

44th GNGTS National Conference
Udine, 10-13 February 2026



Workshop di Microzonazione Sismica: dalla ricerca scientifica a nuovi standard, pratiche e linee guida

Metodologie Machine Learning a supporto della MS

Gabriele Fiorentino, Federico Mori, Giuseppe Naso

Ricercatore, CNR-IGAG

Consiglio Nazionale delle Ricerche

Istituto di Geologia Ambientale e Geoingegneria

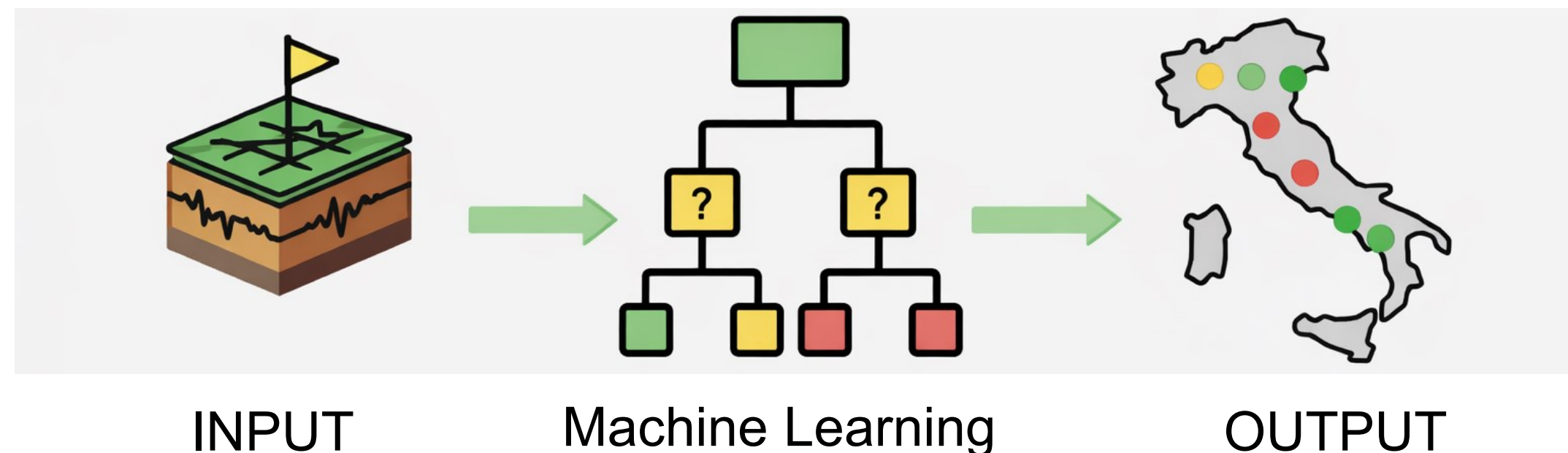
Gabriele.fiorentino@cnr.it

1

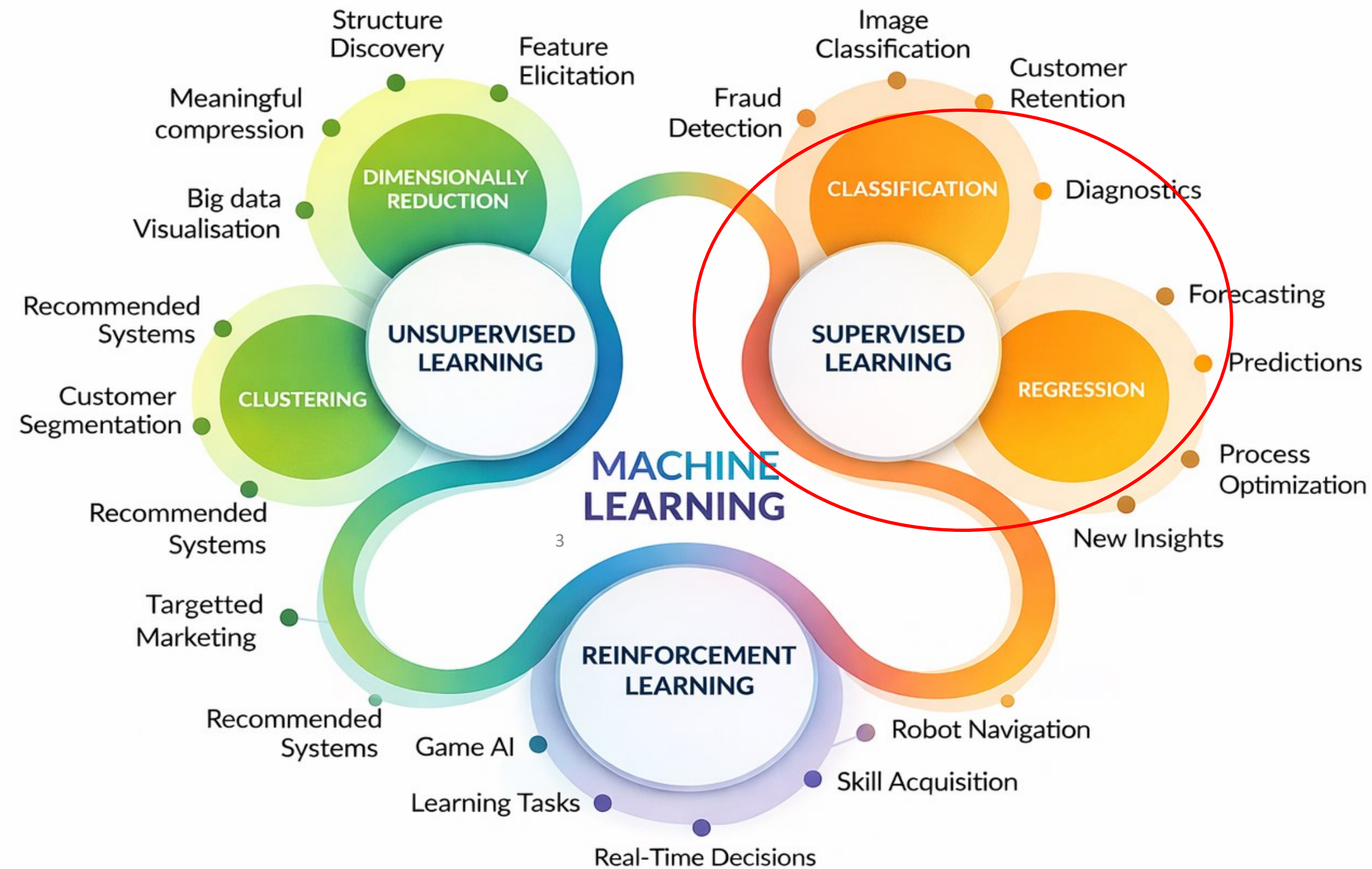


Machine Learning

- Il Machine Learning (ML) è una metodologia di IA che :
 - impara automaticamente dai dati (**data-driven**)
 - Fornisce predizioni senza essere esplicitamente programmato per farlo
- **Tecniche statistiche e probabilistiche** per:
 - **Classificare** (output categoriale, es. Suolo B),
 - **Regressione** (output numerico, es. Vs30=650 m/s)
- E' una metodologia **non parametrica**
- **È possibile spiegare l'importanza** dei predittori



Machine Learning



Come funziona (cenni)

Training

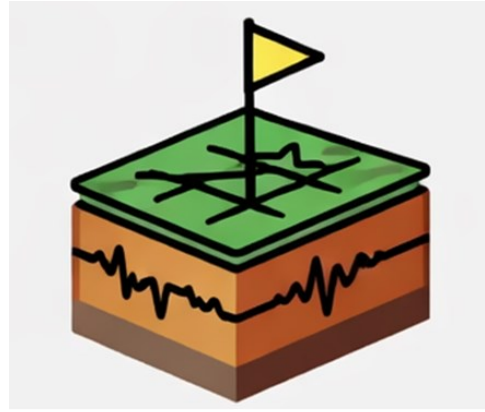
- Il modello ML **impara dai dati**
- Sottoinsieme dei dati → **training set**
- Al modello vengono forniti:
 - **input** (es. flatfile, parametri, dati grezzi da prove)
 - **output corretti** (nel caso supervised),
- L'algoritmo cerca di:
 - individuare **relazioni, pattern**
 - ottimizzare i parametri per ridurre l'errore
- **Rischio di overfitting:**
 - Il modello impara troppo bene dai dati di training, per cui si adatta bene a questi dati
 - Fallisce nel dare buone predizioni

4

Testing

- Si valuta quanto il modello funziona davvero: “// *modello funziona anche su dati nuovi?*”
- Si usa un insieme di dati mai visti prima dal modello, detto **test set**
- Qui non si impara nulla di nuovo:
 - i parametri del modello sono fissi
 - si misura solo la **performance**
- Le metriche di test (errore, accuratezza, ecc.) indicano:
 - Affidabilità → il risultato è coerente?
 - Robustezza → Stabile?
 - capacità di generalizzazione (vs Overfitting)

Perché utilizzare il Machine Learning con dati della Microzonazione Sismica?



