



UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI GEOSCIENZE

Direttore Prof. Nicola Surian

TESI DI LAUREA MAGISTRALE IN

GEOLOGIA AMBIENTALE E DINAMICA DELLA TERRA

**“VALUTAZIONE DELLA SUSCETTIBILITÀ DA FRANA DEI
MONTI LESSINI TRAMITE TECNICHE DI XAI
(EXPLENABLE AI)”**

**“LANDSLIDE SUSCEPTIBILITY ASSESSMENT OF THE LESSINI
MOUNTAINS USING XAI (EXPLANABLE AI) TECHNIQUES”**

Relatore: Prof. Mario Floris

Correlatore: Prof. Ascanio Rosi

Laureanda: Simona Salvetti

Matricola n. 2097827

ANNO ACCADEMICO 2023/2024

Abstract

Questo lavoro di tesi si concentra sulla mappatura della suscettibilità da frana dei Monti Lessini. L'obiettivo principale è quello di identificare e analizzare i principali fattori geologici e geomorfologici che contribuiscono all'instabilità dei versanti, in modo da fornire una base utile per la gestione del territorio e la previsione dei rischi geologici. Per raggiungere tale obiettivo verrà impiegato il modello di machine learning Random Forest, noto per la sua capacità di analizzare grandi dataset non lineari, e le tecniche del Partial Dependence Plot. Questi strumenti consentiranno di valutare nel dettaglio il contributo di ogni singolo fattore predisponente sulla stabilità dei pendii. Il risultato finale di questo lavoro è una mappa di suscettibilità molto accurata, che rappresenta uno strumento importante per comprendere le dinamiche franose dell'area studio e per supportare le decisioni relative alla pianificazione territoriale e agli interventi di mitigazione e prevenzione dei rischi geologici.

INDICE

1. INTRODUZIONE	1
2. INQUADRAMENTO DELL'AREA IN STUDIO	3
2.1. INQUADRAMENTO GEOGRAFICO	3
2.2. INQUADRAMENTO GEOLOGICO	6
2.3. SUCCESSIONE LITOSTRATIGRAFICA	11
3. MATERIALI E METODI	17
3.1. MATERIALI	17
3.1.1. INVENTARIO DEI FENOMENI FRANOSI IN ITALIA (IFFI)	17
3.1.2. SLOPE UNITS (SU)	21
3.1.3. VARIABILI PREDISPONENTI	26
3.2. METODI	38
3.2.1. MODELLI PER LA VALUTAZIONE DELLA SUSCETTIBILITÀ DA FRANA	38
3.2.2. WORKFLOW	43
3.2.2.1. PREPARAZIONE DEL DATASET	43
3.2.2.2. ANALISI PRELIMINARE	44
3.2.2.3. ANALISI FINALE E PARTIAL DIPENDENCE PLOTS	46
3.2.2.4. MAPPA DI SUSCETTIBILITÀ	47
4. RISULTATI	48
4.1. ANALISI DEL DATASET	49
4.2. RISULTATI ANALISI PRELIMINARE	54
4.3. RISULTATI ANALISI FINALE	66
4.3.1. INTERPRETAZIONE DEI RISULTATI	86
4.4. MAPPE DI SUSCETTIBILITÀ DA FRANA	90

5. DISCUSSIONI	96
6. CONCLUSIONE	101
7. BIBLIOGRAFIA	104
8. ALLEGATI	113

1 Introduzione

Con il termine frana si fa riferimento ad un movimento gravitativo di una massa di roccia, terra o detrito lungo un versante (Cruden, 1991). È un fenomeno naturale che ha un impatto significativo sulla popolazione, sia in termini di perdite economiche, quindi danni alle infrastrutture, alle attività economiche e proprietà private, ma anche in termini di vite umane. Per esempio Ubydul et al. (2016) dichiara che nel corso di 19 anni (1995-2014) in Europa si sono verificati 1370 morti e 784 feriti causati da 476 frane; secondo il lavoro Pertly (2012) dal 2004 al 2010, nel mondo sono state registrate 2620 fenomeni franosi letali, che hanno causato un totale di 32.322 vittime. Inoltre, i dati riportati dell'IPCC (2023) sottolineano come il cambiamento climatico abbia incentivato il manifestarsi di eventi meteorologici estremi che influenzano significativamente la stabilità dei versanti.

Questo contesto enfatizza la necessità di effettuare l'analisi del rischio dei fenomeni franosi, per prevederne l'accadimento, l'evoluzione e le possibili conseguenze, in modo da agire sul territorio per mitigare e ridurre l'impatto dell'evento, se questo si dovesse verificare. I tre parametri su cui si basa l'analisi del rischio, introdotti da (Varnes, 1984), sono: la pericolosità, la vulnerabilità e il valore degli elementi a rischio. La pericolosità di frana è la probabilità di occorrenza di un fenomeno di una certa intensità (I), in una determinata area, in un periodo di tempo stabilito. In base a questa definizione è necessario determinare la probabilità che un fenomeno si verifichi sia in termini temporali che spaziali. L'obiettivo di questa tesi di laurea è l'elaborazione di una mappa di suscettibilità di frana (LSM) dell'area dei Monti Lessini, nelle Prealpi Venete. Quindi è stata fatta l'analisi spaziale della pericolosità, ovvero l'identificazione delle zone in cui è presente una maggiore o una minore probabilità di accadimento di un fenomeno franoso.

Per determinare statisticamente la suscettibilità da frana bisogna suddividere lo spazio di indagine in delle unità di mappatura base; Reichenbach et al. (2018) hanno raccolto tutte le unità presenti in letteratura e le hanno riassunte nei seguenti sette gruppi: pixel o celle di griglia, unità di terreno, unità di condizione unica, slope units (SU), unità geoidrologiche, unità topografiche e unità politiche o amministrative. In questa lavoro di tesi si è scelto di lavorare con le SU ovvero le unità di terreno

delimitate da linee di drenaggio e spartiacque, utilizzando esclusivamente le informazioni ottenute da un modello digitale di elevazione (DEM). Hanno una forte relazione con la topografia e sono interessanti per un'ampia gamma di applicazioni. Facilitano la gestione di dati eterogenei, risolvono problemi legati all'accuratezza della posizione, aggregando dati distribuiti in modo significativo (Alvioli et al. 2016).

L'estrema varietà dei fenomeni franosi, delle loro cause e dei loro possibili effetti rende complessa l'elaborazione di una metodologia universale per la valutazione della suscettibilità e del relativo rischio. Esistono molti metodi basati su diversi criteri di approccio che permettono di ottenere questo risultato. Negli ultimi anni è stata introdotta l'intelligenza artificiale (AI) per la valutazione della suscettibilità, aumentando l'accuratezza dei risultati e la capacità previsionale dei modelli.

Per la determinazione della suscettibilità dell'area studio, è stato allenato un modello di machine learning denominato Random Forest. L'analisi si sviluppa partendo dall'inventario dei fenomeni franosi dell'IFFI (Inventario dei Fenomeni Franosi in Italia), da cui si ricavano i punti in frana all'interno dell'area studio. Per la valutazione della probabilità spaziale di accadimento vanno considerati i fattori predisponenti per l'innescio del fenomeno. In questo caso sono state scelte 16 variabili basate sulle condizioni ambientali della zona. Questi fattori sono stati combinati in diversi modi, al fine di ottenere la configurazione migliore in termini di efficienza del modello.

2 Inquadramento dell'area in studio

2.1 Inquadramento geografico

L'area studio (Figura 1) si sviluppa su una superficie di 476 km^2 tra le province di Verona e Vicenza. Comprende la parte orientale dei Monti Lessini e quattro importanti vallate che si estendono ramificandosi verso sud, fino al limite della Pianura Veneta, all'altezza dell'abitato di San Bonifacio. Il confine SE si colloca nelle vicinanze dei Colli Berici, mentre quello orientale esclude di poco il territorio comunale di Montecchio Maggiore e la Val d'Agno. Il bordo settentrionale, invece, è posizionato in prossimità del confine regionale con il Trentino-Alto Adige.

I territori coinvolti in questo studio, pertanto, sono i Monti Lessini e, procedendo da Ovest verso Est, le seguenti piane vallive: la Valle d'Illasi, la Val Tramigna, la Val d'Alpone e la Valle del Chiampo.

I Monti Lessini vengono descritti dal Prof. Ugo Sauro, grande studioso e conoscitore di questo territorio, come un altopiano modellato in una digitazione di dorsali divergenti da nord, ramificatisi ed attenuatisi in un contatto dolce confinante con l'alta Pianura Veneta (Sauro, 1973) collocati all'apice dell'ampio cuneo descritto dal versante meridionale delle Alpi centro-orientali verso Sud (Sauro, 1973).

A livello paesaggistico la regione dei Lessini risulta variegata e mutevole, questo permette di suddividerla in due settori distinti: i Lessini centro-occidentali (Lessini Veronesi), compresi tra la Valle dell'Adige e la Valle d'Illasi, ed i Lessini orientali (Lessini Vicentini) che comprendono Val Tramigna, la Val d'Alpone, la Valle del Chiampo e la Val d'Agno, la quale non è oggetto di questa tesi.

Le differenze morfologiche dei due settori trovano riscontro nelle diverse litologie che le caratterizzano e nelle diverse evoluzioni tettoniche che hanno subito, come vedremo in seguito.

La Valle d'Illasi, la più occidentale tra quelle oggetto di studio si genera nei pressi del Monte Carega (2259 m) e si estende, da Nord verso Sud, per 35 km, lungo il corso del torrente Progno d'Illasi, che è un affluente di sinistra del fiume Adige. Il flusso di questo corso d'acqua è sporadico e molto limitato nel tempo. La vallata si trova quasi interamente all'interno della provincia di Verona, escluso il Gruppo del Carega, un massiccio delle Piccole Dolomiti, compreso all'interno dei confini

regionali del Trentino Alto-Adige. La Valle d'Illasi, come detto in precedenza, funge da spartiacque tra i due settori centro-occidentale e orientale dei Monti Lessini (Failla et al., 1998)

La Val Tramigna, sempre collocata in provincia di Verona, si estende per circa 20 km lungo il percorso del torrente Tramigna, la cui genesi è associata ad una risorgiva nel Comune di Cazzano Tramigna, denominata come Lago della Mora. Il corso d'acqua è un affluente del fiume Alpone che a sua volta confluisce nell'Adige. La piccola piana valliva prosegue fino al raggiungimento del Comune di Soave dove il territorio si apre verso la Pianura Veronese.

