UNIVERSITÀ
DI PAVIA

Dipartimento di scienze della terra e dell'ambiente

Direttore: Prof. S. Seno

Corso di laurea magistrale in Geoscienze per lo Sviluppo Sostenibile

APPROCCIO DI ANALISI DELLE INTERAZIONI MULTI-HAZARD NATURALI IN ITALIA

Relatore: Dott. Zucca Francesco

Co-relatore: Dott. Novellino Alessandro

Tesi di laurea Magistrale

in Geoscienze per lo Sviluppo Sostenibile

di ANDREA MARENZI

matr. 544591

Anno Accademico 2024/2025

Dichiaro che la presente tesi è stata svolta presso il Dipartimento di Scienze della Terra e dell'Ambiente su tema da me assegnato.



dott. Francesco Zucca

INDICE

INTRODUZIONE 3

1. EVOLUZIONE DEGLI STUDI SUGLI APPROCCI MULTI-HAZARD 6

2. TIPOLOGIE DI INTERAZIONE TRA HAZARD 8

2.1 Eventi Compound 11

2.2 Eventi Simultanei 11

2.3 Eventi Consecutivi 12

3. L'ITALIA COME CASO DI STUDIO 14

4. DATASET UTILIZZATI 17

4.1 Dataset delle frane 17

4.2 Dataset delle alluvioni 18

4.3 Dataset degli incendi boschivi 18

4.4 Dataset delle ondate di calore 19

5. METODOLOGIA 22

5.1 Costruzione dell'inventario multi-hazard 22

5.2 Struttura dell'inventario 22

5.3 Hazard considerati 24

5.4 Costruzione del dataset degli eventi fisici 25

5.5 Integrazione delle ondate di calore 26

5.6 Dimensione complessiva del dataset 26

5.7 Definizione delle soglie spazio-temporali 27

5.7.1 Soglie adottate per la combinazione landside-flood 27

5.7.2 Soglie adottate per la combinazione heatwave-wildfire 28

5.8 Implementazione operativa delle soglie 31

5.9 Significato e limiti delle soglie adottate 32

6. ANALISI E IMPLEMENTAZIONE TRAMITE GIS E PYTHON 32

6.1 Input 33

6.2 Pre-processing 33

6.3 Logica dell'analisi 34

6.4 Funzioni Implementate 35

6.5 Output 36

7. RISULTATI E DISCUSSIONI 37

8. CONCLUSIONI 45

9. APPENDICE 47

10. BIBLIOGRAFIA E SITOGRAFIA 55

INTRODUZIONE

I disastri naturali rappresentano una delle principali sfide per la gestione del territorio e per la sicurezza delle comunità umane. Essi derivano dall'impatto di fenomeni naturali potenzialmente dannosi, quali frane, alluvioni, incendi boschivi, tempeste e ondate di calore, quando questi interessano territori esposti e caratterizzati da elementi vulnerabili causando vittime, evacuazioni di popolazione o danni a infrastrutture e abitazioni.

Negli ultimi decenni, l'aumento della frequenza e dell'intensità degli eventi estremi, insieme alla crescente pressione antropica sugli ecosistemi e agli effetti dei cambiamenti climatici, ha evidenziato la necessità di migliorare la comprensione dei processi naturali potenzialmente dannosi e delle loro interazioni con i sistemi socio-ambientali.

Nel contesto degli studi sui disastri naturali, è fondamentale distinguere tra i concetti di *hazard* e *risk*, che rappresentano due componenti concettualmente differenti ma strettamente correlate nell'analisi delle catastrofi naturali. Il termine *hazard* (pericolosità) indica un processo naturale o antropico potenzialmente dannoso (UNDRR, 2017), caratterizzato in termini probabilistici, da una probabilità di accadimento definita nello spazio e nel tempo. Gli hazard comprendono una vasta gamma di fenomeni, tra cui frane, alluvioni, incendi boschivi, tempeste e ondate di calore, e rappresentano la componente fisica alla base del potenziale danno (UNDRR, 2017; IPCC, 2021).

Il concetto di *risk* (rischio) invece si riferisce alla possibilità che un hazard produca conseguenze negative su persone, infrastrutture, beni economici o ecosistemi. Il rischio dipende quindi non solo dalla presenza di un hazard, ma anche dal grado di vulnerabilità e dal livello di esposizione degli elementi presenti sul territorio. In questo senso, il rischio può essere interpretato come il risultato dell'interazione tra pericolosità, vulnerabilità ed esposizione, secondo un approccio ampiamente adottato negli studi di *disaster risk reduction* e nella pianificazione territoriale (UNDRR, 2017; IPCC, 2021).

Nel contesto della gestione territoriale, la distinzione tra hazard e risk assume un ruolo fondamentale, poiché consente di separare l'analisi dei processi naturali dalla valutazione dei loro impatti sul sistema socioeconomico. Ad esempio, un evento di piena fluviale rappresenta un hazard naturale, ma diventa un rischio solo quando interessa aree esposte e caratterizzate da elementi vulnerabili.

Tradizionalmente, gli studi sulla pericolosità naturale hanno analizzato le diverse fenomologie in modo separato, considerando ciascun fenomeno come un processo indipendente. Tuttavia, negli ultimi anni numerose ricerche hanno evidenziato come i disastri naturali siano spesso il risultato di

interazioni tra più hazard, che possono verificarsi simultaneamente o in sequenza temporale, generando effetti concatenati o addirittura amplificati.

Questo ha portato allo sviluppo del concetto di *multi-hazard*, che descrive situazioni in cui più hazard si verificano nello stesso contesto spazio-temporale oppure risultano tra loro interconnessi attraverso relazioni di causalità, amplificazione o dipendenza reciproca (Claassen et al., 2023; Ferrario et al., 2025). In un sistema multi-hazard, un evento può innescare o favorire il verificarsi di altri fenomeni naturali, dando origine a catene di eventi estremi.

Un esempio tipico è rappresentato da precipitazioni intense che possono innescare simultaneamente alluvioni e frane, oppure da periodi di siccità prolungata che aumentano come evidente la probabilità di sviluppo di incendi boschivi. In altri casi, gli hazard possono verificarsi in successione temporale, dando origine a sequenze di eventi interconnessi, spesso descritte in letteratura come *hazard cascades* o *hazard chains* (Kappes et al., 2012), che amplificano gli impatti complessivi dei fenomeni naturali.

L'approccio multi-hazard rappresenta oggi uno dei principali sviluppi nella ricerca sulla gestione dei rischi naturali, poiché consente di analizzare in modo più realistico la complessità dei sistemi naturali e delle loro interazioni con il territorio. Questo approccio assume un'importanza crescente anche alla luce dei cambiamenti climatici in atto, che possono aumentare la probabilità di eventi estremi combinati o concatenati.

Negli ultimi anni, lo sviluppo di metodologie quantitative ha inoltre favorito l'applicazione di tecniche avanzate per l'analisi degli eventi multi-hazard. In particolare, Ferrario et al. (2025) propongono un approccio basato su tecniche di machine learning per identificare eventi estremi multi-hazard sia nel clima attuale sia in scenari climatici futuri, evidenziando il potenziale di queste metodologie nel migliorare la comprensione delle interazioni tra diversi fenomeni estremi.

Parallelamente, anche le politiche di gestione territoriale stanno progressivamente adottando un approccio integrato alla gestione dei rischi. Un esempio applicativo a scala regionale è rappresentato dal Programma Regionale Integrato di Mitigazione dei Rischi (PRIM) sviluppato dalla Regione Lombardia, che promuove una visione sistemica della sicurezza territoriale basata sull'integrazione di diverse tipologie di rischio naturale, tecnologico e sociale. Tale approccio mira a superare l'analisi dei singoli rischi, considerando congiuntamente pericolosità, vulnerabilità ed esposizione, al fine di sviluppare strategie di prevenzione e mitigazione fondate su una gestione coordinata delle diverse fonti di pericolosità presenti sul territorio.

Alla luce di queste considerazioni, lo studio delle dinamiche multi-hazard rappresenta un elemento chiave per migliorare la comprensione dei processi naturali estremi e per supportare strategie di pianificazione territoriale e gestione del rischio più efficaci.

In questo quadro teorico e applicativo si inserisce il presente lavoro di tesi, che si propone di analizzare le relazioni spazio-temporali tra diversi hazard naturali sul territorio italiano, a scala nazionale**,** attraverso la costruzione di un inventario multi-hazard basato su eventi osservati. In particolare, il lavoro integra dataset relativi a frane, alluvioni, incendi boschivi e ondate di calore in un intervallo temporale compreso tra il 2016 e il 2025, al fine di costruire un inventario geospaziale nazionale di eventi.

I dati sono stati elaborati mediante sistemi informativi geografici (GIS) e successivamente analizzati attraverso procedure sviluppate in ambiente Python, che hanno consentito di identificare in modo automatico relazioni spazio-temporali tra eventi appartenenti a diverse tipologie di hazard con l'obiettivo di contribuire a comprendere le interazioni tra fenomeni naturali estremi e fornire una base metodologica replicabile per futuri studi di multi-hazard e per lo sviluppo di strumenti di supporto alle decisioni basati su tecniche di clustering e machine learning.

1.EVOLUZIONE DEGLI STUDI SUGLI APPROCCI MULTI-HAZARD

Negli ultimi decenni, gli studi sugli hazard naturali hanno conosciuto una significativa evoluzione, passando da approcci focalizzati sull'analisi di singoli fenomeni a metodologie sempre più integrate, in grado di considerare le interazioni tra diversi processi naturali. In una fase iniziale della ricerca scientifica, infatti, l'attenzione si è concentrata prevalentemente su specifiche tipologie di hazard, come frane, alluvioni e terremoti, attraverso lo sviluppo di modelli e metodologie dedicate alla comprensione dei singoli fenomeni.

Il concetto di multi-hazard affonda le proprie radici negli studi pionieristici di Hewitt e Burton (1971), che introdussero l'idea di *all-hazards-at-a-place*, evidenziando come in uno stesso territorio possano coesistere più fenomeni potenzialmente dannosi. Successivamente, a partire dagli anni Novanta, tale concetto è stato progressivamente recepito anche nelle politiche internazionali di gestione del rischio, come emerge nei documenti delle Nazioni Unite (UNEP, 1992; UN, 2002), fino a diventare un elemento centrale nei framework di disaster risk reduction, come l'Hyogo Framework for Action (UNISDR, 2005).

A partire dagli anni Duemila, numerosi studi hanno contribuito allo sviluppo di metodologie operative per l'analisi multi-hazard (Delmonaco et al., 2006), fino a lavori di sintesi come quello di Kappes et al. (2012), che offre un quadro sistematico delle principali sfide concettuali e metodologiche legate all'analisi congiunta di più hazard naturali. A partire dagli anni 2010, la ricerca si è progressivamente orientata verso lo studio delle relazioni tra hazard diversi, con una crescente attenzione alle loro interazioni spaziali e temporali.

In questo ambito, Gill e Malamud (2014) propongono una classificazione delle principali tipologie di interazione tra fenomeni naturali, distinguendo tra eventi in cui un hazard può innescarne un altro (*triggering*), situazioni in cui più eventi si verificano simultaneamente nello spazio e/o nel tempo (*compound events*) e sequenze di eventi interconnessi che possono produrre effetti progressivamente amplificati (*cascading events*). Questo contributo ha rappresentato un passaggio importante nella sistematizzazione teorica degli studi multi-hazard.

Un ulteriore sviluppo della ricerca è rappresentato dagli studi dedicati alla quantificazione delle relazioni tra hazard. In questo contesto, lo studio di Tilloy et al. (2019) analizza 14 hazard naturali, classificati in tre principali categorie: geofisici, atmosferici e idrologici, fornendo una visione sistematica delle possibili interazioni tra fenomeni.

Le interrelazioni tra tali hazard sono investigate attraverso tre principali approcci metodologici. Il primo è rappresentato dalle analisi statistiche basate su dati osservati, particolarmente adatte agli hazard atmosferici e idrologici, per i quali sono disponibili serie temporali estese e variabili continue. Il secondo approccio è costituito dai modelli fisici, che permettono di simulare i processi naturali attraverso equazioni che descrivono il comportamento dei sistemi, risultando fondamentali per hazard come tsunami, storm surge e alluvioni. Infine, i modelli concettuali consentono di rappresentare in maniera semplificata sistemi complessi, soprattutto in ambito idrologico, schematizzando le relazioni tra processi come precipitazione, infiltrazione e deflusso.

In linea con quanto evidenziato da Tilloy et al. (2019), tali approcci rappresentano i principali strumenti attraverso cui possono essere sviluppati gli studi multi-hazard, permettendo di analizzare e quantificare le interazioni tra fenomeni naturali.In particolare, essi evidenziano come gli studi multi-hazard possano essere affrontati attraverso metodologie differenti: analisi statistiche basate su dati osservati, modelli fisici dei processi naturali e modelli concettuali, ciascuna con specifici vantaggi e ambiti di applicazione nella rappresentazione dei sistemi complessi.

Negli ultimi anni, la crescente disponibilità di dati geospaziali e climatici ha favorito lo sviluppo di nuovi approcci basati sull'integrazione di grandi dataset di eventi naturali. In questo contesto si inserisce il lavoro di Claassen et al. (2023), che propone una metodologia per la costruzione di database globali di eventi multi-hazard, consentendo di analizzare le relazioni spazio-temporali tra diversi hazard su scala globale. Tale approccio risulta particolarmente rilevante per il presente lavoro, che si basa proprio sull'analisi delle relazioni spazio-temporali tra eventi osservati.

Parallelamente, la ricerca più recente sta esplorando l'impiego di tecniche avanzate di analisi dei dati e di machine learning per l'identificazione di eventi multi-hazard. Un esempio significativo è rappresentato dallo studio di Ferrario et al. (2025), che propone un approccio basato su algoritmi di apprendimento automatico per individuare eventi estremi multi-hazard sia nel clima attuale sia in scenari climatici futuri.

Nel complesso, l'evoluzione degli studi sugli approcci multi-hazard evidenzia un chiaro passaggio da impostazioni inizialmente settoriali e prevalentemente qualitative verso metodologie sempre più integrate, quantitative e data-driven,ovvero fondate sull'analisi sistematica di grandi quantità di dati geospaziali e temporali, che consentono di identificare e modellare le relazioni tra hazard in modo oggettivo e replicabile. Questo cambiamento ha consentito di spostare l'attenzione dalla semplice analisi dei singoli fenomeni alla comprensione delle loro interazioni spazio-temporali, che rappresentano oggi uno degli aspetti centrali nella ricerca sugli hazard naturali e sulla gestione del rischio.

2.TIPOLOGIE DI INTERAZIONI TRA HAZARD

Uno degli aspetti centrali negli studi multi-hazard riguarda la comprensione delle diverse modalità con cui più fenomeni naturali possono interagire tra loro nello spazio e nel tempo. Nella letteratura scientifica più recente, l'approccio multi-hazard non viene infatti interpretato semplicemente come la presenza simultanea di diversi pericoli naturali all'interno di una stessa area geografica, ma come l'insieme delle relazioni spazio-temporali che collegano eventi appartenenti a differenti tipologie di hazard. In questa prospettiva, la dimensione spaziale e quella temporale assumono un ruolo fondamentale nell'identificazione e nella classificazione delle interazioni tra fenomeni naturali.

.