INPUT



OUTPUT

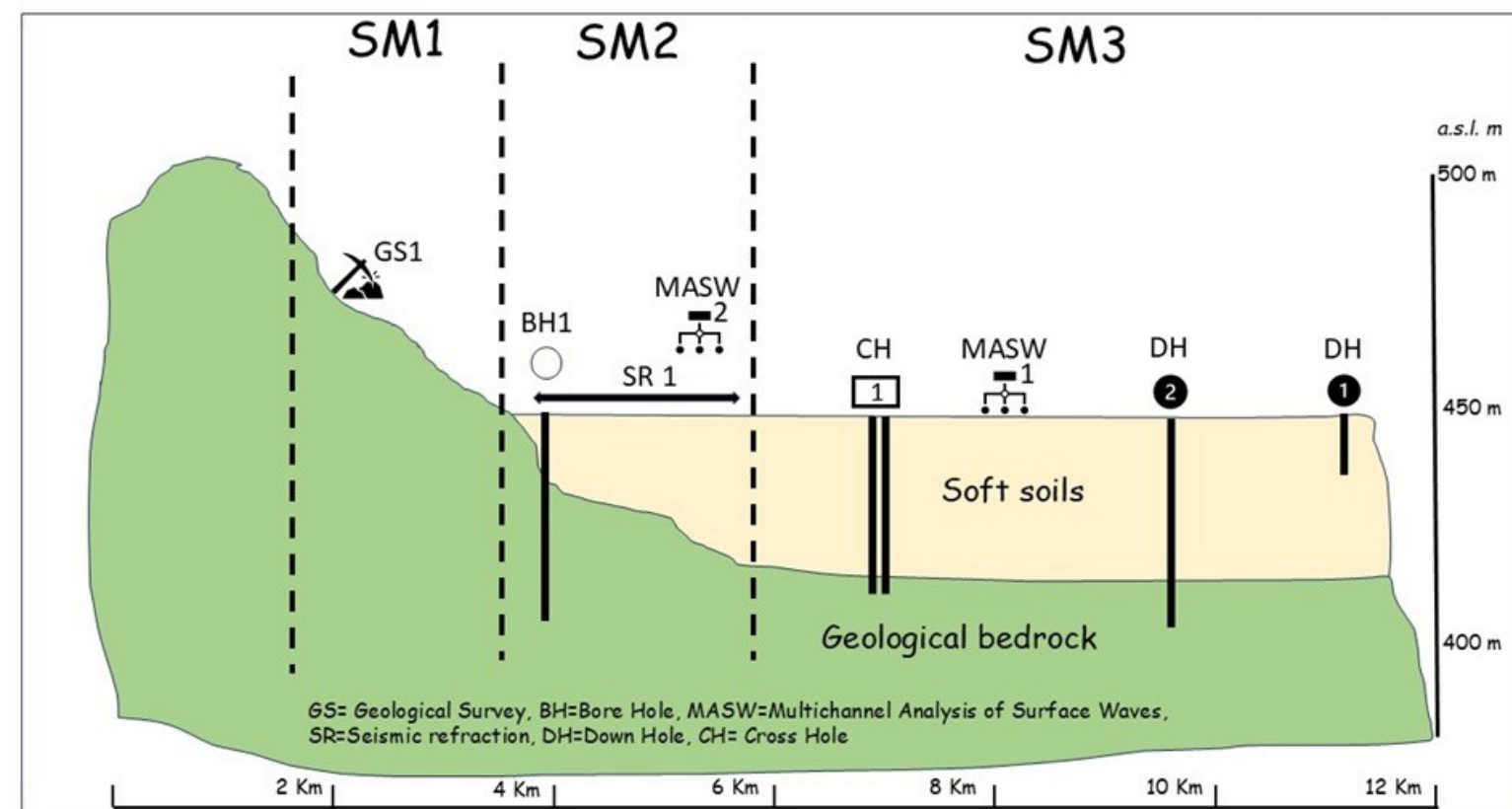
- Microzonazione Sismica (MS) → grandi quantità di dati da diverse fonti (es. HVSR, DH, CH, MASW, ERT etc...)
- Dati MS possono essere integrati con altri (geologici, geotecnici, geomorfologici, DEM)
- Spesso eterogenei per scala spaziale, risoluzione e livello di affidabilità (Comuni diversi, elaborazioni fatte da professionisti diversi)
- Il ML può fornire predizioni o interpretazioni dei dati non visibili con tecniche tradizionali
(es. regressioni parametriche lineari o non lineari)
- Obiettivo non è solo stimare un parametro in modo deterministico, ma descrivere l'accuratezza della predizione attraverso **variabilità e incertezza**

Requisiti minimi dei dati per un buon modello ML

- 1. Quantità di dati:** Il numero minimo da rappresentare le variabilità
- 2. Qualità dei dati:** Dati ben organizzati con etichette corrette → Assenza outlier estremi o valori mancanti
- 3. Variabilità informativa:** I dati devono contenere abbastanza variabilità per consentire al modello di apprendere relazioni utili (no dati duplicati o troppo simili).
- 4. Predittori rilevanti:** Le variabili devono avere una buona relazione col target da prevedere (altrimenti il modello sarà poco predittivo) → sorta di «trial and error»
- 5. Assenza di forte collinearità:** Troppe variabili correlate tra loro rendono instabili i coefficienti → questo è un problema del modello e di chi sceglie le variabili
- 6. Distribuzione coerente tra training e test:** La distribuzione dei dati nei set di training e test deve essere simile (es. alleno su pianure alluvionali e testo su contesti montani)
- 7. Label bilanciate** (per classificazione): L'accuratezza può essere ingannevole. È importante sapere cosa il modello sbaglia oltre al *quanto*

ML come strumento operativo per la MS

- Uno degli scopi della MS è valutare la pericolosità sismica locale tramite l'individuazione di zone a comportamento sismico omogeneo.
- 3 tipi di zone: stabili, stabili con amplificazione, instabili
- Nella pratica, per le prime due zone le analisi valutano i fattori di amplificazione, gli spettri di risposta utili alla progettazione o verifica di strutture o infrastrutture
- Per le zone instabili, oltre allo scuotimento si valuta la probabilità di accadimento di frane e liquefazioni
- In questo contesto, i parametri per noi più interessanti sono parametri quali: V_s , f_0 , h_{800} , parametri geotecnici