La Val d'Alpone è descritta da Salgato (2019) e dal P.A.T.I. d'Alpone (2011). Costituisce la parte più orientale della provincia di Verona e si sviluppa per 25 km lungo il corso del torrente Alpone. Il corso d'acqua nasce dalle pendici del Monte Purga di Bolca (925 m), nello specifico nella località Scaronsi, all'interno del Comune di Vestenanova, e sfocia nel fiume Adige all'altezza dei Albaredo. È caratterizzato da una portata regolare durante i mesi invernali ed estivi, mentre diventa più impetuoso durante le piogge autunnali e primaverili; nel corso dei secoli, infatti, sono stati costruiti argini sopraelevati rispetto al piano campagna, al fine di contenere gli eventi di piena che si possono verificare.

La vallata del Chiampo, che si estende nella parte orientale dei Monti Lessini (Prealpi Venete) rientra quasi interamente nel territorio della Provincia di Vicenza. Ha una forma stretta e allungata per un'estensione di circa 30 km e una lunghezza di appena 6 km.

Il corso d'acqua di discrete dimensioni che attraversa la valle è il torrente Chiampo, che si origina sul Monte Gramolon, che con i suoi 1600 m è la vetta più alta della zona; confluisce poi nel fiume Fratta-Gorzone che sfocia direttamente nell'Adriatico in prossimità di Chioggia (Colleselli et al., 1976; Barbieri et al., 1973).

I principali torrenti che attraversano queste valli sono: torrente Progno d'Illasi, torrente Tramigna, torrente Alpone e torrente Chiampo. Tutti e quattro appartengono al bacino idrografico del fiume Adige. La cui sorgente si trova a quota 1586 m s.l.m., in prossimità del Passo Resia, e si estende per circa 415 km lungo il Trentino Alto-

Adige e il Veneto; sfocia poi nel Mar Adriatico, nella località di Rosolina Mare, ricoprendo una superficie totale di 12100 km².

La caratteristica distintiva dei territori vallivi descritti in precedenza è l'economia agricola basata sulla viticoltura. Queste sono zone di produzione del Vino di Soave, del Valpolicella DOC, dell'Amarone della Valpolicella e del Reticioto DOC.



Figura 1: immagine satellitare che permette di collocare l'area studio nel più ampio contesto delle Prealpi Venete e della Pianura Veneta. Sono evidenziati i Monti Lessini e le piane vallive che si estendono da NNW a SSE: la Val d'Ilasi, la Val Tramigna, la Val D'Alpone e la Valle del Chiampo.

2.2 Inquadramento geologico

I Monti Lessini sono rappresentati da un tavolato carbonatico caratterizzato da una monoclinale con pendenza degli strati da Nord a Sud, che attualmente si sviluppa con una inclinazione di circa 5° verso Sud, decrescendo da una massima elevazione settentrionale di 1800 m a un minimo di 40 m al limite meridionale del promontorio lessineo, prima della sua immersione sotto la copertura sedimentaria quaternaria dell'alta Pianura Veronese (Zampieri, 2000). A livello strutturale fanno parte dell'unità tettonica della Alpi Meridionali (Sudalpino), ovvero la porzione di rilievi alpini che si collocano a Sud rispetto alla Linea Insubrica o Lineamento Periadriatico. Si tratta di un importante lineamento tettonico, attivo dall'Oligocene, che si estende per circa 700 km dal Piemonte fino alla Slovenia, con una direzione principale E-W. È una struttura tettonica complessa che si divide in diversi settori con differente orientazione e caratteristiche cinematiche. I segmenti che la costituiscono, procedendo da Ovest a Est sono: la Linea del Canavese (orientazione SW-NE) e la Linea del Tonale (andamento W-E), ovvero una zona di faglia a scala crostale che immerge ripidamente a nord, caratterizzato sia da back thrusting che da movimento strike – slip destro (Müller et. al., 2001); continuando verso oriente si trovano la Linea delle Giudicarie (ad andamento SSW – NNE), che rappresenta un considerevole cambiamento della direzione dello strike del Lineamento Insubrico; come ultimo si trova la Linea Pusteria-Gailtal-Caravanche (orientata WNW-ESE) (Castellarin & Vai, 1982).

Il settore del Sudalpino, pertanto, può essere suddiviso in delle “unità neotettoniche” che si differenziano tra loro per caratteristiche strutturali e morfologiche e per comportamento neotettonico (Zanferrari et al., 1982). L'area dei Monti Lessini presenta uno stile di deformazione essenzialmente di tipo rigido (Zanferrari et. al., 1982), i cui principali lineamenti strutturali (Figura 2) si snodano rispecchiando l'andamento dei sistemi collocati ai margini occidentali ed orientali dei Lessini: il Sistema delle Giudicarie e la Linea Schio-Vicenza.

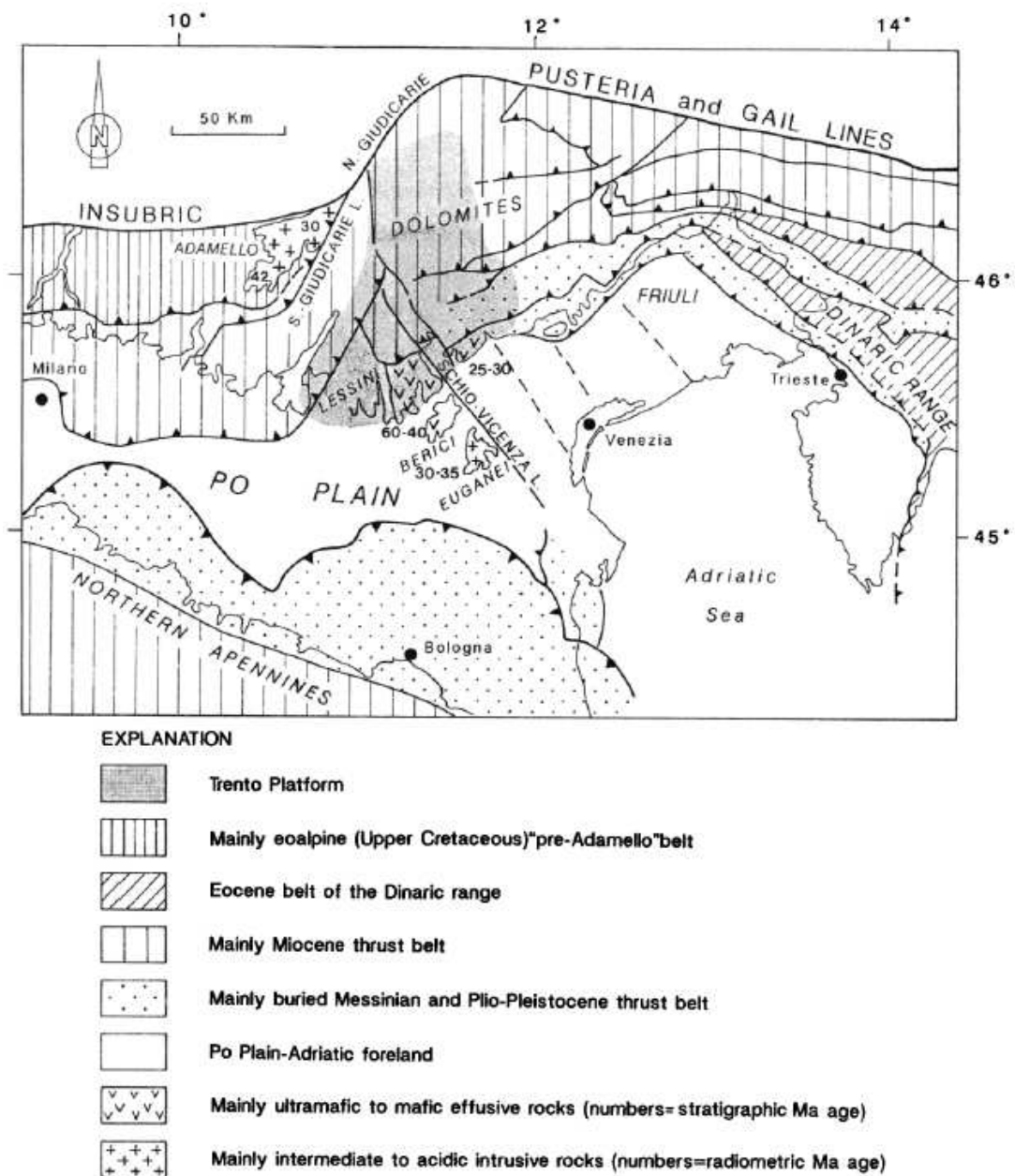


Figura 2: Assetto strutturale dell'area centro-orientale del sud alpino. In Grigio è evidenziata l'area della Piattaforma carbonatica Trentina, di cui anche i Monti Lessini. (Zampieri 1995, modificato da Castellarin et al. 1992).

La linea delle Giudicarie separa la Piattaforma Trentina, di cui i Monti Lessini sono l'affioramento più meridionale, dal Bacino Lombardo. Durante il Giurassico e il Cretaceo è stata caratterizzata da una tettonica di tipo estensionale, a cui è seguito un movimento di tipo transpressivo sinistro, durante il Miocene.

La faglia Schio-Vicenza, invece, separa nettamente l'area montana dei Lessini dalla Pianura alluvionale Veneta (Zampieri, 1995). È un elemento tettonico di particolare rilievo per via della sua elevata estensione e dell'entità dei suoi rigetti.

Ha agito come elemento cinematico trascorrente destro a partire dall'Oligocene-Miocene fino al Tortoniano medio-superiore, assumendo successivamente un carattere definitivo di trascorrente sinistra a partire dal Messiniano fino alle fasi deformative più recenti (Cantelli et al., 1994).

La linea tettonica più importante presente nel territorio oggetto di studio è la faglia di Castelvero (Figura 3). Si colloca all'interno della Val d'Alpone e separa i Lessini centro-occidentali, caratterizzati principalmente da depositi di tipo sedimentario, dai Lessini orientali a carattere magmatico. Si tratta di una faglia subverticale, diretta, con immersione tendente ad Est con direzione NNW-SSW (Carton et al., 1985).

