

Anàlisi dels factors condicionants d'incendis forestals i predicció del risc a Catalunya

Uoc

Xavier Pascuet Capdevila

Àrea 1

Tutora de TFM

Anna Muñoz Bollas

**Professor responsable de
l'assignatura**

Albert Solé Ribalta

08/01/2022

**Universitat Oberta
de Catalunya**



Aquesta obra està subjecta a una llicència de [Reconeixement-NoComercial-SenseObraDerivada 3.0 Espanya de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/3.0/es/)

Fitxa del Treball Final

Títol del treball:	Anàlisi dels factors condicionants d'incendis forestals i predicció del risc a Catalunya.
Nom de l'autor/a:	Xavier Pascuet Capdevila
Nom del Tutor/a de TF:	Anna Muñoz Bollas
Nom del/de la PRA:	Albert Solé Ribalta
Data de lliurament:	01/2023
Titulació o programa:	Màster en Ciència de Dades
Àrea del Treball Final:	Àrea 1
Idioma del treball:	Català
Paraules clau	Geographic Information Systems, Wildfire, Neural Network, Catalunya.
Resum del Treball	
<p>L'objectiu d'aquest treball es analitzar les dades històriques dels incendis forestals a Catalunya. S'utilitzen diversos conjunt de dades d'incendis, meteorològiques, orogràfiques i de vegetació, amb els quals es pretén obtenir les variables determinants dels incendis.</p> <p>A partir d'aquest anàlisi i mitjançant tècniques d'aprenentatge automàtic es pretén construir una xarxa neuronal que predigui el risc d'incendis forestals, per a un dia determinat.</p> <p>Amb aquests model entrenat sobre les dades històriques, s'han generat mapes de risc d'incendi per a 4 dies diferents.</p> <p>A més, s'utilitzaran eines de teledetecció per obtenir les àrees cremades en els 2 grans incendis forestals del 2022 i quantificar el carboni emès i la biomassa perduda.</p>	

Abstract

The aim of this work is to analyze historical data on forest fires in Catalonia. Various sets of fire, meteorological, orographic and vegetation data are used, with which it is intended to obtain the determining variables of the fires.

Based on this analysis and using machine learning techniques, the aim is to build a neural network that predicts the risk of forest fires, for a given day.

With this model trained on the historical data, fire risk maps were generated for 4 different days.

In addition, remote sensing tools will be used to obtain the areas burned in the 2 major wildfires of 2022 and quantify the carbon emitted and biomass lost.

Índex

1.	Introducció	1
1.1.	Context i justificació del Treball	1
1.2.	Motivació	2
1.3.	Objectius del Treball	2
1.4.	Impacte en sostenibilitat, ètic-social i de diversitat	4
1.5.	Enfocament i mètode seguit	5
1.6.	Planificació del Treball	6
2.	Estat de l'art	9
3.	Implementació del treball	16
3.1.	Origen de dades	16
3.2.	Descripció de les dades	17
3.2.1	Incendis	17
3.2.2.	Orogràfiques.....	17
3.2.3.	Meteorològiques.....	17
3.2.4.	Vegetació.....	17
3.3.	Preparació de les dades.....	18
3.3.1.	Neteja de les dades	18
3.3.2.	Correlacions	18
3.3.3.	Reclassificació de cobertes del sòl.....	18
3.3.4.	Generació del conjunt d'exemples.....	21
3.4.	Estadística inferencial	22
3.4.1.	Variable altitud.....	23
3.4.2.	Variable pendent	24
3.4.3.	Variable orientació	25
3.4.4.	Variable temperatura diària màxima	26
3.4.5.	Variable component u del vent	27
3.4.6.	Variable component v del vent.....	28
3.4.7.	Variable humitat específica diària mínima	29
3.4.8.	Variable humitat relativa diària mínima	30
3.4.9.	Variable coberta del sòl	31
3.5.	Importància de variables.....	31
3.6.	Creació de la xarxa neuronal	34
3.7.	Generació de mapes per a dies concrets	34
3.8.	Anàlisi dels Grans Incendis forestals del 2022	38
3.8.1.	Incendi Baldomà 15/06/2022	39
3.8.2.	Incendi El Pont de Vilomara 17/07/2022	42
4.	Conclusions	45
5.	Futures línies de treball	46
6.	Glossari	47
7.	Bibliografia	48
8.	Annexos.....	50

Llista de Figures

Figura 1: Dades incendis forestal 1981-2021	1
Figura 2: Metodologia de la cascada	5
Figura 3: Diagrama de Gantt	8
Figura 4: Mapa de variables biofísiques de l'arbrat. Biomassa	13
Figura 5: Mapa Forestal de España	14
Figura 6: Matriu de correlacions	18
Figura 7: Primera predicció pel 15 de juny de 2022	20
Figura 8: Conjunt d'exemples	21
Figura 9: Diagrama de caixes altitud	23
Figura 10: Histogrames altitud	23
Figura 11: Diagrama de caixes pendent	24
Figura 12: Histogrames pendent	24
Figura 13: Diagrama de caixes orientació	25
Figura 14: Histogrames orientació	25
Figura 15: Diagrama de caixes temperatura diària màxima	26
Figura 16: Histogrames variable temperatura diària màxima	26
Figura 17: Diagrama de caixes component u del vent	27
Figura 18: Histogrames component u del vent	27
Figura 19: Diagrama de caixes component v del vent	28
Figura 20: Histogrames component v del vent	28
Figura 21: Diagrama de caixes humitat específica diària mínima	29
Figura 22: Histogrames humitat específica diària mínima	29
Figura 23: Diagrama de caixes humitat relativa diària mínima	30
Figura 24: Histogrames humitat relativa diària mínima	30
Figura 25: Diagrama de barres coberta del sol	31
Figura 26: Diagrama de barres coeficients assignats regressió	32
Figura 27: Diagrama de barres coeficients assignats regressió (absoluts)	32
Figura 28: Diagrama de barres coeficients assignats regressió (absoluts)	33
Figura 29: Exemple 9 bandes del ràster	35
Figura 30: Predicció risc d'incendi 8/03/2012	36
Figura 31: Predicció risc d'incendi 20/09/2017	37
Figura 32: Predicció risc d'incendi 2/01/2021	37
Figura 33: Predicció risc d'incendi 15/06/2022	38
Figura 34: Mapa Severitat Incendi de Baldomà	41
Figura 35: Mapa Severitat Incendi El Pont de Vilomara	44

Llista de Taules

Taula 1: Incendis forestals per causa 1995-2021	9
Taula 2: Incendis forestals per mesos 1995-2021	11
Taula 3: Biomassa consumida per intensitat de foc	14
Taula 4: Factor d'emissió del gasos.	15
Taula 5: Dades d'origen	16
Taula 6: Reclassificació cobertes del sòl	19
Taula 7: Mostra taula d'atributs dels punts	21
Taula 8: Emissions coníferes Incendi Baldomà	39
Taula 9: Resum coníferes Incendi Baldomà	39
Taula 10: Emissions frondoses Incendi Baldomà	39
Taula 11: Resum frondoses Incendi Baldomà	40
Taula 12: Emissions estrat arbustiu Incendi Baldomà	40
Taula 13: Resum estrat arbustiu Incendi Baldomà	40
Taula 14: Resum total Incendi Baldomà	40
Taula 15: Emissions coníferes Incendi El Pont de Vilomara	42
Taula 16: Resum coníferes Incendi El Pont de Vilomara	42
Taula 17: Emissions frondoses Incendi El Pont de Vilomara	42
Taula 18: Resum frondoses Incendi El Pont de Vilomara	42
Taula 19: Emissions estrat arbustiu Incendi El Pont de Vilomara	43
Taula 20: Resum estrat arbustiu Incendi El Pont de Vilomara	43
Taula 21: Resum total Incendi El Pont de Vilomara	43

1. Introducció

1.1. Context i justificació del Treball

Els incendis forestals són un dels principals agents alterants del paisatge forestal a Catalunya. Impliquen una reducció de la superfície vegetal, amb els conseqüents riscos d'erosió, pèrdua d'habitat per a fauna, emissió de carboni a l'atmosfera, causen nombrosos danys materials i fins i tot humans.

Dos dels factors més determinats per a l'aparició d'incendis són la temperatura i precipitació (influeix directament en la humitat del combustible). En el context actual de canvi climàtic, amb augment de temperatures i de fenòmens extrems com les sequeres, els incendis estan augmentant en tot l'àmbit mundial [1]. No obstant aquest augment no està succeint a Catalunya, en els darrers anys.

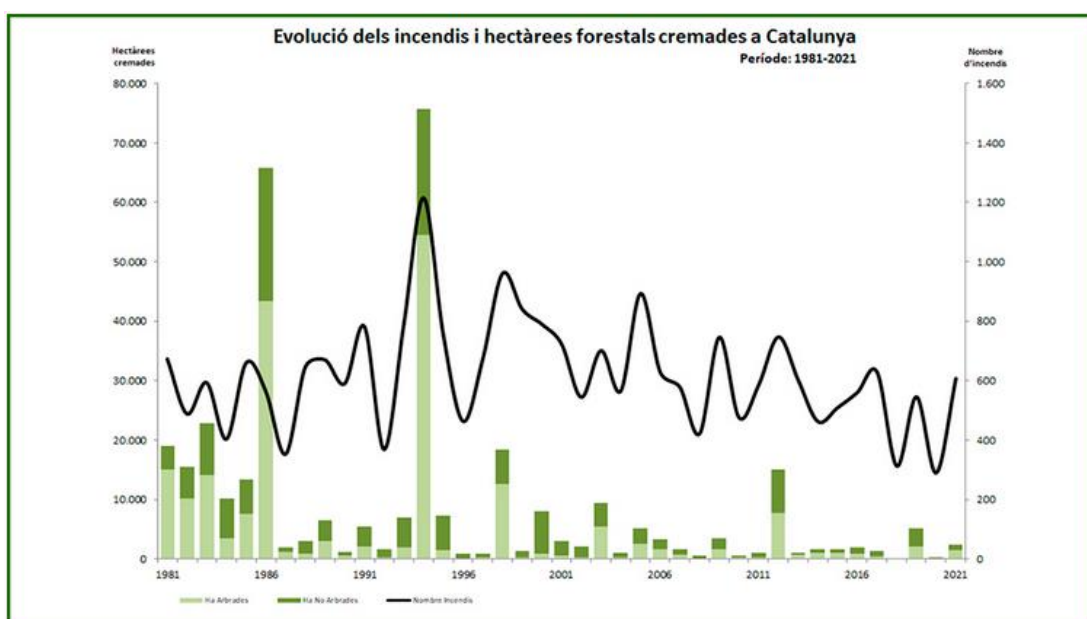


Figura 1: Dades incendis forestal 1981-2021. Font: Departament d'Acció Climàtica, Alimentació i Agenda Rural

Això és deu principalment a dos factors:

- La millora de l'extinció. En els últims anys ha augmentat molt la despesa en extinció d'incendis, així com s'han modernitzat els mitjans i s'ha professionalitzat molt el personal que treballa en ella.
- Les mesures de prevenció: S'han implementat diverses mesures. La prohibició de fer foc en terreny forestal entre el 15 de març i el 15 d'octubre, el Pla Alfa, que restringeix les activitats en terreny forestal en períodes d'alt risc d'incendi, són uns alguns exemples.

Per tal de poder fer una bona prevenció sobre els períodes i llocs d'elevat risc és imprescindible identificar les variables determinants dels incendis. L'obtenció de mapes de risc d'incendis també permet distribuir els mitjans d'extinció de manera efectiva.

El mapeig de la severitat dels incendis mitjançant eines de teledetecció es clau per tal de poder realitzar una bona restauració de les àrees cremades, ja que ofereix resultats fiables i ràpids[2] de quines són les zones més afectades i el grau d'afectació. A més també permet mesurar i qualificar els seus impactes.

1.2. Motivació

Des de petit que m'han fascinat els boscos, com a ecosistema complex i ric, fruit d'això vaig decidir estudiar enginyeria forestal. Els incendis, com a un dels grans fenòmens pertorbadors en l'àmbit mediterrani desperten molt interès dins de l'enginyeria forestal.

Posteriorment a finalitzar els estudis d'enginyeria em van interessar els Sistemes d'Informació Geogràfica i em vaig formar en aquest àmbit. Com a conseqüència vaig començar a programar i, donat que l'estadística sempre m'havia interessat, he cursat aquest màster en ciència de dades.

Així doncs, aquest treball em permet combinar 3 àrees de coneixement que m'apassionen: l'enginyeria forestal, els sistemes d'informació geogràfica i la ciència de dades. També em permetrà conèixer quins han estat els incendis històrics a Catalunya i quines són les variables més determinants.

A més, recentment m'he incorporat a treballar al Servei de Prevenció d'incendis, del Departament d'Acció Climàtica Alimentació i Agenda Rural [3] com a tècnic. Al servei s'elabora un mapa de perill d'incendis diàriament, que es la base en la qual s'adopta el pla Alfa, que restringeix l'accés al medi natural en períodes d'alt risc.

Aquest mapa s'obté mitjançant un model físic a partir de dades topogràfiques, de vegetació, satel·litals, meteorològiques, mostres d'humitat en la vegetació fets a camp, i les prediccions meteorològiques dels models AROME i ECMWF.

En un futur, la idea es substituir aquest model per una xarxa neuronal. Encara que actualment no disposo del hardware ni del temps necessari elaborar un model de xarxa neuronal que inclogui el gran volum de dades amb el que es treballa actualment, aquesta xarxa podrà servir de base de partida per a un projecte futur.

1.3. Objectius del Treball

Aquest treball pretén analitzar les dades dels incendis a Catalunya durant els últims anys, i en funció d'això elaborar un mapa de risc utilitzant una xarxa neuronal. A més és pretén quantificar el carboni emès en els dos grans incendis forestal d'aquest estiu.

1. Objectius generals

- Desenvolupar un projecte d'informació geogràfica
- Investigar l'estat de l'art de models predictius d'incendis
- Obtenir i integrar les dades procedents de diverses fonts
- Documentar el treball mitjançant una memòria
- Realitzar una presentació que resumeixi els principals punts del projecte
- Exposar el projecte davant d'un tribunal

2. Objectius específics

- Realitzar una anàlisi geoespacial utilitzant eines SIG
- Preparar i ajustar una xarxa neuronal sobre dades geogràfiques
- Crear un model predictiu del risc diari d'incendis
- Mapejar la severitat i quantificar el carboni emès en els 2 grans incendis de l'estiu del 2022
- Implementar un visor d'informació geogràfica que mostri la probabilitat d'ignició dels diferents punts del mapa de Catalunya

1.4. Impacte en sostenibilitat, ètic-social i de diversitat

Pel que fa a les dimensions de la competència transversal UOC "Compromís ètic i global" el treball impacta en:

Sostenibilitat

Els incendis suposen una important emissió de carboni a l'atmosfera majoritàriament en forma de CO₂, gas que és el principal responsable del canvi climàtic, això estaria amb concordança amb l'objectiu de Desenvolupament Sostenible (ODS) 2030, ONU. 13 "Climate Action"

També suposen una pèrdua de biodiversitat, tant de flora com de fauna, això s'enquadraria amb l'ODS 15 "Life on land".

Aquest impacte positiu no és inherent al propi treball, però el treball és una eina que facilitarà la prevenció dels incendis i avaluar la severitat dels incendis permetran també millorar la planificació de la rehabilitació post-incendi.

Comportament ètic i responsabilitat social

Els incendis comporten grans pèrdues materials i a vegades fins i tot pèrdues humanes, Així que, donat que el treball facilitarà la seva prevenció, podem afirmar que impacta ens els ODS 5 - No poverty i ODS 8 - Decent work and economic growth.

Diversitat i drets humans

Aquest TFM, per la seva temàtica no té de manera inherent cap impacte en aquests aspectes. Es tracta d'un treball tècnic que no impacta en aspectes de gènere, diversitat o drets humans.

Per tal de que no tingui un impacte negatiu, en el redactat s'ha tingut en compte la perspectiva de gènere, fent ús del llenguatge inclusiu i evitant reforçar biaixos i estereotips.

1.5. Enfocament i mètode seguit

En primer lloc, es fa una recerca bibliogràfica sobre metodologies per estimar el risc d'incendi en una determinada àrea i en unes determinades condicions.

Tot seguit, partint de dades obertes es buscaran les variables més rellevants.

Amb elles s'entrenarà un model per predir el risc d'incendi per a un determinat dia.

Es publicaran els mapes online a la plataforma CARTO.

En una segona part, partint de la zona aproximada de 2 grans incendis del 2022 i utilitzant l'API del Sentinel HUB s'obtindrà l'àrea exacta (resolució de 20 metres) i la severitat del incendi utilitzant l'índex NBR. A partir de la severitat i amb el mapa de variables biofísiques de l'arbrat de Catalunya (ICGC i CREAM)[4] es quantificarà el carboni emès i la biomassa perduda.

En quant el mètode a seguir, considero que per la naturalesa del projecte el més adequat és el de cascada ja que atorga un gran control sobre les diferents fases del projecte ideal per al TFM ja que hi ha un calendari estricte per entregar les diferents PAC's. A més, és fàcil d'utilitzar i s'adapta bé al TFM ja que els requeriments inicials variaran poc.

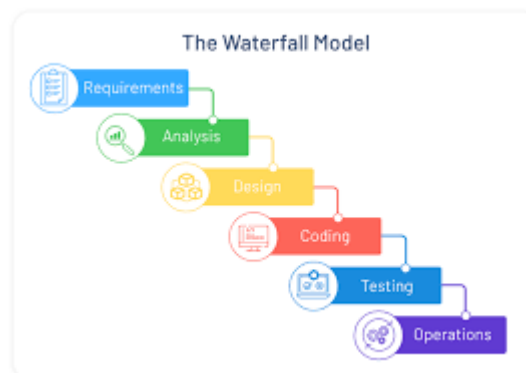


Figura 2: Metodologia en cascada. Font: [5]

1.6. Planificació del Treball

Per a realitzar el projecte es necessiten diferents recursos, tant de software com de hardware:

Software

- Spyder i Jupiter Notebooks: IDE's de desenvolupament de python.
- Llibreries Python: Les llibreries necessàries per al tractament de dades i per treballar amb informació geogràfica: numpy, pandas, matplotlib, geopandas, fiona, raster.
- Rstudio: IDE de desenvolupament de R.
- Llibreries R: Les llibreries necessàries per treballar amb informació geogràfica: raster i rgdal.
- QGIS: Programari SIG per a visualitzar i editar dades espacials.
- Powerpoint: Per a realitzar la presentació.
- CARTO: Eina SIG per a la presentació de mapes interactius.
- Github: Plataforma per allotjar projectes utilitzant el sistema de control de versions Git.