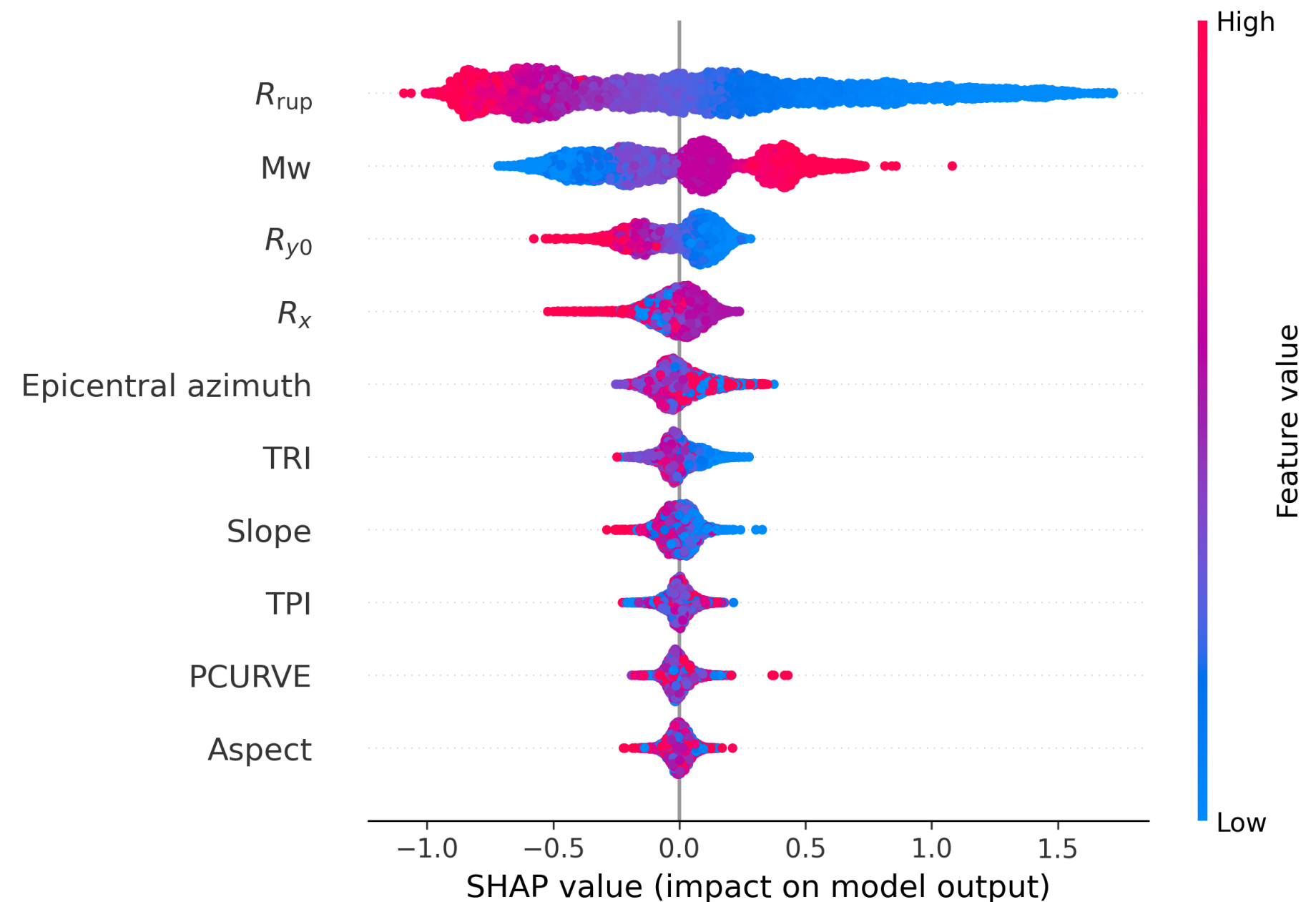
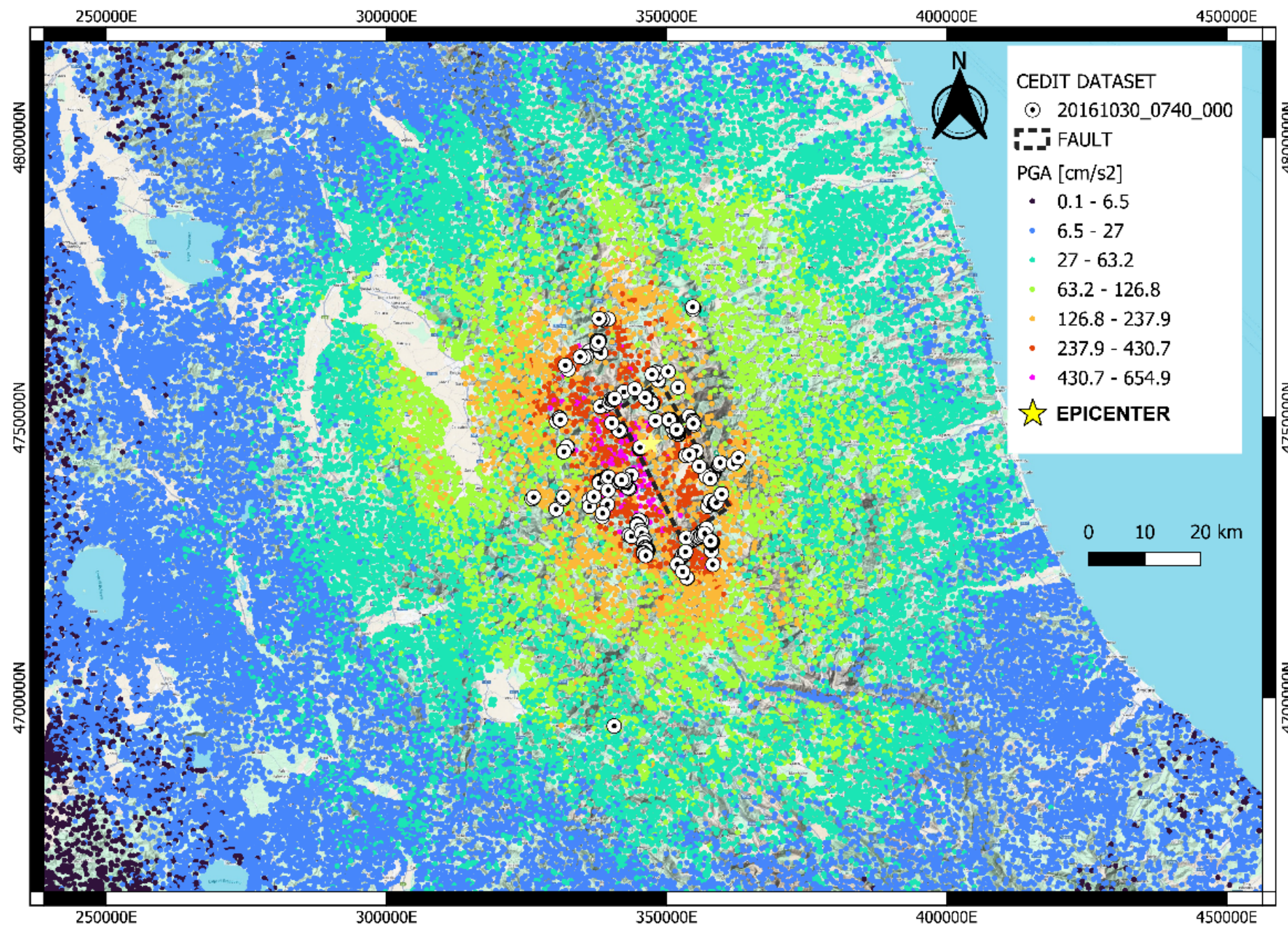


