
This is the **published version** of the bachelor thesis:

Peng Lin, Chengjie. *Simulació d'incendis forestals en el cloud mitjançant Docker i Kubernetes*. Treball de Final de Grau (Universitat Autònoma de Barcelona), 2026 (Enginyeria Informàtica)

This version is available at <https://ddd.uab.cat/record/334437>

under the terms of the  license.

Simulació d'incendis forestals en el cloud mitjançant Docker i Kubernetes

Chengjie Peng Lin

30 de Juny de 2026

Resum– Els simuladors de propagació d'incendis forestals com FARSITE són deterministes i no representen la incertesa associada a la propagació real del foc. Aquest treball proposa una metodologia per transformar aquest comportament en una estimació probabilística mitjançant l'execució massiva de simulacions amb variacions controlades dels paràmetres d'entrada. L'aportació principal és el disseny d'una infraestructura cloud-native totalment automatitzada, reproducible i escalable que encapsula FARSITE i orquestra tot el cicle de vida de les simulacions al cloud. Aquesta plataforma permet executar centenars o milers de simulacions de manera paral·lela i agregar-ne els resultats en mapes de probabilitat de crema. Les proves realitzades mostren que el sistema escala correctament, genera resultats coherents amb la dinàmica de propagació del foc i constitueix una eina operativa i traçable de suport a la presa de decisions en la gestió d'incendis forestals.

Paraules clau– Incendis forestals, FARSITE, Modelització probabilística, Infraestructura cloud-native, Orquestració de sistemes, ensemble.

Abstract– Wildfire spread simulators such as FARSITE are deterministic and do not reflect the inherent uncertainty of real fire propagation. This work proposes a methodology to transform this behavior into a probabilistic estimation through the massive execution of simulations with controlled variations of the input parameters. The main contribution is the design of a fully automated, reproducible, and scalable cloud-native infrastructure that encapsulates FARSITE and orchestrates the entire simulation lifecycle in the cloud. This infrastructure enables the parallel execution of hundreds or thousands of simulations and aggregates their outputs into burn probability maps. The results show that the system scales effectively, produces outputs consistent with wildfire propagation dynamics, and constitutes an operational and traceable decision-support tool for wildfire management.

Keywords– Wildfire, FARSITE, Probabilistic modeling, Cloud-native infrastructure, Systems orchestration, ensemble.



1 INTRODUCCIÓ

ELS incendis forestals representen una de les amenaces ambientals i socials més greus a nivell mundial, i el canvi climàtic intensifica aquest risc [1]. Les condicions de sequera, les temperatures elevades i els vents forts fan que el risc d'incendis forestals sigui especialment alt en moltes regions durant els mesos d'estiu [2, 3]. El 2024 va ser el pitjor any registrat a escala global, amb almenys 13,5 milions d'hectàrees cremades [4].

Disposar d'eines que permetin predir cap a on avançarà un incendi és essencial per la gestió operativa de les emergències, decidir quines zones evacuar, on concentrar els efectius d'extinció o on establir tallafocs preventius. Aquestes decisions depenen directament de la capacitat de simular la propagació del foc amb prou fiabilitat i amb temps suficient per actuar.

Existeixen simuladors de propagació d'incendis, com FARSITE [5], dissenyats precisament per a aquesta tasca. Aquests simuladors reben com a entrada un mapa del paisatge (elevació, pendent, tipus de vegetació), les condicions meteorològiques hora a hora (vent, temperatura, humitat) i un punt o perímetre d'ignició, i produeixen com a sortida la predicció de com s'expandirà el foc al llarg del temps. Tanmateix, aquests simuladors són deterministes:

- E-mail de contacte: chengjie.peng@autonoma.cat
- Menció: Computació
- Treball tutoritzat per: Carles Carrillo i Silvana Silva Pereira
- Curs 2025/26

donat un mateix conjunt d'entrades, el simulador produeix sempre la mateixa sortida. El model no incorpora cap component aleatori, de manera que dues execucions amb els mateixos fitxers retornen exactament el mateix perímetre.

En canvi, la realitat no funciona així, el vent canvia constantment, la humitat del combustible varia localment, i les espurnes s'inicien de forma estocàstica. Una sola simulació no pot capturar aquesta incertesa.

Aquesta limitació és la que motiva el present treball. En lloc de confiar en una única predicció, que pot quedar invalidada per una mesura que ja no és representativa, es proposa executar el simulador centenars o milers de vegades en un entorn cloud amb variacions controlades dels paràmetres d'entrada i agregar-ne els resultats en un mapa de probabilitat de crema. D'aquesta manera, una sortida binària ("crema"/"no crema") es transforma en una estimació probabilística ("80 % de probabilitat de cremar"). L'objectiu no és reproduir el perímetre final un cop extingit l'incendi, sinó estimar com s'expandiria el foc en condicions naturals i quantificar la probabilitat associada a cada zona del territori. Així, els equips d'extinció disposarien d'un mapa de risc que indica cap a on tendeix a propagar-se el foc i amb quina probabilitat, de manera que puguin prioritzar la contenció amb la màxima antelació possible.

Passar d'aquesta idea a un sistema utilitzable en la pràctica no és trivial, executar de forma fiable, reproducible i econòmicament viable centenars o milers de simulacions per a cada incendi requereix una capa d'enginyeria que cap simulador monolític ofereix per si mateix. Cal contenidoritzar el simulador, provisionar un clúster de còmput efímer i a la mida del problema, orquestrar el llançament massiu en paral·lel, agregar els resultats i destruir-ho tot en acabar sense oblidar res que continuï facturant.

L'aportació central d'aquest treball és una infraestructura cloud-native completament automatitzada que materialitza l'enfocament probabilístic com a producte operatiu, i no només com a experiment puntual. La validació de la metodologia probabilística sobre casos reals serveix com a prova que aquesta infraestructura funciona a escala.

2 ESTAT DE L'ART

La simulació de la propagació d'incendis forestals és un camp amb diverses eines consolidades: FARSITE, FlamMap, Phoenix RapidFire i Prometheus, entre d'altres [6]. Totes elles modelen l'expansió del foc a partir de dades topogràfiques, de vegetació i meteorològiques, i totes comparteixen el mateix problema del determinisme. Per a aquest treball s'ha triat FARSITE com a simulador de referència, i és el sistema utilitzat en tots els experiments.

La solució al problema del determinisme no passa necessàriament per millorar el simulador en si, sinó per canviar l'enfocament computacional. En l'article de Carmel et al. [7] els autors utilitzen simulacions d'un model de propagació del foc en què diversos paràmetres d'entrada (condicions meteorològiques, localització d'ignició, data de simulació) es seleccionen aleatòriament per generar múltiples escenaris possibles; els resultats mostren que les zones amb alta freqüència de foc simulat coincideixen amb els incendis

històrics reals de la zona, demostrant que aquest enfocament pot produir mapes de risc fiables a alta resolució. Aquesta aproximació és especialment rellevant en paisatges mediterranis, on el règim de propagació és particularment sensible al vent i a la topografia [8].

Més enllà del mostreig aleatori simple, la literatura recent aborda la incertesa en simulació d'incendis des de dues grans famílies d'algorismes. La primera correspon a les tècniques evolutives, en què s'utilitzen algorismes genètics i altres estratègies metaheurístiques per ajustar els paràmetres del simulador de manera iterativa a partir de l'evolució observada del foc. Aquest enfocament tracta la simulació com un problema d'optimització inversa, on es busca minimitzar la discrepància entre el perímetre simulat i el perímetre real observat [9].

La segona família agrupa les tècniques d'assimilació de dades, que integren observacions reals del foc dins del propi procés de simulació. En aquests esquemes, el model s'executa de manera seqüencial i es corregeix iterativament mitjançant variants del filtre de Kalman, com l'*Ensemble Kalman Filter* (EnKF), introduït sobre FARSITE per Srivas et al. [10, 11]. Aquestes tècniques permeten actualitzar l'estat del simulador a mesura que es disposa de noves observacions del perímetre, com imatges de satèl·lit o dades aerotransportades.

