve at air temperature T
- $T[°C]$ air temperature

The second half of the procedure uses a digital elevation model, the day of the year and the latitude of the target area to quantify the net radiation that reaches the surface. Since we're interested in providing hourly estimates of fire danger, we follow the procedure to compute the radiation variables for hourly or shorter periods.

The extraterrestrial radiation R_a represents the amount of radiation that impacts a given point of the top of Earth's atmosphere. If this quantity is referred to a surface perpendicular to the Sun rays, it is named solar constant and has a value of $0.0820 MJ/m^{-2}min^{-1}$. Depending on the geographic position, season and time of day, this value will decrease depending on the incidence angle of the sun rays. To this end, if J is the day of the year[1], the inverse relative distance between Earth and the Sun d_r is given by:

$$d_r = 1 + 0.033\cos\left(\frac{2\pi}{365}J\right) \tag{6.6}$$

The sunset hour angle depends on the latitude and on the day of the year. It can be calculated using the former and the solar declination:

$$\omega_s = \arccos[-\tan(\phi)\tan(\delta)] \tag{6.7}$$

The solar declination δ in radians is given by:

$$\delta = 0.409\sin\left(\frac{2\pi}{365}J - 1.39\right) \tag{6.8}$$

The solar time angles at the beginning and end of the hourly period are given by:

$$\begin{cases}\omega_1 = \omega - \frac{\pi t_1}{24}\\ \omega_2 = \omega + \frac{\pi t_1}{24}\end{cases} \tag{6.9}$$

where:

- ω $[rad]$ solar time angle at midpoint of hourly or shorter period
- t_1 $[hour]$ length of the calculation period

The solar time angle ω at the midpoint of the period can then be obtained as:

$$\omega = \frac{\pi}{12}\{[t + 0.06667(L_m - L_z) + S_c] - 12\} \tag{6.10}$$

where:

- t standard clock time at the midpoint of the period [hour]

---

[1] The day of the year is counted from 1, 1 January, to 366, 31 December of leap years.

- L_z longitude of the centre of the local time zone [degrees east of Greenwich]
- L_m longitude of the measurement site [degrees east of Greenwich],
- S_c seasonal correction for solar time [hour]

The seasonal correction for solar time depends on the day of the year:

$$\begin{cases} S_c = 0.1645\sin(2b) - 0.1255\cos(b) - 0.025\sin(b) \\ b = \frac{2\pi(J-81)}{364} \end{cases} \tag{6.11}$$

For hourly or shorter periods, the extraterrestrial radiation can be calculated as follows:

$$R_a = \frac{12 \cdot 60}{\pi} G_{sc} d_r \{(\omega_2 - \omega_1)\sin(\phi)\sin(\delta) + \cos(\phi)\cos(\delta)[\sin(\omega_2) - \sin(\omega_1)]\} \tag{6.12}$$

where:

- $R_a [MJ\ m^{-2} hour^{-1}]$ extraterrestrial radiation in the hour (or shorter) period
- $G_{sc} = 0.0820\ MJ\ m^{-2} min^{-1}$ solar constant
- d_r inverse relative distance Earth-Sun
- δ solar declination
- ϕ latitude
- ω_1 solar time angle at beginning of period
- ω_2 solar time angle at end of period

The number of daylight hours depends on the day of the year and on the latitude. Thus, it can be calculated using the sunset hour angle ω_s calculated in equation 6.7:

$$N = \frac{24}{\pi}\omega_s \tag{6.13}$$

The portion of the extraterrestrial radiation that reaches the surface depends not only on the geographic location and on the season, but also on the actual sunshine duration. Since we are mainly targeting a regional or national spatial scale, calibrated values of the coefficients required to calculate the solar radiation are unavailable. Thus, we use the recommended coefficients $a_s = 0.25$ and $b_s = 0.50$:

$$R_s = \left(0.25 + 0.50\frac{n}{N}\right) R_a \tag{6.14}$$

where N is the maximum possible duration of sunshine in hours we calculated in 6.13 and n is the actual duration of sunshine. The FAO procedure [3] relies on the usage of a Campbell-Stokes sunshine recorder for direct measurement of the actual duration of sunshine. In our case, we have no way of including cloud cover in this calculation, but we can exploit knowledge of the topography based on the DEM to determine whether a given pixel, at a given time of day of a given time of the year is illuminated or in shade. To do so, we follow the process shown in [42], which is based on equations developed by [52].

Let us consider a surface located at a latitude ϕ, when the local declination is δ. The angle between the normal to the surface and the horizontal plane (slope) is β, and the displacement between the horizontal projection of the normal from the local meridian (azimuth) is γ. If ω is the solar hour angle at the target time, the cosine of the angle between the incoming beam radiation and the normal to the surface θ is:

$$\begin{aligned}\cos\theta = \sin\delta[\sin\phi\cos\beta - \cos\phi\sin\beta\cos\gamma] \qquad (6.15)\\ + \cos\delta[\cos\phi\cos\beta\cos\omega + \sin\phi\sin\beta\cos\gamma\cos\omega \qquad (6.16)\\ + \sin\beta\sin\gamma\sin\omega] \qquad (6.17)\end{aligned}$$

While the slope angle β can be directly obtained using GDAL's `gdaldem` function, some work must be done to obtain the γ angle. `gdaldem` outputs the aspect angle, which has the same significance as γ, but is 0 when the surface is facing North and is defined between 0 and 360 degrees. To obtain γ from the aspect angle we need only subtract 180 degrees:

$$\gamma = aspect - 180^\circ \qquad (6.18)$$

A scheme of the illumination geometry of a south-facing surface (i.e. $\gamma = 0$) with a β sloped is represented in 2.3.

