Aix Marseille Université

NNT : 0000AIXM0000

THÈSE DE DOCTORAT

Soutenue à Aix-Marseille Université
Le 08 avril 2025 par

Jorge CASTEL-CLAVERA

Analyse et modélisation des dynamiques spatio-temporelles des incendies de forêt en région méditerranéenne : vers de nouveaux indicateurs de danger

Discipline

Sciences de l'environnement

Spécialité
Écologie

École doctorale

ED 251 – Sciences de l'Environnement

Laboratoire/Partenaires de recherche
Institut National de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE)

Composition du jury

Vincent MORON Professeur – AMU	Président du jury
Florent MOUILLOT Directeur de recherche – IRD	Rapporteur
Giada ADELFIO Professeure des universités – Università degli Studi di Palermo	Rapportrice
Imma OLIVERAS MENOR Directeur de recherche – IRD	Examinatrice
Maxime SOMA Ingénieur de recherche – INRAE	Examinateur
Jean-Luc Dupuy Directeur de recherche – INRAE	Directeur de thèse
François Pimont Ingénieur de recherche – INRAE	Co-encadrant de thèse (Membre invité)
Thomas Opitz Directeur de recherche – INRAE	Co-encadrant de thèse (Membre invité)

INRAE

Affidavit

I, undersigned, Jorge Castel-Clavera, hereby declare that the work presented in this manuscript is my own work, carried out under the scientific supervision of Jean-Luc Dupuy, in accordance with the principles of honesty, integrity and responsibility inherent to the research mission. The research work and the writing of this manuscript have been carried out in compliance with both the French national charter for Research Integrity and the Aix-Marseille University charter on the fight against plagiarism.

This work has not been submitted previously either in this country or in another country in the same or in a similar version to any other examination body.

Place Avignon, date 24 janvier 2025

Signature :



Cette œuvre est mise à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.

Liste de publications et/ou brevets et participation aux conférences

1) Liste des publications et/ou brevet réalisées dans le cadre du projet de thèse :
   a) Castel-Clavera, J., Pimont, F., Opitz, T., Ruffault, J., Barbero, R., Allard, D., Dupuy, J-L., 2025. A comparative analysis of fire-weather indices for enhanced fire activity prediction with probabilistic approaches. *Agricultural and Forest Meteorology* 361. https://doi.org/10.1016/j.agrformet.2024.110315.
   b) Castel-Clavera, J., Pimont, F., Opitz, T., Ruffault, J., Rivière, M., Dupuy, J-L., 2022. Disentangling the Factors of Spatio-Temporal Patterns of Wildfire Activity in South-Eastern France. *International Journal of Wildland Fire* 32. https://doi.org/10.1071/WF22086.
   c) Pimont, F., Ruffault, J., Opitz, T., Fargeon, H., Barbero, B., Castel-Clavera, J., Martin-StPaul, N., Rigolot, E., Dupuy, J-L., 2022. Future Expansion, Seasonal Lengthening, and Intensification of Fire Activity under Climate Change in Southeastern France. *International Journal of Wildland Fire* 32. https://doi.org/10.1071/WF22103.
   d) En préparation: Castel-Clavera, J., Pimont, F., Opitz, T., Ruffault, J., Olioso, A., Penot, V., Rodríguez-Suquet, R., Dupuy, J-L., no year. Fire danger prediction enhanced through remote sensing in Mediterranean France.
2) Participation aux conférences et écoles d'été au cours de la période de thèse :
   a) Présentation orale 9th ICFFR (International Conference on Forest Fire Research). Communication écrite : Castel-Clavera, J., Pimont, F., Opitz, T., Ruffault, J., et Dupuy, J-L. 2022. Modelling the Influence of Regional Landscape Drivers on Spatio-Temporal Patterns of Wildfire Activity. In *Advances in Forest Fire Research 2022*, par Domingos Xavier Viegas and Luís Mário Ribeiro, 1228-33, 1.ª Edição. Imprensa da Universidade de Coimbra. https://doi.org/10.14195/978-989-26-2298-9_186.
   b) Présentation orale au séminaire du réseau d'animation scientifique RisqFor (département ECODIV, INRAE) sur l'évaluation et la gestion des risques en forêt.

Résumé

L'activité des incendies dans la région méditerranéenne française est influencée par les conditions météorologiques, comme l'ont montré des recherches antérieures. Les facteurs humains jouent un rôle crucial en raison de la forte densité de population et de l'utilisation et occupation des sols. Cependant, la combinaison de tous les facteurs d'incendie présente des limites persistantes. L'identification des facteurs climatiques et non climatiques qui régissent les régimes d'incendies pose de nombreux défis dans le sud-est de la France. Cette thèse vise à définir des indicateurs de danger d'incendie en France Méditerranéenne, utilisables par des acteurs scientifiques et opérationnels. Trois objectifs spécifiques sont définis pour répondre à l'objectif principal du projet. Premièrement, identifier les facteurs d'incendie spatio-temporels explicites en testant leur contribution aux régimes récents dans le sud-est de la France. Deuxièmement, examiner si les données météorologiques, autres que celles actuellement utilisées par les services opérationnels, peuvent améliorer la prédiction des activités incendiaires à l'échelle quotidienne. Troisièmement, déterminer si des données télédétection optique peuvent compléter ou remplacer les variables météorologiques et/ou paysagères dans la prévision des incendies. L'étude intègre des variables non climatiques dans le cadre de modélisation *Firelihood*, un modèle bayésien d'activité des feux, qui intègre composantes spatio-temporelles. Ensuite, avec une nouvelle méthode d'évaluation, on analyse les variables météo-feu comme prédicteurs de l'activité d'incendie. Enfin, de nouveaux modèles, intégrant des données de télédétection optique, mesurent la contribution des variables météo-feu, de télédétection et paysage aux prévisions d'incendies. Les résultats montrent que des indices feu-météo adéquats améliorent les prévisions d'activité de feu, mais les effets spatio-temporels sont essentiels pour garantir des prévisions précises. Plusieurs facteurs de paysage ont contribué autant que les indices météo-feu à la distribution des occurrences de feu. La télédétection optique améliore surtout pour la prévision des grands incendies. Les patrons temporels d'incendie sont principalement influencés par les changements du climat-feu et des effets inexpliqués. Ces résultats permettront d'améliorer la prévision du danger d'incendie à court et long terme, perfectionnant les classes de danger opérationnelles et créant des modèles régionaux spécifiques.

Abstract

Weather conditions significantly influence fire activity in the French Mediterranean region as shown in previous research. Human factors also play a crucial role due to high population density and anthropogenic land use and land cover. However, the combination and representation of all fire factors in a single framework presents persistent limitations. Identifying climatic and non-climatic drivers of fire regimes entails numerous challenges in South-eastern France. This thesis aims to define new fire danger indicators in Mediterranean France for scientific and operational use. Three specific objectives are defined to address the major aim of the project. First, the identification of the explicit spatio-temporal fire factors by testing their contribution to recent fire regimes in South-Eastern France. The second objective is to investigate if the weather information, other that the currently used operationally, can improve the fire activity predictability at the daily scale. The third objective is to determine if specific optical remote sensing data can improve or replace weather and/or landscape variables on the predictions of fire activities. First, the study incorporates non-climatic variables into the *Firelihood* framework, a Bayesian model of fire activity, with enhanced spatio-temporal components to assess unexplained variations. After, with the newly developed evaluation method, it evaluates meteorological variables and fire-weather indices as predictors of fire occurrence and size. Finally, new models, integrating optical remote sensing data, explore the contribution of weather, remote sensing, and landscape variables to fire predictions. Results reveal that adequate fire-weather indices improve fire activity predictions but are insufficient on their own. Spatial and temporal effects are essential for predictions accuracy. Several landscape drivers contributed as much as fire-weather to the distribution of fire occurrence but less to large fires, whereas optical remote sensing mainly predicts large fires. Temporal fire regime patterns are primarily influenced by fire-weather changes and unexplained effects. These findings enhance short-term fire danger forecasting in Mediterranean France improving operational danger classes and creating specific regional models. In addition, the *Firelihood* framework will be used to create a French national risk assessment and long-term future projections.

Remerciements

I would like to express my deepest gratitude to my thesis director, Jean-Luc Dupuy, for the given opportunity and guidance. I am very grateful to my thesis mentors François Pimont and Thomas Opitz, also to my colleague Julien Ruffault, for their invaluable patience, constructive feedback, and intangible support. I also could not have undertaken this journey without my thesis committee, who generously provided knowledge and expertise. Additionally, this thesis would not have been possible without the generous support from Provence Alpes Côte d'Azur, who financed part of my research contract. I would like to acknowledge all the coauthors from the scientific articles for their contributions to the research, and the Office National de Forêt for the data they have provided, including the bunt area perimeters, plant sensibility map, and their advice on the landscape factors.

I would like to extend my sincere thanks to those persons from academia who inspired my career through their advice, support, and encouragement in order of appearance. Santiago Guiu Torres, natural sciences teacher in La Salle Tarragona. The professors from Universitat de Lleida Cristina Vega García, Álvaro Aunós Gómez, Domingo Molina Terrén, and Víctor Resco de Dios. Finally, thanks to Josep Maria Serra Díaz for introducing me to the research team that hosted my thesis work.

I would be remiss in not mentioning my colleagues from URFM for the best reception, work ambiance and the help, and my family and close ones for your unconditional and wholehearted support.

Table des matières

Introduction **15**

1. Background – State of the art 20

1.1. Overview of main factors of forest fires: weather, fuel, landscape, and human activities. 20

1.2. Fire as a natural process in Mediterranean and beyond. 23

1.3. Fire excluded from the life cycle. 24

1.4. The French Mediterranean region and its fire regime 25

1.4.1. The study area 25

1.4.2. Fire records in Mediterranean France 30

1.5. Perceiving fire as a threat 32

1.6. Methods for fire danger prediction 33

1.6.1. Frequently used fire danger prediction methods 36

1.6.2. Fire danger prediction through remote sensing 37

1.6.3. Fire danger prediction and early alarm systems 39

1.7. Fire danger prediction in France 40

1.8. Limitations for fire weather index approaches 42

1.9. An example of the operational application of a FOP from Canada 44

1.10. A probabilistic framework for fire activities in Mediterranean France 45

2. Influence of the non-climatic factors over the spatiotemporal patterns of wildfire activity and its long-term changes **47**

2.1. Article: Disentangling the factors of spatio-temporal patterns of wildfire activity in south-eastern France 50

2.2. Overview of the results 78

3. Enhancing the short-term prediction of fire danger through fire-weather indices and meteorological variables. **79**

3.1. Article: A comparative analysis of fire-weather indices for enhanced fire activity prediction with probabilistic approaches 81

3.2. Overview of the results 136

3.3. Remaining unexplained spatial and temporal variability 136

4. Advancements from optical Remote Sensing for spatial and temporal fire activity and danger prediction. 138
4.1. Optical remote sensing data (MODIS and Sentinel2) 139
4.1.1. MODIS data presentation 139
4.1.2. Sentinel data presentation and treatment 140
4.2. Article: Fire danger prediction enhanced through optical remote sensing in Mediterranean France 144
4.3. Overview of the results 176
5. Discussion 178
5.1. Insights from modeling 178
5.1.1. Weather indices 178
5.1.2. Land Use and Land Cover (LULC): 179
5.1.3. Optical Remote Sensing: 179
5.1.4. Relative contributions of factors to spatio-temporal variability of fire activity 180
5.1.5. Fire regime trends 181
5.2. Methodological advancements 182
5.2.1. Variables' selection and performance evaluation 183
5.2.2. Contributions of different predictor types to improved model performance 184
5.2.3. Handling different spatial and temporal scales in models 186
5.3. Potential Applications 187
5.3.1. Operational short-term forecasting 187
5.3.2. Hazard Mapping and long-term past and future trends 190
5.4. Future improvements and extensions 191
5.4.1. Agricultural abandonment cartography and other LULC enhancement 191
5.4.2. Use of remote sensing to obtain fuel conditions. 192
5.4.3. Mechanistic modelling for fuel moisture content 193
5.4.4. Use reforecast data instead of reanalysis weather data. 194
5.4.5. Exploring the drivers of winter fires 194
5.4.6. More accurate assessment of the impact of climate change 195
5.4.7. Further perspectives 196

6. Conclusion **198**

Bibliography **201**

Supplementary materials **214**

A. Exploratory work for fire danger prediction using MODIS composite data. 215

1. Article: Disentangling the factors of spatio-temporal patterns of wildfire activity in South-eastern France 219

2. Article: A comparative analysis of fire-weather indices for enhanced fire occurrence prediction with probabilistic approaches 230

3. Article: Fire danger prediction enhanced through optical remote sensing in Mediterranean France 241

Introduction

Wildfires are an increasing threat in many regions of the world due to climate change, human activities, and changes on land use and land cover (M. W. Jones et al., 2022). Wildfires are able to burn vast areas, which entails threats to biodiversity, human health, infrastructure, and the economy. In addition, large quantities of carbon dioxide are emitted to the atmosphere by burning vegetation. The drought conditions in France are on the rise, nonetheless, fire activity has tended to decrease over the recent decades thanks to efficient prevention, management, and firefighting efforts (Curt & Frejaville, 2018). In southeastern Europe, this decreasing trend is predicted to reverse (Pimont et al., 2023). Future projection studies have shown that fire activities may double under the +4ºC global warming reference scenario for France (Pimont et al., 2022), while the frequency of extreme wildfire events is expected to increase (Duane et al., 2021; Ruffault et al., 2020). Due to this complex and varying situation it is necessary to increase the prediction capabilities of wildfire occurrences and their potential intensity.

Forest fires (or wildfires) are part of the natural ecosystem dynamics on the Mediterranean basin, as well human populations have been confronted to the existence of fire since very early forms of settlements (Roebroeks & Villa, 2011). Fires are a fundamental element of the Mediterranean ecosystems (M. W. Jones et al., 2022; Kelly et al., 2020; Pausas & Keeley, 2019), which acts as a regulator of plant populations, for example promoting the regeneration of regular forest stands, or opening patches on forest masses that will be covered in pastures for herbivores. However, humans have also profoundly altered the landscape through their many uses of the land, and their activities have a major impact on the climate. More specifically, fire control policies have led to fire exclusion, which, with lack of fuel management, creates an accumulation of forest fuels that continue to dry during long drought periods (Hurteau et al., 2024; Piñol et al., 2005).

Wildfires are a complex phenomenon influenced by a range of variables, including fuel (quantity, type, and moisture content), meteorological conditions (such as wind speed, humidity, and temperature), topography, and human activities. The measurements that quantify all available fire-related variables come from weather observations and remote

sensing. Weather is a major factor controlling fire ignition and behavior. It acts through the direct effect of wind on combustion and heat transfer processes and through the indirect effects of wind, air temperature, air humidity and precipitations, on fuel moisture (Argañaraz et al., 2018; Chuvieco et al., 2009; Nolan et al., 2016; Sharples et al., 2009). Remote sensing is a direct source of information regarding the fuel presence and moisture content, thanks to the interactions between the plant structures and the electromagnetic radiation (Gale et al., 2021; Rao et al., 2020; Wang et al., 2013; Yebra et al., 2018). Human uses of the land drive the ignitions patterns and seasonality as most fires in the Mediterranean have human origin (Balch et al., 2017; Costafreda-Aumedes et al., 2017; de Rigo et al., 2018; Ganteaume et al., 2013).

There has been an increase of population in the west of the Mediterranean basin. Therefore, large urbanistic developments have expanded the urban areas towards non-managed forested areas. This exposes the infrastructure and the population to fire risks, while also increasing the exposure of forests to human sources of ignition. Ecosystem services[1] (Millennium Ecosystem Assessment, 2005) also raise concern as large and recurrent fires may alter the vegetation cover and compromise the benefits of the ecosystems. Large and intense fire events are a threat which entails losses to human life, housing, and infrastructure. In addition, in the recent decades there has been an increase of large fire events regardless of improved fire extinction policies that succeed to reduce the total number of fires. This is partially attributed to climate change, because of the trend towards warmer, dryer and longer summers in the Mediterranean region (Barbero et al., 2019; Galizia et al., 2023; Schelhaas et al., 2003). This increasing exposure to wildfires and their impact means that fire hazard management and fire danger prediction are increasingly important issues to reduce the risk (definitions in section 1.5).

Wildfire prevention and firefighting require spatial predictions of fire danger to anticipate the fire activity (i.e., number and size of fires) and the fire behavior (rate of spread, energy released, etc.). Such a predictions are required on daily basis, typically for forecasting the occurrence of fire events and their potential magnitude, but also on a longer term (years, decades) to evaluate the fire hazard. Fire hazard may be viewed as the

[1] Ecosystem services are the benefits on economical, health and quality of life dimensions that the natural ecosystems offer to the large population even without this population being directly connected to the natural areas.

temporal integration of the fire danger in a given location. Regarding climate change, projections of the overall exposure to fires is also important for adapting public policies (Pimont et al., 2022 & 2023).

A commonly used approach for fire danger prediction is the use of fire-weather indices, which indicate the extent to which meteorological conditions are favorable for ignition and fire spread. Many fire-weather indices have been already developed all over the world by adapting them to the specific climatic conditions of the respective regions and/or countries of interest. In France, as well as in most European countries, operational authorities adopted the Canadian Fire Weather Index System as a fire danger rating system. In addition, a range of tools have been developed to produce fire hazard maps that account for the potential intensity and behavior of the fire as a function of the average local atmospheric conditions. These fire-weather indices and maps are used for operational applications as the final value used to assess fire danger.

Several limitations arise from using fire-weather indices for fire danger prediction. First, an index is an abstraction of the fire danger level which not necessarily reflects or directly translates into observed fire activity. As a result, a fire-weather index is not directly related to a probability of fire occurrence nor to potential fire behavior. Second, the non-probabilistic approaches do not account for the uncertainty associated to a measured or calculated danger. Acknowledging this uncertainty is crucial for decision taking in fire management and comparing future scenarios for long term prevision. Third, the fire-weather indices do not include all other non-climatic fire factors as land use and land cover or fuel types and condition, which may entail estimation biases of fire activities. For these reasons, other agencies (e.g., in Canada), do not only rely on fire weather to rate the danger, but also on more sophisticated predictive systems for fire occurrence and behavior. To date, such systems are missing in France.

Probabilistic frameworks have demonstrated their potential on predicting fire activities (Brillinger et al., 2006; Preisler et al., 2004; Woolford et al., 2021; Worsnop et al., 2021). Such frameworks, developed and used in the scientific community, translate the measured fire factors into fire probabilities to bypass the limitations of the fire-weather indices on fire danger predictions. The statistical models can be informed with multiple drivers of fire activity; nevertheless, finding an appropriate combination of the various fire factors of different type and obtained from different data sources is not straightforward. These methods are expected to improve the fire danger estimations in

the short term, and in the long term they can be used to calculate projections of future fire trends.

At the beginning of this thesis, we had a model developed within a probabilistic framework similar to those cited before. This framework uses more recent estimation techniques (e.g., Integrated Nested Laplace Approximation) which demonstrated potential in accurately predicting fire activities in southeastern France. This model primarily utilized an operational weather index as a predictor, along with spatio-temporal random effects to implicitly account for those landscape and seasonal effects that could not be directly included in the model (Pimont et al., 2021). This approach was already a significant improvement compared to using the weather index outside a modeling framework, whether in operational settings or scientific studies.

This thesis aimed to address further methodological and cognitive questions. The three methodological questions are: Can we develop new and robust methods for fire danger assessment and fire hazard mapping based on probabilistic predictions of fire activity that integrate all known fire factors for which data can be retrieved? Can we reduce unexplained random effects that have been identified in previous research? How can we identify and select key variables and assess the predictive performance gains when incorporating them into the model? The three cognitive questions are: What new knowledge about the phenomenon of wildfires can be derived from the above model-building process and the simulation results of this model? Can we determine and compare the relative contributions of different factors (weather, landscape, human) across various spatio-temporal scales? Can we understand past fire regime changes and anticipate the impact of climate change on fire activity, among other insights?

To address these six questions, we pursued three specific objectives, which define the structure of the thesis presentation. The first one was to identify the implicit spatio-temporal factors underlying the estimated random effects observed in previous-studies by testing the contribution of a range of landscape and human activity variables to recent fire regimes at decadal temporal scale in South-Eastern France. The second objective was to investigate if the weather indices or weather variables, other than the ones currently used as operational tools or in previous research, can improve the fire activity predictability of the model at the daily scale. The third objective of this thesis was to determine if optical remote sensing data capturing vegetation changes at seasonal scale (structure and hydric properties) can replace weather and/or landscape variables in a

model based on drought indices and random effects, or at least if optical remote sensing data improves the predictability of the model and provides new insight on the factors of fire activity.

The following chapters expose the answers and solutions found to address the questions and objectives mentioned above. In Chapter 1 the reader will find the background and the state of the art of fire activity prediction. This chapter presents the factors that are known to drive wildfires. Then it explains why it is needed to understand and predict fire events and the methods that serve these objectives around the world. It exposes limitations of fire activity prediction followed by strategies that are developed worldwide to overcome them. Finally, the reader will be introduced to the regional context of Mediterranean France, followed by the probabilistic framework and the use of spatio-temporal random effects in this framework to account for non-explained variability of fire activity.

Chapter 2 shows the study of landscape factors (human activities, land use and land cover) as drivers of the spatial distribution of fire activities, and an analysis of the interannual evolution of the fire regime. Chapter 3 is centered on comparing the predictors based on fire weather, for which a new methodology was proposed to improve the comparison of the performance of the different predictors. In chapter 4, the main subject is the analysis of insights from optical remote sensing data, the hypothesis being that remote sensing can explain to some extent the fuel moisture content and overall fuel availability, thus presenting the potential to obtain more precise spatial or seasonal information in predictions.

1. Background – State of the art

This chapter introduces the driving factors of wildfires and the ecological processes derived from wildfires. It continues by exposing the need of understanding and predicting fire events, as well as the methods that are used around the world to achieve this objective. Then, it brings up limitations of fire activity predictions followed by the strategies that are proposed worldwide to overcome them. After the global context, the reader is introduced to the regional context of Mediterranean France, followed by the probabilistic framework and the spatiotemporal effects that account for non-explained variability of fire activity.

1.1. Overview of main factors of forest fires: weather, fuel, landscape, and human activities.

The factors that control wildfires vary depending on the spatial and temporal scales. Moritz et al. (2005) illustrate these three scales of space and time (Fig. 1). The smallest spatiotemporal scale is the flame. From a physics point of view a flame is the combination of three elements: (i) an oxidizer, (ii) fuel and (iii) heat (or other ignition sources). Particularly, for a forest fire those elements correspond to oxygen for the oxidizer, plant biomass for the fuel, and the source of heat are natural or human ignitions.

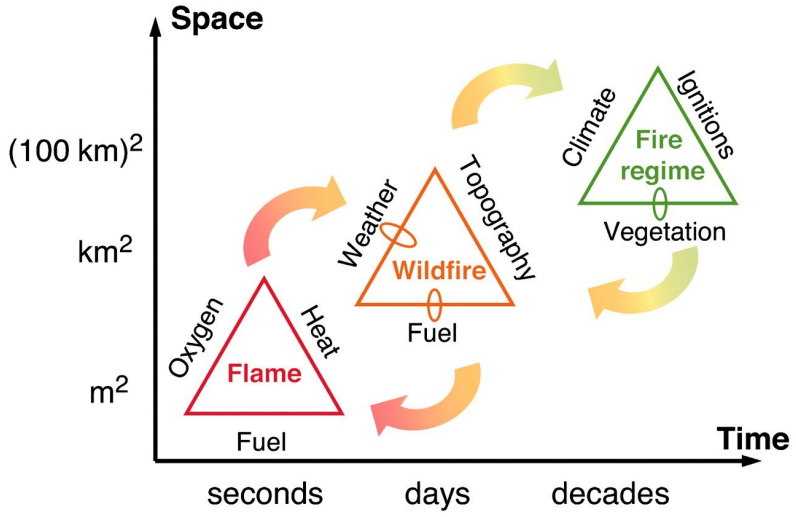


Figure 1: From Moritz et al., (2005). Triangle of fire on three spatio-temporal scales and the main fire factors on each.

At a larger scale, a behavior of a wildfire is affected by the fuel available at a given place and time, the topography, and the daily atmospheric conditions (or weather). The daily atmospheric conditions can facilitate fuel availability (dryness, drought, heat) and fire

spread (wind). The topography affects the rate of fire spread as flames will move faster when going uphill. Certain characteristics of the terrain may change the shape of the fire front. The aspect of the terrain (azimuth orientation of the hill) influences the local microclimate, e.g., the south-exposed sides are warmer and drier. At the wildfire scale, the fire itself interacts with the fuel and the atmosphere changing these conditions in a way that can intensity the fire behavior (i.e., the energy released by the fire per time unit), as well as the fire severity (fuel consumption and fire damage). The combination of the three factors (atmospheric conditions, topography, and fuel) controls the viability of the wildfire event, its intensity, and its final burnt area.

Four types of wildfires can be distinguished depending on the fuel that burns and the intensity (Bennett et al., 2010). (i) Ground fires burn the fuels, or organic matter, that constitute the soil such as peat and litter. (ii) Surface fires are those that burn the lower layers of vegetation, namely herbaceous and shrub biomass. (iii) Canopy fires occur when a surface fire is intense enough, or there is enough vertical fuel continuity, to transmit the flames to the tree canopy. (iv) Finally, selfsustained canopy fires are the most intense variation of the canopy fires. In this case flames are virulent enough to burn the tree canopy without the energy input of a surface fire.

Similarly, there are three types of wildfires depending on the main driver of fire expansion (Lecina-Diaz et al., 2014). (i) Wind wildfires are driven by the local and instantaneous wind conditions, either dominant regional winds or anabatic and katabatic currents. (ii) Topographical wildfires follow the shapes of the terrain. Fire advance faster on uphill progressions, also certain topographical characteristics may concentrate the intensity of the flames or divide the fire front in multiple fronts. (iii) Finally, convective wildfires are usually the largest fire events. Because of their size and high release of energy those fires can change the atmospheric conditions inside the plume. They become almost detached form the daily weather conditions.

At the largest scale, the fire regime (Burley, 2004) summarizes the main characteristics of fires on a region in terms of frequency, size, intensity, and seasonality. The fire regime also describes the type of fires that occur depending on the fuel and intensity or depending on the main driver of fire expansion. The fire regime depends on the climate of the region, the landscape factors, fragmentation of the natural ecosystems and the rate and sources of ignitions. There are interactions between the fire regime and the vegetation, meaning that the fire regime can modify its own driver. Among these scales,

the present study places itself within the wildfire (occurrences and burnt area) and fire regime scales. The scale of the behavior of individual fires, their rate of spread, intensity and severity is not included in the scope of this work.

The weather and climatic data for this study is obtained through a reanalysis gridded data base named SAFRAN (Vidal et al., 2010). The spatial resolution of this weather data is of 8km. At the scale of individual fires, the fire behavior depends on fuel load and structure, which can be measured in the field or remotely with airborne LiDAR(Martin-Ducup et al., 2025). At the landscape scale, these local properties of structure, shape and uses are summarized in simple fuel types, describing vegetation present on the land. The landscape factors are extracted through Land Use and Land Cover (LULC) data, from Corine Land Cover (CLC; https://land.copernicus.eu/en/products/corine-land-cover). In its finest classification, it contains the information about the occupation of the land and the human activities that take place, including types of vegetation, agriculture, and artificial surfaces. In addition, other sources from the French National Geographic Institute (IGN) and the National Institute of Statistics and Economical Studies (Insee) complete this information. Those include digital elevation models (DEM), the Wildland-Urban Interface (WUI) and data of infrastructures and population.

Dead Fuel Moisture Content (DFMC) of finer fuels affects the rate of spread, which spreads more rapidly when fine fuels are drier. DFMC of thicker fuels affects the fuel availability to fire, so that fuel consumption, and in turns, fire intensity and severity increase when thick fuels are dry. DFMC depends on the weather conditions, the fuel structure, and the species composition (Wotton & Beverly, 2007). DFMC of finest fuels (<6mm) will change faster than for thicker fuels, because the passive equilibrium moisture content with the atmosphere will take longer as fuels are thicker. The dead moisture contents of the different elements are accounted for in different operational fire indices further described in detail in section 1.6 (FCFDG, 1992; Jolly et al., 2024).

Fuel moisture content of live fuels (LFMC) is also important especially in shrublands or understory in dry conditions (Pimont et al., 2019). LFMC has with active systems of moisture regulation so that reactions to the changes on atmospheric conditions differ from DFMC. LFMCs will vary with the species composition because of the different wood structures and drought-resistant mechanisms between plant species and can be estimated with simple models (Jolly et al., 2024).

Ignitions sources are maybe the most difficult element to study and to determine after a wildfire event. Nonetheless, a large share of known fire ignitions if the French Mediterranean region, and in Europe, have human origins (de Rigo et al., 2018; Ganteaume et al., 2013). Ignitions are mostly brought by accident or negligence, while arson and lightning each represent a small part of the total fire ignitions. LULC can provide valuable information about potential human ignition sources. For example, the agricultural lands, due to machinery operations and negligence when burning vegetable residues without precautions. Roads are another example because they concentrate the occurrence of accidents involving flammable substances and fire ignitions. Furthermore, roads give accessibility for recreational activities in wildland environments.

1.2. Fire as a natural process in Mediterranean and beyond.

A natural disturbance (Erik A. Beever et al., 2019) is a process that, without human intervention, causes a sudden change in the spatial pattern of the ecosystem. It can be biological or non-biological and implies large mortality rates of living organisms. Natural disturbances are part of the ecosystem dynamics, playing a major role on the characteristic structure of given populations (G. M. Jones & Tingley, 2022; M. W. Jones et al., 2022; Kelly et al., 2020; Pausas & Keeley, 2019). The frequency, the intensity and the type of disturbances are key characteristics of an ecosystem. Wildfires are a common natural disturbance for many forests, shrublands and grasslands all over the Earth. In the boreal forests of North America and Siberia very large wildfires (>100.000ha) take place every season, as in the African savannas or Australia. Mediterranean forests, including the Mediterranean basin and other regions of the Earth (South Africa, Chile, California) are subjected to this disturbance.

Many Mediterranean plant species have functional traits that allow them to tolerate fire (Ne'eman et al., 2012). There are different strategies to deal with fire that can be classified in three main groups: resistance, resilience, and avoidance. In addition, the pyrophytes are those species that have traits to be more favorable and available for fire, as the presence of flammable resins or the accumulation of dry matter. Resistance strategies are characterized by the protection of the vital tissues of the tree (Fernandes et al., 2008). An example of resistance is the thick bark of very low combustibility that provides high thermic insulation in trees as pines, sequoias, or the cork oaks. The resilience is the tolerance of the disturbance and its consequences. Resilient species can

regenerate the population after the episode of mortality (González-De Vega et al., 2016). Examples of resilience are seen in *Quercus ilex* and *Q. coccifera* which resprout from its roots after a fire; or in the serotinous pinecones of *Pinus halepensis*. Serotinous pinecones stay closed, hanging on the branches, and open only under extreme high temperatures and low humidity. This strategy ensures the release of viable pine seeds just after a fire, when the seeds will not encounter any competition (Romero & Ganteaume, 2020). Species as *Cistus sp.* are specialists on colonizing bare soils after fires, using the nitrogen rich ashes, and resisting the drought and heat. Moreover, the blooming of their seeds is triggered by fire. The species that manifest avoidance strategies place themselves away from fire, either spatially (isolation, humid areas), seasonally (annual plans adapting the phenotypical stages to grow outside the fire season), or individually (species with structures that limit the fire occurrence and dispersion).

Thanks to the mentioned strategies there is a continuity of vegetation cover after a wildfire. This ensures the life cycle of a forest, inducing the natural succession of the forest stands decimated by a wildfire. Nevertheless, changes in the fire frequency, intensity and size creates instability and entrains degradation of the ecosystem. For example, if fires become too frequent, the production of seeds and other resources needed for regeneration is not completed between disturbances. As a result, the ecosystem enters a process of regression. This will lead to a transition from forests to shrublands or chaparral, and eventually to bare soils or grasslands. The ecological regression is also brought by the seed dispersal limit, when fire perimeters are too large to permit recolonization from unburnt areas (Acácio et al., 2009).

1.3. Fire excluded from the life cycle.

The management of natural resources has a very long history, dating from the beginning of the agriculture almost 10.000 years ago (Bulliet, Richard W. et al., 2008). The management has progressively taken over the natural dynamics on many ecosystems inducing changes on the forest structure and composition (Bradshaw, 2004). In many cases forest lands where depleted during wood-dependent episodes as railroad constructions. Before the intensification of agriculture most of the European surface was used for agricultural purposes. Forests used to be delegated only to the worst quality and inaccessible sites. The industrialization and intensification of crops have later led to a large abandonment of agricultural lands. More recently, a lot of Mediterranean forests are

in a situation of abandonment, including planted forests. In addition, recent agricultural abandonment means an expansion of non-managed forested lands (Bou Dagher Kharrat et al., 2022; Valladares et al., 2014), which creates a dense accumulation of fuels with continuity over the landscape.

These interactions between human activities and natural ecosystems, or management, alters the natural fire regimes. In Europe most fires have a human-related cause, either by accident, negligence, or arson; natural sources of ignition represent a small part of the total fire activity (Ganteaume et al., 2013). Moreover, suppression strategies have evolved and more efficient than ever, but the suppression of all fires avoids the consumption of fuels which then accumulate (Hurteau et al., 2024; Piñol et al., 2005). On big terms, the European fire regimes tend to a reduced number of fires because of the control and prevention strategies. The big fire events are more frequent and more likely to overcome the capacities of firefighting (Barbero et al., 2015). Furthermore, the Mediterranean climatic region is expanding, reaching ever northern latitudes and higher altitudes (Alessandri et al., 2014). As a result, new forest lands not adapted neither managed for fire prevention become available fuel.

1.4. The French Mediterranean region and its fire regime

1.4.1. The study area

The majority of summer fire activity in Metropolitan France takes place in the French Mediterranean region. The Mediterranean region is a tight space that presents societal and economical challenges owing to the dense human population, the industrial activities, and the touristic interests. In this area, a large fire database held by the French public authorities (Promethée, see 1.4.2) is available to support this research. The region is composed of 15 French *départements* (NUTS3 divisions) situated on the south-east of the country (Fig. 2). This area includes the island of Corsica and all the coast of the continental French land from its Spanish border to the Italian border. Furthermore, the study area can be divided with an ecological criterion, giving 28 Sylvo-Eco-Regions (SERs) defined by the IGN (Fig. 3); this last subdivision will be used in the study.

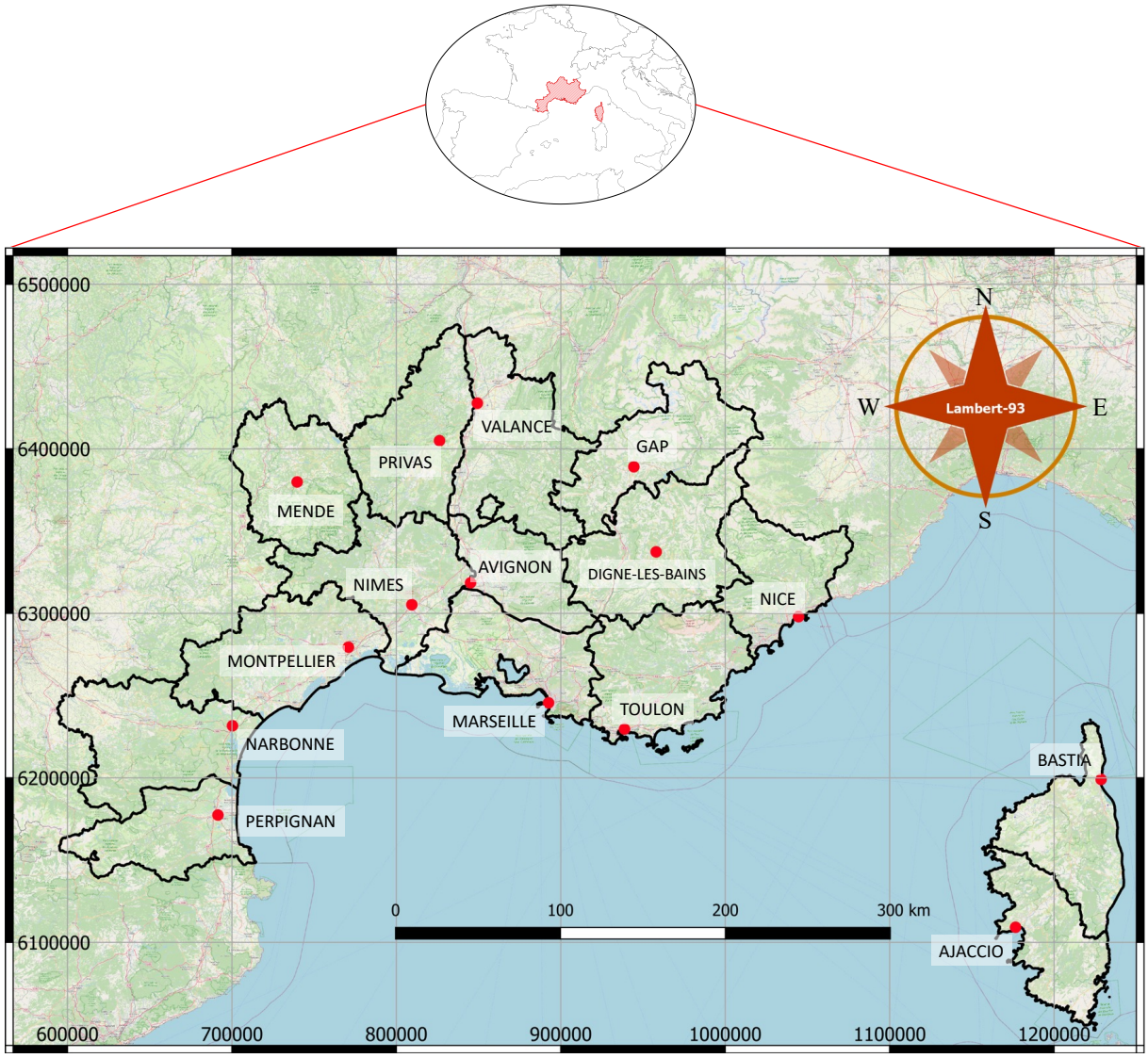


Figure 2: Map of the study region, where black lines delimitate the NUTS3 units and red points indicate main cities to give geographical references. The background is the Open Street Map (https://www.openstreetmap.org/).

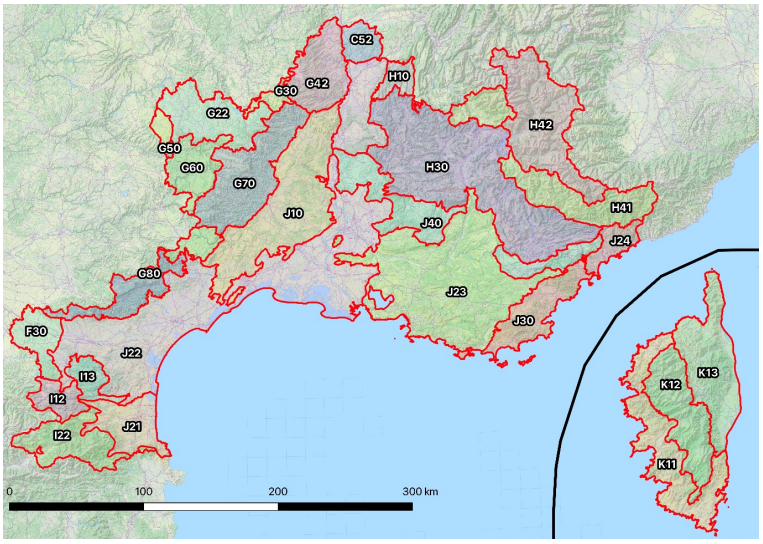


Figure 3: Map of the 28 Sylvo-Eco-Regions (SER) of the study area with their key code.

As the region has been shaped by human uses and human population dynamics (Rick et al., 2020). The intensification of agriculture and minor dependency on wood have led to an expansion of forests and shrublands. The region is economically strategic for the national and international trade with important harbors and industries. In addition, it is a major destination for tourists, attracted by the major cities and the natural scenarios. Main roads, highways and railways run across the area; furthermore, a dense secondary (and smaller) road network covers the land. A dedicated network forest roads with restricted access is in place for the surveillance, management, and suppression of wildfires, to an extent that even remote areas are generally accessible, despites of extensive wildlands and topography.

The main land-use and land-covers in the study area are vegetated areas, either agriculture or wildland. A clear geographical separation is observed between the wildland areas and the populated (or urban) areas. The distribution of agricultural lands is closer to the urban areas, which follow the valley of the Rhone River and the lower and flatter lands. Urban areas are also concentrated along the coastline (except for the Camargue). The study area is one of the Metropolitan French territories with the most surface covered by natural spaces, from which many are forest managed in a non-intensive manner.

In the last decades the French Mediterranean landscape has been dynamic, although the absolute changes in land use and land cover are not big. The changes observed within the pixels of the Défense de la Forêt Contre les Incendies (DFCI) grid (400ha/pixel) from 1993 to 2018 do not overcome an absolute value of 45ha; those largest magnitudes are only seen in single SERs. Shrubland area (Fig. 5A) and agricultural area (Fig. 5D) have suffered the most extended loses; while urban (Fig. 5B), mixed forest (Fig. 5E) and broadleaved forest (Fig. 5C) areas have the largest extent of positive change.

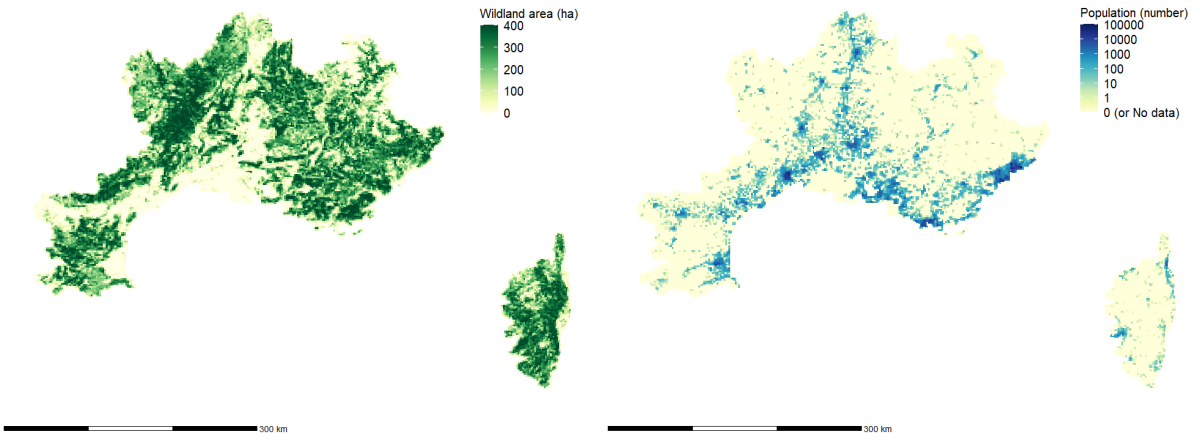


Figure 4: Spatial distribution of wildland area and population both in units per DFCI pixel.

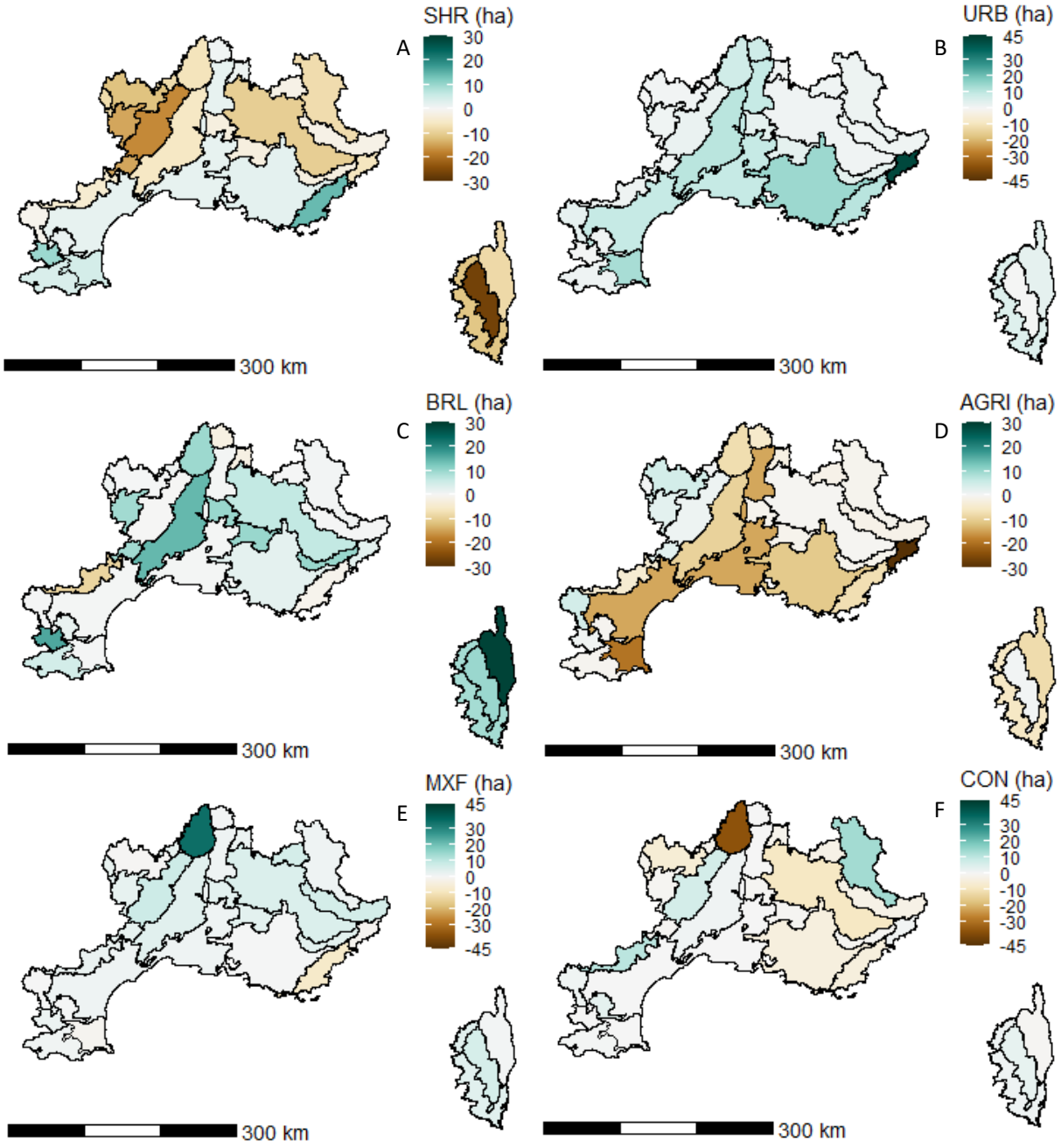


Figure 5: Distribution of total absolute change in mean area per pixel for the period 1993-2018 represented in the Sylvo-Eco-Regions. Where SHR = shrubland; URB = urban areas; BRL = broadleaved forest area; AGRI = agricultural area; MXF = mixed forest area; CON = coniferous forest area.

The study area is mainly characterized by mild cold winters and hot summers, while precipitations are concentrated on the spring and autumn with very dry summers (Fig. 6). Nevertheless, the study area is large so, some regions differ from this standard. For example, the Alps (Fig. 6E) exhibit colder mean temperatures and more stable precipitations through the year. Furthermore, the valley of the Rhone experiences a strong dominant wind from the north-east (with the vernacular name of Mistral). The

recent decade is affected by severe droughts, accumulating several years of hydric deficit and summers hotter than average. Because of the climatic heterogeneity of the study area, Curt et al. (2014) established the pyroclimatic regions by crossing the climate and fire activities. These pyroclimatic regions are zones with homogeneous fire regime, climate, and their temporal trends. They found three pyroclimatic regions which classify the Promethée area in "*high fireprone maritime mountains*", "*moderate fireprone Mediterranean plains and foothills*" and "*low fire-prone temperate mountains*". Those three pyroclimatic regions can be further sub-divided in 10 units for geographic differentiation.

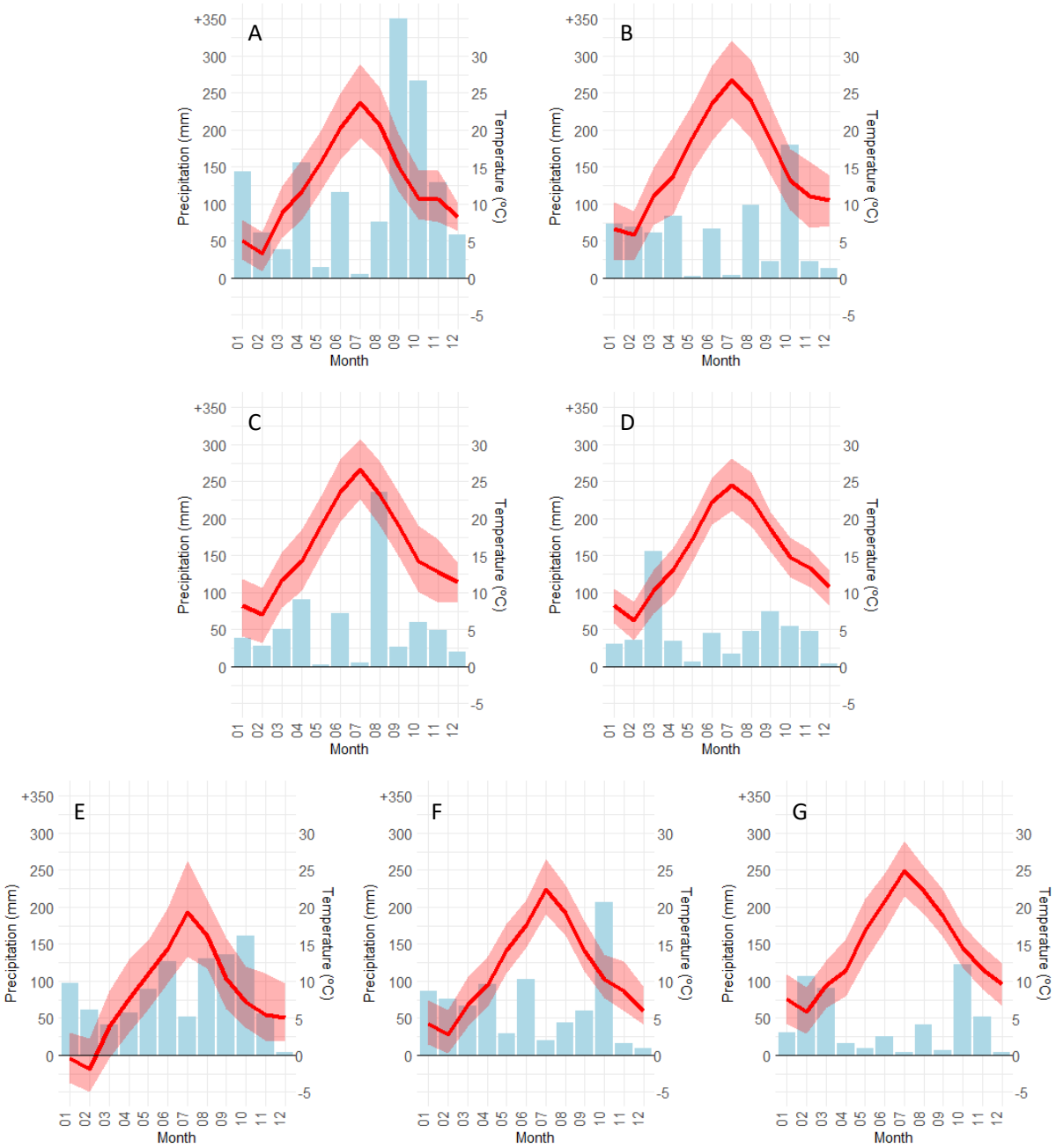


Figure 6: Climographs of seven places scattered on the study area. Weather data sourced from SAFRAN database. The climograph correspond to A – Ardèche's forests, B – City of Avignon, C – City of Montpellier, D – City of Portbou, E – Alps mountains, F- Mountains in Préaples d'Azur and G – Northern Corsica.

1.4.2. Fire records in Mediterranean France

The Promethee database (http://www.promethee.com/) registers all fire activity since 1973. Its aim is to homogenize the records of fire activity between the different French institutions concerned by this phenomenon. The database consists of a 2x2 km-pixel grid (referred as the DFCI grid). Each fire is logged to the corresponding pixel where the ignition was reported along with the date, final burnt area, and ignition source (if known). There have been changes on the methods and the quality of data recordings over time, which will condition the following research. Recently, this dataset has been closed, and substituted by the BDIFF, managed by the IGN, but the present work has been carried out with data formerly extracted from Promethee, with complementary information dataset of burn perimeter obtained with courtesy of the French National Forest Service (ONF).

Despite of the availability of a dataset that intends to log all fire events, regardless of the size or cause, we consider that the registers of fires smaller than 1ha in the Promethee database are tight to too much uncertainty. Some of these minor incidents might not be well logged (imprecise size or location) or omitted in the database in some NUTS3. In addition, in south-eastern France there is a strong surveillance and early attack policy. Establishing a minimal surface (1ha) is a warranty that these constrains are avoided. In this way, only the occurrence of fires that could bring significant risks are studied. Those fires that started under conditions that were not suitable for combustion and fire expansion, including fires that were extinguished with the intervention of first fire responders, are not introduced in the study.

Using a pixel grid has limitations when we study fire occurrences and fire sizes: (i) the lack of knowledge about the precise location of the fire ignition point. (ii) The bad characterization of large fires, as they are likely to exceed the pixel borders and cross different LULC types. To overcome these limitations, we used a burn perimeter database from ONF, that consists of a large compilation of the burnt perimeters in the form of a vectorial shapefile (Fig. 7), from now on referred as fire perimeter data. A small subset has been already used by (Ganteaume & Barbero, 2019). The work here presented uses for the first time the perimeters data with research purposes on the Promethée area. In the fire perimeter data, not all burnt areas are measured with the same method. Between

the possible methods of measurement, there are satellite imagery, field measurement, and contours with known or approximated location for some minor events.

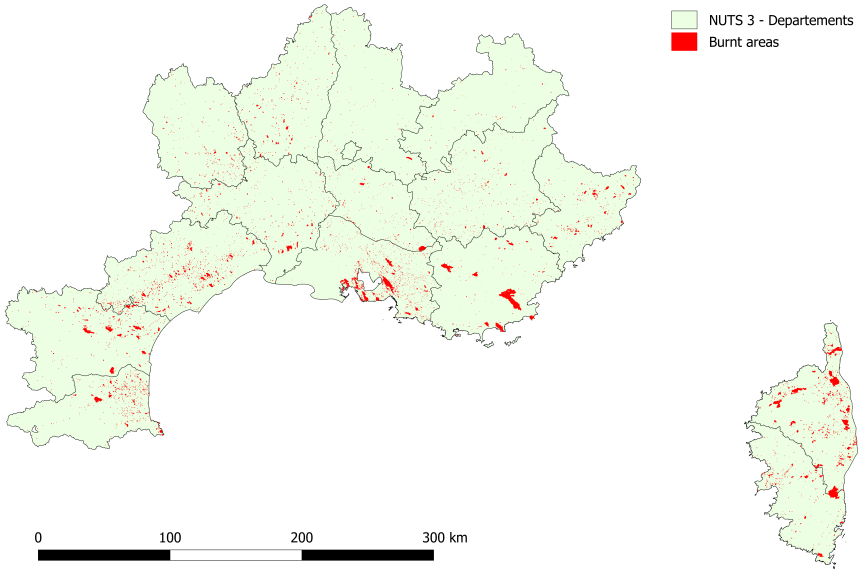


Figure 7: Fire perimeters data presented over the study area.

The seasonality of fire activities is concentrated during the months of summer and early fall in a consistent manner throughout the years (Fig. 8A). The winter months suffer from fewer burnt area (Fig. 8A). The years 2003 and 2017 are highly remarkable due to the record heatwaves, which exceptional fire activity (Fig. 8B). More recently, in 2021, Gonfaron fire showed an abnormal virulent behavior, reaching almost 7.000ha, from which 5.000ha on the first day (https://www.var.gouv.fr). The conditions during Gonfaron fire created a very fast and intense fire of unpredictable behavior and many fire jumps. The events that took place on 2003 (Fig. 8B), with over 800km^2 of burnt area, initiated a change on the policies facing wildfires. An earlier change of fire policy was done in the decade of the 1990s and it has been shown that it induced changes in the fire regime (Ruffault & Mouillot, 2015). Also, societal chances can induce changes on the fire regime and the main drivers of fire activities (Pausas & Fernández-Muñoz, 2012).

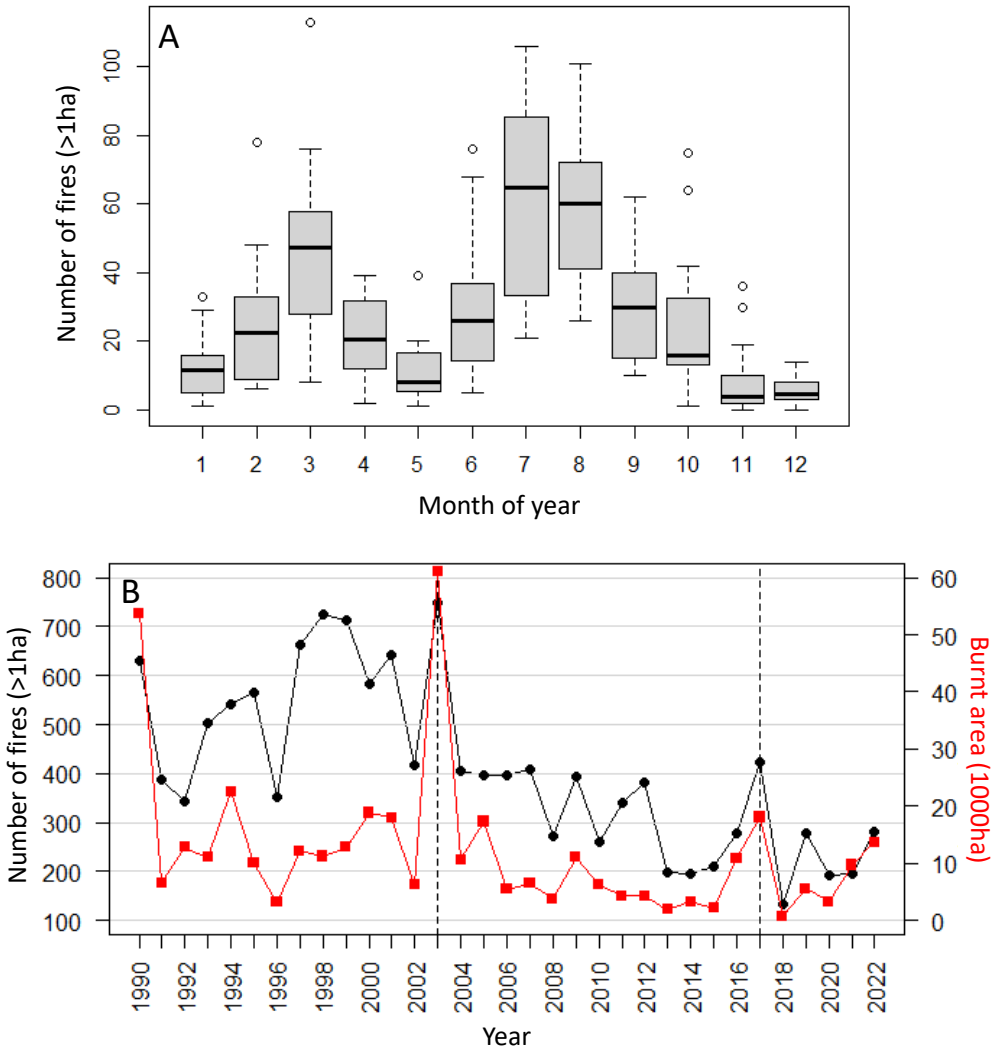


Figure 8: Temporal distribution of the mean number of fires per month for the period 1993-2022 (B). And temporal distribution of the total number of fires and total burnt area per year (B). In the plot B, vertical dotted lines indicate the years 2003 and 2017, both years where remarkably dry and are considered as extreme within their decade.

1.5. Perceiving fire as a threat

The definition for threat, danger and risk are sometimes controversial on the literature. Here, we rely on Baijnath-Rodino et al. (2020) for simplicity. A threat is any phenomena that could potentially bring loses for a given value; this can be a material damage, human life, ecosystem services or economical activities. Danger is defined by the probability of a perturbation to appear/occur; danger can therefore quantify the appearance or not-appearance of the threat. Risk is the product of the danger and the values at stake, given their vulnerability, in other works it accounts for the perturbation and its consequences. The concept of risk englobes a dimension that quantifies the potential loses in the societal and economical terms.

As said before, the Mediterranean regions are highly populated with large quantities of Wildland-Urban Interface (WUI), and the occurrence of large fires are frequent. As a consequence, there is thread of material damages to houses, businesses facilities and other infrastructure, in addition to the possible losses of life and ecosystem services. Many ecosystems face the possibility of not being able to regenerate after the perturbation, leading to the loss of biodiversity and ecological regression processes. The combination of all these social and environmental processes creates a general apprehension to the potential damages caused by fire. This perception of threat and risk needs to be addressed with correct decision making at many levels (forest and landscape management or crisis situations management). It requires an accurate prediction of the potential danger and risk. In this regard, many administrations and managing agencies have developed systems to understand the fire danger. Their objective is to predict the future occurrence of fires and how big and intense they may become.

1.6. Methods for fire danger prediction

Fire danger predictions are expected to give answers to several questions or problematics that arise from the need to manage and prevent the fire danger (Taylor et al., 2013). First, the analysis of past fire activities shows the specific factors that drove the occurrence and behavior of the observed fires. Next, the knowledge built during the exploration of past activities brings the possibility of predicting future fire activities. The predictions can be done at different temporal scales, e.g., daily to weekly forecasts, seasonal previsions, interannual future trends (Di Giuseppe et al., 2024). Finally, fire predictions should inform about the fire suppression difficulty, or in other words, the fire intensity, but as mentioned before, this aspect of fire danger is beyond the scope of the present work. The knowledge of the fire factors and their measurement is useful to quantify the potential for fire activity in a given spatio-temporal unit (see BOX 1), establishing the fire danger levels. In general terms we find three methodologies to express fire danger: indices, fire occurrence prediction and fire behavior prediction. Those can be combined to obtain complex systems of fire danger prediction.

BOX 1

A spatio-temporal unit is the numerical abstraction of the parameters inside a predefined area during a defined time step (Fig. B1.1). Usually, the spatial units are regular and continuous divisions of the territory, taking the shape of a grid in which each unit is called a pixel. The pixels of a grid are squares that have a known side dimension, expressed in units of distance on the ground. In other cases, the spatial units are based on administrative borders, such as communes, counties, etc. Those can also be based on territorial units based on ecological criteria, for example, the French SylvoEcoRegions. The time steps are the regular subdivisions of the temporal dimension. They correspond to the standard time measurements (Hours, days, weeks, etc.).

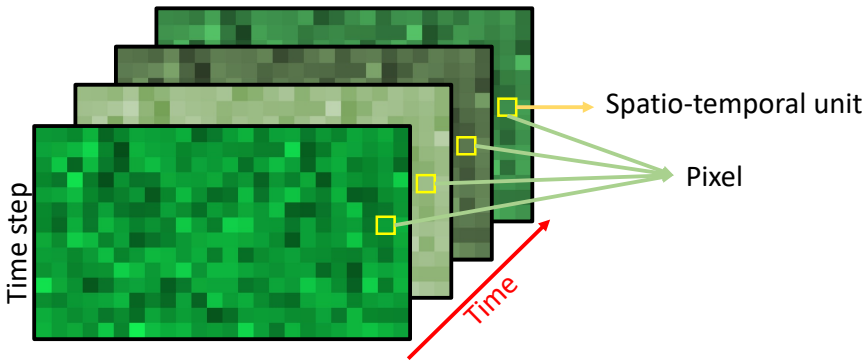


Fig. B1.1: Graphical representation of fours observations of the same space at four defined time steps. Yellow squares locate one same pixel observed four times. Each observation gives place to a spatio-temporal unit.

An index is an ordered numerical value that evaluates a given phenomenon, magnitude, or dimension. It is provided by a specific calculation method nourished with several indicators or predictors. When the index is conceived to evaluate fire danger it is called a fire index. In this case, the index is usually informed with measurable characteristics of the ecosystem and the atmosphere which are related to the occurrence and intensity of wildfires. Most fire indices are fire-weather based, quantifying the meteorological conditions that benefit the occurrence and intensity of fires by rising the fuel availability and helping the advance of flames.

Fire Occurrence Prediction (FOP), defined in Woolford et al. (2011) and Woolford et al. (2021), is a methodology that uses modelling techniques to train a model with historical fire observations and fire factors data. The resulting model, once validated, can be used to emit short-term forecasts and/or long-term predictions of fire activity. The FOP predictor variables are measured fire factors directly introduced in the training or pre-analyzed

through a chosen index compatible with the regional context. This method allows for the combination of fire factors from different sources (weather, landscape, field measurements). Fire behavior prediction (FBP), defined by the Forestry Canada Fire Danger Group (1992), is a method that informs about the potential intensity, type, extent and/or burnt area of the fire. The most powerful FBPs will be able to predict the exact perimeter of a given fire through physics based methods, for example FIRETEC (Pimont et al., 2016). Other simpler FBP methods only aim to predict basic characteristics, as the burnt area or the potential flame intensity.

Fire danger rating systems (FDRS) are more complex methods rely on the combination of several components, as fire indices and even integrated FOPs and FBPs. Each component describes a fire factor based on its specific drivers. These complex methodologies offer a more detailed insight into the fire danger situation on the studied spatio-temporal unit. An example is the Canadian Forest Fire Danger Rating System (CFFDRS), whose development was started in 1968 and is still progressing (Fig 8).

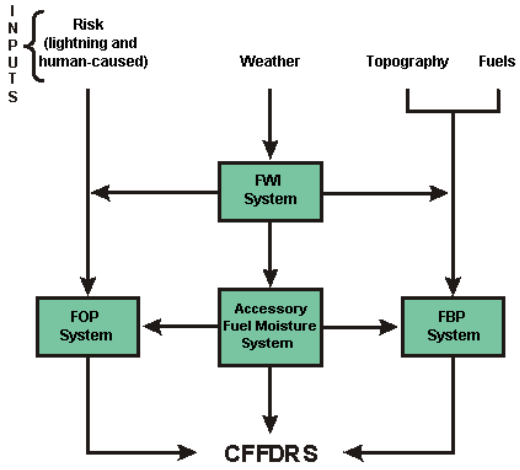


Figure 9: Scheme of the Canadian Forest Fire Danger Rating System (CFFDRS), sourced from https://cwfis.cfs.nrcan.gc.ca/. Where FWI – Fire weather index, FOP – Fire occurrence prediction, and FBP – Fire behavior prediction. The FOP system and the Accessory Fuel Moisture System are under consideration and development phases.

Numerous strategies combine the measurements of fire factors available through a methodological framework that best fits a specific region and period. All strategies can be classified with different criterion: (i) A classification by the geographical area of origin is necessary when the index, FOP or FDRS has been calibrated for specific climatic and landscape conditions. (ii) A second classification regarding the factors that serve as inputs and the type of data that is used (field measurements, gridded databases, remote sensing, etc.). (iii) And a third classification regarding the type of output. Every method gives a defined output which illustrates fire danger through a specific magnitude (e.g., non-

dimensional value, probabilities, predictions of intensity, prediction of behavior). Correctly identifying the categories, within the three classifications, to which an index, FOP or FDRS belongs is the first step to decide the methods that might, or might not, be suitable for a chosen study region and period.

1.6.1. Frequently used fire danger prediction methods

Baijnath-Rodino et al. (2020) identified 24 fire indices and danger rating systems worldwide that are most frequently used for operational or scientific purposes. It is almost impossible to list and study all the fire indices that have ever been developed. Here we have identified the meaningful examples from an international context that will be studied in Chapter 3. A well-known FDRS is the Canadian Forest Fire Danger Rating System (CFFDRS) that includes a fire-weather index, a FOP system, and a fire behavior prediction system. The components of the CFFDRS are based on weather data, risk assessments, topography, and fuel measurements. Recently, an FDRS has been updated for Australia (Hollis et al., 2024), it was preceded by the McArthur Mk5 Fire danger Meter (McArthur, 1967). The Australian FDRS uses fire spread models for each type of wildland environment (forests, grasslands, shrublands…) to obtain a fire behavior index. In the United States we find the US National FDRS (US-NFDRS; (Jolly et al., 2024; Schlobohm & Brain, 2002), a national wide system that needs to adapt to a large range of biomes and so fire regimes (Walding et al., 2018). The US-NFDRS has four outputs: the ignition component (probability that a fire occurrence is important enough to need suppression), spread component (speed of the fire front movement), energy release component (estimation of the energy per area that is available to the flames) and the burning index (effort for fire extinction). A new modification is the addition of the Severe Fire Danger Index to forecast the extreme events (Jolly et al., 2024). Although the three systems described above use different computation methods, they have very similar inputs information requirements including weather, fuel status and landscape.

The Canadian Fire Weather Index System (CFWIS) is maybe one of the most known and used worldwide, even outside of its calibration ecosystem, the boreal forests of Canada (Fig. 9). It is based on meteorological inputs with which it calculates three moisture codes. The moisture codes describe the status of different fuel types of the forest. After they combine into two fire behavior indices, and a final index that englobes all, the “Canadian Fire-Weather Index” (CanFWI). The three moisture codes are: (i) Drought Code (DC) to

quantify the theoretical water deficit of the soil, and therefore the water stress of the vegetation. The DC is an indicator of the drought conditions at seasonal scale. (ii) the Duff Moisture code (DMC) to represent the calculated moisture content for the first organic matter layer on the soil. And (iii) the Fine Fuel Moisture Code (FFMC) that illustrates the moisture content of fine dead fuels. The FFMC indicates the fuels moisture conditions on a small temporal scale. The fire behavior indices are: (i) Initial Spread Index (ISI) that approaches the fire spread rate based on the FMC and the wind speed; and (ii) the Build Up Index (BUI) rating the quantity of available fuel based on DC and DMC.

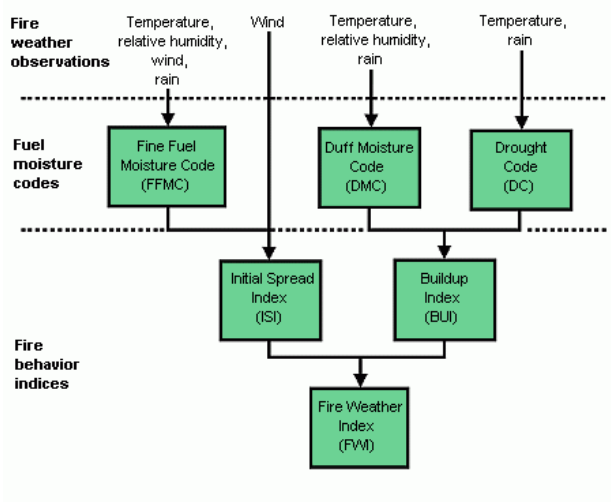


Figure 10: scheme of the Canadian Fire Weather Index System (CFWIS), sourced from https://cwfis.cfs.nrcan.gc.ca/.

1.6.2. Fire danger prediction through remote sensing

Remote sensing englobes all techniques for observation of the Earth surface through imagery in a large spectrum of electromagnetic radiation. The images can be taken from the ground, from manned or unmanned aircrafts, and from satellites. These observations are passive when they collect the reflectance of radiation; or active when a signal is sent to read its specific reflectance. For large spatial and temporal scales, the use of satellite remote sensing is the most adequate because of the regularity of the observations and the coverage of vast regions of the Earth, if not the full planet. Many satellites orbit the Earth and take constant passive and active measurements. The Moderate-Resolution Imaging Spectroradiometer (MODIS), for example, orbits the plated onboard of two satellites since more than 20 years ago. Also, Sentinel-2 satellites have been launched in 2015 and 2017. MODIS and Sentinel-2 are equipped with optic passive sensors that provide measurements in 36 and 13 spectral bands respectively. Remote sensing products are

used to predict the LFMC and other fuel characteristics as surface temperature, dry matter content or still fuel structure (Quan et al., 2021; Rao et al., 2020; Yebra et al., 2018).

The optic electromagnetic spectrum ranges from the Infrared to the Ultraviolets. Within this range, specific wavelengths interact differently with plant tissues, water, and soil (Wong, 2023). Direct measurements may already identify the presence of fuels, their water content, and even plant phenological stage or health state. The NIR (Narrow Infrared) spectral band, for example, is reflected by the chlorophyll so, it is an approximative measure of the vegetation greenness. Also, the SWIR (Short-wave Infrared) spectral band is negatively related to water content, being an indicative of dryness. Sensors from different satellites may not consider the same wave length thresholds to make the distinction between the spectral bands. The combination of spectral bands into remote sensing indices is a method to isolate the signal of a searched trait. Examples of those indices are the Normalized Difference Moisture Index (NDMI) and the Normalized Difference Vegetation Index (NDVI), further details are given in BOX 2.

BOX 2

The NDVI combines Red and NIR spectral bands, stablishing a difference between the leaf's greenness (a greener leaf absorbs more red light, as this spectrum is used for the photosynthesis) and the photosynthetic activity known through the NIR (higher photosynthetic activity reflects more NIR radiation). The NDVI is a quantification of the green vegetation, the interpretation of its values can be done following Table B2.1.

Table B2.1: Interpretation of NDVI values for the three most observed intervals. The missing intervals can be assumed as the intermediate states between the intervals defined here.

Interval	Meaning
(0 , 0.1]	Mostly base soil
(0.2 , 0.3]	Thin vegetation, bushes and meadows
(0.6 , 0.8]	High density forest cover

The NDMI combines NIR and SWIR, it defines the difference between the dry matter content of the leaves known through the NIR2 as it interacts with the leaf's internal structures but not with water, and the water content of the vegetation known through the interaction of SWIR with water and the spongy mesophyll structure of leaves. NDMI is an indicator of vegetation coverage and plant hydric stress, and its values can be interpreted following Table B2.2.