L'aplicació operativa d'aquests enfocaments sobre FARSITE ha estat demostrada en entorns reals mostrant el seu potencial en suport a la gestió d'emergències [12, 13]. En aquest treball es segueix aquesta línia, centrant-se en un enfocament basat en mostreig aleatori amb pertorbacions modelades per distribucions normals dels paràmetres d'entrada.

3 METODOLOGIA

El treball segueix una metodologia experimental. Es parteix de la hipòtesi: l'execució massiva d'un simulador amb pertorbacions dels paràmetres d'entrada pot produir una estimació probabilística que s'ajusti millor a la realitat. A partir d'aquesta hipòtesi es dissenyen experiments per validar-la o refutar-la.

El procés experimental s'estructura en tres fases. En primer lloc, s'executa una simulació determinista amb les dades originals i es documenta la discrepància respecte al perímetre real observat. En segon lloc, s'executen N simulacions amb pertorbacions aleatòries dels paràmetres d'entrada; cada simulació produeix un mapa binari en què cada cel·la indica si el foc hi ha arribat o no. Un cop completades les N simulacions, per a cada cel·la es compta en quantes simulacions el foc hi ha arribat i es divideix pel nombre total, obtenint així una estimació de probabilitat de crema i generant el corresponent mapa de risc. En tercer lloc, es comparen els mapes resultants amb els perímetres reals observats i tot aquest procés es repeteix per a valors creixents de N (30, 50, 100, 200, 500, ...) per analitzar-ne la convergència i es documenta amb les seves condicions inicials i els paràmetres de pertorbació, de manera que sigui completament reproducible.

Els experiments es duen a terme sobre casos reals d'in-

cendis mediterranis documentats, cadascun dels quals inclou el fitxer de paisatge (.lcp), la configuració de simulació amb les sèries meteorològiques (.input) i els perímetres d'ignició i observat en format vectorial (.shp). Aquesta diversitat permet comprovar que l'enfocament no depèn d'un únic incendi i estudiar com varia el comportament del sistema segons les característiques de cada escenari.

TAULA 1: CONJUNT DE CASOS DEL DATASET UTILITZAT.

Cas	Localització	Any
Case 1	Trás-os-Montes, Portugal	2013
Case 2	Miranda do Douro, Portugal	2012
Case 6	Dalmàcia, Croàcia	2012
Case 7	Arkàdia, Grècia	2011

La Taula 1 resumeix el conjunt de casos del dataset. Al cos principal s'utilitzen el Case 2 i el Case 7 per il·lustrar els resultats, escollits per cobrir extrems oposats de la ràtio entre el perímetre d'ignició i el foc final.

Les pertorbacions del model s'apliquen sobre dos components principals: el vent (velocitat i direcció) i el perímetre d'ignició. L'origen d'aquestes dades és un element rellevant, perquè condiciona el grau de confiança associat a la simulació.

Les sèries meteorològiques que rep FARSITE provenen de registres d'estacions fixes situades a distàncies de l'ordre del quilòmetre del front actiu, o bé, en un escenari operatiu, d'un pronòstic meteorològic per a les hores immediatament posteriors al moment de la decisió. Cap de les dues fonts descriu amb precisió la dinàmica local del vent al front del foc: la mesura puntual no incorpora l'efecte del relleu, com la canalització per valls o la redirecció per vessants, ni la pertorbació atmosfèrica generada per la columna tèrmica del mateix incendi. El simulador tracta el valor d'entrada com a magnitud exacta i basa totes les decisions de propagació en aquest valor, de manera que una desviació de pocs graus en la direcció o d'una fracció en la velocitat pot desplaçar el front simulat cap a una zona diferent i alterar substancialment el perímetre predit.

La metodologia proposada actua precisament sobre aquesta vulnerabilitat. En lloc d'alimentar el simulador amb el valor mesurat com a única entrada, es genera un conjunt de variants plausibles al voltant d'aquell valor i s'executen centenars o milers de simulacions, cadascuna amb una mostra independent del conjunt. La sortida no és una predicció única, sinó un mapa de probabilitat en què cada parcel·la del terreny porta associada la fracció d'escenaris en què el foc hi ha arribat. Les zones amb probabilitat alta, considerant el conjunt de variants, constitueixen conclusions robustes respecte a la incertesa de la mesura, mentre que les que cremen només sota un subconjunt particular emergeixen com a zones sensibles, on la predicció depèn críticament del valor concret introduït. El resultat substitueix una resposta determinista per una distribució explícita de prediccions, i és precisament aquesta estructura probabilística la que aporta valor operatiu al producte final.

3.1 Pertorbació del vent

El fitxer d'entrada de FARSITE conté un bloc WIND_DATA amb un registre per a cada interval horari de la simulació. Cada registre especifica la velocitat del vent (mph) i la seva direcció en graus ($0^\circ = \text{N}$, $90^\circ = \text{E}$, $180^\circ = \text{S}$, $270^\circ = \text{W}$). Per a cada simulació es genera una pertorbació independent aplicada a cadascun d'aquests registres.

Direcció. Per a cada registre horari k , la direcció pertorbada θ'_k s'obté afegint un desplaçament δ_k mostrejat d'una distribució normal de mitjana zero:

$$\theta'_k = \theta_k + \delta_k, \quad \delta_k \sim \mathcal{N}(0, \sigma_d^2), \quad \sigma_d = 20^\circ$$

on θ_k és la direcció original registrada a l'estació i $\sigma_d = 20^\circ$. La distribució normal de mitjana zero modela la incertesa de mesura: el vent real pot anar en qualsevol direcció respecte al valor registrat, amb desviacions més probables prop de zero.

Velocitat. Per a cada registre horari k , la velocitat pertorbada v'_k s'obté multiplicant la velocitat original v_k per un factor m_k mostrejat d'una distribució log-normal de mediana unitat:

$$v'_k = v_k \cdot m_k, \quad m_k \sim \text{LogN}(0, \sigma_v^2), \quad \sigma_v = 0.20,$$

amb $\sigma_v = 0.20$. S'utilitza una distribució log-normal en lloc de normal per dos motius: garanteix que el multiplicador sigui sempre positiu (la velocitat del vent no pot ser negativa) i reflecteix que les variacions de velocitat són de naturalesa multiplicativa. Cada registre rep mostres independents, de manera que la pertorbació captura la variabilitat temporal del vent hora a hora. La llavor del generador aleatori s'inicialitza a partir del número de simulació, garantint la reproduïbilitat de qualsevol run individual [12].

3.2 Pertorbació del perímetre d'ignició

El perímetre d'ignició de la simulació també té incertesa associada: és una captura presa en un moment concret, que pot no coincidir exactament amb l'instant d'inici de la simulació, a més de l'error inevitable de digitalització de la línia sobre el mapa. Pinto et al. [13] i Allaire et al. [14] inclouen aquesta condició inicial entre els paràmetres pertorbats sobre FARSITE.

Aplicar un desplaçament fix sobre el contorn no escala correctament: una mateixa magnitud representa una alteració enorme per a un perímetre petit i imperceptible per a un de gran. Per garantir que la magnitud relativa de la pertorbació sigui consistent entre casos, s'opta per un model d'homotècia respecte al centroide del polígon. Cada vèrtex v es desplaça segons:

$$v' = C + (v - C)(1 + f), \quad f \sim \mathcal{N}(0, \sigma_{\text{ign}}^2),$$

on v' és la posició del vèrtex després de la pertorbació, C és el centroide d'àrea del polígon (calculat per la fórmula de *shoelace*) i f és el factor d'escala aleatori. La formulació és proporcional per construcció: un factor $\sigma_{\text{ign}} = 0.10$ representa un 10 % del perímetre tant si fa 100 m com si en fa 2 km, seguint la mateixa lògica multiplicativa de la velocitat del vent. El polígon s'eixampla o s'encongeix uniformement, sense distorsionar-ne la forma.