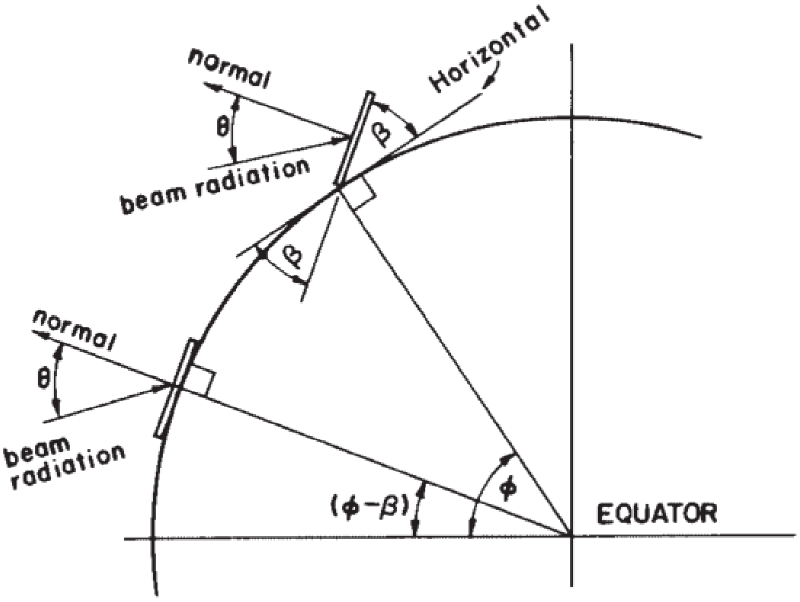

Figure 6.8. Scheme of the geometry of the incoming radiation on an inclined surface for a south-facing slope (i.e. $\gamma = 0$), adapted from [42]. β represents the slope, ϕ represents the latitude and θ is the angle between the normal to the local surface and the incoming beam radiation.

The complementary angle to the beam incidence angle θ represents the solar elevation with respect to the surface. We are more interested in this angle than in θ itself, because when it is positive and sufficiently large we can assume that the surface is directly illuminated by the Sun rays. Let the solar elevation angle be ϵ. We can calculate it as:

$$\epsilon = 180^\circ - \theta = 180^\circ - \arccos(\sin\delta[\sin\phi\cos\beta - \cos\phi\sin\beta\cos\gamma] \quad (6.19)$$

$$+ \cos\delta[\cos\phi\cos\beta\cos\omega + \sin\phi\sin\beta\cos\gamma\cos\omega \quad (6.20)$$

$$+ \sin\beta\sin\gamma\sin\omega]) \quad (6.21)$$

For each hour of potential sunlight N, and therefore for each solar hour angle ω between the sunset hour angle ω_n and the sunrise sunset angle $\omega_n - \frac{\pi}{N}$, we will check if this value is higher than 5° for each individual pixel and its topography. If it is, we will add 1 hour to n, otherwise we will not. The total count will give us the approximate total amount of sunlight hours accounted for shade caused by the local topography.

The clear-sky solar radiation can be obtained by setting $n = N$ and applying a correction for the local elevation:

$$R_{so} = (0.75 + 2 \cdot 10^{-5} z) R_a \quad (6.22)$$

The net shortwave radiation R_{ns} can be simply obtained by using an albedo or canopy reflection coefficient α and calculating its one's complement to obtain the absorption coefficient:

$$R_{ns} = (1 - \alpha) R_s \quad (6.23)$$

For the reference canopy, we set $\alpha = 0.23$. The net longwave radiation R_{nl} for ho hourly periods can be obtained using the Stefan-Boltzmann equation and the hourly mean temperature, while accounting for the actual duration of sunshine by using the ratio between the clear sky and actual solar radiation:

$$R_{nl} = \sigma_{hour} T^2_{hour,K} (0.34 - 0.14\sqrt{e_a}) \left(1.35 \frac{R_s}{R_{so}} - 0.35\right) \quad (6.24)$$

where σ_{hour} is the value of the Stefan-Boltzmann constant converted from $[MJ/(m^2\ day)]$ to $[MJ/(m^2\ hour)]$:

$$\sigma_{hour} = \frac{\sigma}{24} = \frac{4.903 \cdot 10^{-9} MJ/(m^2\ day)}{24 hours/day} = 2.403 \cdot 10^{-10} MJ/(m^2\ hour) \quad (6.25)$$

Thus, the net radiation is defined as the difference between the incoming net shortwave radiation and the outgoing net longwave radiation:

$$R_n = R_{ns} - R_{nl} \quad (6.26)$$

The soil heat flux G for hourly or shorter periods cannot be reliably correlated with air temperature, and is therefore estimated as a fraction of the net radiation:

$$\begin{cases} G_{hr} = 0.1\ R_n, & \text{during daytime} \\ G_{hr} = 0.5\ R_n, & \text{during nighttime} \end{cases} \quad (6.27)$$

Finally, we can use the Penman-Monteith equation adjusted for hourly timesteps is the following:

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma \frac{37}{T+273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34 u_2)} \quad (6.28)$$

where:

- ET_0 reference evapotranspiration [mm hour $^{-1}$];
- R_n net radiation at the crop surface [MJ m^{-2} hour^{-1}];
- G soil heat flux density [MJ m^{-2} hour^{-1}];
- T mean hourly air temperature at 2 m height [°C];
- u_2 mean hourly wind speed at 2 m height [m s^{-1}] at the given time
- e_s saturation vapour pressure [kPa] at the given time
- e_a actual vapour pressure [kPa] at the given time
- $e_s - e_a$ saturation vapour pressure deficit [kPa]
- Δ slope of the vapour pressure curve [kPa °C^{-1}];
- γ psychrometric constant [kPa °C^{-1}].