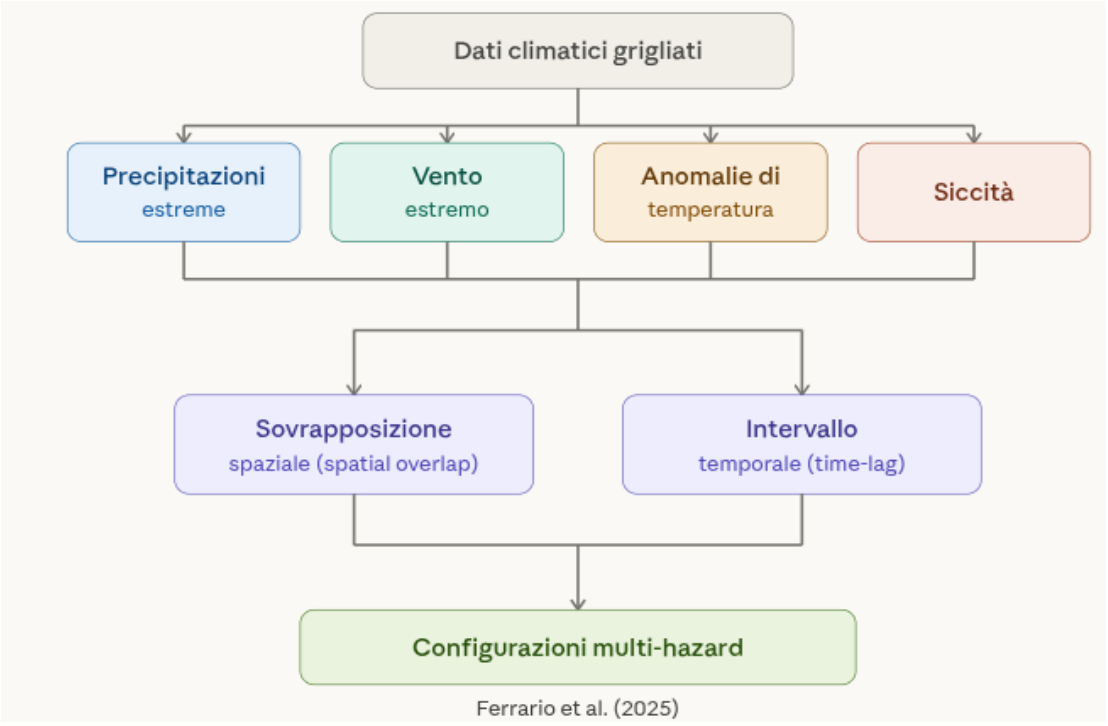


Figura 1: *Framework per l'identificazione delle configurazioni multi-hazard basato su dati climatici grigliati e relazioni spazio-temporali.*

Nei lavori più recenti, l'analisi multi-hazard è spesso sviluppata a partire da hazard di tipo climatico e atmosferico, che rappresentano i principali driver degli eventi estremi. In particolare, Ferrario et al. (2025) considerano quattro categorie principali di hazard climatici: precipitazioni estreme, vento estremo, anomalie di temperatura (ondate di calore) e siccità. Questi fenomeni vengono analizzati attraverso dati climatici grigliati e successivamente combinati mediante procedure basate su sovrapposizioni spazio-temporali (spatial overlap) e intervalli temporali (time-lag) per identificare configurazioni multi-hazard (figura 1).

In un contesto più ampio, Claassen et al. (2023) estendono il concetto di multi-hazard includendo diverse tipologie di hazard naturali, tra cui ancora i fenomeni climatici, ma anche idrologici e geofisici. In questo quadro, eventi quali precipitazioni intense possono rappresentare il driver di fenomeni secondari, come alluvioni e frane, evidenziando come le interazioni tra hazard possano manifestarsi attraverso relazioni di innesco (triggering), amplificazione (amplifying) o successione temporale (consecutive).

Alla luce di questi approcci, le interazioni multi-hazard possono essere analizzate attraverso l'identificazione di specifiche configurazioni spazio-temporali tra eventi. In letteratura, tali configurazioni comprendono diverse tipologie di interazione (Claassen et al., 2023; Ferrario et al., 2025), tra cui eventi compound, eventi consecutive, nonché relazioni di tipo triggering e amplifying.

Tuttavia, ai fini del presente lavoro, si è scelto di adottare una semplificazione operativa, selezionando tre principali configurazioni di interazione: compound events, simultaneous events e consecutive events (Tabella 1). Questa scelta è motivata dalla necessità di tradurre il quadro teorico della letteratura in criteri applicabili a un inventario nazionale di eventi osservati.

Il presente lavoro utilizza eventi reali osservati appartenenti a diverse tipologie di hazard, tra cui frane, alluvioni, incendi boschivi e ondate di calore. Al fine di rendere omogenea la rappresentazione spaziale dei diversi fenomeni e consentire un'analisi integrata, ciascun evento è stato rappresentato attraverso una geometria puntuale, corrispondente al centroide della geometria originaria, oltre che da una data e da una tipologia di hazard (Figure 1 e 2).

La scelta di utilizzare i centroidi non ha lo scopo di descrivere in modo completo la geometria reale degli eventi, ma rappresenta una semplificazione metodologica adottata per rendere confrontabili dataset eterogenei e facilitare l'analisi spazio-temporale. In particolare, la trasformazione delle geometrie originarie (poligonali o raster) in punti consente di ridurre la complessità geometrica dei dati e di uniformare la struttura dell'inventario multi-hazard.

È importante sottolineare che tale approccio comporta inevitabilmente una perdita di informazione spaziale. L'utilizzo di un singolo punto per rappresentare fenomeni areali, come frane, alluvioni e incendi boschivi, o fenomeni continui come le ondate di calore, non consente di descrivere l'estensione reale e la variabilità spaziale degli eventi. Tuttavia, questa semplificazione è stata adottata in modo consapevole, in quanto funzionale agli obiettivi del presente studio.

Dal punto di vista metodologico e computazionale, la rappresentazione puntuale consente infatti di trattare in modo uniforme fenomeni caratterizzati da geometrie e scale spaziali molto diverse. In particolare, l'utilizzo dei centroidi permette di applicare in maniera efficiente procedure automatiche di analisi sviluppate in ambiente Python, basate su metriche di distanza, buffer spaziali e relazioni temporali tra eventi. La gestione di geometrie complesse, come poligoni estesi o campi raster, avrebbe infatti comportato un aumento significativo della complessità computazionale, rendendo più oneroso e meno robusto il processamento dei dati.

Inoltre, la riduzione delle geometrie a punti consente di migliorare la stabilità e la replicabilità dell'analisi, evitando ambiguità legate all'intersezione tra poligoni o alla sovrapposizione di celle raster. Tale scelta permette quindi di definire criteri operativi chiari e coerenti per l'identificazione delle interazioni tra hazard, basati su soglie spazio-temporali applicabili in modo sistematico all'intero dataset.

La struttura finale del dataset, basata su eventi puntuali caratterizzati da coordinate spaziali, data e tipologia di fenomeno, consente quindi di analizzare in modo sistematico le relazioni spazio-temporali tra eventi differenti (Figura 2). Tuttavia, tale approccio non permette di rappresentare i processi fisici alla base dei fenomeni né di identificare in modo diretto relazioni causali tra hazard diversi.

In particolare, nel presente lavoro non sono disponibili informazioni sufficienti per individuare in maniera robusta relazioni di tipo triggering o amplifying, che richiederebbero dati dettagliati sui meccanismi di innesco e sulle dinamiche di propagazione dei fenomeni. Per questo motivo, l'analisi è stata focalizzata su configurazioni direttamente osservabili e quantificabili nello spazio e nel tempo, ovvero compound, simultaneous e consecutive events, coerenti con le informazioni disponibili e traducibili in criteri operativi basati su soglie spazio-temporali.

2.1 COMPOUND EVENTS

Nel presente studio, gli eventi **compound** rappresentano la forma più diretta di interazione multi-hazard in quanto non possiedono separazione né spaziale né temporale tra gli eventi. Essi si verificano quando due o più fenomeni naturali interessano lo stesso settore territoriale e lo stesso intervallo temporale, determinando una sovrapposizione sia spaziale sia temporale tra gli eventi. Questa configurazione è coerente con il concetto di eventi spatio-temporally compound discusso da Ferrario et al. (2025), in cui i diversi hazard vengono combinati attraverso l'intersezione simultanea delle loro dimensioni spaziali e temporali.

Nel contesto degli hazard considerati in questo lavoro, esempi tipici di eventi compound includono precipitazioni intense che generano contemporaneamente alluvioni e frane nello stesso settore territoriale, oppure condizioni di caldo estremo e siccità che favoriscono simultaneamente lo sviluppo e la propagazione di incendi boschivi. In questi casi, l'interazione tra gli hazard produce effetti amplificati rispetto a quelli generati dai singoli fenomeni considerati separatamente.

2.2 SIMULTANEUS EVENTS

Gli eventi **simultaneous** rappresentano una configurazione in cui diversi hazard si verificano nello stesso intervallo temporale ma in aree geografiche differenti. In questo caso, la relazione tra gli eventi è di tipo temporale, mentre la componente spaziale risulta spazialmente discontinua. Questa tipologia non corrisponde direttamente a una categoria formalmente definita nella classificazione di Claassen et al. (2023), ma può essere interpretata in modo analogo al concetto di eventi spatially compound

discusso da Ferrario et al. (2025), riferito a hazard che si verificano nello stesso periodo in luoghi differenti all'interno di una stessa regione.

Dal punto di vista interpretativo, gli eventi simultaneous evidenziano come fenomeni naturali distinti possano verificarsi in modo parallelo all'interno dello stesso contesto meteorologico o climatico. Ad esempio, un sistema perturbato può generare contemporaneamente alluvioni in aree di pianura e frane in settori montani limitrofi. Pur non verificandosi nello stesso punto geografico, tali eventi possono essere considerati parte della stessa configurazione multi-hazard in quanto condividono lo stesso contesto atmosferico.

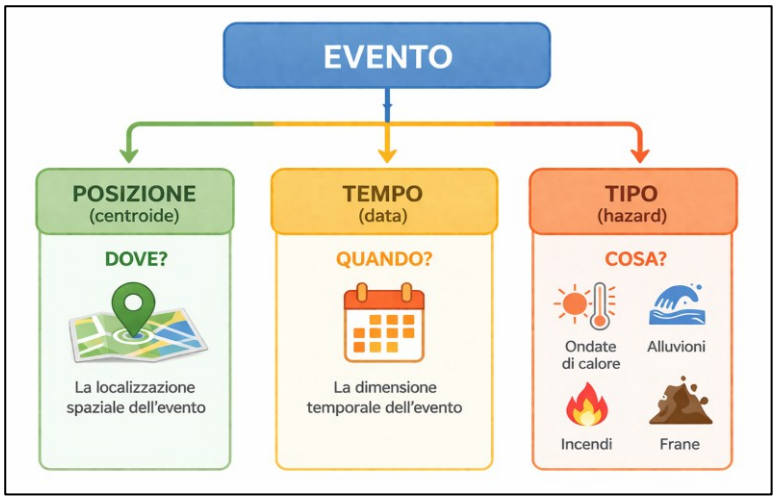
L'identificazione di questa tipologia di interazione richiede la definizione di soglie spaziali e temporali che permettano di stabilire se eventi distinti possano essere considerati simultanei all'interno dello stesso sistema territoriale.

2.3 CONSECUTIVE EVENTS

Gli eventi **consecutive** descrivono situazioni in cui due o più hazard si verificano nello stesso territorio ma in momenti diversi, separati da un intervallo temporale definito. In questo caso, la relazione tra gli eventi è di tipo sequenziale e il primo fenomeno può influenzare le condizioni ambientali favorendo l'innesco del successivo.

Questa configurazione è pienamente coerente sia con la definizione di consecutive hazards proposta da Claassen et al. (2023), sia con l'approccio metodologico di Ferrario et al. (2025), che introducono esplicitamente il concetto di time-lag per identificare sequenze di eventi interconnessi. L'utilizzo di soglie temporali consente infatti di individuare configurazioni in cui un evento modifica temporaneamente il sistema naturale, aumentando la probabilità che un secondo fenomeno si verifichi nello stesso settore territoriale.

Nel contesto del presente lavoro, esempi di eventi consecutive includono precipitazioni intense seguite da frane nei giorni successivi, oppure periodi di siccità prolungata che favoriscono la propagazione di incendi boschivi. In sistemi climatici complessi possono inoltre verificarsi sequenze di eventi estremi, come ondate di calore seguite da precipitazioni intense o alternanze tra periodi di siccità e piogge estreme.



.***Figura 1*** *– Rappresentazione schematica della struttura di un evento all'interno dell'inventario multi-hazard, caratterizzato da attributi spaziali (centroide), temporali (data) e tematici (tipologia di hazard), associati a un identificativo univoco.*

Tipologia	Relazione spaziale	Relazione temporale	Descrizione	Esempio
Compound events	Stesso luogo	Stesso tempo	Eventi simultanei nello stesso punto	heatwave+siccità(stessa area,stesso tempo)
Simultaneous events	Area diversa (entro soglia spaziale)	Stesso Tempo	Eventi contemporanei ma spazialmente separati	Alluvione Nord + incendio Sud (stesso giorno)
Consecutive events	Stesso luogo	Tempo diverso(time-lag)	Eventi in successione nello stesso punto	pioggia → frana (time-lag)

Tabella 1 *– Classificazione delle interazioni multi-hazard sulla base delle relazioni spazio-temporali tra eventi. Fonte: rielaborazione da Claassen et al. (2023) e Ferrario et al. (2025)*

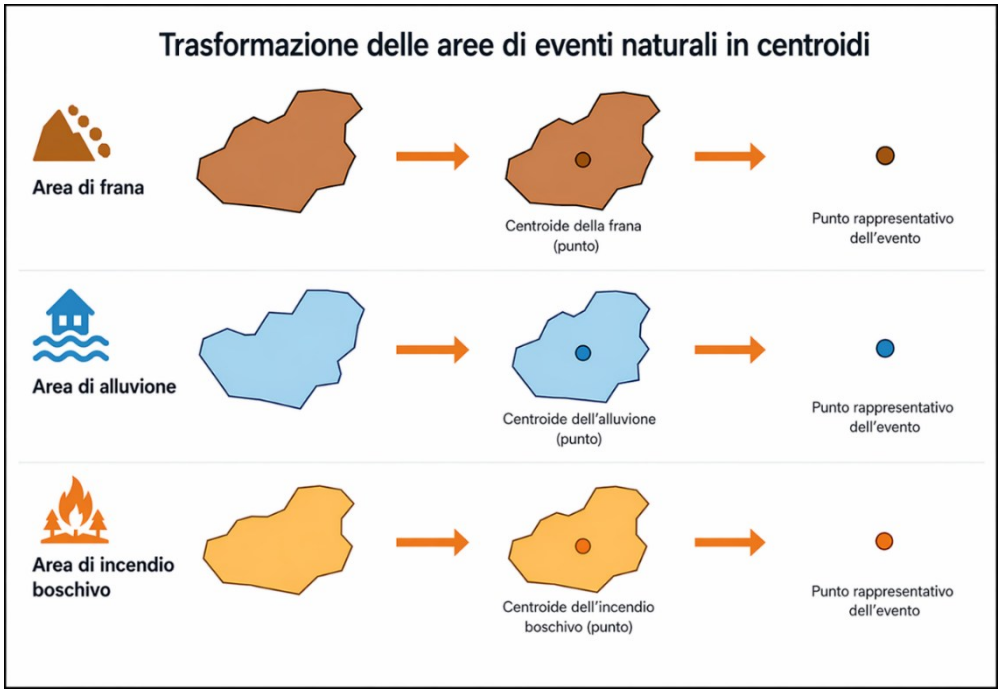


Figura 2 *– Schema di trasformazione delle geometrie degli eventi naturali (frane, alluvioni e incendi boschivi) da rappresentazione areale (poligoni) a rappresentazione puntuale (centroidi).*

3.L'ITALIA COME CASO DI STUDIO

L'Italia è stata scelta come caso di studio perché rappresenta uno dei contesti europei più significativi per l'analisi multi-hazard. La rilevanza del territorio italiano dipende dalla combinazione di alcuni elementi principali: la complessità morfologica e geomorfologica, l'elevata frequenza e varietà dei pericoli naturali e la disponibilità di basi dati ufficiali e osservazionali sufficientemente estese da consentire analisi spaziali e spazio-temporali su scala nazionale. In altri termini, l'Italia non è soltanto un territorio esposto a singoli hazard, ma un sistema territoriale in cui processi naturali differenti coesistono, si distribuiscono in modo eterogeneo e, in molti casi, possono interagire tra loro.

Dal punto di vista fisico, l'Italia è un territorio intrinsecamente fragile. Le caratteristiche morfologiche, geologiche, idrologiche, meteo-climatiche e sismiche del Paese determinano una vera e propria vulnerabilità strutturale del territorio ai fenomeni naturali, alla quale si sovrappone una crescente pressione antropica (ISPRA, 2025a). In particolare, le superfici artificiali sono passate dal 2,7% degli anni Cinquanta al 7,16% nel 2023, aumentando progressivamente l'esposizione di popolazione, edifici e infrastrutture ai pericoli naturali (ISPRA, 2025a).