Table B2.2: Interpretation of NDMI values for five intervals.

Interval	Meaning
(0 , 0.2]	If medium vegetable coverage, high water stress. Or low coverage.
(0.2 , 0.4]	If high coverage, high water stress. Or medium coverage.
(0.4 , 0.6]	High coverage, no stress.
(0.6 , 0.8]	Very high coverage, no stress.
> 0.8	The vegetation coverage is a 100% or the soil is inundated.

1.6.3. Fire danger prediction and early alarm systems

An important differentiation must be done between the prediction or forecast systems here described and the early alarm systems. An index or a fire prediction model is a theoretical approach that gets validated by comparison to historical data. This validation ensures that the prediction or forecast that is emitted from the method is reliable. If validated, the prediction can be used to organize the surveillance and intervention

operational teams to respond more efficiently during the upcoming period. On the other hand, an early alarm system does not predict the potential fire activities. They continuously observe an operational area to detect present fire departures before they are consolidated. It permits to mobilize the operational teams to a very specific location. Early alarm methods are based on the surveillance for hotspots, fume column detection between others (Pettinari & Chuvieco, 2020). The combination of danger prediction and early alarm enables organize authorities to organize early warning of populations and operational services by broadcasting the alert.

1.7. Fire danger prediction in France

South-Eastern France concentrates a large share of the country's fire activities. The fire regime together with the socio-economic challenges in the region creates a need for anticipation and managing. On the second half of the last century, two South-Eastern France specific indices were developed. One by Orieux (1974) as a combination of the calculated drought level and the wind speed in a look up table. And the second by Carrega (1991) using equations estimated with empirical knowledge and hypothesis of fire factors. Later, Drouet & Sol (1993) established a method to evaluate "risk" (following the terminology of this thesis, they evaluated danger). It used a simple equation considering a modified HR that is based on soil temperature, the soil water reservoir, and the wind speed. It was defined as the "indice numerique" (digital index) because it was a first implementation of a methodology based on computational systems (Fig. 9). Its objective was to enhance the estimation of fire danger, integrating the relative humidity in the index.



Figure 11: Representation of the "indice numerique" for the operational forces. The information was transmitted in live from the meteorological agency. Image taken by J.C. Drouet in 1989 on early application of the method published in 1993.

On the present day, several agencies (Meteo-France, ONF, INRAE, SDIS) collaborate to establish a forecast and firefighting strategies. Meteo-France develops the fire danger assessment methodology used for managing and decision making at operational. This methodology is based on the CFWIS and uses its components DC, DMC and FFMC, and the CanFWI. In the original Canadian system, the daily maximum indices (typically at 4 PM solar time) are extrapolated from the weather observed at noon solar time. Météo-France calculates this index using the Canadian algorithm with weather variables at 12 UTC and calls this index IFM12. Météo-France also applies the same algorithm at every hour of the day and takes the maximum value as an index called IFMx. This approach tends to overestimate the IFM and its sub-indices with respect to the original CFWIS. Therefore, the approach of Meteo-France should not be used to calculate other metrics of the Canadian system (such as those of the Fire Behavior Prediction System), as it would result in overestimating the metrics of fire behavior. However, this overestimation does not seem to pose any further problem for rating the daily danger, as long as the thresholds are adjusted to this new scale.

Furthermore, the operational forecast system by Météo-France uses several custom double entry tables to cross the moisture codes between them or with wind speed. From each table they obtain a value ranging [1,5] or [1,6], which is either a final danger output or an input to the next table. At the end, two main discrete numerical metrics are used to

inform about the fire danger in terms of drought, fire spread danger or integrated danger. The first one is a combination of wind and FFMC referred to as "indice d'éclosion propagation" (IEP) and has the same objective as the Initial Spread Index of the original CanFWI. The second one, referred to as "Indice de Sensibilité de la Végétation Vivante" (NSV2) is a combination of DMC and DC. It is supposed to represent Live Fuel Moisture, even if the DMC and DC are used to estimate fuel consumption in the original approach and LFMC is not accounted for in the CFFDRS. Once a map of danger level per area is given for the day, an expert can modify certain results based on expertise and knowledge of the terrain.

In addition to the weather-based methodology, the "Hydric network" (*Réseau hydrique* in French) is a complementary field measure of shrub LFMC in approximately 30 plots scattered on different biomes, some of those plots have a more in-depth control. The National Forest Office (ONF) carries a weekly sampling of moisture content per plot, focusing on the main shrub species on the plot. This data is used to correct the predictions emitted by Meteo-France with the NSV2. Unfortunately, the representation of the danger in each region is heterogeneous, with varying color codes, number of classes, and territorial divisions.

1.8. Limitations for fire weather index approaches

Fire danger prediction is affected by several limitations that have been described and sometimes addressed in the existing literature. The current French method of fire danger prediction in Mediterranean France is, as said before, based on a fire-weather index, namely the CanFWI from the CFWIS. It is important to remind that an index is an abstraction of the danger level, which does not give a simple and direct interpretation of the probability of occurrence and potential extent of a fire event. Moreover, the methods used nowadays do not consider the uncertainty associated with fire activity in the prediction. The uncertainty is an important element for decision making in fire management and the comparison of future scenarios.

Other limitations can be brought by the use of weather-only fire danger rating systems. This last limiting factor may induce a prediction bias of the effect of fire-weather. Although weather has been shown to be a major driver of fire activity (Abatzoglou & Kolden, 2013; Barbero et al., 2015; Turco et al., 2017), other factors that drive fire activity at regional scale are important: (i) The firefighting efforts and capabilities from the region (Pimont

et al., 2021); (ii) the patterns of fire ignitions (Balch et al., 2017); and (iii) the type of management or productivity of the forest (Pausas & Ribeiro, 2013). In Mediterranean France Drouet & Sol (1993) already stated the limitations of using solely meteorological data in their early paper about a fire danger index. Their research pointed that, even if the LFMC is an important fire factor, its quantification is only possible by directly measuring the biomass. Several experiments were carried out by that time, including a promising approach with remote sensing using surface temperature observations.

Several methods can be found in the literature that aim to overcome some of the limitations here presented. The most straightforward option is to recalibrate the classification of a fire weather index by adapting the danger categories to the fire observations on the study area (Alves et al., 2021; DaCamara et al., 2014). A second option is to substitute the components of a FDRS by field observations or modifying them components by site-specific proxies (Garcia-Prats et al., 2015; Yang et al., 2015). These methods avoid using proxies of FMC based on weather observations. Third, remote sensing approaches can provide direct observations of fuel status (moisture content, greenness) and quantities. The remote sensing observations can be introduced into an FDRS or used in probabilistic models. Finally, some studies search for new indicators based on fire-driving hypothesis. For example Rao et al. (2022) developed the concept of Plant-Water Sensitivity.

Still, a major constraint in fire activity prediction comes from the stochastic and rare nature of fire events. This means that even if all the conditions are met for the highest fire danger, an ignition might not happen. The number of fires respect to no-fire days in a small spatial unit is very close to 0. Even if all the fire factors are correctly modelled or measured, there is a random dimension that is not captured by the methods presented above. Probabilistic frameworks are an efficient method to overcome uncertainty associated to stochasticity and rareness of the event (Brillinger et al., 2003; Coughlan et al., 2021; Pimont et al., 2021; Preisler et al., 2004; Preisler & Westerling, 2007; Serra, Saez, Juan, et al., 2014; Serra, Saez, Mateu, et al., 2014; Turner, 2009; Vilar Del Hoyo et al., 2011). A probabilistic framework can be informed with all types of fire factors in a single mathematical expression or equation. It can combine the weather, landscape, remote sensing information, seasonality, human activities, between others. The output of such frameworks is either a probability or an expected number of events per spatial and temporal unit. This can be translated to the probability of fire occurrence or number of

expected fires for a chosen fire size. Using a probabilistic framework, Woolford et al., (2021) developed a FOP system that is integrated in the CFFDRS.

1.9. An example of the operational application of a FOP from Canada

The Woolford et al. (2021) publication is worth of in-depth analysis because of the similarity both in methods and objectives with the research developed in this manuscript. Their objective is to provide a daily FOP system that can be used by the operational forces in Ontario. The fire prediction methodology developed by Woolford et al. (2021) is based on a spatial and temporally explicit general additive model to predict the expected number of fires. They used a pixel grid with 10km cells, each one containing weather and land use and land cover information in addition to the fire records.

They propose a graphical output with a continuous scale that can be also classified in five discrete danger categories. This system is presented at a scale that is useful to the decision makers and operational forces. The units of the output are expressed in number of fires per cell or number of cells needed to observe a fire (if less than 1 fire/dell is expected). This is thought to have an easy interpretation on the field, as it gives a notion of the resources that must be allocated in each territorial unit. Figure 10 shows the graphical output that is presented to the operational forces with the FOP system developed in the study.

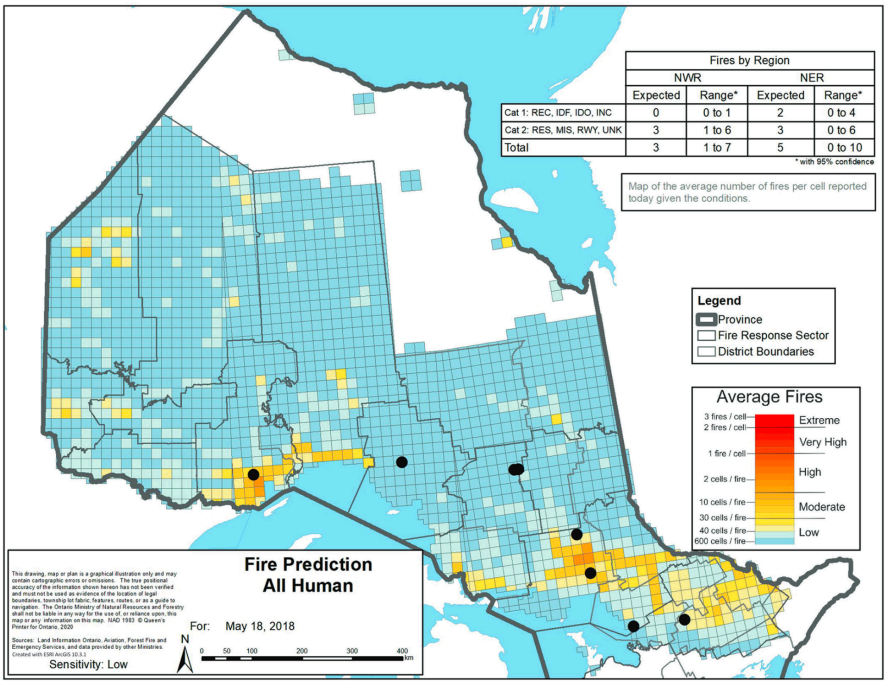


Figure 12: From Woolford et al. 2021. Graphical output of a FOP system in the province of Ontario (Canada). The color scale indicates the FOP prediction and the black dots the fire observations for May 18th, 2018.

1.10. A probabilistic framework for fire activities in Mediterranean France

This thesis uses a Bayesian probabilistic framework to develop a fire activity prediction method in Mediterranean France. This approach was first developed in the frame of Helene Fargeon's PhD (Fargeon 2019) and led to the development of Firelihood (Pimont et al. 2021). The aim of this pioneering work for Mediterranean France was not to predict fire danger, but rather to anticipate the effect of climate change by projecting future fire activity with global warming (Pimont et al. 2022&2023). *Firelihood* models the occurrence and size of fires in two separate components that are mathematically combined afterwards. These two components do not need to be informed with the same predictor variables (although they can). As a result, *Firelihood* can identify specific drivers for the fire occurrence or the fire size. The first version of the model was based on a pixel grid with 8km resolution, corresponding to the SAFRAN climatic database (Vidal et al., 2010).

We define the occurrence as the punctual process of a wildfire ignition that reaches at least 1ha of area. A choice is done to assume that fires that were stopped before reaching 1ha did not occur under conditions that bring significant risks. The challenge lies on fires that reach more than 1ha which take place under fire-prone conditions. As said before, we consider this minimal area because the following reasons: (i) the Promethée data is more exhaustive with fires larger than 1ha; and (ii) there is a strong policy of surveillance and early attack in southern France, firefighting services have the objective of taking down all fires before they reach 1ha.

The occurrence is measured as the expected number of fires in a spatiotemporal unit. This quantity can be aggregated in both spatial and temporal dimensions. The size of each fire is defined by an exceedance probability of a preselected burnt area threshold. Each punctual occurrence process has a score associated to it that expresses the probability of exceedance. It is important not to confuse the probability of exceedance with the intensity of the wildfire. The intensity is defined as the energy released by the fire by unit of area or time and is usually a good predictor of the amount of damage and subsequent effects done to the ecosystem, such as tree mortality.

The *Firelihood* framework benefits from the addition of spatial and temporal effects (STEs) for both occurrence and size predictions (Fig. 11). The STEs account for patterns unexplained by the available explanatory variables. This type of effect that has already

been introduced in Ager et al. (2018), Preisler et al. (2004), Woolford et al. (2009). Here, the STEs include a random spatial effect of spatial coordinates and a random seasonal effect of the week of year. This allows the model to capture residual spatiotemporal variability of fire activity patterns. Further details about the components and functioning of the *Firelihood* model are given in Chapters 2, 3 and 4.

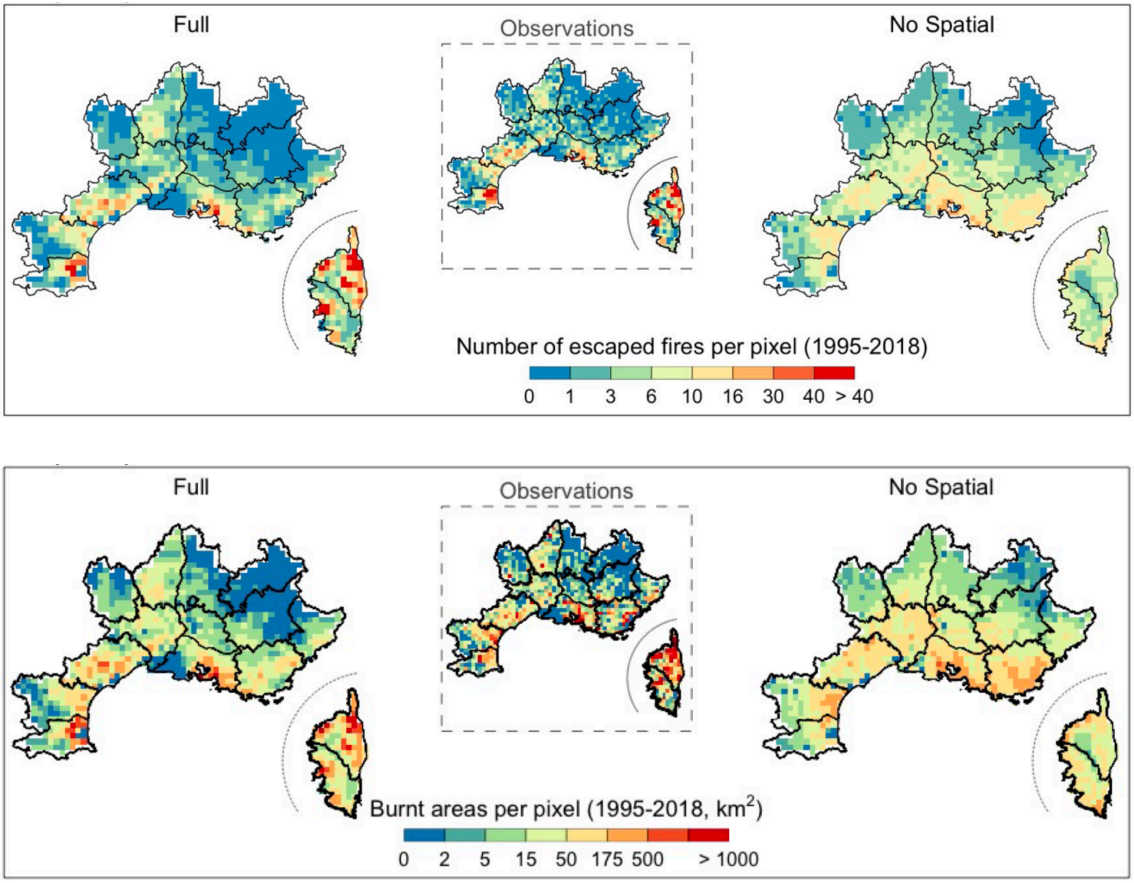


Figure 13: From Pimont et al. (2021). Representation of the occurrence and size model predictions on the SAFRAN pixels with STEs (Full) and without STEs (No Spatial) compared to the respective observations.

Although the potential of the *Firelihood* model for rating fire danger has been suggested (Fargeon, 2019; Pimont et al., 2021), several limitations are identified on the model. First, the model has limited information on land use and land cover, as it is only represented by the forest area per spatial unit. Second, the fire-weather is represented solely with the CanFWI, but in this specific region the index showed biases in the spatial and the temporal dimensions. Third, LFMC is not accounted and DFMC is only included through the CanFWI, contrarily to Rao et al. (2022), for example, who included the FMC used sensing measurements.

2. Influence of the non-climatic factors over the spatiotemporal patterns of wildfire activity and its temporal changes

This chapter introduces the first step of the research, its aim is to identify and quantify the role of the non-climatic factors in spatio-temporal activity and trends of wildland fires. The first version of the model Firelihood (Pimont et al., 2021) demonstrated promising predictive power, but it needed a large contribution of the spatiotemporal effects to represent the observed fire activity (see section 1.10. A probabilistic framework for fire activities in Mediterranean France). This first Firelihood used two variables, the CanFWI and surface of forest area in each pixel/day but did not include human and landscape factors. When adjusted without spatiotemporal effects, this model showed strong spatial and temporal bias found in the predictions, which confirmed that the fire activities in Southern France are driven by other factors than CanFWI and forest area. This is consistent with earlier studies, showing that the spatial patterns of fire activity in the Mediterranean Europe are much driven by patterns of human uses (Moreira et al., 2011), and most of fire ignitions have human causes (Ganteaume et al., 2013). Moreover, the sudden shift on the fire regime on 2003 is still unexplained (Pimont et al., 2021). Those could be attributed to changes on land uses, land cover or climate, or even, as already observed in previous shifts of the fire regime, changes on policy (Ruffault & Mouillot, 2015).

Based on the known origins of fire ignitions and fire activity patterns, one should focus on the landscape and the influence of human activities to explore the non-climatic drivers of fire activity (Costafreda-Aumedes et al., 2017). Therefore, in this study, the predictive model will benefit from landscape and human factors that should enable better predictions compared to a model using only weather, forest cover, and random effects. However, this is not straightforward because the influential variables identified in the literature have never been formally tested in the Mediterranean region of France. In addition, these key variables are often unavailable, and they accounted for through proxies with the necessary spatial resolution either unavailable or unknown.

The landscape and human activities information is extracted from the Land Use and Land Cover (LULC) layers provided by Copernicus (https://land.copernicus.eu) and

French institutions as the Insee and the IGN. Between all sources, they include the natural vegetation covers, with a coarse classification of their type, artificial vegetation covers as different types of agriculture and other green areas and man-made infrastructure as transportation networks and buildings, in addition of the shape of the terrain.

Each variable is quantified at the scale of the DFCI pixel grid (2km pixels), instead of the 8km SAFRAN pixels used in the previous iteration of Firelihood, in order to capture LULC variation at a finer scale than the weather data. The factors that control the ignition and initial spread of a fire are localized on the immediate surroundings of the fire so, the small 2km scale is justified. On the other hand, the drivers of fire size could be described at larger scales, Parisien et al. (2012) already mentioned the interest of finding the good drivers resolution for each fire size. Here, these larger scales are accounted through a spatial smoothing of the data at 4, 8 and 16 km. This part of the study is conducted at yearly scale because LULC variables are observed pluri-annual temporal scale and changes are minor for the temporal span of the data.

According to literature, the effect of certain LULC on fire activity is predictable (Costafreda-Aumedes et al., 2017). For example, the number of fires per pixel might be higher as more wildland area is available on the given pixel. The transport infrastructures are expected to increase fire ignition because they are a source of accidents and an access door for human activities near natural areas. Similarly, the wildland-urban interface (WUI) may play a role on increasing occurrence numbers as it quantifies specific human interactions with forested areas. Agricultural areas may also represent a high ignition potential, due to the machinery, the techniques used a possible accidents or negligent behaviors (Vallet et al., 2022). In addition, depending on the crop type, these areas can also behave as large stocks of fine fuels that propagate fire to or between wildland areas (Cruz et al., 2020). The topographical characteristics described on the pixel, such as aspect and slope, are expected to inform about small variations of the weather conditions and fire behavior, e.g., south exposed slopes being warmer and drier and rougher terrains are more challenging for fire spread control.

In South-Eastern France, the effects of the landscape factors over the fire regime are not well defined (Ruffault et al., 2017). In addition, the relative contribution of landscape factors in comparison to weather factors is unknown. To quantify the effects of these factors in South-Eastern France, as well as their relative contribution in explaining fire activity patterns, the Firelihood model has been modified to incorporate information on

landscape factors. This work is presented in the following article, where we show the improvements on Firelihood brought by the inclusion of landscape factors and enhanced spatial resolution. The new version of Firelihood models the number of fires of 1ha fires, and the probability of exceeding 10ha and 100ha, which can be later translated as the number of 10ha and 100ha fires. Once fitted, the model is used to produce simulations of theoretical scenarios where only climate, LULC or unexplained factors would have been changed at a time. The aim of these simulated scenarios is to attribute observed changes to the different groups of explanatory variables. Moreover, the simulations of the model serve to quantify the relative contributions of the different variables to spatial patterns thanks to a partition of variance.

2.1. Article: Disentangling the factors of spatio-temporal patterns of wildfire activity in south-eastern France

Disentangling the factors of spatio-temporal patterns of wildfire activity in South-eastern France

Running head: patterns of regional wildfire activity

Authors: Jorge Castel-Clavera[1*], François Pimont[1], Thomas Opitz[2], Julien Ruffault[1] and Jean-Luc Dupuy[1]

[1]*URFM, INRAE, Avignon, France*

[2]*BioSP, INRAE, Avignon, France*

[*]Corresponding author: jorge.castel-clavera@inrae.fr

Keywords: Firelihood, Forest fires, Risk management, Modelling, Bayesian, Mediterranean France, Spatial changes, temporal changes

Abstract

Background: Identifying if and how climatic and non-climatic factors drive local changes in fire regimes is, as in many other human-dominated landscapes, challenging in southeastern France where both heterogenous spatial patterns and complex fire trends are observed

Aim: We sought to identify the factors driving the spatial-temporal patterns of fire activity in southeastern France

Methods: We incorporated several non-climatic variables into the probabilistic *Firelihood* model of fire activity and implemented an enhanced spatio-temporal component to quantitatively assess remaining unexplained variations in fire activity.

Key results: Several non-climatic drivers (i.e., orography, land cover and human activities) contributed as much as fire-weather to the distribution of fire occurrence (> 1 ha) but less to larger fires (>10, 100 and 1,000 *ha*). Over the last decades, increased fire-weather induced a strong increase in wildfire

probabilities, which was actually observed on the western part of the region but not so in the east and Corsican Island, most likely due to reinforced suppression policies.

Conclusions: While spatial patterns in fire activity are driven by land-use and land-cover factors, temporal patterns were mostly driven by changes in fire-weather and unexplained effects potentially related to suppression policies but with large differences between regions.

Brief summary

To explain the contrasted fire activity patterns observed in Mediterranean France, we leveraged Bayesian spatio-temporal modeling to disentangle the relative contributions of climatic and non-climatic drivers. We showed that the impact of recent fire-weather increase has caused the increase of fire probability on the western part, but in the east, it has been overcompensated by a reduction of escaped-fire probability.

Introduction

It is well established that wildfire activity has changed in a number of regions of the world during the last decades (Jones et al., 2022). Wildfires result from the combination of many factors: fuel configuration, fire-weather conditions, a human- or naturally-triggered ignition source and the fire suppression and mitigation policies (Costafreda-Aumedes et al., 2017). As climate, vegetation and human settlement patterns are spatially diverse and evolving over time, the spatial and temporal trends of fire activity result from complex interactions between these factors that may sometimes act in opposite directions. Thus, while increased fire-weather have been observed in most areas of the world (Jolly et al. 2015), fire activity has shown a wide range of regional trends due to the influence of non-climatic drivers, including land-use land-cover (LULC) and fire suppression (Vilar et al., 2016; Park et al., 2021). Yet, identifying the drivers of these changes, and quantifying their relative contribution, often proves to be challenging (Pezzatti et al., 2013; Fernandes et al., 2014; Moreno et al., 2014).

In regions such as the Euro-Mediterranean area (EU-Med), where landscapes have a long history of human activities and practices, the spatial patterns of fire activity are strongly driven by human settlement patterns (Moreira et al., 2011). Humans dominate the wildfire regime supplying the vast

majority of ignitions (Ganteaume et al 2013) , but also alter land cover and perform suppression activities, both influencing burnt areas. The study of human-caused fire occurrence has therefore received much attention (Martinez et al 2009, Ganteaume et al., 2013) and modelling efforts (e.g., Olivera et al. 2012, Rodriguez et al. 2014, Costafreda et al. 2017, Ruffault and Mouillot 2017). Results have shown that the "wildland-urban interface" (WUI) and the "wildland-agricultural interface" (WAI), as well as the density of human settlements (such as road density or building density), are among the most important drivers of fire occurrence in the EU-Med (Galiana-Martin et al., 2011; Martin et al., 2019), although the importance of each of these factors varies from one region to another (Moreira et al., 2011; Rodrigues et al., 2014). In comparison, models for fire sizes (or for exceedance probabilities of size thresholds) incorporating LULC factors are less abundant in Europe (e.g., Moreira et al., 2010, Ager et al., 2014, Ruffault and Mouillot 2017). Results suggest that types and continuity of fuel play a major role in fire-spread probabilities but that differences in fire selectivity among flammable vegetation types decrease under severe fire-weather (Barros and Pereira 2014, Fernandes et al., 2016).

Like spatial fire patterns, historical trends in fire activity were largely driven by non-climatic drivers during the last decades in the EU-Med. While fire-weather increased over much of the area (Giannaros et al., 2020), fire activity, including fire numbers and burnt areas, significantly decreased in most regions (Turco et al., 2016, Silva et al., 2019) because of co-evolving human factors that overwhelmed the impact of climate change. Negative trends can be explained, at least in part, by an increased effort in fire management and prevention in several countries, in response to the large and destructive fires that occurred during the 80s and 90s (Moreno et al., 2014, Ruffault and Mouillot 2015), albeit this effect remains difficult to quantify at the local scale.

The stochasticity in the activity of wildfires, showing rare and sparse events since fire occurrence probabilities at a specific location and time remain low even if conditions are favorable, limits the possibility of conducting simple correlative analyses directly on the observations at the local scale. Handling fire stochasticity requires robust statistical frameworks that account for the different factors – fire-weather, LULC, fire prevention and fighting – affecting the probabilities of ignition, initial spread and spread to larger sizes at various spatial and temporal scales, from hourly and daily to yearly (e.g.,

Pimont et al., 2021). Many factors, as fire suppression, culture, social awareness among others, are not easy to incorporate with measurable metrics. High correlation in time and space between such factors can lead to underestimation of credible intervals and to increased risk of confusion between effects (i.e., of wrong attribution of wildfire activities to their causes). Such spatial or temporal effects have been accounted for in a few studies (e.g. Priesler 2004; Woolford 2020); they are likely not spatially uniform in magnitude and even direction, as recently suggested by a thorough analysis of spatio-temporal trends in fire numbers and sizes (Koh et al., 2022). The interaction between spatial and temporal has recently been addressed by Joseph et al. (2019) and Rodrigues et al. (2018).

The French Mediterranean area is a particularly challenging place to disentangle the drivers of spatio-temporal wildfire activity. The coupling of hot and dry summers with strong northly winds, known as 'Mistral' and 'Tramontane', promote the occurrence of large fire events in the region (Ruffault et al., 2016, 2018). However, the climate conditions and human-settlement locations are highly influential, with drier climate in the lowlands near the coast where human population and activities are dense, thereby increasing the probability of fire ignitions (Curt et al., 2016, Pimont et al., 2021). By contrast, lower fire activity (despite local "hot-spots") is observed in hinterland and inland mountains, where climate is less dry. Previous studies provided significant insights into local fire drivers (e.g. Curt et al., 2013, 2016; Ganteaume and Guerra 2018; Ganteaume and Barbero 2019; Fox et al 2015; Ruffault et al 2017) but, to date, a systematic assessment at the regional scale is still missing. In addition, as in other Mediterranean regions, fire-weather has increased in Mediterranean France during the last decades (Fréjaville and Curt 2015; Ruffault et al., 2016, Barbero et al. 2020), but the fire number and, above all, burnt areas have decreased by two folds or more (Curt and Fréjaville 2018). Earlier reductions in fire activity have been attributed to the effect of fire control policies (Fox et al., 2015; Pimont et al., 2021; Ruffault and Mouillot, 2015), but the magnitude of this effect is difficult to estimate because of the potential confusion with changes in climate, landscape factors or other socio-economic drivers.

Here we provide a comprehensive assessment of the spatial and temporal trends of wildfire drivers in southern France. We used counterfactual analyses based on spatiotemporal scenarios constructed thanks to the probabilistic *Firelihood* model (Pimont et al., 2021) to address two objectives. First, we

determined which landscape-related and human factors drive the spatial patterns of fire activity in southern France, assessed their effect on fire probability, and quantified their relative contribution in comparison to fire-weather. Second, for each of these factors, we question whether it contributed to the historical trend in fire activity, and if so, we aimed to quantify this contribution as well as its local variability across the region.

Material and Methods

Overview

Firelihood (FL) is a Bayesian probabilistic framework that models daily occurrence and size of wildfires as a marked point process (points describing occurrences and marks describing sizes) from explanatory variables (Pimont et al., 2021). FL accounts for the stochasticity of fire activity and models the occurrence of fires larger than 1 *ha* (escaped fires) as a Poisson process and the size of these fires as a combination of threshold exceedance probability and piecewise size distributions. FL includes spatio-temporal effects to account for limitations in explanatory variables. Effects are estimated with the R-INLA package, which implements the integrated nested Laplace approximation (Rue et al., 2017) and allows spatially structured Gaussian random effects with Matérn covariance function represented through the stochastic partial differential equation (SPDE) approach (Lindgren and Rue, 2015). The Matérn covariance is a flexible and widely used covariance model in spatial statistics that includes the exponential covariance function as a special case, and it offers the advantage of manageable numerical representations even with more than one thousand pixels as here (Krainski et al 2018).

The first instances of FL (later referred as FL1) included the Fire Weather Index (FWI) and forest area as explanatory variables. Spatio-temporal effects were implemented for the occurrence component and included a seasonal, spatial and annual effect, the latter representing a shift in fire numbers after the 2003 heatwave. Estimations of fire activity were carried out in 8-km pixels. Details on FL1 are provided in Supplementary S2, and the model has been thoroughly evaluated in Pimont et al. (2021). In FL1, the impact of unaccounted factors – corresponding to all unknown explanatory factors – on the estimation of known effects is mitigated by spatio-temporal effects, which additional stochasticity into the model to appropriately capture variability due to unknown factors and statistical uncertainty. Moreover, those

spatio-temporal effects allow a realistic representation of fire activities, so that spatio-temporal analyses are possible despite the stochasticity at play in fire observations (e.g., pixels without observed fires during the observation period).

To further study the evolution of spatio-temporal patterns of fire activity in a more detailed manner, we developed an extension of FL, referred to as FL2, which includes additional LULC at 2-km resolution and refined spatio-temporal effects. The occurrence component of FL2 includes a yearly effect and a spatial effect for local temporal tendencies. Size exceedance models for 10 and 100 ha include a yearly effect as well as spatial effects for SylvoEcoRegions (SERs).

The simulations performed with FL2 allowed to (1) quantify the relative contributions of the different variables to spatial patterns thanks to a partition of variance; (2) predict the evolution of those fire activities from the 1993-2002 decade that would have occurred during the 2009-2018 decade if only climate, LULC or unexplained factors had changed at a time, hence allowing us to attribute observed changes to the different explanatory variables.

Study area and fire activity

The study area consists of 15 NUTS3-level French administrative units located in south-eastern France (Fig. 1, 75.560 km^2), which concentrate the majority of burned area during the summer season in France. It is dominated by a Mediterranean climate and a few alpine stages. This area exhibits contrasted patterns of orography, vegetation, population and climate (Supplementary S1). Non-urban land covers include crop, coniferous and deciduous forests and shrublands. Fire records between 1993 and 2018 were extracted from the Prométhée database at a 2-km resolution, which was also used as the reference grid for our analyses. Observed fires are scattered over the territory with a few hot spots (Fig. 1). Here, we considered fire activity during the warm fire season (from May 25th to October 31st).

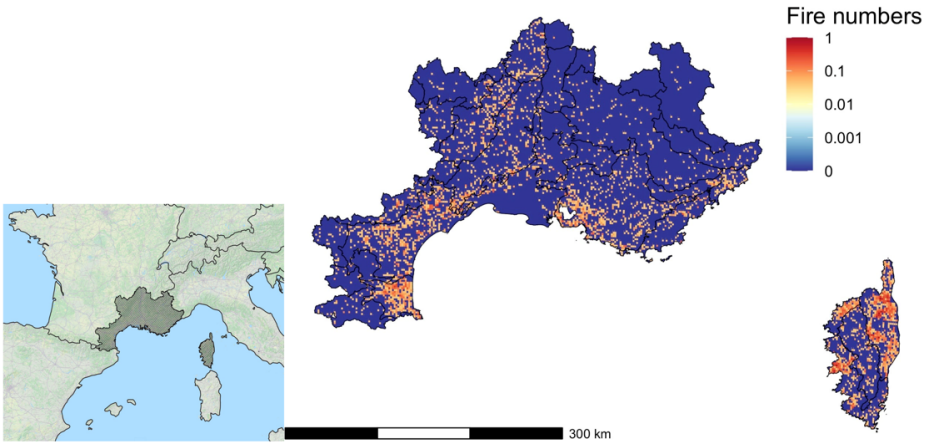


Figure 1: Patterns of fire occurrence over the study period (1993-2018) expressed as the mean number of fires strictly larger than 1ha per year in each 2km pixel of the reference grid (left), with a geographical context map in the right. Note that the scale bar corresponds to the left map.

Explanatory variables

The main explanatory variable was the daily FWI, which represents temporal and spatial variations in meteorological fire danger. FWI was computed over the 2-km reference grid from meteorological variables extracted from the SAFRAN reanalysis at 8-km pixel size (Vidal et al., 2010), using the cffdrs R package (Wang et al., 2017). For each 2-km pixel, we used the SAFRAN data corresponding to the 8-km pixel containing the center of the 2-km pixel of interest, i.e., we did not downscale the climatic data.

LULC data, listed in Table 1, was derived from Copernicus land cover data (https://land.copernicus.eu) and French Forest Service databases (Office National des Forêts, ONF). For some of the variables not available on a yearly basis (see Table 1), we filled the missing values by linear interpolation between available sources. Although LULC changes (See Supplementary S4) may not occur gradually we considered this smoothed approach better to avoid noise introduced in the specific years when data is released, which may accumulate the sudden changes having place during the previous period. Yearly data was extracted for the 2-km reference grid as either the sum of area per pixel, the mean in the pixel or the value at the pixel centroid, depending on the variable.

Smoothed values of LULC at coarser scales (4 to 16 km) were also tested as explanatory variables as an alternative to the default 2-km resolution. We hypothesized that the size threshold exceedance probability could be better explained at coarser scales than at ignition location. The variable "Fuel rating ONF" corresponds to a discrete vegetation sensitivity mapping ranging from 1 to 5, developed by the

French National Forest Service; although it is a static variable it creates no negative impacts on the model, since LULC changes are small over time, and this variable will give the same score to a forested area regardless of its area. Here, each index was weighted according to its corresponding area in the pixel, leading to a score ranging from 0 to 5. In order to improve spatial prediction of exceedance thresholds, we used 28 SERs as explanatory variables (Supplementary Fig. A1) corresponding to homogeneous forest productivity or forest habitats (IGN, 2019).

*Table 1: List of explanatory variables and their characteristics. The "Extraction" field describes how the information was compiled within each grid cell. The "Units" field expresses the represented dimension. Field "Time-step variability" is the period of temporal variation of each variable ("static" is for data that remains constant, either because of the lag of data or because its static nature). The "Source" is the database that provided the data for the variable (for Corine Land Cover (CLC), the class numbers considered are specified). *Interpolation for missing years.*

Variable	**Extraction**	**Units**	**Time step**	**Source**
Wildland area	Pixel area sum	ha	Yearly*	CLC 311-131, 322-324
Wildland area presence	Absence/presence	-	Yearly*	CLC 311-131, 322-324
Shrubland area	Pixel area sum	ha	Yearly*	CLC 322-324
Mix forest area	Pixel area sum	ha	Yearly*	CLC 313
Broadleaved forest area	Pixel area sum	ha	Yearly*	CLC 311
Coniferous forest area	Pixel area sum	ha	Yearly*	CLC 312
Fuel rating ONF	Pixel area weighted sum	Score (0 – 5)	Static	ONF
Aspect	Mean value/pixel	Azimuth (º)	Static	Copernicus
Slope	Mean value/pixel	Gegrees (º)	Static	Copernicus
Wuildland-Urban interface	Pixel area sum	ha	Yearly*	ONF
Agriculture	Pixel area sum	ha	Yearly*	CLC 211, 221-223, 241-244
Urban area	Pixel area sum	ha	Yearly*	CLC 111, 112, 121-124, 132, 133
Population density	Centroid value / Urban area	Num/pixel	Yearly	Insee
Road length	Sum distance/pixel	m/pixel	Static	IGN
Year	-	-	Yearly	-
Fire weather index	Centroid value	Non-dimensional	Daily	SAFRAN, cffdrs

FL2 development

The occurrence component assumed a Poisson distribution for fire numbers and explains the spatio-temporal variation of expected fire numbers:

$$\log N_{i,d}^{1ha} \sim \beta_0 + \underbrace{f_{FWI}(FWI_{i,d}) + f_{WEEK}(WEEK_{i,d})}_{FL1} + \beta(WAp_{i,y}) + \sum_{LULC} f_{LULC}(LULC_{i,y}) + f_{X,Y}(X_i, Y_i) + \underbrace{f_{YEAR}(y) + f'_{X,Y}(X_i, Y_i)(y - 1992)}_{Spatio-temporal\ trends} \quad (1)$$

where $N_{i,d}^{1ha}$ is the expected number of fires larger than 1ha in pixel *i* and day *d,* and the *f*-terms are the various non-linear effects. Subscript "*y*" indicates explanatory variable evolving on a yearly basis.

The effects of daily FWI and week-of-year (seasonal) were the same as in FL1. The wildland area presence ($WAp_{i,y}$) was a binary variable used to differentiate pixels with negligible wildland area from the other pixels. Spatial $(f_{X,Y})$ and spatio-temporal trend components capture the variability in wildfire activity which was not explained by other covariates. Both the spatial effects and the spatially varying change rate $f'_{X,Y}(X_i, Y_i)$ were represented with SPDE. The effect f_{YEAR} accounted for unexplained annual change in fire occurrence for the whole region, whereas $f'_{X,Y}(X_i, Y_i)$ allowed to disentangle regions where the occurrence increases over time ($f'_{X,Y} > 0$) from those where it decreased ($f'_{X,Y} < 0$) – with respect to other covariates.

The size component of FL1 has been improved towards refined exceedance-threshold-probability models for two thresholds *u* of 10 ha and 100 ha. The probability $p_{i,d}^u$ of a fire occurring in pixel *i* and during day *d* to exceed *u* is modelled through a logistic regression model:

$$log\frac{p_{i,d}^u}{1-p_{i,d}^u} = \beta_0^{p,u} + f_{FWI}^{p,u}(FWI_{i,d}) + \sum_{LULCM} f_{LULC}^{p,u}(LULC_{i,y}) + f_{YEAR}^{p,u}(y) + f_{BESAG}^{p,u}(SER) \quad (2)$$

These models, referred to as the size models for simplicity, also include a spatio-temporal component, which was not present in FL1. It consists of a yearly random effect and the spatial effect modelled with the spatially-conditionally autoregressive "Besag" model, allowing dependent effects between adjacent

SERs. The approach was used instead of SPDE because the fire size data is too small to estimate a random effect with Matérn covariance function at relatively high spatial resolution

For both the occurrence and size models, LULC variables were sequentially included according to the parsimony criterions known as Deviance Information Criterion (DIC - Sutanto et al., 2021), a non-dimensional indicator which is lower as higher the quality of the model for a same set of data, and the Area Under the Curve (AUC), scaling from 0.5 to 1, where 0.5 indicates a complete random prediction and 1 a perfect prediction of observation.

FL2 simulations and applications

For a given set of explanatory variables, the occurrence model was used to simulate multiple realizations (here 100) of $N_{i,d}^{1ha}$ for the full study region and period. These potential fire activities were averaged over time to draw a probabilistic occurrence map (Fig.2), which compares well with observations (Fig. 1). It shows that the probability to get a fire is not 0 even if no fires have been recorded over the last thirty years. The occurrence of fires larger than 10 and 100 ha was obtained by combining simulated occurrence with size models ($N_{i,d}^{10ha} = N_{i,d}^{1ha} p_{i,d}^{10ha}$ and $N_{i,d}^{100ha} = N_{i,d}^{1ha} p_{i,d}^{100ha}$).

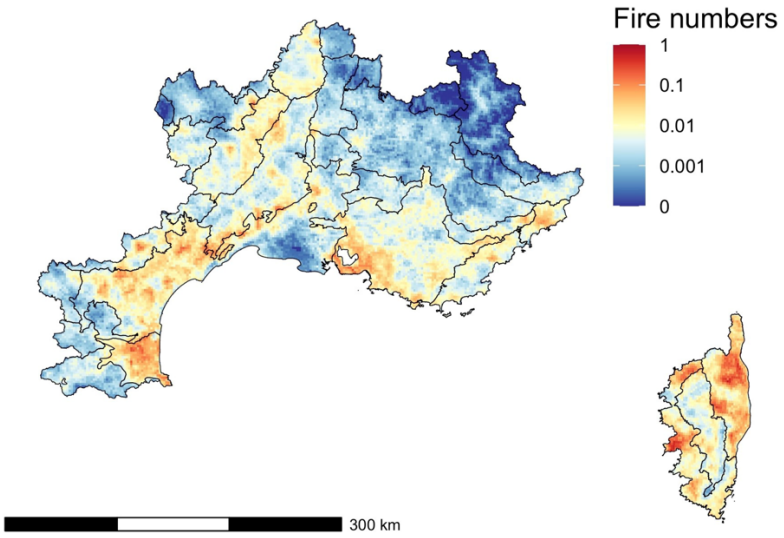


Figure 2: Simulation of 1ha fire numbers for the full study period (1993-2018) expressed as the mean number of fires strictly larger than 1ha per year in each 2km pixel of the reference grid.

In order to disentangle the role of the different types of effects (Fire-weather, LULC and other spatio-temporal effects) in spatial patterns of fire activities, we computed the yearly Poisson intensities related to each type in each pixel, by aggregating daily effects of fire-weather. We then compute the relative

contributions (RC) of the different effects as a partition of spatial variance for each year, which was then averaged for years of the study period. Details of this method, and how it applies to N^{1ha}, N^{10ha} and N^{100ha}, are provided in Supplementary S3.

In order to disentangle the role of the different types of effects in the evolution of spatial patterns, we simulated the fire activities during the recent decade (2009-2018) according to 4 counterfactual scenarios: (i) a scenario "Fire-weather change" where only climate changed (as observed on the recent period) by reproducing the LULC of the early 1993-2002 decade and ignoring spatio-temporal trends (i.e. ignoring $f_{YEAR}(YEAR_i) + f'_{X,Y}(X_i, Y_i)(YEAR_i - 1992)$ in Eq (1)); (ii) a scenario "LULC change" where only LULC changed over time, using the same FWI time series for the recent decade that the one observed during the earlier period and ignoring the spatio-temporal trends; and (iii) a scenario "Other temporal changes" where only the modelled spatio-temporal trends were considered for changes between the two decades, assuming no change in fire-weather and LULC since the 1993-2002 decade; and (iv) a scenario where the three types of variables evolved as observed, corresponding to the simulated actual recent decade (2009-2018). These scenarios were compared with activities simulated in 1993-2002 as a reference. Anomalies between the scenarios to the reference allowed us to attribute the changes, observed between the two decades (scenario *iv*), to the different types of explanatory variables, each type corresponding to one of the three first scenarios. For example, if "Fire-weather change" scenario matches the scenario *iv* it would mean that fire-weather change is responsible for most of the observed changes; this process is supported by the spatial representation of anomalies.

Results

Variable selection and model performance

Variable selection based on DIC and AUC is summarized in Table 2, leading to final models. Significant improvements were obtained when including LULC variables. Fire occurrence patterns were better explained by fine-scale landscape factors, contrary to size models for which 4 km and even 8 and 16 km led sometimes to better predictions. We recall that the partial effects of a variable reflect the effect of this variable with all other variables being fixed. The partial effects of the occurrence model were consistent, with intensity increases associated with fuel type rating, FWI, population and road length

(Fig.3 B,C,H,I). Inverted U-shape responses associated with wildland area, aspect, slope and wildland urban interface (Fig.3 A,D,E,G) were explained by reverse effects of these factors on ignition and initial spread, as already observed for wildland area in Pimont et al. (2021). These covariates have a maximum near respectively 250 ha/pixel, 80ha/pixel, 7.5º to 17.5º, and south and south-east expositions.

The yearly effect (Fig. 3J) confirmed the decay in fire activity observed after the 2003 crisis. The positive spatial effect (Fig. 3K) reveals that unexplained factors led to higher occurrence than expected from FWI and LULC predictors in some sub-regions, e.g., in the island of Corsica, or the inner mountain regions. The spatial distribution of annual trends (Fig. 3L) was contrasted from west to east: positive trends to the west and negative to the east.

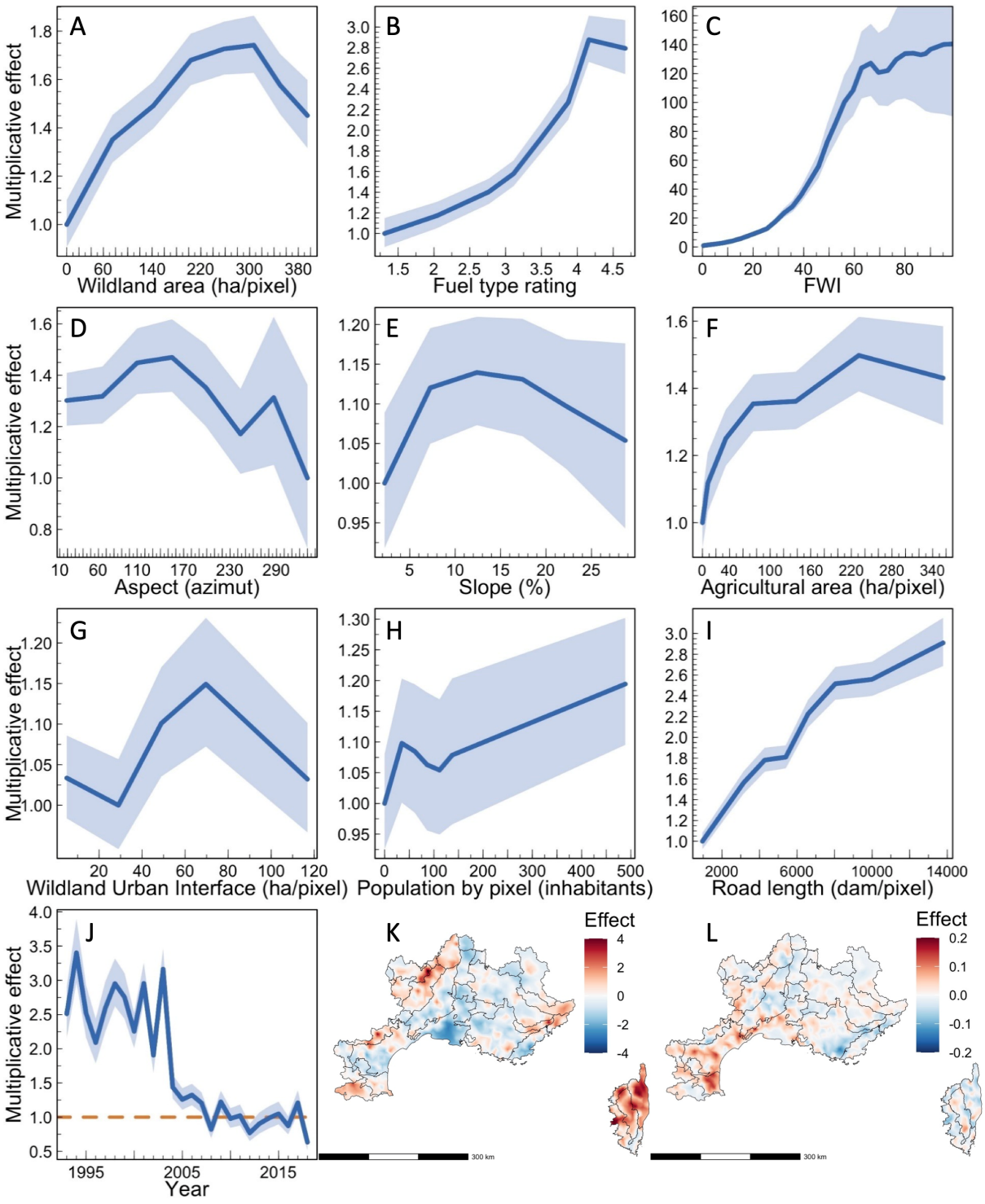


Figure 3: Multiplicative effects (A-J) of the predictor variables in the occurrence model, represented as a graphic line with confidence intervals; the X axis shows the value in each of the variable-specific bins; the Y axis is the value of the multiplicative effect, in other words, the factor by which the expected number of fires changes with respect to the case of a multiplicative effect of 1.. Spatial effect $(f_{X,Y})$ *for the occurrence model from the SPDE model (K); and spatially dependent annual trend* $(f'_{X,Y})$ *for the occurrence model (L). Note that FWI's effect (C) has a larger amplitude compared to the other effects due to its daily scale inherited from the precedent model, Firelihood 1.0 from Pimont et al. (2021).*

Table 2: Variable selection of the occurrence model (OCCURRENCE) and size models for the 10 ha (SIZE 10) and the 100ha (SIZE100) exceedance thresholds. Variables have been added and tested individually. For the sake of brevity, the results for all spatial aggregations have not been presented in the table (only the best one).

Variables included in the model	DIC	AUC (training)	AUC (validation)
OCCURRENCE			
Intercept + WAp + Fuel + WA	58876	0,731	0,736
Intercept + WAp + Fuel + WA + Aspect + Slp	58390	0,743	0,736
Intercept + WAp + Fuel + WA + Aspect + Slp + WUI + Pop + Agri + Roads	57922	0,754	0,751
Intercept + WAp + Fuel + WA + Aspect + Slp + WUI + Pop + Agri + Roads + SPDE + Sp_temp + Years	50131	0,877	0,827
SIZE 10			
Intercept + FWI	5948	0,642	0,571
Intercept + FWI + WA(4km) + Agri (4km) + Con + Brl + Shr	5719	0,684	0,651
Intercept + FWI + WA(4km) + Agri (4km) + Con + Brl + Shr + Pop + Slp + Mxf (8km)	5673	0,694	0,669
Intercept + FWI + WA(4km) + Agri (4km) + Con + Brl + Shr + Pop + Slp + Mxf (8km) + Fuel + Besag + Years	5646	0,706	0,662
SIZE 100			
Intercept + FWI	2139	0,727	0,692
Intercept + FWI + WA(4km) + Agri (4km) + Con + Brl + Shr (4km)	2035	0,767	0,759
Intercept + FWI + WA(4km) + Agri (4km) + Con + Brl + Shr (4km) + Pop + Slp (16km)	2019	0,776	0,776
Intercept + FWI + WA(4km) + Agri (4km) + Con + Brl + Shr (4km) + Pop + Slp (16km) + Fuel + Besag + Years	2002	0,789	0,776

Wap - Wildland area presence, **Fuel** - ONF's fuel rating, **WA** - wildland area, **Aspect** - Aspect of the pixel, **Slp** - Slope, **WUI** - Wildland to Urban interface, **Pop** - Population per pixel, **Agri** - Agricultural area, **Roads** - Road length, **Con** - Coniferous forest area, **FWI** - Fire Weather Index, **BRL** - Broadleaved forest area, **Shr** - Shrubland area, **Mxf** - Mixed forest area, **SPDE** - Spatial model, **Sp_temp** - Spatio-temporal model, **Besag** - Spatial model based on SERs.
The distance indicated in the patenthesis referes to the spatial aggregation used for a given variable; when there is no writen distance, the variable has been used in the original 2km resolution.

The partial effects of both size models were similar (100 ha in Fig. 4 and 10 ha in Supplementary Fig. A2), with monotonic responses. Exceedance probabilities increased with wildland area, fuel type rating, FWI, coniferous and shrubland areas and slope (Fig. 4 A,B,C,D,E,H), and decreased with broadleaved and agricultural areas, and population (Fig. 4 F,G,I). Upper ranges of the drivers were generally associated with saturations. The yearly effect for the 100ha threshold decreased – except the 2003 peak – but this change was not significant as current trends (blue dashed line) were inside the credible interval of early years (Fig. 4J). For the 10ha threshold (Fig. S2K), after a significant decrease until 2003, no clear trend was observed during the recent years. The spatial effect (Fig. 4K) shows a west-east gradient with increased probability in the eastern part and in mountainous regions, which could be explained by operational constraints in suppression policies in remote mountainous regions.

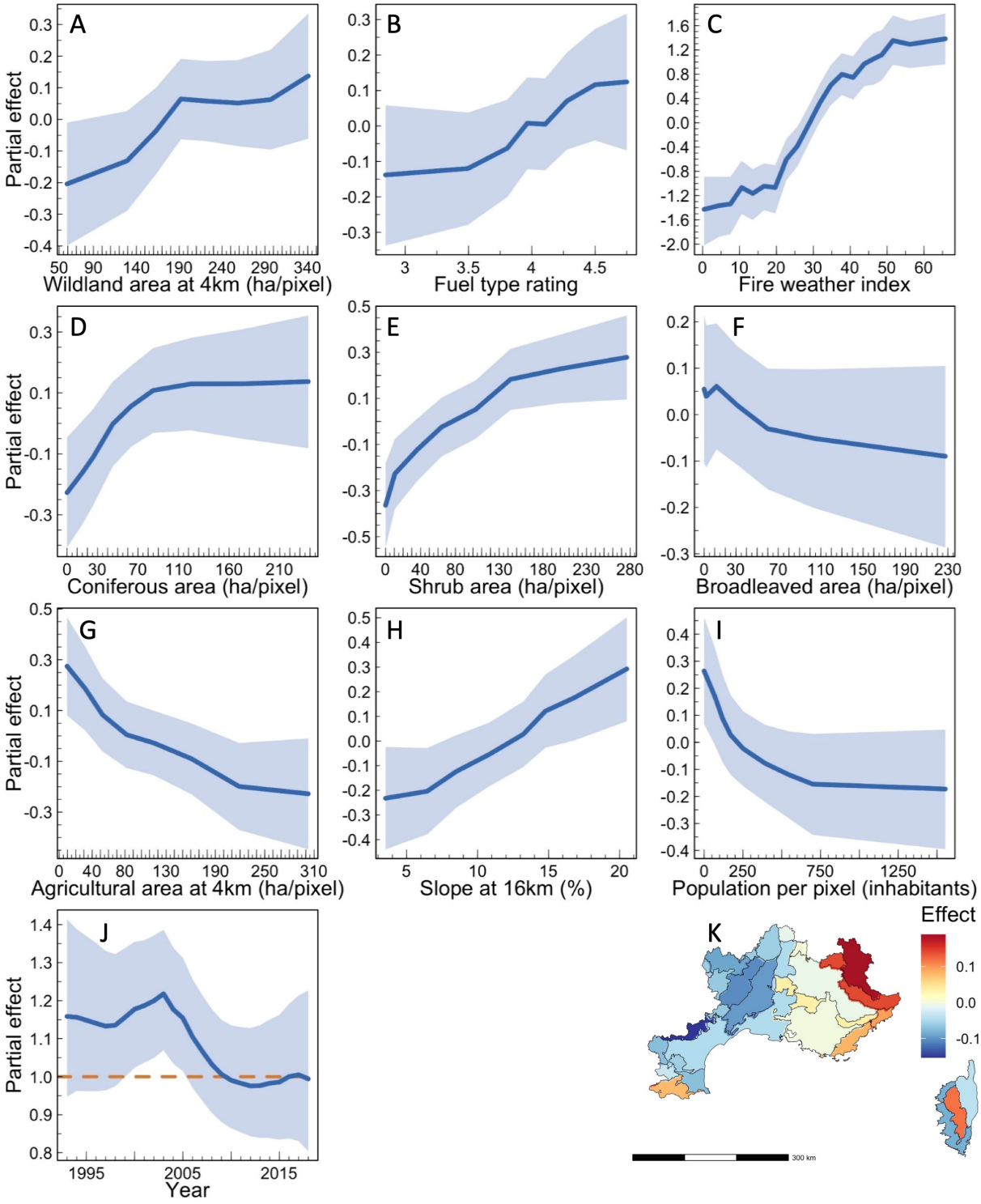


Figure 4: Partial effects (A-J) of the predictor variables for the 100ha exceedance probability model intervals; the X axis shows the value in each bin; the Y axis is the contribution of the predictor variable's value to the logit of the probability of exceedance. Spatial effect by SER from the Besag component in the 100ha exceedance (K).

Factors explaining spatial distributions and their changes

The partition of variance of fire activity allowed us to decompose the spatial variability between four types of effects (Fire-weather, LULC, spatial and temporal) with quite different importance for relative

wildfire risk over the full period (Fig.5). We found that the spatial effect had the biggest contribution to 1ha-number simulations, while LULC and fire-weather explained each one fourth of the total variance. The "temporal effect" in spatial patterns – associated with spatial trends – was less important. For larger fire numbers (10 and 100 ha), the fire-weather explained the largest fraction of spatial variance, followed by the spatial effect, while LULC were important (17-20%). The temporal effect – associated with temporal trends of occurrence – was marginal. These results show that fire-weather and LULC together explained roughly 50% of spatial distribution of 1ha fires and up to 70% for larger fires (10 and 100 ha).

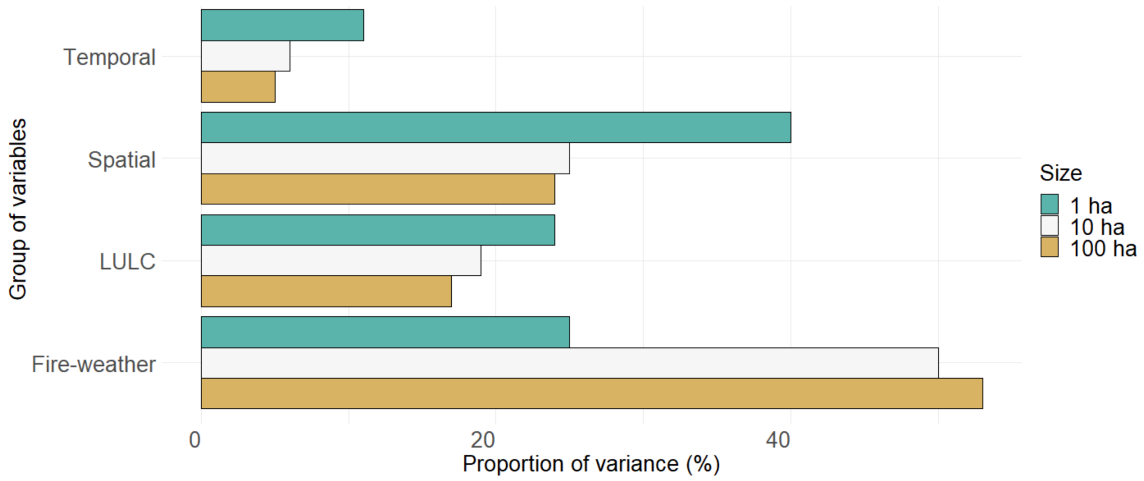


Figure 5: Partition of spatial variance according to each group of explanatory variables, expressed as a proportion for the full model. Spatial variances are computed for log Poisson intensities (Supplementary S2).

The comparison between the earlier and the recent decade for the actual evolution of variables confirmed a decrease in fire numbers for all sizes, ranging between –30 and –45% (Fig. 6). When considering changes in factors one at a time, simulations for the recent decade were more contrasted. The "Fire-weather-change" scenario shows an increase in fire events, while the "LULC-change" scenario did not have a notable effect. The "other-temporal-change" scenario showed a decrease in fire numbers, larger than the "actual" simulations. Hence, the potential increase caused by fire-weather change for the three different fire sizes (1, 10 and 100 ha) was overcompensated by temporal changes which were not explained by LULC variables. We investigated the spatial distributions of these changes by mapping anomaly maps between past (reference period) and recent scenarios (Fig. 7, and supplementary fig. A3 and A4), showing that the spatial distributions were similar throughout fire sizes, with the North-Eastern alpine region showing no noticeable changes. Both the real-present and the other-temporal-changes

scenarios (Fig. 7A,D) showed a widespread decrease over the eastern regions and local increases in the western regions, revealing very heterogeneous trend over the territory. Fire-weather change induced scattered increases by "hot spots", even where reductions are observed (Fig. 7B). Finally, changes in LULC produced marginal changes (Fig. 7C), not explaining the recent fire regime changes.

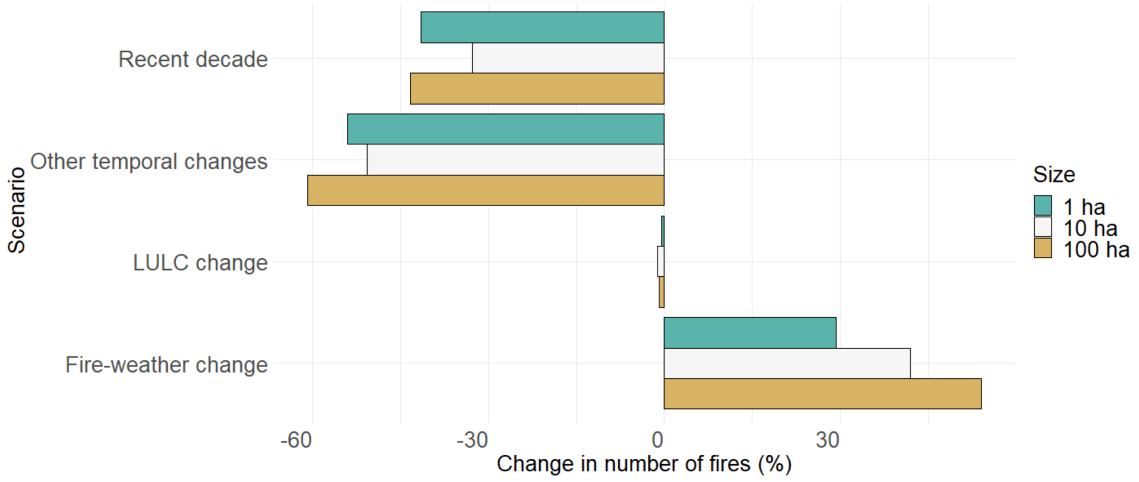


Figure 6: Comparison of simulations of fire activities between the past decade (1990-2003) and the recent decade (2009-2018) scenario and three alternative present scenarios in mean yearly fire numbers per pixel.

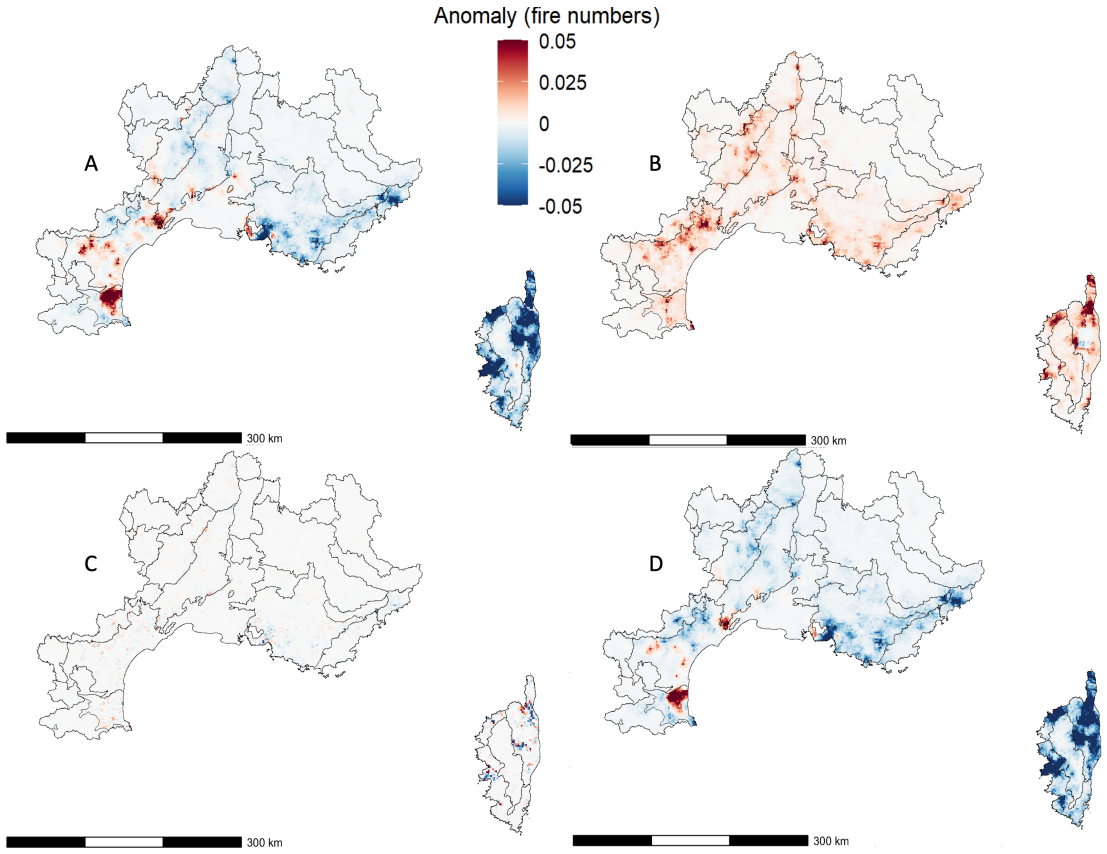


Figure 7: Spatial distribution of anomalies (in mean yearly fire numbers per pixel) between the 1993-2002 and 2009-2018 decades for 1ha fire numbers. Same figure for 10 and 100ha are shown in Supplementary figures A3 and A4. A: Recent decade; B: Fire-weather; C: LULC, D: Other-temporal-changes.

Discussion

Spatio-temporal modeling and analysis

The goal of this work was to understand the underlying drivers of spatial-temporal patterns of fire activities in southeastern France, with a focus on the differences between past and recent fire regimes. For this purpose, we implemented spatio-temporal random effects in the Firelihood probabilistic framework (Pimont et al., 2021). This approach offers at least two important advantages for the modelling of fire activity. First, we found that these random effects provided realistic fire activity scenarios throughout southeastern France, despite fire data sparseness. Second, in a context where unexplained factors depending on the local conditions are very important (Díaz-Avalos et al., 2016), this approach allowed for accurate estimations of the explanatory variables' effects and of residual unexplained spatio-temporal random effects.

Regarding the occurrence model's component, the spatio-temporal effects implemented in FL2 were very sophisticated, as they account for spatial factors, overall temporal changes – seasonal and yearly – and trends in spatial patterns, the later effect allowing spatial patterns to change over time linked to processes not directly related with the included explanatory factors. The spatio-temporal effects of the size models (exceedance probability) were simpler because of the reduced size of the dataset for 10 and 100 ha fires. However, the temporal effects allowed for potential yearly changes and spatial effects structured by SERs and proved to enhance the predictive ability of the model.

LULC factors of fire activities

Overall, the LULC factors retained in our model of fire occurrence (> 1ha) were consistent with the findings of previous modeling studies in southern Europe. Indeed, we found that most prevalent factors in human-caused fire occurrence models were human-related, including population density, dwellings and access networks to forest and natural land areas (Costafreda et al., 2017). However, vegetation cover of different fuel types was not selected in our fire occurrence model, while fuel rating, usually not included in previous studies, was found to be a better predictor. On the contrary, the models for exceedance probability of 10 and 100 ha selected vegetation cover variables as predictors, with either

positive (Shrubland and Conifer) or negative (Broadleaves) effects, which is consistent with previous models for fire size (Diaz-Avalos et al. 2016), and land cover fire-proneness studies in Europe (Moreira et al., 2009). Vegetation cover variables were also found to be more influential on fire sizes than on fire occurrence in the United States (Hawbaker et al., 2013).

Road density had the strongest effect on fire occurrence among all LULC in our model, confirming the key role of accessibility to forest areas found in other regions of Europe (e.g. Vilar et al., 2010; Serra et al., 2014). On the other hand, this factor was not selected in size models, in contrast to previous studies in the United States that selected accessibility (distance to roads) (Dickson et al., 2006; Ager et al., 2013; Hawbaker et al., 2013). As observed by Costafreda et al. (2017), interfaces between forest/wildland, agriculture and urban areas, and vegetation cover, were also found to be highly influential on fire occurrence.

LULC variables played an important role in fire occurrence, explaining ~25% of its spatial distribution, (Fig. 4), with the same order of magnitude as fire-weather, and in accordance with previous results on the western part of the study area (Ruffault and Mouillot 2017). Nonetheless, unexplained spatial factors and their temporal changes still represented half of the spatial variability in occurrence, meaning that these commonly used predictors did not fully explain the fire occurrence hot-spots observed in western areas and Corsica, or cold-spots observed in mountainous areas or alluvial plains (Fig. 2). For the size models, the fire-weather explained the major part of the spatial distribution, and the share of the LULC and unexplained spatial effects were smaller. This might partly be explained by the less sophisticated spatio-temporal effects for exceedances. However, the larger picture that emerges is that fire occurrence was explained equally by fire-weather and by landscape and socio-economic drivers, and once a fire has been ignited, the major driver for its size was the weather during the fire event.

Temporal trends in the fire regime

Fire-weather change alone was found to drive significant increases in potential fire activity, due to a shift of the distribution of the FWI values between the two periods of our study. Around 50% of this increase in fire-weather has already been attributed to anthropogenic climate warming (Barbero et al., 2020). Despite the importance of LULC variables over the spatial distribution of fires, LULC-changes

alone induced marginal changes in fire activity. Temporal changes related to unexplained factors were the main driver, compensating the climate-induced change, contrasting with the findings of Viedma et al. (2018) in a west-central Spanish landscape, who documented virtually no trend in fire-weather, but substantial dynamics in landscape drivers of fire, with significant associated trends in fire activity using a 10km pixel-grid.

Fig. 2J, 3J and A2J give an indication of overall temporal changes associated with unexplained factors, allowing us to conclude that the decay was mostly caused by a major reduction in escaped fire numbers after 2003, and the posterior evolution has been limited. Few variations were detected on exceedance probabilities, especially in the recent years. It is very likely that fire suppression policy reduced the number of escaped summer fires, as suggested for the 1973-2005 period in the western part of the region (Ruffault et al., 2015), or for the 1976-2009 period in most areas of the region (Fréjaville and Curt, 2017); similar trends were reported in Spain by Moreno et al. (2014). Indeed, following the record-breaking year 2003 in terms of burnt area in France, the suppression was reinforced with new fire-fighting material and better training capacities, but such changes were not easy to capture with quantitative variables.

No more reduction has been detected since 2007. In particular, the capability to limit the escaped fires that become large (>100ha) has not been improved, especially during the last decade, matching with Evin et al. (2018), who reported a successful suppression strategy, but no significant trend on larger fires. Those findings may put in question the capacity of the French prevention and suppression system to absorb future increases in fire activity associated with climate change, especially given the fact that large fire numbers are expected to increase faster than 1ha fires (Fargeon, 2019).

The spatial distribution of unexplained changes revealed new interesting patterns, providing more insights into the possible temporal drivers of the observed evolution. The positive effect of the fire-weather was concentrated in the current fire activity hot-spots, whereas the unexplained temporal changes were negative in almost all the eastern part of the region. Positive unexplained temporal trends in the south-west would be, according to fire managers, associated with areas of agricultural abandonment, raising the ignition potential of this area. It is also likely that suppression policies were

less reinforced in these traditionally less fire-prone areas and struggled to face the observed fire-weather increase.

Conclusion

Analyzing factors of fire activities at regional scales is highly challenging because of the stochasticity and the non-stationarity of both these factors and the fire activities. The modeling framework and the simulation plan allowed us to reveal important changes in fire activity and gain insights on some of their drivers. The simulation-based approach allowed us to disentangle the relative contribution of usual explanatory variables to spatial fire activities and to identify differences between fire sizes. We also further explored recent changes in fire regimes and found that the main temporal changes observed over recent decades were driven by unexplained factors suggesting an important contribution of prevention and suppression policies to these trends. Our study further revealed that, if very significant reductions on the number of escaped fires (1ha fires) was observed after the 2003 heatwave, no significant overall reduction could be detected over the last decade, neither on fire numbers nor on the ability of escaped fires to turn into large fires, even if strong regional differences were detected along a west to east gradient. In a context where climate change is expected to further increase fire weather in this region in the next decades (Fargeon et al. 2020), this raises the question as to how long current low levels in fire activity could be maintained by ongoing efforts in fire suppression and prevention.