As specified in the description of each parameter, all meteorological variables should be either measured or converted at 2 m of height to maintain uniformity, and in general the obtained values refer to an extensive surface of green grass that covers the soil which is not suffering from water scarcity conditions. In our case, the reference evapotranspiration is calculated for all the pixels in the area of interest, using the latest available meteorological data at 2 m height and in particular the temperature and wind speed. Even though the vegetation cover of the individual pixel may not correspond at all to the description of the reference plot defined in [3], the evapotranspiration value is still considered representative of the different evaporation and transpiration conditions associated to that area: a pixel hit by an higher amount of radiation, higher temperatures and higher wind speeds will necessarily be more stressed than a pixel characterized by lower mean temperatures and by a topography that causes it to be more hours in shade.

Bibliography

[1] *Resource Acquisition and Transport in Vascular Plants*, 2013. [Online, URL: `https://www.slideshare.net/hayabranko/36-lecture-presentation-16973941`, last accessed October 31, 2022].

[2] A. A. Ager, H. K. Preisler, B. Arca, D. Spano, and M. Salis, *Wildfire risk estimation in the Mediterranean area*, Environmetrics, 25 (2014), pp. 384–396.

[3] R. G. Allen, L. S. Pereira, D. Raes, M. Smith, et al., *FAO Irrigation and drainage paper No. 56*, Rome: Food and Agriculture Organization of the United Nations, 56 (1998), p. e156.

[4] W. A. Allen, H. W. Gausman, and A. J. Richardson, *Mean effective optical constants of cotton leaves*, JOSA, 60 (1970), pp. 542–547.

[5] W. A. Allen and A. J. Richardson, *Interaction of light with a plant canopy*, JOSA, 58 (1968), pp. 1023–1028.

[6] E. Aragoneses, M. García, M. Salis, L. M. Ribeiro, and E. Chuvieco, *Classification and mapping of European fuels using a hierarchical-multipurpose fuel classification system*, Earth System Science Data Discussions, (2022), pp. 1–38.

[7] F. Baret and S. Buis, *Estimating canopy characteristics from remote sensing observations: Review of methods and associated problems*, Advances in land remote Sensing, (2008), pp. 173–201.

[8] K. Berger, C. Atzberger, M. Danner, G. D'Urso, W. Mauser, F. Vuolo, and T. Hank, *Evaluation of the PROSAIL model capabilities for future hyperspectral model environments: A review study*, Remote Sensing, 10 (2018), p. 85.

[9] P. Bowyer and F. Danson, *Sensitivity of spectral reflectance to variation in live fuel moisture content at leaf and canopy level*, Remote Sensing of Environment, 92 (2004), pp. 297–308.

[10] L. S. Bradshaw, *The 1978 national fire-danger rating system: technical documentation*, vol. 169, US Department of Agriculture, Forest Service, Intermountain Forest and Range . . . , 1984.

[11] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, and G. Varoquaux, *API design for machine learning software: experiences from the scikit-learn*

project, in ECML PKDD Workshop: Languages for Data Mining and Machine Learning, 2013, pp. 108–122.

[12] R. E. Burgan, *Behave: fire behavior prediction and fuel modeling system, fuel subsystem*, vol. 167, Intermountain Forest and Range Experiment Station, Forest Service, US . . . , 1984.

[13] ——, *1988 revisions to the 1978 national fire-danger rating system*, vol. 273, US Department of Agriculture, Forest Service, Southeastern Forest Experiment . . . , 1988.

[14] ——, *Monitoring vegetation greenness with satellite data*, vol. 297, US Department of Agriculture, Forest Service, Intermountain Research Station, 1993.

[15] R. E. Burgan, R. W. Klaver, and J. M. Klaver, *Fuel models and fire potential from satellite and surface observations*, International journal of wildland fire, 8 (1998), pp. 159–170.

[16] G. M. Byram and G. M. Jemison, *Solar radiation and forest fuel moisture*, Journal of Agricultural Research, 67 (1943), pp. 149–176.

[17] G. Caccamo, L. Chisholm, R. Bradstock, M. L. Puotinen, and B. Pippen, *Monitoring live fuel moisture content of heathland, shrubland and sclerophyll forest in south-eastern Australia using MODIS data*, International Journal of Wildland Fire, 21 (2011), pp. 257–269.

[18] A. Casajus Valles, M. Marin Ferrer, K. Poljanšek, I. Clark, T. Tagarev, X. Romão, A. de Jager, S. Pelaez, J. Mysiak, G. Papadopoulos, et al., *Conclusions and final recommendations*, Science for Disaster Risk Management 2020-Acting Today, Protecting Tomorrow, (2020).

[19] A. Casu and M. Loi, *A Contribution to Regional Planning Finalized for Fire Resilience*, in International Symposium on New Metropolitan Perspectives, Springer, 2018, pp. 571–578.

[20] P. Ceccato, S. Flasse, and J.-M. Gregoire, *Designing a spectral index to estimate vegetation water content from remote sensing data: Part 2. Validation and applications*, Remote Sensing of Environment, 82 (2002), pp. 198–207.