3.3 Mètriques d'avaluació

Per analitzar la convergència de l'ensemble s'empren la IoU (*Intersection over Union*) i la diferència mitjana de probabilitat per cel·la. Donades dues regions A i B , la IoU es defineix com

$$\text{IoU}(A, B) = \frac{|A \cap B|}{|A \cup B|},$$

és a dir, la proporció d'àrea compartida respecte a l'àrea total que cobreixen. Val 1 si les dues regions són idèntiques i 0 si no se solapen. La diferència mitjana de probabilitat per cel·la mesura si els valors interns deixen de canviar a mesura que s'afegeixen simulacions.

4 ARQUITECTURA DEL SISTEMA I INFRAESTRUCTURA CLOUD

L'execució massiva de simulacions admet dues vies principals: un superordinador o clúster HPC propi i una infraestructura cloud comercial. La primera ofereix més rendiment per nucli i una latència interna baixa, però implica disposar de hardware físic dedicat, gestionar-lo i sol·licitar quotes de càlcul. La segona, és accessible per a qualsevol usuari amb un compte de proveïdor, no requereix cap inversió inicial en hardware i factura només pel temps efectivament consumit, cosa que s'ajusta al patró d'ús d'aquest treball: ràfegues intenses de còmput de curta durada amb períodes llargs d'inactivitat entre experiments.

Sobre aquesta base, l'arquitectura s'estructura en tres capes: contenidorització del simulador, provisió declarativa del clúster i del magatzem, i orquestració automatitzada del cicle complet d'execució.

4.1 Infraestructura com a Codi

Tota la infraestructura de còmput i emmagatzematge es descriu com a codi (*Infrastructure as Code*) mitjançant Terraform [15], una eina per definir i provisionar recursos cloud de forma declarativa. La definició està repartida en mòduls dedicats a la xarxa, al clúster de còmput (EKS) i al magatzem d'objectes (S3). Aquesta separació permet aixecar i destruir l'entorn complet amb una sola ordre (`terraform apply` i `terraform destroy`, respectivament) i garantir que cada experiment es duu a terme sobre una infraestructura idèntica a les anteriors.

El clúster AWS EKS [16] es desplega amb un grup de nodes On-Demand de tipus `c7i-flex.large`. L'elecció d'aquesta família ve determinada per les restriccions de la capa gratuïta d'AWS: `c7i-flex.large` és la única instància de propòsit general admesa dins d'aquest marc. Amb una sol·licitud de recursos d'aproximadament mitja vCPU per simulació, el clúster pot executar fins a una desena llarga de simulacions simultànies. En la pràctica, una execució de mil simulacions sobre els casos provats es completa entre cinc i deu minuts, en funció de la durada simulada i de l'extensió del paisatge, amb un cost aproximat de menys d'un dòlar per experiment (Taula 2).

TAULA 2: DIMENSIONAMENT I TEMPS DEL PIPELINE (CASE 7).

Tipus de node	<code>c7i-flex.large</code> (2 vCPU, 4 GB)
Nodes	8
Pods concurrents	16
1000 runs (temps / cost)	~ 5 min / ~ 0,5 USD
10000 runs (temps / cost)	~ 1 h / ~ 3 USD
Pressupost experimental	< 100 USD

El dimensionament és intencionalment modest, però respon al marc pressupostari del treball i no a una limitació tècnica. Com que la simulació és paral·lela, repartir els runs sobre més nodes en redueix el temps de manera pràcticament proporcional. Amb pressupost suficient, ampliar el clúster només requereix modificar una xifra a la configuració de Terraform, sense cap canvi de disseny.

4.2 Automatització amb GitHub Actions

Executar un ensemble de mil o deu mil simulacions implica molts passos encadenats: aixecar la infraestructura, pujar les dades del cas, llançar centenars de contenidors en paral·lel, esperar que acabin, agregar els resultats i, finalment, destruir-ho tot per no acumular cost. Fer aquesta seqüència manualment és lent i propens a errors. Per això tot el procés s'ha automatitzat en un únic flux de treball de GitHub Actions, una plataforma que permet executar tasques automatitzades definides com a codi, que s'activa amb tres paràmetres (el cas, el nombre de simulacions i si cal destruir la infraestructura en acabar) i orquestra les etapes següents:

1. **Provisió de la infraestructura.** Tota la infraestructura s'aixeca amb `terraform apply`, sense cap clic manual a la consola d'AWS. En descriure-la com a codi, és reproducible i idèntica en cada execució.
2. **Pujada de les dades base.** Només es pugen a S3 els fitxers compartits del cas: el paisatge (`.lcp`), la plantilla d'entrada amb les dades meteorològiques i els perímetres d'ignició i observat. No es pugen els N fitxers d'entrada pertorbats: cada simulació genera la seva pròpia pertorbació dins del contenidor, cosa que redueix aquest pas de dues o tres hores a uns trenta segons.
3. **Llançament de les simulacions.** Es generen N *Jobs* de Kubernetes, un per simulació. Cada contenidor descarrega les dades base, aplica les pertorbacions del vent i del perímetre d'ignició, executa FARSITE i puja només la graella de resultat (`ArrivalTime`) de nou a S3.
4. **Espera amb detecció d'estancament.** El flux vigila el bucket S3 i compta quantes simulacions han dipositat el resultat. Si el recompte no avança durant diversos minuts, prossegueix amb les simulacions completades en lloc de quedar-se bloquejat indefinidament.
5. **Agregació dins del clúster.** La visualització s'executa com un *Job* dins d'EKS: llegeix totes les graelles di-

rectament des de S3 per la xarxa interna d'AWS, n'agrega la probabilitat de crema cel·la a cel·la i genera el mapa final amb els contorns al 10 %, 50 % i 90 %. Així només surt d'AWS el mapa final, evitant baixar desenes de gigabytes de graelles intermèdies.

6. **Destrucció ordenada.** En acabar es destrueix primer la part de còmput (nodes i pla de control d'EKS) i després la resta, atès que mantenir el clúster encès mentre es buida el bucket seguiria facturant.

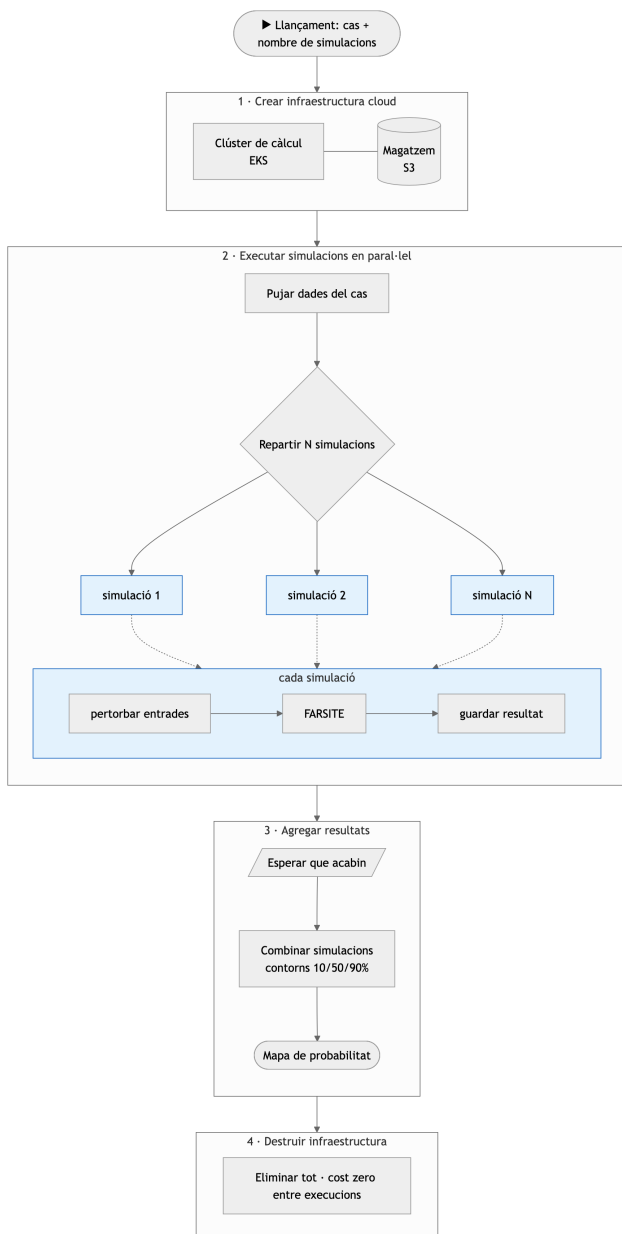


Fig. 1: Pipeline automatitzat d'execució: de la provisió de la infraestructura a la destrucció ordenada.