Conflicts of interest

The authors declare no conflicts of interest

Declaration of Founding

This research did not receive any specific funding

Data availability statement

The data that support this study comes from six sources with different availability conditions:

- Data for CLC available at Copernicus (Land Monitoring Service): https://land.copernicus.eu/
- Fire records available at Prométhée database: https://www.promethee.com/
- Landscape descriptive geographical layers from the French IGN (Institut national de l'information géographique et forestière) at: https://geoservices.ign.fr/catalogue
- Population data available from the French Insee (Institut national de la statistique et des études économiques) at: https://www.insee.fr/fr/statistiques?theme=0
- Climatic SAFRAN data available under demand via SICLIMA Extraction application at: https://www6.paca.inrae.fr/agroclim/Demande-des-donnees-de-Meteo-France
- Fuel classification is not publicly available. It is managed by the French ONF (Office National des Forêts).

References

Ager AA, Buonopane M, Reger A, Finney MA (2013) Wildfire exposure analysis on the national forests in the pacific northwest, USA. *Risk Analysis* **33**, 1000–1020. doi:10.1111/j.1539-6924.2012.01911.x.

Ager AA, Preisler HK, Arca B, Spano D, Salis M (2014) Wildfire risk estimation in the Mediterranean area. *Environmetrics* **25**, 384–396. doi:10.1002/env.2269.

Barbero R, Abatzoglou JT, Pimont F, Ruffault J, Curt T (2020) Attributing Increases in Fire Weather to Anthropogenic Climate Change Over France. *Frontiers in Earth Science*, **8**, 1–11. doi:10.3389/feart.2020.00104.

Barros AMG, Pereira JMC (2014) Wildfire selectivity for land cover type: Does size matter?. *PLoS ONE*, **9**. doi:10.1371/journal.pone.0084760

Costafreda-Aumedes S, Comas C, Vega-Garcia C (2017) Human-caused fire occurrence modelling in perspective: A review. *International Journal of Wildland Fire*, **26**, 983–998. doi:10.1071/WF17026.

Curt T, Borgniet L, Bouillon C (2013) Wildfire frequency varies with the size and shape of fuel types in southeastern France: Implications for environmental management. *Journal of Environmental Management*, **117**, 150–161. doi:10.1016/j.jenvman.2012.12.006.

Curt T, Frejaville T (2018) Wildfire Policy in Mediterranean France: How Far is it Efficient and Sustainable?. *Risk Analysis*, **38**, 472–488. doi:10.1111/RISA.12855.

Curt T, Fréjaville T, Lahaye S (2016) Modelling the spatial patterns of ignition causes and fire regime features in southern France: Implications for fire prevention policy. *International Journal of Wildland Fire*, **25**, 785–796. doi:10.1071/WF15205.

Díaz-Avalos C, Juan P, Serra-Saurina L (2016) Modeling fire size of wildfires in Castellon (Spain), using spatiotemporal marked point processes. *Forest Ecology and Management*, **381**, 360–369. doi:10.1016/j.foreco.2016.09.013.

Dickson BG, Prather JW, Xu Y, Hampton HM, Aumack E.N, Sisk TD (2006) Mapping the probability of large fire occurrence in northern Arizona, USA. *Landscape Ecology*, **21**, 747–761. doi:10.1007/s10980-005-5475-x.

Evin G, Curt T, Eckert N (2018) Has fire policy decreased the return period of the largest wildfire events in France? A Bayesian assessment based on extreme value theory. *Natural Hazards and Earth System Sciences,* **18**, 2641–2651. doi:10.5194/nhess-18-2641-2018.

Fargeon H. (2019) Effet du changement climatique sur l'évolution de l'aléa incendie de forêt en France métropolitaine au 21ème siècle.

Fernandes PM, Barros AMG, Pinto A, Santos JA (2016) Characteristics and controls of extremely large wildfires in the western Mediterranean Basin. *Journal of Geophysical Research: Biogeosciences*, **121**, pp. 2141–2157. doi:10.1002/2016JG003389.

Fernandes PM, Loureiro C, Guiomar N, Pezzatti GB, Manso FT, Lopes L (2014) The dynamics and drivers of fuel and fire in the Portuguese public forest. *Journal of Environmental Management*, **146**, 373–382. doi:10.1016/J.JENVMAN.2014.07.049.

Fox DM, Martin N, Carrega P, Andrieu J, Adnès C, Emsellem K, Ganga O, Moebius F, Tortorollo N, Fox EA (2015) Increases in fire risk due to warmer summer temperatures and wildland urban interface changes do not necessarily lead to more fires. *Applied Geography*, **56**, 1–12. doi:10.1016/j.apgeog.2014.10.001.

Fréjaville T, Curt T (2015) Spatiotemporal patterns of changes in fire regime and climate: defining the pyroclimates of south-eastern France (Mediterranean Basin). *Climatic Change*, **129**, 239–251. doi:10.1007/s10584-015-1332-3.

Fréjaville T, Curt T (2017) Seasonal changes in the human alteration of fire regimes beyond the climate forcing. *Environmental Research Letters*, **12**. doi:10.1088/1748-9326/aa5d23.

Galiana-Martin L, Herrero G, Solana J (2011) A Wildland–Urban Interface Typology for Forest Fire Risk Management in Mediterranean Areas. doi:10.1080/01426397.2010.549218.

Ganteaume A, Barbero R (2019) Contrasting large fire activity in the French Mediterranean. *Natural Hazards and Earth System Sciences*, **19**, 1055–1066. doi:10.5194/nhess-19-1055-2019.

Ganteaume A, Guerra F (2018) Explaining the spatio-seasonal variation of fires by their causes: The case of southeastern France. *Applied Geography*, **90**, 69–81. doi:10.1016/j.apgeog.2017.11.012.

Ganteaume A, Jappiot M (2013) What causes large fires in Southern France. *Forest Ecology and Management,* **294**, 76–85. doi:10.1016/j.foreco.2012.06.055.

Giannaros TM, Lagouvardos K, Kotroni V (2020) Performance evaluation of an operational rapid response fire spread forecasting system in the southeast mediterranean (Greece). *Atmosphere*, **11**, pp. 1–17. doi:10.3390/atmos11111264.

Hawbaker TJ, Radeloff VC, Stewart SI, Hammer RB, Keuler NS, Clayton MK (2013) Human and biophysical influences on fire occurrence in the United States. *Ecological Applications*, **23**, 565–582. doi:10.1890/12-1816.1.

IGN (2019) Résultats par sylvoécorégions - INVENTAIRE FORESTIER. Available at https://inventaire-forestier.ign.fr/spip.php?rubrique253#nh1 (Accessed: 29 October 2021)

Jolly WM, Cochrane MA, Freeborn PH, Holden ZA, Brown TJ, Williamson GJ, Bowman DMJS (2015) Climate-induced variations in global wildfire danger from 1979 to 2013. *Nature Communications*, **6,** 1–11. doi:10.1038/ncomms8537.

Jones MW, Abatzoglou JT, Veraverbeke S, Andela N, Lasslop G, Forkel M, Smith AJP, Burton C, Betts RA, Werf GR, Sitch S, Canadell JG, Santín C, Kolden C, Doerr SH, Le Quéré C (2022) Global and regional trends and drivers of fire under climate change. *Reviews of Geophysics* [Preprint]. doi:10.1029/2020rg000726.

Joseph MB, Rossi MW, Mietkiewicz NP, Mahood AL, Cattau ME, St Denis LA, Nagy RC, Iglesias V, Abatzoglou JT, Balch JK (2019) Spatiotemporal prediction of wildfire size extremes with Bayesian finite sample maxima. *Ecological Applications*, **29**, 1266–1281. doi:10.1002/eap.1898.

Koh J, Pimont P, Dupuy JL, Opitz T (2021) Spatiotemporal wildfire modelling through point processes with moderate and extreme marks. doi:10.48550/arXiv.2105.08004

Krainski E, Gómez-Rubio V, Bakka H, Lenzi A, Castro-Camilo D, Simpson D, Lindgren F, Rue H (2018). Advanced spatial modeling with stochastic partial differential equations using R and INLA.

Lindgren F, Rue H (2015) Bayesian Spatial Modelling with R-INLA. *Journal of Statistical Software* [Preprint].

Martín Y, Zúñiga-Antón M, Rodrigues Mimbrero M (2019) Modelling temporal variation of fire-occurrence towards the dynamic prediction of human wildfire ignition danger in northeast Spain. *Geomatics, Natural Hazards and Risk*, **10**, 385–411. doi:10.1080/19475705.2018.1526219.

Martínez J, Vega-Garcia C, Chuvieco E (2009) Human-caused wildfire risk rating for prevention planning in Spain. *Journal of Environmental Management*, **90**, 1241–1252. doi:10.1016/j.jenvman.2008.07.005.

Moreira F, Andersen A, Anderson SAJ, Ascoli D, Baeza J, Barbati A, Duguy B, Finney M, Gill M, Keane R, Rigolot E, Ryan K, Williams D, Xanthopoulos G (2010) A comparison of landscape planning

approaches and practices for strategic fuel management in southern Europe, Australia, New Zealand and the USA.

Moreira F, Vaz P, Catry F, Silva J.S. (2009) Regional variations in wildfire susceptibility of land-cover types in Portugal: Implications for landscape management to minimize fire hazard. *International Journal of Wildland Fire*, **18,** 563–574. doi:10.1071/WF07098.

Moreira F, Viedma O, Arianoutsou M, Curt T, Koutsias N, Rigolot E, Barbati A, Corona P, Vaz P, Xanthopoulos G, Mouillot F, Bilgili E (2011) Landscape e wildfire interactions in southern Europe: Implications for landscape management. doi:10.1016/j.jenvman.2011.06.028.

Moreno MV, Conedera M, Chuvieco, E, Pezzatti, GB (2014) Fire regime changes and major driving forces in Spain from 1968 to 2010. *Environmental Science and Policy*, **37**, 11–22. doi:10.1016/J.ENVSCI.2013.08.005.

Oliveira S, Oehler F, San-Miguel-Ayanz J, Camia A, Pereira JMC (2012) Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest. *Forest Ecology and Management*, **275**, 117–129. doi:10.1016/j.foreco.2012.03.003.

Park IW, Mann M., Flint LE, Flint AL, Moritz M (2021) Relationships of climate, human activity, and fire history to spatiotemporal variation in annual fire probability across California. *PLoS ONE*, **16.** doi:10.1371/journal.pone.0254723.

Pezzatti GB, Zumbrunnen T, Bürgi M, Ambrosetti P, Conedera M (2013) Fire regime shifts as a consequence of fire policy and socio-economic development: An analysis based on the change point approach. *Forest Policy and Economics*, **29**, 7–18. doi:10.1016/J.FORPOL.2011.07.002.

Pimont F, Fargeon H, Opitz T, Ruffault J, Barbero R, Martin-StPaul N, Rigolot E, Riviére M, Dupuy JL (2021) Prediction of regional wildfire activity in the probabilistic Bayesian framework of Firelihood. *Ecological Applications*, **31**, 1–24. doi:10.1002/eap.2316.

Preisler HK, Benoit JW (2004) A state space model for predicting wildland fire risk. *Proceedings of the Joint Statistics Meetings, Toronto* [Preprint].

Rodrigues M de la Riva J, Fotheringham S (2014) Modeling the spatial variation of the explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression. *Applied Geography*, **48**, 52-63.

Rodrigues M, Jiménez-Ruano A, Peña-Angulo D, de la Riva J (2018) A comprehensive spatial-temporal analysis of driving factors of human-caused wildfires in Spain using Geographically Weighted Logistic Regression. *Journal of Environmental Management*, **225**, 177–192. doi:10.1016/j.jenvman.2018.07.098.

Rodriguez y Silva F, Molina JR, Leal JR (2014) The efficiency analysis of the fire control operations using the VISUAL-SEVEIF tool. *Advances in forest fire research* [Preprint]. doi:10.14195/978-989-26-0884-6.

Rue H, Riebler A, Sørbye SH, Illian JB, Simpson DP, Lindgren FK (2017) Bayesian computing with INLA: a review. *Annual Review of Statistics and Its Application*, **4**, 395-421

Ruffault J, Curt T, Martin-StPaul NK, Moron V, Trigo RM (2018) Extreme wildfire events are linked to global-change-type droughts in the northern Mediterranean. *Natural Hazards and Earth System Sciences*. **18**, 847-856.

Ruffault J, Moron V, Trigo RM, Curt T (2016) Objective identification of multiple large fire climatologies: An application to a Mediterranean ecosystem. *Environmental Research Letters*, **11**. doi:10.1088/1748-9326/11/7/075006.

Ruffault J, Mouillot F (2017) Contribution of human and biophysical factors to the spatial distribution of forest fire ignitions and large wildfires in a French Mediterranean region. *International Journal of Wildland Fire*, **26**, 498–508. doi:10.1071/WF16181.

Ruffault J, Mouillot F, Peters DPC (2015) How a new fire-suppression policy can abruptly reshape the fire-weather relationship. *Ecosphere*, **6**. 1–19. doi:10.1890/ES15-00182.1.

Serra L, Saez M, Mateu J, Varga D, Juan P, Díaz-Ávalos C, Rue H (2014) Spatio-temporal log-Gaussian Cox processes for modelling wildfire occurrence: The case of Catalonia, 1994-2008. *Environmental and Ecological Statistics*, **21**, 531–563. doi:10.1007/s10651-013-0267-y.

Silva JMN, Moreno MV, Le Page Y (2019) Spatiotemporal trends of area burnt in the Iberian Peninsula, 1975–2013. *Reg Environ Change,* **19**, 515–527. https://doi.org/10.1007/s10113-018-1415-6

Sutanto HT, Pramoedyo H, Wardhani WS, Astutik S (2021) The Selection of Bayesian Polynomial Regression with INLA by using DIC, WAIC and CPO, *Journal of Physics: Conference Series*. doi:10.1088/1742-6596/1747/1/012029.

Turco M, Von Hardenberg J, Aghakouchak A, Llasat MC, Provenzale A, Trigo RM (2017) On the key role of droughts in the dynamics of summer fires in Mediterranean Europe OPEN. doi:10.1038/s41598-017-00116-9.

Vidal JP, Martin E, Franchistéguy L, Baillon M, Soubeyroux JM (2010) A 50-year high-resolution atmospheric reanalysis over France with the Safran system. *International Journal of Climatology*, **30**, 1627–1644. doi:10.1002/JOC.2003.

Viedma O, Urbieta IR, Moreno JM (2018) Wildfires and the role of their drivers are changing over time in a large rural area of west-central Spain. *Scientific Reports*, **8**. doi:10.1038/s41598-018-36134-4.

Vilar, L., Gómez, I., Martínez-Vega, J., Echavarría, P., Riaño, D. and Martín, M.P. (2016) 'Multitemporal modelling of socio-economic wildfire drivers in central Spain between the 1980s and the 2000s: Comparing generalized linear models to machine learning algorithms', PLoS ONE, 11(8). doi:10.1371/journal.pone.0161344.

Vilar L, Woolford DG, Martell DL, Martn MP (2010) A model for predicting human-caused wildfire occurrence in the region of Madrid, Spain. *International Journal of Wildland Fire*, **19**, 325–337. doi:10.1071/WF09030.

Wang X, Wotton BM, Cantin AS, Parisien MA, Anderson K, Moore B, Flannigan MD (2017) cffdrs: an R package for the Canadian Forest Fire Danger Rating System, *Ecological Processes*, **6**, 1–11. doi:10.1186/S13717-017-0070-Z/FIGURES/6.

Woolford DG, Martell DL, McFayden CB, Evens J, Stacey A, Wotton BM, Boychuk D (2020) The development and implementation of a human-caused wildland fire occurrence prediction system for the province of Ontario, Canada.

2.2. Overview of the results

The main result from the article shows the importance of landscape factors to explain the spatial patterns of wildfire activities in Mediterranean France. The fire-weather explains a major part of the variance for larger fires. Nonetheless, the spatiotemporal effects are still responsible of a large part of the modelled variance, especially for the occurrence of fire larger than 1ha.

This study has shown how there is an overall reduction of fire numbers for all size categories between the decades of 1993-2002 and 2009-2018, with a strong tilting point after the year 2003. But this temporal evolution of fire activity is very distinct between the east and the west of the study region. While the eastern side and Corsica show a generalized reduction on fire numbers of all sizes, fire numbers increase on the western side. These changes are not explained by LULC or fire weather, instead they may be related to changes in fire management policies that followed the record-breaking fire season of 2003. And contrarily, if only fire-weather drove the temporal evolution of wildfire activity, there would have been an increase on fire numbers in the recent decade. Finally, no significant changes were found within the recent decade.

Altogether, this study has given further proof of the capacity of policy and societal changes to suddenly detach fire activities from climatic and non-climatic drivers. Previous research by Ruffault & Mouillot (2015) found this same effect on the French Mediterranean region. As a result, the fire records before 2003 are less relevant for modelling and prediction of the current fire danger. This factor is considered for the next chapter, where the focus is given to enhance the weather component of the model.

3. Enhancing the short-term prediction of fire danger through fire-weather indices and meteorological variables

Fire prediction based on weather conditions faces several limitations as stated in Chapter 1. Indeed, in the earlier study (Pimont et al. 2021), a single fire weather index has been used (CanFWI), which is also used by operational (as well as subindices). Among the diversity of fire weather indices available (Baijnath-Rodino et al., 2020), it remains unclear whether CanFWI is the best choice. For ex, a seasonal bias between the CanFWI and the fire activity has been identified. The aim of this third chapter is to find which fire-weather indices and weather variables describe the best the climatic fire factors by testing them inside the Firelihood framework. The hypothesis that drives this section of the work is that weather information enhances daily predictability compared to long-term averages, with some combinations performing better than others. Fire weather indices with memory effects are expected to increase the predictability of the models as they reflect the fuel moisture dynamics. Also, the multivariate combination of weather-based variables may be more effective than scalar fire-weather indices.

As seen in previous research (Woolford et al., 2021) this analysis has a high methodological complexity due to the variety of spatial and temporal fire factors. A large part of this chapter consists of the search for an evaluation method that will differentiate the performances between predictors. The model evaluation methods commonly used in this type of study found in the literature are not the most adequate (Phelps & Woolford, 2021). When the phenomena are very rare, the ratio of positive (presence) and negative (absence) observations is very low. This means that the rate of accurate negative predictions by the models is always very high. The variations on the quantity of accurate and false positive predictions does not have much weight on the final evaluated performance. The widely used AUC-ROC, for example, evaluates the capacity of a model to predict both the presence and the absence of the event. Nonetheless, the most relevant measures in fire activities prediction are the rate of real positive predictions and the rate of missed observations. Both measures focus on the positive observations and predictions.

The spatial and temporal effects (STEs) from the model are necessary to reach the best predictability possible, but they adapt to each configuration of predictor variables. This makes the comparison of performances between models not straightforward. A strategy to overcome this constraint is the definition of a baseline model to which all other models will be compared. This baseline model incorporates only the STEs, without any explicit fire-weather information. Following the results of the previous chapter, the study period is reduced to the more recent years (2004-2020) with the same fine spatial scale of 2km. Here the temporal scale is reduced to daily scale to capture the fine daily atmospheric variations through the season. Because the study explores the fire drivers, the climatic data is sourced from the reanalysis SAFRAN grid. Beware that for short term prediction the models should be refitted to reforecast data.

3.1. Article: A comparative analysis of fire-weather indices for enhanced fire activity prediction with probabilistic approaches

A comparative analysis of fire-weather indices for enhanced fire activity prediction with probabilistic approaches

Jorge Castel-Clavera*[a], François Pimont[a], Thomas Opitz[b], Julien Ruffault[a], Renaud Barbero[c], Denis Allard[b], Jean-Luc Dupuy[a]

*Corresponding author: jorge.castel-clavera@inrae.fr

[a] URFM, INRAE, Avignon, France

[b] BioSP, INRAE, Avignon, France

[c] RECOVER, INRAE, Aix-en-Provence, France

Abstract:

Background: Weather conditions play a crucial role in driving fire activity in Mediterranean France. Previous research has demonstrated the influence of these conditions on the likelihood of large fire events over the world. However, certain limitations persist regarding the representation of fire weather in probabilistic models.

Aim: The objective of this paper is to develop an efficient method to rate fire danger by identifying the best representation of weather data for fire activity prediction in Mediterranean France.

Methods: We evaluated the performance of meteorological variables and the most common fire-weather indices (FWIs) worldwide as predictors of fire occurrence and size using the *Firelihood* framework, a probabilistic Bayesian model of fire activity. These models were compared to a fire activity baseline model incorporating only spatial and temporal effects but no explicit fire-weather information to allow for an in-depth study of information not captured by fire-weather indices.

Key results: The results indicate that relying solely on fire-weather indices is insufficient for efficient rating of fire activities. The inclusion of spatial and seasonal effects in the models is crucial for improving the indices' performance. While the Canadian FWI remains the most skillful indicator among the tested indices, using new combinations of several of its subcomponents further increases accuracy. Various performance analyses, including threshold selections, were carried out to comparatively assess those improvements.

Conclusions: The approach shows that probabilistic models informed with appropriately constructed fire-weather indices substantially improve various aspects of fire activity predictions in the Mediterranean area.

Keywords:

Firelihood, fire-weather, spatio-temporal effects, probabilistic, Mediterranean France, fire management.

INTRODUCTION

Operational fire management involves making decisions throughout the fire season regarding the level of preparedness as well as the regional allocation and pre-positioning of surveillance and suppression resources (Taylor et al 2013; Martell 2015). Fire managers hence need reliable information on a daily to weekly basis about the anticipated characteristics of fire occurrence, including the number, size and duration of fires and their potential behavior (Taylor et al. 2013, Xi et al. 2019). The combination of these four traits characterize the fire danger within a given spatio-temporal unit.

Wildfire activity depends on multiscale interactions between climate, vegetation, topography, and humans (Moritz et al. 2005). Substantial evidence now supports the assertation that weather conditions play a major role in driving regional fire activity, affecting both the number and size of fires, across temporal scales ranging from daily to seasonal (Abatzoglou and Kolden 2013, Barbero et al. 2015, Turco et al. 2017). However, fire-climate relationships are not uniform globally, they exhibit significant spatial variations influenced by additional factors such as productivity gradients (Pausas and Ribeiro 2013), ignition patterns (Balch et al. 2017) or human suppression efforts (Pimont et al. 2021). Consequently, predicting fire activity requires context-specific assessment through regional studies (CFSFDG 2021).

Fire danger predictions rely heavily on weather-based operational indices, which are derived from all or a subset of meteorological variables typically including temperature, relative humidity, wind speed and precipitation (Baijnath-Rodino et al. 2020). These indices provide a quantitative assessment of atmospheric and fuel moisture conditions and are generically referred to as Fire Weather Indices (FWIs). They can be either punctual (i.e., based on current daily weather only), such as the Hot-Dry-Windy index (Srock et al. 2018), or with temporal memory, when they cumulate the effect of antecedent weather conditions, such as the Drought Code (DC), a component of the Canadian Fire Weather Index System (CFWIS; Van Wagner

1987). This system includes six indices among which the aforementioned DC and the well-known Fire Weather Index (later referred to as CanFWI). Other examples are the Australian McArthur index (FFDI; Noble et al. 1980), or the American National Fire Danger Ratings System (NFDRS; Deeming et al. 1977). Although developed and calibrated in North America, the CanFWI and its subcomponents are widely used in Europe (and across the world) both for research studies (Dupuy et al. 2020) and in the operational Copernicus system (https://effis.jrc.ec.europa.eu/about-effis/technical-background/fire-danger-forecast).

Despite their widespread use, selecting the most relevant FWIs for specific regions or adapting them to the local context for optimal performance is not straightforward, as FWIs are not physically measurable variables. FWIs were in fact designed as indicators of potential fire behavior (e.g., Sullivan 2017), yet large-scale data sets on the fire rate of spread or the fire intensity are often unavailable. For example, the CanFWI, one output of the CFWIS, has been shown to represent potential fire intensity through a monotonically increasing function linking CanFWI to actual fire intensity (in kW/m) in red pine stands in Ontario, Canada; however, the actual level of fire intensity varies 4 orders of magnitude from 10 to >10000kW/m depending on the local fuel type for the same CanFWI level (Taylor et al. 2022). Likewise, the widely used McArthur Fire Danger Meters (e.g., McArthur (1967, 1977)), later translated into arithmetic equations by Noble et al. (1980), have been designed to represent the fire rate of spread in Australian grasslands or forests (Cruz et al. 2015).

Despite this lack of direct validation, a solid degree of confidence has been achieved regarding the relevance of FWIs due to their strong relationship with large-scale fire activity observations, typically weekly to monthly burnt areas (Krawchuck et al. 2009, Barbero et al. 2014, Turco et al. 2018). However, aggregating climatic data over broad spatio-temporal scales may overlook extreme meteorological events, such as heatwaves or wind events, which are often linked to the occurrence of the largest wildfire events (Ruffault et al. 2020). Moreover, managers need

predictions at finer temporal and spatial resolutions for preparedness and early-warning systems. While refining scales is challenging because individual fire events are rare and stochastic, this unpredictability and associated uncertainty can be effectively addressed with a formalized probabilistic framework (Brillinger et al. 2003, Preisler et al. 2004, Preisler and Westerling 2007, Turner 2009, Vilar et al. 2010, Woolford et al. 2011, Serra et al. 2014ab, Nadeem et al. 2020, Coughlan et al. 2021, Pimont et al. 2021). Such frameworks translate FWIs and Land-Use Land Cover (LULC) variables into event probabilities or expected numbers of fires. They can hence be used as a fire occurrence and fire size components in fire danger rating systems, as the Fire Occurrence Prediction System of the next generation of CFFDRS (CFSFDG 2021), which implements such models (e.g., Woolford et al. 2021).

In Mediterranean France, probabilistic frameworks for fire activity have demonstrated effective predictive skills at spatio-temporal scales that meet operational needs, typically providing daily to weekly predictions over areas ranging from a few square kilometers to several tens of square kilometers (Pimont et al. 2021, Castel-Clavera et al. 2022). However, selecting the most relevant FWIs to inform these models and assessing their predictive performance remains highly challenging (Taylor et al. 2013, Phelps and Woolford 2021). This difficulty arises also from the rarity and stochasticity of fire events as well as the highly seasonally and spatially localized patterns of fire activity. Indeed, such patterns can usually be identified by multi-year aggregation of the fire activity data over a region, regardless of atmospheric variations. As a result, a model that would be solely based on long-term averages of fire activity may already present a decent skill even without considering day-to-day weather variations. This is because its effectiveness would largely arise from the typical fire patterns, and the potential added value of using a specific FWI to explain daily deviations from seasonal norms should be precisely assessed with respect to a spatial-seasonal baseline.

In this paper, we evaluate the performance of the most common FWIs in predicting fire activity in Mediterranean France, while acknowledging for the overwhelming role of the average spatio-temporal fire activity pattern described above. To achieve this, we developed a specific methodology to thoroughly evaluate and compare the models' skills using scoring rules adapted to probabilistic predictions. We further tested our approach for its ability to detect individual fire events and to outperform the CanFWI as a standalone indicator of fire activity, which is used operationally by fire managers in southern France. Ultimately, our study not only compares the performance of various FWIs for the prediction of fire activity, but also provides a detailed analysis of the added value of probabilistic models compared to traditional FWIs for the prediction of the spatiotemporal patterns of fire activity.

MATERIALS AND METHODS

1.1 *Method overview*

To evaluate the performance of FWIs in fire activity predicting, we used the *Firelihood* Bayesian probabilistic framework (Pimont et al. 2021, Castel-Clavera et al. 2022), which simulates daily fire activity in each grid cell of the French Mediterranean Region (see section 2.2.1) using a set of explanatory variables. *Firelihood* typically models the occurrence of fires larger than 1 ha (i.e., fires that escaped initial firefighting control; see details in Pimont et al. 2021) as a Poisson process and the size of these fires is modeled using a peaks-over-threshold (see section 2.3.1). We used a similar model structure to that of Preisler et al. (2004) who estimated the conditional probability of a burnt-area exceedance of 40.5ha given fire occurrence in a frequentist setting with generalized additive models. We on the other hand, estimated the number of fire occurrences and the probability of those exceeding the 80ha threshold in a Bayesian framework. A key aspect of *Firelihood* is its incorporation of spatio-temporal effects (STEs) to account for patterns unexplained by the available explanatory variables, as already shown in Preisler et al. (2004). The STEs include a random spatial effect of spatial coordinates

based on the Matérn covariance and a random seasonal effect of the week of year, to capture residual spatiotemporal variability of fire activity patterns (see section 2.3).

In this study, we incorporated a selection of the most relevant Fire Weather Indices (FWIs) and climate indices available worldwide (see the list and description in section 2.2.1) as explanatory variables in the *Firelihood* framework to evaluate their performance in predicting the number of fires larger than 1 ha and the probability that a fire exceeds 80 ha in size.

This work considers only fires that reaches a minimum of 1 ha (final burnt area > 1ha), because the Promethée data is more exhaustive after 1 ha, some small fire incidents might not be recorded with the correct location and/or size, or not recorded at all. In addition, in south-eastern France there is a strong surveillance and early attack policy. Therefore, we choose to model the occurrence of fires that could bring significant risks and exclude those that happened under conditions that were not suitable for combustion and fire expansion, or fires that were extinct by the intervention of first fire responders. Finally, the firefighting services have the operational objective of stopping fires before they reach 1ha. Therefore, there is a particular challenge in modeling fires larger than 1ha (i.e., escaped fires).

We evaluated 104 different models informed by FWIs and their combinations, both with and without the inclusion of STEs. Model performance in predicting fire activity was assessed and compared using a range of methods and metrics, following the guidelines of Phelps and Woolford (2021). These analyses were carried out at several spatio-temporal scales ranging from pixel-day to zone-week to verify the consistency of scoring, which can vary across scales (Parisien et al. 2011, Wornsop et al. 2021). To enhance the interpretability of our evaluations and quantify how much FWIs improve fire predictions and to explain daily deviations from seasonal norms, we compared their performance to a reference model called the Fire Activity Baseline (FAb). The FAb model is informed solely by STEs, which are functions of pixel

coordinates and the day of the year, and therefore simulate the mean seasonal fire activity for each grid cell.

Finally, we analyzed the performance of our best model for classifying fires into five operational danger classes, to simulate how such probabilistic predictions could be used for operational fire danger assessment. In these analyses, the CanFWI was used as a reference to evaluate potential improvements provided by probabilistic models for the prediction of fire events or for the discrimination of fire events across different levels of danger.

1.2 *Datasets*

1.2.1 Study area

We focused this study on 15 French *départements* (i.e., NUTS3-regions) located in the south-east of the country and covering a total area of 75,560 km^2. The region is characterized by a Mediterranean climate but encompasses a wide variety of terrains and biogeographic zones, from sea-side forests to high-elevation mountains in the Alps. Human settlements are scattered all over the area with higher population densities along the Rhone valley and the Mediterranean coast.

In addition to the administrative division in 15 units, we considered two other divisions: (I) Sylvo-Eco-Regions (SER), which correspond to areas of relatively homogeneous forest conditions with respect to the ecosystem functionality and forest management. (II) Homogeneous zones of fire-climate (ZFC) are a subdivision of the NUTS3 administrative units employed by forest managers to identify area where fire-weather is relatively uniform.

1.2.2 Fire-weather indices

Ten widely recognized daily FWIs were preselected for their relevance and applicability in the Mediterranean context (Table 1), based on an extensive review of FWIs and their usage conducted by Baijnath-Rodino et al. (2020). Although the National Fire Danger Rating System (NFDRS) (REF) meets these criteria, we excluded it from our analysis because its computation relies heavily on local factors, making it less directly applicable in Europe.

FWIs are calculated from several weather variables including temperature, relative humidity, wind speed and, less frequently, precipitation (Table 1). The CanFWI and its three primary components (DC, DMC and FFMC) are likely the most widely used FWI today and rely on all four weather variables. The Hot-Dry-Windy (HDW) incorporates temperature and relative humidity through the vapor pressure deficit (VPD). Note that the original index is calculated with the maximal surface adjusted VPD and U values at any layer of the first 500 vertical meters of the atmosphere but for this study we will only use the ground reanalysis weather data, and therefore call it surface HDW (sHDW). The Chandler Burning Index (CBI) is the only FWI that does not include wind speed. While each index combines those main variables in distinct ways, some of them (e.g., HDW and the Fire Danger Index, FDI) are mathematically very close. FWIs provide a daily rating of fire danger and can be classified in two main categories: *punctual,* which only consider the data of the day of interest, and *with-memory* which also integrates the information of the preceding days, weeks, or months. Precipitation is a factor commonly included in *with-memory* indices, as they usually consider the cumulated precipitation and/or the time since last rainfall to establish simple hydric balances.

Weather data was taken from the SAFRAN meteorological reanalysis available for France (Vidal et al. 2010); this database is provided in an 8km pixel grid at hourly scale. FWI were computed using the R packages *cffdrs* (Wang et al. 2017) and *firebehavioR* (Ziegler et al. 2019), and general-purpose functions from the R statistical software (R Core Team, 2021).

Table 1: List of fire-weather indices tested in this study. The column "Main Inputs" indicates the meteorological variables needed to obtain the subcomponents and eventually the final index, where RH = Relative Humidity, T = Temperature, U = Wind Speed, P = Precipitation. The column "Type" indicates the structure of the index: punctual (Pnt), or with-memory (M).

Index	Short name	Main inputs	Type	Author/s
Chandler Burning Index	CBI	RH, T	Pnt	Chandler et al, (1983)
Hot-Dry-Windy	HDW	RH, T, U	Pnt	Srock et al. (2018)
Fire Danger Index	FDI	RH, T, U	Pnt	Sharples et al. (2009)
Modified FDI	mFDI	RH, T, U, P	M	Sharples et al. (2009)
Forest Mark 5	MARK5	RH, T, U, P	M	McArthur (1967)
Fosberg Fire Weather Index	FFWI	RH, T, U	Pnt	Fosberg (1978)
Modified FFWI	mFFWI	RH, T, U, P	M	Goodrick (2002)
Canadian Fire Weather Index	CanFWI	RH, T, U, P	M	Van Wagner (1977)
Canadian Drought Code	DC	T, P	M	Van Wagner (1977)
Canadian Dough Moisture Code	DMC	RH, T, P	M	Van Wagner (1977)
Canadian Fine Fuel Moisture Code	FFMC	RH, T, U, P	M	Van Wagner (1977)

We produced two sets of indices: one at midday (12h), since it is the reference hour of weather observations used to construct the CanFWI; a second construction is based on selecting the maximum daily value from hourly data, which assures to detect the most fire-prone time of the day but comes with a higher computing cost. For the CanFWI, we also considered a third option: a construction of a daily value showing the best performance in the analyses of Bedia et al. (2014a), as it was also used in the first version of the Firelihood model (Pimont, et al. 2021); the construction is based on a combination of the daily mean, maximum and minimum measurements of variables to get the closest value possible to the reference of CanFWI. This last value considers the mean daily T and U, minimum RH and the 24-hours accumulated P.

1.2.3 Fire data

Fire records were extracted from the Prométhée fire database (database available online http://www.promethee.com/) for the period from 2004 to 2020 in warm-season conditions (25 May to 31 October). This period was selected as a trade-off, so that the dataset was large enough to allow fitting robust models with moderate estimation uncertainty, but that it did not include

fire data corresponding to some fire/weather relationships that would not be representative of the current situation. Indeed, the objective here is to develop prototype probabilistic models that are representative of current fire/weather conditions for operational services. Past studies revealed that the relationship between fire activity and weather has changed significantly over the past decades according to progress in fire management, first in the 1990's (Ruffault and Mouillot, 2015) and more recently, after the reorganization that followed the catastrophic fire season of 2003 (Pimont et al. 2021). For that reason, we discarded the pre-2003 period from the present analyses. This fire dataset (Figure 1) is given on a 2km pixel grid (DFCI; Défense des Forêts Contre les Incendies) and contains more than 13,000 fire outbreaks, of which 3541 burnt more than 1ha and 181 more than 80ha. The events are generally distributed all over the study area with all zones affected by wildfires, but several hot spots with relatively very high fire occurrence density can be identified. The causes of fire are not systematically present in the dataset specially for large fires, almost half of the dataset has an unknown cause; less than 5% of fires are caused by natural circumstances, the rest are human-caused fires including accident, arson, and negligence (Ganteaume and Jappiot., 2013). Therefore, we did not develop distinct models for human and natural causes, as often done in other regions (e.g., Woolford et al. 2021).

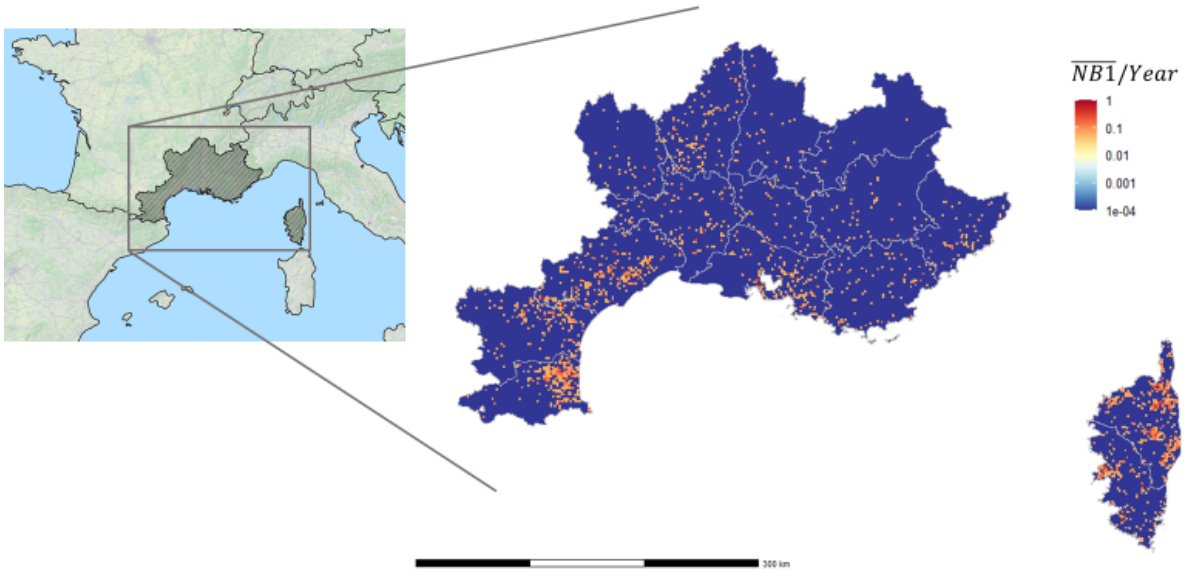


Figure 1: Summer 1ha fire observations on the 2km pixel grid

We joined the daily values of the various FWIs, calculated at the 8km resolution of the weather reanalysis, to the fire observations available at resolution given by the 2km operational grid "DFCI" of the Prométhée database using the closest centroid. Regarding fire occurrence, the counts with value zero corresponding to pixel-days with no fires must also be kept. In addition, the wildland area occupation per pixel (WA), obtained from Copernicus land cover data (https://land.copernicus.eu), is also included in the pixel's information following the same protocol as in Castel-Clavera et al. (2022).

1.3 *Development of probabilistic models based on FWIs*

1.3.1 The "Firelihood" modelling framework

We applied the Bayesian probabilistic framework of *Firelihood* for modelling fire occurrence, as well as the probability of exceedance for an appropriately chosen size threshold, as introduced in Pimont et al. (2021) with several extensions in Castel-Clavera et al. (2022). This approach leverages generalized additive modelling approaches as introduced in Brillinger et al. (2003) for fires. *Firelihood* represents wildfire activity as a marked point process, where points are stochastic and each point describes the occurrence position and time (i.e. pixel and day) of a fire with a minimal surface of 1ha, and to each point a binary mark (i.e., a Bernoulli random variable) is attached that indicates the exceedance of the fire above the size threshold of 80ha, with 1 and 0 indicating exceedance and non-exceedance, respectively. The exceedance indicators are modelled with the corresponding Bernoulli exceedance probabilities. Modelling is conducted in two steps by first modelling the occurrence component, and then the exceedance component by specifying the exceedance probability for each point, where the same or a different set of predictor variables and functional relationships between predictor and response variables can be used within the two components. The effects of the predictor variables are estimated separately for the two components through the R-INLA package in the R software, a tool that leverages a numerically fast implementation of the Integrated Nested Laplace

Approximation (Rue et al. 2017) to obtain accurate posterior estimations of the quantities of interest.

In both modelling steps, the variables incorporated into the predictor belong to one of the following four categories: (i) Land Use and Land Cover (LULC) variables, whose construction and predictive utility was investigated in Castel-Clavera et al. (2022), quantify the presence of different vegetation types and human modifications on the landscape. (ii) The WEATHER category contains the FWIs in Table 1 but also certain raw meteorological variables, and it further includes certain combinations of several of these variables. (iii) Spatial effects (denoted by XY) are designed to capture spatial variation not explained by the other "physical" variables and therefore modify the effect of other variables across space; spatial effects can exhibit spatial autocorrelation between values of adjacent spatial units (either pixels, or other spatial partitions defined *ad hoc*). (iv) Temporal effects (denoted by WEEK) capture seasonal variation not explained by other variables at weekly resolution and therefore correct the effect of other variables that vary throughout the fire season. We refer to the effects in (iii) and (iv) as space-time effects, denoted by STEs. For simplicity, the two space-time effects are always both included when estimated in a model, and they behave as independent additive variables without direct interaction. Given the relatively complex overall structure of our models, we have not explored more sophisticated STEs to keep model fitting and interpretation relatively simple and robust. They capture systematic variations of occurrence and exceedance structures not explained by other predictors. These residual effects do not have a direct physical interpretation, but we can study their estimates to look for the physical predictors that they could represent, and they are also useful for improving predictions.

Finally, a fixed deterministic space-time varying offset (to which we refer simply as ß, see equations 1 and 2) is added to all models on the first step, which consists of the surface of a pixel i on a day d covered by wildland (WA; in ha/pixel) divided by the average WA for all

pixel-day combinations. It expresses if a given pixel has more wildland area than the average (ß>1), average wildland area coverage (ß~1) or less wildland area than the average (ß<1).

$$ß = \frac{WA_{i,d}}{\overline{WA}} \qquad Eq.(1)$$

The first component of the overall model is the "occurrence component" and models the number of fires larger than 1ha, ($N_{i,d}^{1ha}$) in a pixel *i* on day *d* with a Poisson distribution, whose intensity parameter $\lambda_{i,d}^{1ha}$ (i.e., the expected number of fires) is constructed as a function of LULC, WEATHER and STEs (equation 2).

$$log\,\lambda_{i,d}^{1ha} = log ß_{i,d} + \sum_j f_{LULCj}\left(LULCj_{i,d}\right) + \sum_j f_{WEATHERj}\left(WEATHERj_{i,d}\right) + f_{X,Y}(X_i, Y_i) + f_{WEEK}\left(WEEK_{i,d}\right) \qquad Eq.(2)$$

where $N_{i,d}^{1ha} \sim Poisson\left(\lambda_{i,d}^{1ha}\right)$ is the number of fires larger than 1ha, and β is the offset of wildland-area-per-pixel (multiplicative at the scale of $\lambda_{i,d}^{1ha}$ but here log-transformed to match the log-link function of the Poisson intensity $\lambda_{i,d}^{1ha}$); the f components are functions of predictors that can be nonlinear, estimated during model calibration using data and representing effects of the predictor variables from the four above-defined categories: $f_{LULCj}\left(LULCj_{i,d}\right)$ for Land Use and Land Cover; $f_{WEATHERj}\left(WEATHERj_{i,d}\right)$ for Weather variables and FWIs; $f_{X,Y}(X_i, Y_i)$ with X_i, Y_i the pixel coordinates of fire count $N_{i,d}^{1ha}$ for the spatial effect consisting of a spatially structured multivariate Gaussian distribution based on the Matérn covariance function, where the so-called stochastic partial differential equation (SPDE) approach (Lindgren and Rue, 2015) is used to obtain numerically convenient representations; $f_{WEEK}\left(WEEK_{i,d}\right)$ for the seasonal effect encoded using the week number within the year. The main mathematical output of the model ($log\lambda_{i,d}^{1ha}$) is the natural logarithm of the Poisson intensity parameters for each pixel and day in the study domain.

Secondly, the fire size component of our model is specified in the same way as in Preisler et al (2004), with individual fires considered as conditionally independent given their occurrences.

We further mention Koh et al (2023) who propose a model implementing more complex interactions between spatial and temporal patterns of occurrences, moderate and extreme wildfires. The size component of the model (equation 3) expresses the probability $p_{i,d}^{u}$ of a given fire – in a pixel *i* and day *d* – to exceed a burnt area of *u* (value in ha), where the logistic link function is used to model the exceedance probability with the additive predictor:

$$log\frac{p_{i,d}^{u}}{1-p_{i,d}^{u}} = 1 + \sum_{j} f_{LULC,j}\left(LULCj_{i,d}\right) + \sum_{j} f_{WEATHER,j}\left(WEATHERj_{i,d}\right) + f_{WEEK}^{p,u}(WEEK) + f_{SER}^{p,u}(SER) \qquad Eq.(3)$$

where the f components are the non-linear effects of the predictor variables from different categories: $f_{LULC,j}\left(LULCj_{i,d}\right)$ for Land Use and Land Cover, and $f_{WEATHER,j}\left(WEATHERj_{i,d}\right)$ for Weather variables and FWIs. The STEs of the size component include a spatially-conditionally autoregressive $f_{BESAG}^{p,u}$ component based on the 28 Sylvo-Eco-Regions (SER) that encodes dependence between spatially adjacent SERs, and the seasonal component $f_{WEEK}^{p,u}$ is based on the week of the year as before. Here, we use a spatial effect with lower resolution than for the pixel-based variant in the occurrence model for two reasons: the exceedance indicators carry less information than occurrence counts, and they also tend to have more erratic behavior at small spatial scales. Therefore, a spatial model with larger spatial units provides better results in terms of numerical estimation stability and goodness-of-fit, as already found by Castel-Clavera et al. (2022). The mathematical output of the size component is the logit-transformation of the probability $p_{i,d}^{u}$. Using the estimated additive effects and inverting the transformation, we obtain the estimated probabilities $p_{i,d}^{u}$. We can then multiply them with the Poisson intensities $\lambda_{i,d}$ to obtain $p_{i,d}^{u}\lambda_{i,d}$, the expected number of fires $N_{i,d}^{u}$ with size larger than the threshold *u* for a given pixel-day. The size threshold we use in this study is set to u=80 (in ha unit), as this value was identified by Koh et al. (2023) (using statistical extreme-value theory) as the size where the distribution of fires shifts from a power law to the theoretically grounded

generalized Pareto distribution in the studied area. For more detailed information on the *Firelihood* framework, we refer to Pimont et al. (2021) and Castel-Clavera et al. (2022).

Regarding the interpretation of the predictor effects included in the occurrence component, we can invert the log-link function to obtain a partial multiplicative effect $exp\left(f_{VAR}(VAR)\right)$ for each predictor variable VAR. These partial effects are the response functions (or curves) of single predictors when holding the other predictor variables constant. The response of the spatial SPDE effect is not represented as a curve but as a value/pixel, i.e., as a spatial map. The additive effects of the predictor variables in the size component represent the contribution of the value of a variable to the logit of the probability of exceedance of the size threshold. We also represent them with a response curve by directly reporting the estimated curves $f_{VAR}(VAR)$ at the additive scale; note that the interpretation of partial effects after inverting the link function would be more intricate with the logistic link since there would be interaction with the values of the other predictor variables. The response of the spatial Besag effect is represented as a value/SER, again leading to a spatial map.

1.3.2 Set of fitted models based on different Fire Weather Indices

We used the framework described above to study the influence of fire weather on occurrence (Eq. 2) and size (Eq. 3). Probabilistic models with different structures of the additive predictors were developed for this purpose (Table 2). Two relatively simple model configurations without weather-based predictors were fitted and serve as baselines to evaluate the predictive performance of probabilistic models informed with fire-weather indices (FWIs). First, we define a "Null" model in which no predictor effect is estimated and only the wildland area offset ß is present; second, the "Fire activity baseline" (FAb) is defined as a model whose only information are the space-time effects (STEs); it therefore captures only the average seasonal and spatial fire activity patterns, without considering the actual weather observed for a given pixel and day.

We fitted 51 probabilistic models informed with FWIs (Weather-aware models), each of them was estimated with and without STEs to evaluate the role played by STEs in the overall estimation. 104 models were tested in the present study, developed by combining all general model structures described in Table 2.

Table 2: Summary of predictor structures in the probabilistic models used in this study, based on combinations of terms for the occurrence component (eq. 2) and for the size component (eq.3). The "Null" model is a simple reference in which only the presence of vegetation is accounted for; the second reference model is the fire activity baseline FAb informed only by including spatial and seasonal effects (STEs) to the "Null" model. The models informed by FWIs are considered with but also without STEs.

Fitting option	Occurrence 1ha (Eq2)	Size 80ha (Eq3)
Null	ß	1
FAb	$ß + (X_i, Y_i) + WEEK$	$1 + SER + WEEK$
Weather-aware Model	$ß + WA + WEATHER$	$1 + WA + WEATHER$
Weather-aware Model + STEs	$ß + WA + WEATHER + (X_i, Y_i) + WEEK$	$1 + WA + WEATHER + SER + f(WEEK)$

1.4 *Analyses of probabilistic models and fire-weather indices*

1.4.1 Probabilistic model verification and scoring

As explained in the overview, we calculated several complementary metrics to assess and compare the models. First, we verified model calibration using reliability plots to assess the agreement between predicted probabilities and observed frequencies. Second, we checked model parsimony by balancing the model's complexity against its predictive performance. Third, we calculated four scoring metrics to select the models based on their overall performance. Fourth we also evaluated the performance of the most relevant models in detecting observed fire events using Precision-Recall (PR) curves and the F1-score. These metrics are particularly suited to assess both the accuracy and the detection capability of rare events like wildfires. Fifth, we verified that the response functions of the different climatic indices estimated in the models satisfied monotonic behavior as expected and were plausible to

provide models with reliable behaviors. For a deeper analysis of the most relevant models, we analyzed to what extent the spatio-temporal patterns of fire activity were captured by the FWIs, by comparing the STEs of these models with the ones estimated in the FAb model.

Reliability plots

A first simple assessment of model accuracy and sanity check for models consists in checking that event intensities or probabilities estimated by the models are well adjusted to empirical event frequencies in fire observations. This can be addressed with "reliability plots". They compare predicted and empirical fire counts based on binning values of predicted intensities or probabilities into a moderate number of intervals. Examples of reliability plots are shown in Supplementary Material 1. Visual inspection of the models selected for later comparisons showed satisfactory performance since they appropriately reproduce the observed fire probability distribution; the results are shown in the supplementary material.

Predictive performance assessment using scoring functions

When selecting a model, a principle of parsimony must be favored, which refers to the model's simplicity in relation to its ability to accurately describe the data. To prevent overfitting, which occurs when noise is captured instead of the signal arising from patterns in data, the predictive performance of each model must be balanced against its complexity. Therefore, our second step consists in computing penalized log-likelihoods and prediction scores for each model to check whether it is capable to generalize to new data, i.e., to provide useful predictions for data not used during model calibration. For penalizing the log-likelihood, we here use the DIC (Deviance Information Criterion), a generalization of the well-known Akaike Information Criterion for Bayesian models.

To evaluate the predictive performance of the models, we will use the concept of 'proper scoring rules' (Gneiting and Raftery, 2007, Allard et al. 2012). A scoring rule is a function that assigns a numerical value (a score) based on two elements: the predicted probability and the

event that is observed. It measures the deviation between predicted probabilities and the observed values: smaller values indicate better predictions. For each spatiotemporal unit, a score value can be computed, and then an overall score of the prediction is obtained by averaging the individual score values over all units. Among all possible scoring rules, a proper scoring rule is such that the computed scores are lowest when the evaluated probability distribution is equal to the true (but unknown) distribution. Here, we will use the Log-loss score, known to be a proper score (Gneiting and Raftery, 2007), more detailed information in Supplementary Material 2.

The absolute values of these scores have no universal and easily interpretable lower or upper thresholds. This stands in contrast with the Area Under the Receiver-Operating-Characteristic Curve method (AUC-ROC) that rates the model skill with a score between 0.5 and 1, where a value of 0.5 is associated with prediction completely at random and 1 with perfect prediction. To facilitate comparison of a score to a baseline score and measure improvement of a model with respect to the baseline, it is common practice to compute Relative Skill-Scores, as in eq. 4 for the log-loss score. Due to marked seasonal and regional patterns in fire activity and our aim of assessing improvements provided by daily FWIs based on actual weather, we used the score of the model corresponding to the fire activity baseline (FAb model):

$$Skill_{Log-loss} = 1 - \frac{LogLoss_i}{LogLoss_{FAb}} \qquad Eq.4$$

Spatio-temporal scale of scoring

Such prediction scores can be calculated at pixel-day scale corresponding to the spatio-temporal resolution used for estimating the models. Due to the stochastic behavior of fire occurrences and sizes at relatively fine spatial scales, achieving good predictive scores at the pixel resolution of the modelling framework is challenging, as shown in previous work (Pimont al. 2021). Moreover, predicting a relatively high probability of fire in a pixel adjacent to the one where a

fire is observed would still be useful for forest managers and authorities, but this situation would be considered as a bad prediction by a pixel-based prediction score. For example, Bedia et al. (2014) found a lack of correlation between the high skill of a model in terms of Relative Sensitivity Analysis (AUC-ROC) and its capability of predicting and reproducing fire observations. Therefore, we also consider other relevant, coarser spatial scales (pixel, ZFC, NUTS3, whole study area), as well as the weekly temporal scale which is relevant for preparedness. We assess the sensitivity of the skill scores and the ranking of models to spatiotemporal scale to determine the scales providing useful information for fire-danger prediction.

1.4.2 Analysis of the response functions of the different models

The response functions obtained from the models represent the partial effect of each of the predictor variables and provide useful interpretation, but they should also be checked for generally plausible behavior. A useful FWI should give a monotonic effect through all its values and avoid saturations or abrupt changes in its slope. Therefore, studying response functions of FWIs allows us to understand the range of values for which the index provides sensible information.

The response functions of the STEs are also shown. When we include the STEs in a model that also incorporates useful weather information, we expect that STEs have generally less influence on the predictions in comparison to the STEs estimated in the FAb model, i.e., we expect them to have a smaller amplitude. The estimations of the STEs could also change in time or in space when an FWI is included, which informs us how this FWI is spatio-temporally biased.

1.4.3 Comparison of probabilistic models and CanFWI to predict single fire events

To go deeper in the analysis of model predictions, we carried out specific analyses with two probabilistic models of fire activity: one informed by the CanFWI (by analogy with a previous study, Pimont et al. 2021), and one with the best model obtained in the present study (revealed

in results section 3.1 *Models' scoring and selection*) according to the above criteria (see *2.4.1 Probabilistic model verification and scoring).* The approach was then also applied to the CanFWI itself – i.e., not incorporated into a probabilistic model – and provides a reference value to highlight the potential added value of a probabilistic model for predictions.

Predictions of *Firelihood* models and of the raw CanFWI are comparatively assessed through Precision-recall curves (PRC) and F-measures. These metrics are derived from the well-known confusion matrix (Table 3) by confronting the observations with the predictions. The chosen metrics are the Hit Rate (HR, also known as *recall*) to measure how much of the observed fires were correctly detected (i.e., correctly predicted), and the Precision (Pr) to measure how much of the detected fires correspond to observation of a real fire; both of those metrics vary between 0 (worst possible prediction) and 1 (best possible prediction). Based on the notations introduced in Table 3, the metrics were formulated as shown in equations 5 and 6. A numerical not necessarily binary prediction X_i is classified as a detection when the value of this prediction X_i exceeds a fixed threshold for the values of all predictions $\left[X \middle| 1, X_{\max(i)}\right]$. A confusion matrix is created for each of a set of possible thresholds. In this study, we used 1000 threshold values in the interval $\left[X \middle| 1, X_{\max\,(i)}\right]$, and so we obtain 1000 consecutive values of HR and Pr. The precision will always increase as the threshold increases, at the expense of the hit rate.

The two metrics (HR and Pr) calculated for different thresholds are then combined to build (i) the Precision-Recall curve (PRC): the better a model predicts, the larger will be the area under this curve; and (ii) the F_β-measure (equation 7), the harmonic mean of HR and Pr, where the β-coefficient (usually set to 0.5, 1 or 2) is the weight given to HR, such that HR is β times as important as Pr; we aim for the highest value of this harmonic mean. Higher β values attribute more importance to the prediction of observed events but increase the rate of false alarms. Results obtained with different β values are not comparable.

$$HR = \frac{\sum H}{\sum (H + M)} \quad Eq.5 \qquad\qquad Pr = \frac{\sum H}{\sum (H + FA)} \quad Eq.6$$

Table 3: Example of a confusion matrix and the classification of the predictions into the four categories.

Predictions / Observations	Fire (positive observation)	No fire (negative observation)
Fire (positive prediction)	HIT (H)	FALSE ALARM (FA)
No fire (negative prediction)	MISS (M)	TRUE NEGATIVE (TN)

$$F_ß = (1 + ß^2)\frac{Pr \times HR}{(ß^2 \times Pr) + HR} \qquad Eq.7$$

As explained above, these metrics were calculated and then compared for the CanFWI (reference), the probabilistic model informed by the CanFWI, and our best probabilistic model selected according to criteria developed in previous sections 2.4.1. and 2.4.2.

1.4.4 Fire danger rating with models and fire-weather indices

A limitation of the metrics presented in section 2.4.2 is that they are oriented toward the detection of single fire events. In practice, we can discretize these metrics into categories for operational purposes. Here an example is provided with five classes of fire danger (ranging from 1 – the lowest danger, to 5 – the highest danger) attributed to homogeneous Zone of Fire Climate at a daily scale ("ZFC-day"), with breaks set at CanFWI values of 10, 30, 50 and 80, corresponding to the thresholds used by the French national meteorological agency (MeteoFrance) to stablish fire danger at national level (thresholds provided for this study by experts at Office National de Forêts – French National Forest Office – but are not published in scientific literature nor are they available in any official communication report or on a website). Table 4 shows the frequency and percentile at which each class was attributed for the past summer seasons; the lowest class was attributed to 38% of ZFC-days and the highest class to only 0.3% of ZFC-days. A very simple way to evaluate this approach is to compare the rate of observed fires (> 1 ha et > 80 ha) in each class of danger.

To go beyond the evaluation of prediction of single events (section 2.4.3), we tested how probabilistic models perform in rating days and zones in a similar manner to operational

thresholds of CanFWI. For that purpose, we defined five categories of ZFCs-days based on the outputs of probabilistic models. To ensure a fair comparison between approaches, the thresholds on model-predicted values required to define these discrete categories were chosen to obtain the same frequency (in terms of the number of zones-days) as in the operational danger classes. We aggregated outputs of fire activity models (>1ha and > 80ha) at the ZFC-day scales to estimate the expectation of fire numbers in each ZFC-day according to the model. We then rated the ZFC-day according to breaks corresponding to the same percentiles as the operational fire danger classes (Table 4). We then used this approach to compute the rate of observed fires (> 1 ha et > 80 ha) in each of the classes of danger based on expected fire numbers estimated from probabilistic models.

Table 4: Classification of ZFC-days in five danger categories. Where Danger class is the classification category; CanFWI intervals show the breaks used for the classification; N Days indicates the number of ZFC-days classified in each category; Frequency is the percentage of ZFC-days in each danger category; Percentile is the cumulated Frequency.

Danger class	CanFWI intervals	ZFC-Days	Frequency (%)	Percentile
Low	FWI<10	120214	38,1	38,1
Medium	10<FWI<30	122384	38,8	76,9
High	30<FWI<50	60522	19,2	96,1
Very high	50<FWI<80	11617	3,68	99,8
Extreme	FWI>80	935	0,22	100

These rates, calculated for model-based predictions, can be compared with those obtained with the CanFWI. This approach provides a simple way to measure the improvement that can be expected for operational use of probabilistic models instead of fire weather indices.

RESULTS

1.5 *Models' scoring and selection.*

Figure 2 summarizes the performance of the occurrence and size components of the different models according to the log-loss skill score. For easier interpretation of results, the skill values have been renormalized to a scale called the skill-scale, where the FAb has value 0 and the best

model is assigned a value of 100. Models that are worse than the baseline can show unbounded negative values of the skill-scale. The most striking result is that all models including spatio-temporal effects (STEs) obtained systematically higher scores than the FAb model (skill score >0) and higher than the corresponding models without STEs, which often even showed lower scores than the FAb (score<0). For the size component, however, several models without STEs outpaced the FAb, suggesting that the probability of a fire to exceed 80ha is less dependent on the location of the fire and seasonality than the fire occurrence. Another important finding is that for both occurrence and size predictions, the null model (including only forest area) has value close to -200 on the skill scale. This means that the spatio-temporal patterns associated to the FAb capture around 2/3 of the overall spatio-temporal variability of observed fire activity, the corollary being that only an additional 1/3 can be further explained by daily FWIs. This result confirms the very pronounced seasonal and spatial patterns associated with fire activity in the Mediterranean France.

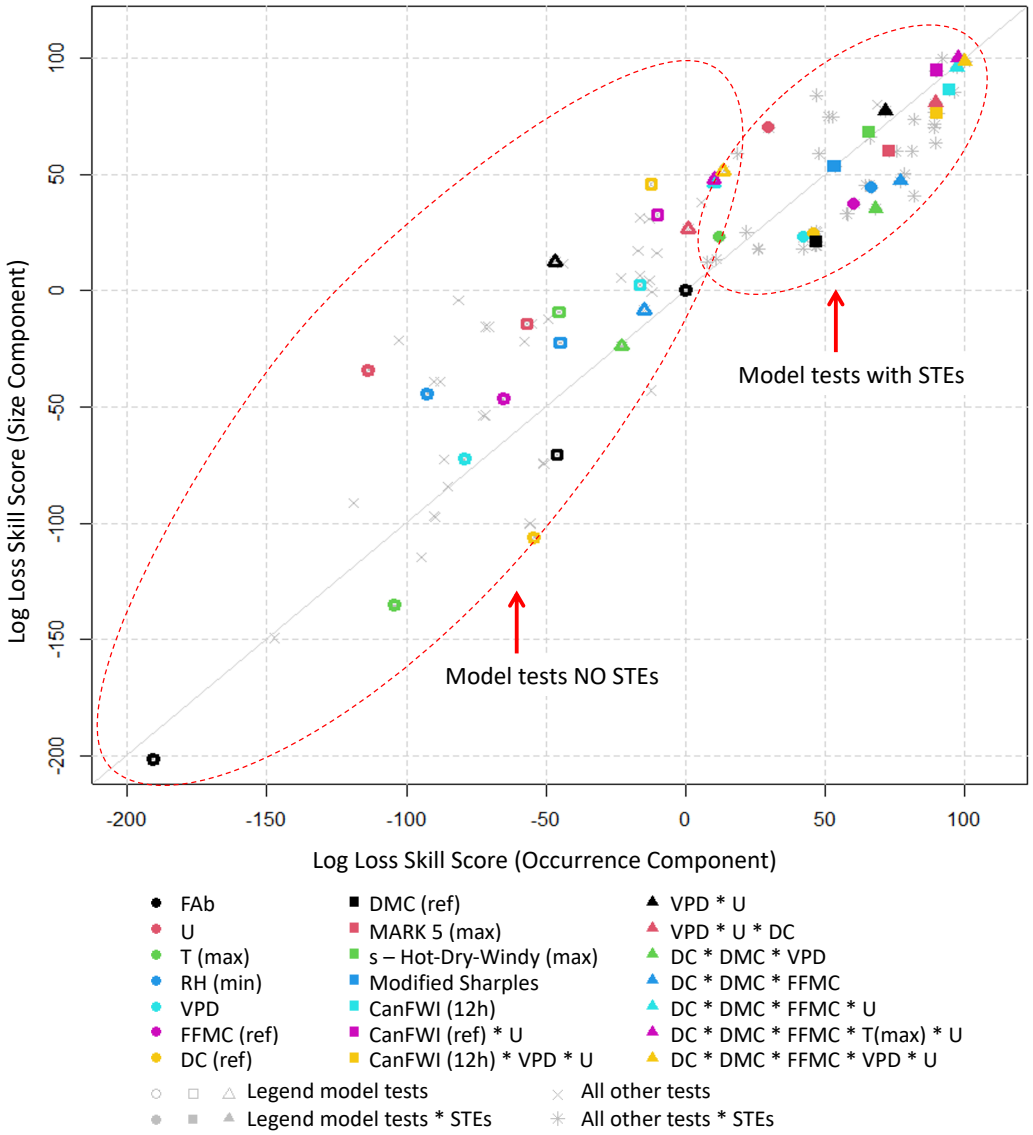


Figure 2:Comparison of the prediction scores on skill scale for the occurrence component (X-axis: Occurrence skill scale) and the size component (Y-axis: Size skill scale) for all models considered. A value of zero corresponds to the FAb; negative values arise if the model performance is lower than the baseline; positive values indicate models with better performance than this baseline. Empty shapes indicate the models tested without space-time effects (STEs)s, while full shapes indicate the same scores but with STEs. Dotted red ovals encompass the spaces of distribution of the scores for models with or without STEs. Grey crosses and stars correspond to other models not referenced in the legend for brevity; their scores are shown in supplementary materials 3. Difference in shapes (rounds, squares, and triangles) were chosen only for the purpose of visualization and a reduced number of different colors. Abbreviations used in the figure legend are as follows: FAb – Fire Activity baseline; U – Wind speed; T – Daily temperature; RH – Relative humidity; VPD – Vapor Pressure Deficit; FFMC – Fine Fuel Moisture Code from the CanFWI system; MARK5 – Forest Mark 5 index; DC – Drought Code from the CanFWI

system; DMC – Dough Moisture Code from the CanFWI system; CanFWI – van Wagner's Canadian Fire Weather Index; "s – Hot – Dry – Windy" – surface Hot Dry Windy index. Also: max –maximal value per day; min – minimal value per day; daily – calculated using Bedia et al. (2014a); 12h – value at noon.

The CanFWI model performed the best among the models based on an index, for both components, and with or without the STEs included (Tables 5 and 6). In addition, a combination of VPD, wind speed and drought code (DC) from CanFWI yielded a very good result. Nevertheless, the best performance in both components was brought by complex combinations of DC*DMC*FFMC*VPD (or Tmax) *U*STEs. Compared to these best performing models which, by definition, have a value on skill scale equal to 100, the CanFWI for the size component (Table 6) provided only an improvement of 3.6 scale units over the FAb model, but an improvement of 70 units when the STEs were included and 95 with both wind and STEs. For the occurrence component (Table 5), including the STEs into the CanFWI model provided an 88 units improvement over the FAb model.

Table 5: Model evaluation (at ZFC/day scale) metrics for the 1ha occurrence component. Names in the table are as follows: LogLoss is the raw output from the score formula, Skill $_{Log\text{-}Loss}$ is the Log Loss skill score based on the FAb model, and Skill Scale Value is the normalization of the Skill Score values between the 0 and the maximal performance value, AUC is the value from an AUC-ROC based on the model fit. In the models, max indicates the maximum daily value, daily indicates the version from Bedia et al. (2014a), min indicates the minimum daily value, and no indication corresponds to the mean daily value. The yellow row corresponds to the FAb. The table is sorted using the occurrence's "Skill Scale Values". All evaluated models are shown in table sup.1 of supplementary materials 3.

Occurrence (1ha)					
Test	**LogLoss**	**Skill$_{Log\text{-}Loss}$**	**Skill Scale Value**	**AUC**	**DIC**
NULL	0,0455	-0,1841	-1,907	0,571	82500
MARK5 max	0,0405	-0,0547	-0,567	0,761	72700
VPD * U	0,0401	-0,0451	-0,468	0,764	72200
Surface HotDryWindy max	0,0401	-0,0438	-0,454	0,769	72300
Modified Sharples	0,0401	-0,0434	-0,450	0,768	72400
CanFWI (ref)	0,0389	-0,0134	-0,139	0,803	71300
CanFWI12h * VPD * U	0,0389	-0,0120	-0,124	0,802	71200
Climatic fire activity baseline	0,0384	0,0000	0,000	0,816	71100
VPD * U * DC	0,0384	0,0008	0,009	0,812	70700
DC * DMC * FFMC * U	0,0380	0,0098	0,102	0,823	70400
DC * DMC * FFMC * VPD * U	0,0379	0,0131	0,136	0,825	70300
Modified Sharples * STeffects	0,0364	0,0514	0,533	0,853	67400
Surface HotDryWindy max * STeffects	0,0360	0,0635	0,658	0,858	67000
VPD * U * STeffects	0,0358	0,0690	0,715	0,860	66900
MARK5 max * STeffects	0,0357	0,0703	0,728	0,866	66900
CanFWI (ref) * STeffects	0,0351	0,0857	0,888	0,874	66400
VPD * U * DC * STeffects	0,0351	0,0863	0,894	0,873	66200
CanFWI12h * VPD * U * STeffects	0,0350	0,0886	0,918	0,875	65900
DC * DMC * FFMC * U * STeffects	0,0348	0,0938	0,972	0,877	65800
DC * DMC * FFMC * VPD * U * STeffects	0,0347	0,0965	1,000	0,879	65600

Table 6: Model evaluation (at pixel/day scale) metrics for the 80ha size component Names in the table are as follows: LogLoss is the raw output from the score formula, Skill $_{Log\text{-}Loss}$ is the Log Loss skill score based on the FAb model, and Skill Scale Value is the normalization of the Skill Score values between the 0 and the maximal performance value, AUC is the value from an AUC-ROC based on the model fit. In the models, max indicates the maximum daily value, daily indicates the version from Bedia et al. (2014a), min indicates the minimum daily value, and no indication corresponds to the mean daily value. The yellow row corresponds to the FAb. The table is sorted using the SIZE's "Skill Scale Value" column. All evaluated models are shown in table sup.2 of supplementary materials 3.

80ha threshold					
Test	**LogLoss**	**Skill$_{Log\text{-}Loss}$**	**Skill Scale Value**	**AUC**	**DIC**
NULL	0,203	-0,292	-2,013	1,000	1428
Modified Sharples	0,162	-0,032	-0,224	0,732	1344
MARK5 max	0,160	-0,020	-0,140	0,748	1327
Surface HotDryWindy max	0,159	-0,013	-0,091	0,751	1326
Climatic fire activity baseline	0,157	0,000	0,000	0,694	1383
FWI (ref)	0,156	0,005	0,036	0,764	1317
VPD * U	0,154	0,018	0,125	0,749	1331
VPD * U * DC	0,151	0,039	0,266	0,755	1325
FWI12h * VPD * U	0,147	0,066	0,458	0,765	1298
DC * DMC * FFMC * U	0,147	0,067	0,463	0,765	1307
DC * DMC * FFMC * Tmax * U	0,146	0,070	0,480	0,770	1306
Modified Sharples * STeffects	0,145	0,077	0,532	0,755	1327
MARK5 max * STeffects	0,144	0,087	0,600	0,764	1313
Surface HotDryWindy max * STeffects	0,142	0,099	0,682	0,771	1309
FWI (ref) * STeffects	0,141	0,102	0,701	0,778	1302
VPD * U * STeffects	0,140	0,112	0,773	0,775	1310
VPD * U * DC * STeffects	0,139	0,117	0,807	0,779	1305
DC * DMC * FFMC * U * STeffects	0,135	0,139	0,958	0,790	1290
FWI12h * VPD * U * STeffects	0,134	0,145	0,999	0,791	1281
DC * DMC * FFMC * Tmax * U * STeffects	0,134	0,145	1,000	0,795	1287

Both scatter plots in Figure 2 are stretched approximately along the 45° diagonal line, revealing that the performance of the models for the two components are fairly well correlated. However, the hierarchy of certain indices and predictor variables changed significantly between occurrence and size components in terms of the Skill Score, e.g., U or RH min were better positioned for the size (Table 6) than the occurrence (Table 5).

1.6 *Response functions to indices and analyses of spatio-temporal effects*

As last step in our model analysis, we inspected the shape of the response functions and maps of the main effects to ensure that the models we recommend are comprehensive and consistent (examples shown in Figure 3). The CanFWI exhibited the largest range for the multiplicative effect among all FWIs (Figure 3), which is consistent with its high performance (section 3.1). Several indices behaved similarly (e.g., daily CanFWI and sHDW max), presenting a nonlinear effect with slower increases in the lower range and faster increases at higher values. The MARK5, KBDI (not shown) and Sharples (not shown) indices presented saturations above some point of the distribution of values, suggesting calibration issues in a Mediterranean climate. However, none of the indices showed abnormal behavior, such as a declining effect.

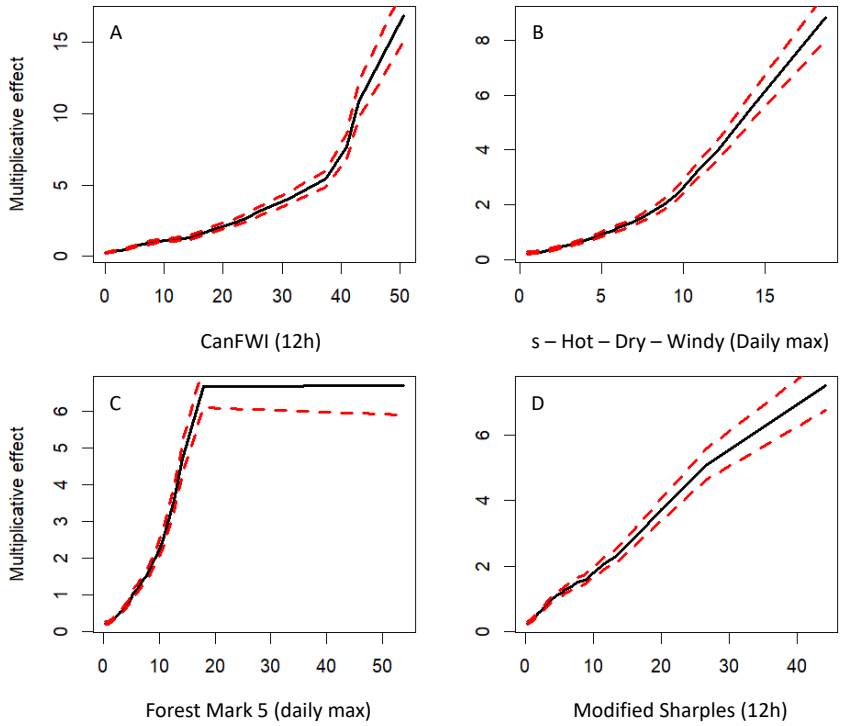


Figure 3: Effects for weather indices (including spatio-temporal effects) fitted in the occurrence model tests. The X axis shows the value in each of the variable-specific bins; the Y axis is the value of the multiplicative effect, in other words, the factor by which the expected number of fires changes with respect to the case of a multiplicative effect of 1.