[21] P. Ceccato, S. Flasse, S. Tarantola, S. Jacquemoud, and J.-M. Grégoire, *Detecting vegetation leaf water content using reflectance in the optical domain*, Remote sensing of environment, 77 (2001), pp. 22–33.

[22] P. Ceccato, N. Gobron, S. Flasse, B. Pinty, and S. Tarantola, *Designing a spectral index to estimate vegetation water content from remote sensing data: Part 1: Theoretical approach*, Remote sensing of environment, 82 (2002), pp. 188–197.

[23] U. A. CERL, *GRASS 4.1 Reference Manual*, US Army Corps of Engineers, Construction Engineering Research Laboratories, Champaign, Illinois, (1993), pp. 1–425.

[24] S. Chaurasia and V. Dadhwal, *Comparison of principal component inversion with VI-empirical approach for LAI estimation using simulated reflectance data*, International Journal of Remote Sensing, 25 (2004), pp. 2881–2887.

[25] A. P. Chessa and A. Delitala, *Il clima della Sardegna*, Chiarella, 1997.

[26] E. Chuvieco, D. Cocero, D. Riano, P. Martin, J. Martinez-Vega, J. De La Riva, and F. Pérez, *Combining NDVI and surface temperature for the estimation of live fuel moisture content in forest fire danger rating*, Remote Sensing of Environment, 92 (2004), pp. 322–331.

[27] E. Chuvieco, I. González, F. Verdú, I. Aguado, and M. Yebra, *Prediction of fire occurrence from live fuel moisture content measurements in a Mediterranean ecosystem*, International Journal of Wildland Fire, 18 (2009), pp. 430–441.

[28] J. Clevers and W. Verhoef, *LAI estimation by means of the WDVI: A sensitivity analysis with a combined PROSPECT-SAIL model*, Remote Sensing Reviews, 7 (1993), pp. 43–64.

[29] B. Combal, F. Baret, M. Weiss, A. Trubuil, D. Mace, A. Pragnere, R. Myneni, Y. Knyazikhin, and L. Wang, *Retrieval of canopy biophysical variables from bidirectional reflectance: Using prior information to solve the ill-posed inverse problem*, Remote sensing of environment, 84 (2003), pp. 1–15.

[30] F. Consortium, *Fireurisk—developing a holistic, risk-wise strategy for european wildfire management*, Environmental Sciences Proceedings, 17 (2022), p. 16.

[31] J. M. Costa-Saura, Á. Balaguer-Beser, L. A. Ruiz, J. E. Pardo-Pascual, and J. L. Soriano-Sancho, *Empirical models for spatio-temporal live fuel moisture content estimation in mixed mediterranean vegetation areas using sentinel-2 indices and meteorological data*, Remote Sensing, 13 (2021), p. 3726.

[32] C. L. Cover, *European Union, Copernicus Land Monitoring Service*, Eur. Environ. Agency, (2018).

[33] H. Croft, J. Chen, Y. Zhang, A. Simic, T. Noland, N. Nesbitt, and J. Arabian, *Evaluating leaf chlorophyll content prediction from multispectral remote sensing data within a physically-based modelling framework*, ISPRS Journal of Photogrammetry and Remote Sensing, 102 (2015), pp. 85–95.

[34] À. Cunill Camprubí, P. González-Moreno, and V. Resco de Dios, *Live fuel moisture content mapping in the Mediterranean Basin using random forests and combining MODIS spectral and thermal data*, Remote Sensing, 14 (2022), p. 3162.

[35] M. Danner, K. Berger, M. Wocher, W. Mauser, and T. Hank, *Fitted PROSAIL parameterization of leaf inclinations, water content and brown pigment content for winter wheat and maize canopies*, Remote Sensing, 11 (2019), p. 1150.

[36] F. Danson, M. Steven, T. Malthus, and J. Clark, *High-spectral resolution data for determining leaf water content*, International Journal of Remote Sensing, 13 (1992), pp. 461–470.

[37] J. E. Deeming, *National fire-danger rating system*, vol. 84, Rocky Mountain Forest and Range Experiment Station, Forest Service, US . . . , 1972.

[38] J. E. Deeming, R. E. Burgan, and J. D. Cohen, *The national fire-danger rating system, 1978*, vol. 39, Department of Agriculture, Forest Service, Intermountain Forest and Range . . . , 1977.

[39] P. Dennison, D. Roberts, S. Peterson, and J. Rechel, *Use of normalized difference water index for monitoring live fuel moisture*, International journal of remote sensing, 26 (2005), pp. 1035–1042.

[40] P. E. Dennison, M. A. Moritz, and R. S. Taylor, *Evaluating predictive models of critical live fuel moisture in the Santa Monica Mountains, California*, International Journal of Wildland Fire, 17 (2008), pp. 18–27.

[41] V. Di Biase and G. Laneve, *Geostationary sensor based forest fire detection and monitoring: An improved version of the sfide algorithm*, Remote Sensing, 10 (2018), p. 741.

[42] J. A. Duffie and W. A. Beckman, *Solar engineering of thermal processes*, Wiley New York, 1980.

[43] E. F. F. I. S. (EFFIS), *Development of a European Fuel Map*, JRC Contract Number 384347, (2017).

[44] M. Enna, *Il peggiore incendio subito dalla Sardegna è rimasto senza colpevoli*. [Online, URL: `https://www.agi.it/cronaca/news/2022-07-22/incendi-sardegna-rogo-montiferru-2021-senza-colpevoli-17526940/`, last accessed January 10, 2023].