La morfologia del Paese contribuisce in modo determinante a questa condizione. L'Italia è caratterizzata da una prevalenza di territori montuosi e collinari, con la presenza delle Alpi a nord e

della catena appenninica che attraversa longitudinalmente la penisola. Questa configurazione favorisce diffusamente i processi di instabilità dei versanti e una risposta idrologica rapida dei bacini in caso di precipitazioni intense. Le aree di pianura e le fasce costiere risultano invece fortemente urbanizzate e densamente infrastrutturate, con un'elevata concentrazione di elementi esposti, rendendo il territorio nazionale particolarmente adatto allo studio dei pericoli naturali multipli.

Il quadro quantitativo conferma l'elevata esposizione. Complessivamente, il 94,5% dei comuni italiani (7.463 comuni) risulta a rischio per frane, alluvioni, valanghe e/o erosione costiera (ISPRA, 2025a). Inoltre, una quota significativa del territorio nazionale è classificata a maggiore pericolosità per frane e alluvioni, evidenziando come la pericolosità naturale interessi diffusamente il territorio italiano e non solo aree marginali.

Tra i principali pericoli naturali, le frane occupano un ruolo centrale. L'Italia si conferma tra i Paesi europei più esposti ai fenomeni gravitativi, con 636.207 frane censite nell'Inventario dei Fenomeni Franosi in Italia (IFFI), con copertura temporale 1116–2024 (ISPRA, 2025a). Circa il 28% di queste è costituito da fenomeni estremamente rapidi, quali crolli e colate di fango e detrito, caratterizzati da elevata distruttività e spesso associati a perdite di vite umane (ISPRA, 2025a).

La mappatura della pericolosità da frana evidenzia ulteriormente la rilevanza del fenomeno. La Mosaicatura nazionale delle aree a pericolosità da frana indica che circa 69.530 km², pari al 23% del territorio nazionale, ricadono in aree a pericolosità da frana o di attenzione, mentre le classi a pericolosità più elevata (P3 e P4) coprono 28.801 km², equivalenti al 9,5% del territorio nazionale (ISPRA, 2025a). Questi dati evidenziano come il dissesto di versante rappresenti una componente strutturale del territorio italiano.

Il peso delle frane emerge anche in termini di impatto umano e sociale. Negli ultimi 50 anni (1974–2023), gli eventi franosi hanno causato 1.060 morti, 10 dispersi, 1.443 feriti e 138.743 evacuati o senzatetto (ISPRA, 2025a). La popolazione a rischio frane è stimata in circa 5,7 milioni di abitanti, di cui 1,28 milioni residenti in aree a pericolosità elevata e molto elevata (P3 e P4), pari al 2,2% della popolazione totale. Nelle stesse aree risultano esposte oltre 582.000 famiglie, più di 742.000 edifici, circa 75.000 unità locali di impresa e circa 14.000 beni culturali (ISPRA, 2025a).

Le alluvioni rappresentano un secondo asse fondamentale della pericolosità naturale italiana. Circa 6,8 milioni di abitanti risultano esposti a rischio alluvioni nello scenario a pericolosità idraulica media, con tempi di ritorno compresi tra 100 e 200 anni (ISPRA, 2025a). Questo dato evidenzia l'elevata esposizione del territorio italiano anche rispetto al rischio idraulico.

Negli ultimi anni, tale fenomeno è stato ulteriormente accentuato dai cambiamenti climatici. Il triennio 2022–2024 è stato caratterizzato da elevate anomalie termiche e da eventi idro-meteorologici di eccezionale intensità; in particolare, nel maggio 2023 le precipitazioni osservate a scala nazionale sono state più del doppio della media mensile e localmente superiori di oltre sei volte rispetto alle medie climatiche (ISPRA, 2025a). L'intensificazione delle piogge brevi e concentrate contribuisce all'aumento delle frane superficiali, delle colate rapide e delle alluvioni, incluse le flash flood.

Un ulteriore hazard rilevante è rappresentato dagli incendi boschivi, particolarmente diffusi nelle regioni mediterranee e nell'Italia centro-meridionale e insulare. Secondo il rapporto ISPRA sugli incendi boschivi, nella stagione 2025 (15 giugno – 15 settembre) sono stati percorsi dal fuoco 78.797 ettari di superficie complessiva, di cui 10.016 ettari di superficie forestale, pari a circa il 13% del totale (ISPRA, 2025b). La categoria forestale maggiormente colpita è quella delle latifoglie sempreverdi, a conferma della vulnerabilità degli ecosistemi mediterranei.

La distribuzione territoriale degli incendi evidenzia una forte concentrazione nel Mezzogiorno e nelle isole, in particolare in Sicilia, Calabria, Campania, Puglia e Sardegna. Le province più colpite risultano Enna e Cosenza, mentre una quota significativa della superficie forestale bruciata ricade all'interno di aree naturali protette (ISPRA, 2025b).

Le ondate di calore completano il quadro della pericolosità naturale. Il triennio 2022–2024 è stato caratterizzato da elevate anomalie termiche e il 2024 è risultato l'anno più caldo registrato in Italia, evidenziando una tendenza verso condizioni climatiche favorevoli a eventi estremi e periodi prolungati di siccità (ISPRA, 2025a).

In definitiva, l'Italia rappresenta un caso di studio particolarmente significativo per l'analisi multi-hazard, in quanto concentra condizioni di elevata complessità morfologica, ampia varietà di pericoli naturali, forte esposizione territoriale e disponibilità di basi dati dettagliate. Nel loro insieme, questi elementi rendono il territorio italiano un vero laboratorio naturale per lo studio delle interazioni tra diversi hazard e per lo sviluppo di analisi multi-hazard su scala nazionale.

4.DATASET UTILIZZATI

Per la costruzione dell'inventario multi-hazard nazionale sono stati utilizzati dataset relativi a quattro principali tipologie di pericoli naturali: frane, alluvioni, incendi boschivi e ondate di calore. La selezione dei dataset è stata guidata dall'obiettivo di utilizzare dati osservazionali affidabili e georeferenziati, caratterizzati dalla presenza di informazioni spaziali e temporali sufficientemente precise da consentire analisi delle relazioni spazio-temporali tra eventi appartenenti a differenti categorie di pericolo.

4.1 DATASET DELLE FRANE

Il dataset relativo alle frane rappresenta una componente fondamentale dell'analisi multi-hazard, considerando l'elevata diffusione dei fenomeni gravitativi sul territorio italiano. In una fase preliminare del lavoro sono stati presi in considerazione diversi inventari nazionali di frane, tra cui l'Inventario dei Fenomeni Franosi in Italia (IFFI) sviluppato da ISPRA, il catalogo ITALICA (Italian rainfall-induced Landslides Catalogue) sviluppato dal CNR-IRPI, il database storico dei terremoti italiani (CFTI) gestito dall'INGV e il catalogo CEDIT relativo alle deformazioni superficiali indotte da terremoti.

Questi database rappresentano le principali fonti informative disponibili a livello nazionale per lo studio dei fenomeni franosi e comprendono sia inventari poligonali delle aree interessate da frane sia cataloghi puntuali di eventi documentati. Tuttavia, molti di questi inventari sono stati progettati principalmente per la mappatura della pericolosità geomorfologica e non sempre presentano una datazione precisa degli eventi o una chiara identificazione dell'evento scatenante.

Per questa ragione, ai fini dell'analisi multi-hazard sono stati selezionati dataset relativi a specifici eventi franosi estremi documentati nella letteratura scientifica recente. In particolare, sono stati utilizzati inventari di frane associati a tre eventi pluviometrici estremi verificatisi in Italia negli ultimi anni: l'evento dell'Emilia-Romagna del maggio 2023 (Berti et al.), l'evento dell'Umbria-Marche del settembre 2022 (Santangelo et al., 2023) e l'evento della Liguria dell'ottobre 2020 (Ardizzone et al., 2024). Questi dataset contengono frane innescate da precipitazioni intense e rappresentano esempi ben documentati di eventi geomorfologici estremi.

L'integrazione di questi inventari ha permesso di costruire un dataset nazionale di frane evento-specifiche caratterizzato da un totale di circa 82.000 fenomeni franosi, ciascuno associato a un evento

meteorologico estremo documentato. L'utilizzo di frane legate a eventi specifici consente di disporre di informazioni temporali affidabili, requisito essenziale per l'analisi delle interazioni spazio-temporali tra diversi hazard naturali.

4.2 DATASET DELLE ALLUVIONI

Il dataset delle alluvioni è stato realizzato utilizzando i dati forniti dal Copernicus Emergency Management Service (EMS), il servizio europeo di mappatura rapida degli eventi estremi basato sull'analisi di immagini satellitari. Il sistema Copernicus EMS produce mappe tematiche relative a diversi tipi di disastri naturali, tra cui alluvioni, incendi, terremoti e frane, con l'obiettivo di supportare le attività di gestione delle emergenze e di monitoraggio dei fenomeni naturali.

Nel presente lavoro sono state selezionate le principali attivazioni EMS relative a eventi alluvionali verificatisi in Italia negli ultimi anni. I dati utilizzati derivano dai layer vettoriali di tipo Observed Event, che rappresentano le aree effettivamente allagate identificate attraverso l'interpretazione di immagini satellitari acquisite durante o immediatamente dopo l'evento.

Questi layer sono caratterizzati da geometrie poligonali che descrivono l'estensione spaziale delle aree inondate. A ciascun evento è associato un codice identificativo EMSR che consente di ricondurre i poligoni alle specifiche attivazioni del servizio Copernicus e quindi alle relative informazioni temporali. L'integrazione dei diversi layer relativi alle attivazioni selezionate ha permesso di costruire un dataset nazionale delle alluvioni composto da circa 17.800 eventi alluvionali, distribuiti su diverse regioni italiane.

4.3 DATASET DEGLI INCENDI BOSCHIVI

Il dataset relativo agli incendi boschivi è stato costruito utilizzando i dati forniti dal sistema EFFIS (European Forest Fire Information System, sviluppato nell'ambito del programma COPERNICUS per il monitoraggio degli incendi a scala europea. EFFIS rappresenta uno dei principali sistemi di

riferimento per l'analisi degli incendi boschivi nel contesto europeo e mediterraneo, fornendo informazioni geospaziali sulle aree percorse dal fuoco.

Il database EFFIS include dati relativi agli incendi boschivi derivati dall'analisi di immagini satellitari e comprende informazioni sulla localizzazione degli eventi, sull'estensione delle aree bruciate e sulle date di inizio e fine incendio. Nel presente lavoro sono stati utilizzati i layer vettoriali relativi alle aree bruciate nel periodo compreso tra 2015 e 2024, che coprono l'intero territorio italiano.

I dati sono rappresentati da geometrie poligonali che descrivono l'estensione delle superfici interessate dal passaggio del fuoco. L'integrazione dei layer annuali ha permesso di costruire un dataset nazionale degli incendi boschivi composto da circa 8.300 eventi distribuiti su tutto il territorio.

La disponibilità di informazioni temporali precise relative all'inizio e alla durata degli incendi rende questo dataset particolarmente adatto allo studio delle relazioni tra incendi boschivi e altri hazard naturali.

4.4 DATASET DELLE ONDATE DI CALORE

Oltre agli hazard di tipo geomorfologico e idrologico, il presente lavoro considera anche le ondate di calore come hazard di tipo climatico, in quanto fenomeni caratterizzati da un forte impatto ambientale e da possibili relazioni con altri processi naturali, in particolare con gli incendi boschivi.

Nel presente studio è stata adottata una definizione operativa di ondata di calore basata sugli indicatori climatici forniti dal Copernicus Climate Change Service (C3S). In particolare, un'ondata di calore è definita come un periodo di almeno tre giorni consecutivi in cui sia la temperatura massima giornaliera (Tmax) sia la temperatura minima giornaliera (Tmin) superano la soglia del 95° percentile della distribuzione climatologica relativa alla specifica giornata, calcolata sul periodo di riferimento 1991–2020. Tale definizione consente di identificare eventi estremi rispetto alla variabilità climatica locale e rappresenta un criterio coerente con gli standard utilizzati negli studi climatici recenti(figura 3).

Il dataset delle ondate di calore è stato costruito a partire da dati climatici forniti dal Climate Data Store (Copernicus Climate Change Service), originariamente distribuiti in formato NetCDF come campi continui su griglia regolare. Tali dati sono rappresentati come raster climatici, in cui il territorio

è suddiviso in celle regolari definite in coordinate geografiche (latitudine-longitudine), con una risoluzione spaziale di 0,25° × 0,25°. Questa risoluzione corrisponde approssimativamente a celle di circa 25–27 km di lato, a seconda della latitudine.

Ogni cella della griglia rappresenta un'unità spaziale su cui vengono calcolati indicatori climatici, tra cui la durata delle ondate di calore. A differenza degli hazard geomorfologici, che sono eventi discreti e spazialmente delimitati, le ondate di calore sono fenomeni continui nello spazio, derivanti da anomalie delle condizioni atmosferiche su larga scala.

Per rendere questi dati utilizzabili nell'ambito dell'analisi multi-hazard, il campo continuo è stato esportato in formato CSV, in cui ogni record rappresenta una singola cella della griglia associata a una specifica data e alla durata dell'evento. Il file risultante contiene quindi, per ciascun evento, le coordinate geografiche (latitudine e longitudine), la data di inizio e la durata dell'ondata di calore.

Successivamente, tramite Python, il dataset è stato convertito in un layer geografico puntuale utilizzando le coordinate presenti nel CSV. In particolare, le coppie latitudine-longitudine sono state trasformate in punti mediante la funzione *points_from_xy*, generando un GeoDataFrame. I punti ottenuti rappresentano il centro geometrico di ciascuna cella della griglia, e sono stati quindi interpretati come centroidi degli eventi di ondata di calore (Figura 4).

Questa trasformazione consente di uniformare la rappresentazione spaziale delle ondate di calore rispetto agli altri hazard analizzati, rendendo possibile l'applicazione di analisi basate su distanza, buffer e relazioni spazio-temporali. Complessivamente, il dataset così ottenuto comprende circa 69.500 eventi di ondata di calore distribuiti sul territorio nazionale nel periodo 2018–2024.

A differenza degli altri hazard considerati, le ondate di calore sono state trattate come dataset separato, in quanto la loro natura continua e la loro origine climatica le distinguono dai fenomeni discreti di tipo geomorfologico e idrologico.