Tot el cicle, de la provisió a la destrucció, passa sense intervenció manual. Canviar de cas o de nombre de simulacions és tan senzill com tornar a llançar el flux amb paràmetres diferents, cosa que ha permès repetir l'experiment sistemàticament sobre els diversos casos del dataset.

5 RESULTATS

Els resultats que es presenten a continuació compleixen una doble funció. Per una banda, validen que la metodologia probabilística produeix mapes coherents amb la propagació natural del foc i que la convergència estadística s'assoleix amb un nombre moderat de simulacions. Per l'altra, confirmen que la infraestructura descrita a la Secció 4 funciona a escala: el sistema ha completat sense intervenció manual desenes d'execucions de fins a 10000 simulacions cadascuna sobre casos diferents del dataset, mantenint la reproduïbilitat dels resultats i el cost dins dels límits previstos.

5.1 Simulació determinista inicial

La metodologia s'ha aplicat a diversos casos reals del dataset. Un d'ells és el Case 7, l'incendi de Dimos Valtetsioy (Arkàdia, Grècia), amb un perímetre final observat de 1761 ha. Abans d'examinar l'ensemble probabilístic, val la pena partir del que produeix FARSITE en règim determinista sobre aquest cas, és a dir, executar una sola vegada amb les dades originals i sense cap pertorbació. La Fig. 2 mostra el perímetre simulat superposat a la imatge i comparat amb el perímetre real observat.

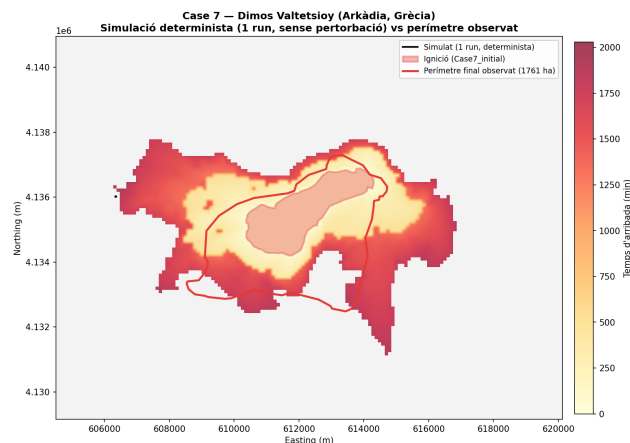


Fig. 2: Simulació determinista del Case 7 amb les dades originals.

El contorn simulat segueix l'orientació general de la propagació i n'aproxima raonablement la magnitud, fet que confirma que FARSITE és un model físicament consistent quan les dades d'entrada descriuen prou bé el cas. La predicció no ofereix cap mesura de confiança: només mirant la línia simulada no es pot saber quines parts del perímetre són robustes i quines respondrien amb força a un canvi petit del vent o de la humitat. Si l'estació hagués registrat un vent 2 mph més fort o 10° desviat, res del mapa permet anticipar si la silueta resultant seria pràcticament la mateixa o significativament diferent.

És aquesta absència d'estimació de confiança el que motiva passar a l'ensemble. Repetir la simulació sobre el ventall de configuracions plausibles dels paràmetres substitueix un contorn únic per un mapa en què cada zona porta associada una probabilitat de crema, i és aquest mapa el que s'analitza a continuació.

5.2 Ensemble probabilístic: Case 7 (Grècia)

La Fig. 3 en mostra el resultat amb 1000 simulacions i les perturbacions aplicades: el mapa de probabilitat de crema amb contorns al 10 %, 50 % i 90 %, i el contorn de punts blancs correspon al perímetre final observat de l'incendi real.

La Fig. 3 es pot interpretar per direccions, i és aquí on aporta la informació més valuosa. A partir del contorn d'ignició, el mapa de probabilitat no creix per igual en totes les direccions. Cap al nord-est la zona d'alta probabilitat gairebé no s'allunya del perímetre inicial, cosa que indica que el foc tenia poca tendència a propagar-se en aquest sentit; de fet, en l'incendi real aquest flanc es va estabilitzar aviat. Cap al nord-oest i cap al sud succeeix el contrari: la probabilitat alta s'allarga molt més enllà del perímetre observat. Aquests són els dos flancs per on el foc tendia a expandir-se amb més força, probablement afavorit per la forma del relleu i pel règim de vent de la zona.

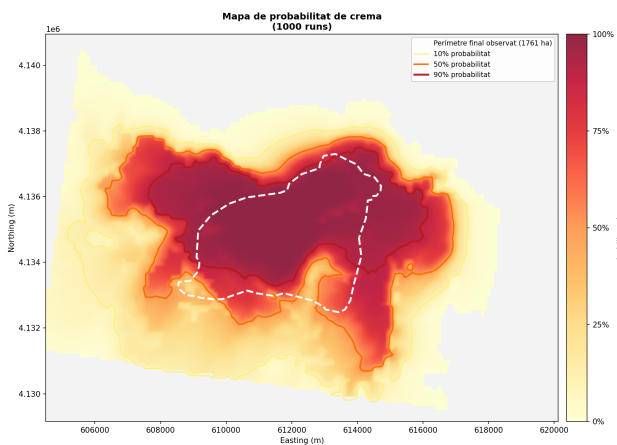


Fig. 3: Incendi real observat i mapa de probabilitat de crema de 1000 simulacions.

5.3 Convergència de simulacions

Per comprovar si augmentar el nombre de simulacions millora el mapa, s'ha repetit l'experiment del Case 7 amb 10000 runs i s'ha comparat amb el conjunt de 1000 runs mitjançant la IoU i la diferència mitjana de probabilitat per cel·la, calculades sobre mapes agregats per a valors creixents de K (30, 50, 100...10000).

Amb 30 runs la IoU respecte al mapa final ja és de 0,90, és a dir, la silueta general apareix aviat; però la probabilitat per cel·la encara es desvia gairebé un 3 %. La convergència real arriba cap als 500 runs, on la IoU supera el 0,99 i la diferència mitjana per cel·la baixa per sota del 0,5 %. Als 1000 runs la diferència es manté per sota del 0,4 % i continua caient cap a zero. La raó és la magnitud de les perturbacions: totes les simulacions parteixen d'unes dades base similars i els desplaçaments són petits, de manera que el conjunt de sortides possibles és estret i la mitjana agregada s'estabilitza ràpidament. Ampliar el rang de les perturbacions faria evolucionar el mapa, però mostrejaria escenaris físicament poc probables; la dispersió addicional seria soroll, no informació nova.

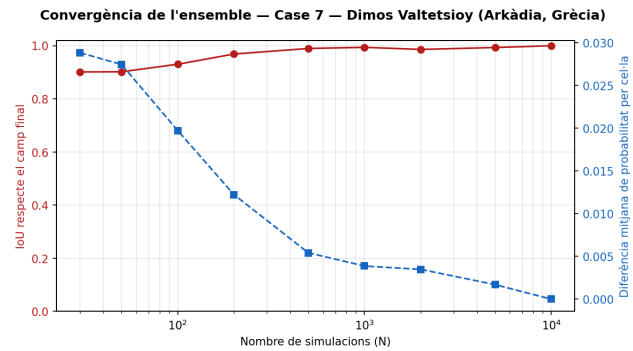


Fig. 4: Convergència de l'ensemble del Case 7 en funció del nombre de simulacions.