[45] J.-B. Feret, C. François, G. P. Asner, A. A. Gitelson, R. E. Martin, L. P. Bidel, S. L. Ustin, G. Le Maire, and S. Jacquemoud, *PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments*, Remote sensing of environment, 112 (2008), pp. 3030–3043.

[46] J.-B. Féret, A. Gitelson, S. Noble, and S. Jacquemoud, *PROSPECT-D: towards modeling leaf optical properties through a complete lifecycle*, Remote Sensing of Environment, 193 (2017), pp. 204–215.

[47] M. A. Finney, C. W. McHugh, I. C. Grenfell, K. L. Riley, and K. C. Short, *A simulation of probabilistic wildfire risk components for the continental United States*.

[48] P. Fiorucci, F. Gaetani, R. Minciardi, and A. Scipioni, *RISICO: A system for dynamic wildfire risk assessment in Italy*, Forest Ecology and Management, (2006), p. S23.

[49] T. H. Fletcher, B. M. Pickett, S. G. Smith, G. S. Spittle, M. M. Woodhouse, E. Haake, and D. R. Weise, *Effects of moisture on ignition behavior of moist California chaparral and Utah leaves*, Combustion Science and Technology, 179 (2007), pp. 1183–1203.

[50] M. A. Fosberg, *Climatological influences on moisture characteristics of dead fuel: theoretical analysis*, Forest Science, 17 (1971), pp. 64–72.

[51] M. García, E. Chuvieco, H. Nieto, and I. Aguado, *Combining AVHRR and meteorological data for estimating live fuel moisture content*, Remote Sensing of Environment, 112 (2008), pp. 3618–3627.

[52] B. Garnier and A. Ohmura, *A method of calculating the direct shortwave radiation income of slopes*, Journal of Applied Meteorology and Climatology, 7 (1968), pp. 796–800.

[53] H. W. Gausman, *Plant leaf optical properties in visible and near-infrared light.*, Graduate Studies, (1985).

[54] L. Giglio, W. Schroeder, and C. O. Justice, *The collection 6 MODIS active fire detection algorithm and fire products*, Remote sensing of environment, 178 (2016), pp. 31–41.

[55] H. T. Gisborne, *The principles of measuring forest fire danger*, Journal of Forestry, 34 (1936), pp. 786–793.

[56] Z. Goncalves and L. Lourenco, *Meteorological index of forest fire risk in the Portuguese mainland territory*, in Proceedings of the International Conference on Forest Fire Research, vol. 7, 1990, pp. 1–14.

[57] V. Gond, D. G. De Pury, F. Veroustraete, and R. Ceulemans, *Seasonal variations in leaf area index, leaf chlorophyll, and water content; scaling-up to estimate fAPAR and carbon balance in a multilayer, multispecies temperate forest*, Tree physiology, 19 (1999), pp. 673–679.

[58] M. González-Sanpedro, T. Le Toan, J. Moreno, L. Kergoat, and E. Rubio, *Seasonal variations of leaf area index of agricultural fields retrieved from Landsat data*, Remote Sensing of Environment, 112 (2008), pp. 810–824.

[59] Google, *Google Street View, SP15, Sardinia.* [Online resource, URL: `https://goo.gl/maps/jaSFv94c9cnprSuk9`, last accessed January 11, 2023].

[60] K. S. Gould, *Nature's Swiss army knife: the diverse protective roles of anthocyanins in leaves*, Journal of Biomedicine and Biotechnology, 2004 (2004), p. 314.

[61] D. Griffiths, *Improved formula for the drought factor in McArthur's Forest Fire Danger Meter*, Australian Forestry, 62 (1999), pp. 202–206.

[62] J. L. Gómez-Dans, *PROSAIL Python Bindings.* [Online, URL: `https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=532`, last accessed October 31, 2022].

[63] D. Haboudane, J. R. Miller, N. Tremblay, P. J. Zarco-Tejada, and L. Dextraze, *Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture*, Remote sensing of environment, 81 (2002), pp. 416–426.

[64] X. Hao and J. J. Qu, *Retrieval of real-time live fuel moisture content using MODIS measurements*, Remote Sensing of Environment, 108 (2007), pp. 130–137.

[65] B. Hapke, D. DiMucci, R. Nelson, and W. Smythe, *The cause of the hot spot in vegetation canopies and soils: Shadow-hiding versus coherent backscatter*, Remote Sensing of Environment, 58 (1996), pp. 63–68.

[66] J. Herman and W. Usher, *SALib: An open-source Python library for sensitivity analysis*, Journal of Open Source Software, 2 (2017), p. 97.

[67] B. Hosgood, S. Jacquemoud, G. Andreoli, J. Verdebout, G. Pedrini, and G. Schmuck, *Leaf optical properties experiment 93 (LOPEX93)*, Report EUR, 16095 (1995).

[68] K. Huemmrich, *PBOREAS TE-18 Geosail Canopy Reflectance Model*, 2000. [Online, URL: https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=532, last accessed October 31, 2022].

[69] ——, *The GeoSail model: a simple addition to the SAIL model to describe discontinuous canopy reflectance*, Remote Sensing of Environment, 75 (2001), pp. 423–431.

[70] S. Jacquemoud, *Inversion of the PROSPECT+ SAIL canopy reflectance model from AVIRIS equivalent spectra: theoretical study*, Remote sensing of environment, 44 (1993), pp. 281–292.

[71] S. Jacquemoud, C. Bacour, H. Poilve, and J.-P. Frangi, *Comparison of four radiative transfer models to simulate plant canopies reflectance: Direct and inverse mode*, Remote Sensing of Environment, 74 (2000), pp. 471–481.