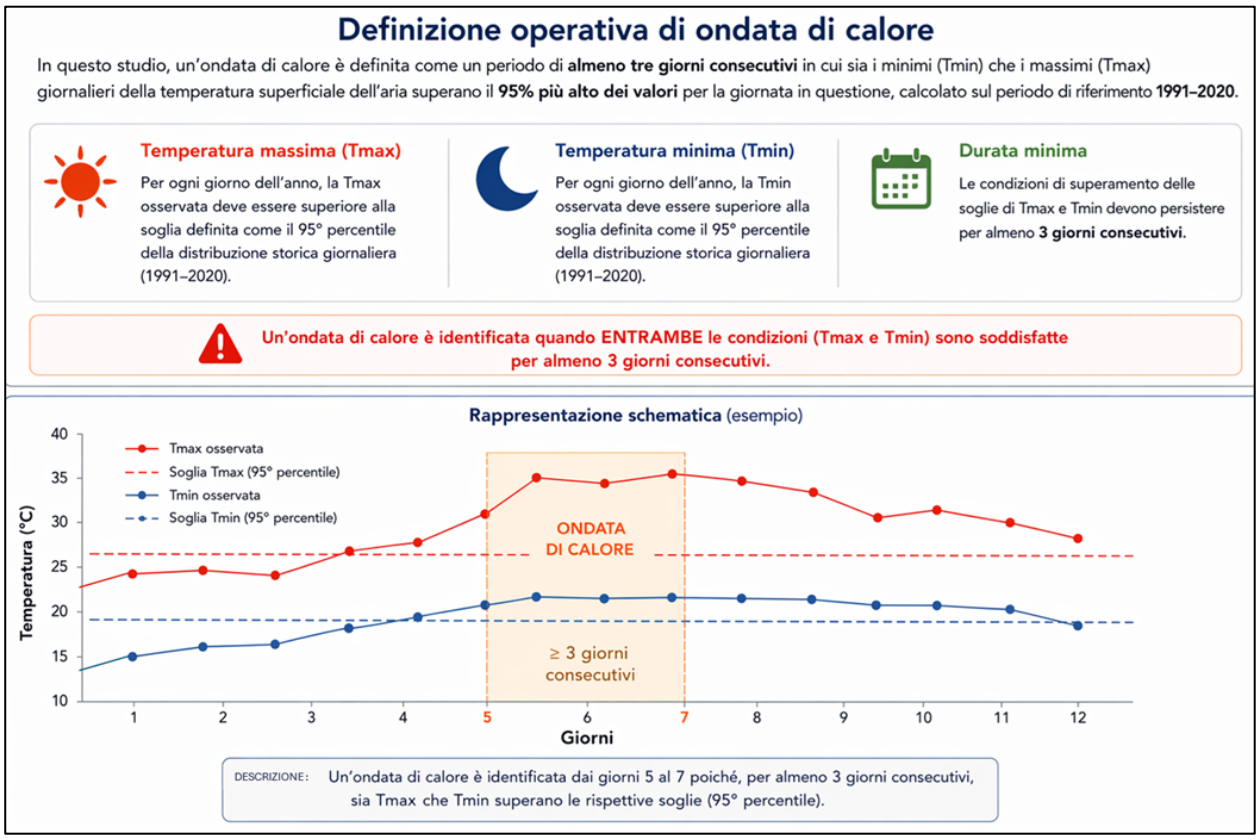


Figura 3 *– Definizione operativa di ondata di calore con grafico esplicativo di individuazione dell'evento.*

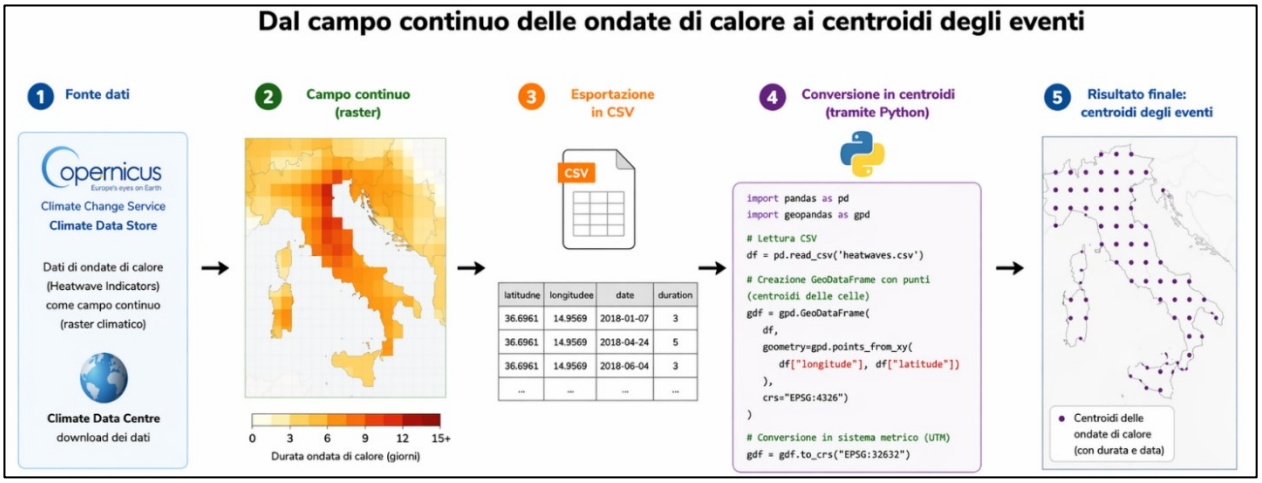


Figura 4 *– Schema del processo di costruzione del dataset delle ondate di calore: dal campo climatico continuo (raster) derivato da Copernicus Climate Data Store, all'esportazione in formato CSV e alla successiva conversione in centroidi tramite Python.*

5.METODOLOGIA

5.1COSTRUZIONE DELL'INVENTARIO MULTI-HAZARD

L'obiettivo principale della fase metodologica è stato la costruzione di un inventario multi-hazard a scala nazionale, basato su eventi naturali osservati e caratterizzati da informazioni spaziali e temporali coerenti. L'inventario costituisce la base dati necessaria per l'analisi delle interazioni tra hazard differenti e rappresenta l'elemento centrale dell'intero flusso di lavoro sviluppato nella presente ricerca.

Il lavoro si inserisce nel contesto delle più recenti ricerche dedicate allo studio dei sistemi multi-hazard e, in particolare, nell'ambito dei progetti promossi dall'Agenzia Spaziale Europea (ESA) dedicati allo studio delle interazioni tra pericoli naturali, tra cui EO4MultiHazards (Earth Observation for Multi-Hazard) e AMHEI (Advancing Multi-Hazard Understanding in Earthquake-Affected Regions). Tali progetti mirano a superare l'analisi tradizionale dei singoli pericoli naturali, promuovendo approcci integrati basati sull'analisi congiunta delle interazioni tra hazard nello spazio e nel tempo.

5.2 STRUTTURA DELL'INVENTARIO

Il principio metodologico alla base della costruzione dell'inventario è stato quello di lavorare esclusivamente su eventi reali osservati, ciascuno rappresentato attraverso tre informazioni fondamentali:

➤ Identificativo univoco (ID), che permette di distinguere ogni evento all'interno del database;

➤ Informazione temporale, rappresentata dalla data di occorrenza dell'evento;

➤ Informazione spaziale, rappresentata dalla posizione geografica dell'evento.

Questa struttura consente di rappresentare ogni evento naturale come un'entità spazio-temporale definita, rendendo possibile il confronto tra eventi appartenenti a tipologie di hazard differenti (figura 1). L'approccio adottato è coerente con la letteratura recente sugli studi multi-hazard, secondo cui tali sistemi possono essere interpretati come configurazioni di eventi estremi collegati attraverso relazioni nello spazio e nel tempo (Claassen et al., 2023; Ferrario et al., 2025).

Per garantire la confrontabilità tra dataset eterogenei e consentire un'analisi integrata delle relazioni spazio-temporali tra hazard, tutte le geometrie degli eventi sono state convertite in rappresentazioni

puntuali mediante il centroide. Tale scelta comporta inevitabilmente una semplificazione della componente spaziale, in quanto fenomeni originariamente rappresentati come poligoni (ad esempio alluvioni o aree bruciate) o come campi continui (nel caso delle ondate di calore) vengono ridotti a una singola posizione rappresentativa (Figure 3 e 4).

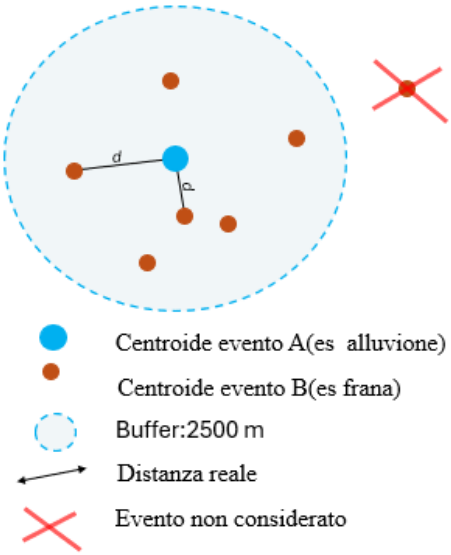
Questa semplificazione implica che la geometria reale degli eventi non venga più esplicitamente considerata nell'analisi, ma venga sostituita dalla posizione del centroide, utilizzata come riferimento spaziale univoco. Di conseguenza, le relazioni spaziali tra gli eventi non sono più descritte in termini di intersezione o sovrapposizione tra aree, bensì attraverso la distanza tra i rispettivi centroidi.

La componente spaziale dell'analisi è quindi trattata mediante un approccio basato su distanze. In una prima fase, viene applicato un buffer spaziale attorno agli eventi di riferimento, con funzione esclusivamente preliminare di preselezione. Il buffer definisce una zona di ricerca entro cui vengono individuate le coppie di eventi potenzialmente vicine, consentendo di ridurre significativamente il numero di confronti e il costo computazionale dell'analisi.

Successivamente, per ciascuna coppia preselezionata, viene calcolata la distanza euclidea reale tra i centroidi degli eventi, espressa in metri. La classificazione finale delle interazioni multi-hazard non dipende quindi dalla geometria del buffer, ma esclusivamente dalla distanza reale tra gli eventi e dal loro ritardo temporale (Figura 5).

Ne consegue che un evento può ricadere all'interno del buffer di ricerca senza essere necessariamente classificato come interazione, qualora la distanza reale tra i centroidi non rientri nelle soglie spazio-temporali definite. Il buffer rappresenta pertanto unicamente uno strumento di selezione preliminare, mentre la validazione e la tipologia di interazione sono determinate da criteri quantitativi basati su distanza e tempo.

SOGLIE SPAZIO-TEMPORALI LANDSLIDE FLOOD



Classificazione	Ritardo temporale (Δt)	Distanza spaziale (d)
Compound	$\Delta t \leq 1$ giorno	$d \leq 250$ m
Simultaneous	$\Delta t \leq 3$ giorni	100 m $< d \leq 2000$ m
Consecutive	$2 < \Delta t \leq 10$ giorni	$d \leq 1000$ m

Figura 5– Classificazione delle interazioni multi-hazard è basata sulla combinazione tra distanza spaziale (d) e ritardo temporale (Δt), mentre il buffer rappresenta una fase preliminare di selezione.

5.3 HAZARD CONSIDERATI

L'inventario multi-hazard costruito nel presente lavoro integra tre principali tipologie di hazard fisici:

➤ Frane

➤ Alluvioni

➤ incendi boschivi

A queste si aggiungono le ondate di calore, considerate come hazard di natura climatica. In particolare, le heatwaves derivano da dataset climatici grigliati (ad esempio prodotti Copernicus basati su dati ERA5), nei quali il territorio è suddiviso in celle regolari di dimensioni dell'ordine di 25–30 km per lato (Figura 4).

Secondo la definizione adottata nello *European State of the Climate 2023*, una ondata di calore è identificata come un periodo di almeno tre giorni consecutivi in cui sia la temperatura minima sia la

temperatura massima giornaliera della superficie dell'aria superano il 95° percentile dei valori relativi allo stesso giorno nel periodo di riferimento 1991–2020.

Tale definizione evidenzia come le ondate di calore rappresentino fenomeni climatici continui nello spazio e persistenti nel tempo, a differenza degli hazard fisici discreti e localizzati (Figura 2). Per questa ragione, e per le differenti modalità di rappresentazione spaziale degli eventi, esse sono state trattate come dataset separato rispetto agli hazard fisici principali.

5.4 COSTRUZIONE DEL DATASET DEGLI EVENTI FISICI

La prima fase della costruzione dell'inventario ha riguardato l'integrazione dei dataset relativi agli hazard fisici. I dataset delle frane, alluvioni e degli incendi boschivi sono stati inizialmente trattati separatamente e successivamente integrati in un unico database nazionale.

Per ciascun dataset sono state eseguite operazioni di:

➤ armonizzazione dei sistemi di riferimento geografico;

➤ uniformazione dei campi attributo;

➤ standardizzazione delle informazioni temporali;

➤ conversione delle geometrie in centroidi puntuali.

Al termine di questa fase è stato possibile costruire un inventario nazionale degli eventi fisici, costituito complessivamente da:

➤ 82.986 eventi franosi

➤ 17.733 eventi alluvionali

➤ 8.381 incendi boschivi

per un totale di circa 109.100 eventi naturali.

Ciascun evento è stato rappresentato attraverso una struttura dati comune costituita da:

➤ ID evento

➤ Data

➤ tipologia di hazard

➤ geometria puntuale

Questa struttura rappresenta il formato base dell'inventario multi-hazard utilizzato nelle successive fasi di analisi.

5.5 INTEGRAZIONE DELLE ONDATE DI CALORE

Oltre agli hazard fisici principali, il lavoro include come già accennato un dataset relativo alle ondate di calore, derivato da un archivio climatico contenente eventi osservati nel periodo 2018–2024. Il dataset comprende complessivamente 69.585 eventi distribuiti sul territorio italiano.

A differenza degli altri hazard considerati, le ondate di calore sono state trattate separatamente in quanto rappresentano fenomeni di natura climatica, caratterizzati da una distribuzione spaziale continua e generalmente descritti tramite dati grigliati (raster), piuttosto che da eventi discreti e localizzati. Questa differenza implica una diversa modalità di rappresentazione e di interpretazione rispetto agli hazard geomorfologici e idrologici, tipicamente descritti tramite geometrie puntuali o areali.

Nonostante questa distinzione, il dataset delle ondate di calore è stato integrato nel flusso di lavoro complessivo attraverso una fase di armonizzazione spaziale e temporale, finalizzata a renderlo confrontabile con gli altri hazard. In questo contesto, l'inclusione delle ondate di calore consente di analizzare possibili relazioni multi-hazard, in particolare con gli incendi boschivi, per i quali esiste una chiara connessione con condizioni di temperatura elevata e prolungata.

5.6 DIMENSIONE COMPLESSIVA DEL DATASET

Considerando l'insieme degli hazard analizzati, il flusso di lavoro ha comportato la gestione complessiva di circa 178.685 eventi naturali, distribuiti tra hazard fisici e hazard climatici.

La costruzione dell'inventario multi-hazard rappresenta quindi il passaggio fondamentale che consente di trasformare dataset eterogenei di eventi naturali in una base dati integrata, strutturata per l'analisi delle relazioni spazio-temporali tra fenomeni naturali differenti.

Questa base dati costituisce il punto di partenza per le successive fasi metodologiche del lavoro, che comprendono la definizione di soglie spazio-temporali e l'implementazione di procedure automatiche

di analisi tramite strumenti GIS e linguaggi di programmazione per l'identificazione delle interazioni multi-hazard.

5.7 DEFINIZIONE DELLE SOGLIE SPAZIO-TEMPORALI

L'identificazione delle interazioni multi-hazard richiede la traduzione delle categorie teoriche discusse nella letteratura in criteri operativi applicabili ai dati osservati. In questo lavoro, le relazioni tra eventi appartenenti a hazard differenti sono state individuate sulla base della loro vicinanza spaziale e della loro relazione temporale, definite rispettivamente tramite uno space lag e un time lag. Tale impostazione è coerente con gli approcci recenti sviluppati nell'ambito degli studi multi-hazard, nei quali gli eventi non vengono considerati semplicemente come co-occorrenze generiche, ma come configurazioni spazio-temporali identificabili attraverso regole esplicite di sovrapposizione o prossimità. Claassen et al. sottolineano infatti l'importanza di introdurre una dimensione temporale esplicita, inclusi i time-lag tra eventi, nella costruzione degli event sets multi-hazard. Ferrario et al., in un contesto diverso ma metodologicamente affine, mostrano inoltre come le interazioni multi-hazard possano essere descritte mediante footprint spazio-temporali e funzioni di overlap tra eventi estremi.

Nel presente studio, le categorie concettuali di **compound**, **simultaneous** e **consecutive** events sono state quindi trasformate in criteri quantitativi applicati alle coppie di eventi presenti nell'inventario multi-hazard.

Questa classificazione è stata implementata nel flusso di lavoro GIS–Python attraverso soglie definite a priori e applicate in modo sistematico alle coppie di eventi appartenenti a diverse tipologie di hazard.

5.7.1 SOGLIE ADOTTATE PER LA COMBINAZIONE LANDSLIDE-FLOOD

Per la coppia frana–alluvione sono state definite tre classi di interazione.

Gli eventi sono stati classificati come compound quando risultano molto prossimi nello spazio e sostanzialmente contemporanei nel tempo. Nel codice utilizzato, questa condizione è soddisfatta quando:

➤ la distanza spaziale tra i due eventi è minore o uguale a 250 m;

➤ il time lag è minore o uguale a 1 giorno.

Formalmente:

$$d \leq 250 \text{ m}, \Delta t \leq 1 \text{ giorno}$$

Questa soglia identifica casi in cui frana e alluvione possono essere considerati parte della stessa configurazione spaziale e temporale, ad esempio in occasione di uno stesso evento pluviometrico intenso che produce fenomeni di esondazione e instabilità di versante nello stesso settore territoriale.

Gli eventi sono stati classificati come simultaneous quando si verificano in un intervallo temporale molto ravvicinato, ma non nello stesso punto. In questo caso, la condizione adottata è:

➤ distanza spaziale maggiore di 100 m e minore o uguale a 2000 m;

➤ time lag minore o uguale a 3 giorni.

Formalmente:

$$100 < d \leq 2000 \text{ m}, \Delta t \leq 3 \text{ giorni}$$

Questa classe rappresenta eventi che possono essere ricondotti allo stesso contesto meteorologico o idrologico, pur non presentando una sovrapposizione spaziale diretta.