The best occurrence model (see section 3.1) was obtained with the daily components of the CanFWI (DC, DMC and FFMC) combined with U and VPD. The response of all those variables when combined in this model (Fig. 4) maintains a plausible shape similar to its response when the variable was used alone. The response to wind speed U is the strongest and then other variables show similar ranges of multiplicative effect. The DMC response reveals saturation at high values suggesting some calibration issue under the south-east France climate. The FFMC has a relatively flat response at lower values and shows a steep increase above 80, revealing high sensitivity when fine fuels get drier.

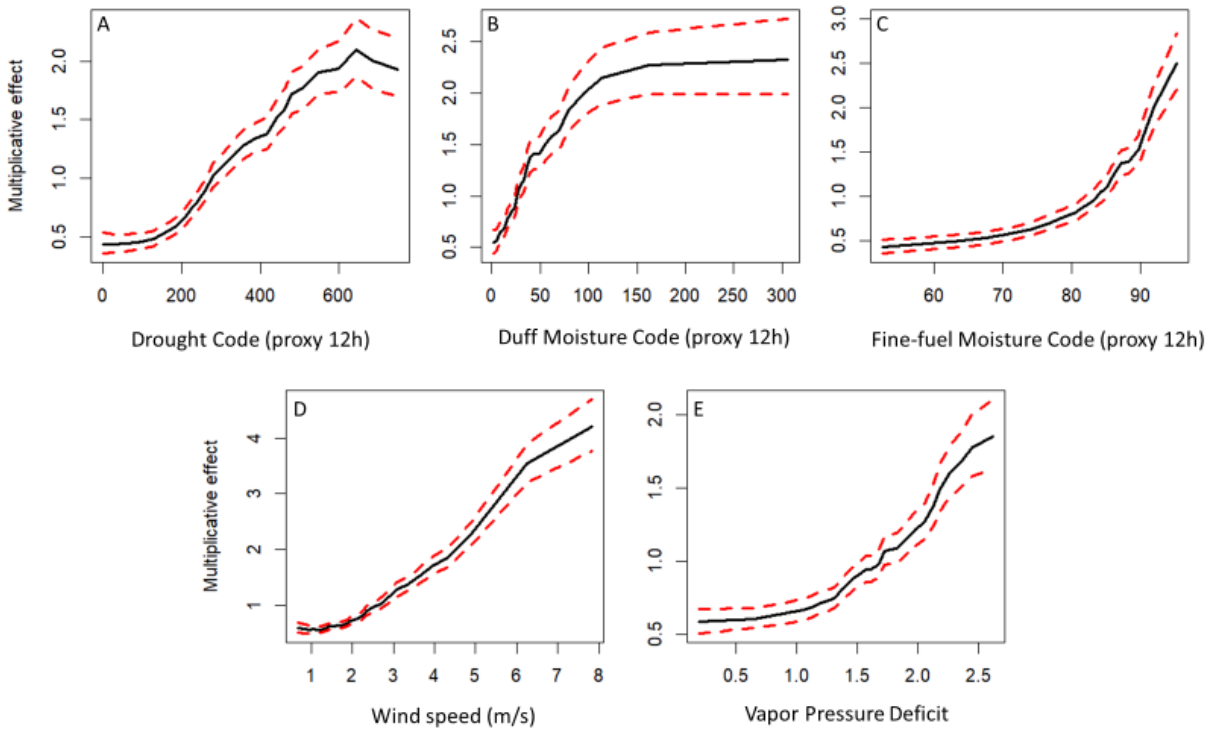


Figure 4: Effects of the occurrence component fitted with DC, DMC and FFMC daily values from the CanFWI (A-C), Wind speed (D) and Vapor pressure deficit (E). The X axis shows the value in each of the variable-specific bins; the Y axis is the value of the multiplicative effect, in other words, the factor by which the expected number of fires changes with respect to the case of a multiplicative effect of 1.

The estimated effects were much weaker for the size (Figure 5) than for the occurrence, with slightly larger uncertainties, but they keep the same trend (with VPD replaced by air temperature in the best-performing size model). The VPD (Fig.4 E) had an effect similar to

FFMC, consistent with the fact that both are related to the short-term moisture content. Temperature in the size component (Fig. 5 E) exhibited a negative slope but with very low magnitude and relatively high estimation uncertainty.

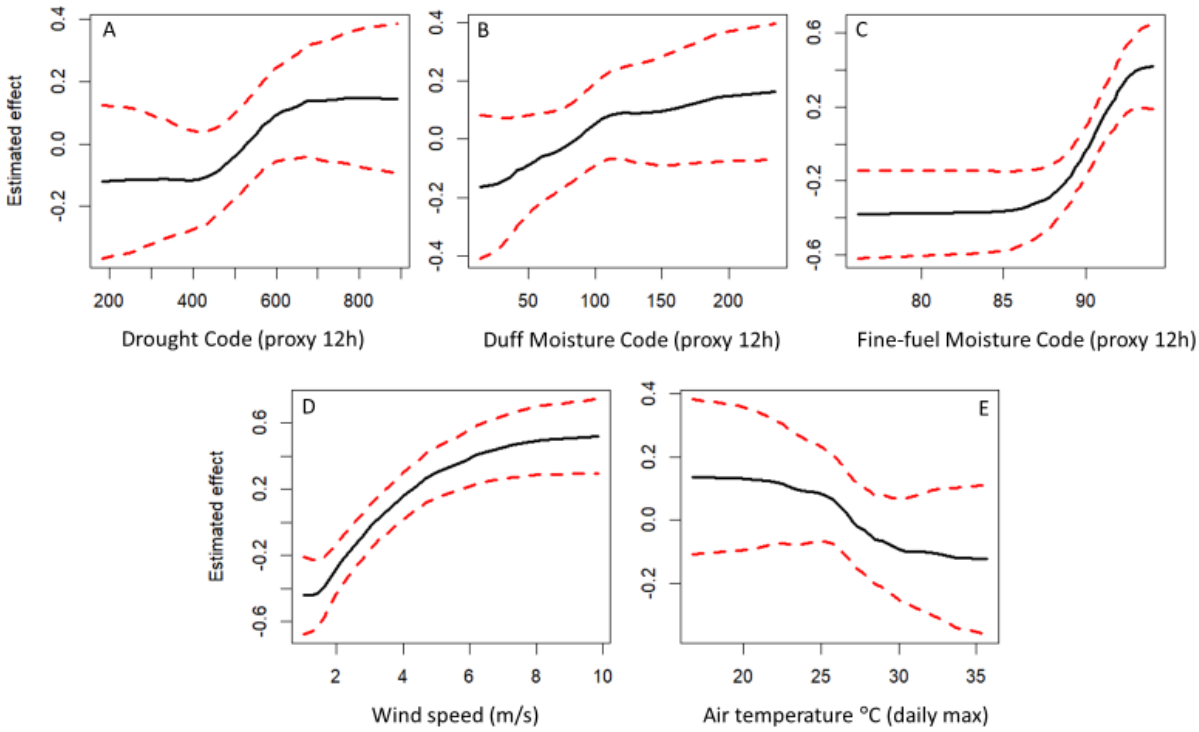


Figure 5: Effects of the size component fitted with DC, DMC and FFMC daily values from the CanFWI (A-C), the Wind speed (D) and the maximum daily temperature (E). The X axis shows the value in each variable-specific bin; the Y axis is the contribution of the predictor variable's value to the logit of the probability of exceedance.

Owing to the critical role played by STEs in model scoring and the importance of the FAb in fire activity patterns, we focused below on the FAb and how corresponding STEs were modified in the model by the inclusion of FWIs.

For the occurrence component, the seasonal effect for the FAb model (Fig. 6 A) captured the seasonal cycle of fire activity with a peak at the end of July. Most of this pattern was explained by the combination of effects of the subcomponents of the CanFWI (Fig. 6 B). One can even notice a small decay in the beginning of August, meaning that fire weather information raised slightly higher danger than present in that period of the season, which might be explained by more intense fire prevention and fighting during the core of the fire season. The overall seasonal

effect for the size (Fig. 6 C) resembles the corresponding seasonal effect for occurrences, but with a weaker – and almost non-significant – response (Fig. 6 D).

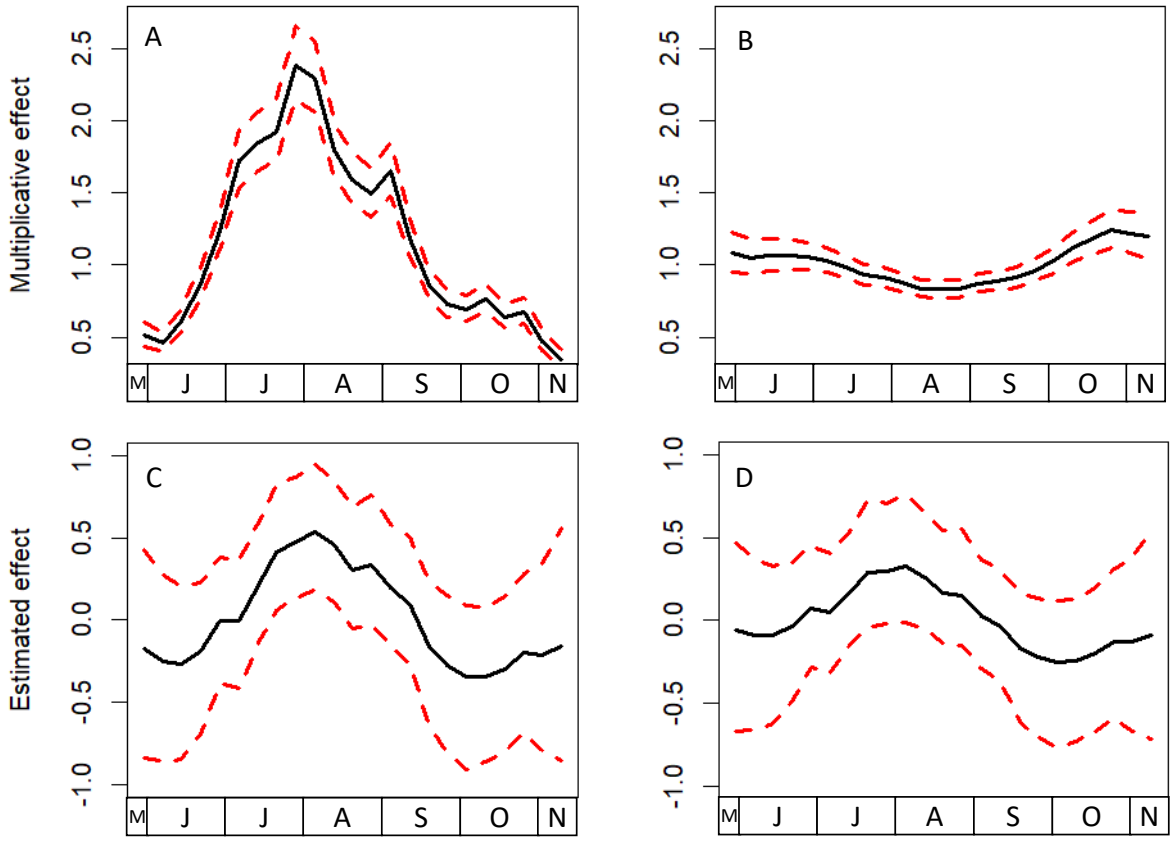


Figure 6: Seasonal effects for four models: (A) Fire Activity baseline for occurrence component; (B) Best fitted model for occurrence component; (C) Fire Activity baseline for size component; (D) Best fitted model for size component. The X axis shows the week of the year. The Y axis of A and B is the factor by which the expected number of fires changes with respect to the case of a multiplicative effect of 1, and the Y of C and D axis is the contribution of the predictor variable's value to the logit of the probability of exceedance. Neutral values are 1 for Occurrence component's multiplicative effect, and 0 for size component's estimated effect.

A similar analysis was carried out for the spatial effect (Fig. 7 and 8). In the FAb model (Fig. 7A), the spatial effect was consistent with the distribution of observed fires as expected (Fig. 1). For the model informed by fire indices, the spatial effect (Fig. 7B) is strongly reduced, the extreme values of the effect are less frequent through all the area; nonetheless, the western side of the Rhone valley and the island of Corsica show higher spatial effects. For the size component, the spatial effect was estimated at the coarser scale of Sylvo-Eco-Regions; the FAb

(Fig. 8A) had a pronounced difference between western (lower effect) and eastern zones (higher effect), with the tilting point being the Rhone valley; Corsica had the highest values on the west, but the range of the effect is generally small. The fire indices produced a shift in the spatial distribution of the effect (Fig. 8B), which was raised in the mountain regions of the north-east (Alps) and lowered in the south-west (mainly in the departements of Bouches-du-Rhone and Var, and a fraction of Vaucluse and Alpes-de-Haute-Provence).

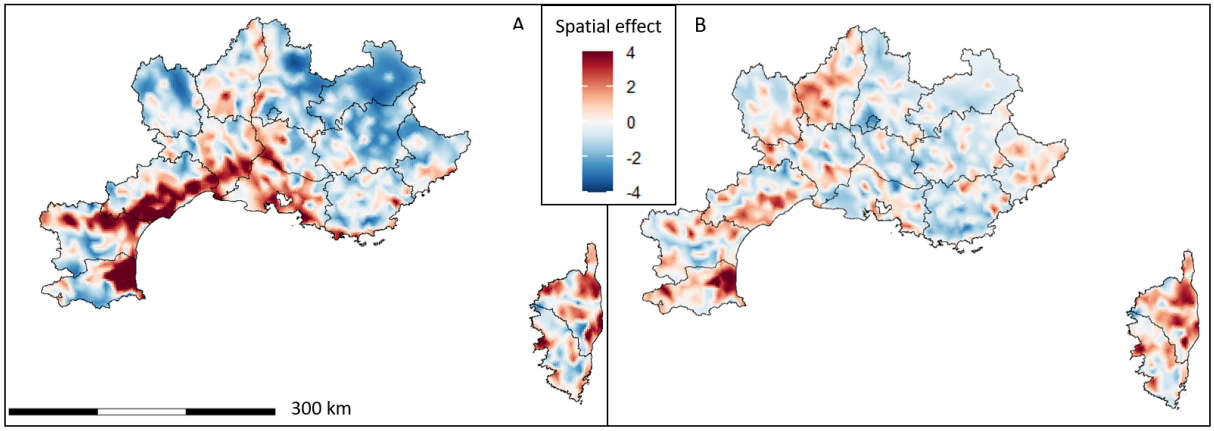


*Figure 7:Spatial effect for two occurrence models: (A) Fire Activity baseline (FAb) for occurrence component; (B) Best fitted model for occurrence component (DC*DMC*FFMC*VPD*U). This map shows that for the same level of predicted danger in given pixel and day, the fire probability will be higher or lower depending on the local environmental conditions.*

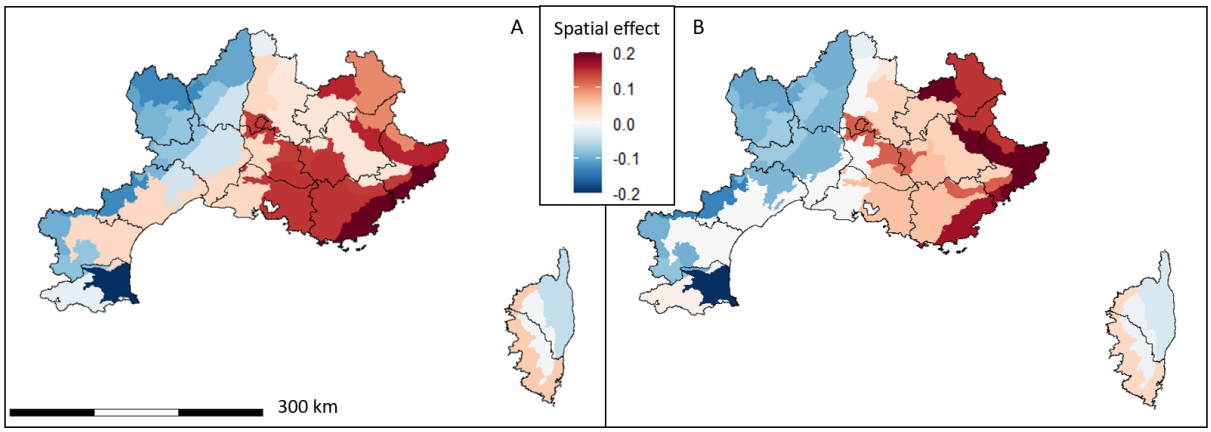


*Figure 8: Spatial effect for two size models: (A) Fire Activity baseline (FAb) for size component; (B) Best fitted model for size component (DC*DMC*FFMC*Tmax*U). This map shows that for the same level of predicted danger for a given day, the fire probability will be higher or lower depending on the local environmental conditions.*

1.7 *Evaluation of model predictability*

As shown in Pimont et al. (2021), the class of models explored here allows estimation of seasonal variations of fire numbers with a good agreement, as well as estimation of confidence interval and coverage probabilities (see examples for four seasons in the supplementary material 4). Here, we provide a more in-depth analysis of predictability using precision-recall curves (PRCs) of the occurrence component (Fig. 9 NB1) at ZFC (homogeneous Zone of Fire Climate) and daily scale and the size component (Fig. 9 P80 and NB80) at a coarser temporal scale (ZFC weekly, to reduce noise associated with low probabilities). The PRCs (Fig. 9) show that, for all predictions, the CanFWI had higher capability of prediction when it was used in the modelling framework than when it was used as an index alone. Indeed, the CanFWI alone exhibited a poor performance: a CanFWI value above 60 (corresponding to severe fire danger) had a hit rate under 10%, hence detecting only a small proportion of actual fires, while the precision values did not increase above 10% until the corresponding recall rates were close to zero. Therefore, even when very few fires were predicted, they still had a probability of more than 90% to be a false alarm. Using CanFWI in the modelling framework decreases that probability to typically 70-80%. The best-performing model (in blue Figure 9) improved the predictions marginally as compared to the CanFWI model for the fire occurrence (NB1: areas under the blue and red curves are very close), but the improvement was usually (Recall>0.2) more significant for the fire size and large-fire occurrence (P80 and NB80).

We can give different weights to the Precision and Hit Rate using the F_β measure. For each β value, the F_β measure plotted as function of the fire predictor (here, the CanFWI or the model predictor including the CanFWI) exhibits a peak value at a given fire predictor value, referred to as optimal threshold hereinafter (see the plots of F_β in Supplementary Materials 6), which is a decision criterion for the occurrence of fire events. The peak values of F_β measures (Table 7) are given by lower optimal thresholds as the β values increase. For the prediction of 1ha fire

numbers (NB1) and the probability of exceeding 80ha (P80) the difference between the optimal thresholds associated to each β value is large, e.g., in the NB1 prediction, the CanFWI varies from a value of 70,8 to 44, and the output of the best fitted model varies from 0,171 to 0.056, for β= 0.5, β=1 and β = 2 respectively. For the prediction of 80ha fire numbers (NB80) these optimal thresholds have a smaller variation: the CanFWI varies from a value of 57,1 to 53,7; and the output of the best fitted model varies from 0,05 to 0,037.

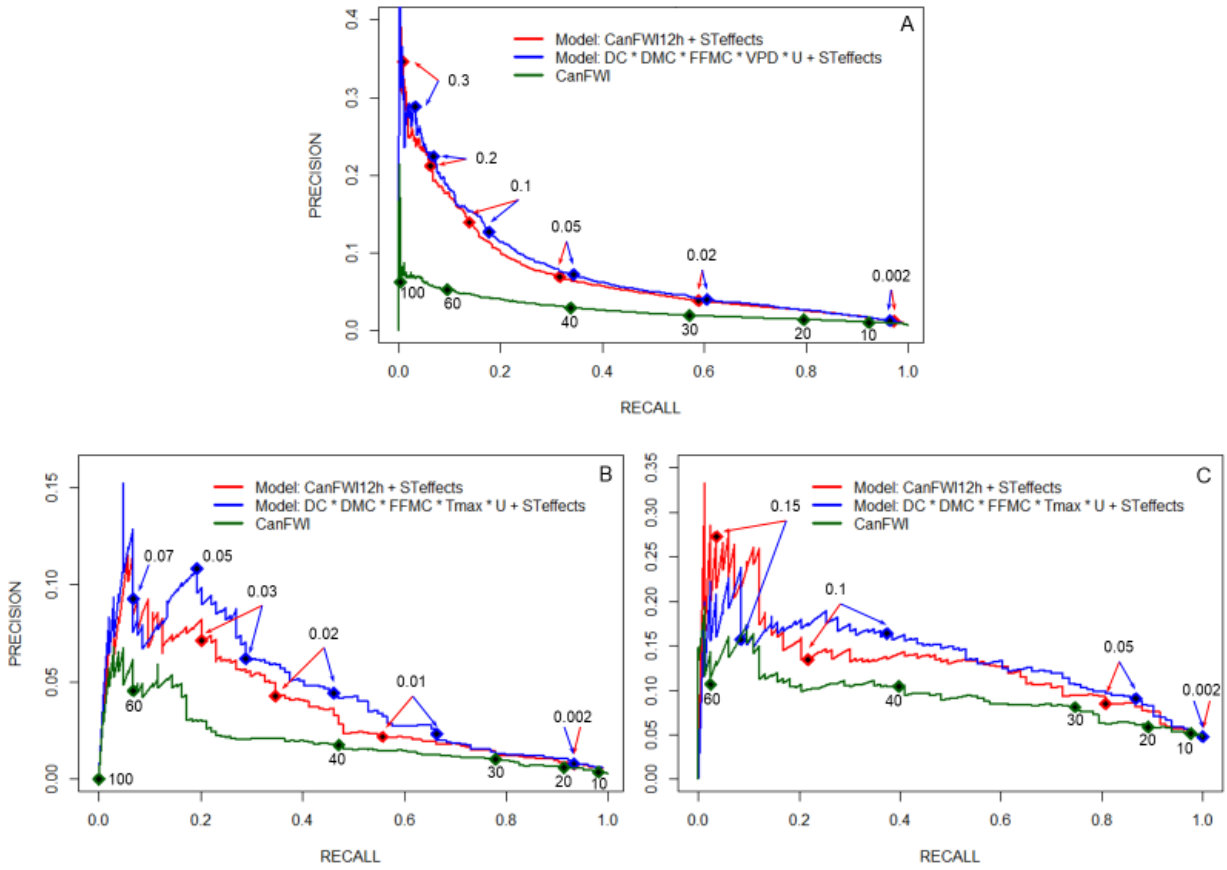


Figure 9: Precision-recall curves for: A – 1ha fires (NB1), B – 80ha fires (NB80) and C – the 80ha threshold exceedance probability (P80). The curves correspond to the CanFWI at 12h taken alone (in green) or included in the model together with spatio-temporal effects (red) and the vest fitted model for each model component (in blue). The black squares and labels show reference values of the CanFWI and the model outputs.

This analysis gives a method to choose a single optimal threshold (decision criterion) depending on the selected ß criterion. When the priority is the detection of fire events (β = 2), maximizing the importance of the Hit Rate and assuming a larger number of false alarms, the optimal

threshold is lowered. On the other hand, the optimal threshold is higher when the precision is given priority (β = 0,5), maximizing the Precision but missing more fire detections. As seen in the PRCs, the models showed a better predictability than the CanFWI alone in all situations, regardless of the β value; and although the more complex model had higher F_β scores, the two models were close to each other.

Table 7: F_β peak scores for three β values (0.5, 1 and 2) for 1ha fires (NB1), 80ha fires (NB80) and the 80ha threshold exceedance probability (P80). Threshold is the value of the fire predictor (model output or CanFWI) that gave the maximal β score in each case. The three tests correspond to, (i) the best fitted model (with STEs), (ii) the model with CanFWI at 12h with STEs, and (iii) the CanFWI at 12h outside the modelling framework.

NB1	β0,5		β1		β2	
	Threshold	Fβ scores	Threshold	Fβ scores	Threshold	Fβ scores
DC * DMC * FFMC * T(max) * U * STEs	0,171	0,159	0,113	0,153	0,056	0,197
CanFWI (12h) * STEs	0,15	0,154	0,092	0,14	0,052	0,186
CanFWI	70,8	0,062	54,4	0,071	44	0,114
NB80	β0,5		β1		β2	
	Threshold	Fβ scores	Threshold	Fβ scores	Threshold	Fβ scores
DC * DMC * FFMC * T(max) * U * STEs	0,05	0,119	0,05	0,139	0,037	0,191
CanFWI (12h) * STEs	0,057	0,102	0,034	0,117	0,023	0,162
CanFWI	57,1	0,066	55,2	0,078	53,7	0,112
P80	β0,5		β1		β2	
	Threshold	Fβ scores	Threshold	Fβ scores	Threshold	Fβ scores
DC * DMC * FFMC * T(max) * U * STEs	0,117	0,2	0,095	0,233	0,07	0,362
CanFWI (12h) * STEs	0,127	0,211	0,084	0,216	0,073	0,349
CanFWI	55,2	0,15	42,1	0,169	31,8	0,287

2.4. Fire danger rating with models and fire-weather indices

Because CanFWI(max) is used by operational services in France, we compare in Table 8 compares how this index perform to classify ZFC-days in fire danger categories (see 2.4.4), with our best probabilistic model. Most ZFC-days (~75%) were classified in danger categories "low" and "medium" while next 20% of ZFC-days were classified in danger category "high". Danger categories "very high" and "extreme" accounted for less than 5% and 0.5% of ZFC-days respectively.

The number of observed fires larger than 1ha (NB1) in the "extreme" category was more than twice higher for the model than for the CanFWI(max) (224 instead of 97) and almost 4 times lower in the "Low" category for the model than for the CanFWI(max) (35 instead of 134). This highlights that the model performs much better than the single index to rate the fire danger in ZFC at a daily scale when using a category-based approach.

*Table 8: Classification of CanFWI and model outputs in five fire-danger categories. Notations in the table are as follows: Danger is the classification category; CanFWI intervals are the breaks used for the classification; Frequency is the percentage of ZFC-days in each danger category; Percen. (Percentile) is the cumulated Frequency; NB1 are the number of >1ha fires found in each danger class classified using the CanFWI outside of any modelling framework and the DC*DMC*FFMC*VPD*U + STeffects model; NB80 are the number of >80ha fires found in each danger class classified using the CanFWI outside of any modelling framework and the DC*DMC*FFMC*VPD*U + STEs model. Percentages between breaks (%) indicate the proportion of fires found in each danger class in comparison to the total number of ZFC-days in each danger class.*

				FWI max		Model: DC*DMC*FFMC* VPD*U * STEs	
Danger	CanFWI interval	Frequency	Percen.	NB1	NB80	NB1	NB80
Low	<10	38,1	38,1	134 (0,11%)	1 (0,001%)	35 (0,03%)	0 (0%)
Medium	10<FWI<30	38,8	76,9	701 (0,57%)	9 (0,007%)	341 (0,28%)	5 (0,004%)
High	30<FWI<50	19,2	96,1	930 (1,54%)	37 (0,061%)	1020 (1,69%)	29 (0,048%)
Very high	50<FWI<80	3,68	99,8	456 (3,93%)	49 (0,42%)	698 (6,01%)	41 (0,35%)
Extreme	FWI>80	0,22	100	97 (10,37%)	20 (2,14%)	224 (24%)	41 (4,39%)

Likewise, the number of observed fires larger than 80ha (NB80) in the "extreme" category was more than twice higher for the model than for the CanFWI(max) (41 instead of 20) and almost twice lower in the low to medium category for the model than for the CanFWI(max) (5 instead of 10).

DISCUSSION

1.8 *Methodological advances*

Eastbaugh et al. (2012) stressed the need for a consistent method to compare the performance of indices. Following Phelps and Woolford (2021), we used the LogLoss scores, along with Precision-Recall curves as they are suitable for rare binary events (Saito and Rehmsmeier, 2015; Sofaer et al. 2018), to assess and compare the performance of a range of models combining fire-weather indices (FWIs) or weather variables as predictors, and including, or not, spatio-temporal effects (STEs). In the occurrence models, we used the Log-Loss score adapted to the Poisson distribution, while the size models were evaluated with the standard Log-Loss scores, which gave very similar results. In addition, we used the Skill Score to compare the performance of each FWI or combination of weather variables to the Fire Activity baseline (FAb) and to the null model. The salient point that motivated the reference to the FAb was the need to properly measure the contribution of meteorological indices or variables to model performance, while including spatio-temporal terms in the model that, alone, could already explain a large part of the variability in fire activity. To our knowledge, this is the first study developing fire occurrence and size prediction models that uses such a specific method for model evaluation and comparison. Moreover, we illustrated the use of the $F_ß$ measure, which is a weighted-harmonic mean of precision and recall, as a method to compare the prediction capability of different probabilistic models or FWIs. This approach allows to balance the trade-off between precision and recall and to select decision thresholds for the prediction of fire events from these models or FWIs.

Our results show that the best performance (max of the $F_ß$ measure) is obtained with ß >1, which corresponds to a priority on the Hit Rate. In this situation, corresponding decision thresholds (on probabilities or FWIs) are lower than those obtained with ß <=1, so that the alarm is activated at lower danger level and less false negatives occurs. For fire danger forecasting application, such a choice is the most relevant as false alarms are preferable over non-detected

fire events. The present analysis shows it also leads to a higher statistical performance of the model.

1.9 *Fire factors and models' performances*

All the FWIs provided consistent responses in terms of fire activity with respect to the expected effects of the weather. However, some indices showed saturation of this response for high index values, indicating unsuitable calibration in the context of south-east France. The responses show low uncertainty for the occurrence of fires (number of fires > 1ha), but high uncertainty for the probability of a fire becoming large (exceeding 80 ha). This is not surprising given the low number of fires > 80 ha. This behavior has already been observed in previous studies (Pimont et al. 2021; Castel-Clavera et al. 2022).

The best model obtained in terms of performance combines the fuel moisture indices of the Canadian Fire Weather System (DC, DMC and FFMC), wind speed and VPD, as well as the spatio-temporal effects (STEs). The CanFWI associated with STEs showed performances close to the best model, which are further improved by including additional variables such as wind and VPD. This indicates that the CFWIS does not completely describe the effects of temperature, moisture, and wind speed on fire activity. For the most part, the models without STEs obtain lower scores than the so-called baseline model (Fire Activity baseline – FAb), i.e., one that only includes the effects of geographical location and seasonality, and much lower scores than the weather-aware models with STEs.

Among the existing indices, the CanFWI led to the best model, notably better than the Australian Mark 5 or the American "Hot Dry Windy" index. Other indices had responses showing very low and/or saturating effect and poor log loss skill scores, meaning that they are not suitable for fire predictions in Mediterranean France. We also observed that raw climatic variables such as T(max) or RH (min) or precipitation, which do not account for cumulative temporal information produced poor predictions by their own, as seen in previous research (Abatzoglou and Kolden, 2013; Stavros et al. 2014, Brown et al. 2022). In contrast, Holsten et

al. (2013) found the RH to be one of the best fire predictors in Germany, and Holden et al. (2018) found that the precipitation within the fire season is the main controller of burnt area.

The results demonstrate the importance of taking STEs into account. Previous research already found the need to include an independent spatial effect (Preisler et al. 2004, Serra et al. 2014ab, Ager et al. 2018, Vecín-Arias et al. 2016, Rodríguez-Pérez et al. 2020, Woolford et al. 2021, Pimentel et al. 2024) and seasonal baselines (Woolford et al. 2009) for fire activity prediction. Although the spatial and temporal effects in the literature are not always considered through the same methodology, Juan-Verdoy (2021) pointed out several advantages of the Stochastic Partial Differential Equation (spatial effect) to enhance the prediction of spatial processes as wildfires, which can be further enhanced with a temporal component (seasonal effect).

The STEs account for patterns that are not quantifiable by the explicit explanatory variables of the models e.g., Gabriel et al. (2017), Opitz et al. (2020), Pimont et al. (2021), Koh et al. (2023) and Orero et al. (2024). They concern either seasonal patterns in human activities (agricultural works, vacation periods…) that complete other predictor variables (agricultural land or the roads per pixel), or spatial patterns that might also be hard to capture with explicit predictors (Lindenmayer et al. 2021, Castel-Clavera et al. 2022), e.g., particularities in firefighting teams of the different administrative regions.

The STEs in the FAb represent an average season throughout the study area. As more predictor variables are added to the model the spatial and temporal patterns of fire activity are more finely captured, identifying the small-scale variations that lead to increased or decreased fire activities. Therefore, when the effects of the weather, particularly using the best models, are used together with the STEs, the effect of the STEs on the model take a smaller magnitude than the corresponding STEs of the baseline model. In other words, the weather is indeed responsible for a significant proportion of the variations in fire activity predicted by the baseline model, but the STEs greatly improve model performance.

These results justify the operational use of the CanFWI rather than another index in this region and suggest that its accuracy can be further improved, albeit at the cost of adding new variables. Indeed, keeping a single scalar to express the meteorological danger may be desirable for certain operational uses. Note that these results relate to summer fires and could differ for winter fires, as suggested in other regions (Trucchia et al. 2022).

1.10 *Earlier studies of index performance*

Despite the wide use of FWIs in modelling studies, we found few scientific studies specifically dedicated to assessing their performance as predictors of fire activity modelling, and sometimes, assessment methods were rudimentary. The HDW index, in its first publication (Srock et al. 2018), was evaluated with only four fire events without any statistical analysis, only by looking at the weeks before and after the fire. The Sharples index (Sharples et al. 2009) was compared visually and with error statistics against other known indices, but there are no validations with fire observations. The Fosberg index (and its modification with the KBDI) was used by Goodrick (2002) in Florida, establishing that this index may not be useful in their environment; the modified Fosberg index had slightly better performance. In the context of Europe, Camia and Amatulli (2009) examined the distribution of the CanFWI components for a dataset of 2816 large fires (> 500ha) in the Mediterranean and found a clear relationship between the ISI (Initial Spread Index – Component of CanFWI based on U and FFMC) and the final fire size.

In Europe, Viegas et al. (1999) carried an early comparison of fire indices although most of the accounted indices have fallen in disuse. Holsten et al. (2013) compared the performance of selected FWIs that are more adapted to central-EU (Germany) climatic conditions as Baumgartner (Baumgartner et al. 1967), Modified M-68 (Landesforst Mecklenburg-Vorpommern, 1999), Angström (Angström, 1942) and Nesterov (Nesterov, 1949) against the CanFWI and single weather variables. The Modified M68, specifically developed for German conditions and used by the German Weather Service, exhibited the best performance, but the CanFWI constituted a valid alternative in most German states according to the authors.

DeAngelis et al (2015) developed a new fire danger index for the Canton Ticino (Switzerland) by implementing niche modelling approaches, and incidentally, compared the predictive performances of a set of meteorological variables and FWIs using the AUC, some are common to our study. The CanFWI performed well, but the use of meteorological variables alone led to similar or better performance. More recently Pagnon-Eriksson et al. (2023) made a comparative study in Sweden, they found as well the CanFWI to be the best predictor amongst 3 other indices for the year 2018; nonetheless, they did not use any modelling framework. Finally, in Mediterranean France, Barbero et al. (2019) found that large fires were associated with both short-term fire-weather conditions and longer-term soil moisture deficits, consistent with the hypotheses underlying the CanFWI and the results of the present study.

1.11 *Towards an operational fire danger rating*

The CanFWI can be used directly as a predictor of fire activity by setting a decision threshold for the occurrence of fires. Our results show, however, that its performance is far inferior to that of the Firelihood model incorporating the CanFWI and the STEs: specifically, for a fixed level of false alarms, the accuracy (percentage of correct fire detection) of the CanFWI is far inferior. Furthermore, the best fitted *Firelihood* model has demonstrated a better performance than the calibration of the CanFWI based on observed fire activities as in Alves et al. (2021) or thresholds on CanFWI used by French operational forces to assess fire danger. This demonstrates the value of using a modelling framework such as Firelihood, incorporating STEs, and opens the prospect of using this type of model rather than FWI alone, or calibrated, for operational purposes. This is consistent with probabilistic models developed for the fire Occurrence Prediction system of CFFDRS (e.g., Woolford et al. 2021).

For operational purposes, it is probably more appropriate to use the model with the best performance, based on new combinations of DC, DMC, FFMC and wind, rather than the model based on the CanFWI, since it does not make its use more complex. The operational use of this model was illustrated in the result section, where it is shown how danger classes established a

priori with CanFWI threshold values can be informed by the best-performing model. The CanFWI alone tends to overestimate the number of days with fire in the low classes and to underestimate them in the high classes, which is partially corrected by the models produced in the present study. Our method for taking into account the differences between regions with a single model thanks to the STEs, without manual recalibrations of the fire weather index. Moreover, identical thresholds can be used by all local administrations.

However, even with the best performing models developed in this study, predictions of day-to-day fire occurrence, or fire size on a weekly scale, in homogeneous zones of fire climate (ZFCs), remain highly uncertain, as indicated by the predictability scores. This means that not even these models, nor the CanFWI or other indices, can be used for accurate high-resolution forecasts of fire activity. On the other hand, they do allow better classification of ZFC-days into danger levels, which is useful for preparing fire-fighting operations. This high level of uncertainty is probably due to the random nature of the causes of fires and their spatial extent. Aggregation on larger spatial and temporal scales reduces uncertainty and remains useful for strategic applications (assessment of past and future trends, for example) and even operational applications (mobilization of resources on a regional scale).

CONCLUSION

We tested a series of fire-weather indices in the French Mediterranean region because the fire prediction methods now in use are subjected to several limitations concerning their accuracy on classifying fire danger. We included spatio-temporal effects in the prediction models to better capture the seasonal and local particularities which are not directly explained by meteorological variables. That led us to the development of a complex model evaluation method, adapted to the stochastic and rare nature of fire occurrence, to warranty fair comparisons of the models to the fire activity baseline.

The fire activity baseline alone, which represents an average fire season for each pixel of the grid, does not capture the daily nor the interannual variations but performs better than weather-based indices in capturing the spatial and seasonal effects of fire. The combination of both, fire activity baseline and fire-weather indices, gives a model that can adjust the short-term seasonal variations at small spatial scale. From our results we recommend continuing to use the Canadian Fire Weather Index and its subcomponents, since it has been the best performer index but combining it with VPD, wind speed and the spatio-temporal effects to enhance the overall predictability of the model.

The developments here, including threshold selection, will help the creation of new fire danger prediction protocols adapted to the needs and demands of the regional French emergency responders. For local operational applications, as the values of Canadian Fire Weather Index are discretized into different classes, applying our results means that the operational thresholds for this index should be adapted (modified) depending on the site and the time in the season.

Founding

This work was supported by the PhD founding given by INRAE (Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement) and Région SUD (Provence-Alpes-Côte d'Azur).

Author contributions

Jorge Castel-Clavera: Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Writing – original draft and Writing – review & editing

François Pimont: Conceptualization, Methodology, Visualization, Supervision, Validation, Writing – review & editing

Thomas Opitz: Conceptualization, Methodology, Supervision, Validation, Writing – review & editing

Julien Ruffault: Conceptualization, Methodology, Writing – review & editing

Renaud Barbero: Methodology, Writing - review & editing

Denis Allard: Methodology, Writing - review & editing

Jean-Luc Dupuy: Conceptualization, Methodology, Supervision Writing - review & editing

Declaration of competing interest

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

REFERENCES

Abatzoglou, J.T., Kolden, C.A., 2013. Relationships between climate and macroscale area burned in the western United States. Int. J. Wildland Fire 22, 1003. https://doi.org/10.1071/WF13019

Allard, D., Comunian, A., Renard, P., 2012. Probability Aggregation Methods in Geoscience. Math. Geosci. 44, 545–581. https://doi.org/10.1007/s11004-012-9396-3

Alves, D., Almeida, M., Viegas, D.X., Novo, I., Luna, M.Y., 2021. Fire Danger Harmonization Based on the Fire Weather Index for Transboundary Events between Portugal and Spain. Atmosphere 12, 1087. https://doi.org/10.3390/atmos12091087

Ångström, A., 1942. The risks for forest fires and their relation to weather and climate (Riskerna för Skogsbrand och deras beroende ar väder och klimat). Svenka skogrårdsföreningens Tidskr. 4, 18 (in Swedish)

Baijnath-Rodino, J.A., Foufoula-Georgiou, E., Banerjee, T., 2020. Reviewing the "Hottest" Fire Indices Worldwide (preprint). Climatology (Global Change). https://doi.org/10.1002/essoar.10503854.1

Balch, J.K., Bradley, B.A., Abatzoglou, J.T., Nagy, R.C., Fusco, E.J., Mahood, A.L., 2017. Human-started wildfires expand the fire niche across the United States. Proc. Natl. Acad. Sci. U. S. A. 114, 2946–2951. https://doi.org/10.1073/pnas.1617394114

Barbero, R., Abatzoglou, J.T., Larkin, N.K., Kolden, C.A., Stocks, B., 2015. Climate change presents increased potential for very large fires in the contiguous United States. Int. J. Wildland Fire 24, 892. https://doi.org/10.1071/WF15083

Barbero, R., Abatzoglou, J.T., Steel, E.A., Larkin, N.K., 2014. Modeling very large-fire occurrences over the continental United States from weather and climate forcing. Environ. Res. Lett. 9, 124009.

Barbero, R., Curt, T., Ganteaume, A., Maillé, E., Jappiot, M., Bellet, A., 2019. Simulating the effects of weather and climate on large wildfires in France. Nat. Hazards Earth Syst. Sci. 19, 441–454. https://doi.org/10.5194/nhess-19-441-2019

Baumgartner, A., Klemmer, L, Raschke, E., Waldmann, G., 1967. Waldbrände in Bayern 1950 bis 1959.Mitteilungen aus der Staatsforstverwaltung Bayerns 36: 57–79.

Bedia, J., Herrera, S., Camia, A., Moreno, J.M., Gutiérrez, J.M., 2014a. Forest fire danger projections in the Mediterranean using ENSEMBLES regional climate change scenarios. Clim. Change 122, 185–199. https://doi.org/10.1007/s10584-013-1005-z

Bedia, J., Herrera, S., Gutiérrez, J.M., 2014b. Assessing the predictability of fire occurrence and area burned across phytoclimatic regions in Spain. Nat. Hazards Earth Syst. Sci. 14, 53–66. https://doi.org/10.5194/nhess-14-53-2014

Brillinger, D.R., Preisler, H.K., Benoit, J.W., 2003. Risk assessment: a forest fire example, in: Institute of Mathematical Statistics Lecture Notes - Monograph Series. Institute of Mathematical Statistics, Beachwood, OH, pp. 177–196. https://doi.org/10.1214/lnms/1215091142

Brown, P., Hanley, H., Mahesh, A., Reed, C., Strenfel, S., Davis, S., Kochanski, A., Clements, C., 2022. Climate-Driven Risk of Extreme Wildfire in California (preprint). Physical Sciences and Mathematics. https://doi.org/10.31223/X5K648

Camia, A., Amatulli, G., 2009. Weather Factors and Fire Danger in the Mediterranean [WWW Document]. JRC Publ. Repos. https://doi.org/10.1007/978-3-642-01754-4

Canadian Forest Service Fire Danger Group (CFSFDG), 2021. An overview of the next generation of the Canadian Forest Fire Danger Rating System.

Carrega, P., 1991. A Meteorological Index of Forest Fire Hazard in Mediterranean France. Int. J. Wildland Fire 1, 79. https://doi.org/10.1071/WF9910079

Castel-Clavera, J., Pimont, F., Opitz, T., Ruffault, J., Rivière, M., Dupuy, J.-L., 2022. Disentangling the factors of spatio-temporal patterns of wildfire activity in south-eastern France. Int. J. Wildland Fire 32, 15–28.

Chandler, C., Cheney, P., Thomas, P., Trabaud, L., Williams, D., 1985. Fire in forestry. Volume 1. Forest fire behavior and effects. Fire Atmosph Sci Res USDA Serv Wash. DC USA.

Coughlan, R., Di Giuseppe, F., Vitolo, C., Barnard, C., Lopez, P., Drusch, M., 2021. Using machine learning to predict fire-ignition occurrences from lightning forecasts. Meteorol. Appl. 28, e1973. https://doi.org/10.1002/met.1973

Cruz, M.G., Gould, J.S., Alexander, M.E., Sullivan, A.L., McCaw, W.L., Matthews, S., 2015. A Guide to Rate of Fire Spread Models for Australian Vegetation. CSIRO Land and Water, Canberra, ACT, and AFAC, Melbourne, Vic.

DaCamara, C.C., Calado, T.J., Ermida, S.L., Trigo, I.F., Amraoui, M., Turkman, K.F., 2014. Calibration of the Fire Weather Index over Mediterranean Europe based on fire activity retrieved from MSG satellite imagery. Int. J. Wildland Fire 23, 945. https://doi.org/10.1071/WF13157

De Angelis, A., Ricotta, C., Conedera, M., Pezzatti, G.B., 2015. Modelling the Meteorological Forest Fire Niche in Heterogeneous Pyrologic Conditions. PLOS ONE 10, e0116875. https://doi.org/10.1371/journal.pone.0116875

Deeming, J.E., Burgan, R.E., Utah), I.F. and R.E.S. (Ogden, Cohen, J.D., 1977. The National Fire-Danger Rating System, 1978. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station.

Dupuy, J., Fargeon, H., Martin-StPaul, N., Pimont, F., Ruffault, J., Guijarro, M., Hernando, C., Madrigal, J., Fernandes, P., 2020. Climate change impact on future wildfire danger and activity in southern Europe: a review. Ann. For. Sci. 77, 35. https://doi.org/10.1007/s13595-020-00933-5

Eastaugh, C.S., Arpaci, A., Vacik, H., 2012. A cautionary note regarding comparisons of fire danger indices. Nat. Hazards Earth Syst. Sci. 12, 927–934. https://doi.org/10.5194/nhess-12-927-2012

Fosberg, M.A., Fosberg, M.A., 1978. Weather in wildland fire management: the fire weather index. Conf. Sierra Nev. Meteorol. June 19–21, 1978, S Lake Tahoe, California, 1–4.

Gabriel, E., Opitz, T., & Bonneu, F., 2017. Detecting and modeling multi-scale space-time structures: the case of wildfire occurrences. Journal de la Société Française de Statistique, 158(3), 86-105.

Ganteaume, A., Camia, A., Jappiot, M., San-Miguel-Ayanz, J., Long-Fournel, M., Lampin, C., 2013. A review of the main driving factors of forest fire ignition over Europe. Environmental management, 51, 651-662.

Garcia-Prats, A., Antonio, D.C., Tarcísio, F.J.G., Antonio, M.J., 2015. Development of a Keetch and Byram—Based drought index sensitive to forest management in Mediterranean conditions. Agric. For. Meteorol. 205, 40–50. https://doi.org/10.1016/j.agrformet.2015.02.009

Gneiting, T., Raftery, A.E., 2007. Strictly Proper Scoring Rules, Prediction, and Estimation. J. Am. Stat. Assoc. 102, 359–378. https://doi.org/10.1198/016214506000001437

Goodrick, S.L., 2002. Modification of the Fosberg fire weather index to include drought. Int. J. Wildland Fire 11, 205. https://doi.org/10.1071/WF02005

Holden, Z.A., Swanson, A., Luce, C.H., Jolly, W.M., Maneta, M., Oyler, J.W., Warren, D.A., Parsons, R., Affleck, D., 2018. Decreasing fire season precipitation increased recent western US forest wildfire activity. Proc. Natl. Acad. Sci. 115. https://doi.org/10.1073/pnas.1802316115

Holsten, A., Dominic, A., Costa, L., Kropp, J., 2013. Evaluation of the Performance of Meteorological Forest Fire Indices for German Federal States. For. Ecol. Manag. 287, 123–131. https://doi.org/10.1016/j.foreco.2012.08.035

Juan Verdoy, P., 2021. Enhancing the SPDE modeling of spatial point processes with INLA, applied to wildfires. Choosing the best mesh for each database ». Communications in Statistics - Simulation and Computation 50, no 10. https://doi.org/10.1080/03610918.2019.1618473.

Koh, J. 2023. Gradient Boosting with Extreme-Value Theory for Wildfire Prediction. Extremes 26, no 2. https://doi.org/10.1007/s10687-022-00454-6.

Krawchuk, M.A., Cumming, S.G., Flannigan, M.D., 2009. Predicted changes in fire weather suggest increases in lightning fire initiation and future area burned in the mixedwood boreal forest. Clim. Change 92, 83–97. https://doi.org/10.1007/s10584-008-9460-7

Landesforst Mecklenburg-Vorpommern, 1999. Durchführungserlass zum Gemeinsamen Waldbrandrunderlass des Ministerium für Landwirtschaft, Umwelt und Verbraucherschutz und Innenministeriums.

Lindenmayer, D., Taylor, C., Blanchard, W., 2021. Empirical Analyses of the Factors Influencing Fire Severity in Southeastern Australia. Ecosphere 12. https://doi.org/10.1002/ecs2.3721.

Martell, D.L., 2015. A Review of Recent Forest and Wildland Fire Management Decision Support Systems Research. Curr. For. Rep. 1, 128–137. https://doi.org/10.1007/s40725-015-0011-y

McArthur, A.G., 1967. Fire behaviour in eucalypt forests, Forestry and Timber Bureau. ed. Commonwealth Department of National Development., Canberra, ACT. 36 pp.

McArthur, A.G., 1977. Grassland Fire Danger Meter Mark V. Country Fire Authority, Melbourne, Vic. (published as a linear slide rule).

Moritz, M.A., Morais, M.E., Summerell, L.A., Carlson, J.M., Doyle, J., 2005. Wildfires, complexity, and highly optimized tolerance. Proc. Natl. Acad. Sci. 102, 17912–17917. https://doi.org/10.1073/pnas.0508985102

Nadeem, K., Taylor, S.W., Woolford, D.G., Dean, C.B., 2019. Mesoscale spatiotemporal predictive models of daily human- and lightning-caused wildland fire occurrence in British Columbia. Int. J. Wildland Fire 29, 11–27. https://doi.org/10.1071/WF19058

Nesterov, V.G., 1949. Combustibility of the forest and methods for its determination (in Russian). USSR State Industry Press.

Noble, I., Bary, G., Gill, A., 1980. Mcarthur Fire-Danger Meters Expressed as Equations. Aust. J. Ecol. 5, 201–203. https://doi.org/10.1111/j.1442-9993.1980.tb01243.x

Opitz, T., Bonneu, F., Gabriel, E., 2020. Point-Process Based Bayesian Modeling of Space–Time Structures of Forest Fire Occurrences in Mediterranean France. Spatial Statistics 40. https://doi.org/10.1016/j.spasta.2020.100429.

Orieux, A., 1974. Conditions météorologiques et incendies en région méditerranéenne. Rev. For. Fr. 122. https://doi.org/10.4267/2042/20873

Pagnon Eriksson, C., Johansson, N., McNamee M., 2023. The performance of wildfire danger indices: A Swedish case study. Safety Science 159. https://doi.org/10.1016/j.ssci.2022.106038.

Parisien, M.-A., Snetsinger, S., Greenberg, J., Nelson, C., Schoennagel, T., Dobrowski, S., Moritz, M., 2012. Spatial variability in wildfire probability across the western United States. Int. J. Wildland Fire 21, 313–327. https://doi.org/10.1071/WF11044

Pausas, J.G., Ribeiro, E., 2013. The global fire–productivity relationship. Glob. Ecol. Biogeogr. 22, 728–736. https://doi.org/10.1111/geb.12043

Phelps, N., Woolford, D.G., 2021. Guidelines for effective evaluation and comparison of wildland fire occurrence prediction models. Int. J. Wildland Fire 30, 225. https://doi.org/10.1071/WF20134

Pimentel, J.S., Bulhões, R.S., et Canas Rodrigues, P., 2024. Bayesian Spatio-Temporal Modeling of the Brazilian Fire Spots between 2011 and 2022. Scientific Reports 14. https://doi.org/10.1038/s41598-024-70082-6.

Pimont, F., Fargeon, H., Opitz, T., Ruffault, J., Barbero, R., Martin StPaul, N., Rigolot, E., Rivière, M., Dupuy, J., 2020. Prediction of regional wildfire activity with a probabilistic Bayesian framework (preprint). Bioengineering. https://doi.org/10.1101/2020.05.20.105767

Pimont, F., Fargeon, H., Opitz, T., Ruffault, J., Barbero, R., Martin-StPaul, N., Rigolot, E., RiviÉre, M., Dupuy, J., 2021. Prediction of regional wildfire activity in the probabilistic Bayesian framework of Firelihood. Ecol. Appl. 31, e02316. https://doi.org/10.1002/eap.2316

Preisler, H.K., Brillinger, D.R., Burgan, R.E., Benoit, J.W., 2004. Probability based models for estimation of wildfire risk. Int. J. Wildland Fire 13, 133. https://doi.org/10.1071/WF02061

Preisler, H.K., Westerling, A.L., 2007. Statistical Model for Forecasting Monthly Large Wildfire Events in Western United States. J. Appl. Meteorol. Climatol. 46, 1020–1030. https://doi.org/10.1175/JAM2513.1

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rao, K., Williams, A.P., Flefil, J.F., Konings, A.G., 2020. SAR-enhanced mapping of live fuel moisture content. Remote Sens. Environ. 245, 111797. https://doi.org/10.1016/j.rse.2020.111797

Rodríguez-Pérez, J.R., Ordóñez, C., Roca-Pardiñas, J., Vecín-Arias, D., et Castedo-Dorado, F., 2020. Evaluating Lightning-Caused Fire Occurrence Using Spatial Generalized Additive Models: A Case Study in Central Spain. Risk Analysis 40, no 7. https://doi.org/10.1111/risa.13488.

Rue, H., Riebler, A., Sørbye, S.H., Illian, J.B., Simpson, D.P., Lindgren, F.K., 2017. Bayesian Computing with INLA: A Review. Annu. Rev. Stat. Its Appl. 4, 395–421. https://doi.org/10.1146/annurev-statistics-060116-054045

Ruffault, J., Mouillot, F., 2015. How a New Fire-Suppression Policy Can Abruptly Reshape the Fire-Weather Relationship. Ecosphere 6, no 10: art199. https://doi.org/10.1890/ES15-00182.1

Ruffault, J., Curt, T., Moron, V., Trigo, R. M., Mouillot, F., Koutsias, N., Belhadj-Khedher, C., 2020. Increased likelihood of heat-induced large wildfires in the Mediterranean Basin. Scientific reports, 10(1), 13790.

Saito, T., Rehmsmeier, M., 2015. The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. PLOS ONE 10, e0118432. https://doi.org/10.1371/journal.pone.0118432

Serra, L., Saez, M., Mateu, J., Varga, D., Juan, P., Díaz-Ávalos, C., Rue, H., 2014. Spatio-temporal log-Gaussian Cox processes for modelling wildfire occurrence: the case of Catalonia, 1994–2008. Environ. Ecol. Stat. 21, 531–563. https://doi.org/10.1007/s10651-013-0267-y

Serra Saurina, L., Sáez Zafra, M., Mateu, J., Juan Verdoy, P., Varga, D., 2014. A spatio-temporal Poisson hurdle point process to model wildfires. Copyr. Stoch. Environ. Res. Risk Assess. 2014 Vol 28 Núm 7 P 1671-1684.

Sharples, J.J., McRae, R.H.D., Weber, R.O., Gill, A.M., 2009. A simple index for assessing fuel moisture content. Environ. Model. Softw. 24, 637–646. https://doi.org/10.1016/j.envsoft.2008.10.012

Sofaer, H.R., Hoeting, J.A., Jarnevich, C.S., 2018. The area under the precision-recall curve as a performance metric for rare binary events. Methods Ecol. Evol. 10, 565–577. https://doi.org/10.1111/2041-210X.13140

Srock, A., Charney, J., Potter, B., Goodrick, S., 2018. The Hot-Dry-Windy Index: A New Fire Weather Index. Atmosphere 9, 279. https://doi.org/10.3390/atmos9070279

Sullivan, A.L., 2017. Inside the Inferno: Fundamental Processes of Wildland Fire Behaviour. Curr. For. Rep. 3, 132–149. https://doi.org/10.1007/s40725-017-0057-0

Taylor, A., Bruce, M., Britton, A., Owen, J., Gagkas, Z., Pohle, I., Fielding, D., Hadden, R., 2022. Fire Danger Rating System (FDRS) Report.

Taylor, S.W., Woolford, D.G., Dean, C.B., Martell, D.L., 2013. Wildfire Prediction to Inform Fire Management: Statistical Science Challenges. Stat. Sci. 28. https://doi.org/10.1214/13-STS451

Turco, M., Rosa-Cánovas, J.J., Bedia, J., Jerez, S., Montávez, J.P., Llasat, M.C., Provenzale, A., 2018. Exacerbated fires in Mediterranean Europe due to anthropogenic warming projected with non-stationary climate-fire models. Nat. Commun. 9, 3821.

Turco, M., Von Hardenberg, J., AghaKouchak, A., Llasat, M.C., Provenzale, A., Trigo, R.M., 2017. On the key role of droughts in the dynamics of summer fires in Mediterranean Europe. Sci. Rep. 7, 81. https://doi.org/10.1038/s41598-017-00116-9

Turner, R., 2009. Point patterns of forest fire locations. Environ. Ecol. Stat. 16, 197–223. https://doi.org/10.1007/s10651-007-0085-1

Vecín-Arias, D., Castedo-Dorado, F., Ordóñez, C., Rodríguez-Pérez, J.R., 2016. Biophysical and lightning characteristics drive lightning-induced fire occurrence in the central plateau of the Iberian Peninsula. Agricultural and Forest Meteorology 225. https://doi.org/10.1016/j.agrformet.2016.05.003.

Viegas, D.X., Bovio, G., Ferreira, A., Nosenzo, A., Sol, B., 1999. Comparative study of various methods of fire danger evaluation in southern Europe. Int. J. Wildland Fire 9, 235-246.

Vidal, J., Martin, E., Franchistéguy, L., Baillon, M., Soubeyroux, J., 2010. A 50-year high-resolution atmospheric reanalysis over France with the Safran system. Int. J. Climatol. 30, 1627–1644. https://doi.org/10.1002/joc.2003

Vilar, L., Woolford, Douglas.G., Martell, D.L., Martín, M.P., 2010. A model for predicting human-caused wildfire occurrence in the region of Madrid, Spain. Int. J. Wildland Fire 19, 325. https://doi.org/10.1071/WF09030

Wagner, C.E.V., n.d. Structure of the canadian forest fire weather index.

Wang, X., Wotton, B.M., Cantin, A.S., Parisien, M.-A., Anderson, K., Moore, B., Flannigan, M.D., 2017. cffdrs: an R package for the Canadian Forest Fire Danger Rating System. Ecol. Process. 6, 5. https://doi.org/10.1186/s13717-017-0070-z

Woolford, D.G., Braun, W.J., Dean, C.B., Martell, D.L, 2009. Site-Specific Seasonal Baselines for Fire Risk in Ontario. GEOMATICA 63, no 4. https://doi.org/10.5623/geomat-2009-0049.

Woolford, D.G., Bellhouse, D.R., Braun, W.J., Dean, C.B., Martell, D.L., Sun, J., 2011. A spatio-temporal model for people-caused forest fire occurrence in the Romeo Malette Forest. J. Environ. Stat. 2, 26.

Woolford, D.G., Martell, D.L., McFayden, C.B., Evens, J., Stacey, A., Wotton, B.M., Boychuk, D., 2021. The development and implementation of a human-caused wildland fire occurrence prediction system for the province of Ontario, Canada. Can. J. For. Res. 51, 303–325. https://doi.org/10.1139/cjfr-2020-0313

Worsnop, R.P., Scheuerer, M., Giuseppe, F.D., Barnard, C., Hamill, T.M., Vitolo, C., 2021. Probabilistic Fire Danger Forecasting: A Framework for Week-Two Forecasts Using Statistical Postprocessing Techniques and the Global ECMWF Fire Forecast System (GEFF). Weather Forecast. 36, 2113–2125. https://doi.org/10.1175/WAF-D-21-0075.1

Xi, D.D.Z., Taylor, S.W., Woolford, D.G., Dean, C.B., 2019. Statistical Models of Key Components of Wildfire Risk. Annu. Rev. Stat. Its Appl. 6, 197–222. https://doi.org/10.1146/annurev-statistics-031017-100450

Yang, Y., Uddstrom, M., Pearce, G., Revell, M., 2015. Reformulation of the Drought Code in the Canadian Fire Weather Index System Implemented in New Zealand. J. Appl. Meteorol. Climatol. 54, 1523–1537. https://doi.org/10.1175/JAMC-D-14-0090.1

Ziegler, J.P., Hoffman, C.M., Mell, W., 2019. firebehavioR: An R Package for Fire Behavior and Danger Analysis. Fire 2, 41. https://doi.org/10.3390/fire2030041

3.2. Overview of the results

This chapter presents a novel method for fire activity prediction models based on the recommendations from Phelps & Woolford (2021). The presented methodology uses the LogLoss score adapted for Poisson processes. LogLoss is a proper scoring rule, meaning that its lowest value is attained when the predicted probability distribution is equal to the observed distribution. LogLoss measures the deviation between predictions – a continuous scale of probabilities – and observations – discrete presence or absence counts – for binary statistical models. Moreover, the performance of each model is expressed as a skill score where the reference performance is the fire activity baseline.

The results confirm the importance of the STEs for fire danger prediction. The STEs where necessary for almost all predictors to overcome the performance of the fire activity baseline. Between all tested predictors, the CFWIS provides the most accurate fire activity predictions through the probabilistic model for the summer season in Mediterranean France. Nonetheless, the performance of the CanFWI was improved with the inclusion of VPD and wind speed. The best performer model included the daily maximal DC, DMC and FFMC components of the CFWIS plus the daily maximal VPD, mean wind speed, and STEs. These shows that, even if the CFWIS provides fair estimations of fire danger, it has a lack of consideration for the effects of wind and short-term drought.

In addition, supplementary results show that there is no significant difference between the use of maximal daily values (so called FWIx in the French operational methodology) or 12h values for each fire weather index. This result is particularly interesting for operational applications (see 1.7. *Fire danger prediction in France*), since the computational costs of calculating the maximal daily index can be up to 10 times higher than for the value at 12h. This is because the maximal daily values are obtained by calculating the index at all hours of the day (e.g., from 8am to 6pm) and selecting the largest final value. In addition, several indices are not meant to be computed in this way. They already account for the theoretical worst daily conditions based on a single defined time weather observation. An example of this type of index structure is the CanFWI, which aims for the most fire prone conditions of the afternoon using the values at midday (12h).

3.3. Remaining unexplained spatial and temporal variability

The CFWIS has shown to be a good predictor of fire occurrence and size danger through the probabilistic framework. Nonetheless, even the most performant indicators need the contribution of the spatio-temporal model components. This indicates that there is still a

significative part of the predicted variance that is not explained by the explicit predictors. A hypothesis that emerges is that weather conditions do not completely capture the fuel status. In the next chapter this lag of explanatory variance will be addressed with remote sensing techniques.

4. Advancements from optical Remote Sensing for spatial and temporal fire activity and danger prediction.

Two main factors driving fire activities have been addressed in chapter 3 (fire weather) and in chapter 1 (human factors and land use - land cover). But none of the before mentioned variables are a direct measure of the fuel physical characteristics before the fire. The LULC data give a hint about the possible fuel types covering the study area and their characteristics, and LULC products may indicate major temporal changes in fuel types over decades. However, LULC data do not provide highly resolved information in both the spatial and temporal dimensions. The phenological state of vegetation through the season, their greenness, and the specific fuel density, as well as their interannual changes, are not provided by the LULC data. The weather indices give estimations of drought and fuel moisture content based only on weather observations. However, those estimations do not account for the presence and composition of fuels. Daily to monthly weather conditions are responsible for the dead fuel moisture content (DFMC) as the hygrometry of those fuels will adapt to the relative humidity of the air at a speed dependent on the fuel size. The vapor pressure deficit (VPD) is one of the weather variables that best fits the DFMC (specially for finer dead fuels) as for the reviewed literature, however, local characteristics of vegetation affect DFMC (Wotton and Beverly, 2007). Moreover, a large part of the forest fuels is green vegetation. The live fuel moisture content (LFMC) is partially controlled by the weather but is influenced by weekly to monthly drought. It also depends on the resistance of the vegetation to drought, their strategies to extract water from dryer soils and other specific properties. The specific vegetation characteristics change the relation between weather conditions and LFMC.

Field measurements are a direct way to know the specific fuel load, structure and the LFMC, but this data acquisition process is highly demanding on resources and obviously cannot offer a systematic coverage of the landscape. Processed based models have been developed that may predict the water dynamics of vegetation from local traits and soil properties (e.g., SurEAU; Ruffault et al., 2022). However, these models still need evaluation and to date they are not adapted to the use at large scale studies because of their computational costs and demand of precision data for their inputs. Remote sensing

presents itself as a system to directly measure vegetation properties of large spatial extents and long periods with adequate resolution. Remote sensing data is sensitive to the presence of vegetation, structure, phenology and can inform about FMC, as explained in section 1.6.2. This is thanks to the different absorbance and reflectance properties of water and dry matter on specific wave lengths from the radio-magnetic radiation.

This chapter develops around the hypothesis that satellite optical reflectance data informs about the structural and hydric effects of vegetation when introduced as factor of a probabilistic fire activity model. Given the inherently spatial nature of remote sensing data, it could improve the model predictions, particularly regarding spatial gradients in fire activity. We did not consider other remote sensing data that are known to respond to vegetation water content variations, such as micro-waves or reflectance in thermal infrared bands, because of the complexities or limitations (spatio-temporal resolution) of these data.

4.1. Optical remote sensing data (MODIS and Sentinel2)

We established a collaboration with the Centre National des Études Spatiales (CNES) framed in the ALEOFEU project (Space Climate Observatory Project). They provided us with pretreated satellite observations from MODIS and Sentinel.

4.1.1. MODIS data presentation

MODIS presented several complications regarding reprojection of the raw reflectance data (personal conversation with CNES), so that they only retrieve a composite product with already calculated indices in 250m pixels and with a 16-day resolution (BOX 2). The provided MODIS data is a collection of elaborated layers with 16 days of temporal resolution and 250m of spatial resolution. Each layer is a temporal composite, it may contain observations from any date within the 16-day time step (Fig. 14). This entails the major limitation for the use of this type of data. From a single layer, not all pixels may contain observations taken before a chosen fire. Therefore, the pixels observed after the fire must be discarded and substituted by observations from the previous layer. This means that the closer observation before the fire may correspond to periods of even 1 month before. For these reasons, we focused the analysis on results based on Sentinel data.

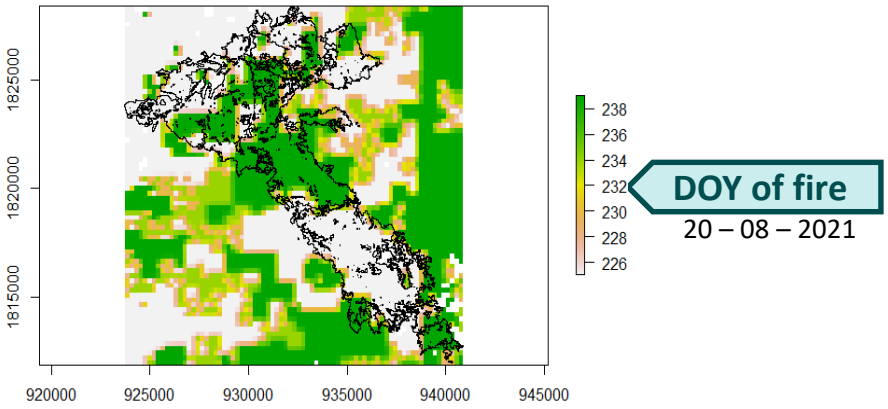


Figure 14: Example of days composing a group of MODIS products layers given for August of 2021. The black perimeter corresponds to the Gonfaron fire.

4.1.2. Sentinel data presentation and treatment

Sentinel-2 data come from two identical satellites that have 10m of spatial resolution. (https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2). Sentinel data was given for the tiles1 within the study area. In one part of the tiles, Sentinel has 5-day temporal resolution. In the other tiles the temporal resolution is irregular, with consecutive 2 – 3-day time-steps between observations due to the superposition of orbits. After selecting the Sentinel pixels that cover wildland areas at 10m resolution. and discarding all others, they were aggregated at 250m resolution to benefit from a faster computation.

It is possible that, even though the spectral bands measure separated wavelengths, their spatial and temporal patterns are similar. This resemblance on the signals will give the same information once introduced in the probabilistic model. Therefore, the first analysis performed with the Sentinel data, once extracted, was to evaluate the correlation between the optical remote sensing variables, including the spectral bands and spectral indices (NDVI and DNMI). This evaluation was done to avoid correlation between predictor variables in the probabilistic models.

The correlation is evaluated on the temporal dimension, separating the data on 9 zones based on the SERs. The SERs are the division of the study area based on ecological properties (see Figure 3). They have been grouped to a simpler division based on a broader criterion. This classification separates the mountain ranges, the coast lands, and two types of inland landscapes. The correlation matrices (Fig. 15) reveal medium and low correlation between the weather indices and the remote sensing data in all regions. There

[1] Sentinel data is cut in tiles of 110km x 110km, the study area used 16 tiles.

is high correlation within bands B2-B3-B4-B5 (from Blue to Red edge-1) and within bands B6-B7-B8-B8A (from Red-edge-2 to NIR2). Also, high correlation is found between NDVI and NDMI. High to moderate correlation is found between the remote sensing indices and all spectral bands. Hence, it won't be appropriate to use the NDVI and NDMI together with the spectral bands as predictors in fire models. In some regions, e.g., *cote est* and *cote ouest*, the correlation is overall higher.

The spatial correlation is evaluated by separating the dataset by month. In this case, the patterns of correlation (Fig.16) are similar to the temporal correlation. Overall correlation is lower for the months of June, July, and August. These preliminary results show which variables can be omitted in the model fitting tests because they carry almost identical information. In addition, as the correlation between variables follows similar patterns in both temporal and spatial dimensions, the choice of variables to include in the spatio-temporal modelling is not conditioned.

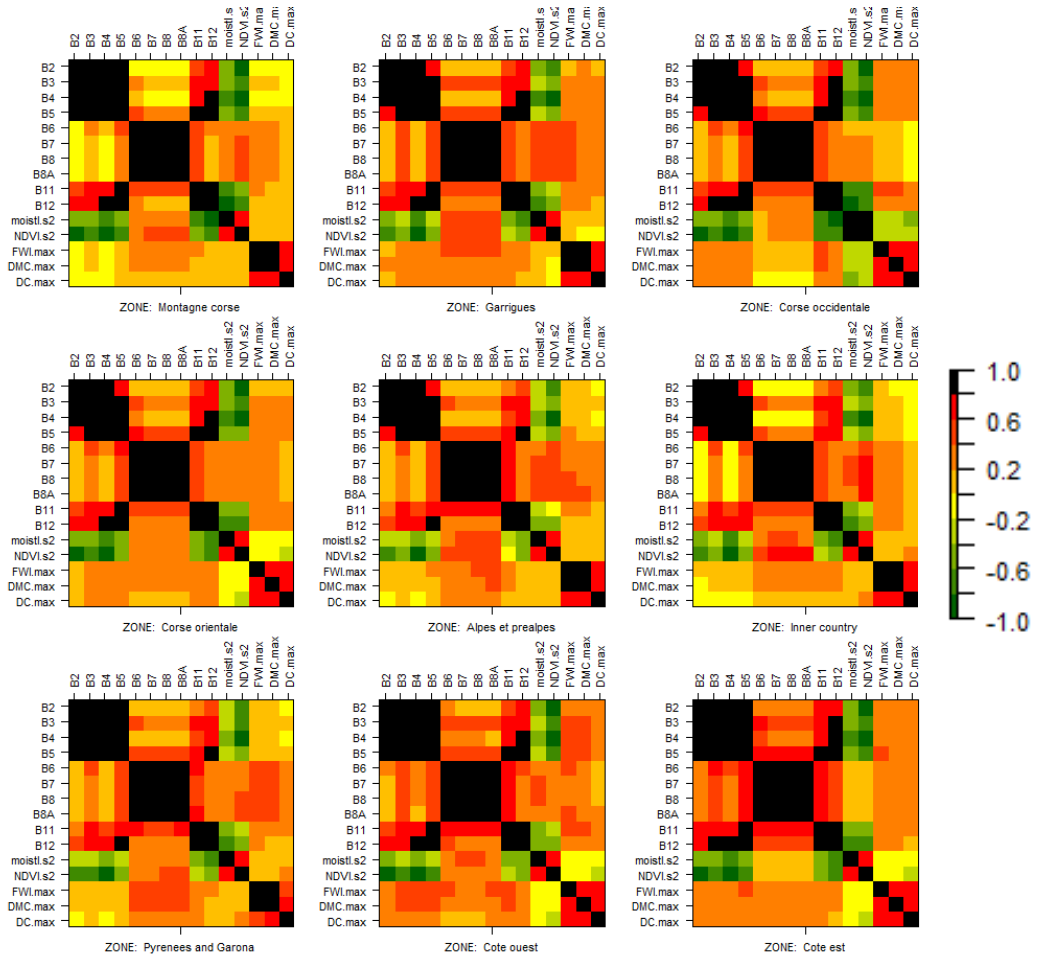


Figure 15: Correlation matrices for temporal series of all Sentinel-2 products and CFWIS moisture codes.