Parametri geotecnici

Parametro geotecnico	Frane sismoindotte	Liquefazione
c' – coesione efficace	✓ Fondamentale (terreni coerenti)	✗ Non rilevante
φ' – angolo di attrito efficace	✓ Fondamentale	✓ Importante
γ / γ_{sat} – peso di volume	✓ Importante	✓ Importante
Spessore copertura instabile	✓ Fondamentale	✗ Non rilevante
Profondità superficie di scorrimento	✓ Fondamentale	✗ Non rilevante
Pendenza del versante	✓ Fondamentale	✗ Non rilevante
Livello di falda / saturazione	✓ Importante	✓ Fondamentale
Permeabilità (k)	✓ Importante	✓ Importante
Densità relativa (D_r)	✗ Non rilevante	✓ Fondamentale
Indice dei vuoti (e)	✗ Non ⁸ rilevante	✓ Importante
Granulometria (D50, % fini)	✗ Secondaria	✓ Fondamentale
Plasticità (LL, PI)	✗ Secondaria	✓ Fondamentale
SPT (N_{60}) / CPT (qc, fs)	✗ Proxy	✓ Proxy chiave
V_s / V_{s30}	✓ Proxy	✓ Proxy
Gmax, curve G/Gmax, D(γ)	✗ Secondaria	✓ Importante