Infine, gli eventi sono stati classificati come consecutive quando risultano relativamente vicini nello spazio ma separati da un ritardo temporale più marcato. In questo caso la soglia adottata è:

➤ distanza spaziale minore o uguale a 1000 m;

➤ time lag maggiore di 2 giorni e minore o uguale a 10 giorni.

Formalmente:

$$d \leq 1000 \text{ m}, 2 < \Delta t \leq 10 \text{ giorni}$$

Questa configurazione consente di intercettare sequenze di eventi in cui un primo fenomeno può aver modificato temporaneamente le condizioni del territorio, favorendo il verificarsi del secondo in un intervallo temporale successivo.

5.7.2 SOGLIE ADOTTATE PER LA COMBINAZIONE HEATWAVE-WILDFIRE

Per la combinazione ondata di calore–incendio boschivo sono state definite soglie differenti, coerenti con la maggiore estensione spaziale delle heatwaves e con il loro carattere temporale prolungato.

La classe compound è stata assegnata quando:

➤ la distanza spaziale tra heatwave e wildfire è minore di 1000 m;

➤ il time lag è minore di 3 giorni.

Formalmente:

$$d<1000 \text{ m}, \Delta t<3 \text{ giorni}$$

Questa soglia tiene conto del fatto che le ondate di calore, essendo derivate da dati climatici a griglia, hanno una risoluzione spaziale più ampia rispetto agli eventi fisici puntuali; la soglia di 1000 m è stata quindi assunta come valore coerente con la scala spaziale del dato climatico e con la necessità di rappresentare un'effettiva prossimità tra condizioni di caldo estremo e incendio. Nella presentazione metodologica del lavoro, tale valore viene infatti esplicitamente collegato alla risoluzione del dataset climatico utilizzato.

La classe simultaneous è stata assegnata quando:

➤ la distanza spaziale è maggiore di 1000 m e minore di 10000 m;

➤ il time lag è minore di 3 giorni.

Formalmente:

$$1000<d<10000 \text{ m}, \Delta t<3 \text{ giorni}$$

Questa condizione descrive incendi e ondate di calore che si verificano nello stesso periodo ma all'interno di un'area geografica più ampia, e che possono quindi essere interpretati come espressione dello stesso contesto atmosferico favorevole.

La classe consecutive è stata infine assegnata quando:

➤ la distanza spaziale è minore di 10000 m;

➤ il time lag è maggiore di 3 giorni e minore di 7 giorni.

Formalmente:

$$d<10000 \text{ m}, 3<\Delta t<7 \text{ giorni}$$

Questa soglia consente di identificare casi in cui una ondata di calore e un incendio non risultano sovrapposti in senso stretto, ma sono comunque sufficientemente vicini nello spazio e nel tempo da poter essere considerati parte di una sequenza multi-hazard.

È fondamentale evidenziare che le soglie spazio-temporali adottate nel presente lavoro non rappresentano una trasposizione rigorosa e letterale delle definizioni teoriche di compound, **simultaneous** e **consecutive** events, ma costituiscono una loro reinterpretazione operativa, resa necessaria dalla struttura e dai limiti intrinseci dei dataset utilizzati.

In termini teorici, tali categorie descrivono configurazioni in cui gli eventi avvengono nello stesso luogo e nello stesso intervallo temporale (**compound**), nello stesso momento ma in aree differenti (**simultaneous**), oppure nello stesso luogo con un ritardo temporale (**consecutive**). Tuttavia, queste definizioni, pur essendo concettualmente chiare, risultano difficilmente applicabili in modo diretto a dataset osservazionali reali, in cui la distribuzione spazio-temporale degli eventi è fortemente condizionata dalla modalità di costruzione degli inventari.

Nel caso specifico delle frane e delle alluvioni, l'analisi delle date uniche ha evidenziato una concentrazione temporale estremamente elevata: a fronte di circa 82.000 frane e 17.864 alluvioni, sono presenti rispettivamente soltanto 15 e 17 date uniche. Questo implica che migliaia di elementi spaziali risultano associati allo stesso giorno, in quanto derivanti da un singolo evento meteorologico estremo. Ad esempio, nel dataset delle frane, un'unica data (01/05/2023) concentra oltre 80.000 eventi, evidenziando chiaramente la natura event-based del database.

In una situazione di questo tipo, l'applicazione rigorosa delle definizioni teoriche che richiederebbero una corrispondenza quasi esatta tra eventi nello spazio e nel tempo avrebbe portato a risultati fortemente limitati o distorti. In particolare, la scarsa variabilità temporale avrebbe reso estremamente rara l'individuazione di relazioni spazio-temporali che soddisfino simultaneamente criteri molto restrittivi sia in termini di distanza sia di intervallo temporale.

Per questo motivo, si è reso necessario introdurre soglie operative più flessibili, in grado di mantenere il significato concettuale delle categorie di interazione, ma adattandolo alla reale distribuzione dei dati. In altre parole, le definizioni teoriche sono state utilizzate come riferimento concettuale, mentre le soglie sono state calibrate in funzione della struttura temporale e spaziale degli inventari disponibili.

Un comportamento opposto si osserva nel caso delle heatwaves e dei wildfire, che presentano una distribuzione temporale molto più ampia e continua, con rispettivamente 995 e 1.420 date uniche. In questo caso, la maggiore densità temporale consente una più elevata probabilità di sovrapposizione entro le soglie definite, rendendo l'individuazione delle interazioni statisticamente più frequente. Tuttavia, anche in questo contesto è stato necessario adattare le soglie spaziali alla natura del dato climatico, caratterizzato da una risoluzione grigliata di circa 25–27 km.

Nel complesso, la reinterpretazione delle soglie non rappresenta quindi una deviazione arbitraria dalle definizioni teoriche, ma una scelta metodologica necessaria per garantire la coerenza tra modello e dati. Le soglie adottate devono essere intese come parametri operativi dataset-dependent, costruiti a partire dal quadro concettuale della letteratura multi-hazard, ma calibrati sulla base della struttura reale delle informazioni disponibili, al fine di evitare una sottostima sistematica delle interazioni e rendere l'analisi effettivamente applicabile.

5.8 IMPLEMENTAZIONE OPERATIVA DELLE SOGLIE

Dal punto di vista operativo, la classificazione è stata implementata nel codice Python attraverso funzioni dedicate che ricevono in input, per ciascuna coppia di eventi, la distanza spaziale tra i centroidi e il time lag tra gli eventi. Per ridurre il numero di confronti possibili e limitare il costo computazionale, è stato preliminarmente applicato un buffer massimo sulle coppie di hazard analizzate: pari a 2500 m per le coppie landslide–flood e a 10000 m per le coppie heatwave–wildfire. Solo le coppie che ricadevano entro questi limiti massimi sono state successivamente valutate rispetto alle soglie più specifiche di classificazione (Figura 5).

Nel caso in cui una coppia soddisfacesse più di una condizione, è stata applicata una gerarchia di priorità che assegna precedenza a:

Compound→Simultaneous→Consecutive

Questa scelta evita ambiguità classificative e consente di attribuire a ciascuna coppia una sola categoria finale di interazione.

5.9 SIGNIFICATO E LIMITI DELLE SOGLIE ADOTTATE

È importante sottolineare che le soglie spazio-temporali adottate in questo studio non devono essere interpretate come valori universali validi per ogni contesto multi-hazard. Al contrario, esse rappresentano scelte operative, definite sulla base della letteratura di riferimento, delle caratteristiche fisiche dei fenomeni analizzati e della struttura spaziale e temporale dei dataset impiegati. Claassen et al. evidenziano infatti che l'introduzione di differenti time-lag modifica in modo sostanziale il numero e la tipologia di eventi multi-hazard identificati, mostrando come i risultati siano sensibili alle soglie adottate

Nel presente lavoro, la definizione delle soglie ha quindi avuto un duplice obiettivo: da un lato, rendere possibile l'identificazione automatica delle interazioni all'interno di un database nazionale molto esteso; dall'altro, mantenere una coerenza fisica e interpretativa rispetto ai fenomeni analizzati. Di conseguenza, il numero di interazioni individuate deve essere interpretato come risultato del framework metodologico adottato: soglie più permissive avrebbero probabilmente aumentato il numero di correlazioni identificate, mentre soglie più restrittive avrebbero portato a una selezione più limitata di eventi.

In questo senso, la definizione delle soglie spazio-temporali costituisce uno dei passaggi metodologici più delicati dell'intero lavoro, poiché rappresenta il punto di collegamento tra la teoria multi-hazard e la sua traduzione in un sistema automatico di identificazione delle relazioni tra eventi.

6. ANALISI E IMPLEMENTAZIONE TRAMITE GIS E PYTHON

L'analisi è stata implementata tramite codice Python (Appendice A)

L'identificazione automatica delle interazioni multi-hazard è stata sviluppata attraverso un workflow integrato basato sull'uso combinato di strumenti GIS per la preparazione dei dati geospaziali e di procedure automatiche implementate in Python nell'ambiente Google Colab. Questa scelta ha consentito di strutturare un flusso di lavoro riproducibile, in grado di gestire dataset eterogenei e di applicare in modo sistematico le soglie spazio-temporali definite nel paragrafo precedente.

6.1 INPUT

Gli input principali dell'analisi sono costituiti da due dataset:

- ➤ il file vettoriale in formato geopackage, contenente l'inventario nazionale degli hazard fisici considerati nello studio, ovvero frane, alluvioni e incendi boschivi;
- ➤ il file tabellare heatwaves_2018_2024.csv, contenente gli eventi di ondata di calore con informazioni geografiche e temporali.

Il file geopackage rappresenta il dataset principale del workflow e contiene, per ciascun evento, almeno le seguenti informazioni: un identificativo univoco (id), la tipologia di hazard (event_type), la data dell'evento (DATE) e, nel caso delle frane, un campo relativo al meccanismo di innesco (Triggering). Le geometrie sono rappresentate in forma puntuale, corrispondente al centroide dell'evento.

Il file CSV delle heatwaves contiene invece le coordinate geografiche dei centroidi degli eventi climatici, la data di inizio (start_date) e la durata (duration). Le ondate di calore sono quindi trattate come un dataset separato, successivamente convertito in formato geospaziale all'interno dell'ambiente Python.

6.2 PRE-PROCESSING

Prima dell'analisi delle interazioni, i dati sono stati sottoposti a una fase di pre-processing finalizzata a renderli coerenti dal punto di vista geometrico, temporale e semantico.

Per il file MULTI_HAZARDS_CORRECT_FINAL_CLEAN.gpkg, il codice verifica innanzitutto la presenza della colonna identificativa id; quando assente o incompleta, essa viene ricostruita a partire dall'indice del GeoDataFrame. Successivamente, i campi testuali vengono normalizzati trasformando i valori in minuscolo e rimuovendo eventuali spazi o incongruenze ortografiche, in particolare nel campo Triggering, così da evitare errori di classificazione dovuti a differenze di scrittura.

Le date vengono convertite in formato datetime mediante la libreria Pandas, utilizzando il campo DATE. I record privi di data valida vengono esclusi dall'analisi. Dal punto di vista spaziale, tutti i dati vengono riproiettati nel sistema di riferimento metrico EPSG:32632, in modo da garantire che i calcoli di distanza siano espressi in metri.

Un’analoga procedura viene applicata al dataset delle heatwaves. Il file CSV viene letto con Pandas, le date di inizio vengono convertite in formato datetime, la durata dell’evento viene trasformata in valore numerico e vengono mantenuti soltanto gli eventi con durata pari o superiore a tre giorni. Le coordinate geografiche vengono poi convertite in punti tramite GeoPandas e anche in questo caso il dataset viene riproiettato in EPSG:32632.

Dopo questa fase, i dati vengono suddivisi in quattro sottoinsiemi distinti:

➤ frane (ls)

➤ alluvioni (fl)

➤ incendi boschivi (wf)

➤ ondate di calore (hw)

Poiché il confronto esaustivo tra tutti gli eventi presenti nei dataset avrebbe generato un numero molto elevato di combinazioni, incompatibile con le risorse di calcolo disponibili in Google Colab, è stato applicato un campionamento casuale di 2500 eventi per ciascuna tipologia di hazard. Il campionamento è stato effettuato con random_state=0, così da rendere il processo riproducibile.

6.3 LOGICA DELL’ANALISI

La logica generale del workflow consiste nel generare coppie di eventi appartenenti a hazard differenti e verificare, per ciascuna coppia, se la relazione tra distanza spaziale e distanza temporale soddisfa le soglie definite per una delle tre classi di interazione: Compound, Simultaneous o Consecutive.

Il confronto non viene effettuato tra tutti gli eventi indistintamente, ma solo tra coppie di hazard considerate rilevanti nel presente lavoro, in particolare:

➤ landslide–flood

➤ heatwave–wildfire

Per ridurre il numero di confronti e migliorare l’efficienza computazionale, il codice applica dapprima un filtro spaziale preliminare mediante la costruzione di un buffer attorno agli eventi del secondo

dataset della coppia. Solo gli eventi che ricadono entro questo buffer vengono considerati come candidati per il calcolo dell'interazione.

Una volta individuate le coppie candidate, vengono calcolati due parametri fondamentali:

➤ la distanza spaziale tra i centroidi dei due eventi;

➤ il time lag, espresso come differenza assoluta in giorni tra le date di inizio dei due eventi.

Sulla base di questi due valori, ogni coppia viene classificata in una delle classi previste. Se una coppia non soddisfa nessuna delle soglie definite, viene esclusa dall'output finale.

Nel caso delle coppie landslide–flood, dopo la classificazione viene inoltre recuperata l'informazione relativa al campo Triggering delle frane, così da distinguere le frane a innesco pluviale da quelle a innesco sismico e attribuire di conseguenza il corretto parent_hazard.

6.4 FUNZIONI IMPLEMENTATE

L'implementazione Python si basa su tre componenti principali.

La prima è costituita dalle funzioni di classificazione:

➤ classify_ls_fl(), dedicata alla classificazione delle coppie frana–alluvione;

➤ classify_hw_wf(), dedicata alla classificazione delle coppie ondata di calore–incendio boschivo.

Queste funzioni ricevono in input la distanza spaziale e il time lag tra due eventi e restituiscono la categoria finale di interazione (Compound, Simultaneous, Consecutive oppure None), applicando le soglie definite a priori.

La seconda componente è la funzione generale calculate_pairs(), che rappresenta il nucleo operativo del workflow. Tale funzione è stata progettata in forma generica, così da poter essere riutilizzata per diverse combinazioni di hazard. In particolare, essa:

- seleziona e rinomina i campi necessari dei due dataset;
- costruisce un buffer attorno agli eventi del secondo dataset;
- esegue una spatial join per individuare le coppie candidate;
- recupera la geometria reale degli eventi per il calcolo della distanza effettiva;
- calcola il time lag tra le date di inizio;

- applica la funzione di classificazione;
- produce un output strutturato contenente gli attributi delle coppie riconosciute.

La terza componente è la procedura finale di consolidamento, che concatena tutte le coppie valide in un unico GeoDataFrame e lo esporta in formato GeoPackage.

6.5 OUTPUT

Il risultato finale del workflow è Il risultato finale del workflow è un file geospaziale vettoriale in formato GeoPackage contenente le coppie di eventi classificate come interazioni multi-hazard. Ogni record dell'output rappresenta quindi una relazione evento–evento, e non un singolo evento isolato.

Per ciascuna interazione identificata, il file finale include:

- l'identificativo dei due eventi coinvolti (id_A, id_B);
- la tipologia di hazard dei due eventi (hazard_A, hazard_B);
- il parent_hazard e l'interacting_hazard;
- la categoria finale di interazione (Interaction);
- un identificativo univoco della coppia (pair_id);
- le date associate ai due eventi (DATE_A, DATE_B);
- la distanza spaziale (dist_m);
- il time lag (time_lag_days);
- la geometria puntuale dell'evento A.

L'output viene salvato in formato GeoPackage nel file MULTI_HAZARDS_pairs_SAMPLE2500_UPDATED.gpkg, così da poter essere successivamente riutilizzato sia per l'analisi dei risultati sia per eventuali sviluppi futuri, come procedure di clustering spazio-temporale o applicazioni di machine learning.

Nel complesso, l'implementazione sviluppata dimostra che è possibile costruire un workflow automatico e riproducibile per l'identificazione di interazioni multi-hazard a partire da dataset geospaziali eterogenei, traducendo le categorie teoriche della letteratura in un sistema operativo di analisi applicabile a casi reali.