Hardware

Ordinador portàtil amb AMD Ryzen 7 4800H 2,9Ghz, 32 GB RAM DDR2 y Ubuntu 22.04

Aquest treball es desenvolupa en 6 fases diferents, corresponents a les 6 PAC's en que està organitzat el curs, en cada fase s'han identificat les tasques a realitzar:

Fase 1: Definició i planificació del treball final (Inici 28/9 – Entrega 9/10)

- Determinar la temàtica del treball final (Inici 28/09 – Final 29/09)
- Descriure de manera concisa en què consistirà el treball final (Inici 30/9 – Final 3/10)
- Raonar la motivació i la justificació del treball escollit. (Inici 3/10 – Final 6/10)
- Definir els objectius de treball (Inici 5/10 – Final 8/10)
- Determinar la metodologia que es seguirà (Inici 6/9 – Final 9/10)
- Realitzar una planificació per al desenvolupament del treball (Inici 6/9 – Final 9/10)

Fase 2: Estat de l'art o anàlisi del mercat (Inici 9/10 – Final 23/10)

- Cercar articles, investigacions que intenten resoldre la mateixa problemàtica (Inici 9/10 – Final 19/10)
- Cercar i seleccionar el software que s'utilitzarà. (Inici 19/10 – Final 23/10)
- Escollir les fonts de dades que s'utilitzaran. (Inici 19/10 – Final 23/10)

Fase 3: Disseny i implementació del treball (Inici 24/10 – Final 23/12)

- Pre-processar les dades (Inici 24/10 – Final 30/10)
- Trobar les variables rellevants (Inici 31/10 – Final 6/10)
- Entrenar la xarxa neuronal (Inici 7/10 – Final 20/10)
- Elaborar el mapa de risc d'incendis (Inici 21/10 – Final 24/10)
- Publicar la cartografia online (Inici 25/10 – Final 28/10)
- Obtindre l'àrea i severitat dels 2 grans incendis de l'estiu 2022 (Inici 29/10 – Final 4/12)
- Creuar les dades amb el mapa de variables biofísiques de l'arbrat (Inici 5/12 – Final 14/12)
- Obtindre les dades de carboni i biomassa emesa en els 2 grans incendis del 2022 (Inici 15/12 – Final 23/12)

Fase 4: Redacció de la memòria (Inici 26/12 – Final 15/12)

- Extreure conclusions sobre el treball realitzat (Inici 26/12 – Final 30/12)
- Proposar les futures línies de treball (Inici 31/12 – Final 4/01)
- Revisió de la memòria final a partir de la documentació creada anteriorment (Inici 4/01 – Final 15/01)

Fase 5: Presentació i defensa del projecte (Inici 16/01 – Final 22/01)

- Resumir els principals aspectes del treball en un Powerpoint (Inici 16/01 – Final 18/01)
- Elaborar la presentació en vídeo (Inici 19/01 – Final 22/01)

Fase 6: Defensa pública (Inici 23/01 – Final 3/02)

- Presentació de treball (Inici 23/01 – Final 3/02)
- Respondre les preguntes dels membres del tribunal (Inici 23/01 – Final 3/02)

A continuació es mostra, mitjançant un diagrama de Gantt, la planificació temporal d'aquestes tasques, en concordança amb les dates d'entrega de les diferents PAC's:

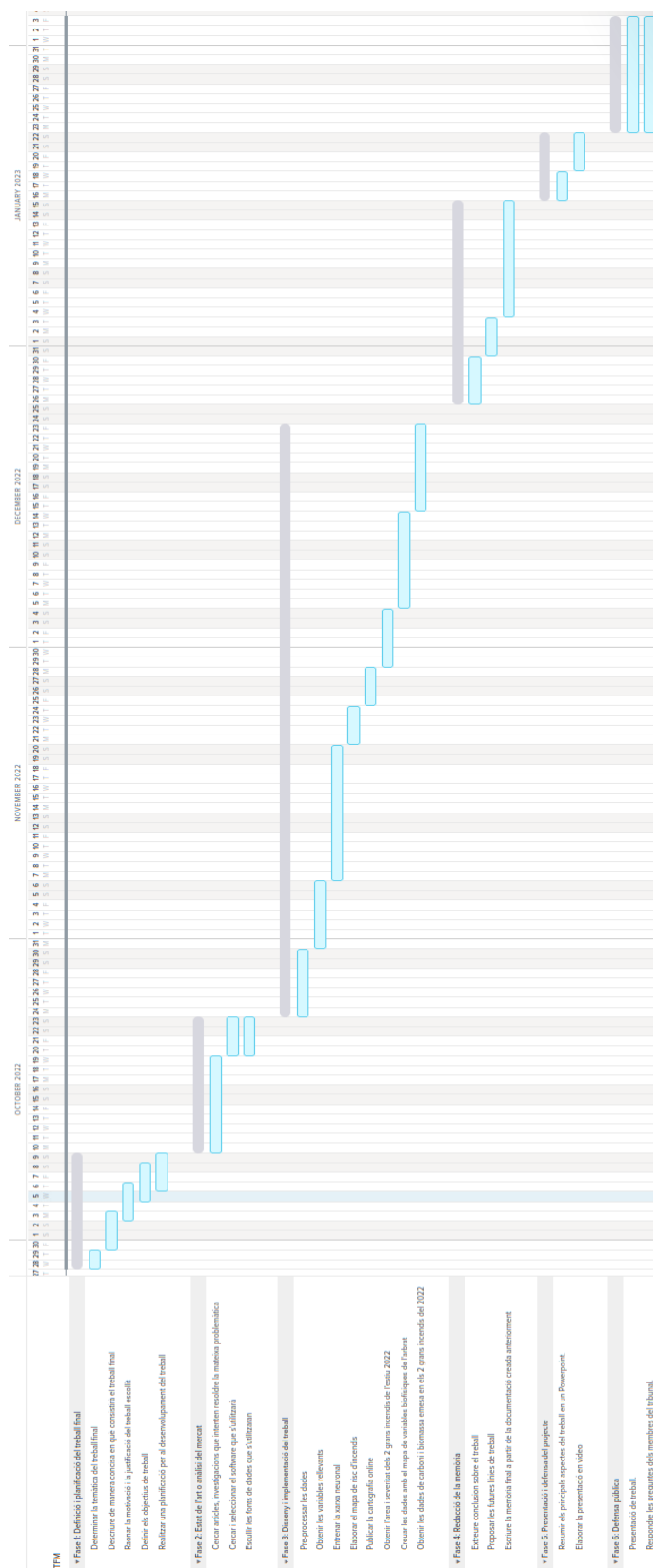


Figura 2: Diagrama de Gantt. elaboració pròpia

2. Estat de l'art

La recerca bibliogràfica per a aquest TFM s'ha realitzat al voltant de 3 punts:

A. Sobre els factors condicionants del incendi

Hi han moltes causes possibles del desencadenament d'un foc, la gran majoria per factor humà (veure taula 1), però són els factors condicionants els que determinen si la ignició inicial es converteix en un gran foc o no. Aquests factors podem agrupar-los en meteorològics, del terreny, de la vegetació, antropogènics. [6]

Causants	Detall	%
CAUSES NATURALS	LLAMP	10,3%
NEGLIGÈNCIES	CREMA AGRÍCOLA	13,3%
	CREMA DE PASTURES	1,5%
	TREBALLS FORESTALS	3,4%
	FOGUERES	4,4%
	FUMADORS	5,8%
	CREMA DE DEXAILLES	1,4%
	ABOCADORS	0,5%
	CREMA DE MATOLLAR	0,6%
	ALTRES NEGLIGÈNCIES	9,0%
	FERROCARRIL	0,7%
ACCIDENTS	LÍNIES ELÈCTRIQUES	5,6%
	MOTORS I MÀQUINES	4,3%
	MANIOBRES MILITARS	0,2%
	D'ALTRES CAUSES (ACCIDENTS)	0,5%
INTENCIONATS	INTENCIONAT	25,3%
CAUSA DESCONEGUDA	CAUSA DESCONEGUDA	12,2%
REVIFATS	REVIFATS	1,0%

Taula 1: Incendis forestals per causa 1995-2021. Font: Departament d'Acció Climàtica, Alimentació i Agenda Rural

Les meteorologia és altament determinant, els grans incendis es produeixen en situacions de calor i de sequera[7]. El vent aporta l'oxigen necessari per a la combustió. Generalment s'utilitzen 5 variables: Temperatura, humitat relativa, velocitat i direcció del vent i precipitacions.

La sequera d'una massa forestal es pot estimar utilitzant l'índex Drought Code [8]. Al País Valencià el 50% de tots els incendis es produeixen a partir del percentil 70 d'aquest índex, i tots els incendis de més de 5000 hectàrees sempre es produeixen per damunt del percentil 75 [9].

La vegetació representa el combustible disponible per al foc. En aquest sentit, alguns tipus de formacions vegetals tenen més facilitat d'ignició que d'altres.

En quant als factors de terreny generalment es consideren l'altitud, la orientació i el pendent [10]. En l'hemisferi nord, les superfícies orientades a sud, al rebre major llum solar presenten humitats més baixes i temperatures més elevades, fet que facilita la ignició. L'altitud també afecta directament a la temperatura així com a la vegetació.

En quan a factors antropogènics, diversos estudis com el de Zhang [11] consideren la distància a nuclis habitats i la distància a carreteres. Donat que més del 80% dels incendis estan directament relacionats amb l'activitat humana, els llocs amb poca presència humana presenten menys incendis.

B. Sobre models de predicció del risc

Pel que fa a la qüestió del desenvolupament de models de predicció en l'àmbit dels incendis d'incendi forestal, hi ha diverses aproximacions. Des de mètodes físics a mètodes estadístics i d'aprenentatge automàtic. En comparació als tradicionals mètodes qualitatius i d'anàlisi estadístic, el mètodes d'aprenentatge automàtic han demostrat la seva habilitat per proporcionar millors resultats per a la predicció espacial d'incendis forestals. [12]

En quant a l'aprenentatge automàtic, en relació als incendis s'ha utilitzat en 6 dominis [6]:

1. Caracterització del combustible, detecció de foc i mapeig
2. Temps de foc i canvi climàtic
3. Ocurrencia de foc, susceptibilitat i risc
4. Predicció del comportament del foc
5. Efectes del foc
6. Gestió del foc

Fixant-nos en la susceptibilitat o risc, entès com la probabilitat espacial o densitat de l'ocurrencia de foc, l'aproximació general és la construcció d'un model utilitzant dades de foc teledetectat amb una combinació de variables del paisatge, clima, estructurals, antropogèniques com a variables exploratòries.

Molts d'aquests treballs han comparat els models d'aprenentatge automàtic amb models d'altres tipus i, en general, els d'aprenentatge automàtic han sigut superiors. Entre els algorismes més utilitzats trobem les Màquines de Suport Vectorial, les Xarxes Neuronals, el Random Forest i el de Màxima Entropia.

Cal destacar el treball de Zhang [11] que compara una xarxa neuronal convolucional (CNN), el Random forest (RF), les màquines de suport vectorial (SVM), una xarxa neuronal de perceptró multicapa (MLP) i la regressió logística de kernel (KLR) per predir la susceptibilitat de foc forestal a la província de Yunnan, a la Xina. Els autors van trobar que CNN millorava a tots els altres algorismes amb una precisió del 87,92 comparada amb RF (84.36%), SVM (80.04%), MLP (78.47%), KLR (81.23%). Anoten que el benefici de la CNN es que incorpora correlacions espacials, per tant pot aprendre característiques espacials. Com a punt negatiu destaquen que els algorismes de xarxes neuronal són difícils d'interpretar en comparació amb els altres.

En una publicació més recent [13], els mateixos autors exploren 4 tipus de xarxes neuronals, de tipus convolucional amb 1 i 2 capes; i de tipus perceptró multicapa també amb 1 i 2 capes, i les testegen en la predicció del risc per a 4 diferents estacions de l'any. La màxima precisió l'ha assolit, de manera significant, la Xarxa convolucional de 2 capes, millorant a les altres 3.

D'entre tots aquest mètodes per a l'estimació del risc, la majoria dels treballs no té en compte el factor estacional, és a dir parteixen de mitjanes anuals que obtenen com a resultat un mapa general de risc anual. La investigació de Zhang [11] agafa com a referència els mesos de primavera, que es quan succeeixen la major part d'incendis en la seva àrea.

En el nostre cas, donat que el clima mediterrani té una estacionalitat molt marcada amb 2 estacions seques (estiu i hivern) i 2 estacions humides, els incendis es concentren en determinats mesos (veure taula 2), per tant l'objectiu és obtenir una model capaç de fer una predicció per a un dia concret amb les seves condicions meteorològiques corresponents.

Mes	%
Gener	3,79%
Febrer	8,06%
Març	10,36%
Abril	6,40%
Maig	8,88%
Juny	13,51%
Juliol	18,36%
Agost	15,18%
Setembre	6,75%
Octubre	3,90%
Novembre	2,78%
Desembre	2,03%

Taula 2: Incendis forestals per mesos 1995-2021. Font: Departament d'Acció climàtica, Alimentació i Agenda Rural

El recent treball de Xavier Ricci [14] és una novetat en aquest àmbit, ja que crea un model de predicció diària utilitzant una xarxa neuronal convolucional en l'àrea de Califòrnia assolint el 95% d'encerts.

C. Sobre mesurament de la severitat i càlcul de la biomassa perduda i emissions

En quant a mètodes per estimar la severitat del foc amb eines de teledetecció, des de fa uns anys el més establert és el que utilitza el Normalized burnt ratio (NBR)[15] definint com a:

$$NBR = (NIR - SWIR) / (NIR + SWIR)$$

On NIR (Near Infrared) correspon a la reflectància de la vegetació en l'espectre electromagnètic de l'infraroig proper. Aquesta reflectivitat és elevada en la coberta vegetal activa; i SWIR (Short Wave Infrared) correspon a la reflectància de la vegetació en l'espectre electromagnètic de l'infraroig d'ona curta. Presenta valors elevats en terrenys secs i superfícies sense coberta vegetal i presenta valors baixos en zones humides, vegetació i neu.

Així, la combinació d'aquestes bandes mitjançant la fórmula anterior ens indicarà on s'han produït canvis en la superfície vegetal i la magnitud d'aquests.

El NBR, per tant, consisteix en un índex normalitzat amb un rang que va del -1 al 1, on la vegetació activa tindrà valors positius per a zones amb vegetació i valors propers a 0 en zones sense vegetació, nues, o cremades.

En el nostre cas, per al càlcul de l'índex s'utilitzarà les imatges del Sentinel 2 on les bandes utilitzades seran la 8 per al NIR i la 12 per al SWIR. S'escull Sentinel donat que presenta una bona resolució (20 metres), acostuma a sobrevolar la mateixa àrea cada 4-5 dies, per tant serà fàcil trobar imatges properes a les dates d'incendi. A més disposa d'una llibreria Python pròpia i d'una API mitjançant el SentinelHub que facilita la cerca d'imatge vàlides i la seva descàrrega.

L'índex final que realment s'utilitzarà consisteix en un diferencial del NBR (se l'anomena dNBR) de la imatge més propera a la data d'inici de l'incendi i la imatge disponible més propera a la finalització de l'incendi.

Teòricament els valors poden anar de -2,0 a +2,0, però les dades vàlides rarament varien molt des de -0,550 fins a 1.350.

En els mapes de severitat on s'aplica el dNBR es categoritza la severitat amb els llindars de [15]:

- No cremat: menor +0.099
- Baixa gravetat: 0.100 a +0.269
- Gravat moderada-baixa: +0.270 a +0.439
- Gravat moderada-alta: de +0.440 a +0.659
- Elevada gravetat: major de +0.660

En el mapa de variables biofísiques de l'arbrat [16], tenim disponible un ràster amb els valors de biomassa en t/ha, amb una resolució de 20 metres, amb el podem conèixer la biomassa preexistent. No obstant, aquest mapa s'elabora a partir de dades LIDAR [17], i només té en compte la vegetació existent a partir dels 3 metres[18], per tant no tenim dades dels matollars ni tampoc del sotabosc existent en formacions boscoses.