[72] S. Jacquemoud and F. Baret, *PROSPECT: A model of leaf optical properties spectra*, Remote sensing of environment, 34 (1990), pp. 75–91.

[73] S. Jacquemoud, F. Baret, and J. Hanocq, *Modeling spectral and bidirectional soil reflectance*, Remote sensing of Environment, 41 (1992), pp. 123–132.

[74] S. Jacquemoud, C. Bidel, and F. Pavan, *ANGERS Leaf Optical Properties Database (2003). Data set. Available on-line [http://ecosis.org] from the Ecological Spectral Information System (EcoSIS)*, 2003.

[75] S. Jacquemoud, W. Verhoef, F. Baret, C. Bacour, P. J. Zarco-Tejada, G. P. Asner, C. François, and S. L. Ustin, *PROSPECT+ SAIL models: A review of use for vegetation characterization*, Remote sensing of environment, 113 (2009), pp. S56–S66.

[76] M. F. Jasinski and P. S. Eagleson, *The structure of red-infrared scattergrams of semivegetated landscapes*, IEEE Transactions on geoscience and remote sensing, 27 (1989), pp. 441–451.

[77] S. Jurdao, E. Chuvieco, and J. M. Arevalillo, *Modelling fire ignition probability from satellite estimates of live fuel moisture content*, Fire Ecology, 8 (2012), pp. 77–97.

[78] T. Kattenborn, *Linking Canopy Reflectance and Plant Functioning through Radiative Transfer Models*, PhD thesis, KIT-Bibliothek, 2019.

[79] B. Koetz, F. Baret, H. Poilvé, and J. Hill, *Use of coupled canopy structure dynamic and radiative transfer models to estimate biophysical canopy characteristics*, Remote Sensing of Environment, 95 (2005), pp. 115–124.

[80] A. Kuusk, *The hot spot effect in plant canopy reflectance*, in Photon-vegetation interactions, Springer, 1991, pp. 139–159.

[81] G. Laneve and L. Fusilli, *High resolution fire hazard index based on satellite images*, Territorium, (2019), pp. 29–41.

[82] G. Laneve, L. Fusilli, and G. Bernini, *Achievements of the PREFER project in the prevention phase of the forest fire management*, in 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), IEEE, 2016, pp. 5765–5768.

[83] G. Laneve, J. Munzer, F. Ferrucci, and F. Batazza, *The development of a fire vulnerability index for the Mediterranean region*, in 2011 IEEE International Geoscience and Remote Sensing Symposium, IEEE, 2011, pp. 4146–4149.

[84] G. Laneve, V. Pampanoni, and R. Uddien Shaik, *The daily fire hazard index: A fire danger rating method for Mediterranean areas*, Remote Sensing, 12 (2020), p. 2356.

[85] H. Lightenthaler, *Chlorophylls and carotenoids: pigments of photosynthetic biomembranes*, Methods in enzymology, 148 (1987), pp. 350–382.

[86] A. S. López, J. San-Miguel-Ayanz, and R. E. Burgan, *Integration of satellite sensor data, fuel type maps and meteorological observations for evaluation of forest fire risk at the pan-European scale*, International Journal of Remote Sensing, 23 (2002), pp. 2713–2719.

[87] X. Ma, A. Huete, N. N. Tran, J. Bi, S. Gao, and Y. Zeng, *Sun-Angle Effects on Remote-Sensing Phenology Observed and Modelled Using Himawari-8*, Remote Sensing, 12 (2020).

[88] E. Marino, M. Guillén-Climent, N. Algeet, J. L. Tomé, and C. Hernando, *Estimation of live fuel moisture content of shrubland using modis and sentinel-2 images*, (2018).

[89] F. Marras, S. De Leo, et al., *L'agricoltura italiana conta. 2020*, (2020).

[90] A. B. Massada, V. C. Radeloff, S. I. Stewart, and T. J. Hawbaker, *Wildfire risk in the wildland–urban interface: a simulation study in northwestern Wisconsin*, Forest Ecology and Management, 258 (2009), pp. 1990–1999.

[91] A. G. McArthur, *Fire behaviour in eucalypt forests*, (1967).

[92] ——, *Weather and grassland fire behavior*, (1967).

[93] J. Muñoz-Sabater, *ERA5-Land hourly data from 1959 to present*, Copernicus Climate Change Service (C3S) Climate Data Store (CDS), (2021).

[94] R. Myneni, Y. Knyazikhin, and T. Park, *MOD15A2H MODIS/terra leaf area index*, FPAR 8-day L4 Global 500 m SIN Grid V006 Data Set, (2015).

[95] V. Nesterov, *Combustibility of forests and methods of its definition*, Goslesbumizdat, Moscow, Leningrad, (1949).

[96] OpenStreetMap contributors, *Planet dump retrieved from `https://planet.osm.org`*. `https://www.openstreetmap.org`, 2017.

[97] V. Pampanoni, G. Laneve, and G. Santilli, *Evaluating Sentinel-3 Viability for Vegetation Canopy Monitoring and Fuel Moisture Content Estimation*, in IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium, IEEE, 2022, pp. 5634–5637.

[98] F. Pastorella, G. Giacovelli, M. Maesano, A. Paletto, S. Vivona, A. Veltri, G. Pellicone, and G. S. Mugnozza, *Social perception of forest multifunctionality in southern Italy: The case of Calabria Region*, Journal of Forest Science, 62 (2016), pp. 366–379.