7. RISULTATI E DISCUSSIONI

Il presente capitolo riporta i risultati dell'analisi multi-hazard ottenuti mediante l'applicazione automatizzata, in ambiente Python su Google Colab, delle soglie spazio-temporali definite a priori al dataset degli eventi naturali.

In particolare, attraverso il codice sviluppato (Appendice A), sono state identificate e classificate le relazioni tra eventi appartenenti a differenti tipologie di hazard, sulla base della loro prossimità spaziale e della loro distanza temporale.

L'analisi è stata condotta su un campione di 2500 eventi randomici per ciascun hazard, permettendo di individuare complessivamente 1955 interazioni multi-hazard.

I risultati quantitativi evidenziano una forte asimmetria tra le diverse tipologie di interazione e tra le combinazioni di hazard considerate. Su un totale di 1955 interazioni multi-hazard identificate, la quasi totalità appartiene alla classe **Simultaneous** (1934 relazioni), mentre le configurazioni **Consecutive** risultano estremamente limitate (20 relazioni) e gli eventi **Compound** sono praticamente assenti (1 solo caso).

Analizzando nel dettaglio la distribuzione per tipologia di hazard, emerge che le interazioni **Simultaneous** sono dominate quasi interamente dalla combinazione landslide–flood, che rappresenta 1884 relazioni, mentre solo 50 relazioni sono associate alla combinazione heatwave–wildfire. Una dinamica analoga si osserva per le interazioni **Consecutive**, dove 19 relazioni riguardano frane e alluvioni e un solo caso è associato alla combinazione heatwave–wildfire. Al contrario, l'unica interazione classificata come **Compound** appartiene alla categoria heatwave–wildfire.

Questa distribuzione evidenzia chiaramente come i risultati siano fortemente influenzati dalla natura dei dataset utilizzati. In particolare, la predominanza delle relazioni landslide–flood riflette sia la maggiore numerosità degli eventi disponibili sia, soprattutto, la loro forte concentrazione temporale all'interno di pochi episodi estremi. Come discusso in precedenza, una parte rilevante delle frane è associata all'evento dell'Emilia-Romagna del 1° maggio 2023, mentre molte delle alluvioni sono collegate agli eventi del 5 maggio 2023, appartenenti allo stesso contesto meteorologico. Questa configurazione genera un elevato numero di relazioni caratterizzate da una distanza temporale molto ridotta, favorendo la classificazione nella categoria **simultaneous.**

Al contrario, la combinazione heatwave–wildfire mostra un numero molto più contenuto di interazioni, ma una maggiore variabilità nella distribuzione temporale, che consente l'individuazione anche di configurazioni consecutive e, in un caso, compound. Questo comportamento riflette la natura intrinsecamente diversa degli hazard climatici rispetto a quelli pluviometrici evento-specifici.

Nel complesso, i risultati confermano come la classificazione delle interazioni multi-hazard non dipenda esclusivamente dalle dinamiche fisiche dei fenomeni naturali, ma sia fortemente condizionata dalla struttura spazio-temporale dei dataset e dalle scelte metodologiche adottate. In particolare, la reinterpretazione operativa delle soglie ha permesso di adattare le definizioni teoriche alle caratteristiche dei dati disponibili, rendendo possibile l'identificazione di pattern multi-hazard significativi, pur nel rispetto dei limiti intrinseci dell'approccio.

La rappresentazione delle interazioni multi-hazard pone una criticità rilevante dal punto di vista visuale. Ogni relazione identificata dall'algoritmo è infatti rappresentata tramite un segmento che connette i centroidi dei due eventi coinvolti, il cui colore varia in funzione della tipologia di interazione (**compound, simultaneous, consecutive**).

Tuttavia, a causa dell'elevato numero di relazioni individuate, la visualizzazione diretta di tutti i segmenti su scala nazionale risulta fortemente caotica e di difficile interpretazione. Come mostrato in Figura 6.

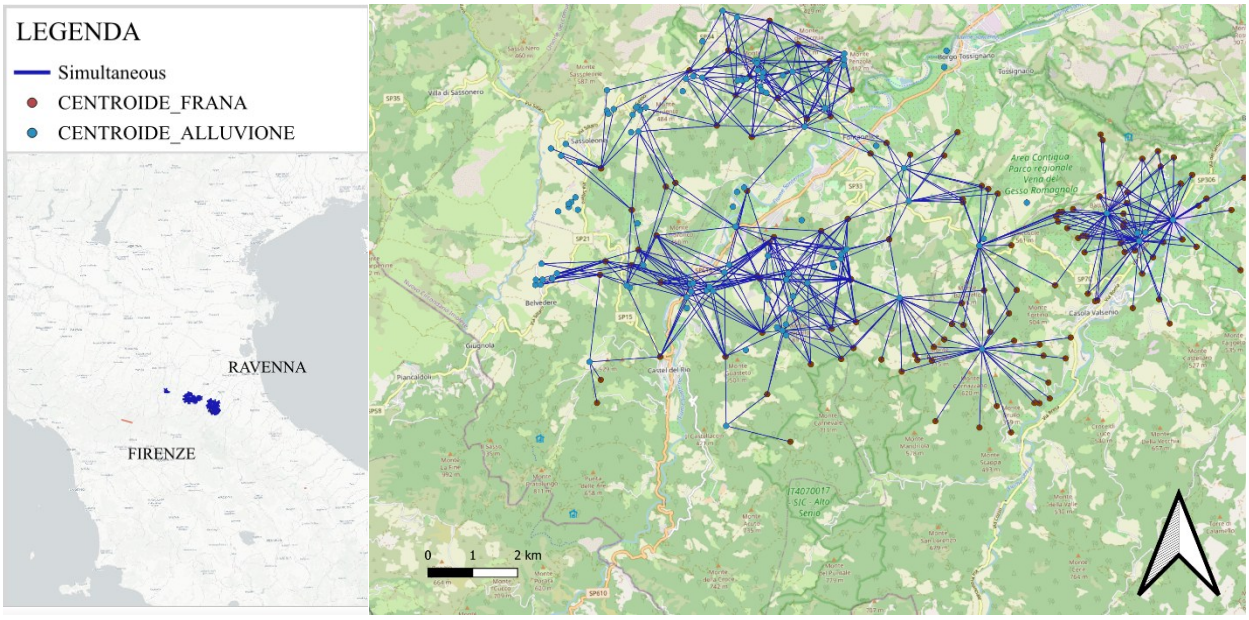


Figura 6 – Visualizzazione tramite GIS delle relazioni multi-hazard generate dall'algoritmo.

La figura mostra una visualizzazione grezza dell'output prodotto dall'algoritmo per le relazioni di tipo **simultaneous** nella combinazione frana–alluvione. In questa rappresentazione, ogni interazione è espressa da un segmento che collega il centroide della frana al centroide dell'alluvione; i punti marroni rappresentano i centroidi delle frane, mentre i punti azzurri rappresentano i centroidi delle alluvioni. Il segmento costituisce quindi la vera unità geometrica della relazione, mentre il colore identifica la classe di interazione assegnata dall'algoritmo.

Dal punto di vista teorico, la categoria **simultaneous** descrive eventi che si verificano nello stesso intervallo temporale, ma non necessariamente nello stesso punto dello spazio. Nella letteratura, essa indica quindi una contemporaneità temporale accompagnata da una separazione spaziale, purché gli eventi appartengano allo stesso contesto generale. Nel presente studio, tuttavia, tale definizione è stata reinterpretata operativamente per renderla applicabile a dataset osservazionali eterogenei e rappresentati tramite centroidi. Per la coppia landslide–flood, una relazione è classificata come **simultaneous** quando la distanza tra i centroidi è maggiore di 100 m e minore o uguale a 2000 m e il *time lag* è minore o uguale a 3 giorni. Il buffer di 2500 m non rappresenta il criterio finale di classificazione, ma solo una fase preliminare di selezione delle coppie candidate.

Questa reinterpretazione si è resa necessaria soprattutto per la struttura temporale dei dataset di frane e alluvioni. Il dataset delle frane è stato costruito integrando inventari relativi a tre eventi pluviometrici estremi recenti, tra cui l'evento dell'Emilia-Romagna del maggio 2023 descritto da Berti et al.; il dataset delle alluvioni deriva invece dai layer Observed Event del servizio Copernicus EMS, selezionando le principali attivazioni EMSR in Italia. Si tratta quindi, in entrambi i casi, di dataset fortemente event-based, in cui moltissimi elementi spaziali sono associati a poche date evento-specifiche.

Nel caso specifico di frane e alluvioni, questa concentrazione temporale è estrema: a fronte di circa 82.000 frane e 17.864 alluvioni, il database contiene rispettivamente solo 15 e 17 date uniche. Inoltre, nel dataset delle frane, una sola data, il 01/05/2023, concentra oltre 80.000 eventi, evidenziando chiaramente la natura evento-specifica dell'inventario. In un contesto di questo tipo, un'applicazione rigidamente letterale delle definizioni teoriche avrebbe prodotto risultati distorti o troppo limitati; per questo le soglie sono state calibrate in modo da mantenere il significato concettuale della contemporaneità, ma adattandolo alla distribuzione reale dei dati.

Nel maggio 2023 l'Emilia-Romagna è stata interessata da precipitazioni molto intense, che hanno prodotto contemporaneamente numerose aree allagate e una diffusissima instabilità di versante. In questo senso, la figura non va letta come l'indicazione di un rapporto causale diretto uno-a-uno tra

una specifica alluvione e una specifica frana, ma come la rappresentazione di un insieme di relazioni spazio-temporali tra fenomeni che appartengono allo stesso episodio idrometeorologico regionale. La forte densità di segmenti che convergono su alcuni centroidi alluvionali suggerisce infatti che singoli eventi di alluvione possano risultare associati a molte frane vicine nel tempo e nello spazio, generando configurazioni uno-a-molti e molti-a-molti. Questo è coerente con la logica del metodo, che classifica relazioni tra coppie di eventi, non proprietà intrinseche del singolo evento.

La figura 6 mostra però anche un limite evidente della rappresentazione lineare. Sebbene il segmento sia la geometria più corretta per rappresentare la relazione calcolata dall'algoritmo, la sovrapposizione di numerose connessioni rende la lettura difficile già a scala locale e quasi impossibile a scala nazionale. Per questo motivo, nelle fasi successive della rappresentazione cartografica si è reso necessario ricorrere a una semplificazione visiva, trasformando ciascun segmento in un punto medio, utilizzato esclusivamente per evidenziare le aree di concentrazione delle interazioni (Figura 9). Tale passaggio non modifica i risultati analitici, ma consente una lettura più chiara dei pattern spaziali.

In sintesi, la figura documenta bene tre aspetti fondamentali: il funzionamento dell'algoritmo, la natura reticolare delle relazioni **simultaneous** e la necessità di distinguere tra output analitico e rappresentazione cartografica finale. Il risultato osservato è quindi metodologicamente coerente: i segmenti mostrano le relazioni effettivamente identificate dal codice, mentre la loro elevata densità riflette la la forte concentrazione temporale dei dataset legati agli eventi estremi del maggio 2023.

Nell'immagine successiva viene presentata una visualizzazione più nel dettaglio (figura 7) delle relazioni, in cui un singolo evento (un'alluvione) è connesso tramite segmenti agli eventi correlati (frane), al fine di evidenziare il funzionamento del metodo e il significato delle diverse tipologie di interazione.

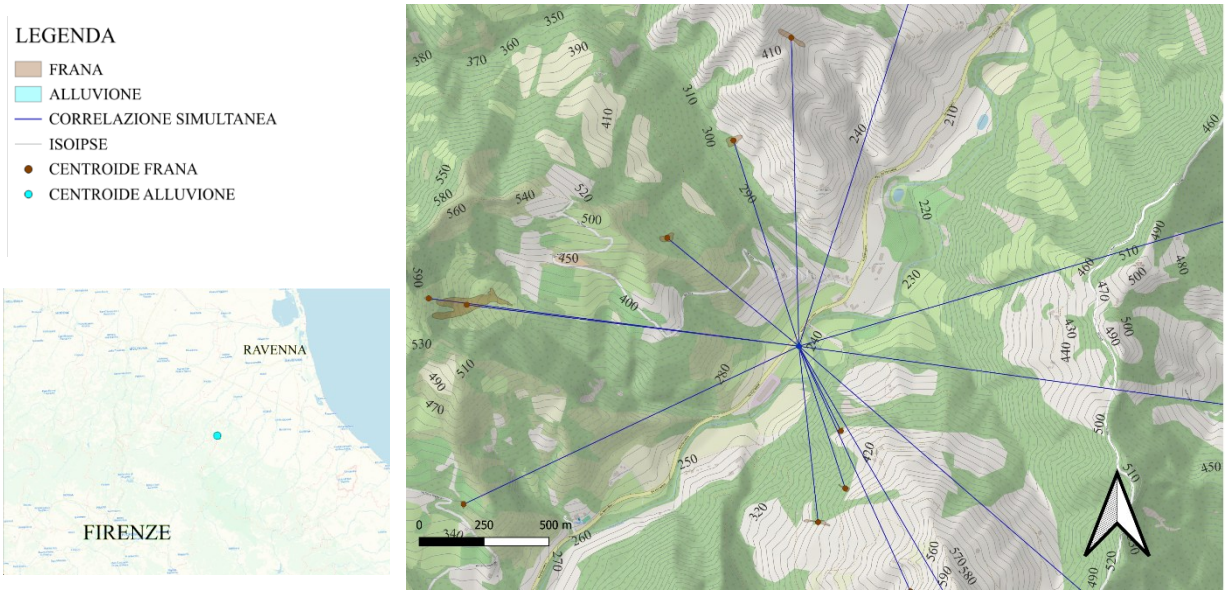


Figura 7 – Visualizzazione GIS delle relazioni multi-hazard generate dall'algoritmo.

La figura mostra un esempio locale di relazioni di tipo **simultaneous** tra alluvioni e frane in un settore appenninico romagnolo, nell'area di Fiumane(Modigliana FC). A differenza della rappresentazione precedente, qui la visualizzazione è proposta a scala più dettagliata e integrata con il modello digitale del terreno (DEM) e con le isoipse, al fine di evidenziare con maggiore chiarezza il contesto morfologico entro cui si sviluppano le interazioni identificate dall'algoritmo.

Dal punto di vista geomorfologico, la figura mostra quindi una relazione funzionale tra morfologia dei versanti, convergenza del drenaggio e localizzazione delle interazioni multi-hazard. Il corso d'acqua rappresenta il principale asse morfologico del sistema, mentre i versanti circostanti costituiscono i settori potenzialmente instabili da cui si originano i fenomeni franosi. In un contesto di piogge intense, la risposta del territorio può essere duplice: da un lato l'aumento del deflusso e l'espansione delle aree allagate nel fondovalle, dall'altro l'innesco di movimenti di versante in settori topograficamente predisposti.

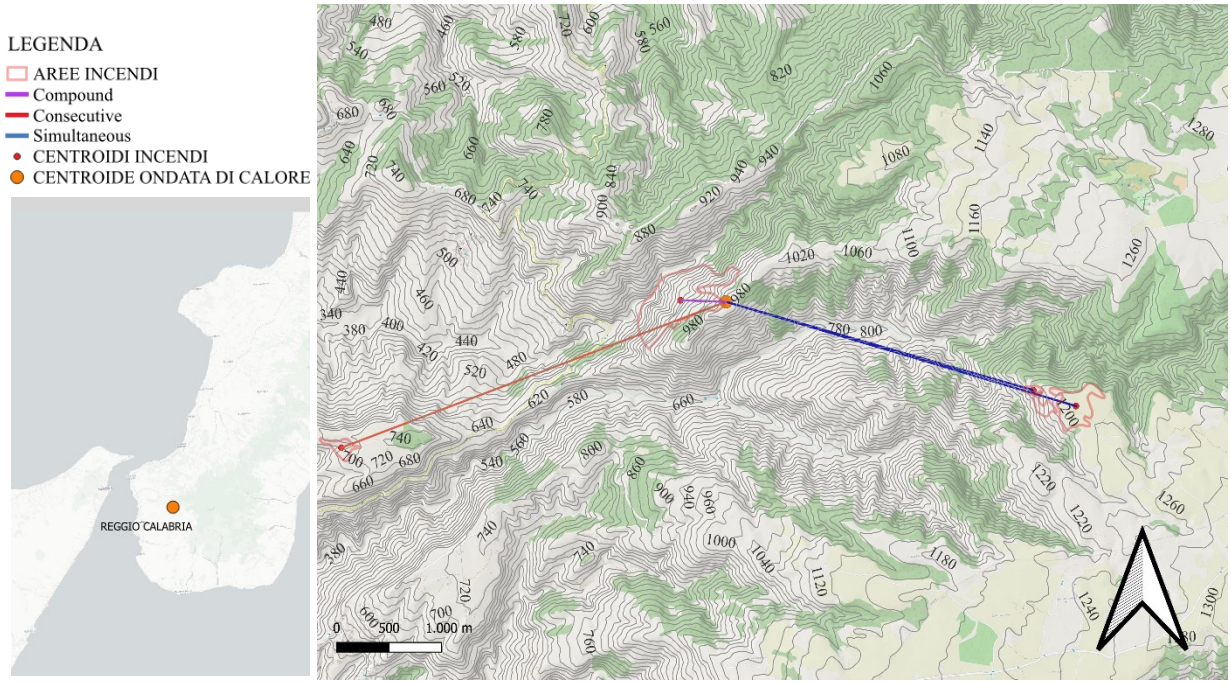


Figura 8 – Visualizzazione dettagliata tramite GIS delle relazioni tra ondate di calore e incendi boschivi generate dall'algoritmo.