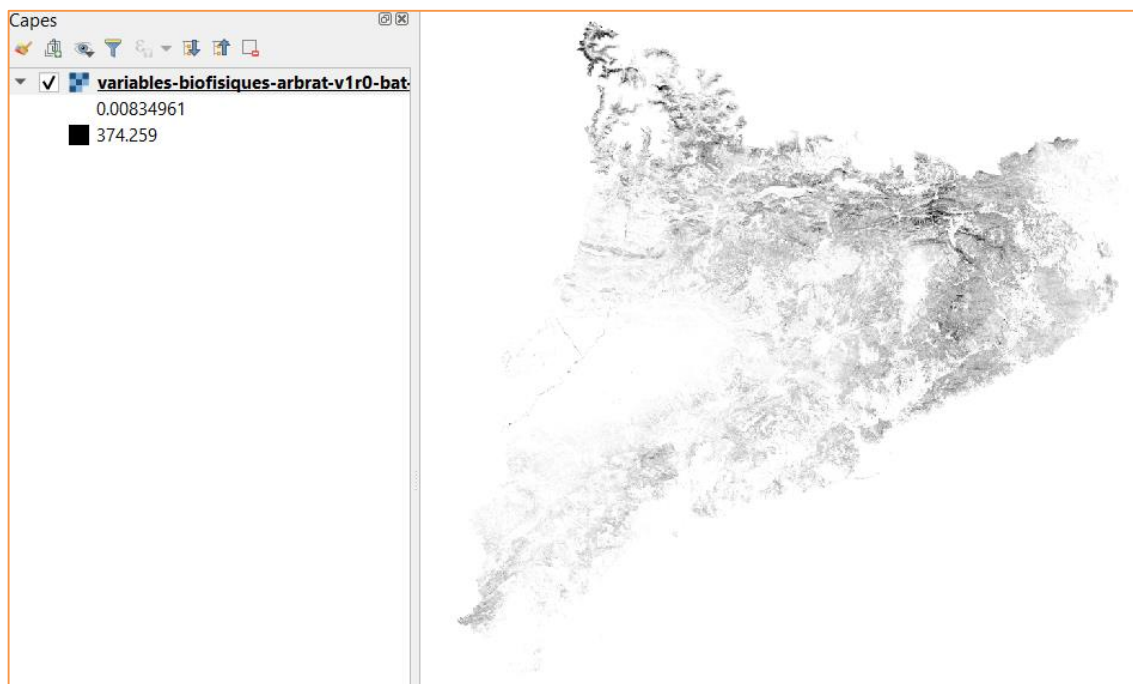


Figura 4: Mapa de variables biofísiques de l'arbrat. Biomassa. Font: ICGC-CREAF

Per al càlcul de la biomassa en matollars purs i en el sotabosc, s'ha partit del Mapa Forestal Nacional [19], que conté l'alçada mitjana de l'estrat arbustiu, la seva tipologia i la fracció de capçades coberta (FCC) . Amb aquestes dades podem aplicar les equacions al·lomètriques proposades per Montero [20], que permeten estimar la biomassa en funció de altres variables, com l'alçada, la fracció de cabuda coberta. Són les següents:

- Tanques, orles, bardes, salzedes, galeries arbustives, etc: $\ln B = 0.949 * \ln(FCC)$
- Taques, cornicabres, garrigues: $\ln B = -3.853 + 1.311 * \ln Hm + 0.726 * \ln(FCC)$
- Garrics i llentiscles: $\ln B = -2.892 + 1.505 * \ln Hm + 0.462 * \ln(FCC)$
- Landes, matollars de erinaceae i afins: $\ln B = -2.921 + 0.984 * \ln hm + 0.747 * \ln(FCC)$
- Estepa i matollars de cistaceae $\ln B = -2.596 + 0.957 * \ln Hm + 0.863 * \ln(FCC)$
- Mescla de matollars de lleguminoses retamoideas: $\ln B = -1.833 + 0.965 * \ln Hm + 0.453 * \ln(FCC)$
- Mescla de matollars de lleguminoses aulagoideas: $\ln B = -2.464 + 0.808 * \ln Hm + 0.761 * \ln(FCC)$
- Lamiàcies i timonedes: $\ln B = -1.877 + 0.643 * \ln Hm + 0.661 * \ln(FCC)$
- Espartars i pasturatges: $\ln B = 0.494 * \ln(FCC)$

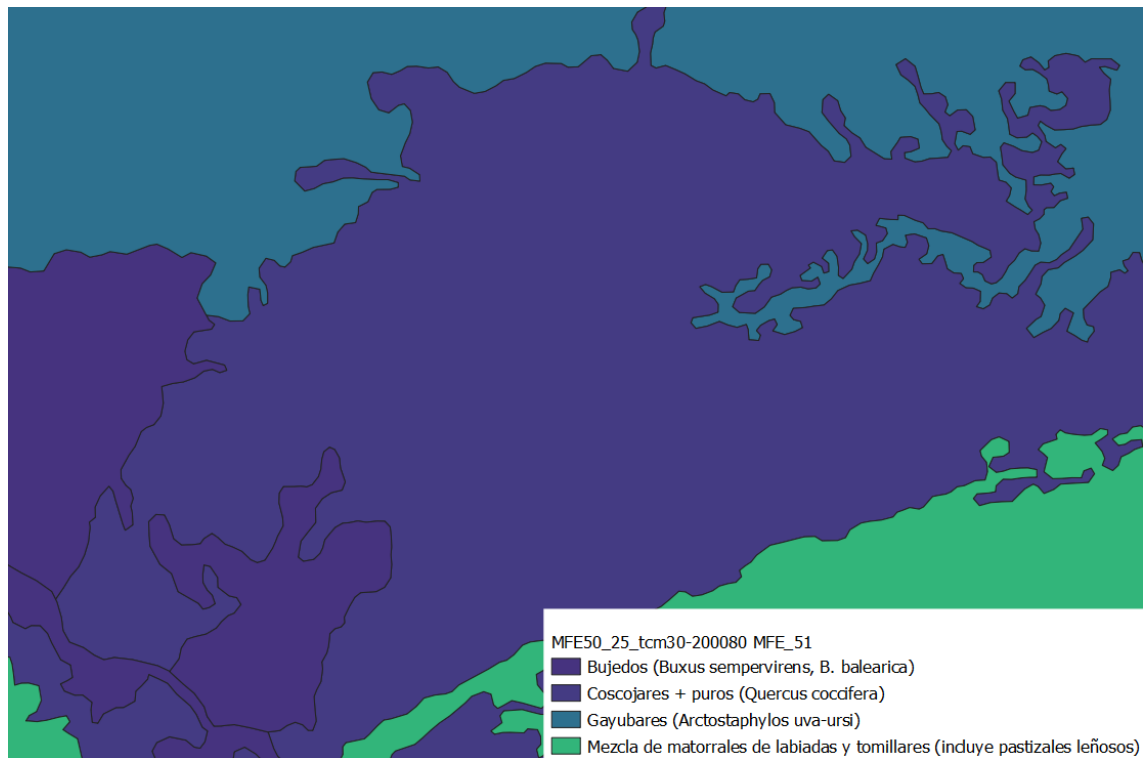


Figura 5: Mapa Forestal Espanyol. **Font:** Ministerio para la Transición Ecológica y el Reto Demográfico».

Un cop, coneixem la severitat i la biomassa preexistent, podem estimar les pèrdues de biomassa [21]. En aquest estudi es publiquen el factors de consum, per a tres tipus de vegetació (Conífera, frondosa i matollar) i 4 nivells d'intensitat:

Nivell d'intensitat	Biomassa consumida		
	Conífera	Frondosa	Matollar
Baix	0,25	0,25	0,71
Moderat	0,47	0,4	0,84
Alt	0,56	0,48	0,89
Molt Alt	0,65	0,56	0,95

Taula 3: Biomassa consumida per intensitat de foc. Font: [16]

Per al càlcul de les emissions, s'utilitzarà l'equació proposada per Wiedinmeyer:

$$Ex = A * B * BC * FEx \text{ [21]}$$

On:

- Ex: Emissions del gas en tones
- A: Àrea cremada en hectàrees
- B: Biomassa preexistent (en tones/hectàrees)
- BC: Factor de biomassa consumida
- FEx: Factors d'emissió del gas(adimensionals):

Gas	Factor
CO2	1.377
CH4	0.0014
N2O	0.00012

Taula 4: Factor d'emissió del gasos. Font: [22]

Amb aquestes dades es possible calcular les estimacions de cada gas, però per tenir una normalització a tones s'acostuma a utilitzar el CO2 equivalent[18]. Això es degut a que permet determinar el potencial d'escalfament global dels gasos d'efecte hivernacle, a partir de la seva equivalència en diòxid de carboni.

El càlcul s'obtindrà a partir de la següent equació:

$$CO2eq = Ex * PEG$$

On:

- CO2eq, és el CO2 equivalent (expressat en tones)
- Ex, són les emissions d'un gas (expressada en tones)
- PEG, el factor Potencial d'Escalfament Global (adimensional). El CH4 té una relació respecte el CO2 de PEG=21 i el N2O un PEG=298.

3. Implementació del treball

3.1. Origen de dades

La següent taula mostra un resum de les dades recollides, s'han seleccionat variables que s'han demostrat significatives en altres treballs i en funció de la seva disponibilitat. S'ha hagut de descartar la inclusió de l'índex Drought Code, que mesura la sequera, ja que no s'ha trobat cap API gratuïta que serveixi per calcular-lo amb un temps computacional assequible.

Tipus de dades	Dades	Resolució	Font
Incendis	Cartografia històric incendis	1:5000	DACC ¹
Orogràfiques	Altitud	5m	ICGC ²
	Pendent	5m	Generació pròpia
	Orientació	5m	Generació pròpia
Meteorològiques	Temperatura	0,25° x 0,25°	Clima data Store ³
	Humitat relativa	0,25° x 0,25°	
	Humitat específica	0,25° x 0,25°	
	Component u del vent	0,25° x 0,25°	
	Component v del vent	0,25° x 0,25°	
Vegetació	Cobertes del sòl	1:5000	ICGC i CREAF ⁴
	Cultius agrícoles amb origen DUN	1:5000	DACC ⁵
	Variable biofísiques de l'arbrat	1:5000	ICGC i CREAF ⁶
	Mapa Forestal Español	1:25000	MITECO ⁷

Taula 5: Dades d'origen

1. <https://agricultura.gencat.cat/ca/serveis/cartografia-sig/bases-cartografiques/boscoss/incendis-forestals/>
2. <https://www.icgc.cat/ca/Descarregues/Elevacions/Model-d-elevacions-del-terreny-de-5x5-m>
3. <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels?tab=overview>
4. <https://www.icgc.cat/Descarregues/Mapes-en-format-d-imatge/Cobertes-del-sol>
5. <https://analisi.transparenciacatalunya.cat/Medi-Rural-Pesca/Mapa-de-cultius-de-Catalunya-amb-origen-DUN-2015/wfzj-8wnc>
6. <https://www.icgc.cat/Administracio-i-empresa/Descarregues/Capes-de-geoinformacio/Mapes-de-variables-biofiques-de-l-arbrat-de-Catalunya>
7. https://www.miteco.gob.es/es/cartografia-y-sig/ide/descargas/biodiversidad/mfe_catalunia.aspx

3.2. Descripció de les dades

3.2.1 Incendis

Les dades dels incendis s'han descarregat de la cartografia que té publicada el Departament d'Acció Climàtica Alimentació i Agenda Rural.

Consta de tots els incendis forestals produïts entre 1986 i 2020. S'han descarregat tots els fitxers shapefile i s'han unit en un sol fitxer geojson.

3.2.2. Orogràfiques

Les dades orogràfiques s'han obtingut a partir del darrer model digital d'elevacions amb resolució 5 metres de l'Institut Cartogràfic i Geològic de Catalunya. A partir d'aquest fitxer, mitjançant eines del programa QGIS s'han generat els mapes de pendent i d'orientació.

3.2.3. Meteorològiques

Les dades meteorològiques s'han obtingut del reanàlisis ERA5 del model ECMWF. Aquest reanàlisis corregeix les dades del model predictiu amb observacions reals, les dades estan disponibles per hora, amb una precisió de 0,25°. A més, disposa d'una API permet fer consultes massives per geolocalització, cosa que facilita molt la feina en termes de temps i de capacitat de l'ordinador.

S'han obtingut les dades horàries a nivell de superfície, del dia de l'incendi per a cada punt d'incendi i del dia escollit aleatòriament per a cada punt de no incendi. Després s'ha calculat el valor diari màxim per al cas de la temperatura, el mitjà per a les components del vent i el mínim per a les humitats.

3.2.4. Vegetació

Les dades de vegetació s'han obtingut d'una banda del mapa de cobertes del sòl del ICGC-CREAF, que s'ha combinat amb el mapa de cultius amb origen DUN, que conté tots els cultius agrícoles declarats a Catalunya, La xarxa neuronal s'ha entrenat únicament amb aquestes dades ja que es consideren suficientment detallades per a la temàtica.

Per a la quantificació de les emissions s'ha utilitzat el Mapa de variables biofísiques de l'arbrat del ICGC-CREAF, elaborat a partir de dades LIDAR i que conté les dades de biomassa en tones per hectàrea; i també s'ha utilitzat el Mapa Forestal Espanyol, elaborat a partir de l'Inventari Forestal Nacional i que conté dades força detallades de l'estrat arbustiu.

3.3. Preparació de les dades

3.3.1. Neteja de les dades

Les dades de l'històric dels incendis a Catalunya s'han netejat amb l'script:

https://github.com/xpascuet/TFM/blob/main/01_data_cleaning.ipynb

3.3.2. Correlacions

S'han buscat si existien correlacions directes entre variables que permetessin descartar-ne alguna. Hi ha correlacions entre les humitats i la temperatura ja que la humitat relativa és la relació entre la específica i la màxima que hi podria haver a una determinada temperatura. No obstant, la correlació directa no és el suficient significativa per descartar alguna variable.

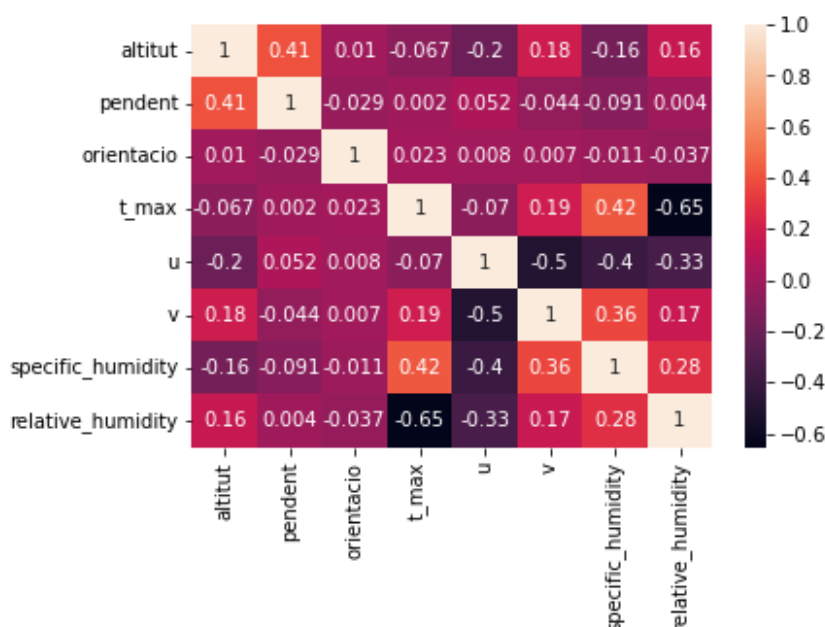


Figura 6: Matriu de correlacions

3.3.3. Reclassificació de cobertes del sòl

S'ha observat que les dades de cobertes del sòl tenen generalment un excés de resolució temàtica per a les necessitats del projecte, per exemple hi ha moltes classes diferents de zones urbanes i de masses d'aigua que no ens aporten informació rellevant al projecte.

S'ha fet una primera reclassificació seguint la taula següent:

Classe	Reclassificació
Conreus herbacis	Agrícola
Horta, vivers i conreus forçats	
Vinyes	
Oliverars	
Altres conreus llenyosos	
Conreus en transformació	
Boscus densos d'acicufolis	Boscus de coníferes
Boscus clars d'acicufolis	Boscus de caducifolis
Boscus densos de caducifolis, planifolis	
Boscus clars de caducifolis, planifolis	Boscus d'esclerofil·les
Boscus densos d'esclerofil·les i laurifolis	
Boscus clars d'esclerofil·les i laurifolis	Matollar
Matollar	
Prats i herbassars	
Bosc de ribera	Bosc de ribera
Sòl nu forestal	Sòl nu
Zones cremades	
Roquissars i congestes	
Platges	
Casc urbà	Urbà
Eixample	
Zones urbanes laxes	
Edificacions aïllades en l'espai rural	
Àrees residencials aïllades	
Zones verdes	
Zones industrials, comercials i/o de serveis	
Zones esportives i de lleure	
Zones d'extracció minera i/o abocadors	
Zones en transformació	
Xarxa viària	
Sòl nu urbà	
Zones aeroportuàries	
Xarxa ferroviària	
Zones portuàries	
Zones humides	Aigua
Embassaments	
Llacs i llacunes	
Cursos d'aigua	
Basses	
Canals artificials	
Mar	

Taula 6: Reclassificació cobertes del sòl

S'ha fet una primera generació dels mapes predictius amb aquesta classificació i s'ha observat que hi ha un elevat risc d'incendi a la zona del pla de Lleida, fet que contrasta amb els pocs incendis presents històricament en aquesta zona. Un exemple d'aquesta primera predicció pel 15 de juny de 2022 està publicat a:
<https://xpasquet.carto.com/builder/91a231b8-f5f4-48e8-923d-4a7a09bb7361/embed>

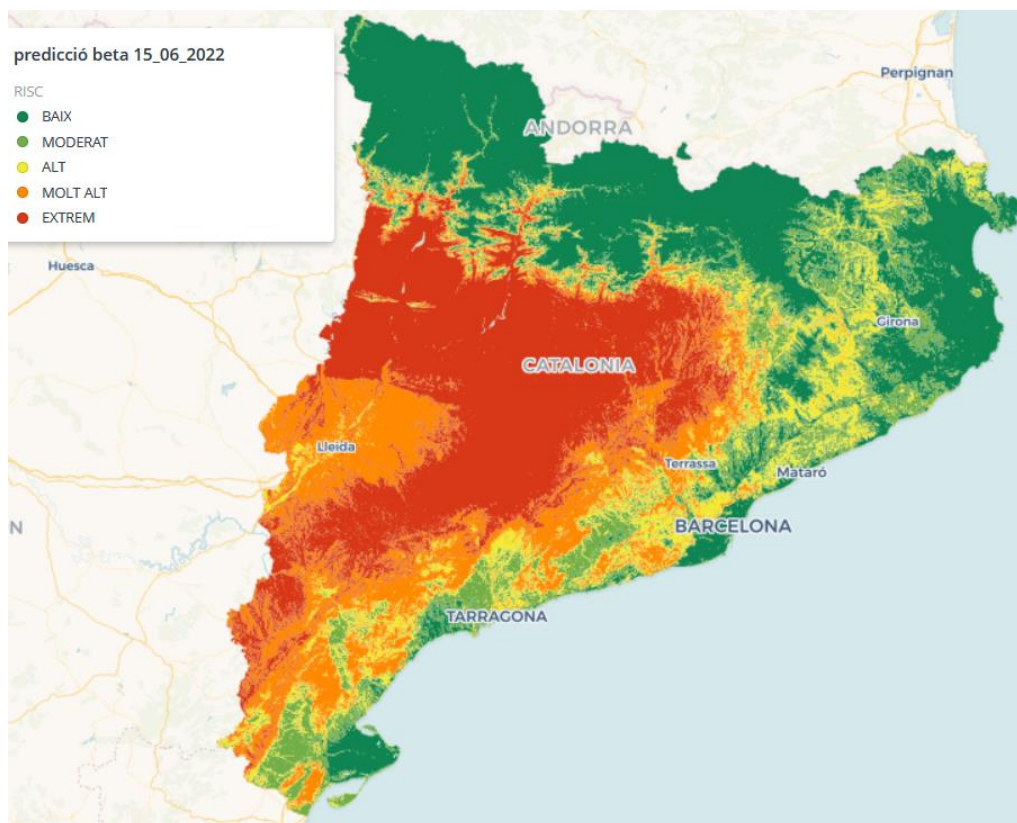


Figura 7: Primera predicció pel 15 de juny de 2022

S'ha cregut necessari fer una nova reclassificació de agrícola a agrícola seca o agrícola regadiu, ja que influeix directament en la humitat del combustible. Per a fer aquesta reclassificació s'ha encreuat el mapa de cobertes del sòl reclassificat anteriorment amb el mapa de cultius amb origen DUN, que conté informació sobre el regadiu.