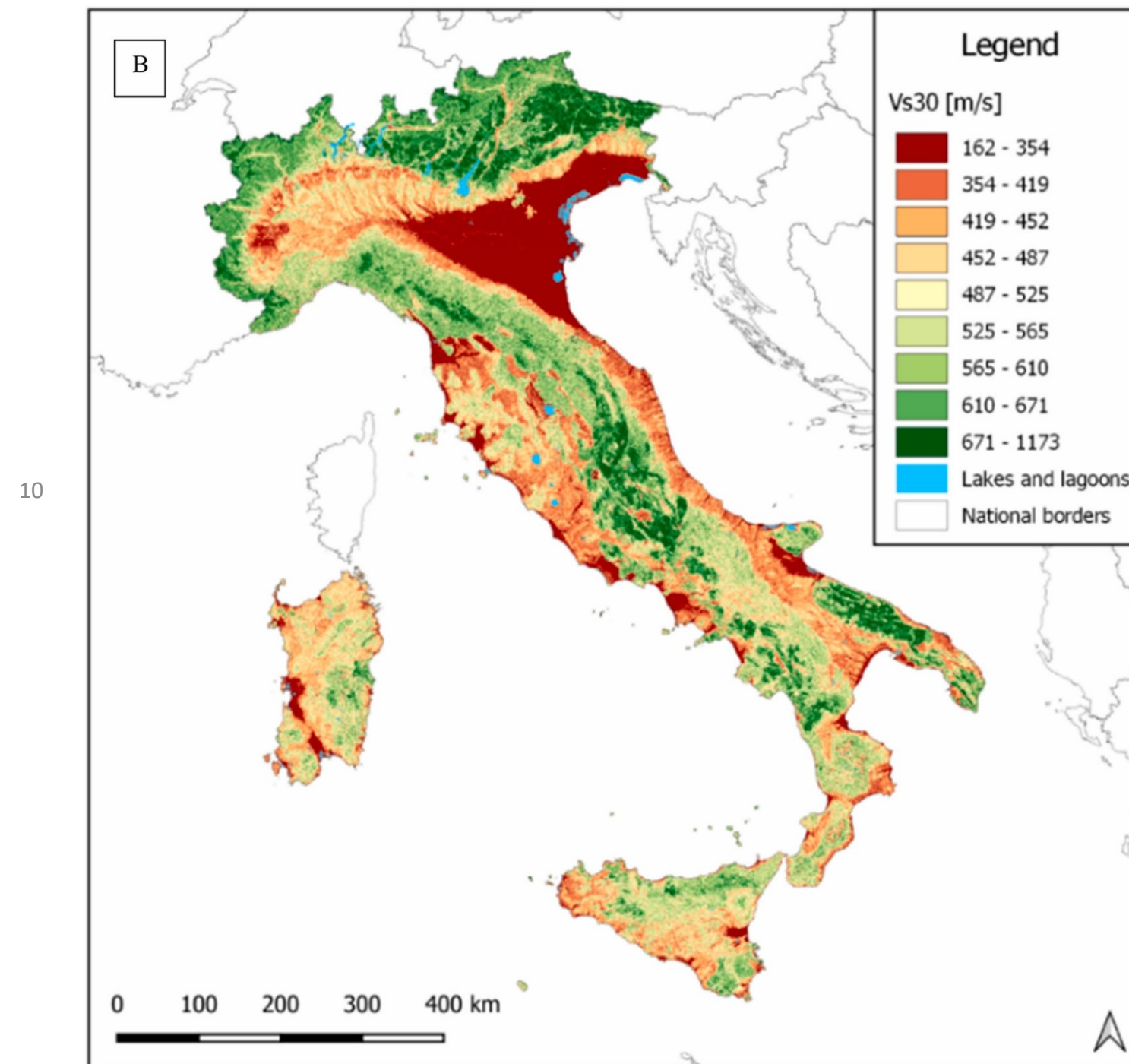
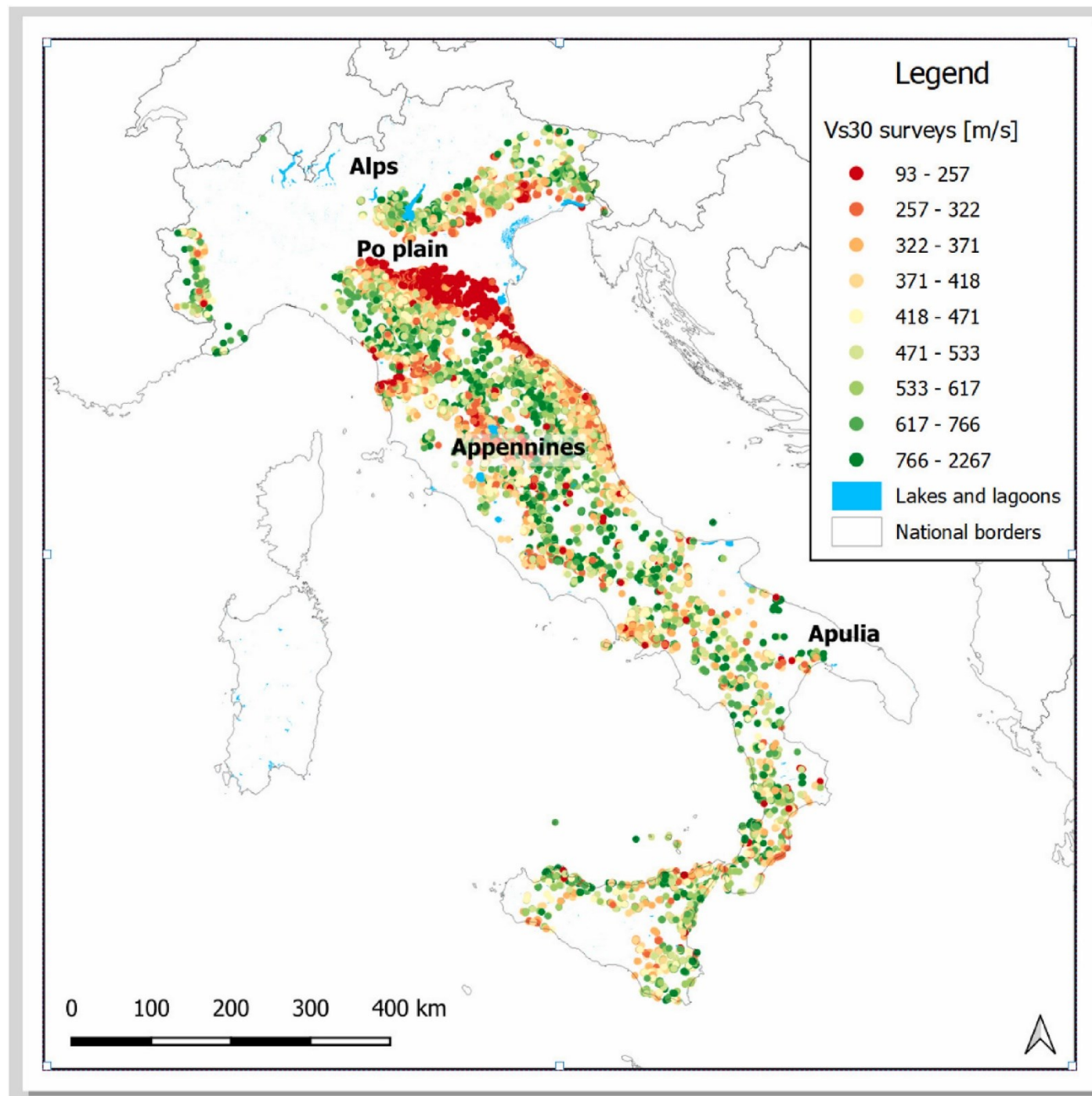
Esempio ML: scenari di scuotimento near-real time per triggering di frane