La figura 8 rappresenta le relazioni spazio-temporali tra ondate di calore e incendi boschivi in un settore dell'Aspromonte, nel sud della Calabria, in prossimità di Reggio Calabria. L'area è caratterizzata da una morfologia montuosa e fortemente articolata, con versanti acclivi, incisioni vallive e coperture forestali tipiche dell'ambiente mediterraneo, condizioni che favoriscono la propagazione degli incendi in presenza di fattori climatici predisponenti.

Nella presente mappa vengono visualizzate tutte e 3 le tipologie di correlazione ovvero: compound, simultaneus e consecutive.

Il primo caso particolarmente significativo è rappresentato dalle relazioni classificate come **compound**. L'ondata di calore si estende dal 3 settembre 2024 all'8 settembre 2024, con una durata complessiva di sei giorni, mentre l'incendio si verifica il 7 settembre 2024, quindi all'interno dell'intervallo temporale dell'ondata di calore. La distanza tra i due eventi risulta pari a circa 436 metri, mentre il timelag è pari a zero. Questo significa che i due eventi non solo sono molto vicini nello spazio, ma si sovrappongono anche temporalmente. In questo caso, la classificazione come *compound* è pienamente coerente sia con la definizione teorica sia con una interpretazione fisica del fenomeno: l'incendio si sviluppa in condizioni di temperatura elevata persistente, con vegetazione già stressata e disidratata, configurando quindi una situazione in cui il contesto climatico contribuisce direttamente all'innesco e alla propagazione del fuoco.

Il secondo caso riguarda invece le relazioni di tipo *consecutive*. In questo caso, l'ondata di calore mantiene la stessa durata (3–8 settembre 2024), mentre l'incendio si verifica il 30 agosto 2024, quindi

alcuni giorni prima. La distanza spaziale è maggiore, pari a circa 3.9 km, e il timelag risulta pari a 4 giorni. In questo caso, gli eventi non sono contemporanei, ma si collocano comunque all'interno dello stesso contesto territoriale. La classificazione come *consecutive* riflette quindi una relazione temporale differita, che non implica necessariamente una causalità diretta, ma suggerisce una possibile sequenza di condizioni ambientali o climatiche che si sviluppano nel tempo.

Infine, il terzo caso è rappresentato dalle relazioni **simultaneous.** In questa situazione, l'ondata di calore si estende sempre dal 3 all'8 settembre 2024, mentre l'incendio si verifica l'8 settembre 2024, quindi in contemporaneità con l'evento climatico. Il time lag è pari a zero, ma la distanza spaziale è più elevata, circa 3.5 km. In questo caso, i due eventi sono contemporanei nel tempo, ma non coincidono spazialmente in modo stretto. La classificazione come **simultaneous** indica quindi che entrambi gli eventi sono inseriti nello stesso contesto climatico o meteorologico regionale, ma non rappresentano lo stesso episodio locale.

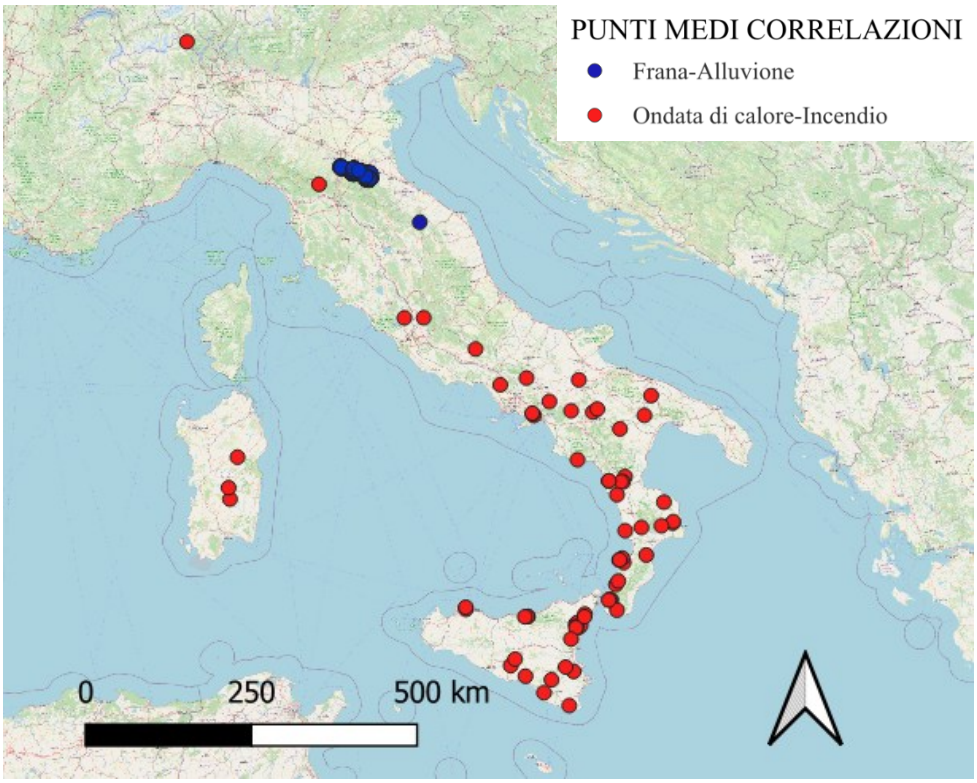


Figura 9-Distribuzione spaziale semplificata delle interazioni multi-hazard (frane–alluvioni e heatwave–wildfire) mediante punti medi delle correlazioni.

La figura 9 rappresenta, come già accennato una semplificazione dell'output del codice multi-hazard, ottenuta attraverso una rielaborazione in ambiente GIS mediante l'utilizzo dei punti medi (midpoint) dei segmenti di correlazione. In particolare, ogni punto rappresenta la posizione media di una relazione spazio-temporale tra due eventi, consentendo di sintetizzare visivamente un numero elevato di connessioni che, nella loro forma originale, risultavano difficilmente leggibili a causa dell'elevata densità di segmenti. Questa operazione di generalizzazione non altera i risultati dell'analisi, ma ha esclusivamente una funzione descrittiva e comunicativa, permettendo di evidenziare le principali aree di concentrazione delle interazioni multi-hazard sul territorio italiano.

Dalla distribuzione spaziale dei punti emergono due pattern principali. I punti blu, che rappresentano le correlazioni tra frane e alluvioni, risultano fortemente concentrati nel settore collinare vallivo romagnolo. Tale distribuzione è coerente con gli eventi estremi verificatisi nel maggio 2023, quando precipitazioni intense e persistenti hanno generato diffuse alluvioni e innescato numerosi fenomeni franosi, configurando un sistema multi-hazard fortemente controllato da un unico evento meteorologico di scala regionale.

Al contrario, i punti rossi, che rappresentano le correlazioni tra incendi boschivi e ondate di calore, mostrano una distribuzione più diffusa lungo l'Italia meridionale e insulare, con una maggiore concentrazione nelle regioni caratterizzate da condizioni climatiche più calde e aride. Questo pattern riflette il ruolo delle alte temperature e della persistenza delle ondate di calore nel favorire l'innesco e la propagazione degli incendi, evidenziando un controllo climatico di tipo termico piuttosto che pluviometrico.

Nel complesso, la figura evidenzia chiaramente come le diverse tipologie di interazione multi-hazard presentino una distribuzione spaziale distinta, legata ai differenti fattori di innesco: da un lato eventi idrometeorologici intensi e concentrati nel Nord Italia, dall'altro condizioni climatiche persistenti e diffuse nel Sud Italia.

8. CONCLUSIONI

Il presente lavoro ha avuto come obiettivo principale la costruzione di un inventario multi-hazard a scala nazionale e l'analisi delle relazioni spazio-temporali tra diversi fenomeni naturali in Italia, integrando dataset eterogenei relativi a frane, alluvioni, incendi boschivi e ondate di calore. Attraverso un approccio basato su dati osservati, elaborati mediante strumenti GIS e procedure automatiche implementate in ambiente Python, è stato possibile identificare e classificare le interazioni tra hazard secondo tre configurazioni principali: compound, simultaneous e consecutive.

Uno dei risultati più rilevanti dello studio riguarda la dimostrazione della possibilità di tradurre un quadro teorico complesso, sviluppato nella letteratura multi-hazard recente, in un sistema operativo basato su criteri quantitativi replicabili. In particolare, l'introduzione di soglie spazio-temporali ha consentito di trasformare concetti teorici in regole applicabili a un dataset nazionale di eventi, rendendo possibile l'identificazione automatica delle relazioni tra fenomeni naturali.

Tuttavia, un aspetto fondamentale emerso dall'analisi riguarda il ruolo determinante dei dataset di input nella distribuzione spaziale delle interazioni individuate. Nel caso delle relazioni tra frane e alluvioni, la forte concentrazione delle correlazioni nella fascia collinare e valliva del margine Appenninico emiliano romagnolo è direttamente legata alla natura event-based del dataset utilizzato, in particolare all'inventario di Berti et al., che documenta un numero molto elevato di frane associate all'evento pluviometrico estremo del maggio 2023. Analogamente, il dataset delle alluvioni, derivato dal Copernicus Emergency Management Service, rappresenta lo stesso evento meteorologico, determinando una sovrapposizione spazio-temporale intrinseca tra i due hazard.

Di conseguenza, le relazioni identificate non devono essere interpretate come una distribuzione naturale dei fenomeni su scala nazionale, ma come il risultato dell'interazione tra condizioni fisiche reali e struttura dei dataset utilizzati. Questo aspetto evidenzia un limite intrinseco degli approcci data-driven: l'analisi riflette inevitabilmente le caratteristiche e le scale dei dati disponibili.

Un comportamento differente emerge invece nell'analisi delle relazioni tra ondate di calore e incendi boschivi. In questo caso, la distribuzione spaziale delle interazioni mostra una maggiore diffusione sul territorio nazionale, con una chiara concentrazione nelle regioni centro-meridionali e insulari. Questo pattern è coerente sia con la distribuzione reale degli incendi, derivata dal sistema European Forest Fire Information System, sia con le condizioni climatiche tipiche delle aree mediterranee, caratterizzate da temperature elevate e periodi prolungati di siccità.

Un ulteriore aspetto rilevante riguarda la natura dataset-dependent delle soglie spazio-temporali adottate. Le soglie utilizzate non rappresentano valori universali, ma parametri operativi calibrati

sulla base delle caratteristiche dei dati disponibili. Questo implica che i risultati ottenuti sono sensibili alle scelte metodologiche e potrebbero variare significativamente in presenza di dataset differenti o di soglie più restrittive o più permissive.

Il metodo presenta alcune semplificazioni importanti. In particolare, l'utilizzo del buffer rappresenta una fase preliminare necessaria per ridurre il numero di confronti tra eventi, ma introduce una componente arbitraria: la scelta della distanza del buffer (ad esempio 2500 m o 10000 m) non deriva da un criterio fisico preciso, ma è una decisione operativa. Questo significa che cambiando il buffer cambiano anche le coppie di eventi analizzate, e quindi i risultati finali.

Un altro limite riguarda la dimensione del campione analizzato. In questo studio sono stati utilizzati 2500 eventi per ciascun hazard per motivi computazionali, ma un'analisi su dataset completi o più estesi potrebbe restituire risultati più robusti e rappresentativi.

Inoltre, la rappresentazione degli eventi tramite centroidi semplifica la realtà, in quanto non considera l'estensione reale dei fenomeni. Di conseguenza, le relazioni individuate devono essere interpretate come una rappresentazione semplificata e non come una descrizione completa dei processi fisici.

Nonostante questi limiti, l'approccio sviluppato rappresenta un primo passo utile per analisi più avanzate. In particolare, i risultati ottenuti possono essere utilizzati come base per tecniche di clustering, ovvero metodi che permettono di individuare automaticamente gruppi di eventi o aree in cui le interazioni multi-hazard si concentrano nello spazio e nel tempo.

Allo stesso modo, questo tipo di dataset può essere utilizzato in modelli di machine learning: algoritmi in grado di apprendere dai dati e riconoscere pattern complessi, con l'obiettivo di sviluppare modelli predittivi capaci di stimare dove e quando potrebbero verificarsi nuove interazioni multi-hazard.