Insieme alla Linea Schio-Vicenza, la faglia di Castelvero ha svolto un ruolo cruciale nella predisposizione delle condizioni necessarie per la formazione del graben Alpone-Agno. Questa struttura si è sviluppata in un contesto di tettonica estensionale, accompagnata da una fase di intensa attività vulcanica mafica e ultramafica (Zampieri, 1995). Il thrust di Marana (E-W) limita a nord l'estensione del graben, che raggiunge una lunghezza totale di circa 35 km.

Anche nella zona dei Lessini veronesi sono riconoscibili due sistemi di graben: il primo, collocato nella parte non-occidentale, ha un'orientazione NNW; mentre il secondo, orientato NNE, si posiziona nella porzione centro-orientale. La faglia di confine principale di questo secondo sistema è la faglia di Boscochiesanuova. Si tratta di una grande frattura che, con notevole estensione (circa 20 km), interessa tutte le formazioni geologiche dei Lessini occidentali (Bosellini et al., 1967).

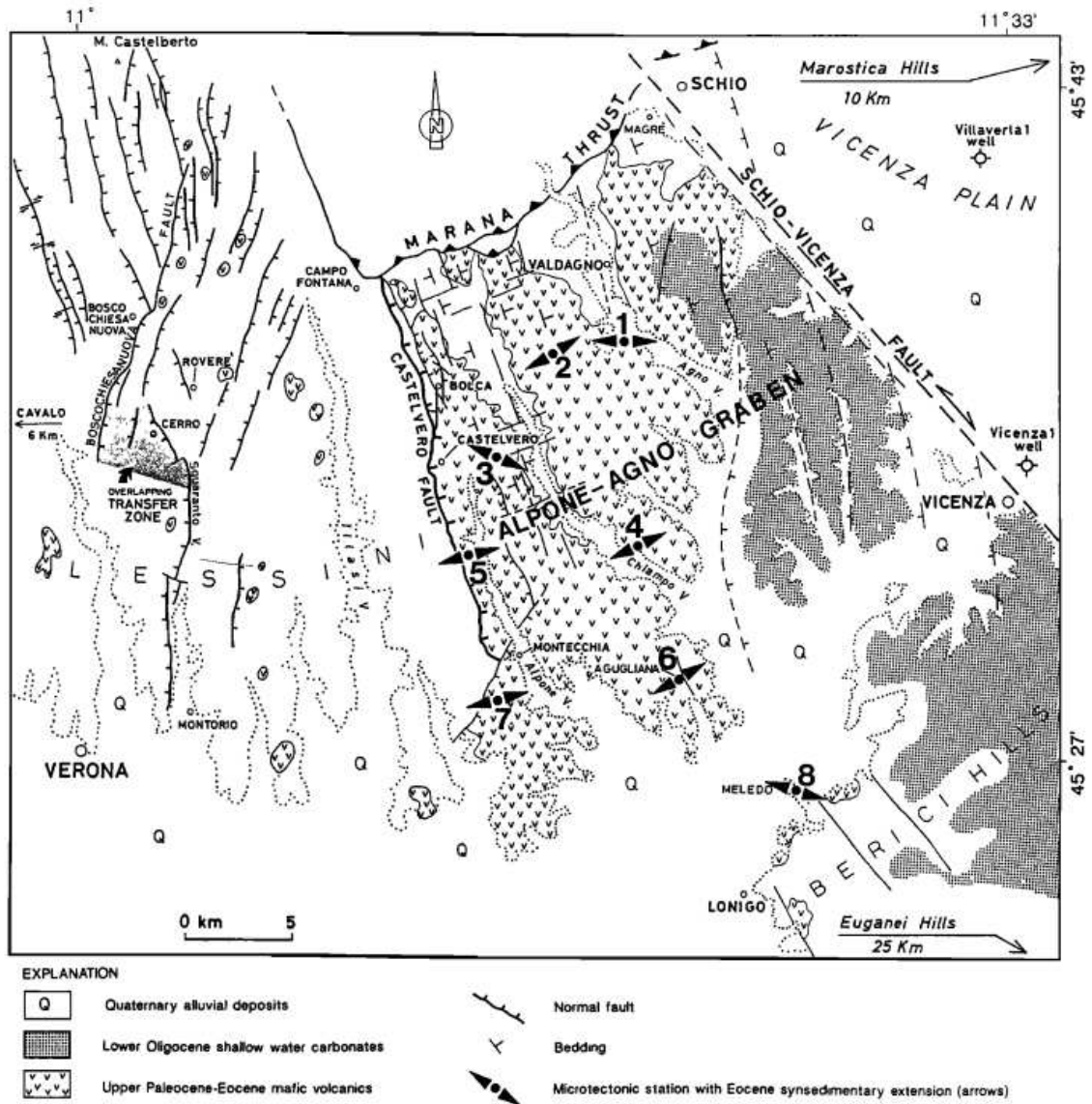


Figura 3: Mappa strutturale dei Lessini Veronesi, in cui sono riconoscibili le strutture tettoniche che li caratterizzano, e dei Lessini Vicentini, in cui sono riconoscibili il graben Alpone-Agno e i depositi Paleogenici di natura vulcanica. (Zampieri, 1995)

Dal punto di vista litologico l’Era Mesozoica (250-65 Ma) segna l’inizio della storia geologica dei Monti Lessini. Durante il Triassico, la maggior parte delle masse continentali odierne erano ancora unite nella Pangea, limitata a oriente dalla Paleotetide. Questo esteso oceano si stava progressivamente chiudendo in favore dell’espansione nella Neotetide, originata da una fase di rifting di età Permiana. I diversi cicli di fluttuazione del livello marino hanno determinato un’estrema varietà di facies sedimentarie. Nell’area studio sono presenti evidenze deposizionali relative al Triassico Superiore, in particolare al Ladinico; questo periodo è caratterizzato da

un'ampia piattaforma carbonatica che si estendeva dall'attuale Lombardia occidentale fino al confine con la Slovenia (Trombetta, 2010).

Verso la fine del Triassico e l'inizio del Giurassico, l'intero scenario paleogeografico del sudalpino cambia radicalmente. La fase di subsidenza differenziale e l'instaurarsi di un regime distensivo attribuito alla formazione del margine passivo sono caratterizzate da elementi paleogeografici e paleostrutturali che si espandono in direzione nord-sud, e sono noti in letteratura come Bacino Lombardo, Piattaforma Trentina, Bacino di Belluno Piattaforma Friulana (Bosellini et al., 1973), come si può osservare in Figura 4.

La Piattaforma di Trento, di cui i Monti Lessini sono la parte più meridionale, è un alto strutturale esteso su un'ampia area dell'Italia nord-orientale. È caratterizzata da una prima fase di sedimentazione di mare sottile (Lias) seguita da una seconda fase di deposizione pelagica (Dogger e Malm) determinata da un parziale sprofondamento della piattaforma stessa; il suo progressivo annegamento non è simultaneo a quello della Piattaforma Friulana che rimane attiva fino all'Eocene.

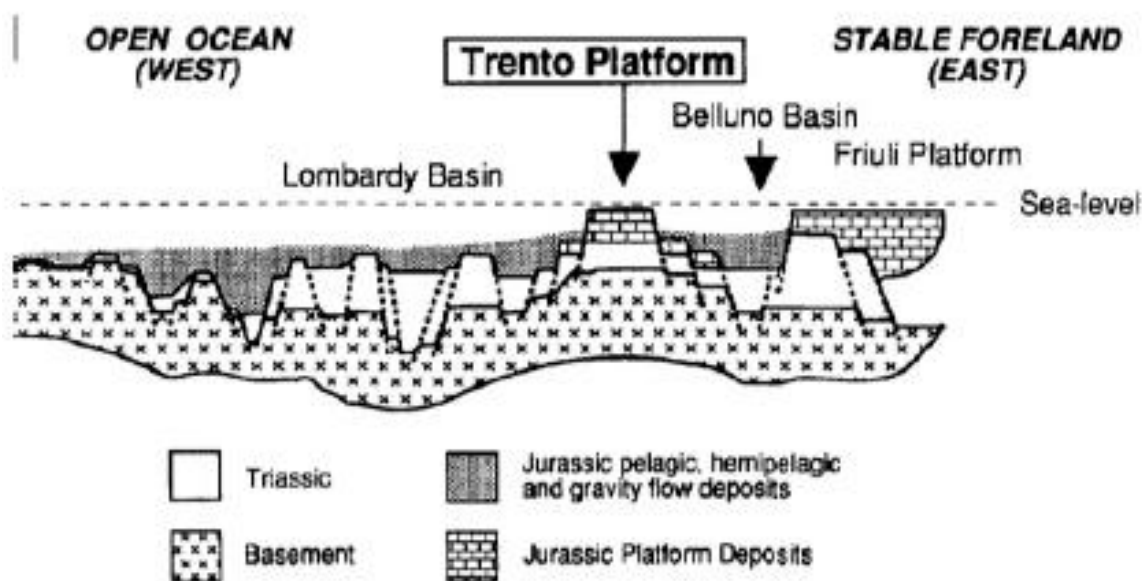


Figura 4: Sezione trasversale attraverso le Alpi Meridionali nel Giurassico, in cui sono visibili i principali elementi paleostrutturali, da Ovest verso Est: Bacino Lombardo, Piattaforma Trentina, Bacino di Belluno e Piattaforma Friulana (Masseti et al., 1998).

Durante il Cretaceo la piattaforma diventa definitivamente un plateau pelagico, in cui si ha una deposizione tipicamente bacinale; in questo contesto si raggiunge la massima profondità marina. Successivamente, infatti, inizia una fase di deformazione

compressiva che porta alla scomparsa dell'Oceano Ligure-Piemontese e alla manifestazione della tettonica alpina. Nell'area dell'attuale Veneto occidentale e Trentino meridionale, come conseguenza di questa progressiva chiusura, si sviluppa una tettonica di tipo estensionale coeva con l'attività magmatica di tipo mafico e ultramafico del Paleogene. Questo magmatismo è associato ad una situazione geodinamica di rift intracontinentale con assottigliamento della crosta, aumento del gradiente termico e risalita di materiale proveniente dal mantello superiore (De Vecchi et al., 1976).

L'attività vulcanica non è limitata a un'unica area, ma mostra un progressivo spostamento da ovest verso est, accompagnato da un'espansione verso nord. Questa migrazione avvenne attraverso cicli eruttivi successivi.