- INPUT → ci sono parametri (predittori) della sorgente e path (M_w , R_{rup} ...), geomorfologici (TRI, Slope, TPI, PCURVE, Aspect)
- Sovrapposizione con il database di frane cosismiche del CEDIT
- Importanza dei parametri geomorfologici dopo quelli di sorgente e path



Esempio di elaborazione dati con ML: Mappa Vs30 con ML (2026)

- Vs30 ha molti limiti, ma 1) è codificata nelle NTC 2) è un dato numeroso
- INPUT→ ci sono parametri (predittori) **numerici** (es. elevation e slope) e **categorici** (es. litologia)
- Il training è stato fatto utilizzando i profili di Vs dal database MS (e da questi il calcolo della Vs30)
- Risultati→ 1) mappa Vs30, 2) mappa dei residui, 3) interpretazione dei risultati (classifica dei predittori)

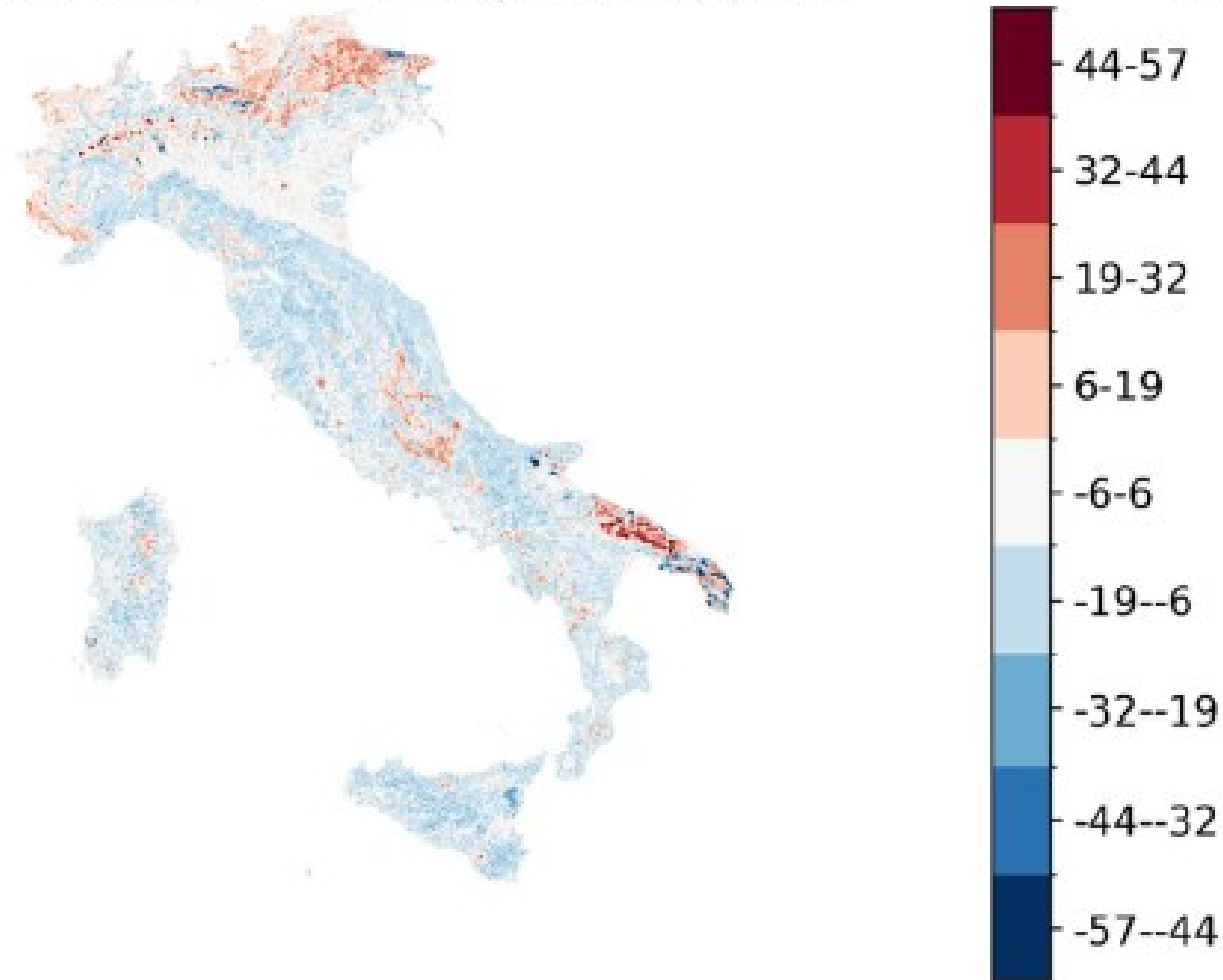


Mori et al. 2026
*Machine learning-
based national Vs30
models and maps for
Italy*

Esempio di elaborazione dati con ML: Mappa Vs30 con ML (2026)