3.3.4. Generació del conjunt d'exemples

Per al conjunt d'exemples s'han escollit mitjançant el programa QGIS de forma aleatòria punts que van ser foc, amb la seva data, les condicions de terreny, la coberta del sòl i seguidament s'han obtingut les condicions meteorològiques del dia.

Per als punts de no foc s'han escollit aleatòriament punts que mai han sigut foc, amb les seves condicions de terreny, la coberta del sòl i les condicions meteorològiques d'un dia aleatori.

altitud	pendent	orientacio	t_max	u	v	speci- fic_hum	rela- tive_hum	coberta_sol	fire
241,84	21,63	14,05	292,32	4,51	-3,70	0,004	27,81	5	1
212,05	34,74	308,23	302,03	0,07	0,14	0,010	43,03	6	1
870,87	31,22	96,62	286,43	0,37	-2,06	0,003	40,74	5	1
288,81	28,28	163,32	290,05	7,06	-7,39	0,005	39,11	5	1
67,97	18,17	146,64	288,04	3,17	2,50	0,006	60,96	4	1
171,14	12,20	235,42	294,52	-3,14	3,49	0,012	78,99	2	1
252,74	13,19	26,50	299,64	1,74	-0,06	0,003	13,52	1	1
189,13	6,07	175,73	299,62	1,16	1,64	0,010	46,47	1	1
818,23	31,11	65,14	306,91	0,42	0,55	0,008	29,26	5	1

Taula 7: Mostra taula d'atributs dels punts

El conjunt consta de un total de 7309 punts, dels quals 3909 són de foc i 4000 són de no foc, pel que es considera que les classes estan suficientment balancejades.

El conjunt s'ha publicat a: <https://xpascuet.carto.com/builder/f4e2eab2-151d-4e6e-ab26-65d23f9a9a01/embed>

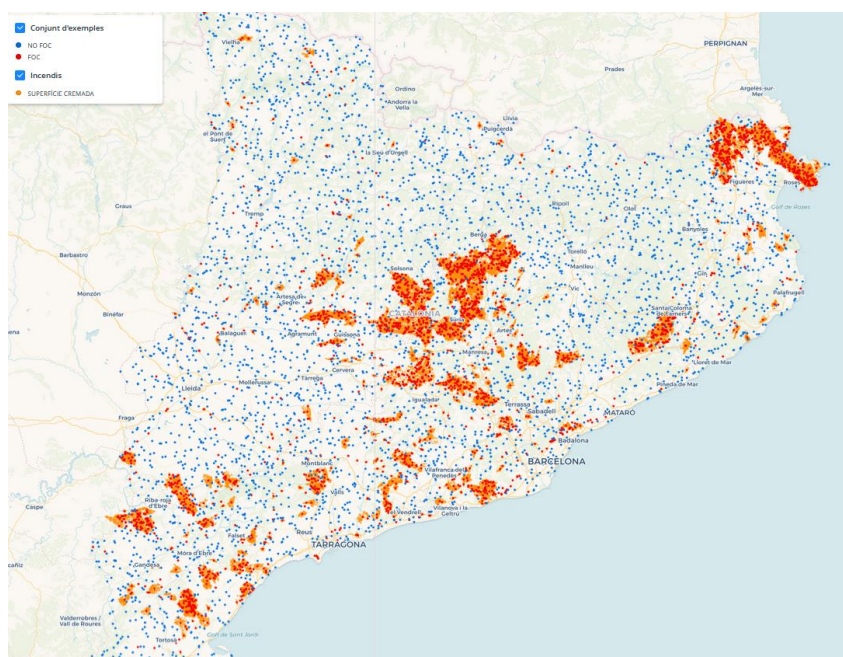


Figura 8: Conjunt d'exemples

3.4. Estadística inferencial

S'ha aplicat estadística inferencial per, a partir de les mostres de focs i de no focs inferir les dades de tota la superfície de Catalunya

Pel que fa a les variables quantitatives, s'han realitzats testos de la T de Student per tal de comparar les mitjanes de 2 mostres independents. La distribució de les variables en la majoria de dades no es realitza de forma normal, però donat que tenim una mostra gran els T-testos es consideren robustos. Primer es realitza un test de Levene per saber si existeix diferències en la variància entre les 2 mostres i per tant, si cal aplicar la correcció de Welch al T-test.

Les hipòtesis contrastades per al t-test de mostres independents són:

- H_0 : No hi ha diferències entre les mitjanes: $\mu_{foc} = \mu_{no-foc}$
- H_a : Si que hi ha diferències entre les mitjanes: $\mu_{foc} \neq \mu_{no-foc}$

S'ha escollit un nivell de significança del 95% i es busca el p-valor.

S'acompanya la informació sobre els testos amb diagrames de caixes i histogrames de distribucions.

Pel que fa a la variable categòrica s'ha realitzat el test de Khi quadrat.

Tot el procés i els gràfics s'han realitzant executant el codi:

https://github.com/xpascuet/TFM/blob/main/03_inferential_statistics.ipynb

3.4.1. Variable altitud

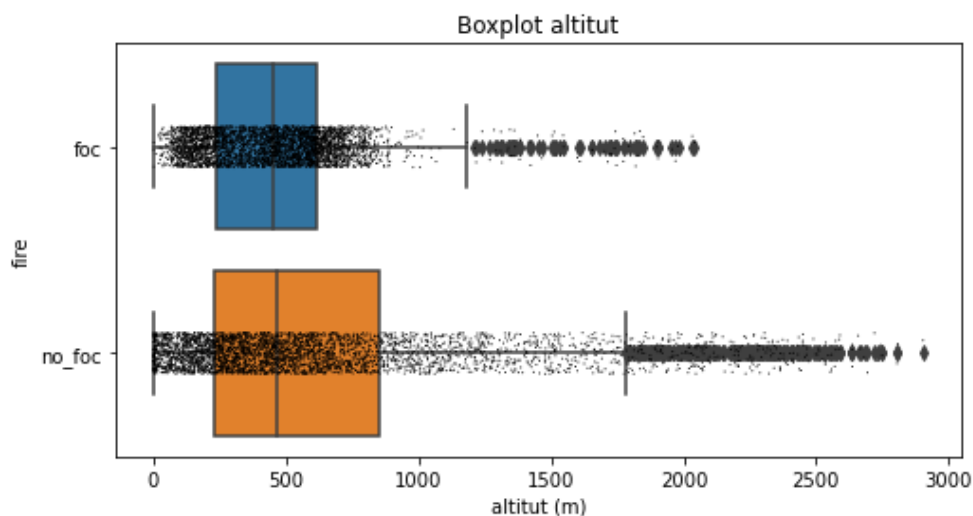


Figura 9: Diagrama de caixes altitud

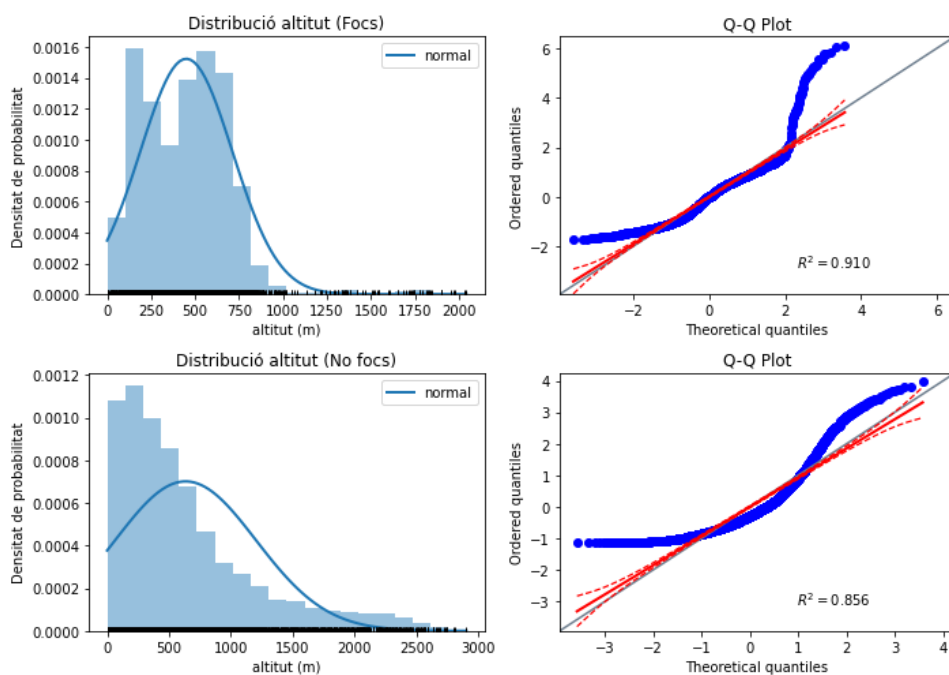


Figura 10: Histogrames altitud

Resultat T- test

El p-valor és $5,199 \times 10^{-174}$. Es rebutja la hipòtesis nul·la. Tenim evidència suficient que la variable altitud és diferent en els focs que en els no focs, amb un nivell de significança del 95%.

3.4.2. Variable pendent

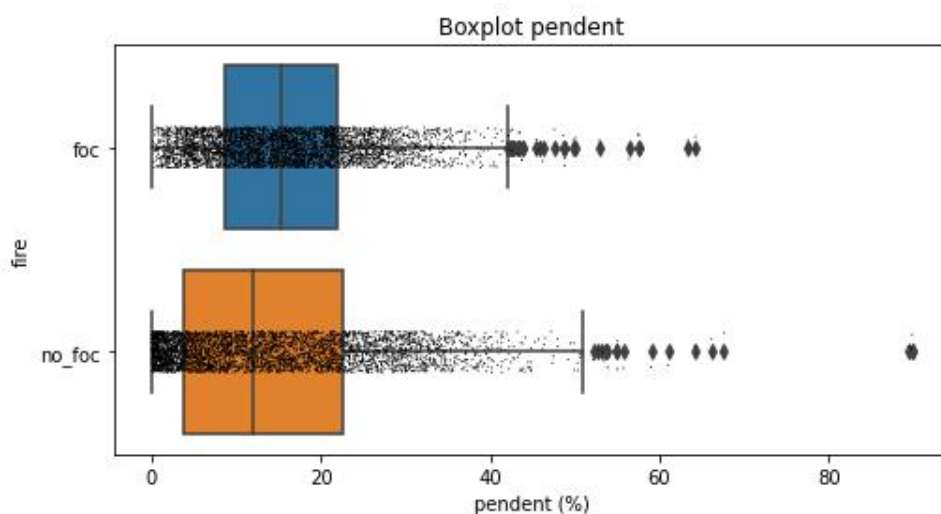


Figura 11: Diagrama de caixes pendent

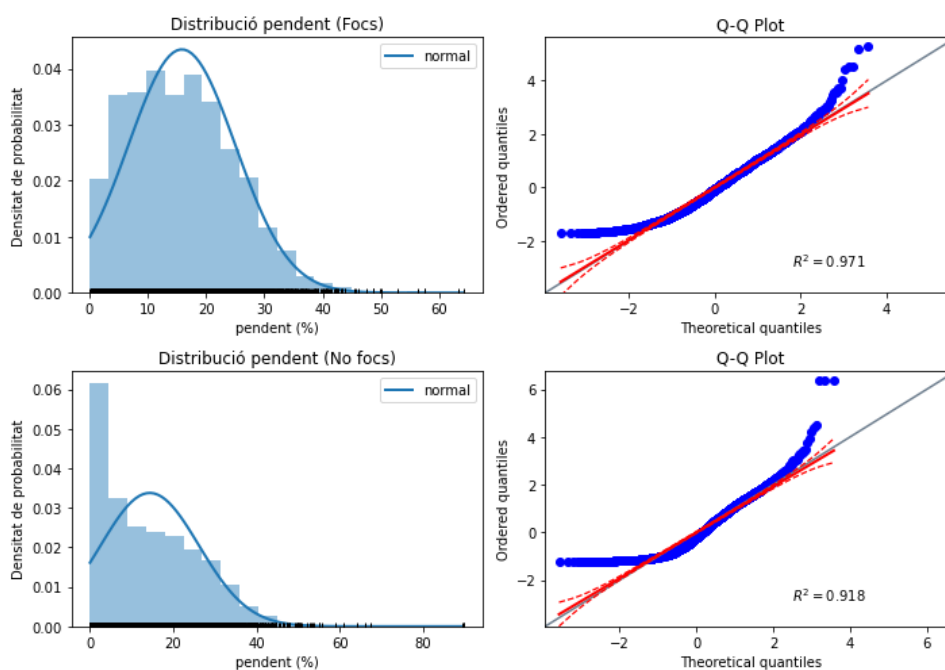


Figura 12: Histogrames pendent

Resultat T- test

El p-valor és $3,522 \times 10^{-10}$. Es rebutja la hipòtesis nul·la. Tenim evidència suficient que la variable pendent és diferent en els focs que en els no focs, amb un nivell de significança del 95%.

3.4.3. Variable orientació

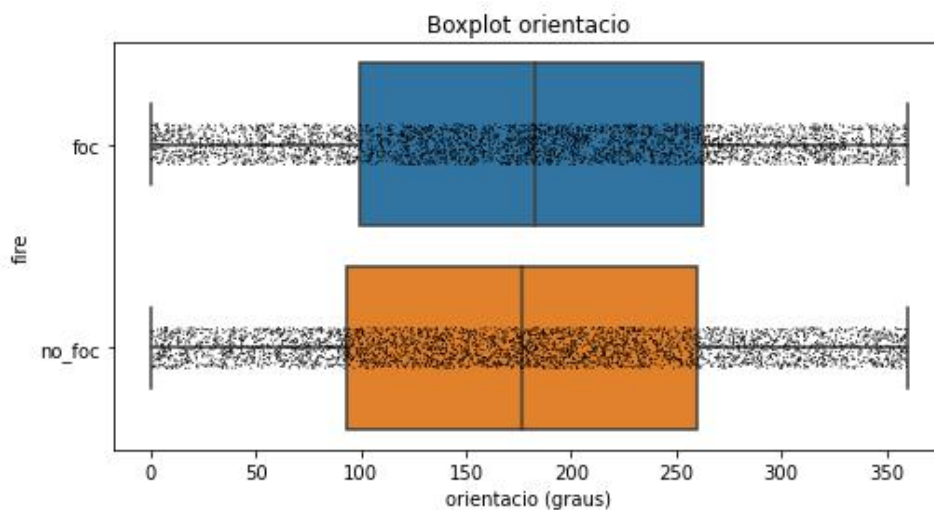


Figura 13: Diagrama de caixes orientació

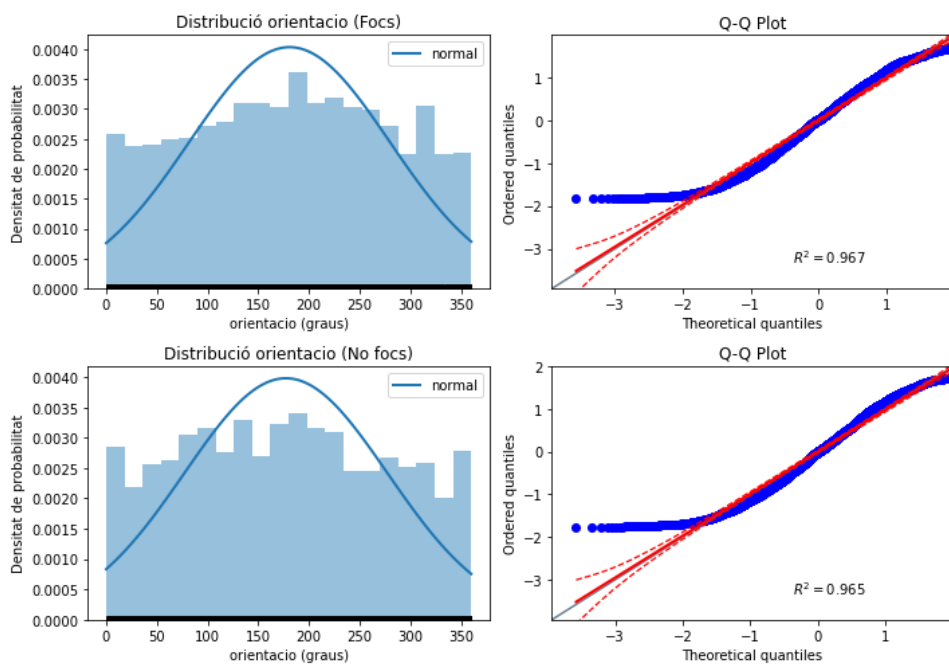


Figura 14: Histogrames orientació

Resultat T- test

El p-valor és 0,111. No es pot rebutjar la hipòtesis nul·la. No tenim evidència suficient que la variable pendent sigui diferent en els focs que en els no focs, amb un nivell de significança del 95%.