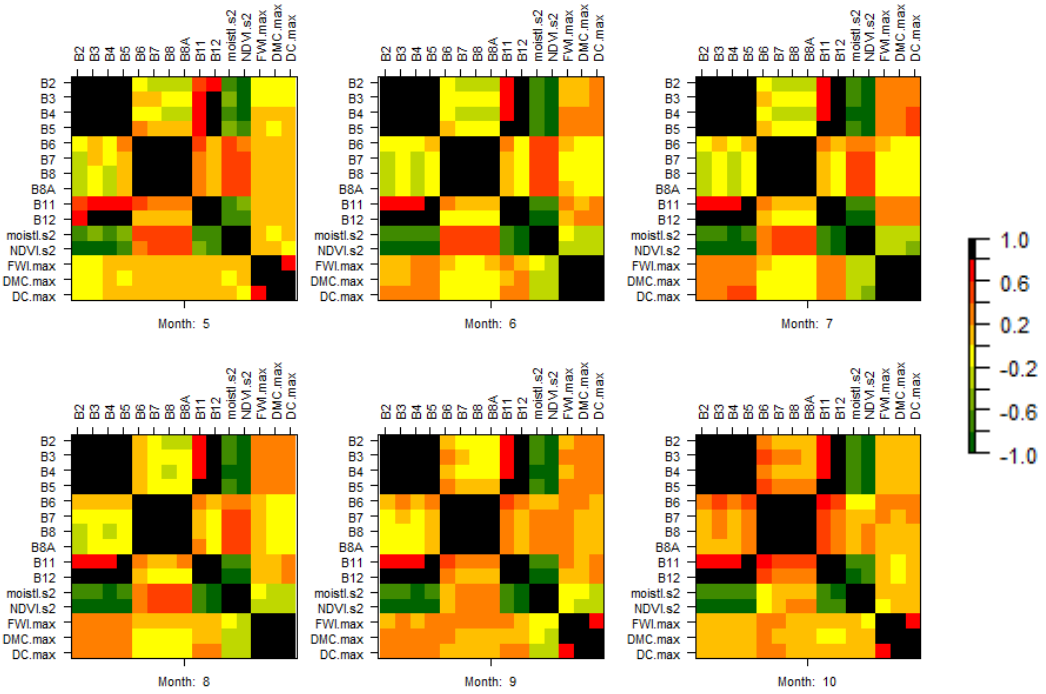


Figure 16: Correlation matrices for spatial distribution of all Sentinel-2 products and CFWIS moisture codes

As said before, several tiles have an irregular temporal resolution. In those cases, the consecutive satellites observations have different observation azimuth angles. This creates two sets of differentiable observations because the reflectance observations are sensible to the view angle relative to the Sun position. In this study, the difference between the orbits will be omitted, because of the low consequences. The observations will be temporally interpolated to obtain a continuous time series. These two sets of observations (visible in Fig 17), although being distinct, follow the same trends in all reflectance bands and indices.

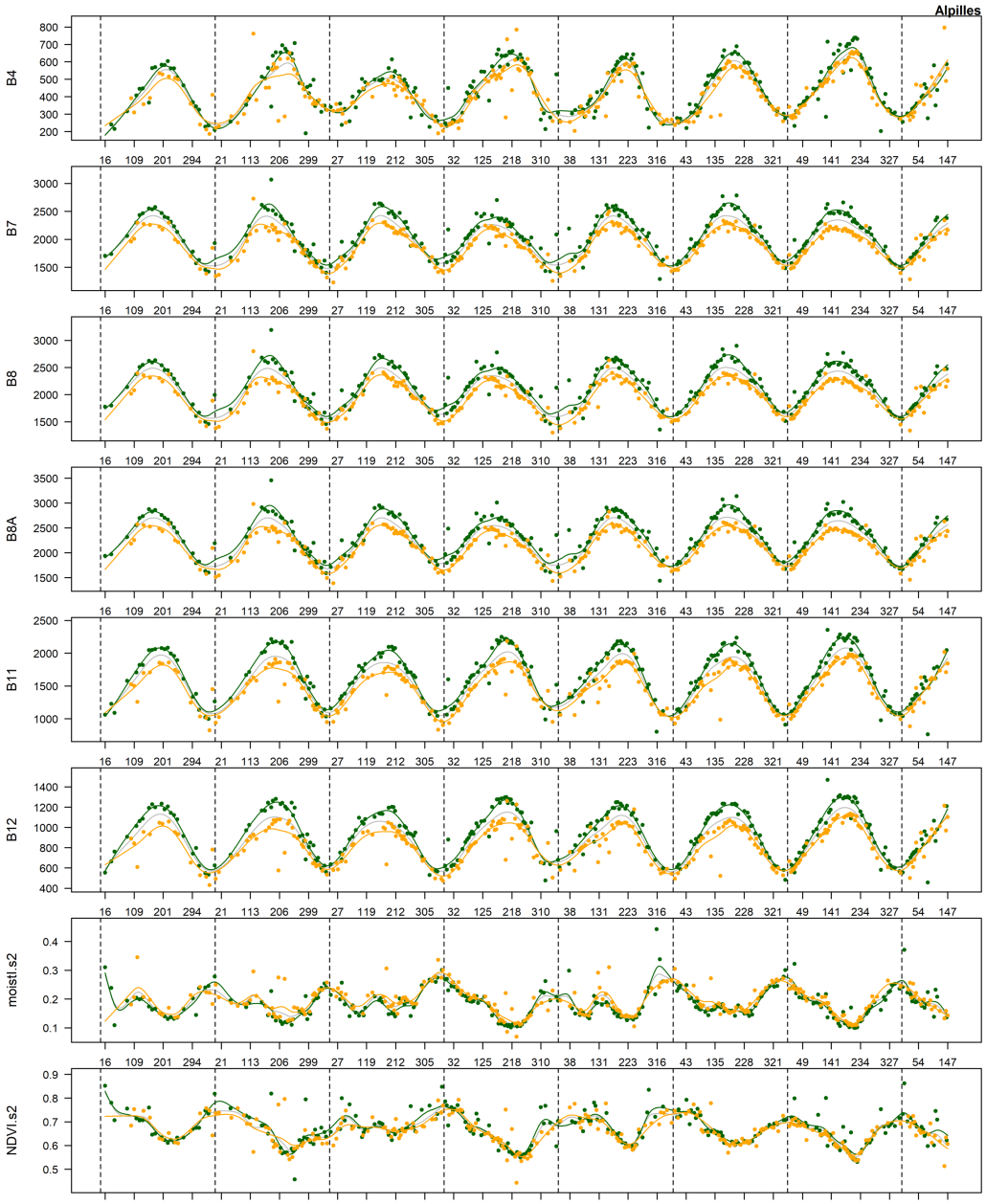


Figure 17: Temporal series and interpolations for Sentiel-2 products extracted to an example DFCI pixel located in the Alpilles mountains. The vertical dotted black lines correspond to January 1st of each year. Dots correspond to the mean observation value per DFCI pixel, yellow and green indicate the two sets of orbits (or viewing angles). Yellow and green lines are the corresponding temporal interpolations of the dots. The light grey line represents the temporal interpolation of all observations without the distinction of the viewing angles.

4.2. Article: Fire danger prediction enhanced through optical remote sensing in Mediterranean France

In the following paper (not yet submitted for review) we develop the work done for integrating optical remote sensing data in the Firelihood probabilistic framework. The study disentangles whether the spatiotemporal patterns of fire activity can be explained by fuel characteristics measured through remote sensing. Moreover, we aim to understand which information is brought by optical remote sensing and at which scale (large scale indicating spatial and/or temporal trends, or small scale indicating precise locations and time periods) or dimension (spatial, temporal or spatio-temporal).

Fire danger prediction enhanced through optical remote sensing in Mediterranean France.

Jorge Castel-Clavera, François Pimont, Thomas Opitz, Julien Ruffault, Albert Olioso, Victor Penot, Raquel Rodriguez-Suquet, Jean-Luc Dupuy

Abstract – there is not still an abstract, but the way it is done will much depend on the proposed journal. I do not think that this is needed for the thesis chapter.

1 Introduction

Fire activity is driven by a complex interaction between climate (and weather), vegetation, topography, human activities and land uses (Moritz et al., 2005). We know that weather is a major player on regional fire activities at daily to seasonal temporal scales (Abatzoglou and Kolden 2013, Barbero et al. 2015, Turco et al. 2017). Fire danger prediction systems are mainly weather-based processes, such as the indices of the Canadian Fire Weather Index System (CFWIS), a component of the Canadian Forest Fire Danger Rating System (CFFDRS), namely the Canadian Fire Weather Index (CanFWI; Van Wagner, 1987) and its moisture codes and fire behaviour indices: Drought Code (DC), Duff Moisture Code (DMC), Fine Fuel Moisture Code (FFMC), Initial Spread Index (ISI) and Build Up Index (BUI). Previous research has shown that the CanFWI shows superiority between several fire-weather indices on the prediction of fire activities (Castel-Clavera et al., 2025). The weather information can be separated in two categories, depending on the temporal scale at which a variable is effective. Short-term weather corresponds to the daily conditions captured through FFMC, wind speed and vapor pressure

deficit (VPD), between others. At longer term, the conditions varying at weekly or monthly scale are captured by surface drought (DMC) and deep drought (DC).

Here we focus on the factors of vegetation, which serves as the fuel for fires, and therefore its quantity (fuel load) and availability (fuel moisture content and dead/live fuel ratio) to burn will determine the fire occurrences and behaviour (surface, rate of spread, severity). Fuel load, which changes with phenology, and human activities exhibit strong spatial variation. Moreover, this variation is not explicitly accounted for in fire weather models. The vegetation moisture content, or live fuel moisture content (LFMC) is a critical factor for the combustibility of fuels, and usually estimated with weather proxies (see moisture codes from CFWIS mentioned above) rather than measured on the ground. Weather proxies have demonstrated to be adequate on their objective to capture the fuel availability at coarse spatial scale (Beverly & Wotton, 2007; Griebel et al., 2023; Nolan et al., 2022; Ruffault et al., 2018) but have shown their limits since they do not reflect the surface states. In fact, the resistance of plant species to drought differs between species and forest types (Pivovaroff et al., 2019; Ruffault et al., 2022; Decarsin et al., 2024) and then statistical relationship between LFMC, weather proxies, and ground LFMC requires to be calibrated for each site and species. Similarly, fire weather indices such as DMC and DC are good proxies for fuel consumption, but they do not account for local response of vegetation to weather, which also depends on soil properties. Finally, fire weather is drastically limited by its spatial resolution (often several km) and does not reflect local properties of vegetation.

Optical remote sensing (visible light, near infrared and short-wave infrared) is sensible to vegetation coverage and biomass, and its water content. One key feature of remote sensing that could become crucial for this study is its high spatial resolution, allowing for direct localized measures of the surface. The data extracted from these measures can be related statistically or physically to one or several biophysical traits.Within the optic electromagnetic spectrum range, specific wavelengths (spectral bands) interact differently with plant tissues, water, and soil (Wong, 2023). Direct measurements of spectral bands may already identify the presence of fuels, their water content, and even plant phenological stage or health state. Moreover, the combination of spectral bands into remote sensing indices is a method to isolate the signal of a searched trait. The NDVI (Normalized Deviation Vegetation Index) is a quantification of the green vegetation. This index combines Red and NIR (Near Infrared) spectral bands, establishing a difference between the leaf's greenness known through the Red (a greener leaf absorbs more red light, as this spectrum is used for the photosynthesis), and the photosynthetic activity known

through the NIR (higher photosynthetic activity reflects more NIR radiation). The NDMI (Normalized Deviation Moisture Index) is an indicator of vegetation coverage and plant hydric stress. This index combines NIR and SWIR (Short-Wave Infrared): it defines the difference between the dry matter content of the leaves known through the NIR as it interacts with the leaf's internal structures but not with water, and the water content of the vegetation known through the interaction of SWIR with water and the spongy mesophyll structure of leaves.

Remote sensing (RS) approaches are used in several parts of the world to address fire danger (Wang et al., 2013; Yebra et al., 2013; Nolan et al., 2016; Yebra et al., 2018; Jia et al., 2019; Tanase et al., 2022), nevertheless the most common approach is to focus on live fuel moisture content (LFMC) prediction or even fuel structure (Pettinari and Chuvieco, 2020; Gale et al 2021). The water content of live vegetation (or LFMC) has a direct effect on the reflectance of several wavelengths and can be directly estimated with remote sensors (Yebra et al., 2013). There are more sophisticated, but complex, techniques to obtain the LFMC from remote sensing information, as the physically based radiative transfer models (Wang et al., 2013; Yebra et al., 2018). Studies as Rao et al. (2020) found an advantage of the simpler empirical models over the complex radiative transfer models. On the other hand, dead fuel moisture content (DFMC) is almost only driven by the weather conditions, specifically the vapor pressure deficit (VPD).

The optical reflectance data is available through a large variety of satellites and sensors, which orbit the Earth at different frequencies and orbits giving a choice on spatial and temporal resolution (Pettinari and Chuvieco, 2020). Sentinel-2 data (2015-2020), which are easily available for operational applications have a spatial resolution of 10m and 20m (size of Sentinel pixels on the ground), and a temporal resolution of 5 or alternating 3 and 2 days, depending on the location on the earth. Hence, optical remote sensing can be viewed as a relevant way to measure fuel properties corresponding to temporal scales related to short-term and long-term drought and seasonal phenological variations. Moreover, the fine spatial resolution offers an opportunity to address spatial variations at a much finer scale than weather data and to account for local factors. As a result, optical reflectance data could bring additional information and could be a relevant alternative to drought indices such as DC and DMC. Few previous research (see for instance the Fire Potential Index (Burgan, 1998) has been found to use remote sensing data directly in fire danger prediction models. We hypothesize that the factors enabling to find the fuel quantity and moisture content should be good predictors of fire. We use an empirical statistical approach to disentangle de relationship of RS reflectance data (Sentinel2) and the fire regime, a method that has been claimed to be scarce but necessary (Chuvieco et al., 2020).

This study is included in the Bayesian probabilistic framework of *Firelihood*, continuing and enriching the work started by Pimont et al. (2021), and enhanced in Castel-Clavera et al. (2022 and 2025), which allows to account for predictors of different natures and operating at various spatio-temporal scales. *Firelihood* is based on the occurrence of fires larger that 1ha and their probability to exceed a preselected size threshold. Previous research with this framework has shown the importance of Land Use and Land Cover (LULC) on explaining yearly spatial fire patterns (Castel-Clavera et al., 2022), in other words, the spatial distribution of fire events is dependent on factors as the density of roads, Wildland-Urban interface (WUI), presence of agricultural areas, between other human factors that bring specific dangers and fire sources.Although the prediction capabilities and pertinence of *Firelihood* have been demonstrated in the earlier publications, we still face several limitations. Among these limitations there is a significant spatial effect (function of geographic coordinates) which accounts for observed variations that are not captured by explanatory variables. If reflectance data provide relevant information to explain high resolution spatial patterns of fuels, it is likely that their inclusion will reduce the magnitude of this spatial effect. Also, remote sensing data could improve knowledge on factors and scales of fire activity and improve the performance of models.

We aim to disentangle whether the spatiotemporal patterns of fire activity can be explained by fuel characteristics measured through remote sensing. As remote sensing data is expected to give more precise spatial measurements than weather methods currently used. One question arising, with important operational implications, is to determine if RS predictors will overperform the CFWIS long-term moisture codes (DC and DMC). If not, to which extent they can improve the model performance when they are added as co-variables to the widely used CFWIS indices. We moreover aim to understand which information is given by RS and in which scale (large scale indicating spatial and/or temporal trends, or small scale indicating precise locations and time periods) or dimension (spatial, temporal or spatio-temporal).

Later, we built a model combining our previous research on LULC and Weather with the *Firelihood* model (Pimont et al., 2021, Castel-Clavera et al., 2022, Castel-Clavera et al., 2025). Here the objective is to include not only the different spatiotemporal scales of fire weather (from daily to interannual) as in Castel-Clavera et al (2025) or of LULC, as in Castel-Clavera (2022), but rather to combine them with relevant reflectance data, which characterizes fuel variations related to drought, phenology, soil properties and species variations.

2 Materials and methods

Study area and fire data

Our study area corresponds to 15 French departments (NUTS3) from the Mediterranean France (south-east), which covers 75,560km^2 (Fig. 1). The region is characterized by the presence of human populations in all the area, the Rhone valley and the coastline are the most populated zones, and the variety of eco-regions from the alpine environment to the coastal forests. The area can be divided with two other criteria aside of the NUTS3: (i) SER – SylvoEcoRegions which correspond to areas of relatively homogeneous forest conditions with respect to the ecosystem functionality and forest management, and (ii) ZFC – Zones of homogeneous fire-climate, a subdivision of the NUTS3 administrative units in which fire-weather is considered as relatively homogeneous; this zonation is used by the forest managers.

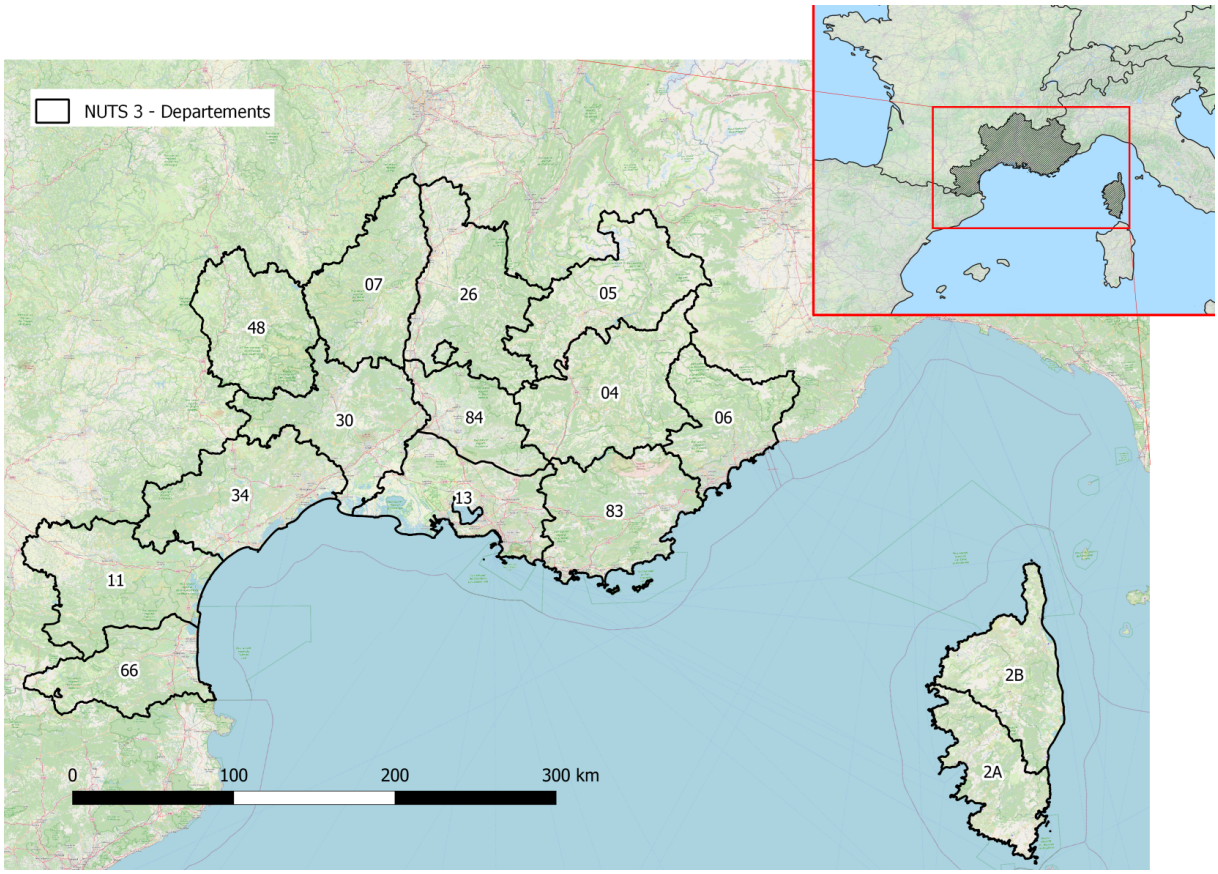


Fig. 1: Study area over the OpenStreetMap divided by the NUTS3 territorial units with their French identification number. The red square shows the position of the study area within the north-west Mediterranean basin.

Fire records were extracted from the Prométhée fire database, which covers all fire activity since 1973 (database available online http://www.promethee.com/). In this study we use the period from 2016 to 2022 to match Sentinel-2 available data in warm-season conditions (25 May to 31 October, Pimont et al. 2021). This fire dataset (Fig. 2) is given on a 2km pixel grid

(DFCI; Défense des Forêts Contre les Incendies) and contains more than 13.000 fire outbreaks, of which 3541 burnt more than 1ha and 181 more than 80ha. We will only consider the registered fires that reached at least 1ha of burnt area because the Promethée data is more exhaustive after 1 ha, some small fire incidents might not be recorded with the correct location and/or size, or not recorded at all. In addition, in south-eastern France there is a strong surveillance and early attack policy, therefore we choose to model the occurrence of fires that could bring significant risks and exclude those that happened under conditions that were not suitable for combustion and fire expansion, or fires that were extinct by the intervention of first fire responders.

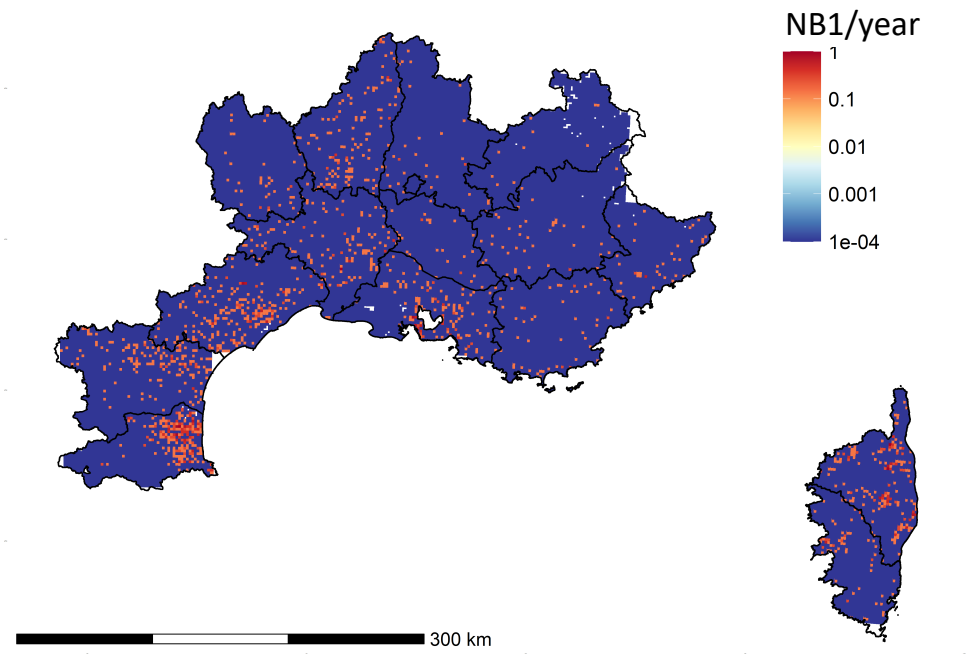


Fig. 2: Spatial distribution of all fire activities larger than 1ha that took place on the summer seasons from 2016 to 2022 (both included). Where NB1 is the number of fires >1ha represented on the DFCI pixel grid. Black lines correspond to NUTS 3 territorial units.

In addition of fire records in the 2km grid, we used fire perimeters provided by the French Forest Service (ONF) in the form of a dataset that registers in a shapefile vector layer the burnt area footprint (Fig. 3). The burnt surfaces registered in this dataset can be either obtained by remote sensing or GPS measurement, although some of them, usually for the very small events, the exact location and/or final burnt surface might be unknown. In some cases, the fire perimeter remains unknown. In this last case it is specified in the metadata.

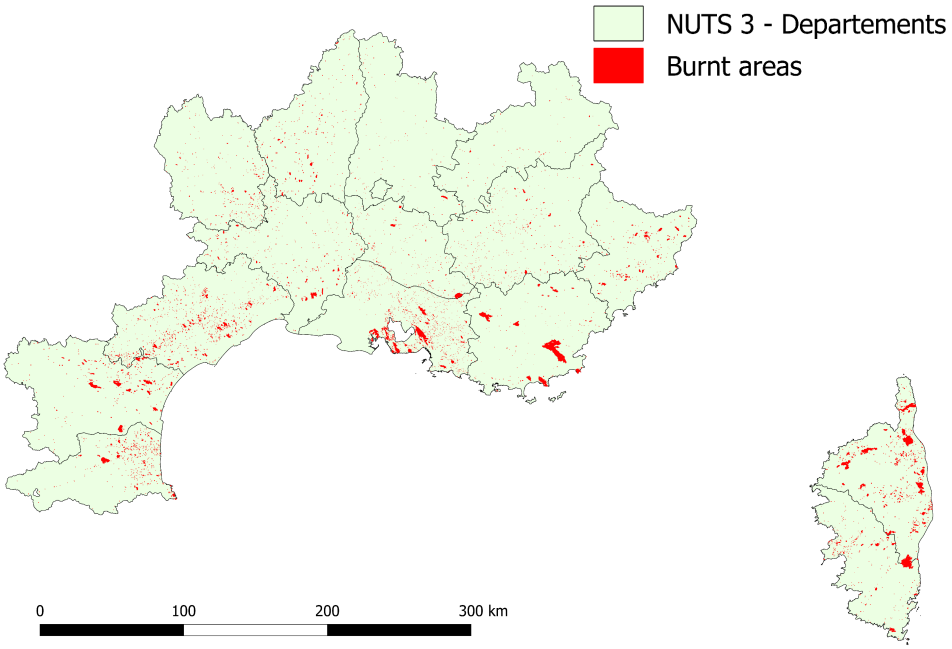


Fig. 3: Burnt areas (in red) database on Promethée database from 2014 to 2020.

In the study area, less than 5% of fires are caused by natural circumstances, the rest are human-caused fires including accident, arson, and negligence (Ganteaume and Jappiot., 2013). The distinction between natural-caused fires and human-caused fires is therefore not justified in the present study.

Data and pre-treatments

The Remote Sensing data for this study are Sentinel spectral bands at a spatial resolution of 5 days combining S2A and S2B. This data features B2 – blue, B3 – green, B4 – red, B5 – red edge 1, B6 – red edge 2, B7 – red edge 3, B8 – NIR, B8A – NIR2, B11 – SWIR and B12 – SWIR2. A cloud mask is applied to delete all pixels covered by clouds, keeping only pixels with visible ground observations, and all data was cropped using a mask of the forested areas in the study region. After applying both masks, the data is aggregated to 250m from its original pixel size of 10m or 20 m to reduce computational costs but keeping only the reflectance coming from vegetated surfaces. The spectral bands can be directly used as a predictor variable, or they can be numerically combined to form a variety of spectral indices. In this study we use the NDVI and the NDMI as they inform both about the vegetation health and the coverage. NDVI is used as a proxy of living biomass and as an indirect proxy of moisture content (González-Alonso, 2007; Yebra et al., 2013). The NDMI estimates the moisture content of living vegetation (Costa-Saura et al., 2021; Monteiro et al., 2024).

To account for the effects of weather over the fire activity we used the Canadian Fire Weather Index (CanFWI) and three subcomponents Drought Code (DC), Duff Moisture Code (DMC) and Fine Fuel Moisture Code (FFMC) from the Canadian Fire Weather Index System (CFWIS; Van Wagner, 1987) as formulated by Van Wagner (1974), as they gave the most accurate predictions of fire activities based on fire-weather in Castel-Clavera et al (2025). The weather data was taken from the SAFRAN meteorological reanalysis available for France (Vidal et al., 2010); this data base is provided in an 8km pixel grid at hourly scale. The indices were computed from these weather data using the R packages *cffdrs* (Wang et al., 2017). Finally, LULC data was derived from the Copernicus land monitoring service (https://land.copernicus.eu/en), National institute for statistics and economical studies (Institut national de la statistique et des études économiques, Insee) and French Forest Service databases (Office National des Forêts, ONF).

In order to fit the probabilistic model, explanatory variables related to satellite raster images have to be matched to the weather data and the fire database, so for each process (occurrence and size) we extracted the mean reflectance value for the concerned area as follows.

The occurrence dataset is based on the DFCI 2km pixel grid used in the Promethée database; each valid image is extracted to the 2km pixels using a mean zonal statistic (in other words, we assign to each pixel/day the mean value of the raster image within its borders). To avoid the appearance of outlier reflectance values we crop the extracted values to their 95% quantiles. Then we interpolated the values for the days without observations using GAMs (Generalised Additive Models) with a day counter as the only independent variable; this day count is fitted as a spline with the degrees of freedom equal to the number of observations -10, as it is empirically observed to give an accurate prediction.

Sentinel-2 images are extracted by tile then merged in a single temporal series before the interpolation. After the data treatment and extraction to the larger 2km pixels, we identified no significant difference between the two different azimuths from which Sentinel-2 satellites can see one same point, the spatial averaging of data already makes disappear the differences between the orbits, therefore all observation points are interpolated as one single time series.

In order to take advantage of the spatial resolution of RS data, we used fire perimeters to extract the relevant reflectance. For that purpose, the size dataset is based on the perimeters of burnt areas. For fires for which the perimeter was unknown, we extracted reflectance data from wildland areas located within a 100ha square around the centroid of the given fire. This rough

approximation will warranty that we are not reading satellite data from a random area, where there might not be even presence of forests, in the middle of a DFCI pixel, but instead we are reading the general status of the vegetation around the area of where the fire might have been potentially ignited. Aside of this precision, the extractions are made to the last valid observation before the fire occurred, using the same mean zonal statistic tool that was used for the occurrence grid.

Model fitting and verification.

Firelihood modelling framework

For modelling fire occurrence, as well as the probability of exceedance for a chosen size threshold we used the Bayesian probabilistic framework of *Firelihood* (Pimont et al. 2021; Castel-Clavera et al. 2022; Castel-Clavera et al., 2025). Similar to what was proposed in Preisler et al., (2004), *Firelihood* is a marked point process where points are stochastic and describe the occurrences (position and time) of a given fire with a minimal surface of 1ha; and its mark is the binary indication of the fire exceeding or not a threshold of 80ha. *Firelihood* includes spatio-temporal effects to capture the variability of fire activities that has not been explained by the predictor variables. The variable selection and model evaluation were carried out following the same approach proposed in Castel-Clavera et al. (2025), which rely on a base model corresponding to the fire activity baseline (FAb), in order to measure the explanation provided by each variable. The FAb is a model is informed only by random spatial and temporal effect, which are functions of pixel coordinates and the week of the year. The FAb simulates the mean seasonal fire activity for each DFCI pixel.

Modelling is conducted in two steps, first the stochastic points are modelled with the occurrence component (eq. 1), and then the exceedance component (eq.2) specifies the size threshold exceedance probability for each point. The set of predictor variables and functional relationships between predictor and response variables used within these two components is not necessarily the same, the effects of the predictor variables are estimated and optimized separately for the two components through the R-INLA package in the R software, a tool that leverages a numerically fast implementation of the Integrated Nested Laplace Approximation (Rue et al., 2017) to obtain accurate posterior estimations of the quantities of interest.

The occurrence component, defined in equation 1, models the number of fires larger than 1ha, ($N_{i,d}^{1ha}$) in a pixel i on day d with a Poisson distribution, whose intensity parameter $\lambda_{i,d}^{1ha}$ (i.e., the expected number of fires) is constructed as a function of LULC (Land Use and Land Cover),

WEATHER, RS (Remote Sensing) and STEs (Spatio-Temporal Effects). Inside the STEs we include (i) the pixel coordinates (X_i, Y_i) of fire count $N_{i,d}^{1ha}$ for the spatial effect consisting of a spatially structured multivariate Gaussian distribution based on the Matérn covariance function, where the so-called stochastic partial differential equation (SPDE) approach (Lindgren and Rue, 2015) is used to obtain numerically convenient representations; and (ii) the seasonal effect $\left(WEEK_{i,d}\right)$ encoded using the week number within the year.

$$log\,\lambda_{i,d}^{1ha} = log ß_{i,d} + \sum_j f_{LULCj}\left(LULCj_{i,d}\right) + \sum_j f_{WEATHERj}\left(WEATHERj_{i,d}\right) + \sum_j f_{RSj}\left(RSj_{i,d}\right) + f_{X,Y}(X_i, Y_i) + f_{WEEK}\left(WEEK_{i,d}\right) \qquad Eq.1$$

where $N_{i,d}^{1ha} \sim Poisson\left(\lambda_{i,d}^{1ha}\right)$ is the number of fires larger than 1ha, and β (equation 2) is the offset of wildland-area-per-pixel (multiplicative at the scale of $\lambda_{i,d}^{1ha}$ but here log-transformed to match the log-link function of the Poisson intensity $\lambda_{i,d}^{1ha}$); the f components are functions of predictors that can be nonlinear, estimated during model calibration using data and representing effects of the predictor variables in each category. The main mathematical output of the model ($log\lambda_{i,d}^{1ha}$) is the natural logarithm of the Poisson intensity parameters for each pixel and day in the study domain.

The fixed deterministic space-time varying offset, used in the occurrence component, consists of the surface of a pixel *i* on a day *d* covered by wildland (WA; in ha/pixel) divided by the average WA for all pixel-day combinations. It expresses if a given pixel has more wildland area than the average (ß>1), average wildland area coverage (ß~1) or less wildland area than the average (ß<1).

$$ß = \frac{WA_{i,d}}{\overline{WA}} \qquad Eq.2$$

The size component (equation 3) expresses the probability $p_{i,d}^{u}$ of a given fire – in a pixel *i* and day *d* – to exceed a burnt area of *u* (value in ha). The exceedance probability, modelled using the logistic link function, is a function of LULC, WEATHER, RS and STEs. In this case, the STEs include a spatially-conditionally autoregressive $f_{BESAG}^{p,u}$ component based on the 28 Sylvo-Eco-Regions (SER) that encodes dependence between spatially adjacent SERs, and the seasonal component is based on the week of the year as for the occurrence equation.

$$log\frac{p_{i,d}^{u}}{1-p_{i,d}^{u}} = 1 + \sum_{j} f_{LULC,j}(LULCj_{i,d}) + \sum_{j} f_{WEATHER,j}(WEATHERj_{i,d}) + \sum_{j} f_{RSj}(RSj_{i,d}) + f_{WEEK}^{p,u}(WEEK) + f_{SER}^{p,u}(SER) \quad Eq.3$$

Where the f components are the non-linear effects of the predictor variables from each Category. Using the estimated additive effects and inverting the transformation, we obtain the estimated probabilities $p_{i,d}^{u}$. We can then multiply them with the Poisson intensities $\lambda_{i,d}$ to obtain $p_{i,d}^{u}\lambda_{i,d}$, the expected number of fires $N_{i,d}^{u}$ with size larger than the threshold u for a given pixel-day. The size threshold we use in this study is set to u=80 (in ha unit), as this value was identified by Koh et al. (2023) (using statistical extreme-value theory) as the size where the distribution of fires shifts from a power law to the theoretically grounded generalized Pareto distribution in the studied area. For more detailed information on the *Firelihood* framework, we refer to Pimont et al. (2021) and Castel-Clavera et al. (2022 and 2025).

Fitted models combining Weather, Remote Sensing and LULC.

First, we establish two baseline models that will serve as a comparison standard for further informed models: (i) the "Null" model is the one without the effect of predictors, only the offset (for occurrence models); and (ii) the "Fire activity baseline" (FAb) is only informed with the STEs, and ß for the occurrence, representing an average spatial and temporal fire activity pattern. We then fit 2 weather-only models: (i) with CanFWI, and (ii) DC, DMC, FFMC, VPD and U for the occurrence component or DC, DMC, FFMC, Tmax and U for the size component, which showed to be the best performer models in Castel-Clavera et al. (2025). The remote sensing (RS) information is fitted in RS-only models to be compared directly to the FAb, and it is used in addition to weather in RS-Weather-aware models which will show the improvements of remote sensing over the previous model versions on fire danger prediction.

The most performant RS and Weather-aware model will also be fitted with LULC variables which have been verified in our previous research (Castel-Clavera et al., 2022). In addition to the wildland area (WA) already present in all other models, the LULC variables are agricultural area, coniferous forest area, broadleaved forest area, mixed forests area, shrubland area, urban area, wildland urban interface area (all areas are expressed in ha/pix), aspect (º azimuth), slope (%), a fuel classification from the French National Forest Office (5 empirical categories), road length (hm/pix), population (inhabitants/pix). These variables were added to the models with a parsimony criterion; for this study, we take the resulting combination of variables from (Castel-Clavera et al., 2022) and tested a large ensemble of variable combinations (see supplementary

materials B). Regarding RS data, we computed spatial and temporal correlations between reflectance bands, revealing high correlations between some of them (see supplementary materials A). For concision, we only tested explanatory variables representative of each group, namely B5 (Red edge), B8A (Near Infrared; NIR) and B11 (Shortwave Infrared; SWIR). The different models were ranked according to their performance and matching parsimony criterion to ensure that variables are adequate.

Table 1: Summary of predictor structures in the probabilistic models used in this study, based on combinations of terms for the occurrence component (eq. 1) and for the size component (eq.2). The "Null" model is a simple reference in which only the presence of vegetation is accounted for occurrence models; the second reference model is the fire activity baseline FAb informed only by adding spatial and seasonal effects (STEs) to the "Null" model. The models informed by FWIs and/or Remote Sensing (RS) are considered with and without STEs.

Fitting option	Occurrence 1ha (Eq2)	Size 80ha (Eq3)
Null	ß	1
FAb	ß + (X_i, Y_i) + $WEEK$	1 + SER + $WEEK$
Weather-only Model	ß + WA + $WEATHER$	1 + WA + $WEATHER$
Weather-only Model + STEs	ß + WA + $WEATHER$ + (X_i, Y_i) + $WEEK$	1 + WA + $WEATHER$ + SER + $f(WEEK)$
RS-only Model	ß + WA + RS	1 + WA + RS
RS-only Model + STEs	ß + WA + RS + (X_i, Y_i) + $WEEK$	1 + WA + RS + SER + $f(WEEK)$
RS and Weather-aware Model	ß + $LULC$ + $WEATHER$ + RS	1 + $LULC$ + $WEATHER$ + RS
RS and Weather -aware Model + STEs	ß + $LULC$ + $WEATHER$ + RS + (X_i, Y_i) + $WEEK$	1 + $LULC$ + $WEATHER$ + RS + SER + $f(WEEK)$

Verification, scoring and evaluation.

We followed the same methods for model verification and scoring as in Castel-Clavera et al. (2025), following the guidelines of Phelps and Woodford (2021), which consisted of assessing the predictive performance and parsimony of the model through the concept of 'proper scoring rules' (Gneiting and Raftery, 2007; Allard et al., 2012).

When selecting a model, a principle of parsimony must be favored, which refers to the model's simplicity in relation to its ability to accurately describe the data. We here use the DIC (Deviance Information Criterion), a generalization of the well-known Akaike Information Criterion for Bayesian models, to balance the predictive performance of each model must be balanced against its complexity, and therefore prevent overfitting. Then we will use Log-Loss

score to evaluate the predictive performance of the model tests, and to facilitate comparison of a score to the FAb score and measure improvement of a model with respect to this baseline, it is common practice to compute Relative Skill-Scores as shown in equation 4. The absolute values of Log-Loss score have no universal and easily interpretable lower or upper thresholds, which stands in contrast with the Area Under the Receiver-Operating-Characteristic Curve method (AUC-ROC; shown for comparison in results) that rates the model skill with a score between 0.5 and 1, where a value of 0.5 is associated with prediction completely at random and 1 with perfect prediction.

$$Skill_{Log-loss} = 1 - \frac{LogLoss_i}{LogLoss_{FAb}} \qquad Eq.4$$

The response functions of each predictor variable in a model represent the partial effect of the variable, providing a useful interpretation of the plausible behaviour of the variable. A good variable, regardless of the explanatory power that brings to the model, should give a monotonic effect through all its values and avoid saturations or abrupt changes in its slope.

2.5. Analysing scales at play in drivers of fire activities

The model development allows to quantify the effect of each factor on fire activity, but the model itself allows to go one step further, as it allows to disentangle the role of these factor in spatio-temporal scales at play in fire activities. For that purpose, we decomposed the contribution that each group of variables gives to the fire occurrence and size prediction following the methods used in Castel-Clavera et al. (2022). Thanks to the Poisson distribution, the simulations of the $\lambda_{i,d}^{1ha}$, $p_{i,d}^{80ha}$ and consequently the $N_{i,d}^{80ha}$ can be partitioned to obtain the relative contributions of Weather (later differentiating long-term weather and short-term weather), Remote Sensing, LULC (including β), STEs to the spatio-temporal distribution of fires (Eq. 5). On a second partition of variance done for the best performer model components, Weather is differentiated in two categories. This ensures that the comparison of Remote Sensing and Weather is done only between factors that drive the long-term fuel moisture content. On one hand, the variables that correspond to short-term (or daily) weather, namely wind speed, VPD, temperature and FFMC. And on the other hand, the variables of long-term weather, which include the Drought Code (DC) and the Duff Moisture Code (DMC).

Then, the contribution of a group of variables $C_{Group\ i}^{\mathrm{N}1ha}$ is defined as the variance of the group's partial $\lambda_{i,d}^{1ha}$ averaged to a given spatial or temporal scale; the same is applicable to $N_{i,d}^{80ha}$ and

$p_{i,d}^{80ha}$, since for very small values of $p_{i,d}^{u}$ we can assume that: $p_{i,d}^{u} \approx e^{log\frac{p_{i,d}^{u}}{1-p_{i,d}^{u}}}$ because $1 - p_{i,d}^{u} \approx 1$ if $p_{i,d}^{u} \approx 0$. Although fixed effects are accounted on the total prediction, their variance is very small and does not vary between models so, it is not displayed on the results corresponding to this section.

$$C_{Group\ i}^{\mathrm{N1ha}} = \overline{var_i\left(log\left(\lambda_{i,d}^{1ha\ Group\ i}{}_{scale}\right)\right)}^{scale} \qquad Eq.5$$

Where $scale$ is the spatio-temporal dimension at which the prediction is aggregated. The contribution of each group is evaluated at four scales: (i) spatio-temporal, (ii) spatial, (iii) seasonal and (iv) interannual. A simple aggregation at the ZFC/day scale is done to obtain the spatio-temporal variances. In this way the contribution by group of variables is given at the same spatio-temporal scale used for model evaluation. To obtain the contribution by group of variables to the spatial distribution of fire activity, the predictions are aggregated by DFCI pixel as the sum of the full temporal series within a given pixel. To obtain the contribution by group of variables to the seasonal distribution of fire activity, the predictions are aggregated by week, as the sum of all the fire activities within a given week accumulated through the years. This aggregation gives the total fire activities of the Promethée area by week. Finally, to obtain the contribution by group of variables to the interannual variation of fire activity, we do an aggregation by year, accounting for the total fire activities of the season in all the study area.

3 Results

Model fits

We first compared the performance of the fitted models based on the Log-loss skill score. A selection of models is shown in tables 2 and 3, the rest of them in supplementary materials B. The Skill score values where normalized to the "skill-scale" from 0 being the skill scale of reference corresponding to the Fire Activity baseline (FAb), to 100 which corresponds to the best performant model of the study; the skill-scale is not bounded below 0. From table 2 we can observe that the null model on the occurrence component (O-Null) has a skill scale of ≈ –140. This means that the spatio-temporal patterns on the occurrence's FAb (O-FAb) capture 60% of the spatio-temporal variability of the fire activity, and only the remaining 40% is explained by the explicit fire predictors. Similarly, table 3 shows that the skill scale of the null model on the size component (S-Null) is ≈ –75. So, the spatio-temporal patterns on the size's FAb (S-FAb) bring 45% of the spatio-temporal variability of large fire activities.

As already shown in Castel-Clavera et al. (2025), models with separated weather indices systematically overperform the same model CanFWI. Similarly, models with spatio-temporal effects (STE) systematically overperform the same model without STEs. Beyond these known features, the gain associated with the different RS predictors can be analysed. Every model of occurrence (Table 2) and size (Table 3) which does not include RS predictors can be improved by the addition of some RS specific predictors, which can be either B8A, B11, NDVI or NDMI. The occurrence models including only RS information are on the lower rankings of skill scale with a large gap with the worst performant weather-only model, e.g., NDVI and CanFWI have almost 50 units of skill score of difference (Table 2), 30 units when these models include STEs. As for the occurrence component, the RS variables are on the lower rankings of skill scale, below the weather-only models (Table 3).

Improvement associated with RS variables in the performance of occurrence models is limited. The improvement in the skill scale given by B8A+B11 for the CanFWI model is around +7 and for CanFWI + STEs is the improvement from B8A+B11 is +3. Moreover, the gain obtained from the RS variables is reduced when considering the combination of the separated components of the CanFWI (DC+DMC+FFMC) with VPD and wind speed. For the size model component, the Remote Sensing data makes a larger improvement (Table 3), reaching 22 in skill scale for the B8A(NIR2), the best performant RS-only model without STEs. For the size component, the improvement brought by the combination of B8A+B11 on the CanFWI model is +20 in skill scale, and +16 if the models include the STEs component. The STEs for the size model are not as fine as in the occurrence so, in this case, the difference of including the mentioned spatial and temporal effects is weaker.

Regarding the addition of Land Use and Land Cover (LULC) variables, the models with WEATHER, RS and LULC are the best performers. For the models including STEs, the LULC information increases the performance by 2 to 4 units of skill Scale. In terms of parsimony (DIC), the most performant model for the occurrence component is the combination of DC * DMC * FFMC * VPD * U * B8A * B11 * LULC * STEs (from now on OM-1). For the size component, the most performant model is the combination of DC * DMC * FFMC * Tmax * U * B8A * B11 * LULC * STEs (from now on SM-1). These models (OM-1 and SM-1) will be described with more detail in the following.

Table 2: Model evaluation (at ZFC/day scale) metrics for the 1ha occurrence component. Names in the table are as follows: LogLoss is the raw output from the score formula and Skill Scale is the normalization of the Skill Score values between the 0 and the maximal performance value, AUC is the value from an AUC-ROC based on the model fit, DIC is the parsimony criterion.

Occurrence 1ha				
model	**LogLoss**	**Skill Scale**	**AUC**	**DIC**
NULL	0,000885	-140,4	0,591	41264
B8A (NIR2)	0,000800	-29,2	0,570	37123
NDMI	0,000789	-14,2	0,627	36606
B11 (SWIR1)	0,000788	-13,1	0,630	36547
B5 (Red edge 1)	0,000788	-12,8	0,626	36553
NDVI	0,000787	-11,8	0,626	36525
Fab	0,000778	0,0	0,771	35365
CanFWI (proxy)	0,000750	36,9	0,780	34781
CanFWI (proxy) * B8A	0,000748	39,8	0,785	34682
CanFWI (proxy) * NDMI	0,000746	42,1	0,791	34583
CanFWI (proxy) * NDVI	0,000745	43,6	0,793	34527
CanFWI (proxy) * B8A * B11	0,000745	44,0	0,794	34527
DC * DMC * FFMC * VPD * U	0,000742	47,6	0,802	34432
DC * DMC * FFMC * VPD * U * NDMI	0,000739	51,3	0,808	34279
DC * DMC * FFMC * VPD * U * NDVI	0,000738	52,5	0,810	34237
DC * DMC * FFMC * VPD * U * B8A * B11	0,000738	52,6	0,811	34231
DC * DMC * FFMC * VPD * U * B8A * B11 * LULC	0,000734	57,4	0,819	34063
B8A (NIR2) * STEs	0,000733	59,6	0,828	33800
B11 (SWIR1) * STEs	0,000730	63,4	0,835	33700
B5 (Red edge 1) * STEs	0,000730	63,9	0,836	33670
NDMI * STEs	0,000729	64,6	0,837	33661
NDVI * STEs	0,000727	66,8	0,839	33588
CanFWI (proxy) * STEs	0,000707	93,9	0,872	32521
CanFWI (proxy) * B8A * STEs	0,000706	95,3	0,874	32446
DC * DMC * FFMC * VPD * U * STEs	0,000705	96,0	0,878	32215
CanFWI (proxy) * NDMI * STEs	0,000705	96,0	0,876	32430
CanFWI (proxy) * NDVI * STEs	0,000705	96,6	0,877	32404
DC * DMC * FFMC * VPD * U * NDMI * STEs	0,000705	96,6	0,881	32162
CanFWI (proxy) * B8A * B11 * STEs	0,000704	96,9	0,876	32379
DC * DMC * FFMC * VPD * U * B8A * B11 * STEs	0,000704	97,2	0,881	32116
DC * DMC * FFMC * VPD * U * NDVI * STEs	0,000704	97,4	0,881	32128
DC * DMC * FFMC * VPD * U * B8A * B11 * LULC * STEs	0,000702	100,0	0,885	31839

Table 3: Model evaluation metrics for the 80ha size component Names in the table are as follows: LogLoss is the raw output from the score formula, Skill Scale is the normalization of the Skill Score values between the 0 and the maximal performance value, AUC is the value from an AUC-ROC based on the model fit, DIC is the parsimony criterion.

Size 80ha				
index	**LogLoss**	**Skill Scale**	**AUC**	**DIC**
NULL * perimeters	0,2201	-73,244	0,500	744
B11 (SWIR1)	0,1693	-2,380	0,694	714
NULL * perimeters * STEs	0,1676	0,000	0,715	718
B5 (Red edge 1)	0,1649	3,784	0,713	707
NDMI	0,1597	11,074	0,723	704
NDVI	0,1566	15,301	0,741	697
B8A (NIR2)	0,1519	21,855	0,756	684
B11 (SWIR1) * STEs	0,1462	29,852	0,793	689
NDMI * STEs	0,1456	30,696	0,789	689
CanFWI (proxy)	0,1455	30,785	0,774	667
B5 (Red edge 1) * STEs	0,1452	31,238	0,796	687
NDVI * STEs	0,1445	32,266	0,796	684
CanFWI (proxy) * NDMI	0,1413	36,611	0,780	664
CanFWI (proxy) * NDVI	0,1407	37,502	0,786	659
CanFWI (proxy) * STEs	0,1348	45,780	0,825	655
B8A (NIR2) * STEs	0,1344	46,274	0,816	659
CanFWI (proxy) * NDMI * STEs	0,1327	48,620	0,830	655
CanFWI (proxy) * NDVI * STEs	0,1327	48,701	0,830	651
CanFWI (proxy) * B8A	0,1326	48,864	0,815	638
DC * DMC * FFMC (proxy) * Tmax * U	0,1325	48,928	0,819	654
CanFWI (proxy) * B8A * B11	0,1314	50,540	0,821	639
DC * DMC * FFMC (proxy) * Tmax * NDMI * U	0,1309	51,138	0,824	655
DC * DMC * FFMC (proxy) * Tmax * NDVI * U	0,1302	52,100	0,828	650
CanFWI (proxy) * B8A * STEs	0,1242	60,483	0,851	628
DC * DMC * FFMC (proxy) * Tmax * U * STEs	0,1235	61,476	0,862	639
CanFWI (proxy) * B8A * B11 * STEs	0,1231	62,074	0,856	629
DC * DMC * FFMC (proxy) * Tmax * U * B8A * B11	0,1226	62,808	0,854	630
DC * DMC * FFMC (proxy) * Tmax * U * NDMI * STEs	0,1220	63,635	0,867	640
DC * DMC * FFMC (proxy) * Tmax * U * NDVI * STEs	0,1218	63,823	0,867	637
DC * DMC * FFMC (proxy) * Tmax * U * B8A * B11 * STEs	0,1144	74,246	0,886	617
DC * DMC * FFMC (proxy) * Tmax * U * B8A * B11 * LULC	0,1027	90,508	0,906	582
DC * DMC * FFMC (proxy) * Tmax * U * B8A * B11 * LULC * STEs	0,0959	100,000	0,926	568

Response functions for Remote Sensing

The OM-1 included the CFWIS drought indices, vapor pressure deficit (VPD) and wind speed (U) as weather information (Fig. 4 D, E, F, G and H). The remote sensing variables that maximized the model performance were the Near Infrared (NIR or band B8A; Fig. 4J) and the Short-wave Infrared (SWIR or band B11; Fig. 4I), both components of the NDMI. NIR and SWIR have effects of opposite slope with similar amplitudes. SWIR has an increasing effect, as the reflectance of SWIR is inversely proportional to water presence. NIR is an indicator of photosynthesis and live biomass, its decreasing effect indicates that higher level of vegetation

greenness and/or live biomass brings lower danger. The LULC variables retained through the parsimony criterion are Wildland area (WA; Fig. 4A), Road length per pixel (Fig. 4B) and the fuel classification from the ONF (Fig. 4C). The WA (Fig. 4A) shows a maximal multiplicative effect between 100 and 200 ha/pixel, below those values the fuel might be too sparse, and over 200ha the results infer that there is low density of ignition sources. The seasonal effect (Fig. 4K) was negligible. The spatial effect of the O-FAb model (Fig. 5A) followed the distribution of observed fires as expected (Fig. 2). The spatial effect of the most performant model (Fig. 5B) is notably reduced compared to the O-FAb (Fig. 5A), meaning that major spatial information was captured by the explicit variables. Nonetheless, this spatial model has still several hot spots, with the higher effect concentrated on the western side of the Rhone valley and the island of Corsica.

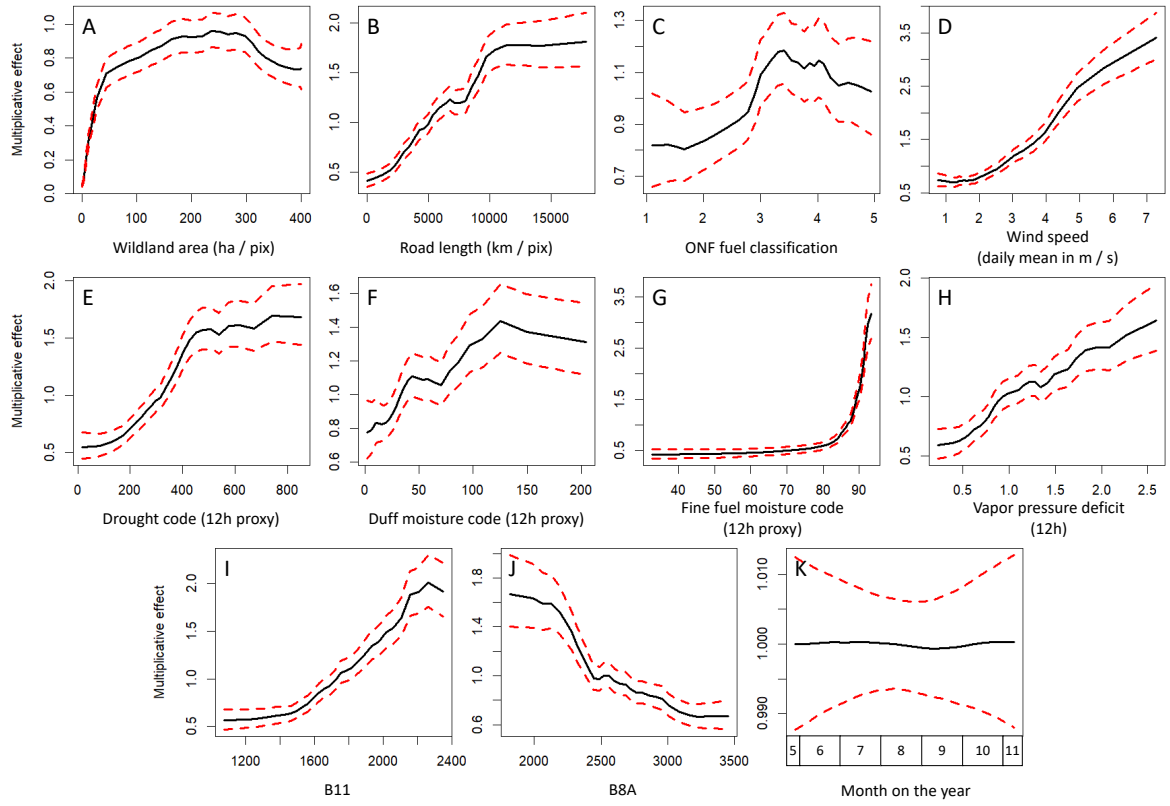


Fig. 4: Effects of the occurrence component fitted with Wildland Area per pixel (A) road length per pixel (B), ONF fuel classification (C), Wind speed (D), DC, DMC and FFMC daily values from the CFWIS (E-G), Vapor pressure deficit (H), SWIR1 (G), NIR2 (J) and STEs (plot K shows the temporal effect). The X axis shows the value in each of the variable-specific bins; the Y axis is the value of the multiplicative effect, in other words, the factor by which the expected number of fires changes with respect to the case of a multiplicative effect of 1.

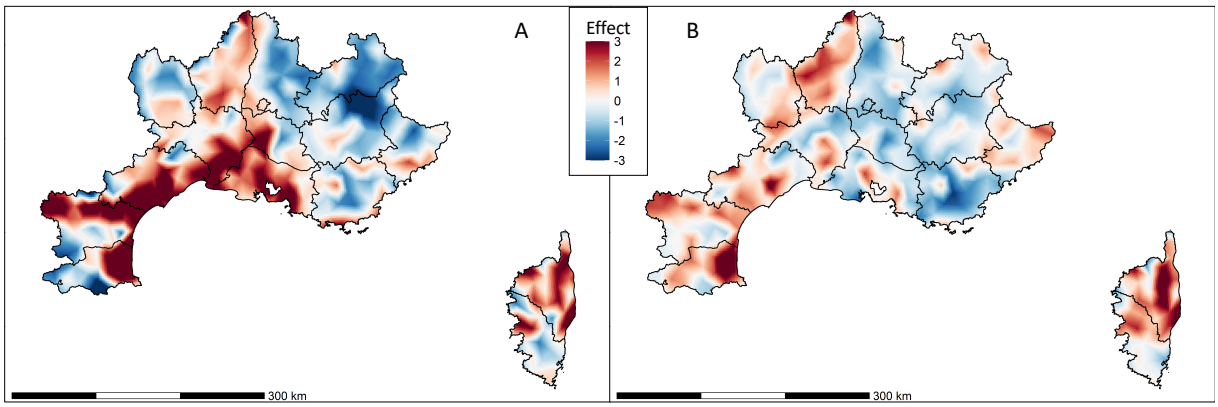


*Fig. 5: Spatial effect for two occurrence models: (A) Fire Activity baseline (FAb) for occurrence component; (B) Best fitted model for occurrence component (DC*DMC*FFMC*VPD*U*B8A*B11). This map shows that for the same level of predicted danger in given pixel and day, the fire probability will be higher or lower depending on the local environmental conditions.*

In the SM-1included the CFWIS drought indices, daily maximal temperature (Tmax), and wind speed (U) as weather information (Fig. 6 I, J, K, L and M). The weather variables of the SM-1 have larger confidence intervals than the OM-1, although the responses follow similar shapes. Aside from the wildland area (WA), SM-1 is fitted with 7 LULC variables (Fig 6 B, C, D, E, F, G and H), which were chosen with the parsimony criterion. WA (Fig. 6A) shows its peak effect at approximately 250 ha/pixel, similarly to the occurrence model, these results indicates that large fire danger increases as more fuel (forest area) is available, but once the wildland area occupies more than half of the pixel the effect either saturates or decreases. The remote sensing components of the OS-1 are the NIR (Fig. 6N) and SWIR (Fig. 6O), both showing decreasing response functions. Although the response of SWIR is not highly significant, it reveals to be important for better model performance in terms of both Skill Scales and parsimony. The seasonal effect (Fig 6P) has a weak effect that decreases over the fire season. This model has a large number and variety of predictor variables among which the seasonal trends might be already captured.

The spatial effect of the size component was estimated at the scale of Sylvo-Eco-Regions; the S-FAb (Fig. 7A) had a pronounced difference between western (lower effect) and eastern zones (higher effect). Nonetheless, the magnitude of the overall effect is small compared to all the other variables of the SM-1. In comparison, the spatial effect for SM-1 (Fig. 7B) was raised in the mountain regions of the north-east (Alps) and Corsica. This effect was lowered in the south-east, namely in the departments of Bouches-du-Rhone and Var, and a fraction of Vaucluse and Alpes-de-Haute-Provence.

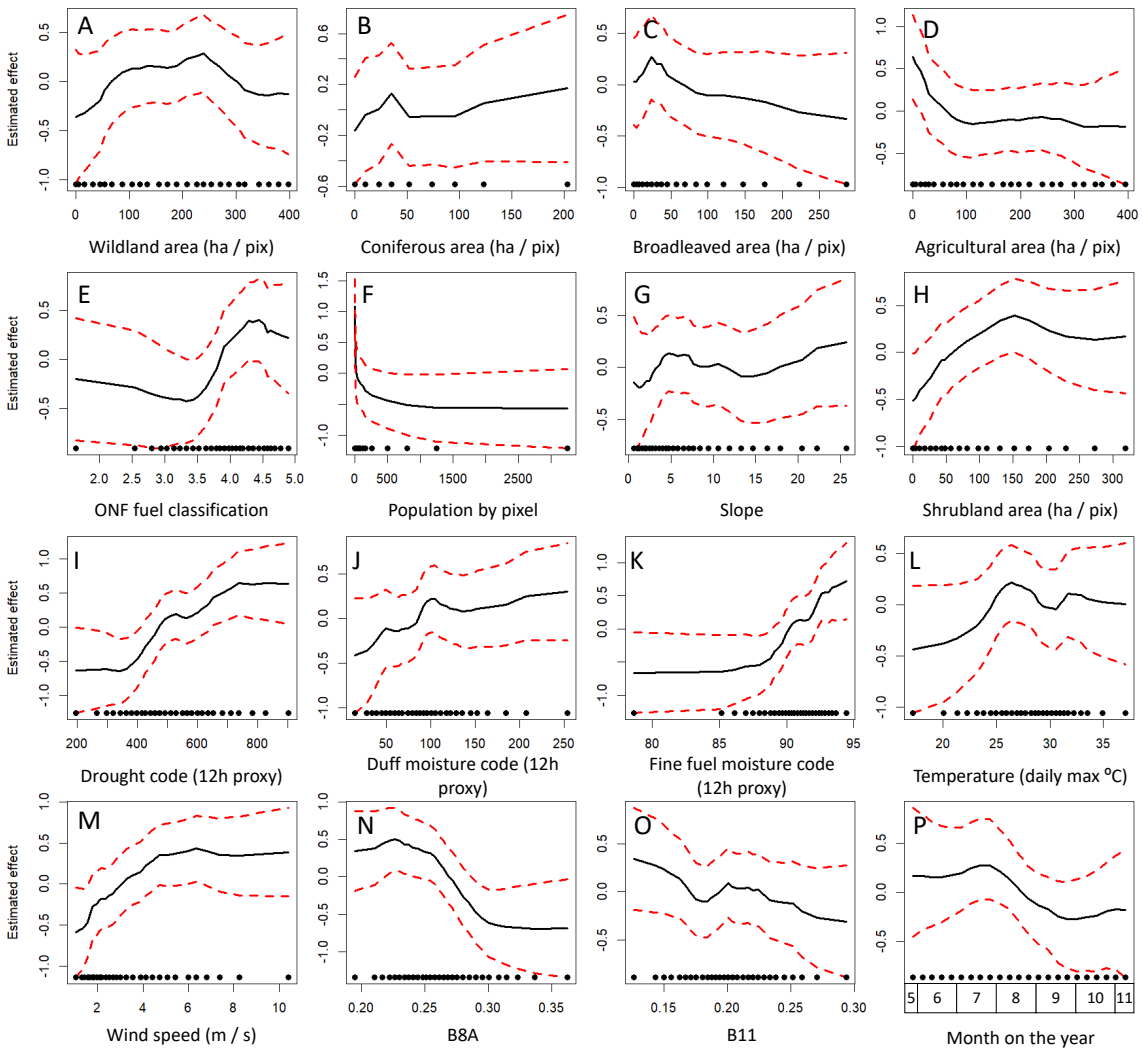


Fig. 6: Effects of the size component fitted with Wildland Area per pixel (A), DC, DMC and FFMC maximal daily values from the CanFWI (B-D), maximal daily air temperature (E), Wind speed (F), B8A (G), B11 (H) and STEs (plot I shows the temporal effect).The X axis shows the value in each variable-specific bin; the Y axis is the contribution of the predictor variable's value to the logit of the probability of exceedance.

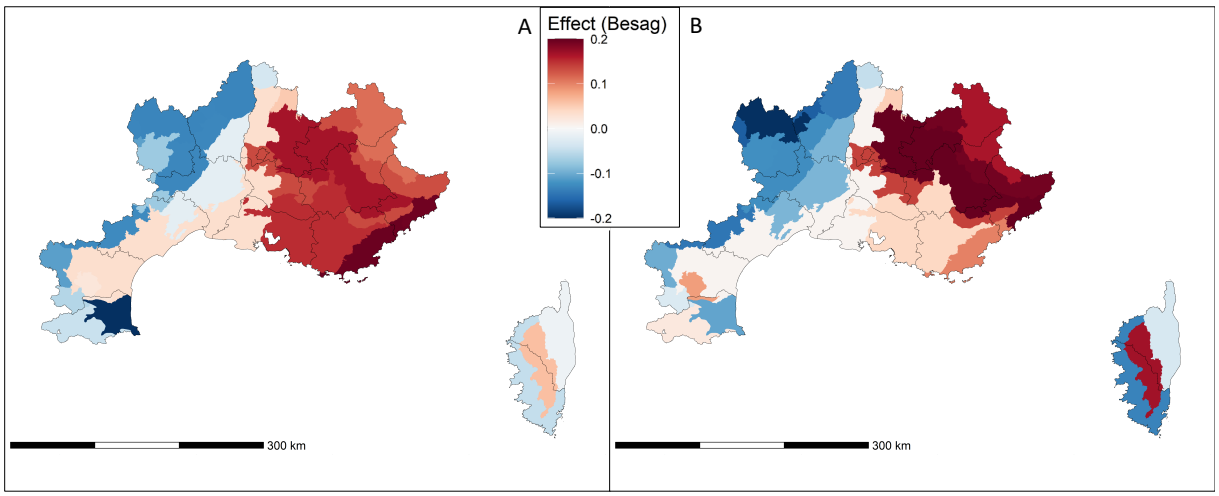


Fig. 7: Spatial effect for two size models: (A) Fire Activity baseline (FAb) for size component; (B) Best fitted model for size component. This map shows that for the same level of predicted danger for a given day, the fire probability will be higher or lower depending on the local environmental conditions.

Partition of simulations

The partition of the simulations shows the contribution of each group of variables to the total variance of the fire danger prediction. We separated the variance in three spatio-temporal dimensions: (i) spatial dimension, (ii) seasonal dimension and (iii) interannual dimension. The first explores the contribution of the variables to the spatial distribution of fire activities. The second and the third dimensions explore the contribution of the variables to the temporal distribution of fire activities at two scales. The seasonal dimension does this at weekly scale, exploring the variability through the summer season. The interannual dimension explores the changes of fire activities between seasons.

This section shows the results for 8 model fits, 4 without STEs and their homologues with STEs. For this section and following Table 1, Model 1 is fitted following the *RS-only Model,* with NDVI as RS predictor variable (including WA). Model 2 is fitted as CanFWI(proxy) * NDMI, corresponding to the *RS and Weather-aware Model* structure with only WA in the LULC category. Model 3 is fitted as DC * DMC * FFMC * VPD (or Tmax) * U * B8A * B11 (Tmax in the size component), corresponding also to the *RS and Weather-aware Model* structure with only WA in the LULC category. Model 4 is fitted as Model 3 but includes further LULC predictor variables. The LULC variables included in Model 4 for the occurrence and the Size components are respectively specified in Fig. 4 and Fig. 6. Models 5 to 8 correspond to models 1 to 4 with added STEs. Therefore, model 8 corresponds to OM-1 and SM-1.

For the spatio-temporal dimension of the occurrence model (Fig. 8A), the remote sensing variables represent a small part of the total variance, especially when combined with further predictors. When included, weather variables take the largest part of the total variance on models without STEs, and on the Models 7 and 8. In general, the absolute variance brought by RS to 1ha occurrence models is reduced when some weather information is added. LULC variance is rather stable through Models 1 to 4 and 5 to 8, with a small increase when variables other than WA are included. The interannual variation (Fig. 8B) is very low and brought almost exclusively by fire weather. On the seasonal dimension of the occurrence component (Fig. 8C), weather is the only group of variables that remain significant. The absolute variance is largely reduced compared to the spatio-temporal variance. The seasonal component of the models seemingly did not have a long enough time series to create a strong effect. Finally, on the spatial dimension (Fig. 8D), the distribution of the variance is similar to the spatio-temporal case, only weather variance is reduced. RS variables did not lose their influence although it remained minor. The total spatial variance was slightly lower that the spatio-temporal variance.

As seen before, the threshold exceedance probability prediction benefits more than the occurrence model from remote sensing information. Fig. 8E shows a larger proportion of spatio-temporal variance explained by RS on P80 models including weather information. In Models 6 and 7, 1/3 and 1/4 respectively of the total variance is explained by the RS variables (NIR2 and SWIR1). The absolute variance brought by RS does not drop when the weather information is complexified. LULC variance is increased by near 3 times for Models 4 and 8, which include all LULC variables. As well, the inclusion of more complex LULC information induces a decrease on the absolute variance brought by the RS variables. In this case (Models 4 and 8) the contribution of the RS variables is between the 10 and 15%. Still, the interannual variation of the size models is very small and brought by the fire weather (Fig. 8F). In the seasonal dimension (Fig. 8G) each model has the total variability largely reduced, and variance is mainly brought by fire-weather with a small contribution of the remote sensing. Models with *RS-only* structure (Models 1 and 5) show very low variance, only when combined with weather the remote sensing variables brought almost the 20% of the seasonal variance. The variance brought by the STEs to the P80 model is concentrated on the seasonal dimension. The Besag component based on the SERs had a small contribution on the 80ha-threshold exceedance probability. On the spatial dimension (Fig. 8H) the variance of the RS group was reduced by more than half in all models, and the variance of the Weather group was reduced by a small fraction. The variance of the LULC group kept its absolute value on the spatial dimension, showing as for the N1 that its significance is solely spatial.

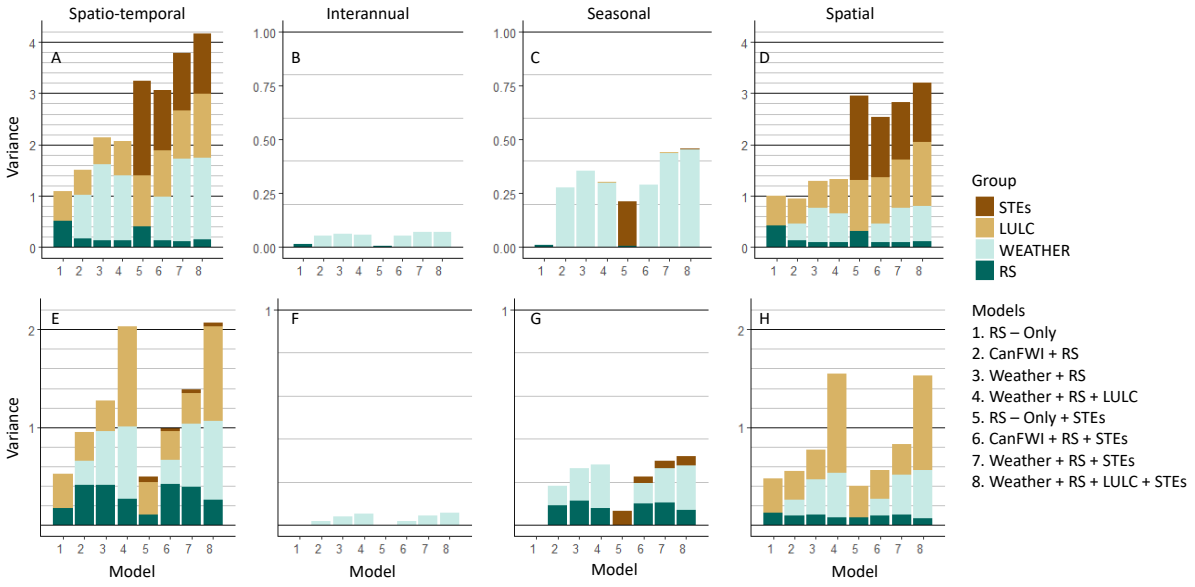


*Fig. 8: Variance of model predictions partitioned (excluding fixed effects for clarity, its variance is constant and smaller than 0.01) for occurrence component (plots A, B, C and D), size component (plots E, F, G and H). Where RS-Only are models without fire-weather information, featuring only the NDVI; CanFWI + RS models combine the effects of CanFWI(proxy) with the remote sensing variables B8A and B11; Weather + RS models correspond to the combination of DC * DMC * FFMC **

*VPD * U * B8A * B11 for occurrence models and DC * DMC * FFMC * Tmax * U * B8A * B11 for fire size threshold exceedance probability; Weather + RS + LULC models correspond to the Weather + RS models with the LULC variables that maximized the performance and parsimony of the occurrence and size components. All models are fitted with WA.*

Reflectance bands and indices show the seasonal variations in vegetation conditions. Weather data also incorporates components with significant daily variations (wind, FFMC, VPD, T) in addition to fuel-related components (e.g., DC, DMC) changing on seasonal scale. To ensure a fair comparison between optical Remote Sensing (RS) and Weather factors in terms of contribution to the prediction variance, the Weather components of the best fitted model have been divided in short-term (U, VPD, FFMC) and long-term (DC, DMC) components. Results show that the short-term weather is the dominant part of the Weather in the occurrence models (N1) as seen in Figure 9 (A, B, C and D). DC and DMC account for the ~25% of the total Weather variance in the spatial and seasonal dimensions (Fig. 9 B, C) and less than 50% on the interannual dimension (Fig. 9D). In the spatio-temporal dimension (Fig. 9A) and especially in the spatial dimension (Fig. 9B), RS and long-term weather show roughly the same contribution to the total variance. Nonetheless, in the seasonal and interannual dimensions (Fig. 9 C, D) the RS variance remains very low. Long-term weather accounts for more than 25% of variance in the seasonal dimension (Fig. 9C) and just under 50% in the interannual dimension (Fig. 9D).