Aquest resultat té implicacions directes per al disseny del sistema. Amb el règim de perturbacions actual, executar 10000 runs no produeix un mapa millor que executar-ne 1000, però multiplica el cost de còmput i el temps d'execució: 1000 runs es completen en uns 5 minuts, mentre que 10000 requereixen aproximadament 1h (temps orientatius que depenen de l'extensió i la durada de cada incendi). En un context operatiu, on el mapa pot orientar decisions d'emergència, el temps de resposta és tant o més rellevant que el cost econòmic, i 1000 runs es revelen com el punt òptim del compromís cost-temps-qualitat amb el règim actual de perturbacions.

5.4 Pertorbació del perímetre d'ignició

Amb la voluntat d'ampliar la cobertura experimental, aquesta prova s'executa sobre un cas diferent del dataset: el Case 2 (Miranda do Douro, Trás-os-Montes, Portugal, agost 2012), un incendi de característiques marcadament diferents pel que fa a topografia, vegetació i mida del perímetre inicial. Aplicar la perturbació sobre un nou escenari permet comprovar que el comportament del sistema es manté consistent més enllà d'un únic cas i acumular evidència sobre la robustesa de l'enfocament. La Fig. 5 i la Fig. 6 mostren la comparació.

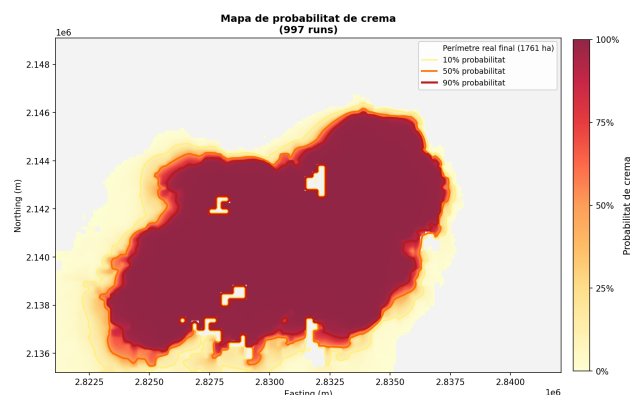


Fig. 5: Case 2 sense perturbació del perímetre d'ignició.

A primera vista els dos mapes són semblants: ocupen aproximadament la mateixa àrea i mantenen la mateixa orientació i forma. La diferència no està en la silueta exterior sinó en l'estructura interna de probabilitat. Sense perturbació, el mapa del Case 2 és gairebé binari: una zona uni-

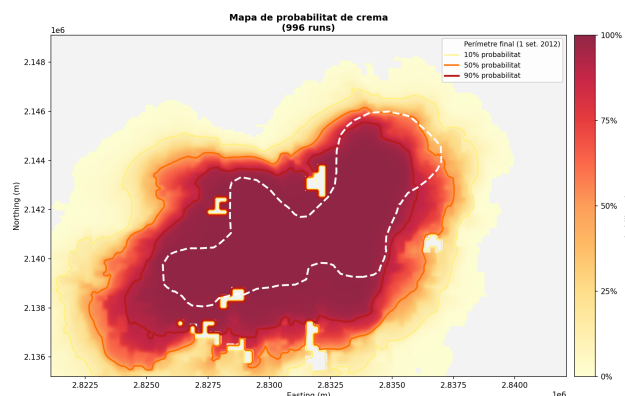


Fig. 6: Case 2 amb pertorbació del perímetre d'ignició ($\sigma_{\text{ign}} = 0,15$).

forme amb probabilitat superior al 90 % i un halo molt estret de probabilitat intermèdia. Aquest tipus de mapa transmet una confiança aparent que no es correspon amb la incertesa real del problema: si la mesura del perímetre inicial té un marge d'error de delineació de l'ordre del 10–15 % i el sistema l'ignora, l'agregat sobreestima la concentració de probabilitat. Amb $\sigma_{\text{ign}} = 0,15$ el mapa adquireix una estructura coherent amb la incertesa real: el nucli del 90 % es redueix a la zona realment robusta i apareixen anells visibles del 50 % i del 10 % que indiquen graus diferents de confiança.

El resultat confirma que la pertorbació de la condició inicial actua sobre l'estructura interna del mapa de manera coherent amb la incertesa física del problema, i ho fa també sobre un cas independent de l'utilitzat per validar la pertorbació del vent. Aquesta consistència entre casos diferents reforça la generalitat de l'enfocament i mostra que els dos eixos de pertorbació (vent i condició inicial) es comporten de manera estable sobre escenaris amb característiques topogràfiques i meteorològiques distintes.

5.5 Zones no-combustibles i interfície urbà-forestal

En revisar els mapes del Case 2 superposats sobre el territori es van detectar zones de forma marcadament geomètrica (rectangular) que no cremaven en cap de les simulacions, apareixent com a buits dins de l'àrea cremada fins i tot envoltades de cel·les amb probabilitat de crema elevada. Calia determinar si es tractava d'un error del simulador o d'un comportament correcte derivat de les dades d'entrada. Per esbrinar-ho es va analitzar directament el fitxer de paisatge (.lcp), l'entrada binària on FARSITE emmagatzema, capa sobre capa, la informació del terreny: elevació, pendent, orientació, cobertura de copes i, la que aquí interessa, el model de combustible de cada cel·la. Un model de combustible és un codi que indica quin tipus de vegetació hi ha en una cel·la i per tant, com s'hi comporta el foc. El paisatge del Case 2 divideix el terreny en una graella de 210×139 cel·les de 100 m de costat i les classifica amb els 13 models de combustible d'Anderson [17], l'esquema estàndard de FARSITE. Els buits s'expliquen perquè 468 cel·les tenen assignat el codi 0, reservat al terreny no-combustible com l'aigua, la roca, el sòl nu o el sòl urbà [18]. Una cel·la mar-

cada així no pot cremar i el simulador la salta sempre, de manera que apareix amb probabilitat zero a tots els runs de l'ensemble. No és un error del simulador, sinó el reflex del mapa de combustible (Fig.7).

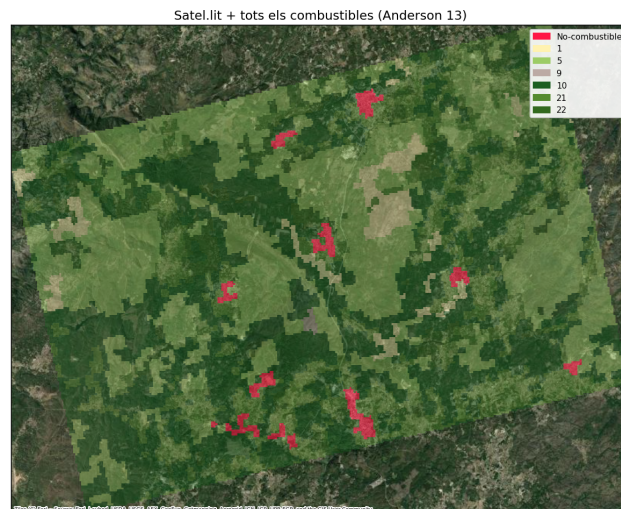


Fig. 7: Cel·les no-combustibles (model 0) del Case 2 superposades sobre imatge satèl·lit.

La forma blocada d'aquestes zones s'explica per la resolució del raster: cada cel·la de 100×100 m rep un únic model de combustible segons la cobertura del sòl dominant, de manera que les vores d'una zona urbana segueixen els límits de la malla i no el contorn real dels edificis. Això explica una observació aparentment contradictòria: al mapa hi ha zones poblades que sí cremen. FARSITE no "veu" pobles, sinó únicament el mapa de combustible rasteritzat. Només el nucli urbà dens, prou compacte per omplir una cel·la sencera, rep el model 0 i actua com a tallafoc; en canvi, l'edificació dispersa amb jardins, horts i arbres entre cases queda dins de cel·les on la vegetació és dominant, i el simulador les crema com si fossin bosc.