3.4.4. Variable temperatura diària màxima

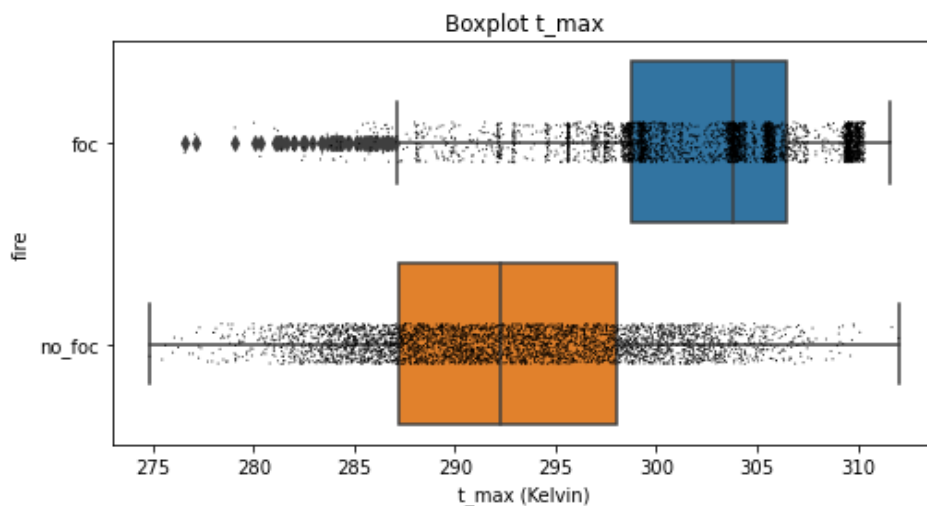


Figura 15: Diagrama de caixes, temperatura diària màxima

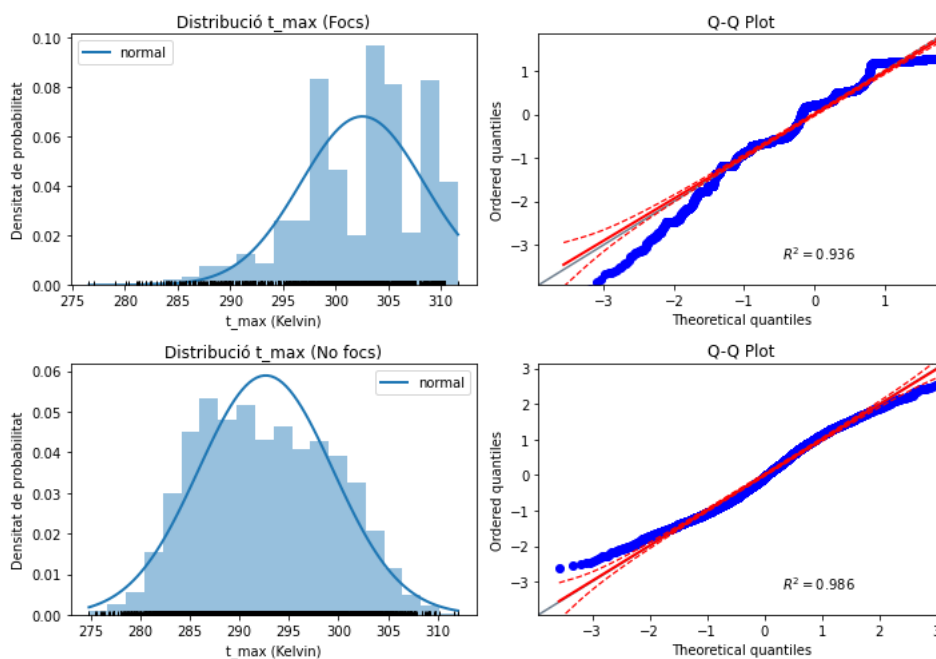


Figura 16: Histogrames temperatura diària màxima

Resultat T- test

El p-valor és 0,0. Es rebutja la hipòtesis nul·la. Tenim evidència suficient que la variable temperatura diària màxima és diferent en els focs que en els no focs, amb un nivell de significança del 95%.

3.4.5. Variable component u del vent

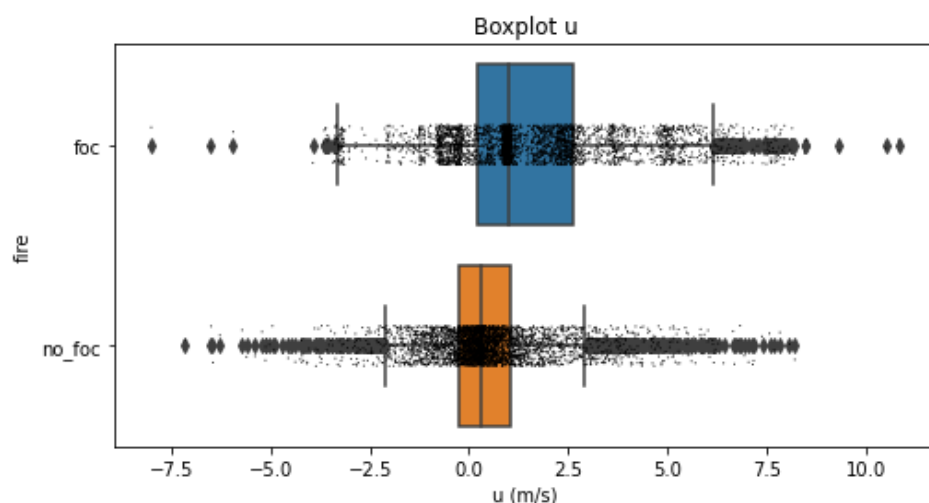


Figura 17: Diagrama de caixes, component u del vent

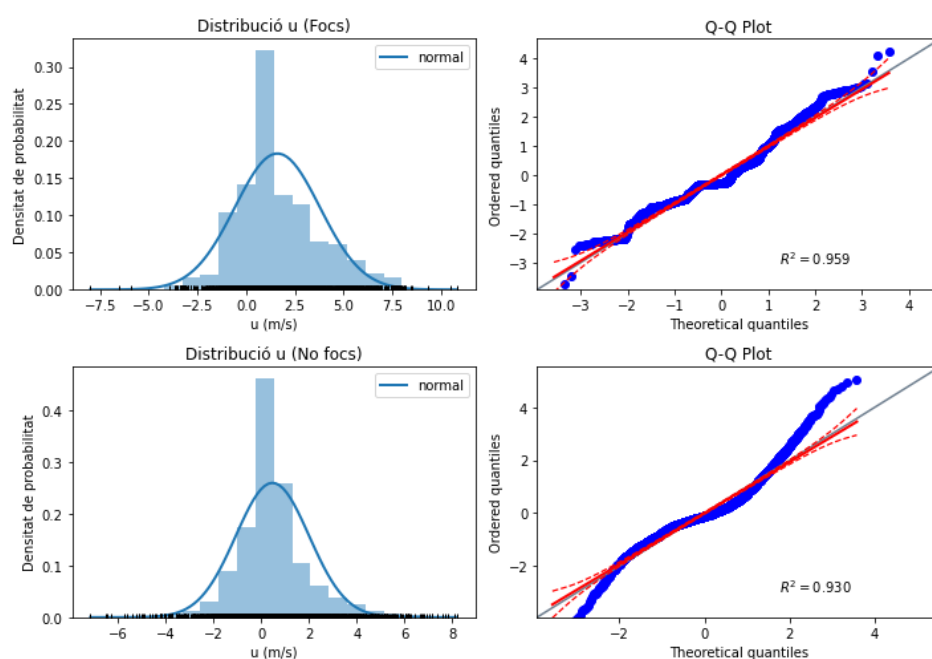


Figura 18: Histogrames component u del vent

Resultat T- test

El p-valor és 1.755×10^{-148} . Es rebutja la hipòtesis nul·la. Tenim evidència suficient que la variable component u del vent és diferent en els focs que en els no focs, amb un nivell de significança del 95%.

3.4.6. Variable component v del vent

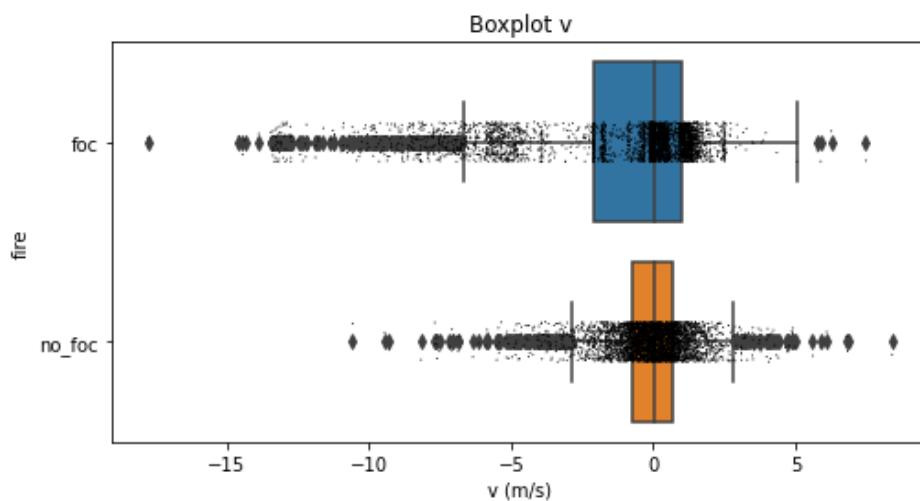


Figura 19: Diagrama de caixes component v del vent

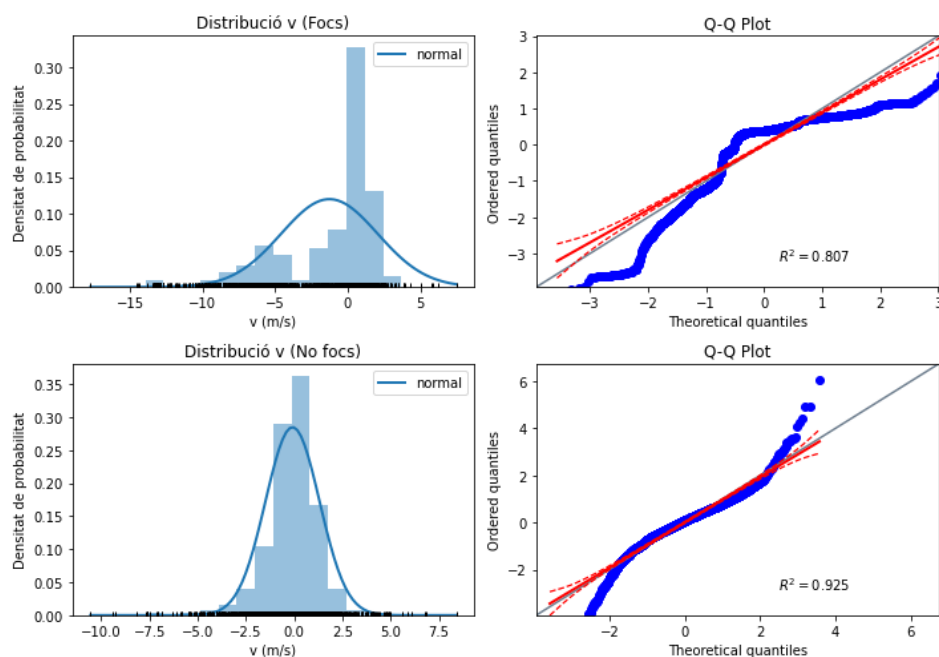


Figura 20: Histogrames component v del vent

Resultat T- test

El p-valor és 1.487×10^{-89} . Es rebutja la hipòtesis nul·la. Tenim evidència suficient que la variable component v del vent és diferent en els focs que en els no focs, amb un nivell de significança del 95%.

3.4.7. Variable humitat específica diària mínima

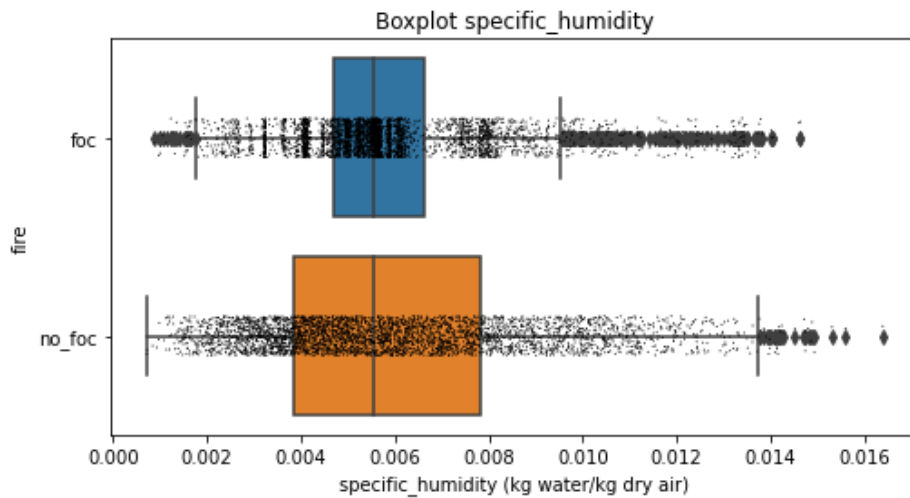


Figura 21: Diagrama de caixes humitat específica diària mínima

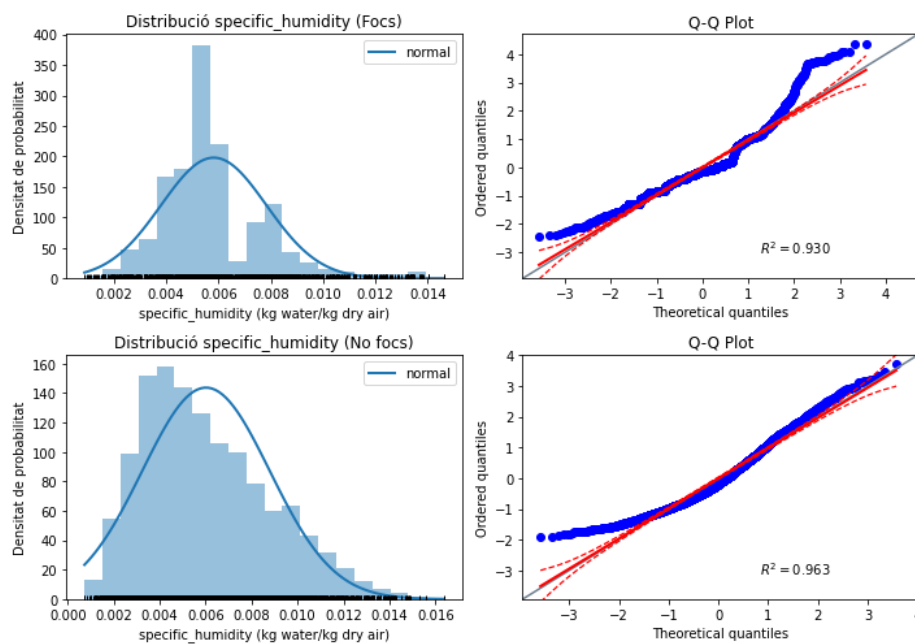


Figura 22: Histogrames humitat específica diària mínima

Resultat T- test

El p-valor és 0,00029. Es rebutja la hipòtesis nul·la. Tenim evidència suficient que la variable humitat específica diària mínima és diferent en els focs que en els no focs, amb un nivell de significança del 95%.

3.4.8. Variable humitat relativa diària mínima

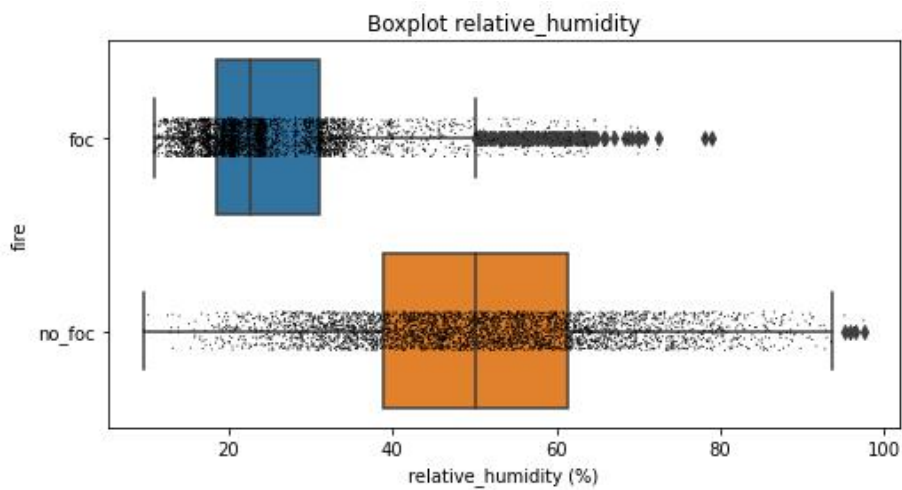


Figura 23: Diagrama de caixes humitat relativa diària mínima

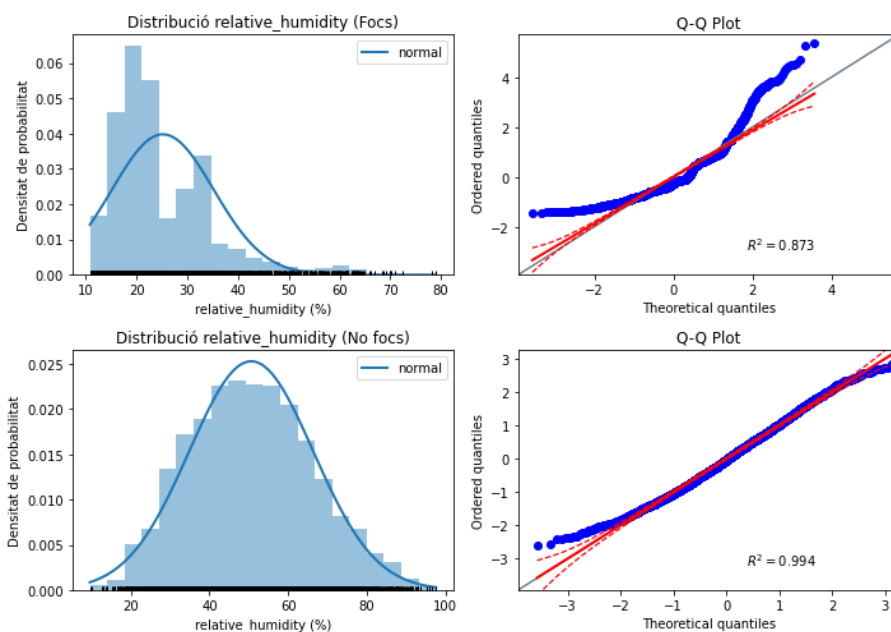


Figura 24: Diagrama de caixes humitat relativa diària mínima

Resultat T- test

El p-valor és 0.0. Es rebutja la hipòtesis nul·la. Tenim evidència suficient que la variable humitat relativa diària mínima és diferent en els focs que en els no focs, amb un nivell de significança del 95%.