9. APPENDICE

CODICE :

```
import geopandas as gpd
import pandas as pd
from pathlib import Path
import shutil
from google.colab import files
from shapely.geometry import LineString

# ------------------------------------------------------------
# 1. PATH E COSTANTI
# ------------------------------------------------------------
MULTI_HAZARDS_PATH = Path("/content/MULTI_HAZARDS_CORRECT_FINAL_CLEAN.gpkg")
HEATWAVES_CSV_PATH = Path("/content/heatwaves_2018_2024.csv")
OUTPUT_PAIRS_PATH  =
Path("/content/MULTI_HAZARDS_pairs_SAMPLE2500_UPDATED_REAL_DURATION_LINES.gpkg")

ID_COL         = "id"
HAZARD_COL     = "event_type"
DATE_COL       = "DATE"
TRIGGERING_COL = "Triggering"

NAME_LANDSLIDE = "landslide"
NAME_FLOOD     = "flood"
NAME_WILDFIRE  = "wildfire"

# ------------------------------------------------------------
# 2. SOGLIE
# ------------------------------------------------------------

# Landslide-Flood

# Landslide-Flood
LSFL_COMPOUND_DIST_MAX = 250
LSFL_SIMUL_DIST_MIN    = 100
LSFL_SIMUL_DIST_MAX    = 2000
LSFL_CONSEC_DIST_MAX   = 1000

LSFL_COMPOUND_TIME_MAX = 1
LSFL_SIMUL_TIME_MAX    = 3
LSFL_CONSEC_TIME_MIN   = 2
LSFL_CONSEC_TIME_MAX   = 10

BUFFER_LSFL = 2500

# Heatwave-Wildfire
HWWF_COMPOUND_DIST_MAX = 1000
HWWF_SIMUL_DIST_MIN    = 1000
HWWF_SIMUL_DIST_MAX    = 10000
HWWF_CONSEC_DIST_MAX   = 10000

BUFFER_HWWF = 10000

# ------------------------------------------------------------
# 3. FUNZIONE PER IL TIME LAG REALE TRA INTERVALLI
# ------------------------------------------------------------
def compute_time_lag_days(start_a, end_a, start_b, end_b):
    """
    Restituisce:
    - 0 se gli intervalli temporali si sovrappongono
```

```
    - altrimenti il numero di giorni tra la fine del primo e l'inizio del
secondo
    """
    if pd.isna(start_a) or pd.isna(end_a) or pd.isna(start_b) or pd.isna(end_b):
        return None

    # Se si sovrappongono
    if start_a <= end_b and start_b <= end_a:
        return 0

    # A finisce prima che inizi B
    if end_a < start_b:
        return (start_b - end_a).days

    # B finisce prima che inizi A
    if end_b < start_a:
        return (start_a - end_b).days

    return None

# ------------------------------------------------------------
# 4. CARICAMENTO MULTI-HAZARDS
# ------------------------------------------------------------
print("Caricamento Multi-Hazards...")
gdf = gpd.read_file(MULTI_HAZARDS_PATH)

# Salvo indice originale del file letto
gdf = gdf.reset_index(drop=False).rename(columns={"index": "orig_fid"})

# ID robusto
if ID_COL not in gdf.columns:
    gdf[ID_COL] = gdf.index
else:
    if gdf[ID_COL].isna().any():
        gdf[ID_COL] = gdf.index

# Normalizzazione testi
gdf[HAZARD_COL] = gdf[HAZARD_COL].astype(str).str.lower().str.strip()

if TRIGGERING_COL in gdf.columns:
    gdf[TRIGGERING_COL] = (
        gdf[TRIGGERING_COL]
        .astype(str)
        .str.lower()
        .str.strip()
        .replace({"eartquake": "earthquake", "none": None, "nan": None, "":
None})
    )
else:
    gdf[TRIGGERING_COL] = None

# Date
gdf["start"] = pd.to_datetime(gdf[DATE_COL], dayfirst=True, errors="coerce")
gdf["end"]   = gdf["start"]   # nel tuo file questi eventi sono puntuali
gdf = gdf.dropna(subset=["start"]).copy()

# CRS
if gdf.crs is None:
    gdf = gdf.set_crs("EPSG:4326")
gdf = gdf.to_crs("EPSG:32632")

# ------------------------------------------------------------
# 5. CARICAMENTO HEATWAVES
```

```
# ------------------------------------------------------------
print("Caricamento Heatwaves...")
df_hw = pd.read_csv(HEATWAVES_CSV_PATH)

# salvo indice originale CSV
df_hw = df_hw.reset_index(drop=False).rename(columns={"index": "orig_fid"})

df_hw["HW_start"] = pd.to_datetime(df_hw["start_date"], errors="coerce")
df_hw["duration"] = pd.to_numeric(df_hw["duration"], errors="coerce")
df_hw = df_hw.dropna(subset=["HW_start", "duration"]).copy()
df_hw["duration"] = df_hw["duration"].astype(int)
df_hw = df_hw[df_hw["duration"] >= 3].copy()

# fine reale heatwave
df_hw["HW_end"] = df_hw["HW_start"] + pd.to_timedelta(df_hw["duration"] - 1,
unit="D")

gdf_hw = gpd.GeoDataFrame(
    df_hw,
    geometry=gpd.points_from_xy(df_hw["longitude"], df_hw["latitude"]),
    crs="EPSG:4326"
).to_crs("EPSG:32632")

# ID logico
gdf_hw["id"] = gdf_hw.index

# ------------------------------------------------------------
# 6. FILTRO HAZARD
# ------------------------------------------------------------
ls = gdf[gdf[HAZARD_COL] == NAME_LANDSLIDE].copy()
fl = gdf[gdf[HAZARD_COL] == NAME_FLOOD].copy()
wf = gdf[gdf[HAZARD_COL] == NAME_WILDFIRE].copy()
hw = gdf_hw.copy()

print("Dataset completi:")
print("  Landslides:", len(ls))
print("  Floods    :", len(fl))
print("  Wildfires :", len(wf))
print("  Heatwaves :", len(hw))

# ------------------------------------------------------------
# 7. SAMPLE 2500
# ------------------------------------------------------------
N = 2500

ls = ls.sample(min(N, len(ls)), random_state=0)
fl = fl.sample(min(N, len(fl)), random_state=0)
wf = wf.sample(min(N, len(wf)), random_state=0)
hw = hw.sample(min(N, len(hw)), random_state=0)

print("\n⚠ USO SAMPLE DA 2500 per ogni hazard")
print("  Landslides (sample):", len(ls))
print("  Floods     (sample):", len(fl))
print("  Wildfires  (sample):", len(wf))
print("  Heatwaves  (sample):", len(hw))

# ------------------------------------------------------------
# 8. CLASSIFICAZIONE
# ------------------------------------------------------------
def classify_ls_fl(dist_m, dt_days):
    if dt_days is None:
        return "None"
```

```
    if dist_m <= LSFL_COMPOUND_DIST_MAX and dt_days <= LSFL_COMPOUND_TIME_MAX:
        return "Compound"

    elif LSFL_SIMUL_DIST_MIN < dist_m <= LSFL_SIMUL_DIST_MAX and dt_days <=
LSFL_SIMUL_TIME_MAX:
        return "Simultaneous"

    elif dist_m <= LSFL_CONSEC_DIST_MAX and LSFL_CONSEC_TIME_MIN < dt_days <=
LSFL_CONSEC_TIME_MAX:
        return "Consecutive"

    return "None"


def classify_hw_wf(dist_m, dt_days):
    if dt_days is None:
        return "None"

    if dist_m < HWWF_COMPOUND_DIST_MAX and dt_days < 3:
        return "Compound"

    elif HWWF_SIMUL_DIST_MIN < dist_m < HWWF_SIMUL_DIST_MAX and dt_days < 3:
        return "Simultaneous"

    elif dist_m < HWWF_CONSEC_DIST_MAX and 3 < dt_days < 7:
        return "Consecutive"

    return "None"

# ------------------------------------------------------------
# 9. FUNZIONE GENERICA CORRETTA
# ------------------------------------------------------------
def calculate_pairs(
    gdf_A, gdf_B,
    hazard_A, hazard_B,
    classifier, buffer_m,
    parent_label, interacting_label,
    start_A_col="start", end_A_col="end",
    start_B_col="start", end_B_col="end",
    idA="id", idB="id",
    orig_fid_A_col="orig_fid", orig_fid_B_col="orig_fid"
):
    if gdf_A.empty or gdf_B.empty:
        return gpd.GeoDataFrame()

    cols_A = [orig_fid_A_col, idA, start_A_col, end_A_col, "geometry"]
    cols_B = [orig_fid_B_col, idB, start_B_col, end_B_col, "geometry"]

    A = gdf_A[cols_A].rename(columns={
        orig_fid_A_col: "orig_fid_A",
        idA: "id_A",
        start_A_col: "start_A",
        end_A_col: "end_A",
        "geometry": "geom_A"
    }).copy()

    B = gdf_B[cols_B].rename(columns={
        orig_fid_B_col: "orig_fid_B",
        idB: "id_B",
        start_B_col: "start_B",
        end_B_col: "end_B",
        "geometry": "geom_B"
    }).copy()
```

```
    # GeoDataFrame di A
    A_gdf = gpd.GeoDataFrame(A, geometry="geom_A", crs=gdf_A.crs)

    # Buffer su B: qui il buffer deve essere la geometria attiva
    B_buf = gpd.GeoDataFrame(B.copy(), geometry="geom_B", crs=gdf_B.crs)
    B_buf["buffer_geom"] = B_buf["geom_B"].buffer(buffer_m)
    B_buf = B_buf.set_geometry("buffer_geom")

    # Spatial join: cerca A dentro i buffer di B
    tmp = gpd.sjoin(
        A_gdf,
        B_buf[["orig_fid_B", "id_B", "start_B", "end_B", "geom_B",
"buffer_geom"]],
        how="inner",
        predicate="intersects"
    ).drop(columns=["index_right"])

    if tmp.empty:
        return gpd.GeoDataFrame()

    # Distanza reale A-B
    tmp["dist_m"] = tmp["geom_A"].distance(tmp["geom_B"])

    # Time lag reale
    tmp["time_lag_days"] = tmp.apply(
        lambda r: compute_time_lag_days(r["start_A"], r["end_A"], r["start_B"],
r["end_B"]),
        axis=1
    )

    # Classificazione
    tmp["Interaction"] = tmp.apply(
        lambda r: classifier(r["dist_m"], r["time_lag_days"]),
        axis=1
    )

    tmp = tmp[tmp["Interaction"] != "None"].copy()
    if tmp.empty:
        return gpd.GeoDataFrame()

    # Filter out invalid or empty geometries before creating LineString
    tmp = tmp[tmp['geom_A'].is_valid & ~tmp['geom_A'].is_empty &
              tmp['geom_B'].is_valid & ~tmp['geom_B'].is_empty].copy()
    if tmp.empty:
        return gpd.GeoDataFrame()

    # Coordinate A e B
    tmp["x_A"] = tmp["geom_A"].x
    tmp["y_A"] = tmp["geom_A"].y
    tmp["x_B"] = tmp["geom_B"].x
    tmp["y_B"] = tmp["geom_B"].y

    # Geometria finale = linea tra A e B
    # Explicitly extract x and y coordinates for robustness
    tmp["geometry"] = tmp.apply(
        lambda r: LineString([
            (r["geom_A"].x, r["geom_A"].y),
            (r["geom_B"].x, r["geom_B"].y)
        ]),
        axis=1
    )
```

```
    # Metadati
    tmp["hazard_A"] = hazard_A
    tmp["hazard_B"] = hazard_B
    tmp["parent_hazard"] = parent_label
    tmp["interacting_hazard"] = interacting_label

    tmp["pair_id"] = (
        tmp["id_A"].fillna(-1).astype(int).astype(str)
        + "__" +
        tmp["id_B"].fillna(-1).astype(int).astype(str)
    )

    # Rinominazione date
    tmp = tmp.rename(columns={
        "start_A": "DATE_A",
        "end_A":   "END_A",
        "start_B": "DATE_B",
        "end_B":   "END_B"
    })

    cols_out = [
        "orig_fid_A", "orig_fid_B",
        "id_A", "id_B",
        "hazard_A", "hazard_B",
        "parent_hazard", "interacting_hazard",
        "Interaction", "pair_id",
        "DATE_A", "END_A", "DATE_B", "END_B",
        "x_A", "y_A", "x_B", "y_B",
        "dist_m", "time_lag_days",
        "geometry"
    ]

    tmp = tmp[cols_out].copy()
    tmp = gpd.GeoDataFrame(tmp, geometry="geometry", crs=gdf_A.crs)

    # Elimina solo veri duplicati
    tmp = tmp.drop_duplicates(subset=["pair_id"])

    return tmp

# ------------------------------------------------------------
# 10. CALCOLO INTERAZIONI
# ------------------------------------------------------------
all_pairs = []

# Landslide - Flood
pairs_ls_fl = calculate_pairs(
    ls, fl,
    hazard_A="landslide",
    hazard_B="flood",
    classifier=classify_ls_fl,
    buffer_m=BUFFER_LSFL,
    parent_label="rainfall",
    interacting_label="flood",
    start_A_col="start",
    end_A_col="end",
    start_B_col="start",
    end_B_col="end",
    idA="id",
    idB="id",
    orig_fid_A_col="orig_fid",
    orig_fid_B_col="orig_fid"
)
```

```
print("\nCoppie Landslide-Flood:", len(pairs_ls_fl))

if not pairs_ls_fl.empty:
    ls_trig = ls[[ID_COL, TRIGGERING_COL]].rename(
        columns={ID_COL: "id_A", TRIGGERING_COL: "Trigger_A"}
    )
    pairs_ls_fl = pairs_ls_fl.merge(ls_trig, on="id_A", how="left")

    def choose_parent(trig):
        if trig == "earthquake":
            return "earthquake"
        elif trig == "rainfall":
            return "rainfall"
        else:
            return "rainfall"

    pairs_ls_fl["parent_hazard"] = pairs_ls_fl["Trigger_A"].apply(choose_parent)
    pairs_ls_fl = pairs_ls_fl.drop(columns=["Trigger_A"])
    all_pairs.append(pairs_ls_fl)

# Heatwave – Wildfire
hw2 = hw.copy()
hw2["start"] = hw2["HW_start"]
hw2["end"]   = hw2["HW_end"]

pairs_hw_wf = calculate_pairs(
    hw2, wf,
    hazard_A="heatwave",
    hazard_B="wildfire",
    classifier=classify_hw_wf,
    buffer_m=BUFFER_HWWF,
    parent_label="temperature",
    interacting_label="wildfire",
    start_A_col="start",
    end_A_col="end",
    start_B_col="start",
    end_B_col="end",
    idA="id",
    idB="id",
    orig_fid_A_col="orig_fid",
    orig_fid_B_col="orig_fid"
)
print("Coppie Heatwave-Wildfire:", len(pairs_hw_wf))

if not pairs_hw_wf.empty:
    all_pairs.append(pairs_hw_wf)

# ------------------------------------------------------------
# 11. MERGE & EXPORT
# ------------------------------------------------------------
if all_pairs:
    pairs_all = gpd.GeoDataFrame(
        pd.concat(all_pairs, ignore_index=True),
        geometry="geometry",
        crs=gdf.crs
    )

    pairs_all.to_file(OUTPUT_PAIRS_PATH, driver="GPKG")

    print("\n🎉 File finale salvato in:")
    print(OUTPUT_PAIRS_PATH)
    print("\nTotale interazioni trovate:", len(pairs_all))
    print("\nDistribuzione Interaction:")
```

```
    print(pairs_all["Interaction"].value_counts())

    zip_path = str(OUTPUT_PAIRS_PATH).replace(".gpkg", "_updated.zip")
    shutil.make_archive(str(Path(zip_path).stem), "zip",
OUTPUT_PAIRS_PATH.parent, OUTPUT_PAIRS_PATH.name)
    files.download(zip_path)

else:
    print("\n⚠ Nessuna interazione trovata con il sample 2500.")
```

10. BIBLIOGRAFIA E SITOGRAFIA

UNDRR (2017). *Terminology on Disaster Risk Reduction*. Geneva: United Nations Office for Disaster Risk Reduction.

IPCC (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press.

Kappes, M. S., Keiler, M., von Elverfeldt, K., & Glade, T. (2012). Challenges of analyzing multi-hazard risk: A review. *Natural Hazards*, 64(2), 1925–1958. https://doi.org/10.1007/s11069-012-0294-2

Claassen, J. N., Ward, P. J., Daniell, J., Koks, E. E., Tiggeloven, T., & de Ruiter, M. C. (2023). A new method to compile global multi-hazard event sets. *Scientific Reports*, 13, 13808.

Ferrario, D. M., Tiggeloven, T., Maraschini, M., Sanò, M., Claassen, J., de Ruiter, M. C., Torresan, S., & Critto, A. (2025). A machine learning methodology to identify current and future multi-hazard extreme events.

Hewitt, K., & Burton, I. (1971). *The hazardousness of a place: A regional ecology of damaging events*. Toronto: University of Toronto Press.

United Nations (1993). *Report of the United Nations Conference on Environment and Development, Rio de Janeiro, 3–14 June 1992. Volume I: Resolutions adopted by the Conference*. New York: United Nations.

United Nations (2002). *Plan of Implementation of the World Summit on Sustainable Development*. Johannesburg: United Nations.

UN/ISDR (2005). *Hyogo Framework for Action 2005–2015: Building the Resilience of Nations and Communities to Disasters*. Geneva: United Nations.

Delmonaco, G., Margottini, C., & Spizzichino, D. (2006). *Report on new methodology for multi-risk assessment and the harmonisation of different natural risk maps*. ARMONIA Project, Deliverable 3.1.

Gill, J. C., & Malamud, B. D. (2014). Reviewing and visualizing the interactions of natural hazards. *Reviews of Geophysics*, 52(4), 680–722.

Tilloy, A., Malamud, B. D., Winter, H., & Joly-Laugel, A. (2019). A review of quantification methodologies for multi-hazard interrelationships. *Earth-Science Reviews*, 196, 102881.

Ferrario, D. M., Tiggeloven, T., Maraschini, M., Sano, M., Claassen, J., Ward, P. J., de Ruiter, M. C., & Couasnon, A. (2025). A machine learning methodology to identify current and future multi-hazard extreme events. *ESS Open Archive*. https://doi.org/10.22541/essoar.175396232.24844978/v1

Copernicus Climate Data Store (CDS). *ERA5 hourly data on single levels from 1940 to present*. Disponibile su: https://cds.climate.copernicus.eu/

Copernicus Emergency Management Service (EMS). *Mapping and Damage Assessment for Emergency Response*. Disponibile su: https://emergency.copernicus.eu/

Copernicus European Forest Fire Information System (EFFIS). *EFFIS – Forest Fire Information System*. Disponibile su: https://effis.jrc.ec.europa.eu/

ISPRA (Istituto Superiore per la Protezione e la Ricerca Ambientale). *Inventario dei Fenomeni Franosi in Italia (IFFI)*. Disponibile su: https://www.isprambiente.gov.it/

CNR-IRPI. *ITALICA – Italian rainfall-induced Landslides Catalogue*. Disponibile su: https://italica.irpi.cnr.it/

INGV (2026). *CFTI – Catalogo dei Forti Terremoti in Italia*. Disponibile su: https://emidius.mi.ingv.it/CFTI5/ (consultato il: 16 aprile 2026).

Università di Firenze / CEDIT. *Catalogo degli effetti cosismici in Italia*. Disponibile su: https://www.cedit.it/