I primi eventi, quasi totalmente sottomarini, si verificarono a partire dal Paleocene superiore, localizzati nelle Valli del Chiampo e d'Agno; solo nell'Eocene inferiore, l'attività si estende al veronese e nei dintorni di Rovereto. Nell'Eocene medio le manifestazioni vulcaniche si spostano anche nei Colli Berici; questo ciclo inizia in condizioni subacquee, come dimostrato dalla presenza abbondante di ialoclastiti e fossili marini, ma si conclude in ambiente subaereo, come dimostra la presenza di flora e fauna all'interno dei tufi e la tessitura delle rocce (Piccoli, 1966). Nell'Eocene superiore, in corrispondenza di una fase di inattività dei Lessini e dei Berici, inizia un'attività eruttiva basica in corrispondenza dei Colli Euganei. Nel Marosticano si effusero grandi quantità di basalti soltanto nell'Oligocene.

I prodotti vulcanici si depositano all'interno del graben di Alpone-Agno con spessori fino a 500 m, a causa dell'elevata subsidenza nell'area.

2.3 Successione litostratigrafica

Le rocce affioranti all'interno dell'area studio sono principalmente di origine sedimentaria e vulcanica, di età compresa tra il Triassico superiore e il Miocene. In superficie è presente anche una copertura quaternaria molto estesa, concentrata soprattutto nei fondivalle e nella parte meridionale dell'area studio, dove inizia la Pianura Padana.

Per la ricostruzione stratigrafica dell'area, sono stati utilizzati i lavori di Bosellini A.

et al. (1967) in merito alle note illustrative della Carta Geologica d'Italia alla scala 1:100.000 (foglio 49 di Verona), e quelli di Antonelli et al. (1990) relativi alla Carta Geologica del Veneto in scala 1:250.000. Inoltre, gli studi specialistici di Sauro U. sui Monti Lessini si sono rivelati fondamentali per identificare la sequenza litologica caratteristica della zona.

Segue una descrizione dettagliata e cronologica delle formazioni che compongono i Monti Lessini, la Valle d'Illasi, la Val Tramigna, la Val d'Alpone e la Valle del Chiampo, ordinata dalla più antica alla più recente:

Dolomia Principale (Triassico Superiore): è costituita da bancate di dolomie bianche e rosate, con intercalazioni di calcari bituminosi. Si originano in ambiente di piattaforma carbonatica soggetta a cicli periodici, che stratificano la Dolomia Principale in layers con caratteristiche deposizionali differenti (ciclo supratidale, ciclo intertidale e ciclo subtibale). L'elevato spessore della formazione (1000 m) è generato dalla lenta ma continua subsidenza della piattaforma carbonatica triassica. All'interno dell'area di studio affiora nell'alta Valle d'Illasi, in corrispondenza del Massiccio del Carega delle Piccole Dolomiti, dove è visibile tutta la sua potenza. Si hanno riscontri anche in corrispondenza dell'alta Valle del Chiampo.

Calcari Grigi di Noriglio (Lias inferiore e medio): nella cartografia più recente è conosciuto come Gruppo dei Calcari Grigi. La sua deposizione avviene nel contesto dell'alto strutturale della Piattaforma Trentina, infatti, in letteratura, viene spesso identificata come la successione carbonatica di mare sottile del Giurassico inferiore. È costituito complessivamente da banchi calcarei compatti e da calcari marnosi, che raggiungono uno spessore di 400-500 m.

Può essere suddiviso in tre membri distinti per i sedimenti caratteristici di ambienti diversi:

- La Formazione di Monte Zugna rappresenta il membro inferiore dei Calcari Grigi. La sedimentazione si realizza in contesto di piana tidale, in cui sono riconoscibili tre cicli.
- Il Calcare Oolitico di Loppio rappresenta il membro medio dei Calcari Grigi. La loro formazione è associata ad una generale tendenza di inabissamento

della Piattaforma trentina, registrato simultaneamente in altri settori della Tetide.

- La Formazione di Rotzo è il membro superiore e più rappresentativo del Gruppo dei Calcari Grigi. È caratterizzata da una deposizione in ambiente di rampa-laguna e dalla presenza di grandi bivalvi anomali, i Lithiotis (bancate a Lithiotis).

All'interno dell'area studio questa Formazione affiora dove la valli hanno maggiormente inciso i rilievi montuosi (Bosellini et al., 1967). Pertanto, si può osservare nelle alte Valli del Chiampo e d'Illasi.

Formazione dell'Oolite di San Vigilio (Lias superiore): è costituita da calcari giallastri e bianchi, affioranti in quasi tutta l'area studio con uno spessore che diminuisce progressivamente da Ovest verso Est, fino a sparire completamente. La sua deposizione avviene in un contesto di mare basso ad alta energia, associato però alla tendenza di annegamento della Piattaforma di Trento.

Rosso Ammonitico Veronese (Malm e Dogger): si tratta della prima Formazione di ambiente pelagico di questa successione stratigrafica; la deposizione dei sedimenti avviene ad una profondità superiore ai 200 m, ma inferiore alla CCD. Queste condizioni si verificano in seguito allo sprofondamento della Piattaforma di Trento, ora Plateau di Trento, determinato da una fase di tettonica distensiva.

È un calcare micritico molto nodulare ricco in ammoniti, mal stratificato con dei calcari marnosi. La presenza di un hardground permette di dividere il Rosso Ammonitico Veronese in due membri: il Rosso Ammonitico Inferiore e il Rosso Ammonitico superiore.

Nella parte dei Lessini compresa entro il Foglio 49 Verona, il Rosso Ammonitico Veronese sovrasta direttamente i Calcari Oolitici di San Vigilio, con una lacuna stratigrafica estesa a quasi tutto il Baiociano, altrove è rappresentato dalla nota Lumachella a Posidonia alpina (Bosellini et al., 1967). I suoi affioramenti sono ampiamente visibili all'interno dell'area studio, nello specifico nelle Valli del Chiampo e d'Illasi, con una potenza che va dai 25 m ai 30 m, determinato da una sedimentazione molto rallentata.

Maiolica o Biancone (Turoniano - Cretaceo Inferiore): questa formazione si origina per la deposizione di fanghi micritici in un contesto di ulteriore approfondimento dei fondali oceanici (circa 400-500 m). Si tratta di un calcare stratificato di colore bianco-grigio, ricco di noduli di selce, con uno spessore variabile di 80 - 150 m.

Complesso dolomitico indifferenziato (Giurassico inferiore – Cretaceo inferiore): sono dolomie e calcari dolomitici giallastri e rosa-violaceo, poco o per nulla stratificati (Bosellini et al., 1967), affioranti nelle Valli d’Illasi e del Chiampo.

Scaglia Rossa Veneta (Cretaceo Superiore): si tratta di una serie bacinale costituita da calcari marnosi rossi a stratificazione scagliosa con la presenza di caratteristici hard-ground. Deriva dall’accumulo di fanghi calcarei costituiti da nanofossili calcarei e foraminiferi planctonici; il suo colore rossastro è dovuto dalla presenza di ematite (Giusberti et al., 2020).

Calcari di Spilecco (Eocene inferiore e Paleocene): è una Formazione poco potente (10-15 m) costituita da calcari marnosi con dei sottili livelli di marne. Affiorano in pochi punti dell’area studio, in particolare in prossimità della località di Bolca e nell’alta valle del Chiampo.

Calcari nummulitici o Calcari di Chiampo (Eocene inferiore e medio): sono calcari marnosi compatti molto ricchi in fossili foraminiferi, tra cui i Nummuliti a cui devono il nome. Hanno uno spessore complessivo di circa 150 m. Sono riconoscibili due membri principali separati da un livello di rocce clastiche stratificate di natura basaltica, derivante del rimaneggiamento di ialoclastiti, breccie d’esplosione e altri prodotti dell’attività effusiva sottomarina (Zampieri, 2022).

La deposizione è avvenuta all’interno del Graben di Alpone-Agno durante una fase di parziale quiete vulcanica; pertanto, all’interno dell’area studio, affiora nelle Valli del Chiampo e nella Valle d’Alpone.

Formazioni Eruttive (Cretaceo superiore – Miocene inferiore): si tratta di tufi, ialoclastiti, lave basaltiche e breccie di esplosione formati in questo periodo di più o meno intensa attività vulcanica, principalmente sottomarina. Il chimismo basaltico delle formazioni è uniforme e costante nel tempo, differentemente dall’eterogeneità dei prodotti vulcanici causata dalla meccanica e dall’ambiente di formazione. Come

spiegato in precedenza le eruzioni sono coeve alla formazione del Graben di Alpone – Agno, e la loro deposizione avviene ciclicamente all'interno dei confini di questa struttura. Considerando l'area studio, infatti, si distingue la regione vulcanica, ad Est della Faglia di Castelvero (Valli d'Alpone e del Chiampo) e la regione a deposizione calcarea, a Ovest della Faglia di Castelvero (Valli d'Illasi e Tramignia).

I prodotti eruttivi sono spesso alterati, argillificati e facilmente erodibili, per questo motivo le rocce vulcaniche non affiorano direttamente alla superficie, ma sono ricoperte da una coltre eluviale e colluviale argillosa formata dai prodotti della degradazione superficiale delle vulcaniti stesse (Barbieri et al., 1973).

Depositi quaternari: le coperture presenti in questa area sono di primaria importanza per quanto riguarda la stabilità dei versanti. Si possono trovare depositi di diverse tipologie:

- *Alluvioni di fondovalle*: sono formati da sedimenti trasportati dai fiumi e accumulati nelle zone più basse della valle. Questi sedimenti, composti da sabbia, ghiaia, limo e argilla, si depositano quando la corrente fluviale rallenta e perde forza, lasciando cadere il materiale che trasportato. Parte della pianura compresa nell'area studio è costituita da alluvioni, generalmente grossolane e terrazzate, appartenenti all'antica conoide dell'Adige e dalle alluvioni più recenti che hanno sbarrato lo sbocco delle valli che scendono dai Lessini (Bosellini et al., 1967).
Le successive alluvioni dei Monti Lessini hanno sovrastato quelle originate dal fiume Adige, distinguendosi da esse anche per la diversa natura litologica: non contengono elementi porfirici tipici del complesso vulcanico atesino.
- *Detrito di versante o depositi colluviali*: consistono in un accumulo di materiale ai piedi di un pendio. I detriti che li compongono si originano principalmente dai rilievi sovrastanti; infatti, non sono soggetti a fenomeni di trasporto, se non il ruscellamento.
- *Depositi eluviali*: si generano per alterazione ed erosione del materiale presente in situ. Non subiscono trasporto, pertanto si accumulano nel luogo di origine.