3.4.9. Variable coberta del sòl

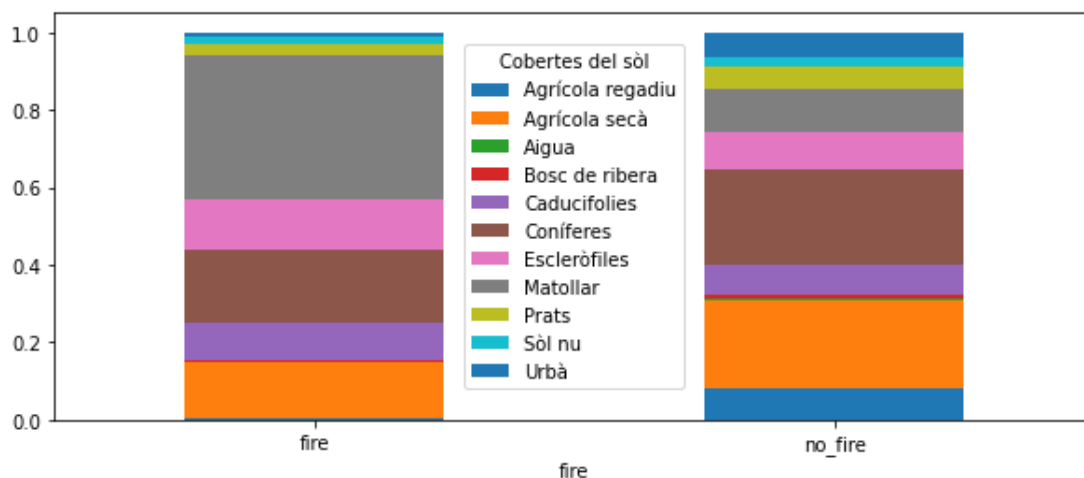


Figura 25: Diagrama de caixes, coberta del sòl

Resultat test Khi quadrat

Es rebutja la hipòtesis nul·la. Tenim evidència suficient que hi ha associació entre foc i la coberta del sòl.

3.5. Importància de variables

En l'apartat anterior hem trobat quines variables tenien diferències significatives en el seu valor mitjà en els casos d'incendis respecte els casos de no incendi, però no permet comparar la importància de cadascuna. Per tal de trobar quines són les variables més importants per a l'existència de foc, s'ha calculat una regressió logística, a partir de les variables quantitatives i s'han extret els coeficients assignats a cada variable.

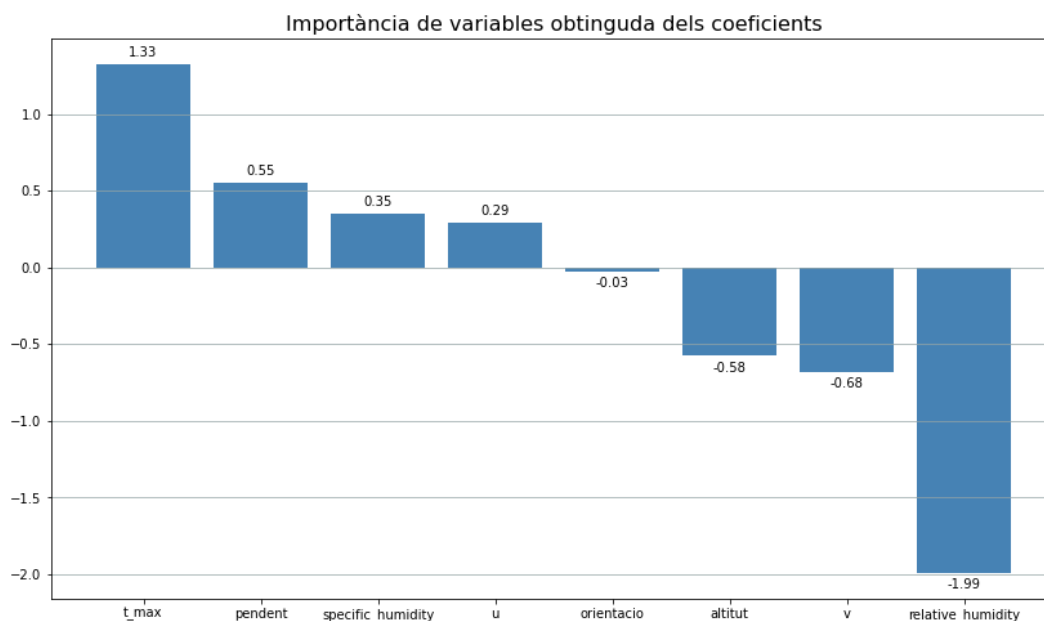


Figura 26: Diagrama de barres coeficients assignats regressió

Així veiem que els coeficients més grans són els més influents en la predicció (tant negatius com positius). Per exemple la humitat relativa té coeficient negatiu en el sentit que com més elevada és, menys probabilitat de foc hi ha.

Una manera més clara de comparar les variables és ordenar els coeficients independentment del seu signe:

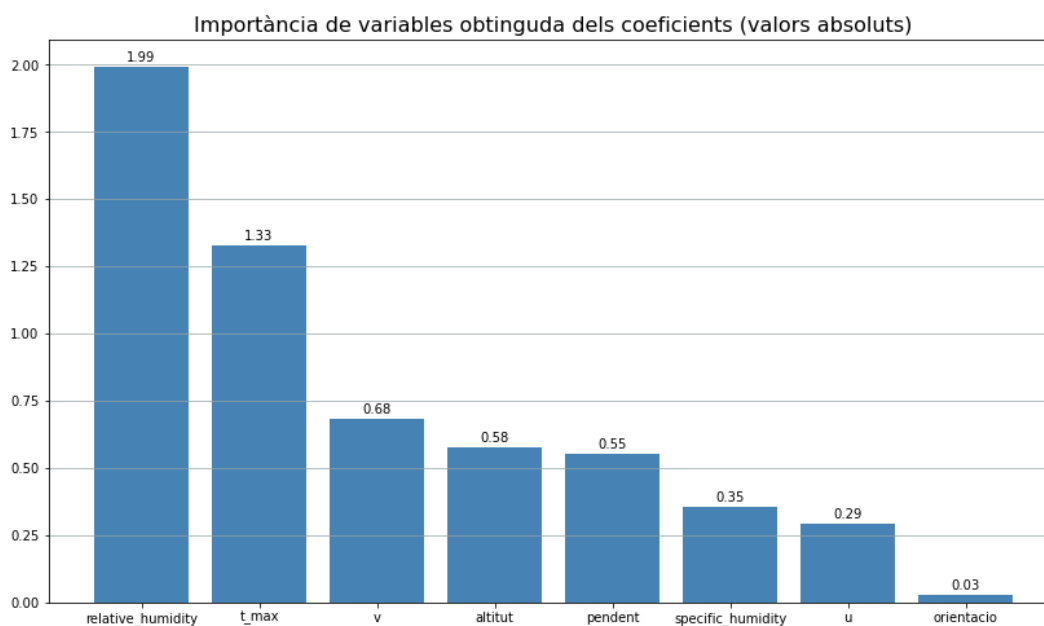


Figura 27: Diagrama de barres coeficients assignats regressió (absoluts)

S'observa com la variable més determinant és la humitat relativa, seguida de la temperatura diària màxima i de la component v del vent. També s'observa com la orientació pràcticament no és determinant.

A més de les variables qualitatives, s'ha decidit binaritzar la variable categòrica "coberta del sòl" creant una variable per cada coberta que tindrà valor 1 quan existeixi i valor 0 quan no. Així s'obté si la presència d'una determinada coberta és determinant o no:

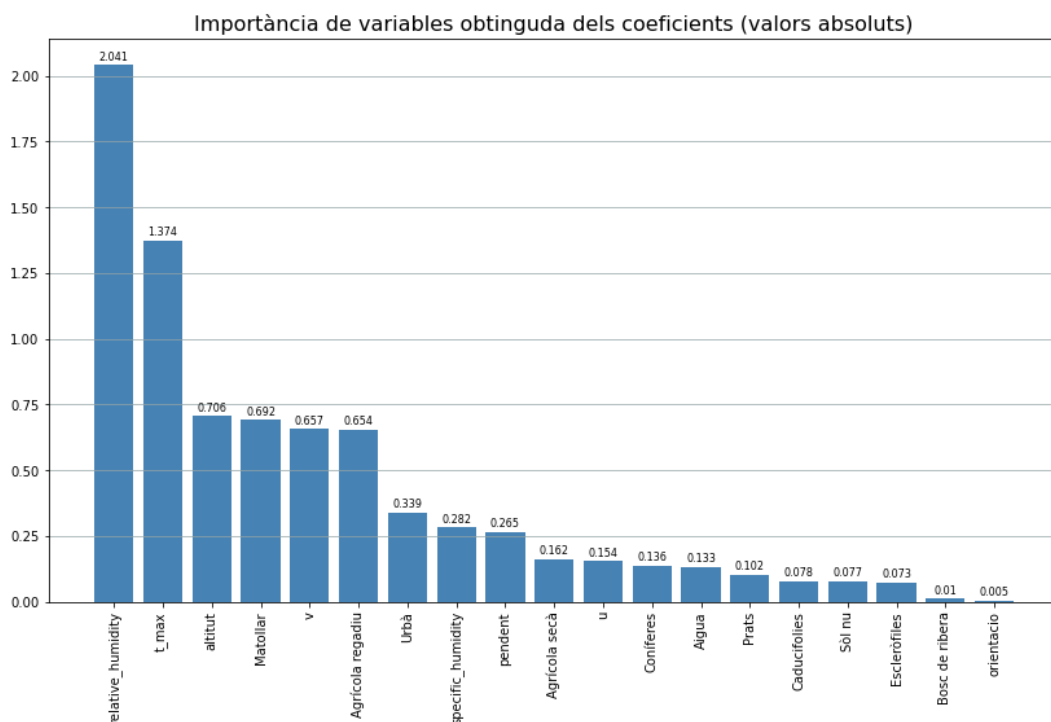


Figura 28: Diagrama de barres coeficients assignats regressió

Veiem que la presència de matollar, així com la presència de ús agrícola de regadiu, són les principals cobertes determinants.

Els càlculs corresponents i els gràfics s'han realitzat amb l'script: https://github.com/xpas-cuet/TFM/blob/main/04_feature_importance.ipynb

3.6. Creació de la xarxa neuronal

S'han normalitzat les dades, prèviament a l'entrenament de la xarxa neuronal seguint una normalització estàndard. Pel que fa a la variable categòrica "coberta del sòl", s'ha creat una variable binària per a cadascuna de les cobertes que pren valor 1 quan és present i valor 0 quan no.

Per al disseny de la xarxa neuronal s'ha escollit un nombre elevat de capes i de neurones i s'han anat reduint fins que no es produïa sobreajustament. La xarxa final té 4 capes intermitges de respectivament 128, 256, 128 i 64 neurones, a part de les capes d'entrada i de sortida.

La xarxa s'ha construït de manera seqüencial, afegint les capes una a una. S'ha utilitzat la funció d'activació ReLU per connectar totes les capes, excepte la última on s'ha utilitzat softmax, ja que és la adient quan la sortida és una distribució de probabilitats. S'han afegit drop-outs entre les capes per prevenir el sobreajustament

S'ha provat la xarxa amb diverses paràmetres de drop-out, fases d'entrenament i algorismes de optimització i restriccions del nucli. El millor resultat s'ha aconseguit amb un drop-out de 0.2, una restricció del nucli de 1, un entrenament de 50 fases i l'algoritme Adamax com a optimitzador.

El conjunt d'exemple s'ha partit en els conjunts d'entrenament i de test amb un 80 i 20 % respectivament.

Sobre les dades de test s'ha obtingut una precisió del 92,8% i una exhaustivitat del 91,8%.

Tot aquest procés s'ha realitzat amb el codi publicat a: https://github.com/xpas-cuet/TFM/blob/main/05_nn_training.ipynb

3.7. Generació de mapes per a dies concrets

S'han escollit diferents dies per generar les prediccions, intentant que n'hi hagués un en cada estació climàtica i amb diferents característiques:

- **8 de març de 2012:** Dia d'episodis simultanis d'incendis al Pirineu alguns de certa importància com Gerri de la Sal (Pallars Sobirà), Calbinyà (Alt Urgell) i Pont de Suert (Alta Ribagorça).
- **20 de setembre de 2017:** Dia humit de tardor sense incendis.
- **2 de gener de 2021:** Dia hivernal sense incendis, amb forta tramuntana al cap de creus.
- **15 de juny de 2022:** Dia d'episodis simultanis d'incendis a gairebé tot Catalunya, alguns de certa importància com el de Baldomà (Noguera), Corbera d'Ebre (Ribera d'Ebre), Castellar de la Ribera (Solsonès).

Per a generar els mapes de risc per a dies concrets, s'han descarregat els corresponents ràsters meteorològics del dataset ERA5 (que són els dependents dels dies). Llavors conjuntament amb les dades orogràfiques i de vegetació (que acceptem com a estàtiques), per a cada dia s'ha creat un ràster multibanda on cada banda és una de les variables, amb la mateixa resolució i abast. La resolució escollida són uns 0,0005° (píxels de 40 x 55 metres aproximadament), que és un punt intermitg entre les resolucions més baixes, que corresponen a les dades meteorològiques (0,25°) i les resolucions més elevades, que corresponen a les dades orogràfiques (5m). A més, el hardware disponible pot treballar bé a aquesta resolució.

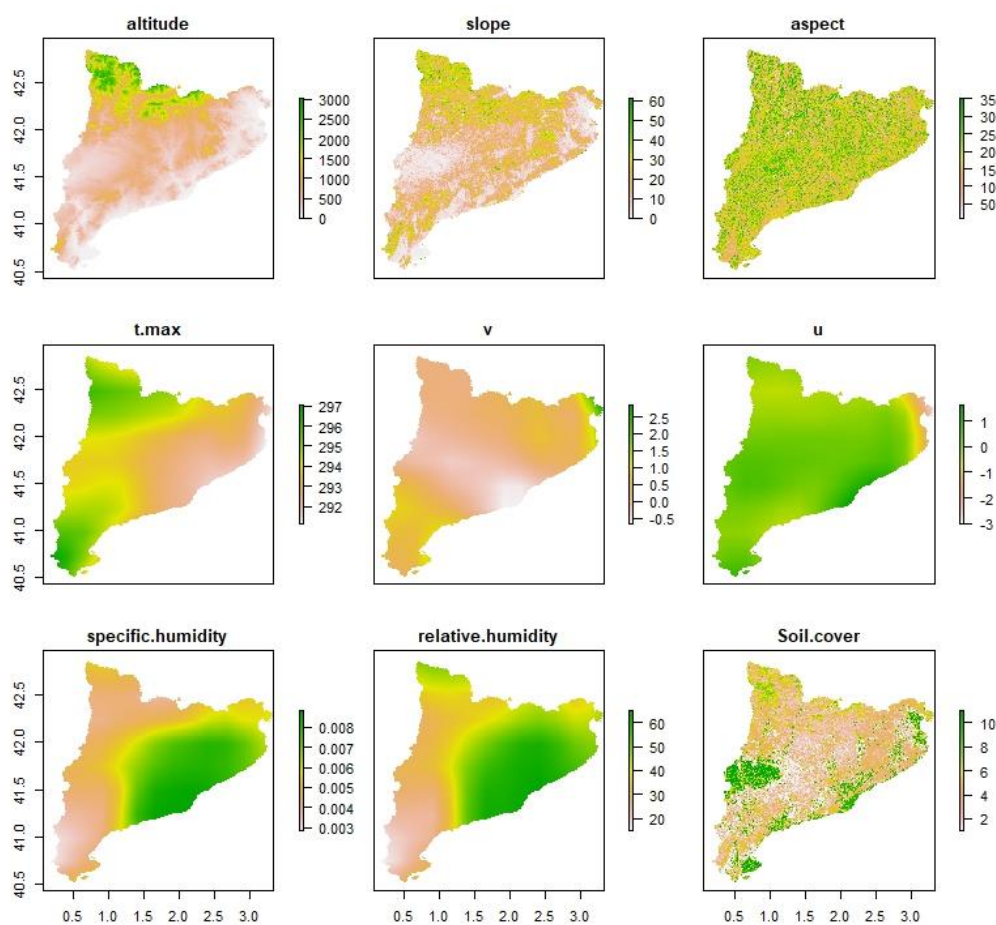


Figura 29: Exemple de les 9 bandes del ràster per al 20/09/2027

Un cop obtingut aquest ràster, es llegeix com si fos una matriu de dades, on cada fila es un píxel. Tot seguit es normalitzen les variables amb els mateixos valors que s'han fet servir per entrenar la xarxa neuronal i es generen les prediccions amb la xarxa entrenada prèviament. El resultat s'exporta també en format ràster.

Tot aquest procés s'ha realitzat amb el codi publicat a: https://github.com/xpas-cuet/TFM/blob/main/06_maps_creation.ipynb

En aquest primer resultat cada píxel té un valor de probabilitat d'incendi entre 0 i 1, s'ha reclassificat en 5 intervals de 0,2 cadascun resultant en 5 classes, ordenades de menor a major: Baix, Moderat, Alt, Molt Alt i Extrem.