6 DISCUSSIÓ

6.1 Interpretació operativa del mapa probabilístic

El mapa de probabilitat de la Fig. 3 va més enllà del perímetre real de l'incendi, però això no representa cap error. FARSITE simula la propagació del foc en condicions naturals i no incorpora cap model d'extinció, de manera que la simulació no s'atura mai per la intervenció humana; l'incendi real, en canvi, es va aturar perquè els equips d'extinció el van contenir. Per aquest motiu és esperable que la zona d'alta probabilitat sobrepassi el perímetre observat: és el resultat que preveu el model, i no una discrepància que calgui corregir. El criteri de validació no és la coincidència de contorns, sinó que el perímetre observat quedi dins de l'envolupant de probabilitat.

L'anàlisi direccional de la Fig. 3 és el que dona valor operatiu al mapa. El seu objectiu no és endevinar el perímetre final un cop extingit l'incendi, sinó mostrar cap a on tendeix a avançar el foc i amb quina probabilitat. Si els equips

d'emergència haguessin disposat d'un mapa d'aquest tipus, haurien pogut identificar ràpidament que el flanc nord-est era de baix risc i concentrar els esforços de contenció als flancs nord-oest i sud, que eren els de propagació més intensa. El mapa funciona com una guia per prioritzar recursos: on actuar primer, on situar els tallafocs i quines zones evacuar amb més antelació.

Una objecció raonable a aquesta lectura seria que el flanc nord-est es va contenir aviat per protegir-hi nuclis de població o infraestructures crítiques, i no perquè el foc hi tingués poc recorregut. Les dades del paisatge permeten descartar aquesta hipòtesi: dins de tota l'àrea afectada, el fitxer de combustible del Case 7 no conté ni una sola cel·la no-combustible (0 de 2928 cel·les a la regió del perímetre), és a dir, cap superfície que el simulador tracti com a inerta (aigua, roca o sòl urbà). La imatge satèl·lit ho confirma visualment, amb un entorn muntanyós i forestal sense pobles ni estructures rellevants ni dins del perímetre ni al seu voltant immediat (Fig. 8). La contenció d'aquell flanc no es pot atribuir a la defensa de béns exposats, cosa que reforça la lectura que el mapa de probabilitat hauria aportat informació directament accionable.

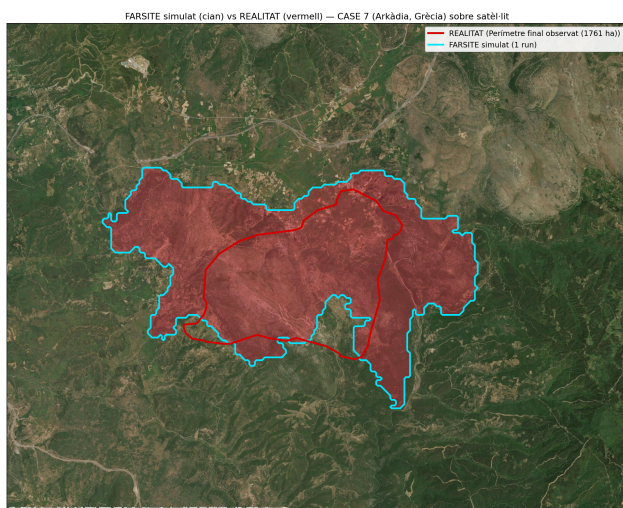


Fig. 8: Case 7 sobre imatge satèl·lit.

6.2 Interfície urbà-forestal

L'observació de les cel·les no-combustibles del Case 2 posa de manifest una limitació coneguda dels simuladors de propagació a la interfície urbà-forestal (WUI) [19, 20]: FARSITE no incorpora cap model de l'entorn construït. Tracta cada cel·la com a combustible de vegetació o com a element inert, i cap de les dues opcions descriu correctament una zona poblada. Per al nucli urbà dens, tractar-lo com a tallafoc infraestima el risc, perquè un incendi real sí que pot destruir edificacions; per a l'edificació dispersa, cremar la cel·la sencera amb física de vegetació no modela ni la ignició estructura a estructura ni la resistència dels edificis. La interpretació del mapa de probabilitat en zones poblades requereix tenir present aquest límit del model.

6.3 Limitacions de les dades d'entrada i del model

Més enllà del comportament a la interfície urbà-forestal, la interpretació dels resultats ha de tenir en compte dues limitacions fonamentals. La primera té a veure amb com es recullen les dades meteorològiques. El vent i la resta de variables atmosfèriques que alimenten el simulador provenen d'estacions fixes repartides pel territori, sovint situades a una distància considerable del front del foc (en alguns casos d'1 a 2 km). La mesura registrada per l'estació no descriu necessàriament el vent local que actua sobre el perímetre actiu: la topografia, amb els seus canals i vessants, i la vegetació poden generar diferències apreciables entre el vent al punt de mesura i el que realment empeny el foc. Es tracta d'una limitació comuna a qualsevol simulador físic que depengui d'entrades meteorològiques puntuals, i no es pot resoldre executant més simulacions; només es pot mitigar amb dades més denses, com ara reanàlisis a alta resolució o una xarxa d'estacions properes al perímetre. Precisament, la pertorbació controlada del vent introduïda en aquest treball cobreix una part d'aquesta incertesa, en explorar el ventall de direccions i velocitats plausibles al voltant del valor mesurat.

La segona limitació és que el simulador, per construcció, no modela la intervenció humana durant l'extinció. Els equips d'extinció prioritzen la defensa de zones poblades i d'infraestructures crítiques per damunt de la contenció en àrees no habitades, de manera que el perímetre real observat sovint reflecteix decisions tàctiques que FARSITE no pot conèixer. Com que la sortida del simulador representa l'evolució del foc en absència total d'intervenció, un mapa amb probabilitat alta en una direcció determinada no és necessàriament erroni encara que no coincideixi amb el perímetre real: pot ser una hipòtesi vàlida de propagació natural que, en la realitat, va ser modulada per l'extinció. Per aquest motiu el mapa de probabilitat de crema s'ha de llegir com una estimació del comportament natural del foc, i no com una predicció directa del perímetre final un cop extingit.

7 CONCLUSIONS

Aquest treball demostra que la combinació d'un simulador determinista amb un mostreig sistemàtic dels seus paràmetres d'entrada és una via efectiva per quantificar la incertesa en la propagació d'incendis forestals. L'aportació principal, més enllà del propi mapa de probabilitat, és la infraestructura cloud-native que la materialitza com a producte operatiu: la contenidorització de FARSITE, la provisió amb Terraform del clúster EKS i del magatzem S3, i l'orquestració amb GitHub Actions de tot el cicle (aixecar, executar en paral·lel, agregar i destruir) permeten passar de l'experiment puntual a un sistema reproducible, traçable i amb un cost acotat per disseny.

Els resultats sobre el Case 7 i el Case 2 confirmen tres propietats del sistema. Primer, els mapes probabilístics resultants són coherents amb la propagació natural del foc i aporten informació direccional accionable que una sola simulació determinista no ofereix. Segon, la pertorbació

del perímetre d'ignició actua de manera complementària a la del vent, reforçant la gradació interna del mapa en escenaris on l'envolupant exterior canvia poc. Tercer, la convergència estadística s'assoleix amb un nombre moderat de *runs* (per sota dels 1000 amb el règim de perturbacions actual), de manera que l'escalat a 10000 simulacions valida la robustesa del pipeline més que millorar el resultat.

Com a línies futures es preveu ampliar els eixos de pertorbació a la resta de variables meteorològiques i substituir el mostreig aleatori simple per esquemes amb justificació estadística més forta (per exemple, sèries temporals amb autocorrelació o estratègies metaheurístiques).

ÚS DE LA IA GENERATIVA

S'ha fet ús d'eines d'intel·ligència artificial generativa com a suport per a aspectes formals i tècnics: millora de la redacció, l'estil i la claredat expositiva, correcció lingüística i ortogràfica, formatació de la bibliografia, traducció entre llengües (castellà, català, anglès) i assistència puntual en la depuració d'errors de compilació i en l'explicació del format dels fitxers d'entrada del simulador.

El disseny de l'enfocament probabilístic, la definició dels objectius, la selecció de paràmetres a pertorbar, el disseny experimental i l'anàlisi dels resultats han estat elaborats íntegrament per l'estudiant.