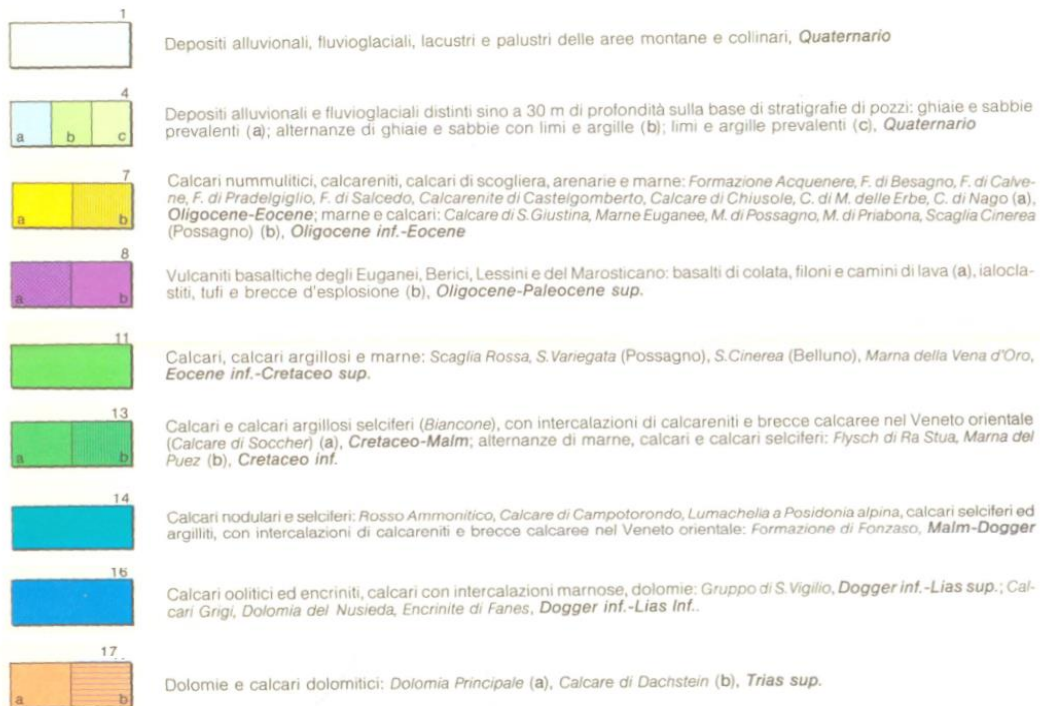
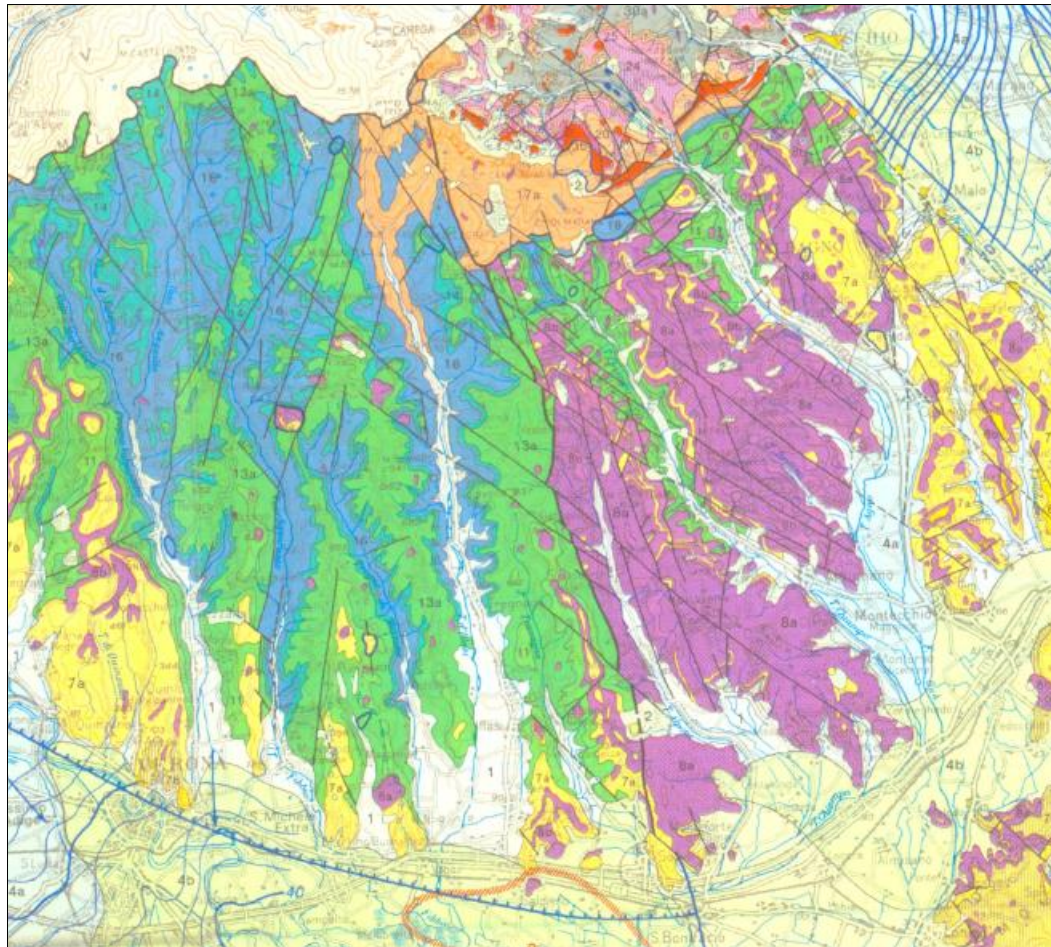


Figura 5: Estratto della Carta Geologica del Veneto alla scala 1:250000 (Antonelli et al., 1988)

3 Materiali e Metodi

3.1 Materiali

3.1.1 Inventario dei Fenomeni Franosi in Italia (IFFI)

Il termine frana fa riferimento ad un movimento gravitativo di una massa di roccia, terra o detrito lungo un versante (Cruden, 1991). L'Italia, a causa della sua geologia e morfologia, è esposta ad un elevato rischio di frana. Per questo motivo, nel 1997, viene fondato il progetto IFFI (Inventario dei Fenomeni Franosi in Italia), i cui documenti sono consultabili e scaricabili grazie alla piattaforma IdroGeo (idrogeo.isprambiente.it/app/iffi). Il progetto IFFI è realizzato da ISPRA (Istituto superiore per la Protezione e la Ricerca Ambientale), dalle Regioni e Province Autonome allo scopo di fornire un quadro dettagliato delle frane presenti sul territorio nazionale. L'obiettivo è fornire una banca dati il più possibile completa, utilizzabile nei vari contesti, pubblici e privati, per la gestione e progettazione del territorio. Ad oggi, in Italia, sono censiti 635026 fenomeni, mentre in Veneto le frane catalogate da IFFI sono 9443, di cui 61 sono soggette ad attività di monitoraggio specifico.

Ogni fenomeno franoso viene rappresentato con tre differenti modalità, in base al contesto, tramite:

- Punto Identificativo del Fenomeno Franoso (PIFF) collocato in corrispondenza del coronamento della frana.
- Poligono se la superficie della frana è cartografabile.
- Linea se la larghezza della frana non è cartografabile.

Per ogni frana sono indicati i dettagli specifici di quel fenomeno, tra cui la collocazione, la litologia caratteristica del dissesto, lo stato di attività e la sua data di attivazione, e per finire, quando possibile, il tipo di movimento.

Di primaria importanza è la classificazione del fenomeno: nel corso del tempo diversi studiosi hanno dato il loro contributo all'individuazione del miglior metodo per classificare i fenomeni; le più importanti sono quelle basate sulle osservazioni visive di (Varnes, 1978), (Cruden & Varnes, 1996) e di (Hungry et al., 2001). Esiste anche una classificazione geotecnica, generalmente utilizzata nel contesto dell'analisi

di stabilità, elaborata da (Hunchinson, 1988).

Il metodo di classificazione più utilizzato è quello di (Varnes, 1978), che prevede una distinzione degli eventi sulla base del tipo di materiale coinvolto e del tipo di movimento.

Tipo di movimento	Materiale		
	Roccia	Detrito	Terra
Crollo	Crollo di roccia	Crollo di detrito	Crollo di terra
Ribaltamento	Ribalt. di roccia	Ribalt. di detrito	Ribalt. di terra
Scivolamento rotazionale	Sciv. Rot di roccia	Sciv. Rot di detrito	Sciv. Roto di terra
Scivolamento traslativo	Sciv. Trasl. di roccia	Sciv. Trasl. di detrito	Sciv. Trasl. di terra
Espansione laterale	Esp. di roccia	Esp. di detrito	Esp. di terra
Colamento	Col. di roccia	Col. di detrito	Col di terra

Tabella 1: Classificazione delle frane secondo (Varnes, 1978)

Come detto in precedenza anche lo stato di attività di una frana è importate al fine del monitoraggio e della progettazione territoriale. (Cruden & Varnes, 1996) hanno proposto queste definizioni per la sua valutazione:

- Attiva: attualmente in movimento.
- Sospesa: non è attualmente in movimento, ma si è mossa nell'ultimo ciclo stagionale.
- Riattivata: nuovamente attiva, dopo essere stata sospesa o inattiva.
- Inattiva: si è mossa per l'ultima volta prima dell'ultimo ciclo stagionale.

Le frane inattive possono essere ulteriormente suddivise in:

- Quiescenti: possono essere riattivate solo per mezzo delle loro cause scatenanti.
- Naturalmente stabilizzate: sono state protette dalle loro cause scatenanti senza l'intervento di eventi antropici.

- Artificialmente stabilizzate: sono state protette dalle loro cause scatenanti per mezzo di misure di stabilizzazione
- Relitte: si sono sviluppate in contesti geomorfologici o climatici diversi da quelli attuali.