Per a publicar els mapes on-line, CARTO no permet el format ràster i limita la mida de les dades, per tant s'ha hagut de reduir la dimensionalitat, deixant-la en uns 0,002^o (píxels de 160 x 220 metres aproximadament) i transformar-los a un format geojson amb l'script R: https://github.com/xpascuet/TFM/blob/main/07_poligonize.R

Els mapes resultant s'han publicat online a: <https://xpascuet.carto.com/builder/fea2609f-09e0-4afe-a0c6-daf4e3f6c828/embed>

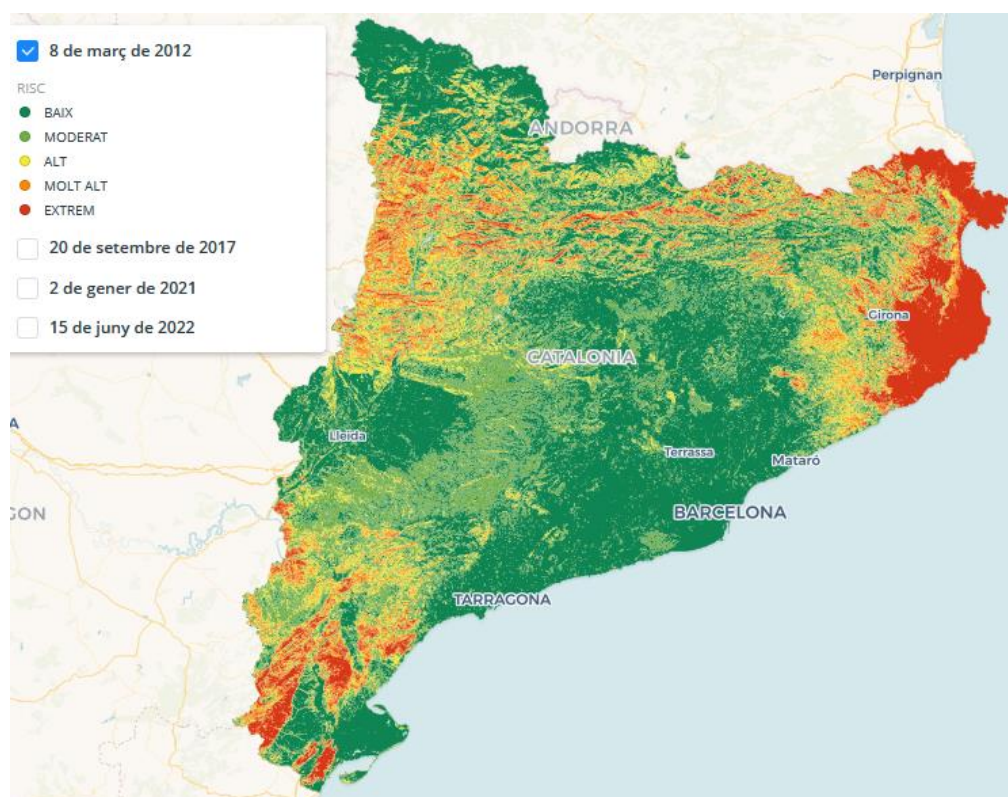


Figura 30: Predicció risc d'incendi 8/03/2012

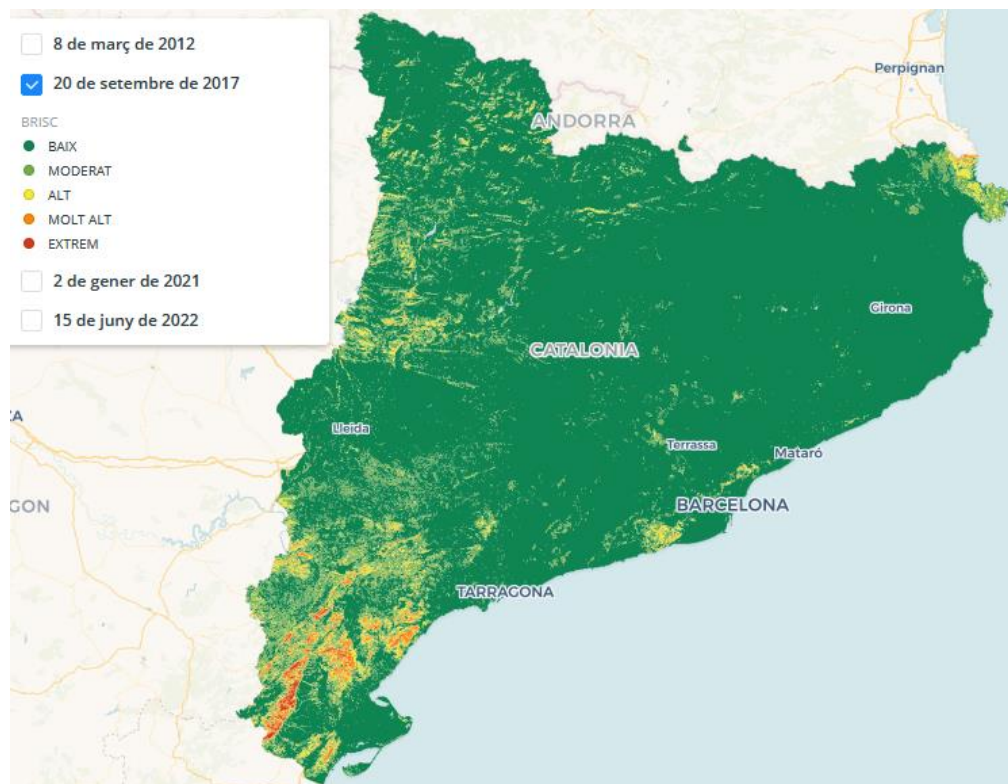


Figura 31: Predicció risc d'incendi 20/09/2017

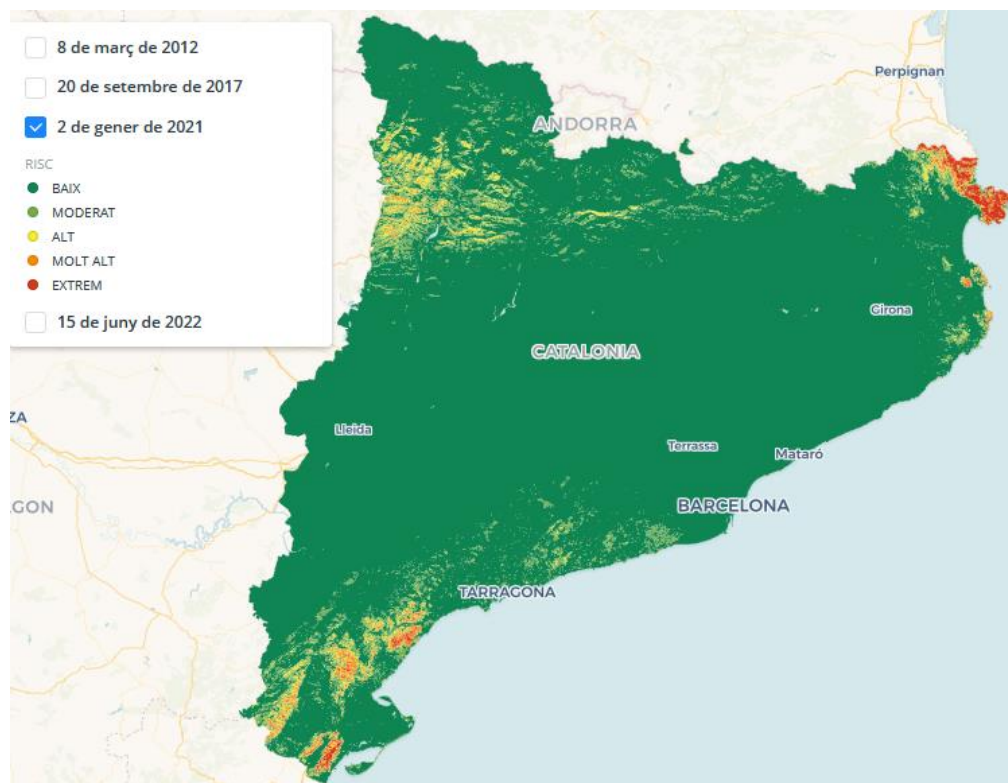


Figura 32: Predicció risc d'incendi 2/01/2021

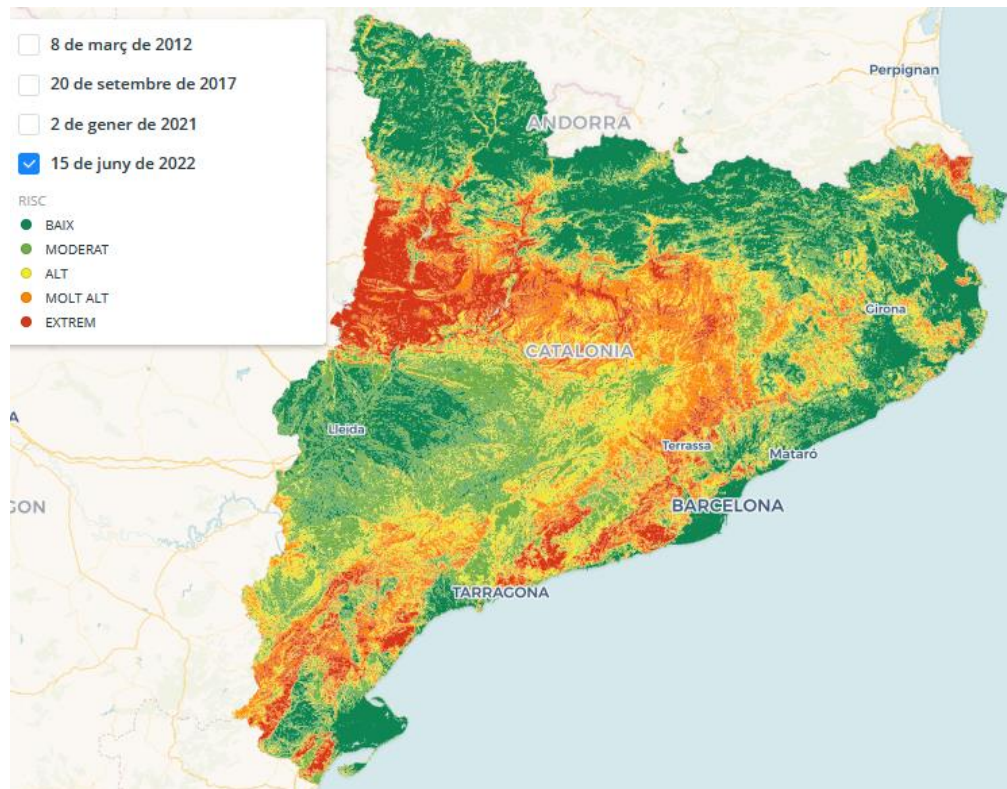


Figura 33: Predicció risc d'incendi 15/06/2022

3.8. Anàlisi dels Grans Incendis forestals del 2022

S'han analitzat els 2 majors incendis del 2022: El del Baldomà i el del Pont de Vilomara

Partint de la superfície aproximada en forma de rectangle on es va produir l'incendi i les dates d'inici i d'extinció, s'han utilitzat les API's del SentinelHUB, primer l'statisticalAPI per saber quines eren les imatges pre i post incendi disponibles més properes i després l'API estàndard per descarregar les imatges i les bandes B08 i B12 corresponents a l'infraroig proper i l'infraroig d'ona curta. Utilitzant un llindar del dNBR de 0,1 s'ha delimitat l'àrea cremada.

S'ha realitzat mitjançant l'script python disponible a: https://github.com/xpascuet/TFM/blob/main/06_get_burned_area_2022.ipynb

Un cop tenint el dNBR, mitjançant el programari QGIS; s'ha classificat la severitat i s'han retallat les capes de cobertes del sòl, biomassa de l'arbrat i el Mapa Forestal de España. S'ha fet la intersecció de les capes i s'han extret les estadístiques zonals de biomassa per a cada combinació de severitat i cobertura del sol, en els cas de la zona arbrada; i per a cada combinació de severitat, alçada de matollar, tipus de matollar i fracció de cabuda coberta en el cas de les zones amb estrat arbustiu.

Amb aquestes dades, s'ha estimat la biomassa total perduda i el carboni alliberat, mitjançant l'script: https://github.com/xpascuet/TFM/blob/main/07_biom_carbon_quant.ipynb

Els resultats numèrics estan detallats a: <https://github.com/xpascuet/TFM/tree/main/dades>

3.8.1. Incendi Baldomà 15/06/2022

1. Coníferes

Superfície = 52,38 hectàrees

Total biomassa preexistent = 1852,4 tones

Total biomassa consumida = 1060,56 tones

		CO ₂ equivalent
Total CO₂	1460,40	1460,40
Total CH₄	1,48	31,18
Total N₂O	0,13	37,93
Total emissions	1462,01	1529,50

Taula 8: Emissions (t)

Severitat		Superfície	%	Biomassa preexistent	Biomassa consumida	CO ₂ equivalent
Baixa		3,31	6,33	132,86	33,21	47,90
Moderada		6,52	12,45	249,45	117,24	169,08
Alta		15,74	30,05	505,03	282,81	407,86
Molt Alta		26,81	51,18	965,06	627,29	904,66

Taula 9: Resum superfície, biomassa preexistent i consumida i emissions per severitat

2. Frondoses

Superfície = 1170,31 hectàrees

Total biomassa preexistent = 34935,04 tones

Total biomassa consumida = 16564,37 tones

		CO ₂ equivalent
Total CO₂	22809,14	22809,14
Total CH₄	23,19	486,99
Total N₂O	1,99	592,34
Total emissions	22834,31	23888,47

Taula 10: Emissions (t)

Severitat	Superfície	%	Biomassa preexistent	Biomassa consumida	CO ₂ equivalent
Baixa	102,16	8,73	4244,89	1061,22	1530,45
Moderada	170,94	14,61	5229,88	2091,95	3016,93
Alta	416,26	35,57	10581,94	5079,33	7325,21
Molt Alta	480,95	41,10	14878,33	8331,87	12015,88

Taula 11: Resum superfície, biomassa preexistent i consumida i emissions per severitat

3. Estrat arbusti

Total biomassa preexistent = 8006,69 tones

Total biomassa consumida = 7024,69 tones

		CO ₂ equivalent
Total CO₂	9673,00	9673,00
Total CH₄	9,83	206,53
Total N₂O	0,84	251,20
Total emissions	9683,68	10130,73

Taula 12: Emissions (t)

Severitat	Superfície	%	Biomassa preexistent	Biomassa consumida	CO ₂ equivalent
Baixa	250,83	10,39	762,82	541,61	781,08
Moderada	570,95	23,65	1852,65	1556,22	2244,32
Alta	931,64	38,59	3246,63	2889,50	4167,12
Molt Alta	660,87	27,37	2144,59	2037,36	2938,21

Taula 13: Resum superfície, biomassa preexistent i consumida i emissions per severitat

Total incendi

Superfície = 2753,06 hectàrees

Total biomassa preexistent = 44794,13 tones

Total biomassa consumida = 24649,61 tones

Total CO₂ equivalent = 35548,70 tones

Mitjana CO₂ equivalent= 12,91 t/ha

Severitat	Superfície	%	Biomassa preexistent	Biomassa consumida	CO ₂ equivalent
Baixa	355,91	12,93	5140,57	1636,04	2359,43
Moderada	674,84	24,51	7331,98	3765,42	5430,33
Alta	1039,98	37,78	14333,59	8251,64	11900,19
Molt Alta	682,33	24,78	17987,99	10996,52	15858,75

Taula 14: Resum superfície, biomassa preexistent i consumida i emissions per severitat

Les dades detallades estan a: https://github.com/xpascuet/TFM/tree/main/dades/resultats_baldoma

El mapa resultant de la severitat està publicat a <https://xpascuet.carto.com/builder/d302e520-9406-4d4c-a07e-0fd0a5693adc/embed>

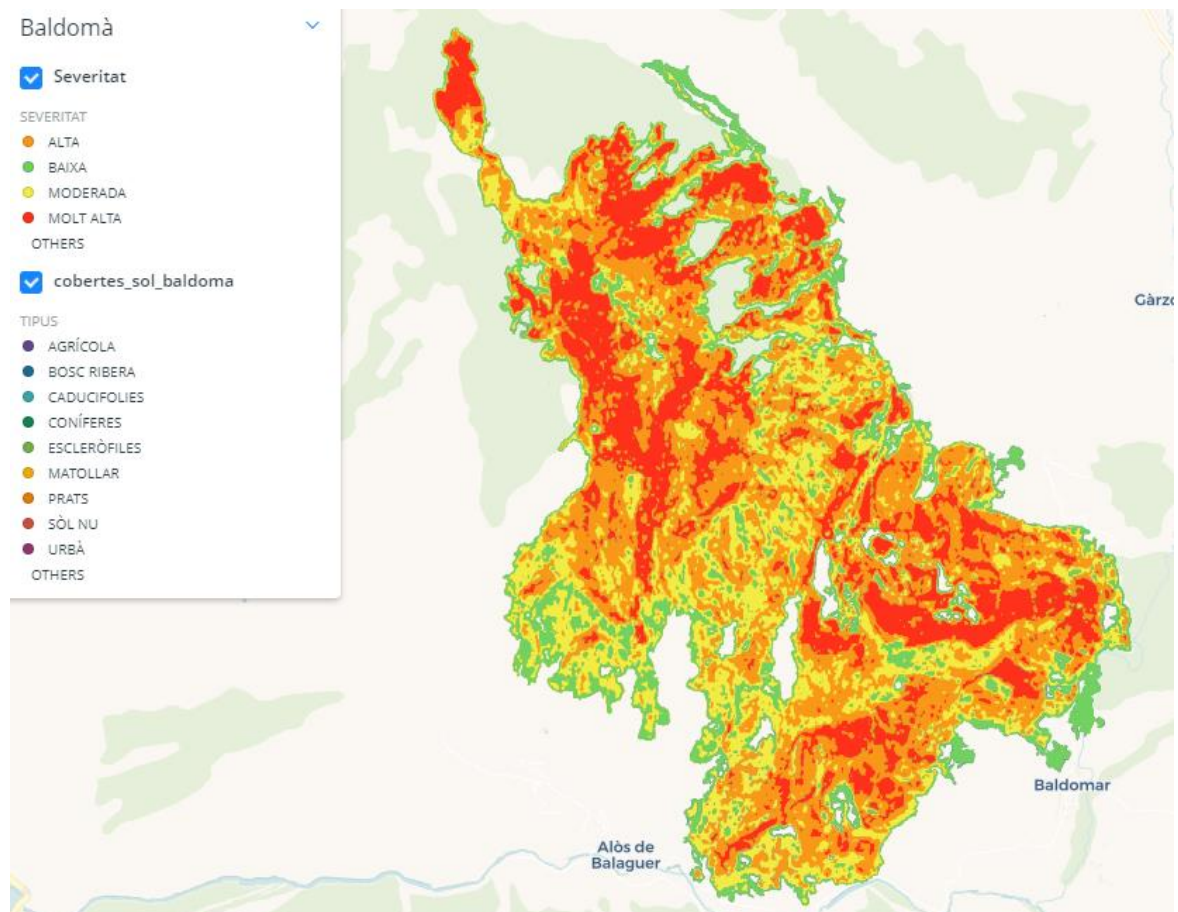


Figura 34: Mapa de Severitat Incendi de Baldomà

3.8.2. Incendi El Pont de Vilomara 17/07/2022

1. Coníferes

Superfície = 765,35 hectàrees

Total biomassa preexistent = 27053,96 tones

Total biomassa consumida = 13667,12 tones

		CO ₂ equivalent
Total CO₂	18819,62	18819,62
Total CH₄	19,13	401,81
Total N₂O	1,64	488,74
Total emissions	18840,39	19710,17