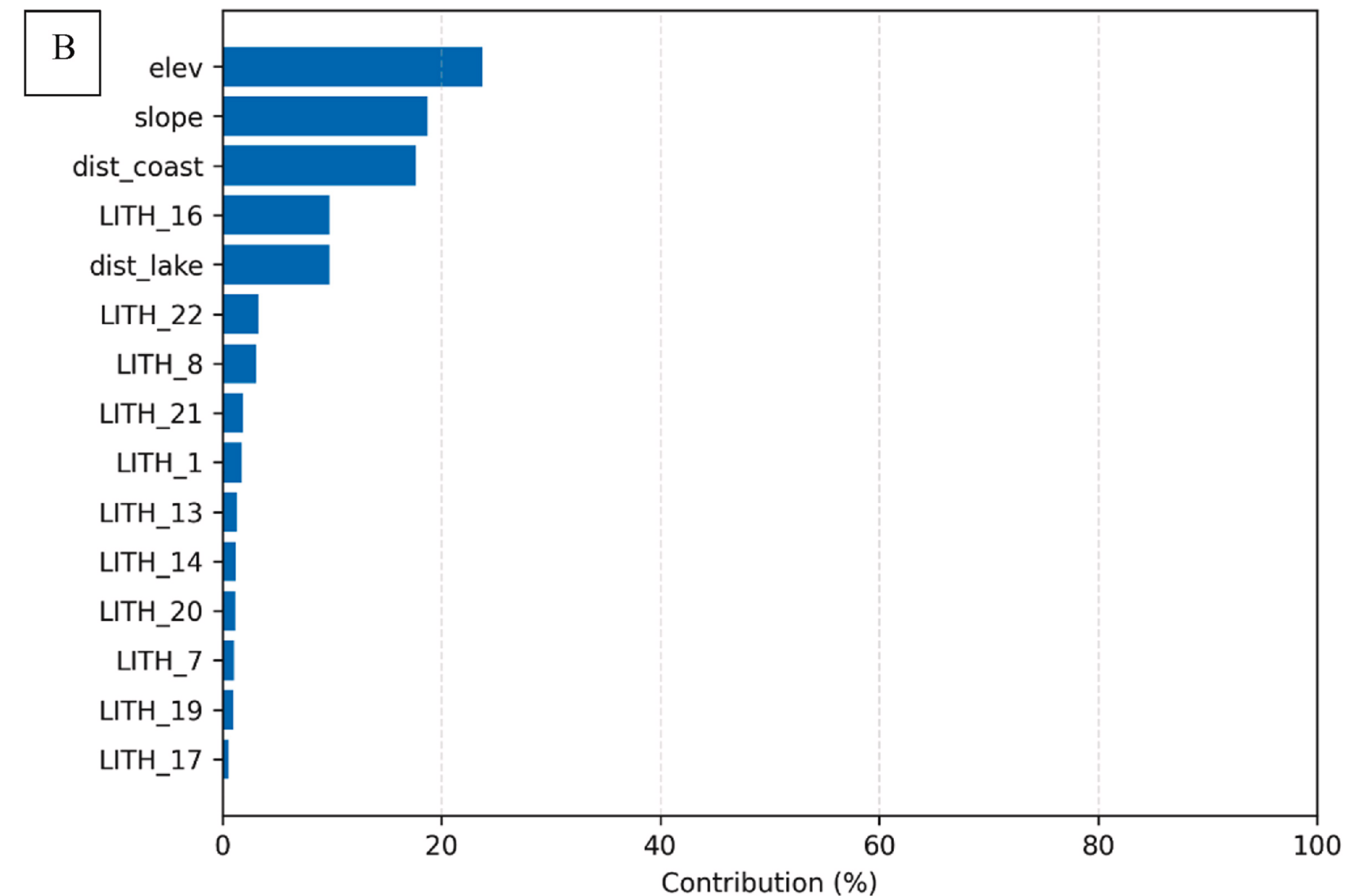
- La mappa dei residui mostra errori massimi di circa 50 m/s
- La maggior parte dei residui è tra ± 20 m/s

C) Residuals – Lithological support



11

- SHAP valuta il contributo al risultato di ciascun parametro (o predittore)
- Per mappa Vs30 Elevation, Slope e distanza dalla costa sono i parametri che più influenzano Vs30



Vs30: limiti concettuali e valore operativo

Limiti

- Parametro **fortemente sintetico**, non rappresenta la complessità stratigrafica né la profondità del bedrock.
- Siti con **uguale Vs30** possono avere risposte sismiche molto diverse, soprattutto alle basse frequenze.
- **Stima spesso incerta**, basata su misure indirette o proxy empirici.
- **Classi normative discrete (A–E)** introducono discontinuità artificiali.

Vantaggi

- È l'**unico parametro realmente numeroso** nei dati di Microzonazione Sismica, quindi trattabile in modo robusto con approcci data-driven.
- Permette una **mappa nazionale continua, replicabile e aggiornabile**.
- Fornisce uno **strumento di riferimento e screening**, non sostitutivo delle analisi di MS locale.

12

Alcuni prodotti possibili / in corso di sviluppo

- Scenari di scuotimento in pianificazione e near-real time
- Applicazione a predizioni probabilistiche di frane e liquefazione
- Indice di qualità delle indagini rivisitato in chiave data-driven
- Mappa nazionale Vs30



Grazie per l'attenzione

Gabriele Fiorentino, Federico Mori, Giuseppe Naso

CNR-IGAG

Consiglio Nazionale delle Ricerche

Istituto di Geologia Ambientale e Geoingegneria

Gabriele.fiorentino@cnr.it



th
44th GNGTS National Conference
Udine, 10-13 February 2026

14

CENTROMS

CENTRO PER LA
MICROZONAZIONE SISMICA
E LE SUE APPLICAZIONI