On the probability of 80ha-threshold exceedance (P80) models, the long-term weather represents the largest share of total Weather variance (Fig. 9 E, F, G and H). For the spatio-temporal dimension (Fig. 9E) the DC+DMC represent the same variance as the RS variables. For the spatial dimension (Fig. 9F) the long-term weather brought largely more variance than the RS variables. In the seasonal dimension (Fig, 9G) the long-term weather variables accounted for one third of the total prediction variance, well above the ~25% accounted by the RS variables. In this case, short-term weather accounts for the ~25% of variance. The interannual variability of the P80 (Fig. 9H) is explained mainly though DC and DMC (~75%), and short-term weather (~25%). Finally, LULC play a larger role in the P80 than the N1 (Fig. 9 A, E) although its variance is only brought in the spatial dimension for both model components (Fig. 9 B, F).

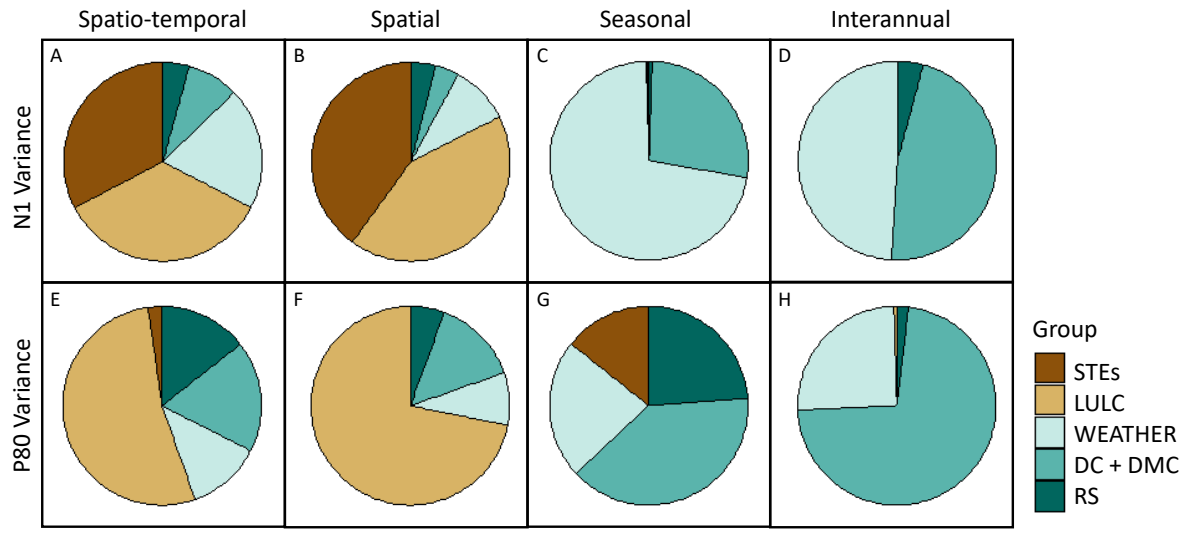


Fig. 9: Partition of variance for the most performant model of each component (OM-1 in boxes A to D and SM-1 in boxes E to H) shown for 4 dimensions Spatio-temporal (boxes A and E), Spatial (boxes B and F), Seasonal (boxes C and G) and Interannual (boxes D and H). Weather has been subdivided in two components, one for the long-term moisture codes (DC + DMC) and one for the short-term weather (WEATHER) which includes all other weather variables form each model component.

4 Discussion

This study had the objective to better explain the observed spatio-temporal patterns of fire activity in South-eastern France by the introduction of optical remote sensing information in a probabilistic modelling framework. Otherwise, part of the variability of these patterns is explained by random spatio-temporal effects included in the model. The main hypothesis is that the optical remote sensing observations will inform about the status and quantity of fuels. This hypothesis is based on previous literature (Guk et al., 2024, Tanase et al. 2022, Chuvieco et al. 2020, Yebra et al. 2018, Yebra et al., 2013). Recent publications show the potential of using optical remote sensing to predict fuel moisture content, and later fire danger (Yebra, et al., 2018). Furthermore, Chuvieco et al. (2020) mentioned the challenges that arise from RS for fire danger applications, for example, the remote sensing information may come from a variety of sensors and types of captured information. In addition to the need for greater efforts on statistical validation approaches, which is addressed in our study. We used the *Firelihood* framework (Pimont, et al., 2021, Castel-Clavera, et al., 2022) and the advances on model evaluation later developed (Castel-Clavera, et al., 2025) to test the explanatory power of optical remote sensing spectral bands as predictors of fire danger. The remote sensing variables used in this study correspond to the optical reflectance bands from Sentinel-2.

Response of weather, landscape and optical remote sensing variables on fire activity prediction.

In our study, the results show that optical remote sensing data improves the performance of models based on weather and wildland area occupation. Due to the temporally reduced dataset for Sentinel data in comparison to our previous research (Pimont et al., 2021; Castel-Clavera et al., 2022) the spatial and temporal effects have been modified to a smother form. The grid for spatial effect in the occurrence component has less dense triangulation and the seasonal effect for both model components have less sensitivity. Therefore, the influence of the STEs was smaller for this study with Sentinel remote sensing data, which does not mean that they are less important to consider in model simulations.

The occurrence models show a small improvement while the size models benefit largely from the studied remote sensing variables. Further exploration has shown that the occurrence component saturates of information when Weather, RS and all LULC variables are included. Therefore, we have limited the number of LULC following the parsimony criteria. This result illustrates the potential limitations of a probabilistic approach when the size of the dataset (here the number of years) is limited.

The OM-1, most performant combination for the occurrence component, featured weather, LULC and remote sensing information, paired with the STEs. The selected RS variables are the SWIR and NIR spectral bands, both components of the NDMI. Those spectral bands interact with the water and the dry matter of plants respectively. The selected LULC variables have shown the same behaviour as in our previous research (Castel-Clavera et al. 2022) despite the change of the study period. The Weather variables that constitute the best performer combination are as well in concordance with Castel-Clavera et al. (2025). In this model the seasonal effect (based on the week of the year) had a flat effect, meaning that all the seasonal variability is captured by the explicit variables included in the model (especially weather variables). It can also reveal one limitation of the approach, as the short study period might not provide enough data for the identification of strong seasonal trends.

The OS-1, most performant combination for the size component, featured weather, LULC and remote sensing information, paired with the STEs. The remote sensing variables retained in this model component are the SWIR and NIR spectral bands, as in OM-1. Nevertheless, the interpretation of their effect might be different as in the occurrence component. The weather variables of the size model have larger confidence intervals than the occurrence due to the smaller dataset and short duration of the time series, nonetheless, they present the expected behaviors as seen in our previous research (Castel-Clavera et al. 2022). Some LULC variables (e.g., coniferous area or slope) and the SWIR show very flat effects with large confidence

intervals, this suggests not to include them in further iterations of the modelling framework. Nonetheless, as they have improved the parsimony criterion (DIC; deviance information criterion) and the model performance, they are kept in this work as representative factors of fire size prediction.

In SM-1 SWIR and NIR have decreasing effects. The NIR keeps the decreasing effect in both model components. The SWIR has a decreasing effect in the size component, while it has a strong increasing effect in the occurrence component. The physical reason of this behavior is not straightforward with the available data and analysis. The SWIR reflectance is higher when the vegetation water content is lower, therefore we assume that the SWIR variable in the occurrence component indicates lower FMC, which entail more fire danger. SWIR has shown to be related to moisture content of plants (Danson and Bowyer, 2004). On the other hand, for the size component, the SWIR reflectance could carry on other phenomenon as its effects is inversed in comparison to the occurrence component. With the advancements of the current research, it is not possible to give a clear physical reason behind the effect of SWIR in SM-1. Furthermore, NIR is a proxy of the photosynthetic rate of fuels so, as the photosynthesis of plants within the pixel is more active, the fire occurrence danger decreases. In both model components, the combination of reflectance bands performed better than remote sensing indices (namely, NDVI and NDMI). However, a drawback from this result is the increased difficulty for the interpretation of the responses given by spectral bands.

Contribution of weather, LULC and RS variables to spatio-temporal variations in fire activity.

For the occurrence models (N1), remote sensing brings a small share of the total prediction variance, which belongs almost entirely to the spatial-interannual dimension with a similar contribution to that of long-term weather. The seasonal variance of occurrence models is not significatively driven by remote sensing. Short-term weather is the main driver of the seasonal variance. A step further for this research is to average the summer spectral reflectance by pixel to have only a precise spatial information. Some authors have already indicated the mostly spatial effect of RS, using mean values for the season. For example, Chéret & Denux (2011) did an analysis of annual variations of the NDVI through the relative greenness (RGRE) to characterize fluctuations on vegetation. Those indicate changes in photosynthesis and indirectly fuel dryness.

For 80ha threshold exceedance predictions (P80), the remote sensing variables bring a larger part of the total variance, still at a similar extent as long term-weather. This result is in

concordance with the model evaluation. In this case, the contribution of the optical remote sensing is shared between the seasonal and spatial-interannual variances. The seasonal variation is mostly explained by long-term weather in the case of the fire sizes, leaving short-term weather and remote sensing to a second and third positions.

For both model components, LULC does not contribute to the temporal variability, as LULC factors are mostly static over the season and introduced at yearly scale in the model fit. Moreover, this study is based on a short time period during which LULC has not experienced large modifications. Nonetheless, the LULC factors are a main source of variability in the spatial-interannual variation, especially for the P80 predictions. As seen in our previous research LULC neither played a crucial role on long-term fire regime changes but was crucial for the spatial distribution of fire activities (Castel-Clavera et al., 2022). Those results recall that the seasonal trends of fire ignitions and initial spread (occurrence of scaped fires) are mostly driven by the short-term weather conditions. While the probability of those fires to become large (P80) is mostly driven by long-term drought conditions (DC, DMC and RS).

Potential and limitations of RS to fire danger prediction

This study is the first application of Sentinel-2 optical remote sensing for fire danger prediction in Mediterranean France. Moreover, the statistical approach used here brings the advantage that the remote sensing data is compared directly to fire activities. This approach provides the response functions of each variable, which indicates the distribution of the effect along the range of the variable. Other approaches rely often on establishing correlations between remote sensing observations and fire records without the support of a modelling background (e.g., Ahmed et al., 2019). Nevertheless, the exploration of the potential application of optical remote sensing into our probabilistic framework entailed several technical and theoretical limitations and assumptions. One of the assumptions done for this work has been to consider that there is no significant difference between the two possible azimuths of observations from the two Sentinel-2 satellites. As a main spatial component is expected from the optical remote sensing data, this fine scale temporal variation is not considered.

Moreover, the optical reflectance data is not a direct measurement of plant traits. Although specific spectral bands interact with a given plant trait (dry matter content, moisture content, photosynthetic activity, etc.) their values must be interpreted with caution. For example, small NDVI values correspond either to large coverages by stressed vegetation, or moderate coverages by sclerophyllous vegetation or chaparral (garrigue) so the values have to be

evaluated considering the type and quantity of vegetation in the pixel. The estimations of live fuel moisture content require calibration of the remote sensing data (Yebra et al., 2013). The work on calibration of remote sensed data with field observations of fuel moisture in the region, is very recent and is not still applicable for the empirical models.

Overall, the effects of optical remote sensing in this research are small. Several perspectives are open for the future uses of RS data in Mediterranean France that could be tested un the near future following our statistical approach. Active remote sensing techniques as microwave are claimed to outperform the optical remote sensing (Konings et al., 2019, Rao et al., 2020). Soil Moisture Active Passive (SMAP) is a combined product between active and passive microwave remote sensing, which has proven a better estimator of LFMC than the optical remote sensing indices (Jia et al., 2019). Rao et al. (2020) exposed, as well, the advantages of empirical models over the mechanistic RTMs (Radiative Transfer Models). On the other hand, Tanase et al. (2022) found active remote sensing (radar) to be less effective than optical remote sensing. That research found the VARI (Visible Atmospherically Resistant Index), EVI (Enhanced Vegetation Index) and NDMI as the best predictors of LFMC. In addition, remote sensing can provide high resolution fuel classification (Yebra, et al., 2023), which could improve the fire danger prediction by complementing the LULC factors.

Acknowledgements

We thank the team on the Laboratoire d'Observation de la Terre from the Centre National d'Études Spatiales (CNES), managed by Raquel Rodriguez-Suquet, and the project AleoFeu for the accessibility and pre-treatments to the Sentinel-2 optical reflectance products used in this study. We thank the Office National des Forêts for providing the fire records dataset that we used in this study.

References

Abatzoglou, J. T., & Kolden, C. A. (2013). Relationships between climate and macroscale area burned in the western United States. International Journal of Wildland Fire, 22(7), 1003. https://doi.org/10.1071/WF13019

Abdollahi, A., & Yebra, M. (2023). Forest fuel type classification: Review of remote sensing techniques, constraints and future trends. Journal of Environmental Management, 342, 118315. https://doi.org/10.1016/j.jenvman.2023.118315

Allard, D., Comunian, A., & Renard, P. (2012). Probability Aggregation Methods in Geoscience. Mathematical Geosciences, 44(5), 545–581. https://doi.org/10.1007/s11004-012-9396-3

Barbero, R., Abatzoglou, J. T., & Brown, T. J. (2015). Seasonal reversal of the influence of El Niño–Southern Oscillation on very large wildfire occurrence in the interior northwestern United States. Geophysical Research Letters, 42(9), 3538–3545. https://doi.org/10.1002/2015GL063428

Burgan, R., Klaver, R., & Klaver, J. (1998). Fuel Models and Fire Potential From Satellite and Surface Observations . International Journal of Wildland Fire 8. https://doi.org/10.1071/WF9980159.

Castel-Clavera, J., Pimont, F., Opitz, T., Ruffault, J., Rivière, M., & Dupuy, J.-L. (2022). Disentangling the factors of spatio-temporal patterns of wildfire activity in south-eastern France. International Journal of Wildland Fire, 32(1), 15–28.

Chéret, V., & Denux, J.-P. (2011). Analysis of MODIS NDVI Time Series to Calculate Indicators of Mediterranean Forest Fire Susceptibility. GIScience & Remote Sensing, 48(2), 171–194. https://doi.org/10.2747/1548-1603.48.2.171

Chuvieco, E., Aguado, I., Salas, J., García, M., Yebra, M., & Oliva, P. (2020). Satellite Remote Sensing Contributions to Wildland Fire Science and Management. Current Forestry Reports, 6(2), 81–96. https://doi.org/10.1007/s40725-020-00116-5

Gale, M. G., Cary, G. J., Van Dijk, A. I. J. M., & Yebra, M. (2021). Forest fire fuel through the lens of remote sensing: Review of approaches, challenges and future directions in the remote sensing of biotic determinants of fire behaviour. Remote Sensing of Environment, 255, 112282. https://doi.org/10.1016/j.rse.2020.112282

Ganteaume, A., Camia, A., Jappiot, M., San Miguel-Ayanz, J., Long-Fournel, M., & Lampin, C. (2013). A Review of the Main Driving Factors of Forest Fire Ignition Over Europe. Environmental Management, 51(3), 651–662. https://doi.org/10.1007/s00267-012-9961-z

Gneiting, T., & Raftery, A. E. (2007). Strictly Proper Scoring Rules, Prediction, and Estimation. Journal of the American Statistical Association, 102(477), 359–378. https://doi.org/10.1198/016214506000001437

Guk, E., Bar-Massada, A., Yebra, M., Scortechini, G., & Levin, N. (2024). Examining the Transferability of Remote-Sensing-Based Models of Live Fuel Moisture Content for Predicting Wildfire Characteristics. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17, 14762–14776. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. https://doi.org/10.1109/JSTARS.2024.3445138

Jia, S., Kim, S. H., Nghiem, S. V., & Kafatos, M. (2019). Estimating Live Fuel Moisture Using SMAP L-Band Radiometer Soil Moisture for Southern California, USA. Remote Sensing, 11(13), Article 13. https://doi.org/10.3390/rs11131575

Koh, J. (2023). Gradient boosting with extreme-value theory for wildfire prediction. Extremes, 26(2), 273–299. https://doi.org/10.1007/s10687-022-00454-6

Konings, A. G., Rao, K., & Steele-Dunne, S. C. (2019). Macro to micro: Microwave remote sensing of plant water content for physiology and ecology. The New Phytologist, 223(3), 1166–1172. https://doi.org/10.1111/nph.15808

Lindgren, F., & Rue, H. (2015). Bayesian Spatial Modelling with R—INLA. Journal of Statistical Software, 63(19). https://doi.org/10.18637/jss.v063.i19

Moritz, M. A., Morais, M. E., Summerell, L. A., Carlson, J. M., & Doyle, J. (2005). Wildfires, complexity, and highly optimized tolerance. Proceedings of the National Academy of Sciences, 102(50), 17912–17917. https://doi.org/10.1073/pnas.0508985102

Nolan, R. H., Resco De Dios, V., Boer, M. M., Caccamo, G., Goulden, M. L., & Bradstock, R. A. (2016). Predicting dead fine fuel moisture at regional scales using vapour pressure deficit from MODIS and gridded weather data. Remote Sensing of Environment, 174, 100–108. https://doi.org/10.1016/j.rse.2015.12.010

Pettinari, M. L., & Chuvieco, E. (2020). Fire Danger Observed from Space. Surveys in Geophysics, 41(6), 1437–1459. https://doi.org/10.1007/s10712-020-09610-8

Phelps, N., & Woolford, D. G. (2021). Guidelines for effective evaluation and comparison of wildland fire occurrence prediction models. International Journal of Wildland Fire, 30(4), 225. https://doi.org/10.1071/WF20134

Pimont, F., Fargeon, H., Opitz, T., Ruffault, J., Barbero, R., Martin-StPaul, N., Rigolot, E., RiviÉre, M., & Dupuy, J. (2021). Prediction of regional wildfire activity in the probabilistic

Bayesian framework of Firelihood. Ecological Applications, 31(5), e02316. https://doi.org/10.1002/eap.2316

Preisler, H. K., Brillinger, D. R., Burgan, R. E., & Benoit, J. W. (2004). Probability based models for estimation of wildfire risk. International Journal of Wildland Fire, 13(2), 133. https://doi.org/10.1071/WF02061

Rao, K., Williams, A. P., Diffenbaugh, N. S., Yebra, M., & Konings, A. G. (2022). Plant-water sensitivity regulates wildfire vulnerability. Nature Ecology & Evolution, 6(3), 332–339. https://doi.org/10.1038/s41559-021-01654-2

Rao, K., Williams, A. P., Flefil, J. F., & Konings, A. G. (2020). SAR-enhanced mapping of live fuel moisture content. Remote Sensing of Environment, 245, 111797. https://doi.org/10.1016/j.rse.2020.111797

Rue, H., Riebler, A., Sørbye, S. H., Illian, J. B., Simpson, D. P., & Lindgren, F. K. (2017). Bayesian Computing with INLA: A Review. Annual Review of Statistics and Its Application, 4(1), 395–421. https://doi.org/10.1146/annurev-statistics-060116-054045

Tanase, M. A., Nova, J. P. G., Marino, E., Aponte, C., Tomé, J. L., Yáñez, L., Madrigal, J., Guijarro, M., & Hernando, C. (2022). Characterizing Live Fuel Moisture Content from Active and Passive Sensors in a Mediterranean Environment. Forests, 13(11), 1846. https://doi.org/10.3390/f13111846

Turco, M., Von Hardenberg, J., AghaKouchak, A., Llasat, M. C., Provenzale, A., & Trigo, R. M. (2017). On the key role of droughts in the dynamics of summer fires in Mediterranean Europe. Scientific Reports, 7(1), 81. https://doi.org/10.1038/s41598-017-00116-9

Van Wagner, C. E. (1974). Structure of the Canadian Forest Fire Weather Index. Environment Canada, Forestry Service.

Van Wagner, C. E. (1987). Development and structure of the Canadian Forest Fire Weather Index System. Minister of Supply and Services Canada.

Vidal, J., Martin, E., Franchistéguy, L., Baillon, M., & Soubeyroux, J. (2010). A 50-year high-resolution atmospheric reanalysis over France with the Safran system. International Journal of Climatology, 30(11), 1627–1644. https://doi.org/10.1002/joc.2003

Wang, L., Hunt, E. R., Qu, J. J., Hao, X., & Daughtry, C. S. T. (2013). Remote sensing of fuel moisture content from ratios of narrow-band vegetation water and dry-matter indices. Remote Sensing of Environment, 129, 103–110. https://doi.org/10.1016/j.rse.2012.10.027

Wang, X., Wotton, B. M., Cantin, A. S., Parisien, M.-A., Anderson, K., Moore, B., & Flannigan, M. D. (2017). cffdrs: An R package for the Canadian Forest Fire Danger Rating System. Ecological Processes, 6(1), 5. https://doi.org/10.1186/s13717-017-0070-z

Yebra, M., Dennison, P. E., Chuvieco, E., Riaño, D., Zylstra, P., Hunt, E. R., Danson, F. M., Qi, Y., & Jurdao, S. (2013). A global review of remote sensing of live fuel moisture content for fire danger assessment: Moving towards operational products. Remote Sensing of Environment, 136, 455–468. https://doi.org/10.1016/j.rse.2013.05.029

Yebra, M., Quan, X., Riaño, D., Rozas Larraondo, P., Van Dijk, A. I. J. M., & Cary, G. J. (2018). A fuel moisture content and flammability monitoring methodology for continental Australia based on optical remote sensing. Remote Sensing of Environment, 212, 260–272. https://doi.org/10.1016/j.rse.2018.04.053

4.3. Overview of the results

This first study with optical remote sensing (RS) for the probabilistic predictions of fire activity in the Promethée area has shown the potential application of this type of data on the modelling framework. The results on Sentinel-2 reveal that the best remote sensing variables are B8A+B11 (NIR and SWIR) for the occurrence (N1) and the size (P80) models. Both N1 and P80 are based on the best performer models shown in Chapter 3. Optical RS effect is more important for fire size predictions than for occurrences. In all cases, the performance brought by the optical remote sensing data did not replace the explanatory power of the weather data. The information brought by the two data sources is complementary.

The partition of variance shows that the satellite data brings mainly explanatory power for the N1 model on the spatial dimension, and few temporal information. In the P80 models the variance of RS is shared between the spatial and the temporal dimensions. The other groups of explicit variables (LULC and Weather) bring more variance to the predictions that optical RS variables. As expected, the variance from LULC is concentrated in the spatial dimension. Short-term weather variables, such as wind, FFMC, temperature or VPD are the main responsible of the seasonal variance for N1 models. In P80 models it is the long-term weather (DC, DMC) and the remote sensing that bring a major part of the variance. Nonetheless, weather variance is larger than the RS variance in both cases. The interannual variance is smaller than the spatial and the temporal variances in all model components, although the differences on fire activity between seasons remain important during this period.

The modelling process presented in this chapter has shown the limits of the available data, regarding the number of explanatory factors that can be used in a model. This is visible through the Skill-Scale of the occurrence models, as the Skill Scale is reduced when all LULC predictors are introduced. Nevertheless, the LULC variables do not show irregular behavior. As the model performance is lower with all LULC variables than without, it is necessary to select the most relevant variables with the parsimony criterion. As shown in the results, the best occurrence model only includes three LULC variables: wildland area, fuel type rating and road length. The main limitation found in this chapter is the length of the study period, which has been reduced to match the availability of Sentinel-2 data. In the future, fitting models with longer Sentinel-2 time series will be

possible. In addition, the tests with the MODIS data available for a longer time, despite the issues described in 4.1., have led to similar results.

5. Discussion

After addressing the objectives defined at the beginning of this work, the results have shown the capabilities of the probabilistic Bayesian framework for fire danger prediction. Each chapter has explored one aspect of the fire factors using adequate and innovative methods. This discussion will go through the most important findings of each chapter giving answer to the cognitive and methodological questions. And finally, it will develop the applications of our fire danger prediction and perspectives for enhancement and further development of our research.

5.1. Insights from modeling

The model building and evaluation processes have given important insights from the fire activities in the study area. The modelling framework is constituted of two separate components, one for the occurrences of fires larger than 1ha (N1), and a second one for the probability of an 80ha threshold exceedance (P80). The factors here studied (weather, fuel, landscape, human) have contributed to the final predictions in different proportions. Beyond the model that were developed over these studies, this work provides knowledge regarding the different drivers of fire occurrences and fire sizes. In addition, the model was used to explore the changes that have been observed in the fire regime during the last decades.

5.1.1. Weather indices

The CanFWI positioned as the best predictor over all tested fire weather indices, which strengthens its position as a key danger index for the French Mediterranean region. Nonetheless, the combination of the fuel moisture indices of the Canadian Fire Weather System (DC, DMC and FFMC), wind speed and VPD (for N1) or Tmax (for P80), as well as the STEs, yields the best fire activity predictions. The combination of DC, DMC, FFMC and wind speed explain a major part of the fire activity variations. While the contribution of VPD or Tmax is smaller, this finding illustrates that a shorter time scale than FFMC response time can play a role to identify fire activities. Furthermore, the results of Chapter 2 demonstrated the importance of taking STEs into account, underlining that weather alone is not sufficient for a spatially consistent prediction of fire danger. The STEs capture either seasonal patterns in human activities or spatial patterns that might also be hard to

capture with explicit predictors. Through the modelling process, it is possible to obtain the independent effects of each predictor variable. Therefore, we were able to disentangle the behavior of each fuel moisture index and meteorological variable. Moreover, those meteorological and weather effects remain stable regardless of the introduction of further factors in the model, which demonstrates their robustness on the prediction of fire activities.

5.1.2. Land Use and Land Cover (LULC):

We found that most prevalent factors in fire occurrences were human-related, including population density, dwellings and access networks to forest and natural land areas (Costafreda-Aumedes et al., 2017). Vegetation cover types were not found to drive fire occurrences, instead, fuel rating was found to be a better predictor. On the contrary, the exceedance probability of 10 and 100 ha was determined by vegetation variables, matching previous results for fire size (Díaz-Avalos et al., 2016), and land cover fire-proneness studies in Europe (Moreira et al., 2009). Vegetation cover variables were also found to be more influential on fire sizes than on fire occurrence in the United States (Hawbaker et al., 2013).

LULC variables played an important role in fire occurrence, explaining ~25% of its spatial distribution in accordance with previous results on the western part of the study area (Ruffault and Mouillot 2017). Nonetheless, unexplained spatial factors and their temporal changes still represented half of the spatial variability in occurrence, meaning that these commonly used predictors did not fully explain spatial and temporal distribution of fire occurrences. The fire-weather explained a major part of the spatial distribution of larger fires, and the share of the LULC and unexplained spatial effects were small. In addition, the LULC factors explain better fire size when considered at coarser scales (>2km). It can be concluded that fire occurrence was explained equally by fire-weather and by landscape and socio-economic drivers, and once a fire has been ignited, the major driver for its size was the weather during the fire event.

5.1.3. Optical Remote Sensing:

The analysis of satellite optical remote sensing effects on fire activity started with the preliminary correlation analysis of the available data to determine which variables were

worthy of further exploration. As a result, we found many of the spectral bands to be highly consistently correlated in the temporal and the spatial dimensions. Therefore, only one of the bands between B2 to B5, B6 to B8A and B11 to B12 was chosen. The tests of optical remote sensing on the probabilistic framework revealed that Near Infrared (NIR) and Short-Wave Infrared (SWIR) spectral bands are better predictors of fire activities than the Normalized Difference Vegetation Index and the Normalized Difference Moisture Index. In contrast, Tanase et al. (2022) found the optical reflectance indices to be the most effective predictors in the study case. The mentioned spectral bands (NIR and SWIR) are particularly effective to explain the spatio-temporal distribution of larger fires. NIR had the same response functions on the size and the occurrence predictions, while the SWIR showed an opposite slope between both predictions. The reason behind the change on variable's behavior between occurrence and size, discussed in Chapter 4, is difficult to disentangle through our probabilistic approach, but raises questions for future research. Although the framework gives robust variable's effect and provides its observed behavior, the assumptions done for the physical factors behind those effects should be confirmed by a specific study. Such a work has not been carried out during the PhD work, but perspectives are discussed in section 5.4.

Although results from MODIS are not shown in Chapter 4 for the reasons related to daily product availability (only a composite product was used, see section 4.1.1), the exploratory analysis on the P80 were carried out. Those analysis showed that MODIS optical reflectance data has potential applicability in the framework (see supplementary materials A). Nonetheless, the data constraints on its temporal resolution discarded the use of the MOD13Q1 and MYD13Q1 products from MODIS on the fire occurrence predictions. Between Sentinel-2 and MODIS, all tested variables show the same behavior despite the difference in the datasets. The main differences are that Sentinel-2 data has higher spatial resolution (before aggregation) with a temporal resolution of multiple days, while MODIS data is available for a longer period at daily scale (although the MODIS product we had available was at 16-day temporal resolution).

5.1.4. Relative contributions of factors to spatio-temporal variability of fire activity

In Chapters 2 and 4 *Firelihood* was used to study the contribution of variables to the prediction variances. In the first case we studied the influence of three groups of variables

on the temporal changes of the fire regime and the interannual and spatial variability of the fire activities. In the second case, the study aimed to find the contribution of each group of variables to the spatial and seasonal distributions of the predicted fire activities.

In both studies the effects of LULC were important for the reproduction of the spatial patterns of fire activities, especially for the P80. Moreover, the weather effects were crucial for the fire danger prediction in both N1 and P80 components. Later in Chapter 4, it was shown that the seasonal distribution of N1 is mostly affected by short-term weather, while for the P80 it is long-term weather that drives the seasonal trends. This is an important finding that helps to understand how the different temporal scales at play are involved.

With respect to earlier research (Pimont et al. 2021), the inclusion of additional drivers allowed to reduce the importance of STEs. This is due to the role of all other predictors which capture the variability of fire activity (see 5.2). Nonetheless, in the present study, the spatial effects (from the STEs) remain a key factor for the fire danger predictions, meaning that there are key fire factors that we are not still able to quantify despite the inclusion of all tested variables. Finally, the analysis on Sentinel-2 data revealed that the effect of optical remote sensing variables is mostly in the spatial distribution of fire activities, as suggested by Chéret & Denux (2011). The remote sensing variables are mainly effective in the spatial and the seasonal distributions of large fire size.

5.1.5. Fire regime trends

During the 1993-2018 study period, there was an observed decay in escaped fire numbers mostly caused by a major reduction after 2003, the posterior evolution has been limited. Few variations were detected on exceedance probabilities, especially in the recent years. Following the record-breaking year 2003 in terms of burnt area in France, the suppression was reinforced with new fire-fighting material and better training capacities, but such changes were not easy to capture with quantitative variables. As suggested for the 1973-2005 period (Ruffault et al., 2015), or for the 1976-2009 period (Fréjaville & Curt, 2017), those changes in firefighting policies can drive the observed changes in fire activities.

The analysis carried in the study reveal that fire-weather change alone drive significant increases in potential fire activity. Around 50% of this increase in fire-weather over

France has already been attributed to anthropogenic climate warming (Barbero et al., 2020). Despite the importance of LULC variables over the spatial distribution of fires, LULC-changes alone induced marginal temporal changes in fire activity. The unexplained factors (captured through the Spatio-Temporal Effects) were the main drivers of temporal changes in fire activity. The spatial distribution of those unexplained changes revealed positive temporal trends in the south-west region of the study area. According to fire managers, those trends are associated with areas of agricultural abandonment, raising the ignition potential of this area. It is also likely that suppression policies were not strong enough -or not enough reinforced- in these traditionally less fire-prone areas and struggled to face the increase driven by fire-weather. The resolution of the LULC data was a limiting factor to carry this type of temporal analysis. The abandoned agricultural lands cannot be detected with the spatial resolution of Corine Land Cover and its classifications.

As said before, we analyzed the period 1993 – 2018 to obtain the mentioned results. Until then, the burnt area had a decreasing trend for the last decade, except for 2017. Nevertheless, during the publication of the work and the continuation of the thesis, we have obtained the fire activities until 2022 which starts to show a trend of rising burnt areas, despite of prevention and suppression policies. Moreover, our most recent results show that the interannual variability of the predicted fire activities is mainly driven by fire-weather. This may create concern around the fire policies, the observed rise of burnt areas could be a sign that the effects of climate change are not compensated any more by fire suppression. Pimont et al. (2022) anticipate, that under constant prevention and suppression policies, the fire activities would intensify with climate change, with expanded spatial distribution and longer fire seasons.

5.2. Methodological advancements

To address the objectives defined for this thesis, we have developed several extensions and improvements to the *Firelihood* modelling framework. The modelling is done with a Bayesian approach based on Lindgren & Rue (2015), at a relatively broad scale (2 km) by separating the processes of occurrence and spread of wildfires. This has advantages compared to other approaches (e.g., machine-learning or generalized additive models) in terms of overfitting risks and interpretation of the outputs. Those advantages identified for Firelihood can be summarized in three main points. First, a better management of the

uncertainty on the predictions than with other methods which assume that the parameters of the model are deterministic. the simulation of fire activities with the Bayesian framework accounts for the possible values of each parameter before applying the Poisson law (Huang et al., 2017). Second, the spatial autocorrelation of the observations is well accounted through the spatial effect included in the model. This model component creates a mesh to discretize the space and compute the influence of neighboring points in the space (Lindgren et al., 2011). Third, the possibility of *a-priori* parametrization of the explanatory variables (parametrization of the model before fitting). This allows to adjust the sensibility of the model while maintaining the interpretability of the results afterwards (Rue et al., 2017).

In previous research, probabilistic methods similar to the ones used in the present work have already shown the capabilities for accurate fire activity prediction (Parisien et al., 2012; Phelps & Woolford, 2021; Preisler et al., 2004; Woolford et al., 2011; Worsnop et al., 2021). The work done in this thesis focused on different aspects of the modelling process. First, there is the problem of predictor variable selection for a better understanding of the trends in fire activities over the last decades in relation to changes in landscape fire factors. Later, the attention was brought to the method of evaluation for the short-term fire danger prediction and the weather drivers of fire activities. The last part of this thesis concentrates on the exploration of alternative methods and data sources, in particular remote sensing, to account for fuel conditions and to quantify the contribution of each type of variable to the final model predictions.

5.2.1. Variables' selection and performance evaluation

AUC is a standard metric to compare models for binary outcomes and has also been repeatedly used in this thesis, but it also has weaknesses in particular when evaluating the predictions of rare and stochastic events, in the context of fire activity dominated by strong seasonal and spatial patterns. In Chapter 2, this metric was used to compare the model performances at the scale of pixel (DFCI) per year. At this large temporal scale, where the patterns of fire activity show only moderate variations in frequency, the AUC gave pertinent scores and rankings. Nonetheless, at pixel per day scale, the patterns of fire activity are highly variable, due to the rareness of the event at this finer scale, which motivates using more specific evaluation methods. Eastaugh et al. (2012) stressed the

need for a consistent method to compare the performance of fire indices. Following Phelps and Woolford (2021), from Chapter 3, all model tests were additionally evaluated using LogLoss scores, along with Precision-Recall curves that are suitable for rare binary events (Saito & Rehmsmeier, 2015; Sofaer et al., 2018). This extended, multi-criteria model evaluation method was used to compare the performance of each informed model to the Fire Activity baseline (FAb; see Chapter 3). By considering several criteria when comparing a large number of possible predictor configurations, we can further check whether the finally selected model performs well for all criteria and can be considered as being truly better than its competitors. This novel approach emphasizes the variations of fire activity at a daily scale, which remained limited with respect to large-scale seasonal and spatial variations, but which are critical for operational applications. This method more accurately measures the contribution of all variables to overall model performance at fine spatiotemporal scales, while keeping seasonal and spatial random effects in the model that, alone, could already explain a large part of the variability in fire activity.

5.2.2. Contributions of different predictor types to improved model performance

Throughout the studies conducted in this thesis, variants of the *Firelihood* model have served to assess the predictive potential of various fire-related predictor variables. Thanks to those tests, several predictor variables have been kept following the parsimony criterion. First, the landscape factors showed their importance on the description of the spatial patterns of fire activity. In terms of model performance, the introduction of landscape factors increased the AUC (performance metric used before the new evaluation protocol) of the occurrence models from 0,73 with a simple model to 0,88 with all LULC variables and STEs. Those increments, although being significant, were lower for the size exceedance thresholds (10ha and 100ha). The Chapters 3 and 4 needed an alternative method for model evaluation due to the smaller temporal scale. The new evaluation method consisted of the comparison of model performance, based on the Log Loss score, against a Fire activity baseline (FAb). This comparison is quantified through the Skill Scale, which attributes a value of 0 to the FAb and a value of 100 to the best-performing fitted model. In our research, the FAb consists of a model fitted only with the STEs in order to represent an average season in the study area. The new explicit predictors added to the model capture with higher precision the small-scale variations on the spatial and

temporal patterns of fire activity. As a result, the magnitude of the STEs is reduced when further relevant fire predictors are used together with the STEs.

Moreover, the modification of the models' weather component showed an improvement of almost 30 units of Skill Scale compared to the Canadian FWI; this difference is of 12 units of Skill scale when STEs are included along with the CanFWI. Furthermore, the use of the best model based on weather and land-use land-cover shows better predictability than the operational methods based on the Canadian FWI. In the most advanced version of *Firelihood*, where we further include optical remote sensing reflectance data, we found benefits on the prediction of the 80ha threshold exceedance probability with ∼10 units of increment on Skill Scale. For the fire occurrence, the increment brought by the optical remote sensing was smaller, ∼5 units on Skill Scale.

Despite all the significant contributions from the different predictor variables, the spatial and temporal random effects remain necessary, even in the models selecting from many weather, land-use land-cover and remote sensing variables, to achieve the best possible fire activity prediction. Previous research on this subject already underlined that unexplained factors depending on the local conditions are very important (Díaz-Avalos et al., 2016). The STEs implemented in the first work in Chapter 2 were quite sophisticated. In the fire occurrence component, they account for spatial factors, overall temporal changes, and trends in spatial patterns. The latter effect allows spatial patterns to change over time. The STEs of the size models were constructed using simpler structures because of the smaller fire dataset. The temporal effects allowed for potential yearly changes, and spatial effects were structured by SERs.

Between Chapter 2 and Chapters 3 and 4, the structure of the spatio-temporal effects was slightly modified. The reason was that the full extent of all of these specialized variants of STEs cannot be jointly implemented in a single model due to its complexity, especially when the model further includes many other components capturing the effects of the various other predictor variables. An excess of complexity creates models whose implementation is long and highly demanding of computational capacities. Possible confusions between the effects of different model components can lead to models showing problems with parameter identification and uncertainty assessment. Moreover, the interpretation of the results is not straightforward or even meaningful. We make a tradeoff between the different forms and complexity levels of the STEs to better adapt

these effects to the specific model purposes and to the amount of information already brought by the other predictor variables included in the model. For example, when the temporal resolution of the model was modified from yearly (to study long-term trends of the fire regime) to daily (to explore short-term fire danger prediction). Moreover, the resolution of the STEs must be adapted to the data availability, as higher resolutions demand longer training datasets. As a result of these tradeoffs, the effect of overall temporal changes based on the year presented in Chapter 2 is no longer used in Chapters 3 and 4. Instead, a seasonal effect (based on the week) is included to account for unexplained variations of activity on a weekly basis. The later can be representative of seasonal changes in prevention or suppression, impact of tourism, or simply seasonal bias in explanatory variables such as the Drought Code (DC from CFWIS). Also, the spatial effect was simplified, not accounting anymore for temporal changes on its distribution.

The probabilistic framework of *Firelihood* has demonstrated to be a powerful and adequate tool to understand and predict the fire activities, but it remains highly demanding on training data. The analysis carried out for Sentinel-2 data has shown the limit of the approach regarding the size of the fire dataset. The model reaches a saturation of explanatory power when too many predictor variables are introduced. In other words, if the model is informed with too many variables for the size of the training data, the predictability decreases, possibly due to overfitting but also to instabilities arising in the numerical optimization procedures required to estimate the high number of parameters in the model. Similarly, the precision of the STEs had to be decreased to avoid overfitting of the probabilistic models of fire occurrence and size.

5.2.3. Handling different spatial and temporal scales in models

The probabilistic framework here presented has the advantage of allowing spatial and temporal aggregations of the simulations, thanks to additivity of Poisson process (i.e., numbers of fire can be summed across scales). The work presented here has been adapted in spatial and temporal scales to specific purposes. In Chapter 2 the models were used at yearly scale to capture the spatial patterns of fire activity and their long-term evolution. A comparison was carried out at decadal scale to disentangle the possible drivers of the multi-year trends, while keeping a spatial scale of high resolution (2km). Later, the model was used to provide better predictors of fire activity for short-term forecasts, which

necessitates a more precise temporal scale. In contrast, those forecasts are less demanding on the spatial scale, as the services of prevention and firefighting are organized at a larger territorial level than a 2km pixel. Therefore, in Chapter 3 and 4, the model simulations were aggregated spatially to the scale of the fire-climatic zones ("Zones Climat-Feu"; ZCFs), a subdivision of NUTS-3 used operationally.

It could seem a waste of calculation efforts or computational resources to fit a model in a fine spatial and temporal scale (pixel / day) and later aggregate the simulations at larger scales. Instead, a modelling process that is done directly at the larger scale could appear more convenient. Nonetheless, this thesis defends the use of fine scales for the model fitting and simulation. The small scales allow to extract as much variability and information as possible from the predictor variables, so that they remain interpretable. The aggregation of meteorological variables at larger temporal scale smooths out informative temporal variation and would overlook extreme meteorological events, such as heatwaves or wind events, which are often linked to the occurrence of the largest wildfire events (Ruffault et al. 2020). Similarly, the aggregation of landscape data at larger spatial scales may hide local particularities, which are crucial for the fire occurrence prediction.

5.3. Potential Applications

The probabilistic models developed here have a variety of possible applications for the operational and the scientific fields. The capability of explaining the fire activities across different spatial and temporal scales enables the method to be used for short-term forecasting and for mapping the long-term trends of fire hazard. In the following, two potential operational applications of the new *Firelihood* versions are presented.

5.3.1. Operational short-term forecasting

Results on Chapter 3 have shown the capability of weather-based probabilistic models to improve operational methods. Section 3.1 showed that simple improvement of operational danger classes could be proposed from probabilistic model. It would result into a better classification of fire spread days. An example application is shown in Figure 18 where the model predictions are compared to the CanFWI for a given day with a high fire activity. The CanFWI classification (Fig. 18B) shows a similar distribution to the

modelled fire occurrence danger (Fig. 18C), most of the differences between both classifications are no larger than 1 class. An advantage of the modelling process, compared to the CanFWI, is that the model classification shows the probability of a fire to exceed a threshold of 80ha (Fig. 18D). Figures 18C and 18D highlight the differences between occurrence of 1ha fires (Fig. 18C), which are widely spread over the plain areas and Mediterranean hinterlands, and the potential for large fires (Fig. 18D) which is concentrated on Var and Alpes Maritime on this specific day. Those differences might be brought by the different effects that the landscape factors, the fire-weather and the STEs have on either the occurrence or the size model components.

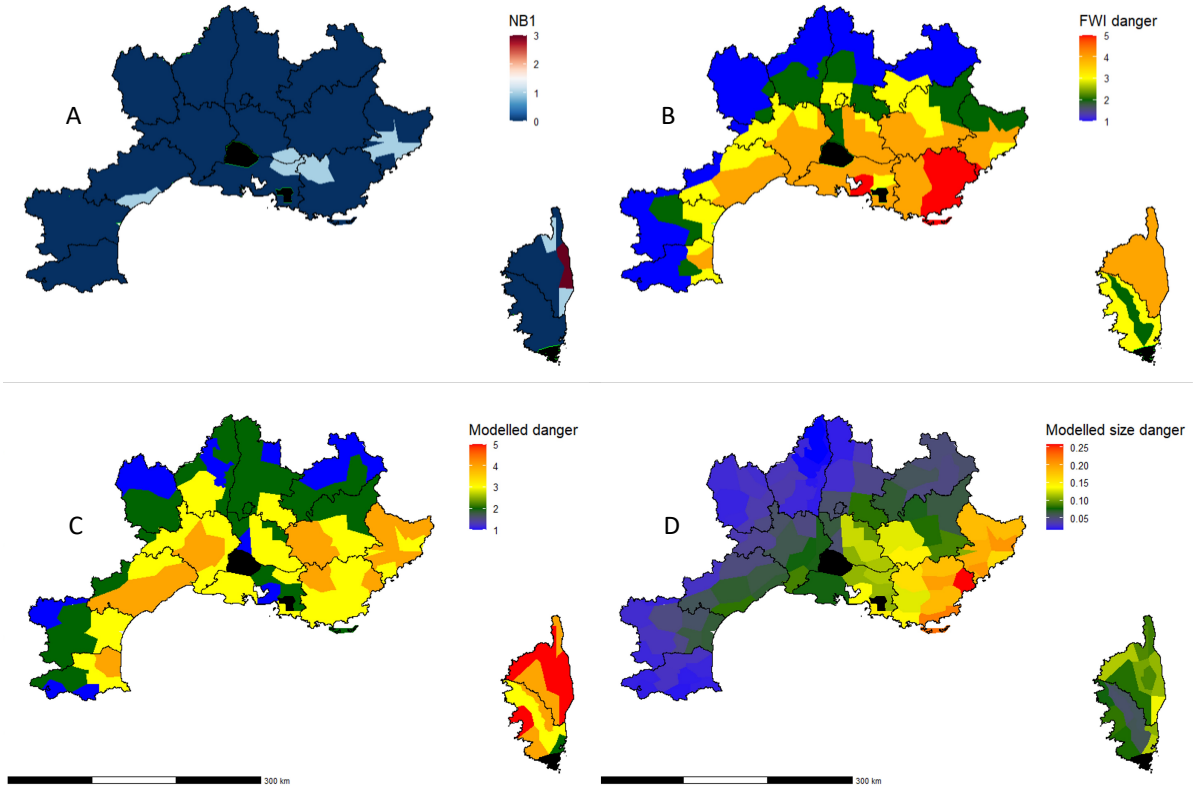


Figure 18: Fire observations and predictions summarized by ZFC for July 27th, 2017. Where: A – Fire (>1ha) observations; B – Distribution of the 5 classes of CanFWI(max) following the classification of Météo-France; C – Distribution of the 5 classes of modeled fire danger with the best fitted weather model (see Chapter 3); D – Distribution of the modeled fire size (80ha exceedance threshold) with the best fitted weather model (see Chapter 3). Black areas correspond to zones with no data due to computational errors. Black lines correspond to the departments (NUTS-3).

The final product which could be presented to operational managers is adaptable to their needs in term of spatio-temporal scales. Despite the type of presentation, the model can be fitted to the local particularities of the operational teams for an enhanced precision. A locally adapted model training entails that responses functions and STEs are also more specific to the designed area. During the development of this thesis, two prototype

products have been proposed to operational stakeholders from two regions within the study area. One proposal is already in development for the *Service Départemental d'Incendie et de Secours* (SDIS) from *Bouches-du-Rhône* (SDIS13). To maximize the precision and utility of the product, the model is trained and simulated for the *Bouches-du-Rhone* department (NUTS-3 unit) and the four adjacent departments. This model is given at the operationally most relevant spatio-temporal scales, namely the ZFCs or the woodland massifs. The output of this product is released as in Figure 18, but in this case, the maps cover only the area of the five departments (NUTS-3) included in for the SDIS-13 (Figure 19).

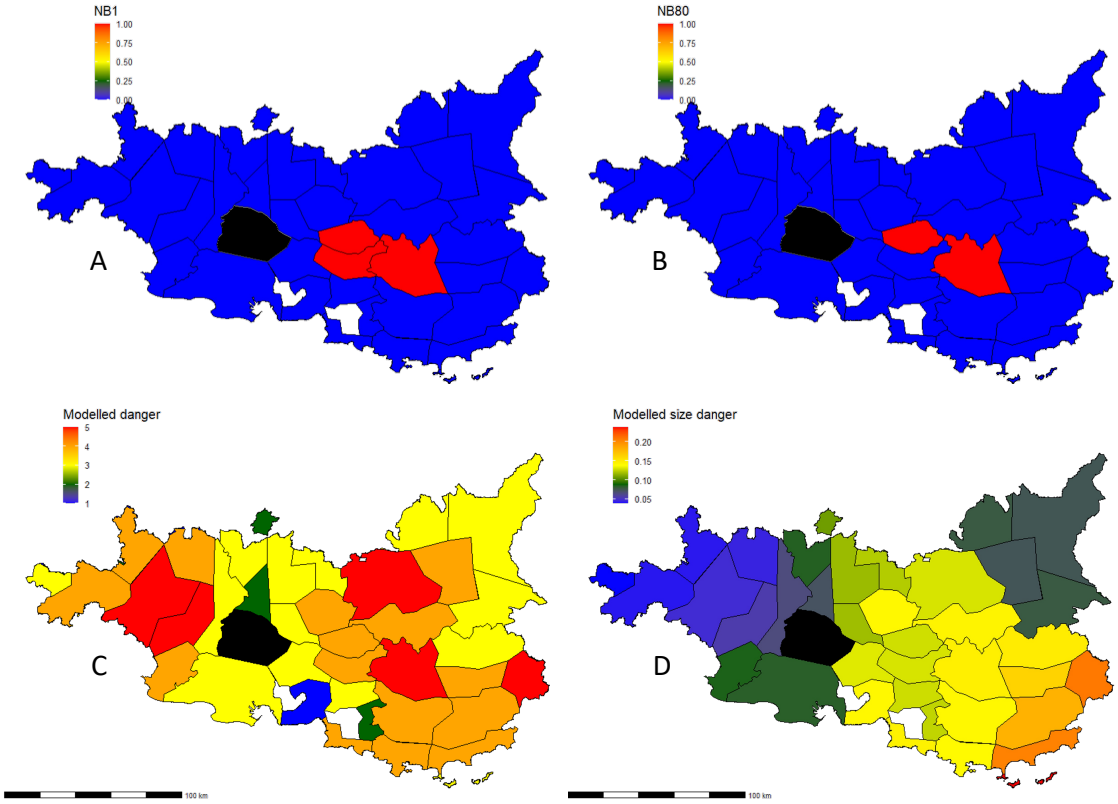
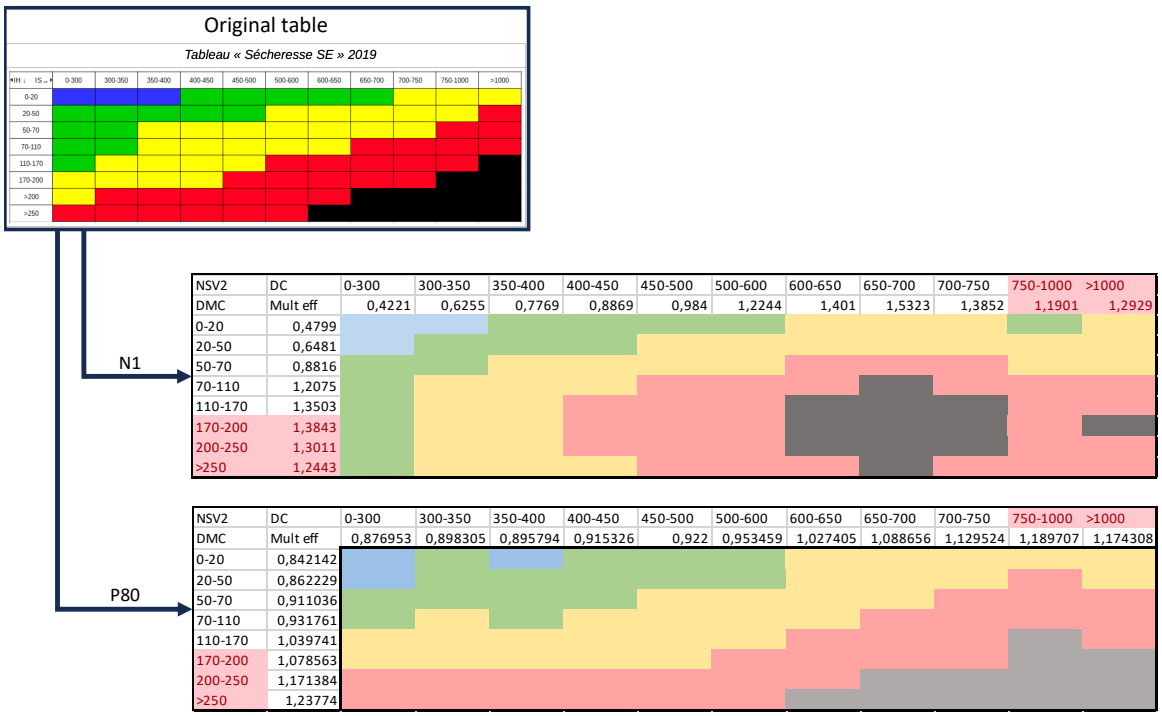


Figure 19: Fire observations and predictions in the Bouches-du-Rhone, Alpes-de-Haute-Provence, Gard, Var, and Vaucluse departments summarized by ZFC (black lines) for July 27th, 2017. Where: A – Occurrence observations of >1ha fires; B – Size observations of >80ha fires; C – Distribution of the 5 classes of fire danger corresponding to the best fitted weather model (see Chapter 3); D – Distribution of the modeled fire size danger (80ha exceedance threshold) with the best fitted weather model (see Chapter 3). Black areas correspond to zones with no data due to computational errors.

A second proposal is developed in the frame of the Aleofeu research project (leaded by DDTM of *Aude* department and funded by CNES). In this case, the proposal is to modify the existing prediction methods with the predictions of the probabilistic models (Figure 20). This can be done by using a classification by quantiles of the number of days that are expected to appear in each danger category (see Chapter 3). As a result, the output of the double entry tables used by the operational forces at the present day is modified

accordingly. The new classification of the output is built around the effect of the variables observed through the empirical models. In addition, this classification also distinguishes the expected number of 1ha fire occurrences from the probabilities of the fires exceeding the 80ha threshold.



NSV2	DC	0-300	300-350	350-400	400-450	450-500	500-600	600-650	650-700	700-750	750-1000	>1000
DMC	Mult eff	0,4221	0,6255	0,7769	0,8869	0,984	1,2244	1,401	1,5323	1,3852	1,1901	1,2929
0-20	0,4799											
20-50	0,6481											
50-70	0,8816											
70-110	1,2075											
110-170	1,3503											
170-200	1,3843											
200-250	1,3011											
>250	1,2443											

NSV2	DC	0-300	300-350	350-400	400-450	450-500	500-600	600-650	650-700	700-750	750-1000	>1000
DMC	Mult eff	0,876953	0,898305	0,895794	0,915326	0,922	0,953459	1,027405	1,088656	1,129524	1,189707	1,174308
0-20	0,842142											
20-50	0,862229											
50-70	0,911036											
70-110	0,931761											
110-170	1,039741											
170-200	1,078563											
200-250	1,171384											
>250	1,23774											

Figure 20: Reclassification of operational double entry tables, keeping the frequency of observation for each danger class in occurrence and size components. In this example the table corresponds to the NSV2 index (in French: Niveau de sensivilité de la vegetation vivante; in English: Sensibility level of life vegetation). Red values of DC and DMC in the reclassified tables correspond to magnitudes of the drought where the variable reaches the maximal effect on fire danger and starts to decrease. "Mult eff" indicates the effect of the variable of the model prediction.

5.3.2. Hazard Mapping and long-term past and future trends

Once a model is built, it can be used to produce simulation of realistic or theoretical scenarios. This is an interesting feature to study the causes of changes on the past fire regimes, or the possible scenarios that we will encounter in the future. Danger and risk managers and policy makers can benefit from the results given by those studies. Those results bring the possibility to explore the potential consequences of changes on any of the fire factors. The model has been used to study the long-term trends (from 1993 to 2019) of the past fire regime, as illustrated in Chapter 3. The study confirmed the large reduction on the 1ha fire numbers, and a small reduction on the large fire numbers. This shift of the fire regime happened suddenly after the year 2003. The study revealed that none of the explicit fire predictors introduced to the model could explain the reduction on

fire numbers, instead, the shift was attributed to the STEs. In contrast, if considered on their own, the changes on fire-weather during the studied period would have led to an increase of fire numbers on all size categories. The STEs are likely to have captured the changes on policy that followed the catastrophic event of year 2003. Using climate change projections, the model can inform about the future trends of fire regimes based on various intensities of climate change. This functionality of the *Firelihood* framework is used in Pimont et al. (2022), where authors show that fire regimes will intensify in the areas that fire is already present nowadays. In addition, the warming climate would bring an expansion of the fire-prone regions and longer fire seasons.

5.4. Future improvements and extensions

Even the most complex models that are presented in this study still show several limitations concerning the explanatory capability of the included variables. The resulting distribution of the STEs, crucial for the predictability of the models, give a hint on the possible factors that have been captured. The weak points of the overall fire activity predictions have been identified thanks to the modelling process presented above. In the following, the technical improvement possibilities are presented together with the perspectives for future model applications.

5.4.1. Agricultural abandonment cartography and other LULC enhancements

The French Mediterranean region has suffered a large abandonment of its agricultural lands as a consequence of the socio-economic trends of the region (Casanova Enault et al., 2021). Recent agricultural abandonment means an expansion of non-managed forested lands (Bou Dagher Kharrat et al., 2022; Valladares et al., 2014), which creates a dense accumulation of fuels with continuity over the landscape. In many cases, this phenomenon is considered to increase the fire activity due to the changes in the landscape structure. In the case of our study area, the local managing authorities from the western regions have shown, in a personal communication, their concern about this process of abandonment.

In this study, the land use and land cover (LULC) information are extracted from the Copernicus service of Corine Land Cover (CLC). This product has large temporal and spatial resolutions, in addition, of a coarse classification of the LULC types. As a result, it does not capture well the agricultural abandonment processes in the study region. Several

tests (results not shown) have been carried out to identify the types of LULC that correspond to changes from agricultural lands to abandoned lands, but none has given pertinent results. Because of these limitations regarding the accuracy of the explanatory variable, it has not been possible to precisely quantify the role of agricultural abandonment in fire activity changes.

In a general context, certain remote sensing products may increase the spatial resolution of fuel classification (Abdollahi & Yebra, 2023). In France, a new product developed by the French National Geographic Institute (IGN; Institut national de l'information géographique et forestière) to describe the LULC will be soon widely available (OGS-CE; https://geoservices.ign.fr/ocsge). This product is built at a better spatial resolution than CLC, and therefore the characterization of the abandoned agricultural lands will benefit from its release. In addition, the OGS-GE can be used to improve the LULC variables that have been used until now, thanks to a better specificity on France. It is thus expected that the role of agricultural changes would be precisely estimated in the near future thanks to Firelihood approach. This is important to anticipate future trends, as well as potential consequences of public policies in terms of land use.

5.4.2. Use of remote sensing to obtain fuel conditions.

As exposed in this thesis, satellite remote sensing is a promising tool to obtain information about fuel conditions for large spatial extents (e.g., Gale et al., 2021; Nolan et al., 2016; Wang et al., 2013; Yebra et al., 2013, 2018). The interaction of the spectral bands with water and vegetation structures reveals the water content on fuels, greenness of live plants and even fuel load and structure. Chapter 4 presented the first use of remote sensing for fire activity prediction with a probabilistic framework in the Promethee area. This was done with optical remote sensing sourced from Sentinel-2 through the project Aleofeu and provided by the CNES. The presented methodology resulted in small improvements brought by the optical remote sensing variables, mainly in the spatial dimension. Remote sensing data could therefore be processed at seasonal scale to obtain a static indicator for the entire season but with fine spatial resolution. This indicator could inform about the overall expected behavior of fuels during the given season and reduce the computational cost of the data treatment.

This opens the hypothesis that optical remote sensing information could further improve the predictions of fire activity if it was used to obtain FMC estimations that later will inform the probabilistic model. This contrasts with the method used in the study, which uses optical remote sensing reflectance data as direct predictors. Furthermore, this additional step on the modelling process would facilitate the interpretation of the effect of the variables on fire occurrence and size predictions. Another advantage of this refinement is that the predicted effects could then be attributed to FMC, and not to any other hidden factor embedded in the optical remote sensing data such as biomass quantity. But as a counterpart, it could bring additional uncertainty into the predictions.

Remote sensing products are not limited to the passive optical sensors. There are other passive sensors that measure, for example, the microwave signals or the thermal signal through the infrared spectrum. The thermal signal is known to well capture the water stress of plants (Liu et al., 2020), with more instantaneous measurements than the optical reflectance. Furthermore, the active remote sensing is an alternative to the optical sensors. This type of remote sensing reads the reflectance of a signal sent beforehand such as the radar. Those alternatives are described with more detail in the discussion of Chapter 4.

5.4.3. Mechanistic modelling for fuel moisture content

Fuel characterization is one important factor for accurate fire activity prediction. In this study, as in many other research and applications, it has been done through weather-based proxies. Later, remote sensing has been tested as an alternative to weather measurements for describing fuel moisture content. This last option has shown a low predictive power on fire danger prediction. In consonance with 5.4.2, remote sensing could provide more information as a predictor of fuel moisture content than directly as a predictor in the probabilistic model. For this reason, a possibility to enhance the performance of *Firelihood* is to use mechanistic variables to express, for example, SWI, DFMC, LFMC. Those variables would substitute the moisture codes and optical remote sensing reflectance bands by specific estimations of the dead and live fuel moisture contents. Mechanistical modelling needs sufficient field observation data for its calibration. In this case, field measurements carried out in the plots from the Reseau Hydrique by the ONF can be used. Current ongoing studies of our research group

(Xiangzhuo et al. in preparation) proposed a model for LFMC at 250m resolution for the Promethée area. This model is trained with a large variety of remote sensing and weather data over the field measurements taken throughout the plots of the Reseau Hydrique. The preliminary results of Xiangzhuo et al. (personal communication) confirm that optical remote sensing does not successfully model the spatio-temporal distribution of LFMC in the study region.

5.4.4. Use reforecast data instead of reanalysis weather data.

Reanalysis climatic data describes the past weather conditions that took place in a spatiotemporal unit. It is built from weather observations which are spatially interpolated to a pixel grid. Because the data represents the real weather conditions, it is useful to explore and understand the drivers of past fire activity, at is has been illustrated in this chapter. Nonetheless, the short-term fire danger predictions are done with weather forecasts. Predicted weather data do not always completely match the posterior observations. Using a model trained with reanalysis weather data for short term predictions with forecast weather data can induce a bias on fire activity prediction. One solution to improve the short-term predictability is to use reforecast data to train supplementary models. This reforecast data is similar to the reanalysis data, although it is based on the historical weather forecasts. It is possible to choose different time lapses of forecast, namely the number of days between the forecasting calculation and the forecasted day. In this thesis the reforecast data has not been used, but all the developed methods are transferable. Such an approach is currently explored in the frame of the FIRE-RES project at the European scale in our research group. It would also be explored for France in the frame of the starting "Climatic Risk" project of PEPR IRIMA, in which our research group and CNRM are involved.

5.4.5. Exploring the drivers of winter fires

During the winter season the fire activity is lower than during the summer season (Fig. 8), but it is significant and raises specific issues in mountainous areas. The drivers of fire occurrence and size are expected to be different during the colder months. In Chapters 3 and 4 the model has been trained specifically for summer fire activities. Provisional tests have been proposed on weather factors and STEs of winter fires to determine the winter

fires. The methods used to fit the probabilistic model are equal between both seasons, only changing the size threshold for the exceedance probability. In the winter season, the exceedance probability is calculated for a threshold of 20ha because of the lower fire numbers smaller fire sizes during winter. These tests have not been explored yet, due to lack of time. The results will be produced and discussed after the end of the thesis project, which may provide useful insights for operational predictions in the same manner as described in 5.3.1.

5.4.6. More accurate assessment of the impact of climate change

In the research team, similar models, based on the same framework presented in this thesis (including Pimont et al. 2022 for South-eastern France), have been used to inform a national inspection report [1] aiming at anticipating changes to address to fire prevention and suppression to adapt to climate changes. The corresponding INRAE study is described in detail in Pimont et al. (2023). Results from Pimont et al., (2022) show an increase of fire activity levels in both the spatial and the temporal dimensions (Figure 21). For the global warming scenario of +4°C we expect that the spatial coverage of high to extreme levels of fire activity increase from the current 8% to 30% (Fig. 21a). For this same scenario, the ~40% of days (67 days) within the summer season are classified as very high to extreme fire activity levels, while the core of the reference fire season is 38 days of high fire activity (Fig. 21b). Overall, the +4°C scenario could entail an increase of +160% for the seasonal burnt area within our study area.

[1] The report is available at: https://www.interieur.gouv.fr/fr/Publications/Rapports-de-l-IGA/Rapports-recents/Politique-de-prevention-et-de-lutte-contre-l-incendie-de-foret-dans-un-contexte-d-extension-et-d-intensification-du-risque-du-au-changement-climatique (Last visited on 3/01/2025).

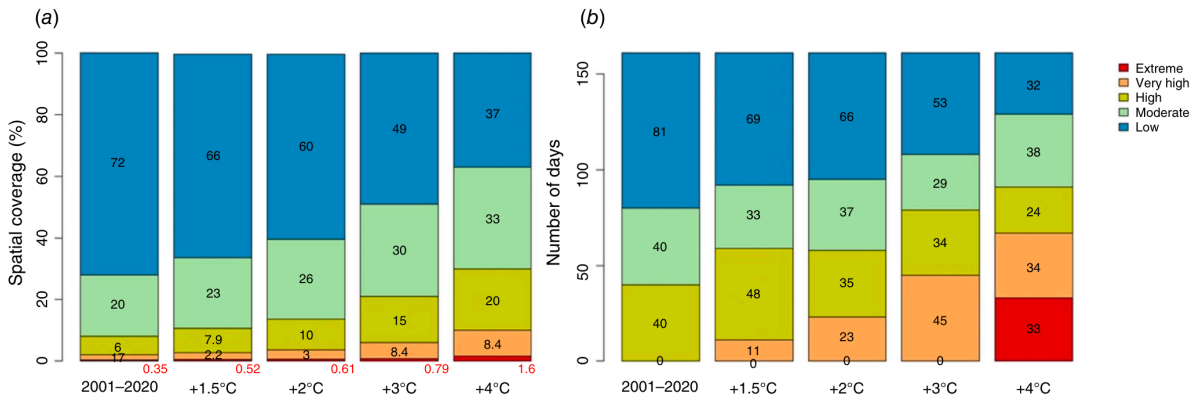


Figure 21: From Pimont et al. (2022). Changes in distribution of (a) spatial, and (b) seasonal fire activity levels for the degrees of global warming. The reference period is 2001 to 2020 which describes the present levels of fire activity. The projections of fire activity levels are shown for four mean global temperatures.

The study of Pimont et al. (2022) was based on the Firelihood framework and used CanFWI as main predictor. We can expect to improve the accuracy of the projections using the models developed in this thesis, especially with the developments from chapter 3. In that chapter we studied the climatic drivers of fire activities, the resulting model accounts for further details on fire-weather. In addition, this newer version of the projections will be calculated using a more recent version of the climate projections – CMIP6 (Eyring et al., 2016) – instead of the CMIP5 (Taylor et al., 2012) used in Pimont et al. 2022). Finally, the projections could include the future trends of winter fires (see 5.4.5), which is an important addition since seasonal lengthening under +4°C indicates that the moderate fire activity levels could extend into the winter season.

5.4.7. Further perspectives

To date, *Firelihood* produced predictions for the fire size associated with ignitions, but it could be extended to any fire characteristic as long as the observations data is available. Fire intensity and fire severity are valuable candidates for this extension. Fire intensity could become available at relevant scales (or spatio-temporal resolution) thanks to more resolved Fire Radiative Power[2] remote sensing products (e.g., Sentinel-3, Meteosat). Fire severity is already available through remote sensing indices such as dNBR (Normalized Burn Ratio). Those possible model outputs need calibration and validation with further

[2] Fire Radiative Power is the measure of the energy (Watts) released by the fire.

study of the available data and predictor variables. Such adaptations will be explored in PEPR IRIMA and European project GREENEO, in collaboration with CNRM.

Firelihood is used in other geographical areas of similar extent as Promethee and larger. The framework is applied for the study of fire activities in the Landes, a woodland dominated region of South-western France. Moreover, it is used at European level, within the European project FIRE-RES, to study the extreme events, and it will be applied at the country level (Metropolitan France) to assess the exposure to fires and risk in all municipalities. Finally, the models will be updated periodically with each release of updated data, including the fire observations, weather, LULC and remote sensing data. This will enhance the accuracy of the models as larger datasets will be used to train the models, but also will take into account the evolution of the fire regime and fire drivers over time. The knowledge built in this thesis may be beneficial for those parallel applications of the *Firelihood*.

6. Conclusion

Wildfires in the French Mediterranean region is a climatic risk with many interactions with human activities, which creates concern to the population and the territorial managers. Therefore, the prediction of fire danger throughout the fire-prone Mediterranean region is key to reduce the consequences of fire activities and adapt the human activities to reduce the human-caused fire ignitions. This thesis presented six questions and three specific objectives aiming to find effective fire danger prediction methods. This led to the exploration of new data sources for the region of study though a probabilistic modelling framework, the development of a new model performance evaluation protocol, and the improvement of a fire activity prediction model.

The thesis has been structured in 4 chapters, which start by the exposition of the background of fire danger and its prediction in chapter 1. This chapter presented the main fire drivers at different spatial and temporal scales and the existing methods for fire danger prediction. In addition, it stated the benefits of using a probabilistic framework both from a theoretical base and through examples. Chapter 2 quantified the importance of different landscape factors (including land use, land cover and human activities) in the spatial distribution of fire activities. This chapter concluded that, despite the importance of LULC in the spatial distribution of fire activities, the evolution of the fire regime between the decade of 1993-2002 and the decade of 2009-2018 is not by changes in the landscape or the fire-weather, but rather by changes in public policy and other societal factors. Chapter 3 presented the development of a new model evaluation methodology, based on the Log Loss score, to compare the performance of fire-weather variables in the prediction of daily fire danger. The outcome of the chapter is that the sub-indices of the CanFWI (FFMC, DMC and DC) in combination with weather variable (VPD, Tmax) are the variables that maximize the prediction performance of fire danger, but that the contribution of LULC and other spatial factor is critical for reproducing consistently observed spatio-temporal patterns of fire activities. Chapter 4 analyzed optical remote sensing data as a driver of fire danger following the methodological advancements from chapter 3. In this chapter we found that the combination of near infrared and short-wave infrared has the largest effect on fire danger prediction.

The model simulations have been used to disentangle the relative contributions of the chosen predictor variables to the fire danger predictions. From this analysis we highlight the importance of the LULC factors to the spatial distribution of fire activity. Also, fire-weather has a larger influence in the 80ha threshold exceedance probability (P80) than in the 1ha fire occurrences (N1). Seasonal variance of N1 is driven by short-term weather factors; in P80 the long-term weather and remote sensing factors contribute more than short-term weather. The interannual variations between the decade of 1993-2002 and the decade of 2009-2018 are driven by non-explained factors captured through the STEs. The interannual variation along the 2014-2022 period is very small and driven by fire weather. The tested optical remote sensing variables brought mainly spatial information for the N1. Optical remote sensing variables contributed more to the prediction of the 80ha threshold exceedance probabilities than to N1. In this case they brought both spatial and seasonal information.

The probabilistic modelling framework has shown its capabilities to combine the effects of highly heterogeneous sources of data (fire-weather, LULC and satellite optical remote sensing). Each source carries specific, but generally interpretable information about the factors that drive fire activities, combining spatial and temporal scales to obtain a comprehensive description of the French Mediterranean fire regimes. In addition, this framework has allowed to include random spatial and temporal effects that complete the modelling regarding explanatory factors which were not straightforward to quantify (e.g., fire suppression) thanks to random spatial effects. Those effects capture the information that has not been described by the explicit variables added to the model.

We developed a specific approach for model evaluation and variable selection adapted to rare and stochastic events in a context where seasonal and spatial patterns are very strong. This method consisted of the comparison between a baseline and each modelling test. The baseline is fitted to describe the average fire season throughout the study area based only on the day of the year (DOY) and the geographical position. It can be viewed as the equivalent of modeled climographs in the context of fire activity. The spatio-temporal random effects have demonstrated to be a very powerful tool to enhance the fire danger predictions at the different scales explored in this study. Nevertheless, to attain a high level of accuracy they need to be trained with a large dataset, which is not always

available. As a result, the spatio-temporal effects have been adapted to the specific situation in each chapter so, we have obtained valuable information from each context.

This work is promising in terms to preoperational results and offers numerous perspectives. The potential applications for the *Firelihood* framework can be separated in two categories. The first one addresses short-term danger forecasting, which could serve to the operational forces. This first group of operational prototypes will provide daily to weekly forecasts of fire danger at a spatial scale that is known by the end users. The second one addresses scientific and policy applications, by providing the long-term trends of the fire regime and the mapping of fire hazard at the present day and in the future. Finally, several perspectives of improvement for the future research have been identified, as well as potential extensions of the probabilistic framework. The acquisition of higher resolution land use and land cover data, the further test of the remote sensing products and the analysis of winter fires are among the future improvements.

Bibliography

This reference list shows the bibliographical references that appear in the text outside the three attached articles.