Nell'area studio l'inventario IFFI include 642 punti in frana, classificati come si può osservare nella Figura 6.

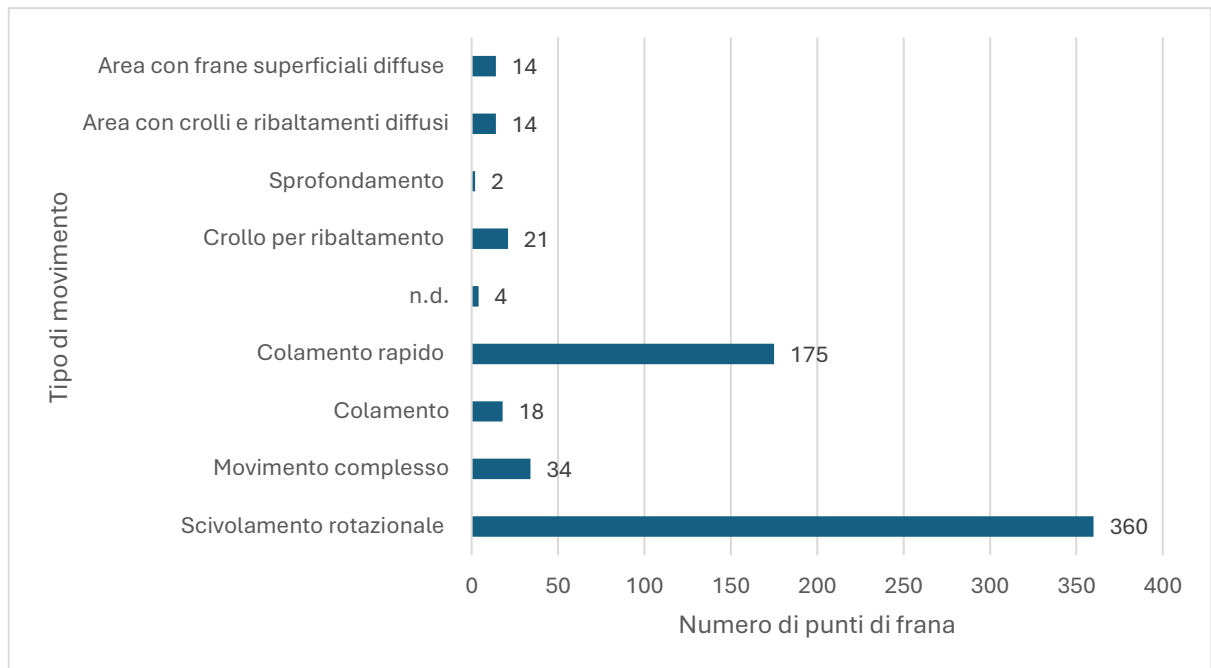


Figura 6: Classificazione dei punti in frana ricavati dall'inventario IFFI all'interno dell'area studio. Si può osservare il numero di punti in frana per ogni tipologia di movimento franoso.

Gli scivolamenti rotazionali, che sono i movimenti gravitativi più abbondanti (56%), consistono nello scorrimento verso valle di una massa di terreno o roccia, lungo una superficie curvilinea di neoformazione. Il corpo di frana in movimento si disloca spesso in più blocchi che creano molteplici scarpate, oltre a quella principale.

Il colamento e il colamento rapido sono rispettivamente il 3% e il 27% dei punti in frana nell'area studio. Assumono una diversa denominazione in base alla tipologia di materiale coinvolto, al contenuto d'acqua e alla velocità (Hungry et al., 2001). I più comuni sono le colate di detrito o debris flow, ovvero un insieme di particelle di sedimento di varie dimensioni e acqua che scorre lungo una regione confinata a forma di canale fino alla sua estremità, dove diventa non confinata e si diffonde formando una massa a forma di ventaglio. La profondità del flusso va da 1 m a

diversi metri, le velocità medie possono raggiungere i 15 m/s e le pendenze del canale variano da 40° nella zona di innesco a circa 3° nella zona di deposizione (Hunter et al. 1994).

Circa il 5% dei punti in frana sono caratterizzati da un movimento complesso, ovvero che presenta due o più tipologie di movimenti in sequenza.

Infine le frane di crollo per ribaltamento sono il 3%, e consistono in una veloce caduta per gravità dell'ammasso roccioso, attorno al proprio asse posto al di sotto del baricentro. Interessa versanti verticali o sub-verticali in roccia, in cui sono presenti fratture preesistenti al crollo che isolano prismi di materiali.

La collocazione degli eventi gravitativi all'interno della zona di indagine è fortemente influenzata dalla geologia. La distribuzione della popolazione dei punti evidenzia, infatti, una densità maggiore nell'area centro-orientale della sezione (Figura 7), dove si impone una litologia di origine vulcanica. Questa abbondanza di fenomeni si trova in corrispondenza dei versanti vallivi delle valli d'Alpone e del Chiampo, collocate all'interno del graben Alpone-Agno. Nel corso del tempo sono stati fatti diversi studi e indagini sulla franosità e sugli eventi collocati in queste vallate; tra questi (Colleselli et al., 1976) e (Barbieri et al., 1973).

Anche nella zona settentrionale dell'area studio, si evidenzia una presenza consistente di punti in frana. Si tratta dell'alta valle d'Illasi ai piedi del Gruppo del Carega.

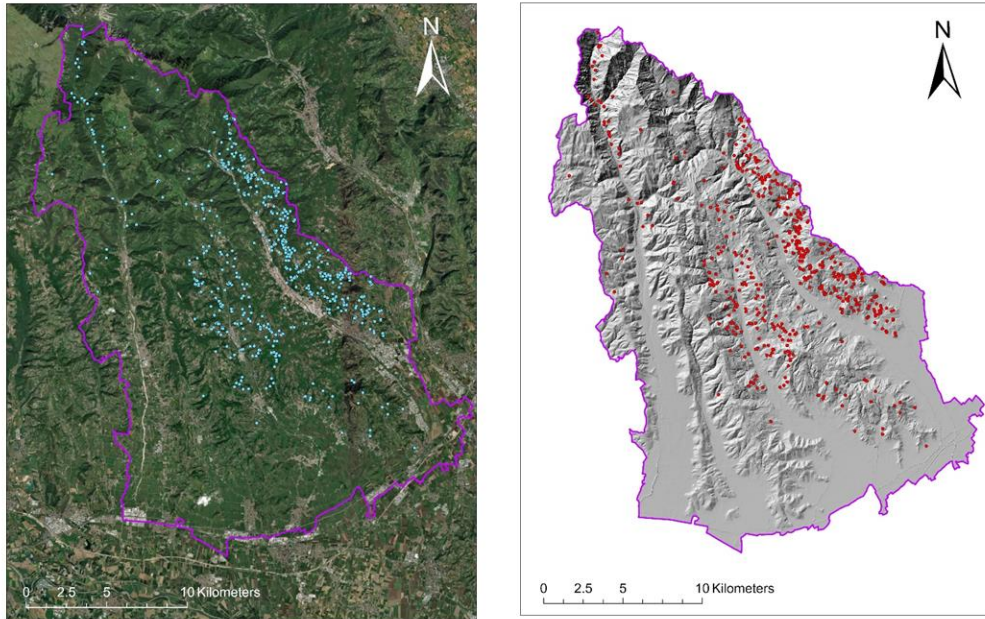


Figura 7: Sulla sinistra l'immagine satellitare della distribuzione dei punti di frana all'interno dell'area studio, ricavati dall'inventario dei fenomeni franosi italiani (IFFI). Sulla destra l'immagine su base hillshade della distribuzione dei punti di frana all'interno dell'area studio, ricavati dall'inventario dei fenomeni franosi italiani.

3.1.2 Slope Units (SU)

La gestione delle frane è un considerevole problema per la società; questi fenomeni naturali hanno un impatto elevato sulla popolazione, sia in termini di perdite economiche, quindi danni alle infrastrutture, alle attività economiche e proprietà private, ma anche in termini di sociali, paesaggistici e storico-culturali. È, dunque, importante indagare e conoscere il problema per poter pianificare gli interventi di previsione e prevenzione necessari (Frattini & Crosta, 2006). L'analisi del rischio ha lo scopo di prevedere l'accadimento, l'evoluzione e le possibili conseguenze di un evento, in modo da poter agire sul territorio per mitigarne e ridurne l'impatto. Nel 1984 Varnes e la commissione tecnico scientifica dell'UNESCO, nell'ambito dell'International Association of Engineering Geology (IAEG), hanno definito il termine rischio come segue:

$$R = H(I)V(I)E \quad (3.1)$$

Con V si intende la vulnerabilità, ovvero il grado di perdita prodotto su un elemento, o su una serie di elementi, esposti al rischio risultante dal verificarsi di un evento dannoso di una data intensità. Il suo valore si esprime in una scala da 0 (nessuna perdita) a 1 (perdita totale). Esprime la correlazione esistente tra l'intensità di un evento e il danno atteso.

Con E si intendono l'esposizione di un elemento o di un insieme di elementi alla probabilità che in un dato periodo di tempo possano essere investiti da un evento di una determinata intensità. Si tratta quindi del valore esposto della zona.

L'Hazard (H) è la pericolosità di frana, ovvero la probabilità che un fenomeno potenzialmente distruttivo di una determinata intensità si verifichi in un dato periodo di tempo ed in una data area (Varnes et al. 1984).

Questa definizione pone l'attenzione su tre concetti fondamentali per la valutazione della pericolosità di frana:

L'intensità dell'evento (I), ovvero l'energia con cui il fenomeno si verifica; quantifica il potenziale dello stesso di provocare danni in termini di impatto diretto e indiretto. È l'energia cinetica del movimento del versante:

$$E_k = \frac{1}{2}mv^2 \quad (3.2)$$

È legata alla massa di materiale spostato e alla velocità, che, avendo dipendenza quadratica, influisce maggiormente sull'intensità della frana.

La componente temporale è associata al momento in cui l'evento gravitativo potrebbe verificarsi. Si fonda sull'analisi della frequenza con cui il fenomeno si è manifestato, determinata attraverso l'osservazione di foto aeree o ortofoto del territorio in esame e mediante lo studio degli annali pluviometrici, per verificare le precipitazioni registrate nel corso degli anni.