AGRAÏMENTS

L'autor agraeix als tutors Carles Carrillo i Silvana Silva Pereira la seva orientació, seguiment i suport al llarg del desenvolupament d'aquest treball. Les seves observacions i recomanacions han estat clau per encaminar correctament el projecte i millorar-ne els diferents aspectes tècnics i metodològics. També es vol expressar agraïment per la seva disponibilitat a l'hora de resoldre dubtes i per les reunions de seguiment, que han permès avançar de manera més estructurada i coherent.

REFERÈNCIES

- [1] IPCC, "Climate Change 2022: Impacts, Adaptation and Vulnerability," IPCC Sixth Assessment Report, 2022. [Online]. Available: <https://www.ipcc.ch/report/ar6/wg2/>.
- [2] European Environment Agency, "Forest fires in Europe," Nov. 18, 2021. [Online]. Available: <https://www.eea.europa.eu/en/analysis/indicators/forest-fires-in-europe>.
- [3] NASA, "Wildfires and Climate Change," May 28, 2025. [Online]. Available: <https://science.nasa.gov/earth/explore/wildfires-and-climate-change>.
- [4] J. MacCarthy, S. Tyukavina, M. Weisse, N. Harris, and J. Richter, "New Data Confirms: Forest Fires Are Getting Worse," World Resources Institute, Aug. 2024. [Online]. Available: <https://www.wri.org/insights/global-trends-forest-fires>.
- [5] M. Finney and P. Andrews, "FARSITE—A program for fire growth simulation," *Fire Management Notes*, vol. 59, no. 2, pp. 13–15, Spring 1999. [Online]. Available: https://research.fs.usda.gov/sites/default/files/2024-01/firelab-finney_and_andrews_1999_fmn_v59_i2_pp13-15.pdf.
- [6] J. Silva, J. Marques, I. Gonçalves, R. Brito, S. Teixeira, J. Teixeira, and F. Alvelos, "A systematic review and bibliometric analysis of wildland fire behavior modeling," *Fluids*, vol. 7, no. 12, p. 374, 2022. [Online]. Available: <https://doi.org/10.3390/fluids7120374>.
- [7] Y. Carmel, S. Paz, F. Jahashan, and M. Shoshany, "Assessing fire risk using Monte Carlo simulations of fire spread," *Forest Ecology and Management*, vol. 257, no. 1, pp. 370–377, Jan. 2009. [Online]. Available: <https://doi.org/10.1016/j.foreco.2008.09.039>.
- [8] M. Salis, B. Arca, F. J. Alcasena, M. Arianoutsou, V. Bacciu, P. Duce, B. Duguy, N. Koutsias, G. Mallinis, I. Mitsopoulos, J. M. Moreno, D. Pérez, I. R. Urbieto, E. Xystrakis, G. Zavala, and D. Spano, "Predicting wildfire spread and behaviour in Mediterranean landscapes," *International Journal of Wildland Fire*, vol. 25, no. 10, pp. 1015–1032, 2016. [Online]. Available: <https://doi.org/10.1071/WF15081>.
- [9] J. Pereira, J. Mendes, J. S. Júnior, C. Viegas, and J. R. Paulo, "A Review of Genetic Algorithm Approaches for Wildfire Spread Prediction Calibration," *Mathematics*, vol. 10, no. 3, art. 300, 2022. [Online]. Available: <https://doi.org/10.3390/math10030300>.
- [10] T. Srivas, T. Artés, R. A. de Callafon, and I. Altintas, "Wildfire spread prediction and assimilation for FARSITE using ensemble Kalman filtering," *Procedia Computer Science*, vol. 80, pp. 897–908, 2016. [Online]. Available: <https://doi.org/10.1016/j.procs.2016.05.328>.
- [11] T. Zhou, L. Ding, J. Ji, L. Yu, and Z. Wang, "Ensemble transform Kalman filter (ETKF) for large-scale wildland fire spread simulation using FARSITE tool and state estimation method," *Fire Safety Journal*, vol. 105, pp. 95–106, 2019. [Online]. Available: <https://doi.org/10.1016/j.firesaf.2019.02.009>.
- [12] D. Moreau, K. Wiebels, and C. Boettiger, "Containers for computational reproducibility," *Nature Reviews Methods Primers*, vol. 3, art. 50, Jul. 2023. [Online]. Available: <https://doi.org/10.1038/s43586-023-00236-9>.
- [13] R. M. S. Pinto, A. Benali, A. C. L. Sá, P. M. Fernandes, P. M. M. Soares, R. M. Cardoso, R. M. Trigo, and J. M. C. Pereira, "Probabilistic fire spread forecast as a

- management tool in an operational setting,” *Springer-Plus*, vol. 5, art. 1205, Jul. 2016. [Online]. Available: <https://doi.org/10.1186/s40064-016-2842-9>.
- [14] F. Allaire, J.-B. Filippi, and V. Mallet, “Generation and evaluation of an ensemble of wildland fire simulations,” *International Journal of Wildland Fire*, vol. 29, no. 2, pp. 160–173, 2020. [Online]. Available: <https://doi.org/10.1071/WF19073>.
- [15] HashiCorp, “Terraform: Infrastructure as Code,” HashiCorp Documentation, 2024. [Online]. Available: <https://developer.hashicorp.com/terraform/docs>.
- [16] Amazon Web Services, “Amazon Elastic Kubernetes Service (EKS) — User Guide,” AWS Documentation, 2024. [Online]. Available: <https://docs.aws.amazon.com/eks/latest/userguide/what-is-eks.html>.
- [17] H. E. Anderson, “Aids to determining fuel models for estimating fire behavior,” USDA Forest Service, Intermountain Forest and Range Experiment Station, Ogden, UT, USA, Gen. Tech. Rep. INT-122, Apr. 1982. [Online]. Available: <https://doi.org/10.2737/INT-GTR-122>.
- [18] J. H. Scott and R. E. Burgan, “Standard fire behavior fuel models: a comprehensive set for use with Rothermel’s surface fire spread model,” USDA Forest Service, Rocky Mountain Research Station, Fort Collins, CO, USA, Gen. Tech. Rep. RMRS-GTR-153, Jun. 2005. [Online]. Available: <https://doi.org/10.2737/RMRS-GTR-153>.
- [19] W. E. Mell, S. L. Manzello, A. Maranghides, D. Butry, and R. G. Rehm, “The wildland–urban interface fire problem — current approaches and research needs,” *International Journal of Wildland Fire*, vol. 19, no. 2, pp. 238–251, 2010. [Online]. Available: <https://doi.org/10.1071/WF07131>.
- [20] I. Mitsopoulos, G. Mallinis, and M. Arianoutsou, “Wildfire risk assessment in a typical Mediterranean wildland–urban interface of Greece,” *Environmental Management*, vol. 55, no. 4, pp. 900–915, 2015. [Online]. Available: <https://doi.org/10.1007/s00267-014-0432-6>.

APÈNDIX

A1. Mapes de probabilitat de casos addicionals

Aquesta secció recull els mapes de probabilitat de crema obtinguts amb el mateix pipeline sobre tres casos addicionals del dataset (Case 1, Case 4 i Case 6). Reforcen que el comportament observat sobre els casos il·lustrats al cos principal (Case 2 i Case 7) es manté de manera consistent en escenaris amb topografia, vegetació i règims meteorològics diferents.