Abatzoglou, J. T., & Kolden, C. A. (2013). Relationships between climate and macroscale area burned in the western United States. International Journal of Wildland Fire, 22(7), 1003. https://doi.org/10.1071/WF13019

Abdollahi, A., & Yebra, M. (2023). Forest fuel type classification: Review of remote sensing techniques, constraints and future trends. Journal of Environmental Management, 342, 118315. https://doi.org/10.1016/j.jenvman.2023.118315

Acácio, V., Holmgren, M., Rego, F., Moreira, F., & Mohren, G. M. J. (2009). Are drought and wildfires turning Mediterranean cork oak forests into persistent shrublands? Agroforestry Systems, 76(2), 389–400. https://doi.org/10.1007/s10457-008-9165-y

Ager, A. A., Barros, A. M. G., Day, M. A., Preisler, H. K., Spies, T. A., & Bolte, J. (2018). Analyzing fine-scale spatiotemporal drivers of wildfire in a forest landscape model. Ecological Modelling, 384, 87–102. https://doi.org/10.1016/j.ecolmodel.2018.06.018

Alessandri, A., De Felice, M., Zeng, N., Mariotti, A., Pan, Y., Cherchi, A., Lee, J.-Y., Wang, B., Ha, K.-J., Ruti, P., & Artale, V. (2014). Robust assessment of the expansion and retreat of Mediterranean climate in the 21st century. Scientific Reports, 4(1), 7211. https://doi.org/10.1038/srep07211

Alves, D., Almeida, M., Viegas, D. X., Novo, I., & Luna, M. Y. (2021). Fire Danger Harmonization Based on the Fire Weather Index for Transboundary Events between Portugal and Spain. Atmosphere, 12(9), Article 9. https://doi.org/10.3390/atmos12091087

Argañaraz, J. P., Landi, M. A., Scavuzzo, C. M., & Bellis, L. M. (2018). Determining fuel moisture thresholds to assess wildfire hazard: A contribution to an operational early warning system. PLOS ONE, 13(10), e0204889. https://doi.org/10.1371/journal.pone.0204889

Baijnath-Rodino, J. A., Foufoula-Georgiou, E., & Banerjee, T. (2020). Reviewing the "Hottest" Fire Indices Worldwide [Preprint]. Climatology (Global Change). https://doi.org/10.1002/essoar.10503854.1

Balch, J. K., Bradley, B. A., Abatzoglou, J. T., Nagy, R. C., Fusco, E. J., & Mahood, A. L. (2017). Human-started wildfires expand the fire niche across the United States. Proceedings of the National Academy of Sciences of the United States of America, 114(11), 2946–2951. https://doi.org/10.1073/pnas.1617394114

Barbero, R., Abatzoglou, J. T., Larkin, N. K., Kolden, C. A., & Stocks, B. (2015). Climate change presents increased potential for very large fires in the contiguous United States. International Journal of Wildland Fire, 24(7), 892. https://doi.org/10.1071/WF15083

Barbero, R., Curt, T., Ganteaume, A., Maillé, E., Jappiot, M., & Bellet, A. (2019). Simulating the effects of weather and climate on large wildfires in France. Natural Hazards and Earth System Sciences, 19(2), 441–454. https://doi.org/10.5194/nhess-19-441-2019

Bennett, M., Fitzgerald, S., Parker R. (B.), Main M., Perleberg, A., Schnepf, C. C., Mahoney R. (2010). Reducing Fire Risk on Your Forest Property. Oregon State University Extension Service.

Bou Dagher Kharrat, M., Martinez De Arano, I., Zeki-Bašken, E., Feder, S., Adams, S., Briers, S., Fady, B., Lefèvre, F., Górriz-Mifsud, E., Mauri, E., Mavsar, R., Muys, B., Nocentini, S., Pettenella, D., Peñuelas, J., Sardans, J., Secco, L., Travaglini, D., Tripodi, G., … Bozzano, M. (2022). Mediterranean Forest Research Agenda 2030. European Forest Institute. https://doi.org/10.36333/rs5

Bradshaw, R. H. W. (2004). Past anthropogenic influence on European forests and some possible genetic consequences. Forest Ecology and Management, 197(1), 203–212. https://doi.org/10.1016/j.foreco.2004.05.025

Brillinger, D. R., Preisler, H. K., & Benoit, J. W. (2003). Risk assessment: A forest fire example. In Institute of Mathematical Statistics Lecture Notes—Monograph Series (pp. 177–196). Institute of Mathematical Statistics. https://doi.org/10.1214/lnms/1215091142

Brillinger, D. R., Preisler, H. K., & Benoit, J. W. (2006). Probabilistic risk assessment for wildfires. Environmetrics, 17(6), 623–633. https://doi.org/10.1002/env.768

Bulliet, Richard W., Crossley, Pamela Kyle, Headrick, Daniel R., Hirsch, Steven W., Johnson, Lyman L., & Northrup, David. (2008). The Earth and Its Peoples. A Global History. Third edition.

Burley, J. (2004). Encyclopedia of Forest Sciences. http://www.sciencedirect.com:5070/referencework/9780121451608/encyclopedia-of-forest-sciences

Carrega, P. (1991). A Meteorological Index of Forest Fire Hazard in Mediterranean France. International Journal of Wildland Fire, 1(2), 79. https://doi.org/10.1071/WF9910079

Casanova Enault, L., Popoff, T., & Debolini, M. (2021). Vacant lands on French Mediterranean coastlines: Inventory, agricultural opportunities, and prospective scenarios. Land Use Policy, 100, 104914. https://doi.org/10.1016/j.landusepol.2020.104914

Chuvieco, E., González, I., Verdú, F., Aguado, I., & Yebra, M. (2009). Prediction of fire occurrence from live fuel moisture content measurements in a Mediterranean ecosystem. International Journal of Wildland Fire, 18(4), 430. https://doi.org/10.1071/WF08020

Costafreda-Aumedes, S., Comas, C., & Vega-Garcia, C. (2017). Human-caused fire occurrence modelling in perspective: A review. International Journal of Wildland Fire, 26(12), 983. https://doi.org/10.1071/WF17026

Coughlan, R., Di Giuseppe, F., Vitolo, C., Barnard, C., Lopez, P., & Drusch, M. (2021). Using machine learning to predict fire-ignition occurrences from lightning forecasts. Meteorological Applications, 28(1), e1973. https://doi.org/10.1002/met.1973

Cruz, M. G., Hurley, R. J., Bessell, R., & Sullivan, A. L. (2020). Fire behaviour in wheat crops – effect of fuel structure on rate of fire spread. International Journal of Wildland Fire, 29(3), 258–271. https://doi.org/10.1071/WF19139

Curt, T., & Frejaville, T. (2018). Wildfire Policy in Mediterranean France: How Far is it Efficient and Sustainable? Risk Analysis, 38(3), 472–488. https://doi.org/10.1111/risa.12855

Curt, T., Fréjaville, T., & Bouillon, C. (2014). Characterizing pyroregions in south-eastern France. In Advances in Forest Fire Research. Chapter 4—Fire Risk Assessment and Climate Change (pp. 1093–1101). Imprensa da Universidade de Coimbra. https://doi.org/10.14195/978-989-26-0884-6_119

DaCamara, C. C., Calado, T. J., Ermida, S. L., Trigo, I. F., Amraoui, M., & Turkman, K. F. (2014). Calibration of the Fire Weather Index over Mediterranean Europe based on fire activity retrieved from MSG satellite imagery. International Journal of Wildland Fire, 23(7), 945. https://doi.org/10.1071/WF13157

de Rigo, D., Liberta', G., Durrant, T., Artes Vivancos, T., & San-Miguel-Ayanz, J. (2018, January 18). Forest fire danger extremes in Europe under climate change: Variability and uncertainty. JRC Publications Repository. https://doi.org/10.2760/13180

Di Giuseppe, F., Vitolo, C., Barnard, C., Libertá, G., Maciel, P., San-Miguel-Ayanz, J., Villaume, S., & Wetterhall, F. (2024). Global seasonal prediction of fire danger. Scientific Data, 11(1), 128. https://doi.org/10.1038/s41597-024-02948-3

Díaz-Avalos, C., Juan, P., & Serra-Saurina, L. (2016). Modeling fire size of wildfires in Castellon (Spain), using spatiotemporal marked point processes. Forest Ecology and Management, 381, 360–369. https://doi.org/10.1016/j.foreco.2016.09.013

Drouet, J. C., & Sol, B. (1993). Mise au point d'un indice numérique de risque météorologique d'incendies de forêts. Dossier: Forêts et incendies 1991-1992. Forêt Méditerranéenne, XIV(2), 155–162.

Duane, A., Castellnou, M., & Brotons, L. (2021). Towards a comprehensive look at global drivers of novel extreme wildfire events. Climatic Change, 165(3–4), 43. https://doi.org/10.1007/s10584-021-03066-4

Erik A. Beever, Suresh Andrew Sethi, Suzanne Prange, & Dominick DellaSala. (2019). Introduction: Defining and interpreting ecological disturbances | U.S. Geological Survey.

Eastaugh, C. S., Arpaci, A., & Vacik, H. (2012). A cautionary note regarding comparisons of fire danger indices. Natural Hazards and Earth System Sciences, 12(4), 927–934. https://doi.org/10.5194/nhess-12-927-2012

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geoscientific Model Development, 9(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Fargeon, H. (2019). Effet du changement climatique sur l'évolution de l'aléa incendie de forêt en France métropolitaine au 21ème siècle.

FCFDG. (1992). Development and structure of the Canadian Forest Fire Behavior Prediction System. Canadian Forest Service.

Fernandes, P. M., Vega, J. A., Jiménez, E., Rigolot. E. (2008). Fire resistance of European pines. Forest Ecology and Management 256. 246-55. https://doi.org/10.1016/j.foreco.2008.04.032

Fréjaville, T., & Curt, T. (2017). Seasonal changes in the human alteration of fire regimes beyond the climate forcing. Environmental Research Letters, 12(3), 035006. https://doi.org/10.1088/1748-9326/aa5d23

Gale, M. G., Cary, G. J., Van Dijk, A. I. J. M., & Yebra, M. (2021). Forest fire fuel through the lens of remote sensing: Review of approaches, challenges and future directions in the remote sensing of biotic determinants of fire behaviour. Remote Sensing of Environment, 255, 112282. https://doi.org/10.1016/j.rse.2020.112282

Galizia, L. F., Barbero, R., Rodrigues, M., Ruffault, J., Pimont, F., & Curt, T. (2023). Global Warming Reshapes European Pyroregions. Earth's Future, 11(5), e2022EF003182. https://doi.org/10.1029/2022EF003182

Ganteaume, A., & Barbero, R. (2019). Contrasting large fire activity in the French Mediterranean. Natural Hazards and Earth System Sciences, 19(5), 1055–1066. https://doi.org/10.5194/nhess-19-1055-2019

Ganteaume, A., Camia, A., Jappiot, M., San Miguel-Ayanz, J., Long-Fournel, M., & Lampin, C. (2013). A Review of the Main Driving Factors of Forest Fire Ignition Over Europe. Environmental Management, 51(3), 651–662. https://doi.org/10.1007/s00267-012-9961-z

Garcia-Prats, A., Antonio, D. C., Tarcísio, F. J. G., & Antonio, M. J. (2015). Development of a Keetch and Byram—Based drought index sensitive to forest management in Mediterranean conditions. Agricultural and Forest Meteorology, 205, 40–50. https://doi.org/10.1016/j.agrformet.2015.02.009

González-De Vega, S., De las Heras, J., & Moya, D. (2016). Resilience of Mediterranean terrestrial ecosystems and fire severity in semiarid areas: Responses of Aleppo pine forests in the short, mid and long term. Science of the Total Environment, 573, 1171–1177. Scopus. https://doi.org/10.1016/j.scitotenv.2016.03.115

Hawbaker, T. J., Radeloff, V. C., Stewart, S. I., Hammer, R. B., Keuler, N. S., & Clayton, M. K. (2013). Human and biophysical influences on fire occurrence in the United States. Ecological Applications, 23(3), 565–582. https://doi.org/10.1890/12-1816.1

Hollis, J. J., Matthews, S., Fox-Hughes, P., Grootemaat, S., Heemstra, S., Kenny, B. J., & Sauvage, S. (2024). Introduction to the Australian Fire Danger Rating System†. International Journal of Wildland Fire, 33(3). https://doi.org/10.1071/WF23140

Huang, J., Malone, B. P., Minasny, B., McBratney, A. B., Triantafilis, J. (2017). Evaluating a Bayesian modelling approach (INLA-SPDE) for environmental mapping. Science of The Total Environment 609. https://doi.org/10.1016/j.scitotenv.2017.07.201.

Hurteau, M. D., Goodwin, M. J., Marsh, C., Zald, H. S., Collins, B., Meyer, M., & North, M. P. (2024). Managing fire-prone forests in a time of decreasing carbon carrying capacity. Frontiers in Ecology and the Environment, e2801. https://doi.org/10.1002/fee.2801

Jolly, W. M., Freeborn, P. H., Bradshaw, L. S., Wallace, J., & Brittain, S. (2024). Modernizing the US National Fire Danger Rating System (version 4): Simplified fuel models and improved live and dead fuel moisture calculations. Environmental Modelling & Software, 106181. https://doi.org/10.1016/j.envsoft.2024.106181

Jones, G. M., & Tingley, M. W. (2022). Pyrodiversity and biodiversity: A history, synthesis, and outlook.

Jones, M. W., Abatzoglou, J. T., Veraverbeke, S., Andela, N., Lasslop, G., Forkel, M., Smith, A. J. P., Burton, C., Betts, R. A., van der Werf, G. R., Sitch, S., Canadell, J. G., Santín, C., Kolden, C., Doerr, S. H., & Le Quéré, C. (2022). Global and Regional Trends and Drivers of Fire Under Climate Change. Reviews of Geophysics, 60(3), e2020RG000726. https://doi.org/10.1029/2020RG000726

Kelly, L. T., Giljohann, K. M., Duane, A., Aquilué, N., Archibald, S., Batllori, E., Bennett, A. F., Buckland, S. T., Canelles, Q., Clarke, M. F., Fortin, M.-J., Hermoso, V., Herrando, S., Keane, R. E., Lake, F. K., McCarthy, M. A., Morán-Ordóñez, A., Parr, C. L., Pausas, J. G., … Brotons, L. (2020). Fire and biodiversity in the Anthropocene. Science, 370(6519), eabb0355. https://doi.org/10.1126/science.abb0355

Lecina-Diaz, J., Alvarez, A., & Retana, J. (2014). Extreme Fire Severity Patterns in Topographic, Convective and Wind-Driven Historical Wildfires of Mediterranean Pine Forests. PLOS ONE, 9(1), e85127. https://doi.org/10.1371/journal.pone.0085127

Lindgren, F., Rue H., Lindst J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: The SPDE approach.

Lindgren, F., Rue H. (2015) Bayesian Spatial Modelling with R - INLA ». Journal of Statistical Software 63. https://doi.org/10.18637/jss.v063.i19.

Liu, N., Deng, Z., Wang, H., Luo, Z., Gutiérrez-Jurado, H. A., He, X., & Guan, H. (2020). Thermal remote sensing of plant water stress in natural ecosystems. Forest Ecology and Management, 476, 118433. https://doi.org/10.1016/j.foreco.2020.118433

Martin-Ducup, O., Dupuy, J.-L., Soma, M., Guerra-Hernandez, J., Marino, E., Fernandes, P. M., Just, A., Corbera, J., Toutchkov, M., Sorribas, C., Bock, J., Piboule, A., Pirotti, F., & Pimont, F. (2025). Unlocking the potential of Airborne LiDAR for direct assessment of fuel bulk density and load distributions for wildfire hazard mapping. Agricultural and Forest Meteorology, 362, 110341. https://doi.org/10.1016/j.agrformet.2024.110341

McArthur, A. G. (1967). Fire behaviour in eucalypt forests: Vol. Leaflet No. 107 (Forestry and Timber Bureau). Commonwealth Department of National Development.

Millennium Ecosystem Assessment (Ed.). (2005). Ecosystems and human well-being: Synthesis. Island Press.

Moreira, F., Vaz, P., Catry, F., & Silva, J. S. (2009). Regional variations in wildfire susceptibility of land-cover types in Portugal: Implications for landscape management to minimize fire hazard. International Journal of Wildland Fire, 18(5), 563–574. https://doi.org/10.1071/WF07098

Moreira, F., Viedma, O., Arianoutsou, M., Curt, T., Koutsias, N., Rigolot, E., Barbati, A., Corona, P., Vaz, P., Xanthopoulos, G., Mouillot, F., & Bilgili, E. (2011). Landscape – wildfire interactions in southern Europe: Implications for landscape management. Journal of Environmental Management, 92(10), 2389–2402. https://doi.org/10.1016/j.jenvman.2011.06.028

Moritz, M. A., Morais, M. E., Summerell, L. A., Carlson, J. M., & Doyle, J. (2005). Wildfires, complexity, and highly optimized tolerance. Proceedings of the National Academy of Sciences, 102(50), 17912–17917. https://doi.org/10.1073/pnas.0508985102

Ne'eman, G., Lev-Yadun, S., Arianotsou M. (2012). Fire-related traits in Mediterranean basin plants.

Nolan, R. H., Resco De Dios, V., Boer, M. M., Caccamo, G., Goulden, M. L., & Bradstock, R. A. (2016). Predicting dead fine fuel moisture at regional scales using vapour pressure deficit from MODIS and gridded weather data. Remote Sensing of Environment, 174, 100–108. https://doi.org/10.1016/j.rse.2015.12.010

Orieux, A. (1974). Conditions météorologiques et incendies en région méditerranéenne. Revue Forestière Française, S, 122. https://doi.org/10.4267/2042/20873

Parisien, M.-A., Snetsinger, S., Greenberg, J., Nelson, C., Schoennagel, T., Dobrowski, S., & Moritz, M. (2012). Spatial variability in wildfire probability across the western United States. International Journal of Wildland Fire, 21, 313–327. https://doi.org/10.1071/WF11044

Pausas, J. G., & Fernández-Muñoz, S. (2012). Fire regime changes in the Western Mediterranean Basin: From fuel-limited to drought-driven fire regime. Climatic Change, 110(1), 215–226. https://doi.org/10.1007/s10584-011-0060-6

Pausas, J. G., & Ribeiro, E. (2013). The global fire–productivity relationship. Global Ecology and Biogeography, 22(6), 728–736. https://doi.org/10.1111/geb.12043

Pausas, J. G., & Keeley, J. E. (2019). Wildfires as an ecosystem service. Frontiers in Ecology and the Environment, 17(5), 289–295. https://doi.org/10.1002/fee.2044

Pettinari, M. L., & Chuvieco, E. (2020). Fire Danger Observed from Space. Surveys in Geophysics, 41(6), 1437–1459. https://doi.org/10.1007/s10712-020-09610-8

Phelps, N., & Woolford, D. G. (2021). Guidelines for effective evaluation and comparison of wildland fire occurrence prediction models. International Journal of Wildland Fire, 30(4), 225. https://doi.org/10.1071/WF20134

Pimont, F., Dupuy, J.-L., Ruffault, J., Rigolot, E., Opitz, T., Legrand, J., & Barbero, R. (2023). Projections des effets du changement climatique sur l'activité des feux de forêt au 21ème siècle: Rapport final [Report, INRAE]. https://doi.org/10.17180/gpdj-xb05

Pimont, F., Fargeon, H., Opitz, T., Ruffault, J., Barbero, R., Martin-StPaul, N., Rigolot, E., RiviÉre, M., & Dupuy, J. (2021). Prediction of regional wildfire activity in the probabilistic Bayesian framework of Firelihood. Ecological Applications, 31(5), e02316. https://doi.org/10.1002/eap.2316

Pimont, F., Parsons, R., Rigolot, E., de Coligny, F., Dupuy, J.-L., Dreyfus, P., & Linn, R. R. (2016). Modeling fuels and fire effects in 3D: Model description and applications. Environmental Modelling & Software, 80, 225–244. https://doi.org/10.1016/j.envsoft.2016.03.003

Pimont, F., Ruffault, J., Martin-StPaul, N. K., & Dupuy, J.-L. (2019). Why is the effect of live fuel moisture content on fire rate of spread underestimated in field experiments in

shrublands? International Journal of Wildland Fire, 28(2), 127–137. https://doi.org/10.1071/WF18091

Pimont, F., Ruffault, J., Opitz, T., Fargeon, H., Barbero, R., Castel-Clavera, J., Martin-StPaul, N., Rigolot, E., & Dupuy, J.-L. (2022). Future expansion, seasonal lengthening and intensification of fire activity under climate change in southeastern France. International Journal of Wildland Fire, 32(1), 4–14. https://doi.org/10.1071/WF22103

Piñol, J., Beven, K., & Viegas, D. X. (2005). Modelling the effect of fire-exclusion and prescribed fire on wildfire size in Mediterranean ecosystems. Ecological Modelling, 183(4), 397–409. https://doi.org/10.1016/j.ecolmodel.2004.09.001

Preisler, H. K., Brillinger, D. R., Burgan, R. E., & Benoit, J. W. (2004). Probability based models for estimation of wildfire risk. International Journal of Wildland Fire, 13(2), 133. https://doi.org/10.1071/WF02061

Preisler, H. K., & Westerling, A. L. (2007). Statistical Model for Forecasting Monthly Large Wildfire Events in Western United States. Journal of Applied Meteorology and Climatology, 46(7), 1020–1030. https://doi.org/10.1175/JAM2513.1

Quan, X., Yebra, M., Riaño, D., He, B., Lai, G., & Liu, X. (2021). Global fuel moisture content mapping from MODIS. International Journal of Applied Earth Observation and Geoinformation, 101, 102354. https://doi.org/10.1016/j.jag.2021.102354

Rao, K., Williams, A. P., Diffenbaugh, N. S., Yebra, M., & Konings, A. G. (2022). Plant-water sensitivity regulates wildfire vulnerability. Nature Ecology & Evolution, 6(3), 332–339. https://doi.org/10.1038/s41559-021-01654-2

Rao, K., Williams, A. P., Flefil, J. F., & Konings, A. G. (2020). SAR-enhanced mapping of live fuel moisture content. Remote Sensing of Environment, 245, 111797. https://doi.org/10.1016/j.rse.2020.111797

Roebroeks, W., & Villa, P. (2011). On the earliest evidence for habitual use of fire in Europe. Proceedings of the National Academy of Sciences, 108(13), 5209–5214. https://doi.org/10.1073/pnas.1018116108

Romero, B., & Ganteaume, A. (2020). Does recent fire activity impact fire-related traits of Pinus halepensis Mill. And Pinus sylvestris L. in the French Mediterranean area? Annals of Forest Science, 77(4), 106. https://doi.org/10.1007/s13595-020-01016-1

Ruffault, J., Curt, T., Moron, V., Trigo, R. M., Mouillot, F., Koutsias, N., Pimont, F., Martin-StPaul, N., Barbero, R., Dupuy, J.-L., Russo, A., & Belhadj-Khedher, C. (2020). Increased

likelihood of heat-induced large wildfires in the Mediterranean Basin. Scientific Reports, 10(1), 13790. https://doi.org/10.1038/s41598-020-70069-z

Ruffault, J., Moron, V., Trigo, R. M., & Curt, T. (2017). Daily synoptic conditions associated with large fire occurrence in Mediterranean France: Evidence for a wind-driven fire regime: DAILY SYNOPTIC CONDITIONS ASSOCIATED WITH LARGE FIRE OCCURRENCE. International Journal of Climatology, 37(1), 524–533. https://doi.org/10.1002/joc.4680

Ruffault, J., & Mouillot, F. (2015). How a new fire-suppression policy can abruptly reshape the fire-weather relationship. Ecosphere, 6(10), art199. https://doi.org/10.1890/ES15-00182.1

Ruffault, J., Pimont, F., Cochard, H., Dupuy, J.-L., & Martin-StPaul, N. (2022). SurEau-Ecos v2.0: A trait-based plant hydraulics model for simulations of plant water status and drought-induced mortality at the ecosystem level. Geoscientific Model Development, 15(14), 5593–5626. https://doi.org/10.5194/gmd-15-5593-2022

Saito, T., & Rehmsmeier, M. (2015). The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. PLOS ONE, 10(3), e0118432. https://doi.org/10.1371/journal.pone.0118432

Schlobohm, P., & Brain, J. (2002, June 30). Gaining an Understanding of the National Fire Danger Rating System. https://www.semanticscholar.org/paper/Gaining-an-Understanding-of-the-National-Fire-Schlobohm-Brain/1b49d5fab8a7d5c06697eac5cdb63fb413aeb463

Serra, L., Saez, M., Juan, P., Varga, D., & Mateu, J. (2014). A spatio-temporal Poisson hurdle point process to model wildfires. Stochastic Environmental Research and Risk Assessment, 28(7), 1671–1684. https://doi.org/10.1007/s00477-013-0823-x

Serra, L., Saez, M., Mateu, J., Varga, D., Juan, P., Díaz-Ávalos, C., & Rue, H. (2014). Spatio-temporal log-Gaussian Cox processes for modelling wildfire occurrence: The case of Catalonia, 1994–2008. Environmental and Ecological Statistics, 21(3), 531–563. https://doi.org/10.1007/s10651-013-0267-y

Sharples, J. J., McRae, R. H. D., Weber, R. O., & Gill, A. M. (2009). A simple index for assessing fuel moisture content. Environmental Modelling & Software, 24(5), 637–646. https://doi.org/10.1016/j.envsoft.2008.10.012

Sofaer, H. R., Hoeting, J. A., & Jarnevich, C. S. (2018). The area under the precision-recall curve as a performance metric for rare binary events. Methods in Ecology and Evolution, 10(4), 565–577. https://doi.org/10.1111/2041-210X.13140

Tanase, M. A., Nova, J. P. G., Marino, E., Aponte, C., Tomé, J. L., Yáñez, L., Madrigal, J., Guijarro, M., & Hernando, C. (2022). Characterizing Live Fuel Moisture Content from Active and Passive Sensors in a Mediterranean Environment. Forests, 13(11), 1846. https://doi.org/10.3390/f13111846

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An Overview of CMIP5 and the Experiment Design. https://doi.org/10.1175/BAMS-D-11-00094.1

Taylor, S. W., Woolford, D. G., Dean, C. B., & Martell, D. L. (2013). Wildfire Prediction to Inform Fire Management: Statistical Science Challenges. Statistical Science, 28(4). https://doi.org/10.1214/13-STS451

Turco, M., Von Hardenberg, J., AghaKouchak, A., Llasat, M. C., Provenzale, A., & Trigo, R. M. (2017). On the key role of droughts in the dynamics of summer fires in Mediterranean Europe. Scientific Reports, 7(1), 81. https://doi.org/10.1038/s41598-017-00116-9

Turner, R. (2009). Point patterns of forest fire locations. Environmental and Ecological Statistics, 16(2), 197–223. https://doi.org/10.1007/s10651-007-0085-1

Valladares, F., Benavides, R., Rabasa, S. G., Díaz, M., Pausas, J. G., Paula, S., & Simonson, W. D. (2014). Global change and Mediterranean forests: Current impacts and potential responses. In D. A. Coomes, D. F. R. P. Burslem, & W. D. Simonson (Eds.), Forests and Global Change (pp. 47–76). Cambridge University Press. https://doi.org/10.1017/CBO9781107323506.005

Vallet, L., Mouillot, F., & Ourcival, J.-M. (2022). Agricultural fires in France: A first national overview from data mining. In D. X. Viegas & L. M. Ribeiro, Advances in Forest Fire Research 2022 (1.a Edição, pp. 969–976). Imprensa da Universidade de Coimbra. https://doi.org/10.14195/978-989-26-2298-9_147

Vidal, J., Martin, E., Franchistéguy, L., Baillon, M., & Soubeyroux, J. (2010). A 50-year high-resolution atmospheric reanalysis over France with the Safran system. International Journal of Climatology, 30(11), 1627–1644. https://doi.org/10.1002/joc.2003

Vilar Del Hoyo, L., Martín Isabel, M. P., & Martínez Vega, F. J. (2011). Logistic regression models for human-caused wildfire risk estimation: Analysing the effect of the spatial

accuracy in fire occurrence data. European Journal of Forest Research, 130(6), 983–996. https://doi.org/10.1007/s10342-011-0488-2

Walding, N. G., Williams, H. T. P., McGarvie, S., & Belcher, C. M. (2018). A comparison of the US National Fire Danger Rating System (NFDRS) with recorded fire occurrence and final fire size. International Journal of Wildland Fire, 27(2), 99. https://doi.org/10.1071/WF17030

Wang, L., Hunt, E. R., Qu, J. J., Hao, X., & Daughtry, C. S. T. (2013). Remote sensing of fuel moisture content from ratios of narrow-band vegetation water and dry-matter indices. Remote Sensing of Environment, 129, 103–110. https://doi.org/10.1016/j.rse.2012.10.027

Wong, C. Y. S. (2023). Plant optics: Underlying mechanisms in remotely sensed signals for phenotyping applications. AoB Plants, 15(4), plad039. https://doi.org/10.1093/aobpla/plad039

Woolford, D. G., Bellhouse, D. R., Braun, W. J., Dean, C. B., Martell, D. L., & Sun, J. (2011). A spatio-temporal model for people-caused forest fire occurrence in the Romeo Malette Forest. Journal of Environmental Statistics, 2(1), 26.

Woolford, D. G., Braun, W. J., Dean, C. B., & Martell, D. L. (2009). Site-Specific Seasonal Baselines for Fire Risk in Ontario. GEOMATICA, 63(4), 355–363. https://doi.org/10.5623/geomat-2009-0049

Woolford, D. G., Martell, D. L., McFayden, C. B., Evens, J., Stacey, A., Wotton, B. M., & Boychuk, D. (2021). The development and implementation of a human-caused wildland fire occurrence prediction system for the province of Ontario, Canada. Canadian Journal of Forest Research, 51(2), 303–325. https://doi.org/10.1139/cjfr-2020-0313

Worsnop, R. P., Scheuerer, M., Giuseppe, F. D., Barnard, C., Hamill, T. M., & Vitolo, C. (2021). Probabilistic Fire Danger Forecasting: A Framework for Week-Two Forecasts Using Statistical Postprocessing Techniques and the Global ECMWF Fire Forecast System (GEFF). Weather and Forecasting, 36(6), 2113–2125. https://doi.org/10.1175/WAF-D-21-0075.1

Wotton, B. M., & Beverly, J. L. (2007). Stand-specific litter moisture content calibrations for the Canadian Fine Fuel Moisture Code. International Journal of Wildland Fire, 16(4), 463–472. https://doi.org/10.1071/WF06087

Yang, Y., Uddstrom, M., Pearce, G., & Revell, M. (2015). Reformulation of the Drought Code in the Canadian Fire Weather Index System Implemented in New Zealand. Journal of Applied Meteorology and Climatology, 54(7), 1523–1537. https://doi.org/10.1175/JAMC-D-14-0090.1

Yebra, M., Dennison, P. E., Chuvieco, E., Riaño, D., Zylstra, P., Hunt, E. R., Danson, F. M., Qi, Y., & Jurdao, S. (2013). A global review of remote sensing of live fuel moisture content for fire danger assessment: Moving towards operational products. Remote Sensing of Environment, 136, 455–468. https://doi.org/10.1016/j.rse.2013.05.029

Yebra, M., Quan, X., Riaño, D., Rozas Larraondo, P., Van Dijk, A. I. J. M., & Cary, G. J. (2018). A fuel moisture content and flammability monitoring methodology for continental Australia based on optical remote sensing. Remote Sensing of Environment, 212, 260–272. https://doi.org/10.1016/j.rse.2018.04.053

Supplementary materials

Supplementary materials from the report:

A. Exploratory work for fire danger prediction using MODIS composite data. 215

A. Exploratory work for fire danger prediction using MODIS composite data.

The MOD13Q1 and MYD13Q1 products presented several limitations regarding the temporal resolution and the composite product. Despite the complexity of the data treatment required by the composites, the analysis on the 80ha threshold exceedance probability was carried out using the fire perimeters (see 1.4.2). The results from this preliminary study are shown in the following.

Figure S1 shows the estimated effects for the most performant models tested with MODIS data WA * DC * DMC * FFMC * NIR * MIR * U * STEs. The CFWIS drought codes correspond to the daily maximal value. NIR is the Near Infrared and MIR the medium Infrared. FFMC, DMC and DC (Fig. S1 A, B and C) show a very similar response to the models with Sentinel-2 data. NIR (Fig. S1D) shows the same decreasing effect that it does in the mentioned Sentinel-2 model. MIR (Fig. S1E) has a bell-shaped effect where either minimal or maximal values have an effect below 0. WA (Fig. S1F) has a flat effect while the seasonal effect (Fig. S1G) decreases through the season.

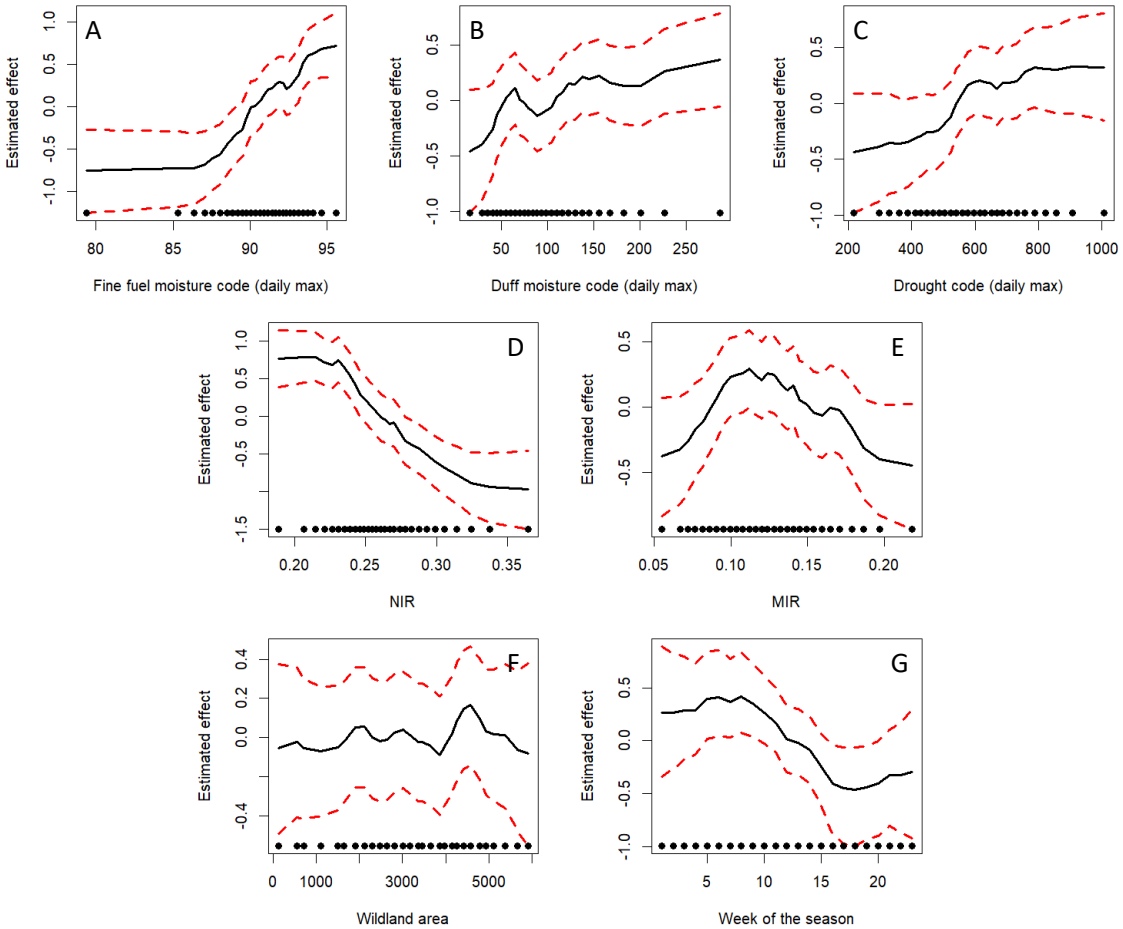


Figure S 1: Effects of the component fitted with DC, DMC and FFMC maximal daily values from the CanFWI (A-C), NIR (D), MIR (E), Wildland Area per pixel (F) and STEs (plot G shows the temporal effect).The X axis shows the value in each variable-specific bin; the Y axis is the contribution of the predictor variable's value to the logit of the probability of exceedance.

Figure S2 shows the effects from the variables of a simpler model using WA * CanFWI(max) * NDVI * STEs. The effect of the NDVI (Fig. S2A) follows a bell shape with a maximal effect between 0,5<NDVI<0,6. The CanFWI (Fig. S2B) has an increasing effect that saturates after CanFWI=70. WA (Fig. S2C) and the season (Fig. S2D) have alsost flat effects with minimal variations.

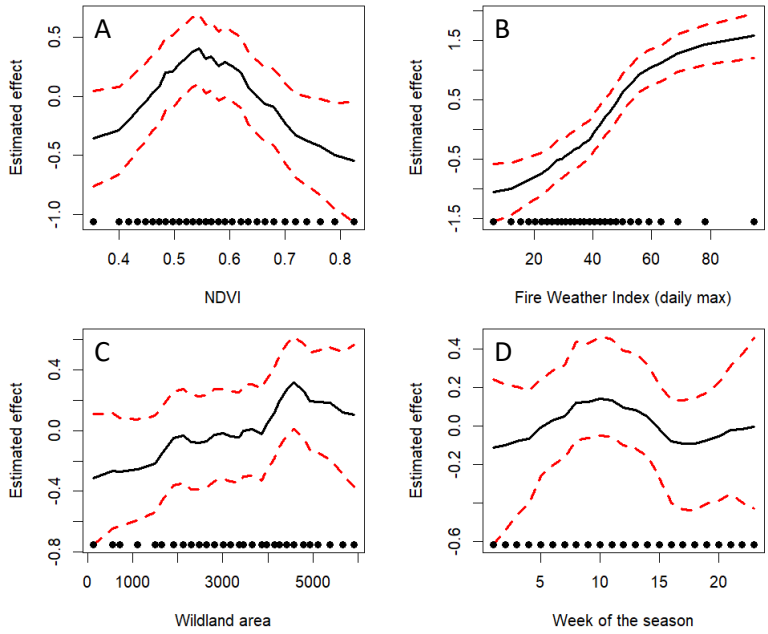


Figure S 2: Effects of the component fitted with NDVI (A), CanFWI (B), Wildland Area per pixel (C) and STEs (plot D shows the temporal effect).The X axis shows the value in each variable-specific bin; the Y axis is the contribution of the predictor variable's value to the logit of the probability of exceedance.

Finally, Table S1 shows the evaluation from all models tested in this study of MODIS data.

Table S 1: (next page) Model evaluation metrics for the size component. The yellow rows correspond to the Null model and the FAb. The table is sorted using the "Skill Scale".

Model	LogLoss	Skill Scale	DIC
NULL	0,2100	-143,9	1761
LAI	0,1803	-65,3	1750
MIR	0,1766	-55,3	1736
RED	0,1752	-51,6	1733
Fpar	0,1731	-46,0	1748
NDVI	0,1689	-34,8	1731
EVI	0,1650	-24,7	1726
NIR	0,1621	-17,1	1703
Fab	0,1557	0,0	1682
LAI * STEs	0,1506	13,5	1682
Fpar * STEs	0,1494	16,8	1683
FWImax	0,1474	22,0	1615
NDVI * STEs	0,1474	22,0	1668
DC * DMC * FFMCmax	0,1471	22,8	1658
MIR * STEs	0,1469	23,4	1667
EVI * STEs	0,1463	24,9	1663
RED * STEs	0,1453	27,5	1659
DC * DMC * FFMCmax * NDVI	0,1427	34,4	1640
FWImax * NDVI	0,1423	35,6	1597
FWImax * NDVI * U	0,1418	36,8	1600
DC * DMC * FFMCmax * EVI	0,1416	37,3	1638
FWImax * EVI	0,1407	39,7	1596
FWImax * EVI * U	0,1398	42,3	1599
NIR * STEs	0,1394	43,2	1621
DC * DMC * FFMCmax * U	0,1379	47,2	1607
DC * DMC * FFMCmax * Tmax * U	0,1374	48,5	1613
DC * DMC * FFMCmax * STEs	0,1369	49,9	1621
DC * DMC * FFMCmax * NIR	0,1362	51,8	1597
FWImax * NIR	0,1360	52,2	1566
FWImax * STEs	0,1349	55,3	1576
FWImax * NIR * U	0,1348	55,3	1567
DC * DMC * FFMCmax * NDVI * U	0,1347	55,8	1591
DC * DMC * FFMCmax * EVI * U	0,1345	56,3	1593
DC * DMC * FFMCmax * Tmax * NDVI * U	0,1344	56,5	1597
DC * DMC * FFMCmax * Tmax * EVI * U	0,1341	57,3	1599
DC * DMC * FFMCmax * EVI * STEs	0,1340	57,7	1606
DC * DMC * FFMCmax * NIR * MIR	0,1340	57,7	1589
DC * DMC * FFMCmax * NDVI * STEs	0,1339	58,0	1605
FWImax * NIR * MIR	0,1332	59,8	1555
FWImax * NIR * MIR * U	0,1319	63,0	1557
FWImax * EVI * STEs	0,1316	63,8	1559
FWImax * NDVI * STEs	0,1312	65,1	1560
DC * DMC * FFMCmax * NIR * U	0,1298	68,6	1556
FWImax * NDVI * U * STEs	0,1297	69,1	1561
FWImax * EVI * U * STEs	0,1295	69,5	1560
DC * DMC * FFMCmax * NIR * STEs	0,1288	71,4	1567
DC * DMC * FFMCmax * NIR * MIR * U	0,1278	73,9	1548
FWImax * NIR * STEs	0,1269	76,4	1526
DC * DMC * FFMCmax * NIR * MIR * STEs	0,1266	77,2	1556
DC * DMC * FFMCmax * U * STEs	0,1261	78,6	1553
DC * DMC * FFMCmax * Tmax * U * STEs	0,1257	79,5	1558
FWImax * NIR * U * STEs	0,1252	81,0	1525
FWImax * NIR * MIR * STEs	0,1243	83,3	1516
DC * DMC * FFMCmax * NDVI * U * STEs	0,1240	84,1	1538
DC * DMC * FFMCmax * EVI * U * STEs	0,1239	84,4	1540
DC * DMC * FFMCmax * Tmax * NDVI * U * STEs	0,1236	85,0	1543
DC * DMC * FFMCmax * Tmax * EVI * U * STEs	0,1235	85,4	1545
FWImax * NIR * MIR * U * STEs	0,1229	87,0	1515
DC * DMC * FFMCmax * NIR * U * STEs	0,1201	94,5	1506
DC * DMC * FFMCmax * NIR * MIR * U * STEs	0,1180	100,0	1495

Supplementary materials from articles:

1. Article: Disentangling the factors of spatio-temporal patterns of wildfire activity in South-eastern France 219

2. Article: A comparative analysis of fire-weather indices for enhanced fire occurrence prediction with probabilistic approaches 230

3. Article: Fire danger prediction enhanced through optical remote sensing in Mediterranean France 241

1. Article: Disentangling the factors of spatio-temporal patterns of wildfire activity in South-eastern France

Title: Disentangling the factors of spatio-temporal patterns of wildfire activity in South-eastern France

Running head: patterns of regional wildfire activity

SUPPLEMENTARY MATERIALS

Authors: Jorge Castel-Clavera[1*], François Pimont[1], Thomas Opitz[2], Julien Ruffault[1] and Jean-Luc Dupuy[1]

[1]URFM, INRAE, Avignon, France

[2]BioSP, INRAE, Avignon, France

[*]Corresponding author: jorge.castel-clavera@inrae.fr

Keywords: Firelihood, Forest fires, Risk management, Modelling, Bayesian, Mediterranean France, Spatial changes, temporal changes

Supplementary material

S1. Description of the study area

The Mediterranean French area has a wide variety of climatic regions (see also Fig. A1 for the SylvoEcoRegions), in Fig. S1.1 we show some examples (for the year 2015) geographically scattered in the area. Overall, precipitation is more common during the colder months, and reduced during the warmer months. Maximal temperatures vary approximately from +40ºC to +25ºC, and minimal temperatures reach locally to -20ºC.

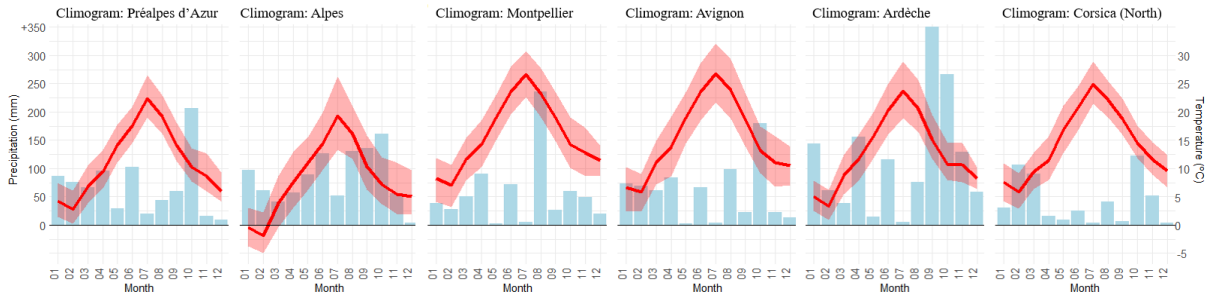


Fig. S1.1: Climograms for the year 2015 corresponding to six SAFRAN pixels located throughout the study area. In order of appearance: Préalpes d'Azur (pixel 8528), Alpes (pixel 7653), Montpellier (pixel 8858), Avignon (pixel 8507), Ardèche (pixel 7794), Corsica (pixel 9638). Precipitation represented in blue columns. Temperature represented with a red line and the maximal-minimal monthly temperature interval.

The spatial distribution of wildland area (Fig. S1.2A) and population (Fig. S1.2B) follow inversed patterns, with few settlements merged between the natural lands. Instead, population gets concentrated near the coast and allong the Rhone river valley.

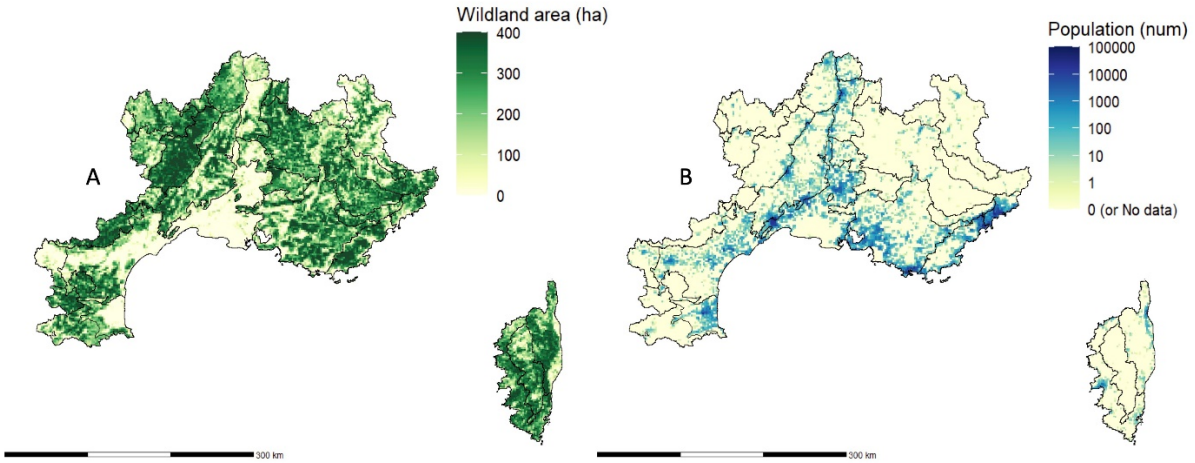


Fig. S1.2: Spatial distribution of Wildland area (A) and inhabitants per pixel (B) inside the study area. Represented in the 2km DFCI grid.

S2. Overview of Firelihood FL1 specifications and estimation (Pimont et al. 2021)

Firelihood is a Bayesian probabilistic framework described in details in Pimont et al. (2021). The model belongs to the class of Generalized Additive Mixed Models (GAMMs). The occurrence component is a Poisson model which predicts the number of fires (N_i) per SAFRAN's pixel and per days:

$$\log N_i \sim \log(WA) + f_{FWI}(FWI_i) + f_{WEEK}(WEEK_i) + \beta(Year > 2003) + f_{X,Y}(X_i, Y_i) \quad \text{(eq.S1)}$$

The explanatory variables were the wildland area (WA) and the FWI. The model had an offset $\log(WA_i)$, a temporal component based on the weekly correction, a post 2003 effect and a spatial component based on the geographical pixel coordinates.

The nonlinear f-functions in the equations are modelled with piecewise constant first-order random walks (rw1) with a fixed number of classes (the above-mentioned intervals) corresponding to a transformation of the continuous variables into discrete classes (categorical variables). One hyperparameter (called *precision*) governs curve smoothness (i.e., the size of the small steps between consecutive segments). For the two-dimensional predictors (X,Y) of the spatial effect, FL1 uses the SPDE approach to provide a numerically convenient representation of the spatial Matérn covariance function. It is applied to a triangulation mesh of the 2D space and specifies a prior distribution corresponding to a Gaussian random field over 2D space, and a the SPDE approach is used. Two hyperparameters are estimated for this random field: precision (to control the spatial variability of field values) and range (to control spatial dependence, *i.e.* the smoothness of the piecewise linear spatial surface).

Estimation is carried out with the integrated nested Laplace approximation (INLA), which is appropriate for regression models with Gaussian process priors for random effects as in our setting. It uses fast and accurate deterministic approximations of the components of the "posterior model", which is estimated by combining the prior model with the observed data thanks to the famous Bayes theorem. INLA output allows simulation from the posterior model, where the stochastic nature of the posterior model is preserved and allows for accurate uncertainty assessments in the applications.

For each simulated fire, its size is predicted from a complex piecewise modelling of the size distribution which built up on three exceedance thresholds (10, 100 abs 1000ha) and Pareto and Generalized Pareto distributions. Three exceedance thresholds are modelled as a probability p_i^u of a given predicted fire i to exceed a certain size u:

$$log\frac{p_i^u}{1-p_i^u} = \beta_0^{p,u} + f_{WA}^{p,u} + f_{FWI}^{p,u} \quad \text{(eq.S2)}$$

The modelling of piecewise Pareto and Generalized Pareto distributions is not reported here, as these components of the *Firelihood* -which allows to simulate burnt areas- have not been modified in the present work, but all details are provided in Pimont et al. (2021). These models were estimated with the same modelling approach as described above for the occurrence component.

S3. Partition of effects involved in spatial patterns

The estimation of relative contribution of FWI, LULC and unexplained spatio-temporal effect to spatial distributions of fire activities can be obtained by partitioning the spatial variance of simulations of the log number of fires, thanks to the additivity of the Poisson process.

If we sum the expected fires of a given year *y* in pixel *i*, we can write:

$$logN_{i,y}^{1ha} = log\left(\sum_{d\ in\ year\ y} N_{i,d}^{1ha}\right) \quad \text{(eq.S3)}$$

Since LULC and spatio-temporal terms are constant for pixel *i* and year *y*, we can factorize them:

$$logN_{i,y}^{1ha} = \beta_0 + log\left(\underbrace{\sum_{d\ in\ year\ y} e^{f_{FWI}(FWI_{i,d})+f_{WEEK}(WEEK_{i,d})}}_{\lambda_{i,y}^{FWI}}\right) + log\left(\underbrace{e^{\beta(WAp_{i,y})+\sum_{LULCM} f_{LULCM}(LULCM_{i,y})}}_{\lambda_{i,y}^{LULCM}}\right) + log\left(\underbrace{e^{f_{X,Y}(X_i,Y_i)+f_{YEAR}(y)+f'_{X,Y}(X_i,Y_i)(y-1992)}}_{\lambda_{i,y}^{Spatio-temporal}}\right) \quad \text{(eq.S4}$$

So that, $logN_{i,y}^{1ha} \propto log\left(\lambda_{i,y}^{FWI}\right) + log\left(\lambda_{i,y}^{LULCM}\right) + log\left(\lambda_{i,y}^{Spatio-temporal}\right)$

We then propose to estimate the relative contribution (RC) percentage of each component as:

$$RC_{FWI}^{1ha} = 100\ \frac{\overline{var_i\left(log\left(\lambda_{i,y}^{FWI}\right)\right)}^{y}}{\Delta} \quad \text{(eq.S5)}$$

$$RC_{LULCM}^{1ha} = 100\ \frac{\overline{var_i\left(log\left(\lambda_{i,y}^{LULCM}\right)\right)}^{y}}{\Delta}$$

$$RC_{Spatio-temporal}^{1ha} = 100\ \frac{\overline{var_i\left(log\left(\lambda_{i,y}^{Spatio-temporal}\right)\right)}^{y}}{\Delta}$$

Where var_i is the variance over all pixels i and the bar denotes the average over years and $\Delta = \overline{var_i\left(log\left(\lambda_{i,y}^{FWI}\right)\right)}^{y} + \overline{var_i\left(log\left(\lambda_{i,y}^{LULCM}\right)\right)}^{y} + \overline{var_i\left(log\left(\lambda_{i,y}^{Spatio-temporal}\right)\right)}^{y}$

For 10 and 100 ha, we applied a similar methodology (for 100 ha):

$$logN_{i,y}^{100ha} = log\left(\sum_{d\ in\ year\ y} N_{i,d}^{1ha}\, p_{i,d}^{u}\right) \quad \text{(eq.S6)}$$

Assuming that $p_{i,d}^{u}$ is small, we can write:

$$p_{i,d}^{u} \approx e^{\beta_0^{p,u} + f_{FWI}^{p,u}(FWI_{i,d}) + \sum_{LULCMs} f_{LULCM}^{p,u}(LULCM_{i,y}) + f_{YEAR}^{p,u}(y) + f_{BESAG}^{p,u}(SER)}$$

Grouping together terms varying on the daily basis (FWI terms for occurrence and exceedance and seasonal corrections):

$$logN_{i,y}^{100ha} = log\left(\sum_{d\ in\ year\ y} N_{i,d}^{1ha}\, p_{i,d}^{u}\right) \qquad \text{(eq.S7)}$$

$$= \beta_0 + \beta_0^{p,u} + log\left(\underbrace{\sum_{d\ in\ year\ y} e^{f_{FWI}(FWI_{i,d}) + f_{WEEK}(WEEK_{i,d}) + f_{FWI}^{p,u}(FWI_{i,d})}}_{\lambda_{i,y}^{FWI}}\right)$$

$$+ log\left(\underbrace{e^{\beta(WAp_{i,y}) + \sum_{LULCM} f_{LULCM}(LULCM_{i,y}) + f_{LULCM}^{p,u}(LULCM_{i,y})}}_{\lambda_{i,y}^{LULCM}}\right)$$

$$+ log\left(\underbrace{e^{f_{X,Y}(X_i,Y_i) + f_{YEAR}(y) + f'_{X,Y}(X_i,Y_i)(y-1992) + f_{YEAR}^{p,u}(y) + f_{BESAG}^{p,u}(SER)}}_{\lambda_{i,y}^{Spatio-temporal}}\right)$$

So that the same metrics for relative contribution can be built for $logN_{i,y}^{100ha}$ and $logN_{i,y}^{10ha}$ than for $logN_{i,y}^{1ha}$.

S4. Distribution and changes of Land-Use and Land-Cover by SylvoEcoRegion in the Promethee area.

The French Mediterranean landscape has been dynamic over the study period. Here we illustrate the observed changes in terms of Land-Use and Land-Cover (LULC) from 1993 to 2018, represented as the mean area per pixel within each SylcoEcoRegion (SER) in figures S4.1 and S4.2.

The largest changes do not exceed the 45ha in average per pixel of 400ha, although those largest magnitudes are only seen in single SERs. Shrubland area (Fig. S4.1 A – B) and agricultural area (Fig. S4.2 K – L) have suffered the most extended loses; while urban (Fig. S4.1 E – F), mixed forest (Fig. S4.2 G – H) and broadleaved forest (Fig. S4.2 I – J) areas have the largest extent of positive change.

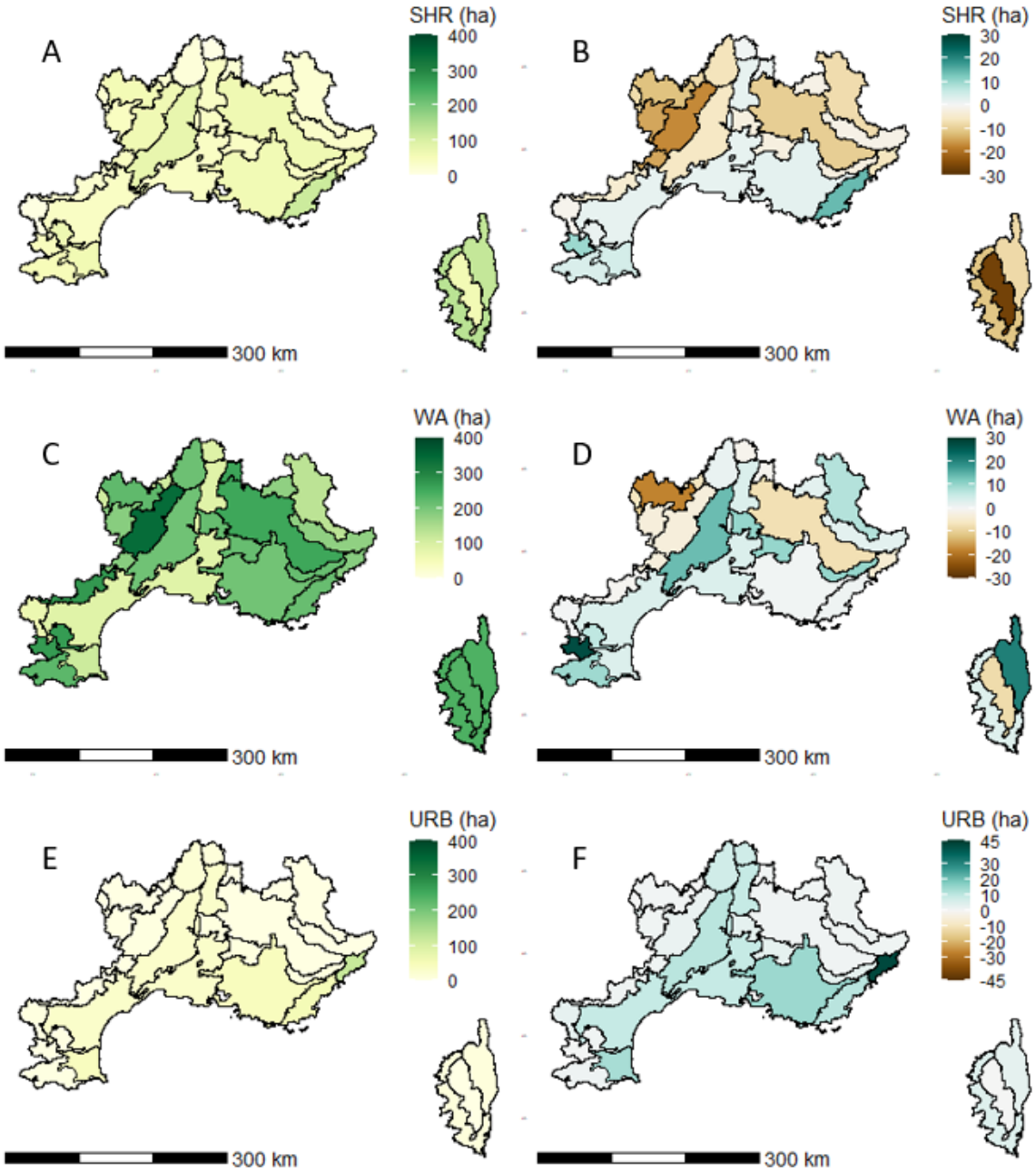


Fig. S4.1: Distribution of mean area per pixel of LULC variables by SERs in 1993 (left column) and their total absolute change in mean area per pixel for the period 1993-2018. Where: SHR = shrubland area; WA = wildland area: URB = urban area.

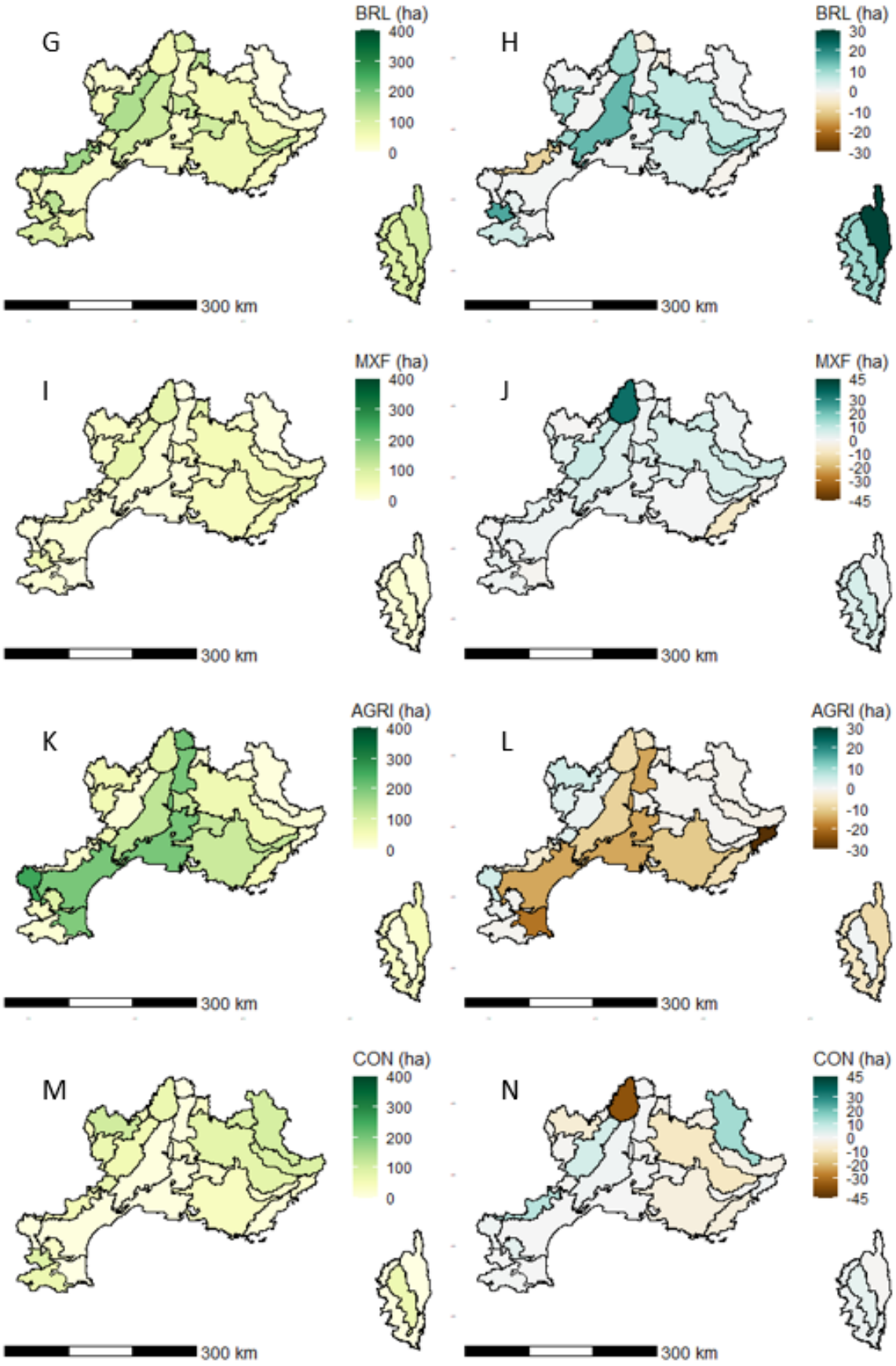
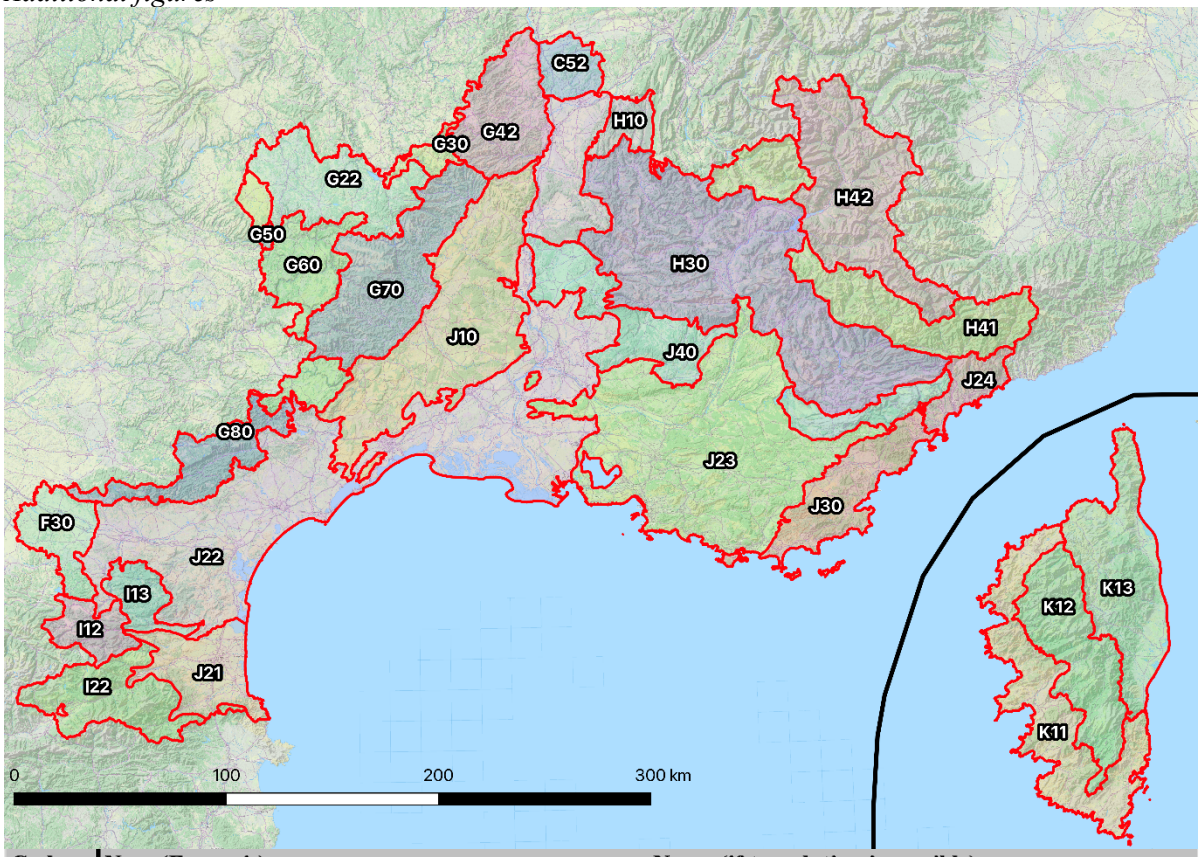


Fig. S4.2: Continuation: Fig S4.1. Where BRL = broadleaved forest area; MXF = mixed forest area; AGRI = agricultural area; CON = coniferous forest area.

Additional figures



Code	Nom (Français)	Name (if translation is possible)
K11	Corse occidentale	West Corsica
K12	Montagne corse	Mountaiins of Corsica
K13	Corse orientale	East Corsica
G42	Monts du Vivarais et du Pilat	-
F30	Coteaux de la Garonne	Hillsides of Garonne river
H41	Alpes intermédiaires du Sud	Southern Mid-Alps
H42	Alpes internes du Sud	Southern Inner-Alps
I13	Corbières	-
G50	Ségala et Châtaigneraie auvergnate	-
I12	Pyrénées cathares	Cathare Pyrenees
J23	Provence calcaire	Sandstone Provence
J10	Garrigues	-
J24	Secteurs niçois et préligure	Nice and preligure sectors
G60	Grands Causses	-
I21	Haute-chaîne pyrénéenne	Upper pyrenean chain
I22	Pyrénées catalanes	Catalan Pyrenees
J21	Roussillon	-
G70	Cévennes	-
H10	Préalpes du Nord	Northern Pre-Alps
J22	Plaines et collines rhodaniennes et languedociennes	Rhone and Languedoc plains and hills
J30	Maures et Esterel	-
G22	Plateaux granitiques du centre du Massif central	Granite plateaus in the center of the Massif Central
G80	Haut-Languedoc et Lévézou	Upper Languedoc and Lévézou
H22	Alpes internes du Nord	Northern Inner-Alps
C52	Plaines et piémonts alpins	Alpine plains and foothills
G30	Massif central volcanique	Volcanic Massif Central
J40	Préalpes du Sud	Southern Pre-Alpes
H30	Alpes externes du Sud	Southern Outer-Alpes

Fig. A1. Map of the 28 Sylvo-Eco-Regions (SER) of the study area with the key code and name. The perimeter of the shape corresponds to the border Promethee zone.

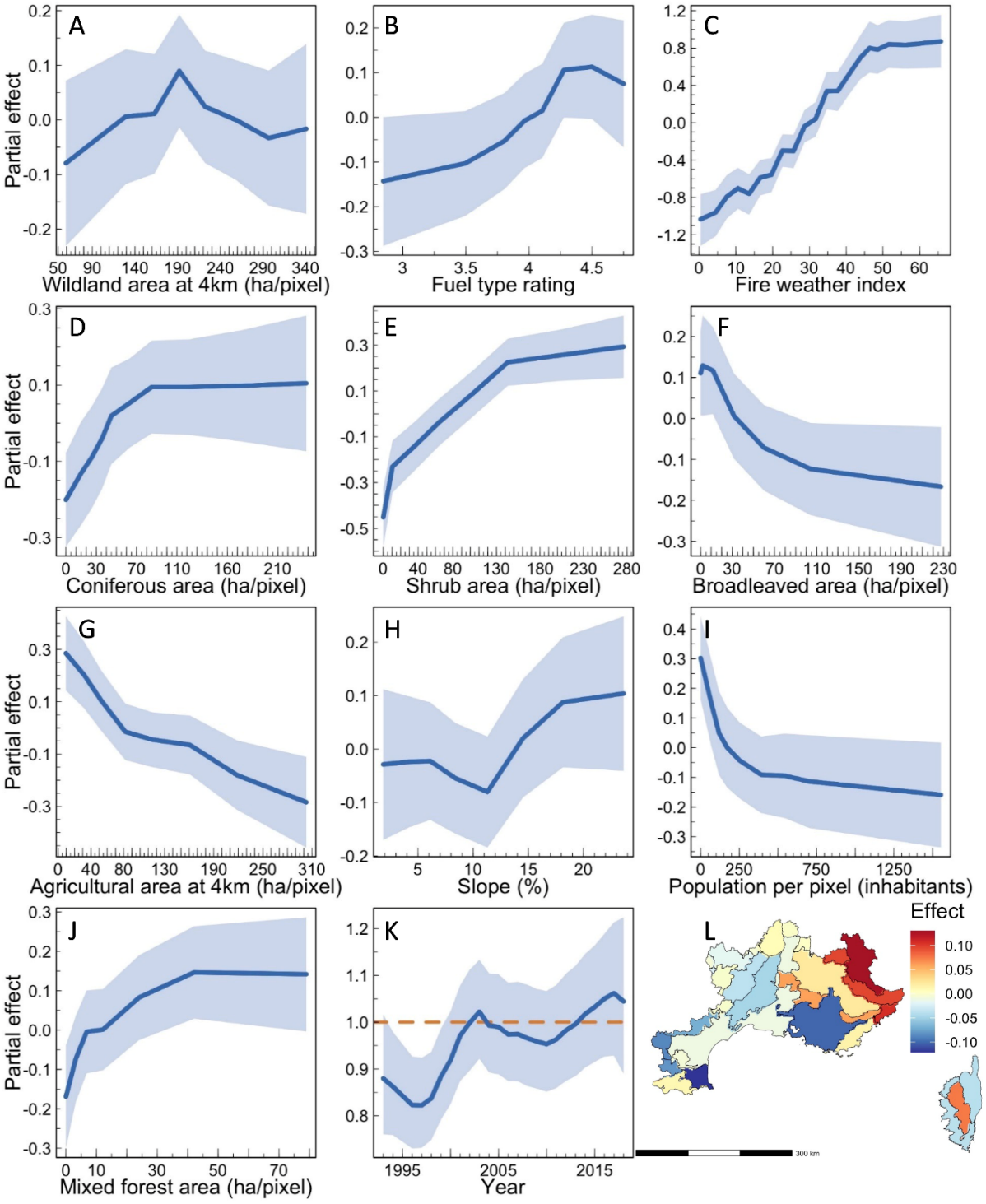


Fig. A2. Partial effects of the predictor variables for the 10ha exceedance probability model intervals; the X axis shows the value in each pixel but only inside the 95% of the data distribution, cutting the extreme values from below and above; the Y axis, Partial effect, is the contribution of fire the predictor variable's value to the probability of exceedance. And the spatial effect of the SylvoEcoRegions. Spatial effect from the Besag component in the 10ha exceedance model. The magnitude effect is drawn in a color coding in each pixel, the labels inside each polygon (thin black lines) corresponds to the code of the SylvoEcoRegions. The magnitude of the effect will be positive when the model underpredicts the exceedance probability, and negative when the model overpredicts this probability.