La tecnologia InSAR è un altro strumento fondamentale per l'analisi del pericolo temporale, grazie alla sua capacità di acquisizione dei dati continua e regolare. È un sistema di monitoraggio ad alta risoluzione, che permette di osservare le deformazioni del suolo a scala globale (Costantini et al., 2021).

La componente spaziale o suscettibilità è la probabilità che una frana avvenga in una determinata area. Supponendo che le frane si verifichino in futuro a causa delle stesse condizioni che le hanno prodotte in passato, la valutazione della suscettibilità può essere utilizzata per prevedere le posizioni geografiche delle future frane (Guzzetti et al., 2006), tramite la creazione di una mappa di suscettibilità di frana (LSM). Questo principio è valido solo entro determinati limiti, i dati storici che si devono utilizzare, infatti, devono rispecchiare la situazione attuale. Producendo una LSM si stima la probabilità spaziale relativa di un evento franoso, che fornisce uno strumento prezioso in supporto alle decisioni politiche in merito alla riduzione del rischio (Jacobs et al., 2020).

Un importante prerequisito per la modellazione della suscettibilità di frana è la selezione delle unità di base della mappa; con questo termine si indica una porzione della superficie terrestre che contiene un insieme di condizioni del terreno differenti dalle unità adiacenti, con confini definibili (Hansen, 1984). Alla scala dell'analisi, un'unità di base rappresenta un dominio che massimizza l'omogeneità interna e l'eterogeneità tra le unità (Guzzetti et al., 1999).

Secondo (Guzzetti et al., 1999) e (Reichenbach et al., 2018) tutti i metodi per l'identificazione delle unità base, presenti in letteratura, possono essere riassunti nei seguenti 5 gruppi:

- Celle di una griglia o pixel: con questo metodo, l'area viene suddivisa in celle quadrate di uguali dimensioni. La grandezza delle celle è determinata in base alla risoluzione spaziale del Digital Elevation Model (DEM).
A ciascuna cella viene attribuito un valore per ogni fattore predisponente preso in considerazione nello studio della suscettibilità di frana.
Secondo Martinello et al. (2021) questo metodo offre buone prestazioni quando si conosce esattamente la posizione di innesco di una frana. Sono quindi troppo localizzate per esprimere condizioni di instabilità relative a fenomeni franosi di dimensioni elevate. Le mappe prodotte con questo approccio pixel-based tendono a frammentare il territorio in piccole porzioni indipendenti tra loro, creando difficoltà nella lettura e nell'interpretazione. Appaiono come mosaici di celle singole con specifici livelli di suscettibilità, che non garantiscono una coerenza spaziale o una connessione logica tra i

pixel adiacenti su uno stesso pendio.

I pixel, come unità di base, sono utilizzati circa dell'84 % degli articoli (Reichenbach et al., 2018).

- Unità di terreno: sono le unità di base più utilizzate dai geomorfologi, si fondano sull'osservazione che, negli ambienti naturali, le interazioni tra materiali, forme del terreno e processi dinamici determinano confini che spesso riflettono a condizione unica.
- Metodo a poligoni con condizioni costanti o UCU (Unique Conditions Units): l'area viene divisa in poligoni derivanti dall'intersezione delle diverse mappe disponibili, ciascuno con valori costanti per ogni variabile. Questo è un processo di classificazione di più fattori legati all'instabilità del pendio che vengono combinati in una mappa o in un unico strato che rappresenta le condizioni complessive.
- Slope units (SU): secondo Alvioli et al. (2016) sono le unità di terreno che vengono delimitate da linee di drenaggio e spartiacque, utilizzando esclusivamente le informazioni ottenute da un modello digitale di elevazione (DEM). Hanno una forte relazione con le condizioni idrogeologiche, geomorfologiche e topografiche che modellano il paesaggio naturale; infatti, massimizzano l'omogeneità geomorfologica all'interno e massimizzano anche l'eterogeneità geomorfologica tra domini confinanti.

Con il metodo delle SU viene facilitata la gestione di dati eterogenei e la risoluzione di problemi legati all'accuratezza della posizione. Siccome l'intera cinematica della frana (zone di innesco, propagazione e accumulo), può avvenire all'intero della SU, con questa tecnica si producano risultati molto accurati (Martinello et al., 2021); anche per la previsione di fenomeni franosi molto estesi.

Una dinamica controversa riguardo alle Slope Units è la loro definizione cartografica: Carrara (1988) propone una suddivisione manuale del paesaggio in SU, partendo da una mappa topografica. Ovviamente, questa strategia non è applicabile su larga scala e può generare una elevata incertezza. Per questa ragione numerosi studiosi hanno lavorato per automatizzare questo processo. Non è un'opera facile, perché non esiste un'unica possibile suddivisione del

paesaggio in Slope Units; dipende da molteplici fattori, tra cui le dimensioni e la complessità dell'area studio, la qualità e la risoluzione dei dati disponibili sull'elevazione del terreno (Alvioli et al., 2016). Inoltre, i criteri adottati potrebbero influire negativamente sull'efficacia predittiva dei modelli di suscettibilità, portando a unità troppo grandi che non sono adatte ad individuare la zona di innesco e a discriminare la dinamica specifica della frana (Martinello et al., 2021).

Alvioli et al. (2016) hanno proposto un nuovo software per l'automatizzazione del processo di suddivisione del terreno in SU, partendo da un modello digitale di elevazione (DEM) e da un insieme di dati di input. Il software, chiamato *r.slopeunits*, è sviluppato in Python per GRASS GIS. In letteratura le SU come unità di base sono utilizzate in circa il 5.1% degli articoli; esistono però diversi studi che confrontano i risultati ottenuti dalla modellazione con pixel e Slope Units.

Le Slope Units sono state utilizzate come unità di base del terreno per generare la mappa di suscettibilità dell'area selezionata per questa tesi di laurea.

Lo shapefile delle SU del Veneto, elaborate con il software *r.slopeunits*, è di disponibile gratuitamente sul portale del gruppo di ricerca geomorfologico, con sede a Perugia, dell'Istituto di Ricerca per la Protezione Idrogeologica (IRPI). Questo istituto fa parte del Dipartimento Scienze del Sistema Terra e Tecnologie per l'Ambiente, del Centro Nazionale delle Ricerche (CNR).

Le unità di base individuate nella Regione sono 10252, ovviamente localizzate in prossimità dei rilievi. Per gli scopi di questa tesi, sono state selezionate inizialmente 520 SU, tutte contenute parzialmente o completamente all'interno dell'area studio. Per minimizzare la possibilità di errori nell'elaborazione finale della LSM, si è deciso di escludere le unità collocate sul contorno dell'area selezionata, la cui superficie inclusa all'interno della sezione era troppo piccola per essere rilevante. Il numero totale di SU finale (Figura 8) ammonta a 491 che occupano circa 397 km^2 , ovvero l'83% della superficie dell'area studio, la cui estensione totale è di 476 km^2 .

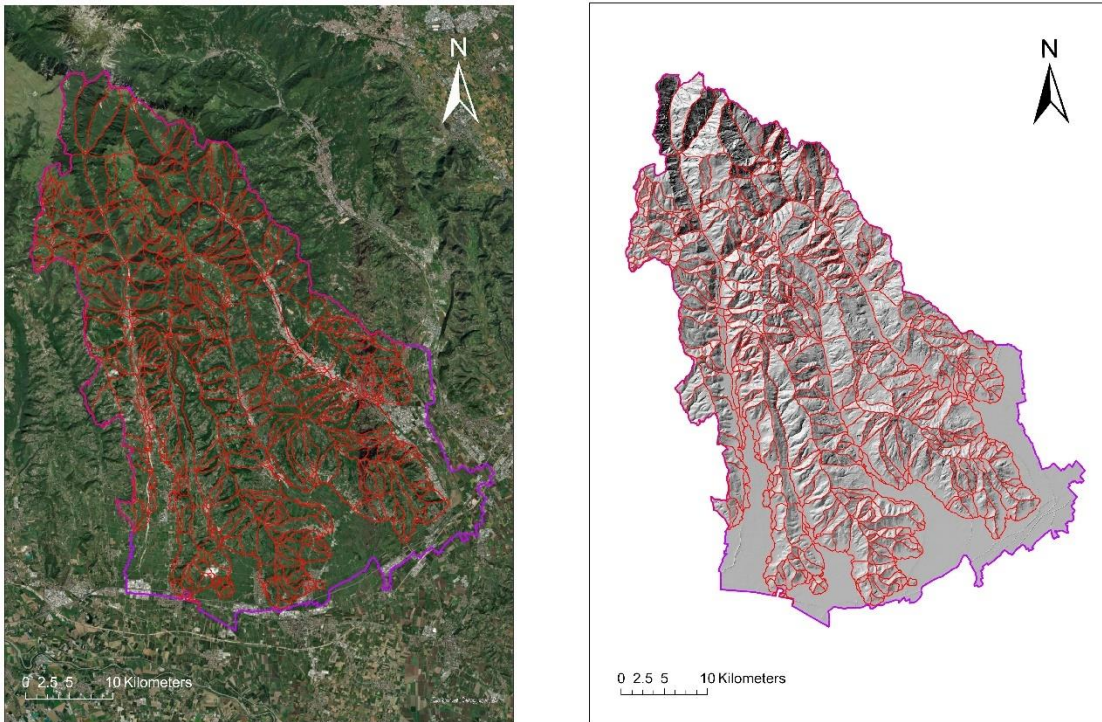


Figura 8: Sulla sinistra l'immagine satellitare della distribuzione delle Slope Units dell'area studio, ricavate dal portale del gruppo di ricerca geomorfologico dell'Istituto di Ricerca per la Protezione Idrogeologica (IRPI); Sulla destra l'immagine su base hillshade della distribuzione della Slope Units all'interno dell'area studio, ricavate dal portale del gruppo di ricerca geomorfologico dell'Istituto di Ricerca per la Protezione Idrogeologica (IRPI)

- Unità topografiche: sono sottogruppi in cui possono essere divise le SU, generati dall'intersezione tra le curve di livello e i confini dei tubi di flusso ortogonali alle curve di livello (O'Loughlin, 1986).

3.1.3 Variabili predisponenti

La distribuzione delle frane sul territorio è condizionata da una serie di fattori esterni o interni che agiscono sul versante in modo predisponente o innescante.