Taula 15: Emissions (t)

Severitat	Superfície	%	Biomassa preexistent	Biomassa consumida	CO ₂ equivalent
Baixa	83,35	0,80	4158,68	1039,67	1499,37
Moderada	163,73	1,58	6091,33	2862,92	4128,80
Alta	400,10	3,85	12867,18	7205,62	10391,66
Molt Alta	118,17	1,14	3936,77	2558,90	3690,35

Taula 16: Resum superfície, biomassa preexistent i consumida i emissions per severitat

2. Frondoses

Superfície = 358,20 hectàrees

Total biomassa preexistent = 13678,32 tones

Total biomassa consumida = 6590,71 tones

		CO ₂ equivalent
Total CO₂	9075,41	9075,41
Total CH₄	9,23	193,77
Total N₂O	0,79	235,68
Total emissions	9085,43	9504,86

Taula 17: Emissions (t)

Severitat	Superfície	%	Biomassa preexistent	Biomassa consumida	CO ₂ equivalent
Baixa	28,02	7,82	1247,96	311,99	449,94
Moderada	47,39	13,23	1893,78	757,51	1092,45
Alta	127,64	35,63	4740,96	2275,66	3281,87
Molt Alta	155,15	43,31	5795,63	3245,55	4680,60

Taula 18: Resum superfície, biomassa preexistent i consumida i emissions per severitat

3. Estrat arbusti

Total biomassa preexistent = 4384,45 tones

Total biomassa consumida = 3810,28 tones

		CO ₂ equivalent
Total CO₂	5246,76	5246,76
Total CH₄	5,33	112,02
Total N₂O	0,46	136,26
Total emissions	5252,55	5495,04

Taula 19: Emissions (t)

Severitat	Superfície	%	Biomassa preexistent	Biomassa consumida	CO ₂ equivalent
Baixa	151,25	11,13	567,35	402,82	580,93
Moderada	299,33	22,03	955,40	802,53	1157,38
Alta	619,06	45,55	1894,80	1686,38	2432,02
Molt Alta	289,31	21,29	966,90	918,56	1324,70

Taula 20: Resum superfície, biomassa preexistent i consumida i emissions per severitat

Total incendi

Superfície = 1.549,33 hectàrees

Total biomassa preexistent = 45.116,74 tones

Total biomassa consumida = 24.068,11 tones

Total CO₂ equivalent = 34.710,07 tones

Mitjana CO₂ equivalent = 22,40 t/ha

Severitat	Superfície	%	Biomassa preexistent	Biomassa consumida	CO ₂ equivalent
Baixa	271,14	17,50	5973,99	1754,48	2530,24
Moderada	349,16	22,54	8940,50	4422,97	6378,63
Alta	638,74	41,23	19502,94	11167,66	16105,55
Molt Alta	290,29	18,74	10699,30	6723,01	9695,65

Taula 21: Resum superfície, biomassa preexistent i consumida i emissions per severitat

Les dades detallades estan a: https://github.com/xpascuet/TFM/tree/main/dades/resultats_pont_vilomara

El mapa resultant de la severitat està publicat a: <https://xpascuet.carto.com/builder/d55b83cc-1875-44c1-9359-53b4c5cf7d6b/embed>

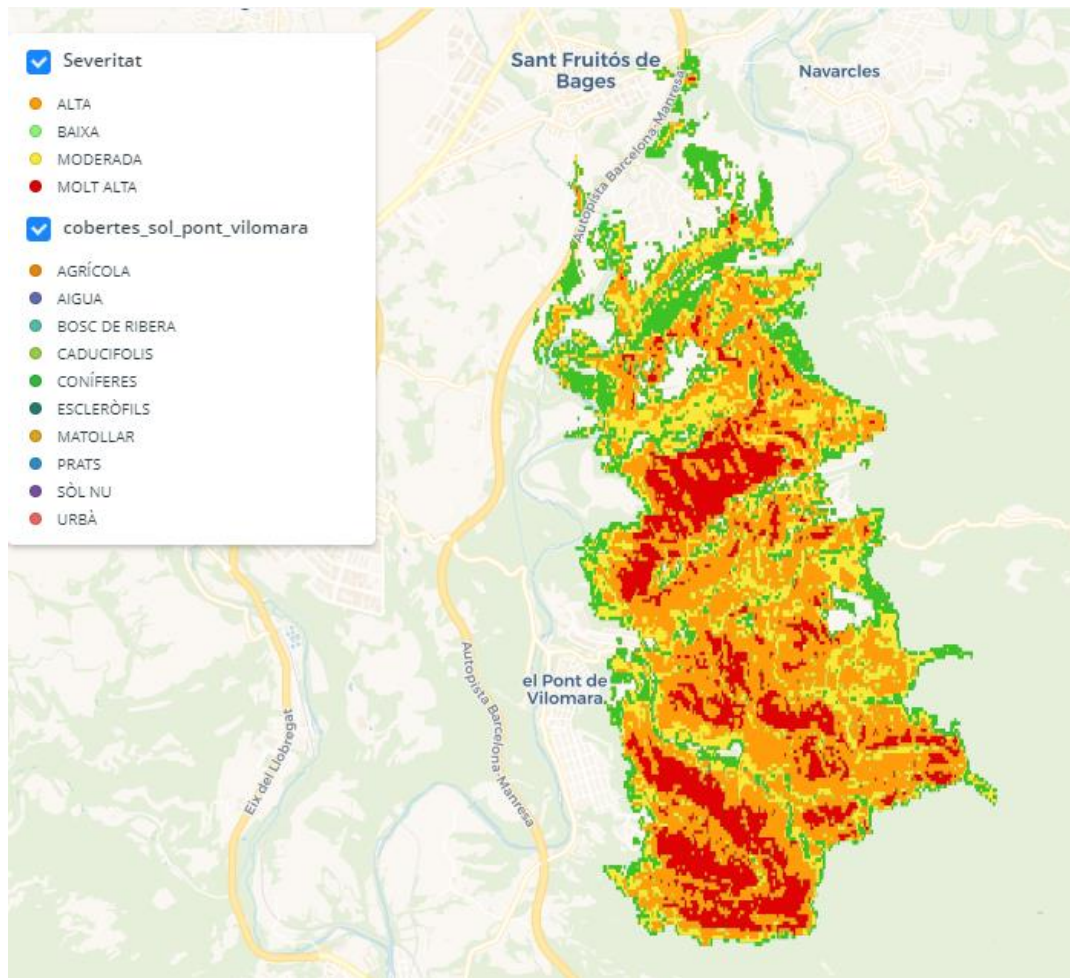


Figura 34: Mapa Severitat Incendi El Pont de Vilomara

4. Conclusions

Aquest projecte consta d'unes 3 parts més o menys diferenciades: l'obtenció dels factors determinants dels incendis, la creació d'un model predictiu del risc utilitzant xarxes neuronals i l'anàlisi dels grans incendis forestals del 2022.

Pel que fa als factors determinants, m'ha permès realitzar una anàlisi geoespacial utilitzant eines SIG. Com a resultat s'ha obtingut que la coberta del sòl, l'alçada, la humitat relativa, la humitat específica la temperatura, la velocitat del vent, i el pendent són determinats per a l'aparició d'un incendi. En canvi, la orientació no s'ha demostrat que sigui determinant.

Pel que fa a la creació d'un model productiu del risc, la conclusió és la idoneïtat dels algorismes de xarxes neuronals per a predir el risc d'incendis, el model utilitzat ha permès obtenir una predicció de forma ràpida i sense la necessitat de dades de difícil obtenció, amb un alt encert (92,8% de precisió). A més, m'ha permès implementar un visor online a la plataforma CARTO que mostra el risc dels diferents punts del mapa de Catalunya.

Aquest punt també m'ha permès prendre consciència de la importància de tenir dades de qualitat a l'hora de tenir un bon model, les primeres prediccions amb el mapa de cobertes del sòl, donaven resultats poc satisfactoris i la simple diferenciació de la coberta agrícola entre agrícola de secà o agrícola de regadiu ha millorat molt els resultats.

El hardware i el software ha estat un limitant important, els rasters d'entrada a la xarxa neuronal, sobre els que generar les previsions, s'han hagut de dimensionar a uns 0,0005° (píxels de 40 x 55 metres aproximadament) per tal que al generar les prediccions no es sobrecarregués la RAM de l'ordinador, perdent bastanta resolució pel que fa a les dades orogràfiques i de cobertes del sòl. A més, el rasters obtinguts com a resultat, s'han hagut de tornar a redimensionar a resolució de 0,002° (píxels de 160 x 220 metres aproximadament) i passar-los a un format de polígons per tal de que la plataforma CARTO permetés publicar-lo.

Pel que fa a l'anàlisi dels grans incendis forestals del 2022, s'ha comprovat la idoneïtat de les eines de teledetecció per tal de quantificar la biomassa perduda i el carboni emès en incendis forestal. De manera ràpida i utilitzat programari lliure s'ha pogut obtenir una estimació de forma ràpida i poc costosa, sense necessitat de desplaçar-se a la zona de l'incendi i realitzar inventari sobre el terreny. A més, s'ha pogut mapejar la severitat i identificar les zones que s'han cremat amb més intensitat.

5. Futures línies de treball

Pel que a les futures línies de treball, aquesta memòria podria expandir-se de formes diverses:

- Aplicar la metodologia utilitzada per a mapejar la severitat i quantificar la biomassa pèrdua i el carboni emès a altres incendis.
- Utilitzar en un futur (quan estigui processat) el vol LIDAR3 de l'ICGC, s'han utilitzat dades que provenen del vol LIDAR2, que té una densitat de 0,5 punts/m². Actualment s'ha realitzat el vol 3 amb una densitat de 5 punts/m², quan s'hagi publicat es podria quantificar la biomassa i les emissions de carboni de forma més precisa.
- Utilitzar la mateixa metodologia de partir de dades meteorològiques diàries per generar models predictius diaris en zones de clima mediterrani similar.
- Incorporar a aquest model dades de més qualitat, sobretot hi ha molt marge de millora en les dades meteorològiques. S'ha partit del Re-anàlisi ERA5 del model ECMWF, que tenen una resolució de 0,25° (uns 20 x 27 km a les nostres latituds). Aquestes dades s'han redimensionat a 0,0005°, augmentant-ne molt la resolució però d'una manera lineal, sense tenir en compte les característiques del terreny, com altituds, vents dominants, etc. El Servei Meteorològic de Catalunya disposa de la Xarxa d'Estacions Meteorològiques Automàtiques amb dades molt més precises, es podrien crear mapes continus a partir d'elles tenint en compte les característiques del terreny i obtenir unes dades meteorològiques molt més fiables.
- Incorporar alguna variable que quantifiqui la sequera, com el Drought Code [8]. Al País Valencià el 50% de tots els incendis es produeixen a partir del percentil 70 d'aquest índex, i tots els incendis de més de 5000 hectàrees sempre es produeixen per damunt del percentil 75 [9]. La idea inicial del projecte era utilitzar-lo però no es va trobar cap API gratuïta que permetés calcular-lo dins del termini del projecte.
- Sobre aquestes mateixes dades, utilitzar un model de xarxa neuronal que inclogui capes convolucionals. Aquest projecte agafa les dades ràster i les aplanava abans d'introduir-les al model, d'aquesta manera correlacions entra píxels propers no es tenen en compte. En canvi, amb xarxes convolucionals les entrades poden ser matrius bidimensionals, d'aquesta manera el model tindria en compte la relació entre píxels propers, per això actualment són l'algoritme més utilitzat en classificació d'imatges.

6. Glossari

Ràster: Imatge digital representada en un malla de caselles(píxels), cadascuna d'ells té un únic valor uniforme que representa un atribut geogràfic per aquesta unitat d'espai. Cada casella es referència per les seves coordenades x i y.

Shapefile: Format vectorial on es guarda la localització d'elements geogràfics i els atributs associats a ells, propietat e la companyia ESRI.

Geojson: Format estàndard obert dissenyat per representar elements geogràfics i els seus atributs. Basat en el format JSON.

LIDAR: Sistema làser que permet mesurar la distància entre el punt d'emissió fins a un objecte o superfície. En superfícies sense vegetació s'obté un únic retorn, en canvi en superfícies amb vegetació el sensor es capaç de registrar diversos retorns per a un mateix pols de llum, ja que es reflectit parcialment per la vegetació. Aquesta capacitat multiretorn permet descriure la estructura de la vegetació capturant informació tridimensionals dels diferents estrats i del sòl.

Component u del vent: És la velocitat horitzontal del vent, en direcció cap a l'est.

Component v del vent: És la velocitat horitzontal del vent, en direcció cap al sud.

Xarxes Neuronals Convolucionals: Tipus de xarxa neuronal que utilitza capes ocultes de tipus convolucional, caracteritzades per poder processar matrius de 2 dimensions. És l'algoritme més utilitzat per a problemes de classificació d'imatges.

Relu: Funció d'activació de capes de neurones que transforma els valors d'entrada anul·lant els negatius i deixant els positius tal com entren.

Softmax: Funció d'activació de capes de neurones que transforma les sortides a una representació en forma de probabilitats, de manera que el sumatori de totes les sortides es 1.

7. Bibliografia

- [1] Y. Liu, J. Stanturf, S. Goodrick. *Trends in global wildfire potential in a changing climate* For. Ecol. Manage., 259 (4) (Feb. 2010), pp. 685-697
- [2] Llorens, R; Sobrino, J.; Fernàndez, C.; Fernández-Alonso, J.; Vega, J. A..(2021) *A methodology to estimate forest fires burned areas and burn severity degrees using Sentinel-2 data*. Application to the October 2017 fires in the Iberian Peninsula. International Journal of Applied Earth Observations and Geoinformation
- [3] http://sac.gencat.cat/sacgencat/AppJava/organisme_fitxa.jsp?codi=12988
- [4] <https://www.icgc.cat/Administracio-i-empresa/Descarregues/Capes-de-geoinformacio/Mapes-de-variables-biofisiques-de-l-arbrat-de-Catalunya>
- [5] *How to Use the Waterfall Method in Any Project: actiTIME Guide* URL: <https://www.actitime.com/project-management/waterfall-model>
- [6] Jain, P; Coogan, S.; Subramanian S.G.; Crowley, M.; Taylor, S.; Flannigan, M.. *A review of machine learning applications in wildfire science and management*, 2020
- [7] Quílez, R. (2020) *Incendios forestales. Análisis y operaciones para su combate*. Fundación de ecología del fuego y gestión de incendios forestales Pau Costa.
- [8] Van Wagner, C.E. (1987). *Development and Structure of the Canadian Forest Fire Weather Index System*.
- [9] Quílez, R.; Van den Berg, H.; Vinué, H.; Oliver, J.; Coll, E. *Análisis histórico del Drought Code en la Comunidad Valenciana para la obtención de niveles de riesgo de incendios forestales*.
- [10] Hong, H.; Tsangaratos, P.; Ilia, I.; *Applying genetic algorithms to set the optimal combination of forest fire related variables and model forest fire susceptibility based on data mining models. The case of Dayu County, China*. Science of The Total Environment
- [11] Zhang G.; Wang M.; Liu K.. *Forest Fire Susceptibility Modeling Using a Convolutional Neural Network for Yunnan Province of China*, 2019
- [12] Bar Massada, A., A.D. Syphard, S., I. Stewart, and V.C. Radeloff. 2013. *Wildfire ignition-distribution modelling: a comparative study in the Huron–Manistee National Forest, Michigan, USA*. International Journal of Wildland Fire 22(2): 174–183.
- [13] Zhang G.; Wang M.; Liu K. *Deep neural networks for global wildfire susceptibility modeling*, Ecological Indicators, Volume 127, 2022

- [14] Ricci, X. *Creación de un modelo de predicción de riesgos de incendios forestales usando una red neuronal convolucional sobre datos históricos de meteorología de California. UOC 2022.*
- [15] Key, C. H., & Benson, N. C. (2005). Landscape assessment: ground measure of severity, the composite burn index; and remote sensing of severity, the normalized burn ratio. *FIRE-MON: Fire effects monitoring and inventory system, 2004.*
- [16] <https://www.icgc.cat/Administracio-i-empresa/Descarregues/Capes-de-geoinformacio/Mapes-de-variables-biofisiques-de-l-arbrat-de-Catalunya>
- [17] <https://www.geospatialworld.net/blogs/what-is-lidar-technology-and-how-does-it-work/>
- [18] A. Just, A; J. Gilabert, J.; Corbera, J. (2021) *DARP-ICGC Estimació de biomassa i carboni alliberat en incendis forestals (2021)*
- [19] https://www.miteco.gob.es/es/cartografia-y-sig/ide/descargas/biodiversidad/mfe_catalunia.aspx
- [20] Montero, G., Pasalodos-Tato, M., López-Senespleda, E., Onrubia, R., & Madrigal, G. (2013). *Ecuaciones para la estimación de la biomasa en matorrales y arbustados mediterráneos. 6CFE01-139.*
- [21] De Santis, A.; Asner, G. P.; Vaughan, P. J.; & Knapp, D. E. (2010). *Mapping burn severity and burning efficiency in California using simulation models and Landsat imagery. Remote Sensing of Environment, 114(7), 1535-1545.*
- [22] Wiedinmyer, C.; Quayle, B.; Geron, C.; Belote, A.; McKenzie, D.; Zhang, X.; Wynne, K. K. (2006). Estimating emissions from fires in North America for air quality modeling. *Atmospheric Environment, 40(19), 3419-3432.*

8. Annexos

El conjunt de codis utilitzats, així com els resultats dels càlculs amb detalls estan publicats a:

<https://github.com/xpascuet/TFM>