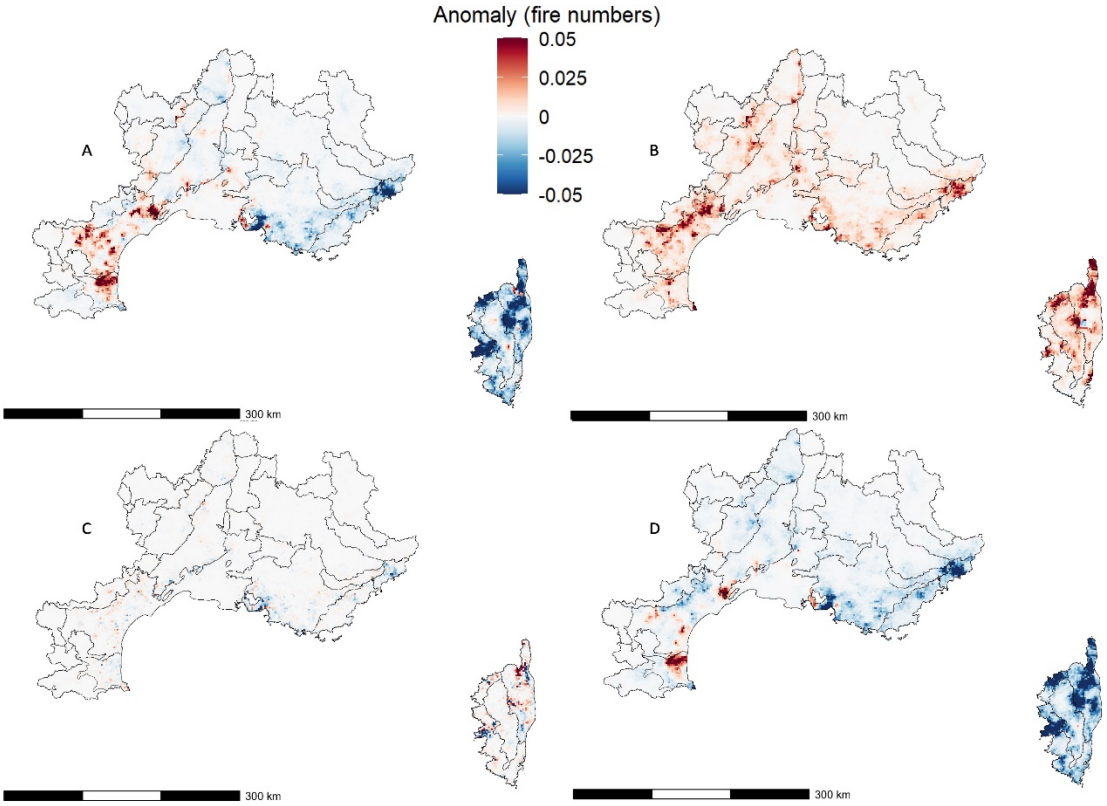


Fig. A3. Spatial distribution of the changes (anomalies) between the first and second decades of the comparison. For illustration porpoises, here represented the scenarios for 10ha fires.

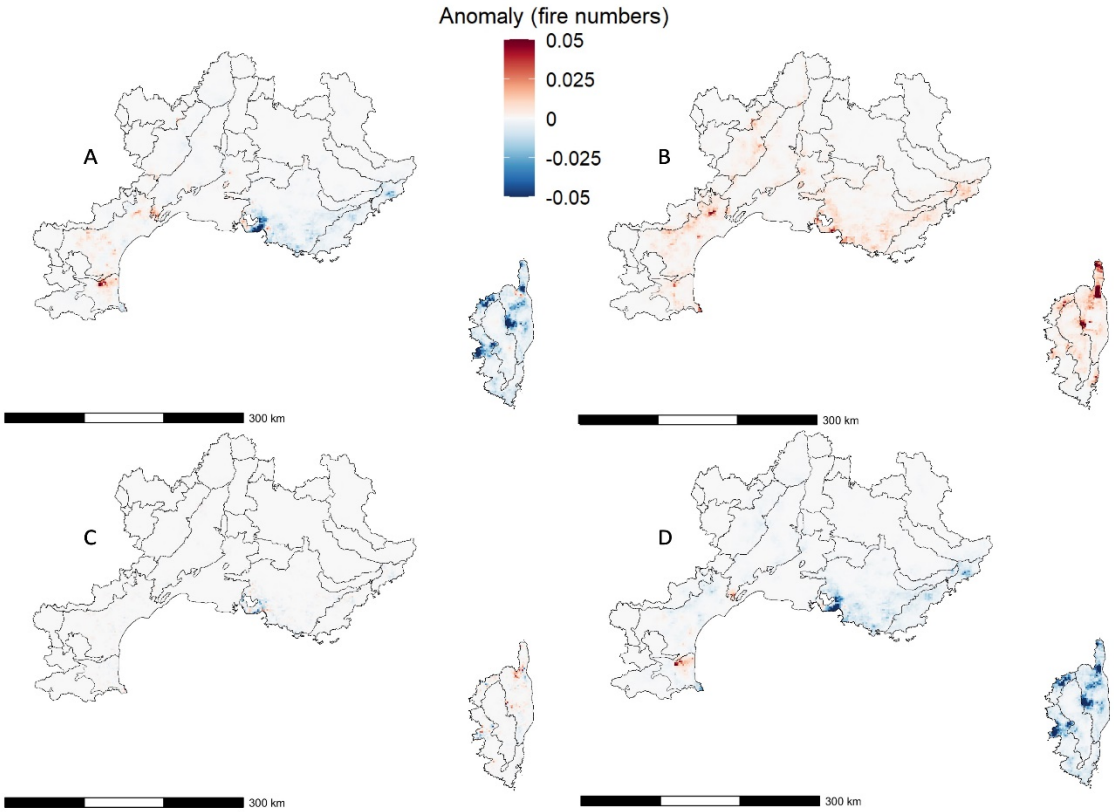


Fig. A4. Spatial distribution of the changes (anomalies) between the first and second decades of the comparison. For illustration porpoises, here represented the scenarios for 100ha fires.

2. Article: A comparative analysis of fire-weather indices for enhanced fire occurrence prediction with probabilistic approaches

Jorge Castel-Clavera*[a], François Pimont[a], Thomas Opitz[b], Julien Ruffault[a], Renaud Barbero[c], Denis Allard[b], Jean-Luc Dupuy[a]

*Corresponding author: jorge.castel-clavera@inrae.fr

[a] URFM, INRAE, Avignon, France

[b] BioSP, INRAE, Avignon, France

[c] RECOVER, INRAE, Aix-en-Provence, France

A. Reliability plots for model tests

Reliability plots are a tool to compare the predicted probabilities against the observed probability for discrete intervals of the predicted probability. From this kind of graphic, we can deduce two main traits: (i) Accuracy if the predictions and observations have a similar distribution, giving a line close to the 45º slope. (ii) Specificity if the model tends to predict values at the extremes of the probabilities (values close to 0 or 1, avoiding the middle terms). That would mean that the model is not ambiguous.

The examples in Fig.A.1 show the reliability plots for 2 models which even though do not have the same performance in terms of $Skill_{Log-loss}$ they have a very close behavior in accuracy and specificity. The frequencies (predicted and observed) do not reach high values because of the rarity of the fire events in space and time. Both models represented have good accuracy although the higher frequencies of fire numbers – both 1ha fires (NB1) and 80ha fires (NB80) – are underestimated (the observed frequency is higher than the predicted) – this was a common finding for all models – because the models do not completely predict the extreme fire events. The 80ha thresholds exceedance probability (P80) is overestimated for the lower probabilities

and underestimated for the higher probabilities. Regarding the specificity, the quantiles of predicted frequency are mostly at the lower values of the spectrum, with one category representing the extreme values, meaning that the model does not give ambiguous predictions.

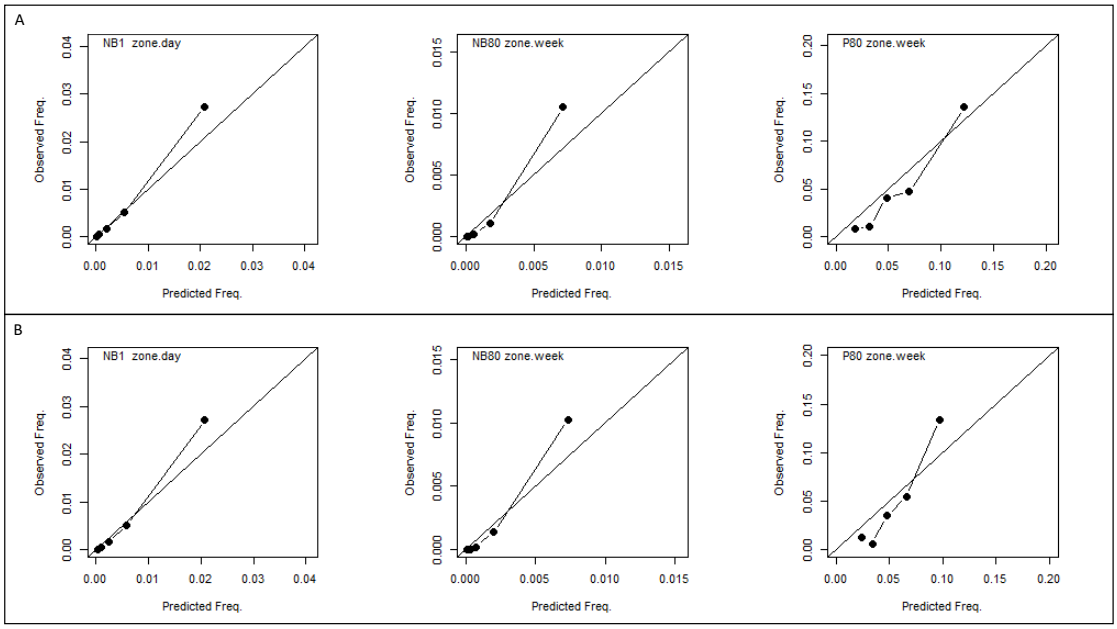


*Fig.(A.1): Reliability plots for two models (CanFWI(12h) – Box A –, and DC*DMC*FFMC*VPD*U – Box B –) with STEs. The predicted frequencies are compared to the observed frequencies of fire numbers 1ha fires (NB1) and 80ha fires (NB80) and the probability of a 1ha-fire becoming a >80ha-fire.*

B. Log Loss with formulas

LogLoss is a measure of the distances between predictions and observations for binary statistical models. Where observations are either presence or absence (0 or 1) and the predictions are continuous probabilities between 0 and 1 (0% and 100% probability of presence), the Log Loss is the difference between the predicted probability (p) and the observation (y) as expressed in the Equation B.1:

$$LogLoss_i = -\{[y_i \times \ln(p_i)] + [(1 - y_i) \times \ln(1 - p_i)]\} \qquad Eq.(B.1)$$

The score that corresponds to a model is the mean of all observation-prediction pairs (Equation B.2):

$$LogLoss = -\frac{1}{N}\sum_{i=1}^{N}\{[y_i \times \ln(p_i)] + [(1 - y_i) \times \ln(1 - p_i)]\} \qquad Eq.(B.2)$$

The LogLoss score can be adapted to probabilistic frameworks with a Poisson distribution for the occurrence component evaluation (Equation B.3):

$$LogLoss = \frac{1}{N}\sum_{i=1}^{N} - \log\big(P(y_i|\lambda_i)\big) \qquad Eq.(B.3)$$

Where N is the number of observations, y_i is the observed value for the observation *i*, λ_i is the predicted value corresponding to the observation *i*. Therefore, $P(y_i|\lambda_i)$ is the probability mass function of observing y_i given λ_i in a Poisson distribution.

C. Full table of index evaluation

Table (C.1): Model evaluation (at ZFC/day scale) metrics for the 1ha occurrence component. Names in the table are as follows: LogLoss is the raw output from the score formula, Skill Score is the Log Loss skill score based on the FAb model, and Skill Scale Value is the normalization of the Skill Score values between the 0 and the maximal performance value, AUC is the value from an AUC-ROC based on the model fit. In the models, max indicates the maximum daily value, daily indicates the version from Bedia et al. (2013), min indicates the minimum daily value, and no indication corresponds to the mean daily value. The yellow rows correspond to the FAb. The table is sorted using the occurrence's "Skill Scale Values".

Occurrence (1ha)					
Model	**LogLoss**	**Skill Log-Loss**	**Skill Scale Value**	**AUC**	**DIC**
NULL	0,0455	-0,1841	-1,91	0,57	82530
WA	0,0439	-0,1419	-1,47	0,60	75987
Precipitation	0,0428	-0,1147	-1,19	0,66	74975
U	0,0426	-0,1097	-1,14	0,67	74723
T (max)	0,0423	-0,1007	-1,04	0,69	74268
Modified Fosberg (max)	0,0422	-0,0990	-1,03	0,70	72846
T (mean)	0,0419	-0,0913	-0,95	0,71	73895
RH (min)	0,0418	-0,0892	-0,92	0,71	73980
Chandler (max)	0,0417	-0,0869	-0,90	0,72	73895
KBDI (12h)	0,0417	-0,0869	-0,90	0,72	73898
KBDI (max)	0,0417	-0,0865	-0,90	0,72	73888
Chandler (12h)	0,0417	-0,0846	-0,88	0,72	73809
Sharples	0,0414	-0,0785	-0,81	0,72	73612
VPD	0,0413	-0,0764	-0,79	0,73	73416
FFMC (max)	0,0411	-0,0699	-0,72	0,74	73281
Fosberg (12h)	0,0411	-0,0692	-0,72	0,73	73354
FFMC (12h)	0,0411	-0,0692	-0,72	0,74	73282
Fosberg (max)	0,0410	-0,0681	-0,71	0,73	73359
FFMC (ref)	0,0408	-0,0631	-0,65	0,75	73068
Modified Fosberg (12h)	0,0406	-0,0559	-0,58	0,75	72924
MARK5 (max)	0,0405	-0,0547	-0,57	0,76	72696
DC (max)	0,0405	-0,0538	-0,56	0,76	72816
DC (12h)	0,0405	-0,0538	-0,56	0,76	72814
MARK5 (12h)	0,0404	-0,0531	-0,55	0,76	72655
DC (ref)	0,0404	-0,0524	-0,54	0,76	72769
DMC (max)	0,0403	-0,0493	-0,51	0,77	72529
DMC (12h)	0,0403	-0,0491	-0,51	0,77	72517
HotDryWindy (12h)	0,0402	-0,0474	-0,49	0,76	72495
VPD * U	0,0401	-0,0451	-0,47	0,76	72235
DMC (ref)	0,0401	-0,0446	-0,46	0,78	72337
HotDryWindy (max)	0,0401	-0,0438	-0,45	0,77	72320
Modified Sharples	0,0401	-0,0434	-0,45	0,77	72368
VPD * U * T (max)	0,0400	-0,0425	-0,44	0,76	72116
CanFWI (max)	0,0393	-0,0224	-0,23	0,80	71597
DC * DMC * VPD	0,0393	-0,0221	-0,23	0,80	71557
CanFWI (12h) * VPD	0,0390	-0,0165	-0,17	0,80	71367
CanFWI (12h) * T(max)	0,0390	-0,0159	-0,17	0,80	71381

CanFWI (12h) * HR (min)	0,0390	-0,0158	-0,16	0,80	71378
CanFWI (12h)	0,0390	-0,0157	-0,16	0,80	71383
DC * DMC * FFMC	0,0390	-0,0143	-0,15	0,81	71320
CanFWI (ref)	0,0389	-0,0134	-0,14	0,80	71302
CanFWI (12h) * U	0,0389	-0,0126	-0,13	0,80	71248
DC * DMC * FFMC * T (max)	0,0389	-0,0125	-0,13	0,81	71285
CanFWI (12h) * VPD * U	0,0389	-0,0120	-0,12	0,80	71155
DC * DMC * FFMC * VPD	0,0389	-0,0117	-0,12	0,81	71190
DC * DMC * FFMC * HRmin	0,0388	-0,0100	-0,10	0,81	71178
CanFWI (ref) * U	0,0388	-0,0098	-0,10	0,81	71167
FAb	0,0384	0,0000	0,00	0,82	71136
VPD * U * DC	0,0384	0,0008	0,01	0,81	70701
DC * DMC * VPD * U	0,0382	0,0053	0,05	0,82	70543
WA * STEs	0,0381	0,0075	0,08	0,82	69154
DC * DMC * FFMC * T (max) * U	0,0380	0,0098	0,101	0,82	70429
DC * DMC * FFMC * U	0,0380	0,0098	0,10	0,82	70430
T (mean) * STEs	0,0380	0,0105	0,11	0,82	69020
T (max) * STEs	0,0380	0,0116	0,12	0,83	68996
DC * DMC * FFMC * VPD * U	0,0379	0,0131	0,14	0,83	70284
Modified Fosberg (max) * STEs	0,0377	0,0179	0,19	0,83	67584
Precipitation * STEs	0,0376	0,0210	0,22	0,83	68661
KBDI (max) * STEs	0,0374	0,0250	0,26	0,84	68398
KBDI (12h) * STEs	0,0374	0,0251	0,26	0,84	68425
U * STEs	0,0373	0,0286	0,30	0,83	68351
VPD * STEs	0,0368	0,0407	0,42	0,85	67984
DC (ref) * STEs	0,0367	0,0441	0,46	0,85	67878
DMC (max) * STEs	0,0367	0,0449	0,46	0,85	67802
DC (12h) * STEs	0,0367	0,0450	0,47	0,86	67838
DC (max) * STEs	0,0367	0,0451	0,47	0,85	67837
DMC (ref) * STEs	0,0367	0,0452	0,47	0,86	67757
Sharples * STEs	0,0367	0,0452	0,47	0,84	67738
DMC (12h) * STEs	0,0367	0,0453	0,47	0,86	67792
Modified Fosberg (12h) * STEs	0,0366	0,0461	0,48	0,85	67614
Fosberg (max) * STEs	0,0365	0,0494	0,51	0,85	67512
Fosberg (12h) * STEs	0,0365	0,0505	0,52	0,85	67536
Modified Sharples * STEs	0,0364	0,0514	0,53	0,85	67355
FFMC (max) * STEs	0,0363	0,0556	0,58	0,86	67535
FFMC (12h) * STEs	0,0363	0,0557	0,58	0,86	67540
FFMC (ref) * STEs	0,0362	0,0583	0,60	0,86	67412
Chandler (max) * STEs	0,0360	0,0621	0,64	0,86	67158
HotDryWindy (max) * STEs	0,0360	0,0635	0,66	0,86	66968
Chandler (12h) * STEs	0,0360	0,0635	0,66	0,86	67135
HotDryWindy (12h) * STEs	0,0360	0,0636	0,66	0,86	67115
RH (min) * STEs	0,0359	0,0640	0,66	0,86	67093
DC * DMC * VPD * STEs	0,0359	0,0657	0,68	0,87	66945
VPD * U * T (max) * STEs	0,0359	0,0660	0,68	0,86	66918

VPD * U * STEs	0,0358	0,0690	0,71	0,86	66847
MARK5 (max) * STEs	0,0357	0,0703	0,73	0,87	66886
MARK5 (12h) * STEs	0,0356	0,0727	0,75	0,87	66831
DC * DMC * FFMC * STEs	0,0356	0,0742	0,77	0,87	66703
DC * DMC * FFMC * VPD * STEs	0,0355	0,0756	0,78	0,87	66608
DC * DMC * FFMC * T (max) * STEs	0,0354	0,0782	0,81	0,87	66593
CanFWI(max) * STEs	0,0354	0,0789	0,82	0,87	66650
CanFWI(ref) * STEs	0,0351	0,0857	0,89	0,87	66436
CanFWI (12h) * T (max) * STEs	0,0351	0,0858	0,89	0,87	66225
CanFWI (12h) * STEs	0,0351	0,0859	0,89	0,87	66225
VPD * U * DC * STEs	0,0351	0,0863	0,89	0,87	66208
DC * DMC * FFMC * HR (min) * STEs	0,0351	0,0863	0,89	0,88	66148
CanFWI(12h) * U * STEs	0,0351	0,0864	0,89	0,87	66120
CanFWI (ref) * U * STEs	0,0351	0,0868	0,90	0,87	66283
CanFWI (12h) * VPD * STEs	0,0351	0,0869	0,90	0,88	66108
CanFWI (12h) * VPD * U * STEs	0,0350	0,0886	0,92	0,88	65920
CanFWI (12h) * HR (min) * STEs	0,0349	0,0910	0,94	0,88	65871
DC * DMC * VPD * U * STEs	0,0348	0,0927	0,96	0,88	65783
DC * DMC * FFMC * U * STEs	0,0348	0,0938	0,97	0,88	65824
DC * DMC * FFMC * T (max) * U * STEs	0,0348	0,0942	0,98	0,88	65816
DC * DMC * FFMC * VPD * U * STEs	0,0347	0,0965	1,00	0,88	65644

Table (C.2): Model evaluation for the 80ha size component. Where: LogLoss is the raw output from the score formula, Skill Score is based on the FAb model, and Skill Scale Value is the normalization of the Skill Score values between the 0 and the maximal performance value set to 100, AUC is the value from an AUC-ROC based on the model fit. In the tests max indicates the maximal daily value, daily indicates the version from Bedia et al. (2013), min indicates the minimal daily value, and no indication corresponds to the mean daily value. The yellow rows correspond to the FAb. The table is sorted using the SIZE's "Skill Scale Value" column.

80ha threshold					
Model	**LogLoss**	**Skill Log-Loss**	**Skill Scale Value**	**AUC**	**DIC**
NULL	0,203	-0,292	-2,013	1,000	1428
WA	0,191	-0,216	-1,493	0,604	1419
T(max)	0,188	-0,196	-1,353	0,598	1420
T (mean)	0,183	-0,166	-1,146	0,617	1415
DC (ref)	0,181	-0,154	-1,060	0,632	1409
DC (max)	0,180	-0,145	-1,001	0,636	1407
DC (12h)	0,180	-0,145	-0,999	0,636	1407
KNDI (max)	0,179	-0,141	-0,973	0,631	1406
KBDI (12h)	0,179	-0,140	-0,969	0,631	1406
Precipitation	0,178	-0,133	-0,914	0,634	1411
DMC (12h)	0,174	-0,108	-0,744	0,665	1391
DMC (max)	0,174	-0,107	-0,740	0,666	1391
VPD	0,174	-0,105	-0,725	0,668	1390
DMC (ref)	0,173	-0,102	-0,704	0,669	1389
FFMC (max)	0,169	-0,078	-0,539	0,694	1374
FFMC.12h	0,169	-0,078	-0,535	0,693	1374

FFMC (ref)	0,168	-0,068	-0,466	0,703	1368
RH (min)	0,167	-0,064	-0,443	0,709	1362
Chandler (12h)	0,166	-0,057	-0,392	0,715	1358
Chandler (max)	0,166	-0,057	-0,392	0,715	1358
U	0,165	-0,049	-0,341	0,715	1357
DC * DMC * VPD	0,163	-0,034	-0,237	0,689	1372
Modified Sharples	0,162	-0,032	-0,224	0,732	1344
Modified Fosberg (12h)	0,162	-0,032	-0,218	0,736	1342
Modified Fosberg (max)	0,162	-0,031	-0,214	0,736	1342
Fosberg (max)	0,161	-0,022	-0,155	0,740	1334
Fosberg (12h)	0,161	-0,022	-0,155	0,740	1334
MARK5 (12h)	0,160	-0,021	-0,142	0,748	1328
MARK5 (max)	0,160	-0,020	-0,140	0,748	1327
HotDryWindy (12h)	0,160	-0,018	-0,122	0,747	1329
HotDryWindy Max	0,159	-0,013	-0,091	0,751	1326
DC * DMC * FFMC	0,159	-0,012	-0,084	0,714	1355
Sharples	0,158	-0,006	-0,039	0,755	1327
DC * DMC * FFMC * VPD	0,157	-0,001	-0,007	0,710	1355
FAb	0,157	0,000	0,000	0,694	1383
CanFWI (12h)	0,157	0,003	0,024	0,765	1315
CanFWI (ref)	0,156	0,005	0,036	0,764	1317
DC * DMC * FFMC * T (max)	0,156	0,007	0,046	0,732	1348
CanFWI (max)	0,156	0,008	0,054	0,767	1312
CanFWI (12h) * T(max)	0,156	0,009	0,064	0,767	1315
VPD * U * T (max)	0,155	0,017	0,115	0,756	1330
WA * STEs	0,154	0,018	0,125	0,701	1377
VPD * U	0,154	0,018	0,125	0,749	1331
T (mean) * STEs	0,154	0,020	0,135	0,701	1378
DC * DMC * FFMC * HR (min)	0,153	0,024	0,163	0,729	1338
CanFWI (12h) * VPD	0,153	0,025	0,172	0,760	1312
KBDI (12h) * STEs	0,153	0,026	0,178	0,707	1374
KBDI (max) * STEs	0,153	0,026	0,181	0,707	1374
DMC (12h) * STEs	0,153	0,028	0,193	0,710	1369
DMC (max) * STEs	0,153	0,028	0,195	0,710	1369
DMC (ref) * STEs	0,152	0,030	0,209	0,711	1368
VPD * STEs	0,152	0,033	0,230	0,715	1365
T (max) * STEs	0,152	0,034	0,232	0,706	1375
DC (ref) * STEs	0,152	0,036	0,246	0,711	1368
Precipitation * STEs	0,151	0,037	0,253	0,710	1371
DC (max) * STEs	0,151	0,037	0,254	0,712	1367
DC (12h) * STEs	0,151	0,037	0,256	0,712	1367
VPD * U * DC	0,151	0,039	0,266	0,755	1325
CanFWI (12h) * U	0,150	0,045	0,310	0,762	1304
CanFWI (12h) * HR (min)	0,150	0,046	0,314	0,768	1298
CanFWI (ref) * U	0,150	0,047	0,327	0,763	1305
FFMC (12h) * STEs	0,150	0,048	0,329	0,732	1352

FFMC (max) * STEs	0,150	0,048	0,330	0,733	1352
DC * DMC * VPD * STEs	0,149	0,051	0,353	0,724	1355
FFMC (ref) * STEs	0,149	0,054	0,374	0,739	1347
DC * DMC * VPD * U	0,149	0,055	0,381	0,755	1320
RH (min) * STEs	0,147	0,065	0,445	0,746	1338
Chandler (12h) * STEs	0,147	0,066	0,454	0,748	1336
Chandler (max) * STEs	0,147	0,066	0,454	0,748	1336
CanFWI (12h) * VPD * U	0,147	0,066	0,458	0,765	1298
DC * DMC * FFMC * U	0,147	0,067	0,463	0,765	1307
DC * DMC * FFMC * STEs	0,146	0,068	0,471	0,744	1339
DC * DMC * FFMC * T (max) * U	0,146	0,070	0,480	0,770	1306
DC * DMC * FFMC * VPD * STEs	0,146	0,073	0,503	0,741	1339
DC * DMC * FFMC * VPD * U	0,145	0,075	0,515	0,764	1306
Modified Sharples * STEs	0,145	0,077	0,532	0,755	1327
Modified Fosberg (max) * STEs	0,144	0,085	0,588	0,759	1323
Modified Fosberg (12h) * STEs	0,144	0,085	0,589	0,760	1323
MARK5 (12h) * STEs	0,144	0,087	0,599	0,764	1313
MARK5 (max) * STEs	0,144	0,087	0,600	0,764	1313
DC * DMC * FFMC * T (max) * STEs	0,144	0,087	0,600	0,760	1328
DC * DMC * FFMC * HR (min) * STEs	0,143	0,092	0,634	0,756	1324
HotDryWindy (12h) * STEs	0,142	0,095	0,659	0,768	1312
HotDryWindy (max) * STEs	0,142	0,099	0,682	0,771	1309
U * STEs	0,141	0,101	0,700	0,762	1323
CanFWI (ref) * STEs	0,141	0,102	0,701	0,778	1302
CanFWI (12h) * STEs	0,141	0,104	0,714	0,778	1300
CanFWI (max) * STEs	0,140	0,107	0,736	0,781	1297
Fosberg (max) * STEs	0,140	0,108	0,748	0,774	1307
Fosberg (12h) * STEs	0,140	0,108	0,748	0,774	1307
CanFWI (12h) * VPD * STEs	0,140	0,110	0,762	0,777	1299
CanFWI (12h) * T (max) * STEs	0,140	0,111	0,766	0,783	1298
VPD (max) * U * STEs	0,140	0,112	0,773	0,775	1310
VPD * U * T (max) * STEs	0,139	0,116	0,799	0,781	1306
VPD * U * DC * STEs	0,139	0,117	0,807	0,779	1305
Sharples * STEs	0,138	0,122	0,841	0,785	1299
DC * DMC * VPD * U * STEs	0,138	0,124	0,856	0,779	1303
CanFWI (12h) * HR (min) * STEs	0,137	0,125	0,865	0,785	1286
CanFWI (12h) * U * STEs	0,136	0,137	0,946	0,789	1284
CanFWI (ref) * U * STEs	0,136	0,137	0,947	0,791	1285
DC * DMC * FFMC * U * STEs	0,135	0,139	0,958	0,790	1290
DC * DMC * FFMC * VPD * U * STEs	0,135	0,143	0,984	0,789	1289
CanFWI (12h) * VPD * U * STEs	0,134	0,145	0,999	0,791	1281
DC * DMC * FFMC * T (max) * U * STEs	0,134	0,145	1,000	0,795	1287

D. Model predictability

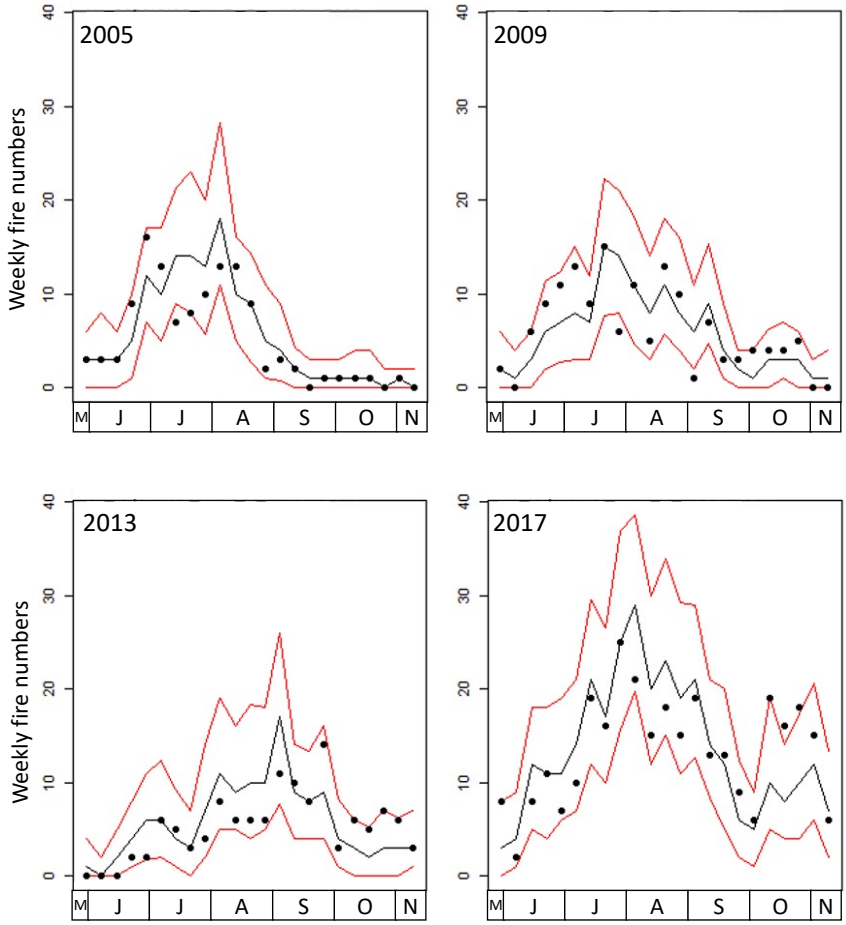


*Fig. (D.2): Weekly fire numbers for 4 years: simulated (lines; black = mean of 150 simulations, red = 95% quantiles of the 150 simulations) and observed (points). The model here represented is DC*DMC*FFMC*VPD*U*STEs.*

The inter annual variability in the seasonality of fire activity was well captured by the simulations of a model including both Weather variables and the STEs (Fig.B.2). In contrast, the FAb model would not reproduce the features of each particular year such as a late fire season in 2013 or an unusual high peak during the last 5 weeks of the season of 2017. As expected, the FAb model, which reproduces mean seasonal activity, only fitted well to moderate (mean) fire seasons such as in 2009.

E. F-beta curves for P80 and NB80

The F-beta curves presented in the main text of the paper are extended in this section. Figure D.3 presents, additionally to the CanFWI(12h) model, the best fitted model (DC*DMC*FFMC*VPD*U *STEs). We see how moth models have a similar trend in the response, but the best fitted model gives a higher score, and most importantly, it does not decrease as fast as the simpler model.

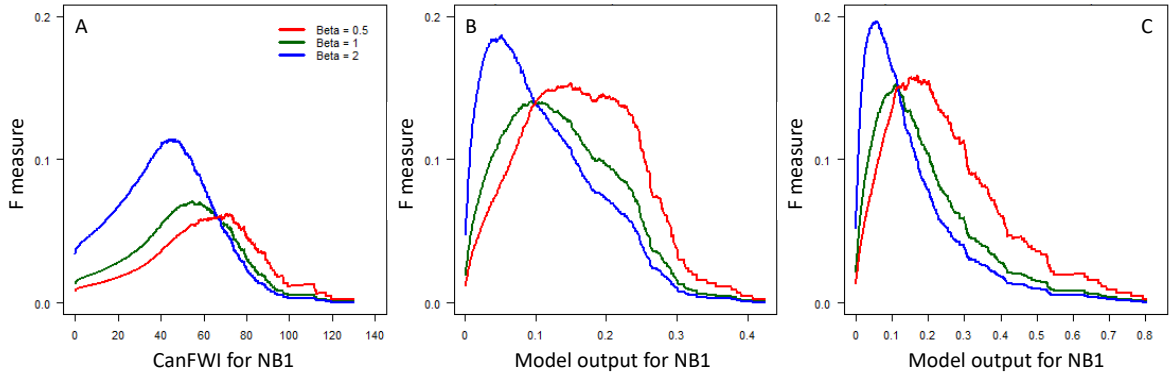


*Fig. (D.3): Fb measure for two occurrence components (NB1) models (CanFWI-12h and DC*DMC*FFMC*VPD*U) with STEs and the CanFWI with no modelling framework.*

Figure D.4 compares the CanFWI with no modelling framework (Fig.D.4A), a simple model (CanFWI*STEs; Fig.D.4B) and the best fitted model (DC*DMC*FFMC*VPD*U*STEs; Fig.D.4C) on their capability to predict the P80 (Probability of a given fire to exceed the 80ha size threshold). Although the differences are smaller than for the occurrence component, the models reach a higher F value for all ß values. The $F_ß$ values are highest for the DC*DMC*FFMC*VPD*U*STEs model, and as seen for the occurrence component, the differences between models are small. Combining the predictions of the occurrence component and the SC we obtain the expected numbers of >80ha fires (NB80). F_β values are higher for the modelling frameworks than for the CanFWI. The DC*DMC*FFMC*T*U*STEs model has a better F_β for all beta values than CanFWI(12h) *STEs. Therefore, the results between OC and SC are consistent in terms of predictability.

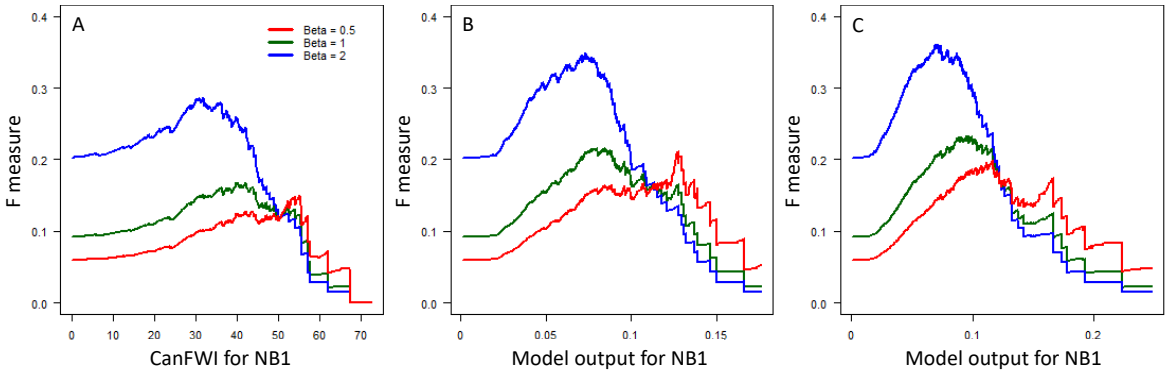


*Fig. (D.4): F_β measures for the 80ha threshold exceedance probability (P80) as predicted through the CanFWI with no modelling framework and two models (i) CanFWI(12h) + STEs and (ii) DC*DMC*FFMC*Tmax*U + STEs.*

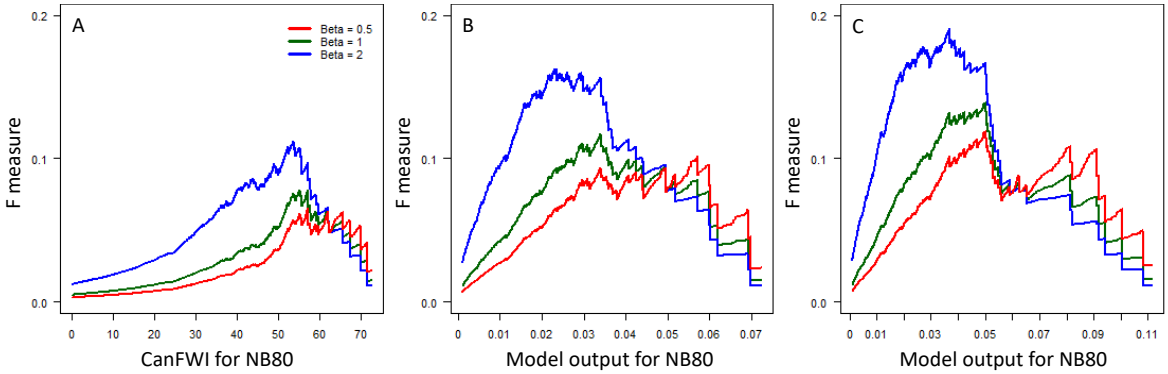


*Fig. (D.5): F_β measure for the expected number of >80ha fires prediction (NB80) as predicted through the CanFWI with no modelling framework and two models (i) CanFWI(12h) + STEs and (ii) DC*DMC*FFMC*Tmax*U + STEs.*

3. Article: Fire danger prediction enhanced through optical remote sensing in Mediterranean France.

A. Correlation matrices with optical remote sensing variables

In the context of the thesis report, the content of this supplementary material has been explained in the point 4.1.2 of the 4th thesis chapter.

B. Model evaluation for the full list of tested models

Table 1S Model evaluation (at ZFC/day scale) metrics for the 1ha occurrence component. Names in the table are as follows: LogLoss is the raw output from the score formula and Skill Scale is the normalization of the Skill Score values between the 0 and the maximal performance value, AUC is the value from an AUC-ROC based on the model fit, DIC is the parsimony criterion.

Occurrence 1ha				
model	**LogLoss**	**Skill Scale**	**AUC**	**DIC**
NULL	0,0008853	0,00	0,591	41264
B8A (NIR2)	0,0008005	29,16	0,570	37123
NDMI	0,0007891	34,20	0,627	36606
B11 (SWIR1)	0,0007882	34,63	0,630	36547
B5 (Red edge 1)	0,0007880	34,76	0,626	36553
NDVI	0,0007873	34,84	0,626	36525
FAb	0,0007782	39,00	0,771	35365
CanFWI (proxy)	0,0007501	46,91	0,780	34781
CanFWI (proxy) * B8A	0,0007479	47,76	0,785	34682
CanFWI (proxy) * NDMI	0,0007461	48,14	0,791	34583
CanFWI (proxy) * NDVI	0,0007450	48,60	0,793	34527
CanFWI (proxy) * B8A * B11	0,0007447	48,74	0,794	34527
DC * DMC * FFMC * VPD * U	0,0007419	49,96	0,802	34432
DC * DMC * FFMC * VPD * B8A * U	0,0007411	50,29	0,803	34395
DC * DMC * FFMC * VPD * NDMI * U	0,0007391	51,27	0,808	34279
DC * DMC * FFMC * VPD * NDVI * U	0,0007382	51,71	0,810	34237
CanFWI (proxy) * NDVI * LULC	0,0007381	51,75	0,814	34214
DC * DMC * FFMC * VPD * B8A * B11 * U	0,0007381	51,78	0,811	34231
CanFWI (proxy) * B8A * B11 * LULC	0,0007379	51,90	0,814	34213
DC * DMC * FFMC * VPD * U * NDVI * fuelONF	0,0007376	52,03	0,812	34197
DC * DMC * FFMC * VPD * U * B8A * B11 * fuelONF	0,0007376	52,04	0,812	34203
DC * DMC * FFMC * VPD * U * B8A * B11 * LULC5	0,0007367	52,46	0,815	34170
DC * DMC * FFMC * VPD * U * NDVI * LULC5	0,0007363	52,64	0,816	34153
DC * DMC * FFMC * VPD * U * B8A * B11 * LULC4	0,0007345	53,41	0,819	34063
DC * DMC * FFMC * VPD * U * NDVI * LULC4	0,0007343	53,62	0,819	34057
DC * DMC * FFMC * VPD * U * B8A * B11 * LULC2	0,0007341	53,92	0,821	34055
DC * DMC * FFMC * VPD * U * NDVI * LULC2	0,0007340	54,02	0,820	34045
DC * DMC * FFMC * VPD * U * B8A * B11 * LULC3	0,0007337	54,34	0,821	34036
DC * DMC * FFMC * VPD * U * NDVI * LULC3	0,0007336	54,50	0,821	34028
B8A (NIR2) * STEs	0,0007328	54,82	0,828	33800
DC * DMC * FFMC * VPD * U * NDVI * LULC	0,0007328	54,82	0,824	33988
DC * DMC * FFMC * VPD * U * B8A * B11 * LULC	0,0007326	54,91	0,825	33978
B11 (SWIR1) * STEs	0,0007299	56,40	0,835	33700
B5 (Red edge 1) * STEs	0,0007295	56,58	0,836	33670
NDMI * STEs	0,0007289	56,83	0,837	33661
NDVI * STEs	0,0007273	57,56	0,839	33588
DC * DMC * FFMC * VPD * U * NDVI * LULC * STEs	0,0007098	65,77	0,886	31668
DC * DMC * FFMC * VPD * U * B8A * B11 * LULC * STEs	0,0007091	66,06	0,886	31653
DC * DMC * FFMC * VPD * U * NDVI * LULC3 * STEs	0,0007066	67,13	0,886	31703
CanFWI (proxy) * STEs	0,0007066	67,14	0,872	32521
DC * DMC * FFMC * VPD * U * B8A * B11 * LULC3 * STEs	0,0007062	67,36	0,886	31691
CanFWI (proxy) * NDVI * LULC * STEs	0,0007057	67,56	0,882	31952
CanFWI (proxy) * B8A * STEs	0,0007056	67,60	0,874	32446
CanFWI (proxy) * B8A * B11 * LULC * STEs	0,0007054	67,70	0,882	31928
DC * DMC * FFMC * VPD * U * STEs	0,0007050	67,87	0,878	32215
CanFWI (proxy) * NDMI * STEs	0,0007050	67,91	0,876	32430
CanFWI (proxy) * NDVI * STEs	0,0007045	68,12	0,877	32404
DC * DMC * FFMC * VPD * NDMI * U * STEs	0,0007045	68,15	0,881	32162
DC * DMC * FFMC * VPD * B8A * U * STEs	0,0007045	68,19	0,879	32169
CanFWI (proxy) * B8A * B11 * STEs	0,0007044	68,27	0,876	32379
DC * DMC * FFMC * VPD * U * NDVI * LULC5 * STEs	0,0007042	68,37	0,883	32022
DC * DMC * FFMC * VPD * B8A * B11 * U * STEs	0,0007041	68,42	0,881	32116
DC * DMC * FFMC * VPD * NDVI * U * STEs	0,0007039	68,51	0,881	32128
DC * DMC * FFMC * VPD * U * B8A * B11 * LULC5 * STEs	0,0007039	68,53	0,883	32007
DC * DMC * FFMC * VPD * U * NDVI * fuelONF * STEs	0,0007034	68,74	0,882	32087
DC * DMC * FFMC * VPD * U * B8A * B11 * fuelONF * STEs	0,0007031	68,87	0,882	32079
DC * DMC * FFMC * VPD * U * NDVI * LULC2 * STEs	0,0007029	68,96	0,885	31807
DC * DMC * FFMC * VPD * U * NDVI * LULC4 * STEs	0,0007024	69,17	0,885	31845
DC * DMC * FFMC * VPD * U * B8A * B11 * LULC2 * STEs	0,0007023	69,21	0,886	31808
DC * DMC * FFMC * VPD * U * B8A * B11 * LULC4 * STEs	0,0007020	69,34	0,885	31839

Table 2S: Model evaluation (at pixel/day scale) metrics for the 80ha size component Names in the table are as follows: LogLoss is the raw output from the score formula, Skill Scale is the normalization of the Skill Score values between the 0 and the maximal performance value, AUC is the value from an AUC-ROC based on the model fit, DIC is the parsimony criterion.

Probability 80ha				
model	**LogLoss**	**Skill Scale**	**AUC**	**DIC**
NULL * perimeters	0,220	-73,24	0,500	744
B11 (SWIR1)	0,169	-2,38	0,694	714
FAb	0,168	0,00	0,715	718
B5 (Red edge 1)	0,165	3,78	0,713	707
NDMI	0,160	11,07	0,723	704
NDVI	0,157	15,30	0,741	697
B8A (NIR2)	0,152	21,86	0,756	684
B11 (SWIR1) * STEs	0,146	29,85	0,793	689
NDMI * STEs	0,146	30,70	0,789	689
FWIproxy	0,146	30,79	0,774	667
B5 (Red edge 1) * STEs	0,145	31,24	0,796	687
NDVI * STEs	0,144	32,27	0,796	684
FWIproxy * NDMI	0,141	36,61	0,780	664
FWIproxy * NDVI	0,141	37,50	0,786	659
FWIproxy * NDMI * U	0,141	37,63	0,786	666
FWIproxy * NDVI * U	0,140	38,81	0,792	661
DC * DMC * FFMCproxy	0,139	39,40	0,795	666
DMC * FFMCproxy * NDMI * U	0,138	41,38	0,795	667
DMC * FFMCproxy * NDVI * U	0,137	42,40	0,801	662
FWIproxy * STEs	0,135	45,78	0,825	655
B8A (NIR2) * STEs	0,134	46,27	0,816	659
FWIproxy * NDMI * STEs	0,133	48,62	0,830	655
DC * DMC * FFMCproxy * NDMI * U	0,133	48,67	0,815	656
FWIproxy * NDVI * STEs	0,133	48,70	0,830	651
FWIproxy * B8A	0,133	48,86	0,815	638
DC * DMC * FFMCproxy * Tmax * U	0,133	48,93	0,819	654
DC * DMC * FFMCproxy * NDVI * U	0,132	49,26	0,818	652
FWIproxy * NDMI * U * STEs	0,131	50,41	0,836	656
FWIproxy * B8A * U	0,131	50,46	0,821	639
FWIproxy * B8A * B11	0,131	50,54	0,821	639
FWIproxy * NDVI * U * STEs	0,131	50,92	0,836	652
DC * DMC * FFMCproxy * Tmax * NDMI * U	0,131	51,14	0,824	655
FWIproxy * B8A * B11 * U	0,130	51,86	0,827	641
DC * DMC * FFMCproxy * Tmax * NDVI * U	0,130	52,10	0,828	650
DC * DMC * FFMCproxy * STEs	0,130	52,76	0,836	652
DMC * FFMCproxy * NDMI * U * STEs	0,129	54,17	0,841	655
DMC * FFMCproxy * NDVI * U * STEs	0,129	54,51	0,843	652
DC * DMC * FFMCproxy * B8A * U	0,125	58,72	0,841	632
DC * DMC * FFMCproxy * B8A * B11 * U	0,124	60,36	0,847	633
FWIproxy * B8A * STEs	0,124	60,48	0,851	628
DC * DMC * FFMCproxy * Tmax * U * STEs	0,124	61,48	0,862	639
DC * DMC * FFMCproxy * NDMI * U * STEs	0,123	61,75	0,859	641
DC * DMC * FFMCproxy * NDVI * U * STEs	0,123	62,01	0,860	639
FWIproxy * B8A * B11 * STEs	0,123	62,07	0,856	629
FWIproxy * B8A * U * STEs	0,123	62,35	0,858	628
DC * DMC * FFMCproxy * Tmax * U * B8A * B11	0,123	62,81	0,854	630
DC * DMC * FFMCproxy * Tmax * NDMI * U * STEs	0,122	63,63	0,867	640
DC * DMC * FFMCproxy * Tmax * NDVI * U * STEs	0,122	63,82	0,867	637
FWIproxy * B8A * B11 * U * STEs	0,122	63,86	0,862	629
DC * DMC * FFMCproxy * Tmax * U * B8A * B11 * * fuelONF	0,119	67,75	0,864	622
DC * DMC * FFMCproxy * B8A * U * STEs	0,117	71,07	0,876	617
DC * DMC * FFMCproxy * B8A * B11 * U * STEs	0,116	72,42	0,880	618
DC * DMC * FFMCproxy * Tmax * U * B8A * B11 * * LULC5	0,115	73,59	0,877	611
DC * DMC * FFMCproxy * Tmax * U * B8A * B11 * STEs	0,114	74,25	0,886	617
DC * DMC * FFMCproxy * Tmax * U * B8A * B11 * LULC2	0,113	76,28	0,881	609
DC * DMC * FFMCproxy * Tmax * U * B8A * B11 * fuelONF * STEs	0,110	79,82	0,895	608
DC * DMC * FFMCproxy * Tmax * U * B8A * B11 * LULC5 * STEs	0,107	84,56	0,905	598
DC * DMC * FFMCproxy * Tmax * U * B8A * B11 * LULC2 * STEs	0,105	87,58	0,908	594
DC * DMC * FFMCproxy * Tmax * U * B8A * B11 * LULC	0,103	90,51	0,906	582
DC * DMC * FFMCproxy * Tmax * U * B8A * B11 * LULC * STEs	0,096	100,00	0,926	568

amU
Aix Marseille Université

NNT : 0000AIXM0000

THÈSE DE DOCTORAT (rapport français)

Soutenue à Aix-Marseille Université
Le 08 avril 2025 par

Jorge CASTEL-CLAVERA

Analyse et modélisation des dynamiques spatio-temporelles des incendies de forêt en région méditerranéenne : vers de nouveaux indicateurs de danger

Discipline
Sciences de l'environnement

Spécialité
Écologie

École doctorale
ED 251 – Sciences de l'Environnement

Laboratoire/Partenaires de recherche
Institut National de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE)

Composition du jury

Vincent MORON Professeur – AMU	Président du jury
Florent MOUILLOT Directeur de recherche – IRD	Rapporteur
Giada ADELFIO Professeure des universités – Università degli Studi di Palermo	Rapportrice
Imma OLIVERAS MENOR Directeur de recherche – IRD	Examinatrice
Maxime SOMA Ingénieur de recherche – INRAE	Examinateur
Jean-Luc Dupuy Directeur de recherche – INRAE	Directeur de thèse
François Pimont Ingénieur de recherche – INRAE	Co-encadrant de thèse (Membre invité)
Thomas Opitz Directeur de recherche – INRAE	Co-encadrant de thèse (Membre invité)

INRAE

1. Résumé introductif et objectifs

L'activité des incendies dans la région méditerranéenne française est pilotée par les conditions météorologiques, comme l'ont montré des recherches antérieures. Les facteurs humains jouent également un rôle crucial en raison de la forte densité de population, de l'utilisation et de l'occupation des sols (zones urbaines, routes, agriculture), ainsi que des activités humaines liées à la prévention et à la lutte. Cependant, l'identification des facteurs climatiques et non climatiques d'incendie dans le sud-est de la France reste difficile. De plus, la combinaison de ces différents facteurs de nature différente et la quantification de leur effet respectif présente des limites persistantes. Cette thèse vise à comprendre les interactions entre ces différents facteurs, afin de définir des indicateurs de danger d'incendie en France Méditerranéenne basés sur la prédiction probabiliste, utilisables à la fois par des acteurs scientifiques et opérationnels. Cet objectif principal peut être décliné en trois sous-objectifs : Premièrement, identifier les facteurs spatio-temporels explicites qui déterminent le danger d'incendie en testant leurs contributions aux régimes d'incendie récents en France méditerranéenne. Deuxièmement, examiner si des données météorologiques autres que celles actuellement utilisées opérationnellement, peuvent améliorer la prédiction des activités d'incendie à l'échelle quotidienne. Troisièmement, déterminer si des données de télédétection optique obtenues par satellite peuvent améliorer ou remplacer les variables météorologiques et/ou paysagères (occupation et utilisation des sols, rugosité du relief, facteurs anthropiques) dans la prévision des incendies.

2. Contexte scientifique – État de l'art

2.1. Les principaux facteurs déterminant l'activité des feux de forêt.

A l'échelle du régime de feu, soit à grande échelle spatio-temporelle, les facteurs de feu sont le climat, les caractéristiques du paysage et la fréquence des ignitions. Le régime de feu décrit les traits principaux des feux pour une région géographique donnée, à savoir leur fréquence, leur taille, leur intensité, leur saisonnalité et les causes des départs de feu. A l'échelle de l'incendie, ces facteurs sont les conditions météorologiques journalières, la quantité et le type de combustible ainsi que sa disponibilité au feu, et la topographie locale. La prise en compte de ces différents facteurs est nécessaire dans les travaux de modélisation et de prédiction du danger de feu.

La teneur en eau des combustibles végétaux morts (DFMC, pour Dead Fuel Moisture Content en anglais), influence la vitesse de propagation du feu, qui se propage plus rapidement lorsque ces combustibles fins (<6 mm) sont plus secs. La teneur en eau des combustibles vivants (LFMC, pour Live Fuel Moisture Content en anglais) est également importante, en particulier dans les zones arbustives ou dans le sous-bois en conditions de déficit hydrique (Pimont et al., 2019). Les combustibles vivants disposent de systèmes actifs de régulation de l'eau. Leurs réactions aux changements des conditions atmosphériques diffèrent de celles des DFMC. Le LFMC varie selon les espèces en présence en raison des différentes structures végétales et des mécanismes de résistance à la sécheresse propres à chaque espèce végétale. Les sources d'ignition sont l'élément le plus difficile à déterminer après le passage du feu ; mais on sait que la grande majorité des incendies (95%) en France méditerranéenne est d'origine humaine.

Les incendies sont une perturbation naturelle courante dans les écosystèmes forestiers, arbustifs et les prairies partout dans le monde. La plupart des espèces végétales méditerranéennes présentent des mécanismes d'adaptation, résistance et résilience au feu. Néanmoins, les augmentations de la fréquence, de l'intensité et de la taille des incendies créent une instabilité et entraînent une dégradation de l'écosystème. Par exemple, si les incendies deviennent trop fréquents, le temps nécessaire aux processus de régénération forestière (production de graines et d'autres ressources nécessaires à la régénération) peut devenir trop court entre les perturbations. L'écosystème forestier peut alors entrer dans un processus de régression, qui conduit à une transition vers des systèmes arbustifs ou des maquis, et finalement vers des prairies ou des sols nus. La régression écologique est également accentuée par la limite de dispersion des graines.

2.2. La région méditerranéenne française et les données de feu.

Dans ce travail, nous nous intéressons aux 15 départements constituant la région méditerranéenne française, qui concentre la majorité des activités de feu de la France Métropolitaine. La région méditerranéenne française est un espace qui présente des enjeux sociétaux et économiques forts en raison de la forte densité de population, de nombreuses activités industrielles et des intérêts touristiques. Dans cette zone, une importante base de données sur les incendies (Promethée) a été développée et entretenue par les autorités publiques françaises, ce qui permet de soutenir cette recherche. La base de données Prométhee (http://www.promethee.com/) enregistre toute l'activité des

incendies depuis 1973. L'objectif de Promethée est de répertorier et fournir de manière homogène les relevés d'activité des incendies entre les différentes institutions françaises concernées par ce phénomène. La base de données est structurée sous la forme d'une grille de pixels de 2x2 km (grille DFCI ; pour Défense des Forêts Contre les Incendies). Chaque incendie est enregistré dans le pixel correspondant à l'endroit où le départ de feu a été signalé, avec la date, la surface brûlée finale et la source d'ignition (si connue). Récemment en 2024, cet ensemble de données a été clôturé et remplacé par la Base de Données sur les Incendies de Forêts en France (BDIFF), gérée par l'IGN. Cependant, le présent travail a été réalisé avec les données extraites précédemment de Prométhee, complétées par des informations sur les périmètres brûlés obtenus grâce à l'Office National des Forêts (ONF).

2.3. Prédiction du danger de feu en France

En France, Météo-France développe la méthodologie d'évaluation du danger d'incendie qui est utilisée pour la prise de décision opérationnelle. Cette méthodologie est basée en grande partie sur le *Canadian Fire-Weather Index System* (CFWIS) et utilise ses composantes *Drought Code* (DC), *Duff Moisture Code* (DMC) et *Fine Fuel Moisture Code* (FFMC), ainsi que l'indice de danger final du système, le *Canadian Fire Weather Index* (CanFWI), ou Indice Feu-Météo Canadien en français. Ces indices ont été recombinés pour former des indices opérationnels français, comme par exemple l'IEP ou NSV2 de Météo-France. Un indice feu-météo est une valeur numérique ordonnée qui évalue le danger de feu à travers des observations météorologiques. Ce système fondé essentiellement sur la météorologie n'intègre pas les autres facteurs de feu importants pour définir le danger, comme les caractéristiques locales du territoire et du combustible. De plus, les indices correspondent à une quantification abstraite du niveau de danger qui ne donne pas une interprétation quantitative de la probabilité d'occurrence et de la taille potentielle d'un feu. De plus, la méthode ne considère pas l'incertitude associée à la prédiction des activités de feu, un élément important lors de la prise de décisions pour la gestion du feu ou pour la comparaison des scénarios futurs de régimes de feu.

2.4. Un cadre probabiliste pour les feux de forêt en France méditerranéenne

Cette thèse utilise un cadre probabiliste bayésien pour construire une méthode de prédiction de l'activité de feu en France méditerranéenne. Tout d'abord, l'analyse des

incendies passés permet de déterminer les facteurs spécifiques qui ont contrôlé les occurrences et le comportement des feux. Ensuite, la prédiction doit informer de la difficulté attendue pour la lutte contre l'incendie. Le cadre probabiliste constitue une méthode efficace pour surmonter l'incertitude associée au caractère stochastique et à la rareté des incendies. Un cadre probabiliste permet de combiner différents types de facteurs d'incendies dans une seule expression ou équation mathématique. En conséquence, ces cadres permettent de combiner des facteurs d'incendie provenant de différentes sources (météorologiques, paysagères, ou relevés de terrain). De plus, la télédétection par satellite peut fournir des données pour décrire la teneur en eau des combustibles, leur structure ou encore leur composition. Les modèles probabilistes associés fournissent soit par une probabilité, soit par un nombre attendu d'événements.

L'étude probabiliste sur les incendies en France méditerranéenne initiée dans les travaux de thèse d'Hélène Fargeon (Fargeon 2019) a donné lieu à *Firelihood* (Pimont et al., 2021). *Firelihood* modélise l'occurrence et la taille des incendies en deux composantes distinctes. Le cadre formel de *Firelihood* permet l'ajout d'effets spatiaux et temporels (*Spatio-temporal effects*, STEs) pour les prédictions d'occurrence et de taille. Les STEs capturent des processus non expliqués par les variables disponibles. Le modèle résultant, une fois validé, peut être utilisé pour émettre des prévisions à court terme ou des projections à long terme sur l'activité des incendies. Le cadre probabiliste utilisé dans cette thèse a l'avantage de permettre l'agrégation spatiale et/ou temporelle des simulations des activités de feu. En effet, grâce à la propriété additive du processus de Poisson utilisé dans ce cadre, nous pouvons sommer les nombres de feux prédits sur différences échelles.

3. Influence des facteurs non-climatiques sur la distribution spatiotemporelle de l'activité de feux et leurs changements à longue terme.

Ce chapitre a pour objectif d'identifier et de quantifier le rôle des facteurs non climatiques dans l'activité spatio-temporelle des feux de végétation ainsi que dans leurs tendances temporelles à long terme. Pour explorer les facteurs non climatiques de l'activité des feux, il est essentiel de se concentrer sur le paysage et l'influence des activités humaines. Par conséquent, dans cette étude, le modèle prédictif bénéficiera de l'inclusion des facteurs liés au paysage et aux activités humaines sous la forme de données d'usage et occupation du sol, ce qui devrait permettre de meilleures prédictions. Les changements

à long terme seront étudiés entre les décennies 1993-2002 et 2009-2018. Nous avons observé une réduction globale du nombre d'incendies pour toutes les catégories de taille entre les deux décennies, avec un changement brusque après l'année 2003.

La première version du modèle *Firelihood* (Pimont et al., 2021), qui montre un potentiel prédictif prometteur, reposait sur une contribution importante des STEs pour représenter l'activité de feu observée. Les STEs sont composés d'un facteur spatial aléatoire (les coordonnées spatiales de chque pixel) et d'un effet saisonnier aléatoire (la semaine de l'année). Les STEs correspondent aux facteurs qui ne sont pas décrits de forme directe avec les prédicteurs introduits sur le modèle. Ces éléments permettent à *Firelihood* de capturer la variabilité spatio-temporelle résiduelle des activités de feu. L'inclusion de facteurs liés au paysage et aux activités humaines cherche à expliciter les effets qui jusqu'à maintenant sont capturés uniquement par les STEs. De plus, dans ce chapitre, les STEs comprennent aussi un facteur qui décrit l'évolution de la composante spatiale au cours des années.

Les résultats principaux du chapitre montrent l'importance des facteurs de paysage dans la compréhension des patrons spatiaux d'activité de feu en France méditerranéenne. Nous avons retenu huit facteurs humains et du paysage dans chaque composante du modèle probabiliste (occurrence et taille). Pour les occurrences des feux supérieures à 1ha, les principaux facteurs non-climatiques sont la classification des types de combustible (produite par l'Office National des Forêts), la densité de routes et l'étendue des zones agricoles. Pour les tailles de feux qui dépassent 10 et 100ha, ce sont l'étendue des zones arbustives et agricoles, ainsi que la pente du terrain.

Les facteurs climatiques expliquent en grande partie la variance des grands feux (10 et 100ha). Néanmoins, les STEs sont aussi responsables d'une forte proportion de la variance des prédictions, spécialement pour les occurrences de feux d'un hectare, ce qui veut dire qu'une grande partie des patrons d'activité ne sont expliqués ni par la météo ou le climat, ni par les facteurs usuels de l'activité des feux. L'évolution temporelle de l'activité des incendies à long terme est très différente entre l'est et l'ouest de la région étudiée. Alors que le côté est de la partie continentale et la Corse montrent une réduction généralisée du nombre d'incendies, ce nombre augmente du côté ouest. Ces changements ne s'expliquent pas par l'occupation des sols ou les conditions météorologiques liées aux incendies. En revanche, si seuls les facteurs climatiques avaient influencé l'évolution temporelle de l'activité des feux de forêt, il y aurait eu une augmentation généralisée du

nombre de feux au cours de la dernière décennie. Une interprétation plausible du fait que les variables testées expliquent peu la tendance temporelle observée pourrait être l'effet de la lutte contre les incendies. Il convient cependant de noter que ce travail a été réalisé à partir des données feux jusqu'en 2021 ? et n'intègre donc pas les évolutions les plus récentes de l'activité des feux.

Ce travail a donné lieu à une publication scientifique et une présentation orale en conférence internationale. Pour plus d'informations consultez les références ci-dessous :

- Castel-Clavera, J., Pimont, F., Opitz, T., Ruffault, J., Rivière, M., Dupuy, J-L., 2022. Disentangling the Factors of Spatio-Temporal Patterns of Wildfire Activity in South-Eastern France. *International Journal of Wildland Fire* 32. https://doi.org/10.1071/WF22086.
- Présentation orale 9th ICFFR (International Conference on Forest Fire Research). Communication écrite : Castel-Clavera, J., Pimont, F., Opitz, T., Ruffault, J., et Dupuy, J-L. 2022. Modelling the Influence of Regional Landscape Drivers on Spatio-Temporal Patterns of Wildfire Activity. In *Advances in Forest Fire Research 2022*, par Domingos Xavier Viegas and Luís Mário Ribeiro, 1228-33, 1.ª Edição. Imprensa da Universidade de Coimbra. https://doi.org/10.14195/978-989-26-2298-9_186.

4. Analyse détaillée des facteurs de la météorologie

L'étude de Pimont et al. (2021) utilise un seul indice météorologique d'incendie (CanFWI) dans le cadre de *Firelihood*. Parmi la diversité des indices feu-météo disponibles, il n'est pas certain et il n'a pas été démontré que le CanFWI soit le meilleur choix. Par exemple, un biais saisonnier entre le CanFWI et l'activité d'incendie a déjà été identifié. L'objectif de ce chapitre est de déterminer quels indices feu-météo et quelles variables météorologiques décrivent le mieux les facteurs climatiques liés à l'activité journalière des feux de forêt en France méditerranéenne. Les STEs sont encore nécessaires pour atteindre la meilleure capacité de prédiction possible des modèles probabilistes, mais ils s'adaptent à chaque configuration des variables prédictives. Autrement dit, lorsqu'on change de variable prédictive, par exemple en passant d'un indice feu-météo à un autre, les estimations des effets aléatoires (STEs) changent. Cela rend la comparaison des performances entre les modèles complexe. Une stratégie pour surmonter cette contrainte est la définition d'un modèle de référence auquel tous les autres modèles seront comparés. Ce modèle de référence, estimé en intégrant uniquement les STEs, vise à décrire l'activité moyenne attendue au cours d'une saison et selon l'endroit du territoire, indépendamment de la météorologie réelle qui se produit un jour donné. Il peut être vu comme un équivalent au climatogramme en météorologie, sauf

qu'il décrit ici l'activité des feux. Les comparaisons se font sur la base du *LogLoss* score, qui est une mesure de performance des modèles probabilistes adaptable aux évènements rares et stochastiques. Cette approche innovante permet de mieux identifier la contribution des facteurs météorologiques quotidiens dans un contexte où les variations spatio-temporelles de l'activité des feux sont largement dominées par leurs variations saisonnières et spatiales.

Les résultats confirment l'importance des STEs pour la prédiction du danger d'incendie. Les STEs étaient nécessaires pour presque tous les prédicteurs testés afin de surpasser la performance du modèle de base. Dans le cadre de cette étude, le CFWIS fournit les prédictions les plus précises de l'activité des incendies par modélisation probabiliste pour la saison estivale en France méditerranéenne. Néanmoins, les performances du CanFWI ont été améliorées avec l'inclusion du VPD (déficit de pression de vapeur d'eau de l'atmosphère) et de la vitesse du vent, bien que le CanFWI intègre déjà l'humidité relative de l'air et la vitesse du vent. Le modèle le plus performant de l'étude incluait les composantes DC, DMC et FFMC du CFWIS, ainsi que le VPD maximal journalier, la vitesse moyenne du vent et les STEs. Cela montre que, même si le CFWIS offre des estimations acceptables du danger d'incendie, on peut construire des combinaisons des effets du vent et de la sécheresse à court terme plus représentative de l'activité des feux et qui peuvent différer selon qu'on s'intéresse aux petits ou aux grands feux. De plus, les résultats obtenus dans ce chapitre montrent qu'il n'existe pas une différence significative entre les prédictions faites avec un indice feu-météo type « maximal journalier » et sa valeur calculée à midi.

Ce travail a donné lieu à une publication scientifique. Pour plus d'informations consultez le matériel ci-dessus :

- Castel-Clavera, J., Pimont, F., Opitz, T., Ruffault, J., Barbero, R., Allard, D., Dupuy, J-L., 2025. A comparative analysis of fire-weather indices for enhanced fire activity prediction with probabilistic approaches. *Agricultural and Forest Meteorology* 361. https://doi.org/10.1016/j.agrformet.2024.110315.

5. Apports de la télédétection optique par satellite dans la prédiction spatiale et temporelle des activités d'incendie

Aucune des variables mentionnées dans les travaux précédents n'est une mesure directe des caractéristiques physiques du combustible avant l'incendie. Les données d'usage et d'occupation des sols donnent une indication générale sur les types de

combustibles susceptibles de couvrir la zone d'étude et leurs caractéristiques. Cependant, ces données ne fournissent pas d'informations très détaillées dans les dimensions spatiales et temporelles. L'état phénologique de la végétation au fil de la saison, son état hydrique, ainsi que les variations interannuelles de l'abondance des combustibles, ne sont pas fournis par les données d'usage et d'occupation des sols. Les indices météorologiques donnent des estimations de la sécheresse et de la teneur en eau du combustible. Cependant, ces estimations ne prennent pas en compte la présence et la composition des combustibles. Ce chapitre s'articule autour de l'hypothèse selon laquelle les données de réflectance optique mesurée par des satellites tels que ceux de MODIS ou Sentinel-2 peuvent apporter de l'information sur les effets structurels et hydriques de la végétation lorsqu'elles sont introduites comme facteur dans un modèle probabiliste d'activité d'incendie. Compte tenu de la nature spatiale des données de télédétection, elles pourraient améliorer les prédictions du modèle probabiliste, en particulier en ce qui concerne les gradients spatiaux de l'activité d'incendie.

Cette étude est une première utilisant la télédétection optique par satellite pour la prédiction probabiliste de l'activité d'incendie en France méditerranéenne. L'étude a démontré le potentiel d'application des données de télédétection optique pour ces prédictions. Les résultats montrent que les meilleures variables de télédétection parmi les variables étudiées sont l'infrarouge proche et l'infrarouge d'onde courte (NIR et SWIR). Les effets obtenus avec les variables de la télédétection optique sont plus importants pour les prédictions de taille des feux que pour les prédictions des occurrences. Dans tous les cas, les performances apportées par les données de télédétection optique ne supplantent pas l'apport des données météorologiques, même en ce qui concerne la sécheresse de moyen et long terme. Les informations fournies par ces deux sources de données sont complémentaires.

La partition de la variance montre que, pour les occurrences de feux, les données de télédétection par satellite apportent principalement de l'information concernant la dimension spatiale. En revanche, pour les tailles des feux, la variance de la télédétection est partagée entre les dimensions spatiale et temporelle. Les autres groupes de variables explicites apportent plus de variance aux prédictions des activités de feu que la télédétection optique. Les variables d'utilisation et occupation des sols apportent presque toute sa variance sur la dimension spatiale. Les variables météorologiques à court terme, comme le vent, le FFMC, la température ou le VPD, sont les principales responsables de la

variance saisonnière pour les occurrences. Dans les modèles de taille, ce sont les variables météorologiques saisonnières (DC, DMC, variant typiquement de la semaine au mois) et les données de télédétection qui apportent une grande partie de la variance.

Ce travail donnera lieu à une publication scientifique. Pour plus d'informations consultez le chapitre 4 du manuscrit de thèse pour trouver la proposition d'article :

- Castel-Clavera, J., Pimont, F., Opitz, T., Ruffault, J., Olioso, A., Penot, V., Rodríguez-Suquet, R., Dupuy, J-L., in prep. Fire danger prediction enhanced through remote sensing in Mediterranean France.

6. Discussion et conclusion

Le cadre de modélisation probabiliste a montré sa capacité à combiner les effets de plusieurs sources de données de nature très différentes et de quantifier les effets de chacun sur l'activité des feux, en distinguant ceux de l'occurrence des feux de plus de 1ha et ceux conduisant à des grands feux (>80ha). Chaque source apporte des informations spécifiques sur les facteurs qui influencent les activités d'incendie pour obtenir une description complète des régimes d'incendie. De plus, les effets spatio-temporels (STEs) décrits antérieurement, sont un élément clé pour capturer la variabilité des activités des feux de forêt. Étant donné la complexité des STEs, les travaux présentés ont nécessité des apports méthodologiques pour garantir une comparaison juste de l'apport de chaque variable à la prédiction de danger. Nous avons développé une nouvelle méthode d'évaluation des modèles adaptée au cadre *Firelihood*. Cette méthode consiste en une comparaison entre un modèle de base, intégrant seulement les STEs et aucun facteur explicite, et chaque test de modélisation. Malheureusement, pour atteindre un niveau élevé de précision, ces effets spatio-temporels nécessitent une grande quantité de données pour leur entraînement, qui ne sont pas toujours disponibles. En conséquence, les effets spatio-temporels ont été adaptés aux objectifs spécifiques de chaque chapitre.

Pour l'étude des patrons spatiaux des activités des feux et leur évaluation à long-terme, nous avons choisi l'échelle annuelle. Ceci nous a permis de comparer les activités de feux pendant deux décennies, pour chercher les causes des changements multi-annuels, en gardant une haute résolution spatiale (soit 2km). Ensuite, nous avons utilisé *Firelihood* pour obtenir les meilleurs prédicteurs possibles de l'activité de feu journalière. Néanmoins, d'après les services de lutte et prévention contre les incendies, les prédictions journalières ne nécessitent pas une haute résolution spatiale. En conséquence, les

prédictions dans les pixels (2 km)ont été spatialement agrégées au sein des Zones homogènes de météorologie du feu, tout en gardant leur résolution journalière.

La calibration d'un modèle à une échelle spatiale et temporelle fine (pixel/jour) pour ensuite agréger les simulations à des échelles plus larges ne constitue pas un gaspillage de ressources de calcul. Un processus de modélisation effectué directement à une échelle plus large pourrait alors paraître plus pratique. Néanmoins, cette thèse défend l'utilisation d'échelles fines pour l'ajustement et la simulation des modèles. Les petites échelles permettent d'extraire autant de variabilité et d'information que possible des variables avalisées. L'agrégation des variables météorologiques à une échelle temporelle plus large négligerait des événements météorologiques extrêmes, souvent liés à la survenue des plus grands incendies de forêt. De même, l'agrégation des données paysagères à des échelles spatiales plus larges pourrait rendre invisibles certaines particularités locales, essentielles pour la prévision des incendies.

Ce travail offre de nombreuses perspectives, tant du point de vue de futures recherches que des applications opérationnelles. Les applications potentielles du cadre *Firelihood* peuvent être séparées en deux catégories. La première concerne la prévision du danger à court terme, pour améliorer les prévisions quotidiennes ou hebdomadaires du danger d'incendie à une échelle spatiale adéquate pour utilisateurs finaux. La seconde catégorie de produits, orientée vers des applications scientifiques et de politiques publiques, fournit des cartographies des tendances à long terme des régimes d'incendie. Cette application permet aussi la cartographie des risques d'incendie projetés dans le futur. Finalement, plusieurs perspectives d'amélioration pour les travaux futurs ont été identifiées, ainsi que des extensions potentielles du cadre probabiliste. L'acquisition de données sur l'utilisation et l'occupation des sols à plus haute résolution, des tests supplémentaires avec des produits de télédétection et l'analyse des incendies d'hiver font partie de ces améliorations futures.