[99] R. D. Peters and S. D. Noble, *Spectrographic measurement of plant pigments from 300 to 800 nm*, Remote sensing of environment, 148 (2014), pp. 119–123.

[100] Physics Catalyst, *Anatomy of flowering plants Notes.* [Online; last accessed October 31, 2022].

[101] S. Plummer, *The Angular Vegetation Index: an atmospherically resistant index for the second along track scanning radiometer (ATSR-2)*, in Proc. Sixth Int. Symp. Physical Measurements and Signatures in Remote Sensing, Val d'Isere, France., 1994.

[102] H. K. Preisler, D. R. Brillinger, R. E. Burgan, and J. Benoit, *Probability based models for estimation of wildfire risk*, International Journal of wildland fire, 13 (2004), pp. 133–142.

[103] E. Prikaziuk and C. van der Tol, *Global sensitivity analysis of the SCOPE model in Sentinel-3 Bands: thermal domain focus*, Remote sensing, 11 (2019), p. 2424.

[104] X. Quan, B. He, X. Li, and Z. Tang, *Estimation of grassland live fuel moisture content from ratio of canopy water content and foliage dry biomass*, IEEE Geoscience and Remote Sensing Letters, 12 (2015), pp. 1903–1907.

[105] X. Quan, M. Yebra, D. Riaño, B. He, G. Lai, and X. Liu, *Global fuel moisture content mapping from MODIS*, International Journal of Applied Earth Observation and Geoinformation, 101 (2021), p. 102354.

[106] B. Raoult, C. Bergeron, A. L. Alós, J.-N. Thépaut, and D. Dee, *Climate service develops user-friendly data store*, ECMWF newsletter, 151 (2017), pp. 22–27.

[107] D. Riaño, P. Vaughan, E. Chuvieco, P. J. Zarco-Tejada, and S. L. Ustin, *Estimation of fuel moisture content by inversion of radiative transfer models to simulate equivalent water thickness and dry matter content: analysis at leaf and canopy level*, IEEE Transactions on Geoscience and Remote Sensing, 43 (2005), pp. 819–826.

[108] K. Richter, F. Vuolo, G. D'Urso, and M. Palladino, *Evaluation of near-surface soil water status through the inversion of soil–canopy radiative transfer models in the reflective optical domain*, International journal of remote sensing, 33 (2012), pp. 5473–5491.

[109] Roberta Secci, *Il 2021 è stato l'anno peggiore per gli incendi in sardegna dal 1998.* [Online, URL: `https://www.agi.it/cronaca/news/2022-06-18/sardegna-incendi-anno-dati-record-17137552/` last accessed October 31, 2022].

[110] R. C. Rothermel, *A mathematical model for predicting fire spread in wildland fuels*, vol. 115, Intermountain Forest & Range Experiment Station, Forest Service, US . . . , 1972.

[111] M. Salis, A. A. Ager, B. Arca, M. A. Finney, V. Bacciu, P. Duce, and D. Spano, *Assessing exposure of human and ecological values to wildfire in Sardinia, Italy*, International Journal of Wildland Fire, 22 (2013), pp. 549–565.

[112] M. Salis, L. Del Giudice, B. Arca, A. A. Ager, F. Alcasena-Urdiroz, O. Lozano, V. Bacciu, D. Spano, and P. Duce, *Modeling the effects of different fuel treatment mosaics on wildfire spread and behavior in a Mediterranean agro-pastoral area*, Journal of environmental management, 212 (2018), pp. 490–505.

[113] A. Saltelli, *Making best use of model evaluations to compute sensitivity indices*, Computer physics communications, 145 (2002), pp. 280–297.

[114] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola, *Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index*, Computer physics communications, 181 (2010), pp. 259–270.

[115] J. San-Miguel-Ayanz, T. Durrant, et al., *Advance EFFIS report on forest fires in Europe, Middle East and North Africa 2021*, (2021).

[116] L. N. Sardegna, *Auto s'incendia e le fiamme dilagano nella campagna tra Bonarcado e Santu Lussurgiu.* [Online, URL: https://www.lanuovasardegna.it/oristano/cronaca/2021/07/23/news/auto-s-incendia-e-le-fiamme-dilagano..., last accessed January 10, 2023].

[117] F. P. Schoenberg, R. Peng, Z. Huang, and P. Rundel, *Detection of non-linearities in the dependence of burn area on fuel age and climatic variables*, International Journal of Wildland Fire, 12 (2003), pp. 1–6.

[118] W. Schroeder, P. Oliva, L. Giglio, and I. A. Csiszar, *The New VIIRS 375 m active fire detection data product: Algorithm description and initial assessment*, Remote Sensing of Environment, 143 (2014), pp. 85–96.

[119] Q. Shu, X. Quan, M. Yebra, X. Liu, L. Wang, and Y. Zhang, *Evaluating the sentinel-2a satellite data for fuel moisture content retrieval*, in IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium, IEEE, 2019, pp. 9416–9419.

[120] A. Simard, *The moisture content of forest fuels–I. A review of the basic concepts. Canadian Department of Forest and Rural Development, Forest Fire Research Institute*, tech. rep., Information Report FF-X-14, Ottawa, Ontario, 1968.

[121] S. Singh, *Understanding the role of slope aspect in shaping the vegetation attributes and soil properties in Montane ecosystems*, Tropical Ecology, 59 (2018), pp. 417–430.

[122] I. M. Sobol, *Sensitivity analysis for non-linear mathematical models*, Mathematical modelling and computational experiment, 1 (1993), pp. 407–414.