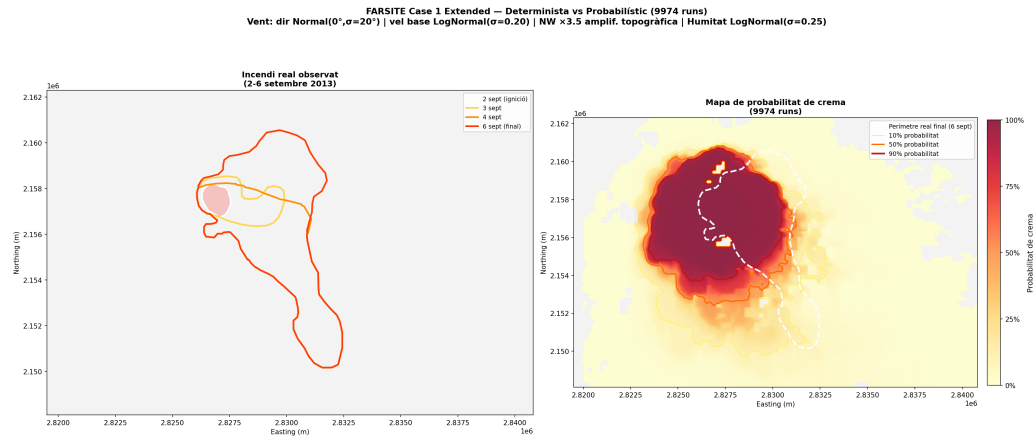


Fig. 9: Mapa de probabilitat de crema del Case 1 (Trás-os-Montes, Portugal, 2013).

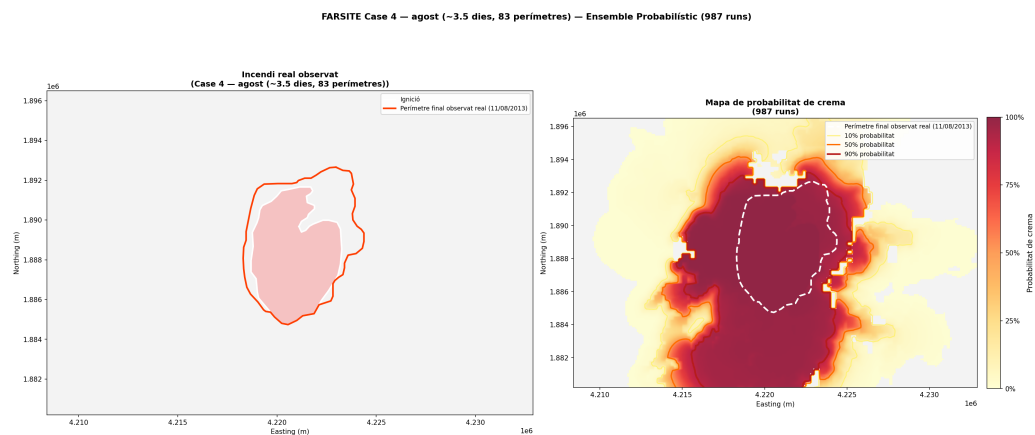


Fig. 10: Mapa de probabilitat de crema del Case 4 (Europa, 2013).

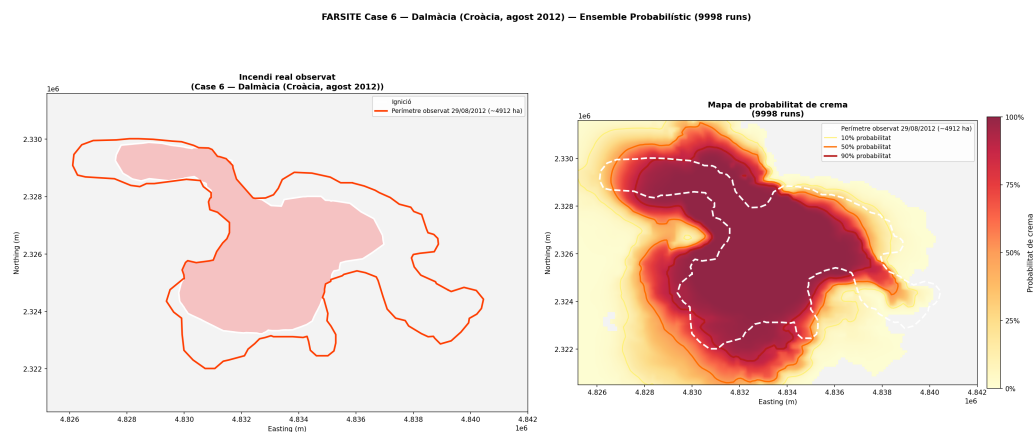


Fig. 11: Mapa de probabilitat de crema del Case 6 (Dalmàcia, Croàcia, 2012).

A2. Representació tridimensional del terreny

A partir del fitxer de paisatge (.lcp) es pot reconstruir una representació tridimensional del relleu, útil per il·lustrar visualment l'orografia sobre la qual es propaga el foc. Aquesta visualització ajuda a interpretar la influència del pendent i de la canalització per valls que es discuteixen al cos principal.

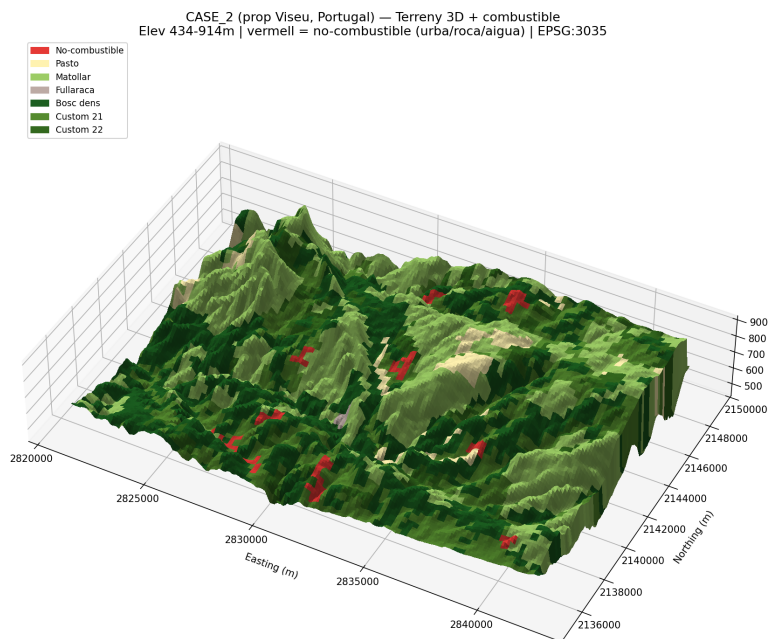


Fig. 12: Representació tridimensional del terreny d'un dels casos, generada a partir del fitxer de paisatge (.lcp).

A3. Ensemble de 10000 simulacions del Case 7

La Fig. 13 mostra el mapa probabilístic del Case 7 obtingut amb 10000 simulacions, referenciat a la secció de convergència del cos principal. Comparat amb el de 1000 runs (Fig. 3), els dos mapes són visualment indistingibles, fet ja quantificat per la mètrica IoU i la diferència mitjana de probabilitat per cel·la (Fig. 4).

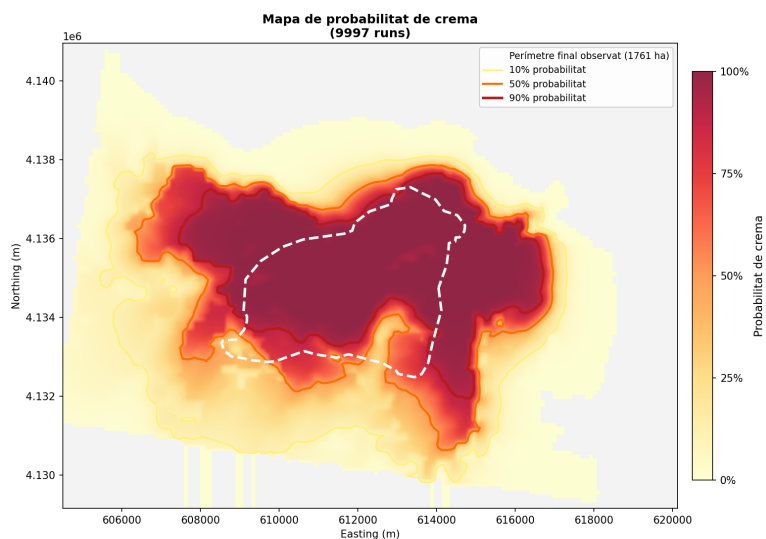


Fig. 13: Ensemble de 10000 simulacions del Case 7.