[123] ——, *Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates*, Mathematics and computers in simulation, 55 (2001), pp. 271–280.

[124] D. Sonntag, *Important new values of the physical constants of 1986, vapour pressure formulations based on the ITS-90, and psychrometer formulae*, Zeitschrift für Meteorologie, 40 (1990), pp. 340–344.

[125] D. Stow, M. Niphadkar, and J. Kaiser, *Time series of chaparral live fuel moisture maps derived from MODIS satellite data*, International Journal of Wildland Fire, 15 (2006), pp. 347–360.

[126] G. Suits, G. Safir, and A. Ellingbroe, *Prediction of directional reflectance of a corn field under stress*, Nasa. Manned Spacecraft Center 4th Ann. Earth Resources Program Rev., Vol. 2, (1972).

[127] C. Tucker, *The NDVI has been proven to be well correlated with various vegetation parameters, such as green biomass*, Rouse et al, (1979), pp. 605–790.

[128] C. J. Tucker, *Remote sensing of leaf water content in the near infrared*, Remote sensing of Environment, 10 (1980), pp. 23–32.

[129] C. J. Tucker and P. Sellers, *Satellite remote sensing of primary production*, International Journal of Remote Sensing, 7 (1986), pp. 1395–1416.

[130] F. P. L. (US), *Wood handbook: wood as an engineering material*, The Laboratory, 1987.

[131] C. Van Wagner et al., *Structure of the Canadian forest fire weather index*, vol. 1333, Environment Canada, Forestry Service Ontario, 1974.

[132] R. Velez, *Manual de Operaciones Contra Incendios Forestales*, Min. Agricultura, ICONA, Madrid, (1993).

[133] F. Ventura, V. Marletto, and F. Zinoni, *Un metodo per il calcolo dell'indice meteorologico del rischio di incendio forestale*, Sherwood, 68 (2001), pp. 13–18.

[134] W. Verhoef, *Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model*, Remote sensing of environment, 16 (1984), pp. 125–141.

[135] W. Verhoef, *Theory of radiative transfer models applied in optical remote sensing of vegetation canopies*, Wageningen University and Research, 1998.

[136] W. Verhoef and H. Bach, *Simulation of hyperspectral and directional radiance images using coupled biophysical and atmospheric radiative transfer models*, Remote sensing of environment, 87 (2003), pp. 23–41.

[137] F. Warmerdam, *The geospatial data abstraction library*, in Open source approaches in spatial data handling, Springer, 2008, pp. 87–104.

[138] M. Weiss, F. Baret, S. Garrigues, and R. Lacaze, *LAI and fAPAR CYCLOPES global products derived from VEGETATION. Part 2: validation and comparison with MODIS collection 4 products*, Remote sensing of Environment, 110 (2007), pp. 317–331.

[139] A. Wexler, *Vapour pressure formulation for water in the range 0*, (1976).

[140] J. Xie, T. Qi, W. Hu, H. Huang, B. Chen, and J. Zhang, *Retrieval of Live Fuel Moisture Content Based on Multi-Source Remote Sensing Data and Ensemble Deep Learning Model*, Remote Sensing, 14 (2022), p. 4378.

[141] M. Yebra and E. Chuvieco, *Generation of a species-specific look-up table for fuel moisture content assessment*, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2 (2009), pp. 21–26.

[142] ——, *Linking ecological information and radiative transfer models to estimate fuel moisture content in the Mediterranean region of Spain: Solving the ill-posed inverse problem*, Remote Sensing of Environment, 113 (2009), pp. 2403–2411.

[143] M. Yebra, E. Chuvieco, and D. Riaño, *Estimation of live fuel moisture content from MODIS images for fire risk assessment*, Agricultural and forest meteorology, 148 (2008), pp. 523–536.

[144] M. Yebra, P. E. Dennison, E. Chuvieco, D. Riaño, P. Zylstra, E. R. Hunt Jr, F. M. Danson, Y. Qi, and S. Jurdao, *A global review of remote sensing of live fuel moisture content for fire danger assessment: Moving towards operational products*, Remote Sensing of Environment, 136 (2013), pp. 455–468.

[145] M. Yebra, G. Scortechini, A. Badi, M. E. Beget, M. M. Boer, R. Bradstock, E. Chuvieco, F. M. Danson, P. Dennison, V. Resco de Dios, et al., *Globe-LFMC, a global plant water status database for vegetation ecophysiology and wildfire applications*, Scientific data, 6 (2019), pp. 1–8.

[146] D. Zanaga, R. Van De Kerchove, D. Daems, W. De Keersmaecker, C. Brockmann, G. Kirches, J. Wevers, O. Cartus, M. Santoro, S. Fritz, M. Lesiv, M. Herold, N.-E. Tsendbazar, P. Xu, F. Ramoino, and O. Arino, *ESA WorldCover 10 m 2021 v200*, Oct. 2022.

[147] P. J. Zarco-Tejada, C. Rueda, and S. Ustin, *Water content estimation in vegetation with MODIS reflectance data and model inversion methods*, Remote Sensing of Environment, 85 (2003), pp. 109–124.

[148] L. Zhu, G. I. Webb, M. Yebra, G. Scortechini, L. Miller, and F. Petitjean, *Live fuel moisture content estimation from MODIS: A deep learning approach*, ISPRS Journal of Photogrammetry and Remote Sensing, 179 (2021), pp. 81–91.

[149] P. J. Zylstra, *Forest flammability: modelling and managing a complex system*, PhD thesis, UNSW Sydney, 2011.