Sono definite cause innescanti quelle condizioni naturali o artificiali che portano le forze instabilizzanti agenti sul pendio a prevalere su quelle stabilizzanti, causando il fenomeno gravitativo. I fattori scatenanti sono, dunque, il contributo fondamentale nel momento di iniziale cedimento del pendio (Lee et al., 2014). Le piogge e i

terremoti sono due dei trigger più comuni che contribuiscono maggiormente al verificarsi delle frane nel mondo (Bai et al., 2020).

Una delle fasi cruciali nella costruzione di un LSM altamente efficiente è quella di selezionare i fattori predisponenti più rilevanti e influenti per quell'area (Mingyong et al., 2022). Ovvero quelle variabili che rendono un pendio più o meno incline al verificarsi di un fenomeno franoso.

Si possono osservare numerosi fattori predisponenti nel mondo de fenomeni franosi, e sono molteplici anche gli studi che li hanno analizzati e comparati; per questo motivo è fondamentale comprendere come le singole variabili influenzino i processi gravitativi del territorio oggetto di studio, in modo da selezionare quelle da integrare nell'indagine.

Sulla base delle condizioni ambientali dell'area e della letteratura scientifica, come Catani et al. (2013) e Reinchenbach et al. (2018), sono stati individuati 17 parametri predisponenti per questa tesi di laurea, divisi in variabili continue e categoriche.

Le variabili continue sono un tipo di variabile numerica o quantitativa che può assumere una distribuzione continua di valori all'interno di un intervallo specifico. Possono essere descritte come variabili che possono assumere tutti i valori compresi in un intervallo reale, il che significa che possono assumere un'infinità non numerabile di valori diversi. Quelle utilizzate in questo caso studio sono: la quota, la pendenza del versante (slope angle), l'orientazione del versante (aspect), la curvatura totale, la curvatura di profilo e planare, la distribuzione delle piogge, la densità di drenaggio, la distanza dai fiumi, il Topographic Wetness Index (TWI), il Topographic Position Index (TPI), il Normal Difference Vegetation Index (NDVI), la distanza dalle strade e il Topographic Roughness Index (TRI).

Le variabili categoriche sono un tipo di variabile utilizzato in statistica e data science per rappresentare dati qualitativi o nominali. Queste variabili possono essere definite come una classe o una categoria di dati che non possono essere quantificati in modo continuo, ma solo in modo discreto (le variabili discrete possono assumere un insieme finito o numerabile di valori). Quelle utilizzate in questo caso studio sono: la litologia, l'uso del suolo (land use) e il Material Strenght Ranking (MSR).

Tutti i layer descritti in seguito sono stati acquisiti dal Geoportale della Regione del Veneto. Il modello digitale di elevazione (DEM) con risoluzione spaziale 25 m,

sempre scaricabile gratuitamente dal geoportale, è stato utilizzato per calcolare le variabili topografiche, attraverso gli specifici strumenti del Toolbox di ArcGIS Pro. L'unica eccezione riguarda il fattore piogge, fornito dall'Agenzia Regionale per la Prevenzione e Protezione Ambientale del Veneto (ARPAV).

In seguito, tutte le mappe dei fattori di condizionamento sono state ricampionate con una risoluzione di 25 m per pixel.

I fattori morfometrici considerati sono:

Elevazione: in letteratura, i dati relativi all'elevazione sono comunemente utilizzati per la valutazione della suscettibilità e per la costruzione della LSM (Reichenbach et al., 2018). Questo fattore riflette la variazione di altezza presenti all'interno dell'area studio; in questo caso si passa da un minimo di circa 21 m ad un massimo di 1973 m, come si può osservare in Figura 9.

L'elevazione ha un impatto significativo sulle caratteristiche topografiche, che contribuiscono alle differenze spaziali di molti processi geomorfologici, così come alla distribuzione della vegetazione. L'altitudine influenza anche in gran parte il clima, compresa la quantità, l'intensità e la distribuzione delle precipitazioni (Meena et al., 2022).

Slope: è la variabilità della pendenza che si riferisce al tasso di variazione della quota del terreno in uno spazio differenziale, basato sul valore di pendenza estratto per ciascun punto del terreno (Mingyong et al., 2022). È un fattore molto importante per determinare la suscettibilità, in quanto un pendio con un'elevata pendenza è maggiormente soggetto al collasso e quindi allo sviluppo di una frana (Meena et al., 2022). In questo caso studio la pendenza minima riscontrata è pari a 0°, mentre quella massima a 77° (Figura 9).

Aspect: indica la direzione geografica verso cui è orientata una pendenza topografica, ed è misurato in senso orario rispetto al Nord (De Malaperdas & Panagiotidis, 2018); in questo sistema un angolo di 0° corrisponde a un'esposizione verso Nord, 90° verso Est, 180° verso Sud e 270° verso Ovest (Figura 10). Inoltre, l'aspect del pendio è un fattore importante per la determinazione dell'umidità topografica, influenzata dall'esposizione solare e dalle precipitazioni, che può fornire indicazioni sulla direzione di una frana. (Chen & Chen, 2021).

Curvatura: è il fattore predisponente che descrive come varia la pendenza di un versante, attraverso la sua geometria concava o convessa.

La curvatura è caratterizzata da un valore negativo che identifica la presenza di una superficie convessa, mentre, al contrario, una curvatura caratterizzata da un valore positivo identifica la presenza di una superficie concava. Il valore uguale a zero indica una superficie pianeggiante (Calligaris, 2016). In questo caso studio, il massimo valore che identifica una superficie convessa è -72, mentre il massimo valore che identifica una superficie concava è +72 (Figura 10).

Curvatura planare: è una tipologia di curvatura calcolata ortogonalmente alla linea di massima pendenza del versante. Regola la convergenza (valore negativo) e la divergenza (valore positivo) del flusso lungo un versante. In questo caso studio il massimo valore che identifica convergenza è -39, mentre il massimo valore che identifica divergenza è +47 (Figura 11).

Curvatura di profilo: è una tipologia di curvatura calcolata lungo la linea di massima pendenza del versante. Influisce sull'accelerazione (valore positivo) o decelerazione (valore negativo) del flusso attraverso la superficie. Per questo motivo è un indicatore dell'erosione e del movimento dei sedimenti (Chen & Chen, 2021). In questo caso studio il massimo valore che identifica decelerazione è -47, mentre il massimo valore che identifica l'accelerazione è +45 (Figura 11).

Topographic Roughness Index (TRI): è l'indice di rugosità topografica. Secondo Riley et al (1999) rappresenta l'eterogeneità topografica del terreno, attraverso il calcolo della differenza di elevazione tra pixel adiacenti in un DEM. Man mano che la pendenza del terreno cambia, il TRI diminuisce, in relazione al movimento dello slope. Il range di valori del parametro in questo caso studio è 0 m – 0.88 m (Figura 12).

Topographic Position Index (TPI): l'indice di posizione topografica misura la differenza tra l'elevazione di un punto centrale e quella media dei punti circostanti all'interno di un raggio prestabilito. I valori di TPI positivi indicano che il punto centrale è situato più in alto rispetto alle aree limitrofe, come le creste, mentre valori negativi suggeriscono una posizione inferiore rispetto alla media, come nelle valli (De Reu et al., 2013); i valori di TPI più bassi indicano quindi la presenza di

caratteristiche topografiche più basse rispetto alle zone circostanti, il che può correlarsi a una maggiore probabilità di occorrenza di frane (Meena et al., 2022). Il range di valori per questo studio varia tra -87 e + 126 (Figura 12).

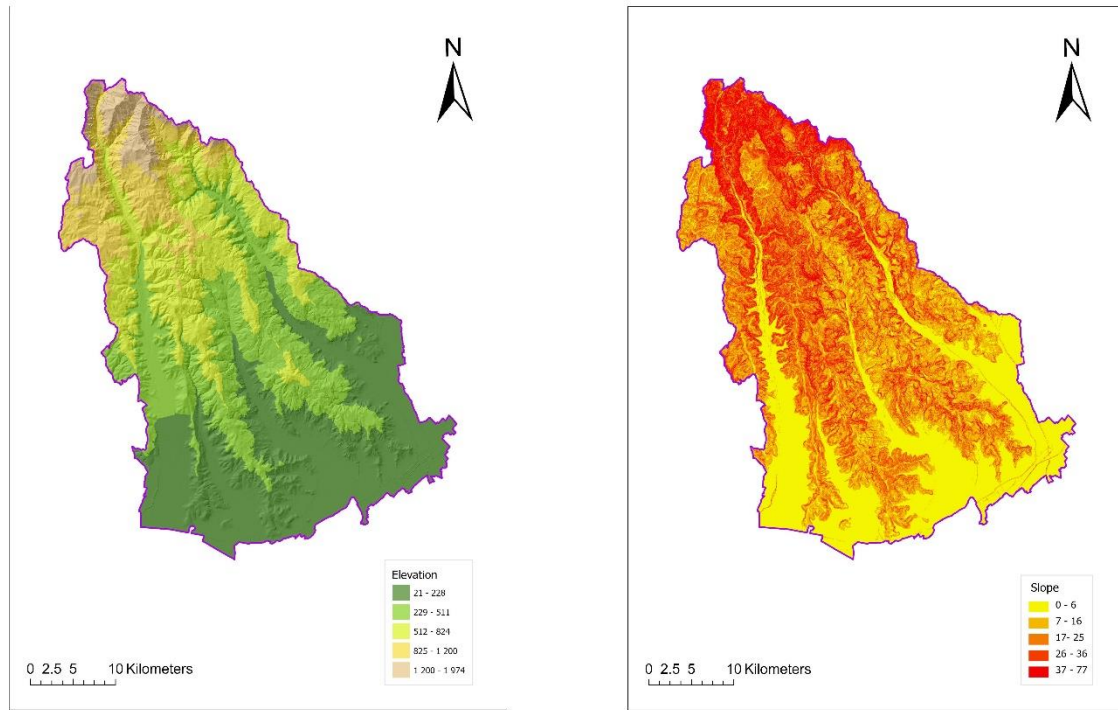


Figura 9: Sulla destra la mappa della variabile predisponente continua quote; è divisa in 5 classi il cui valore minimo 21 m e il valore massimo 1973 m). Sulla destra la mappa della variabile predisponente continua Slope; è divisa in 5 classi il cui valore minimo 0° e il valore massimo 77°.

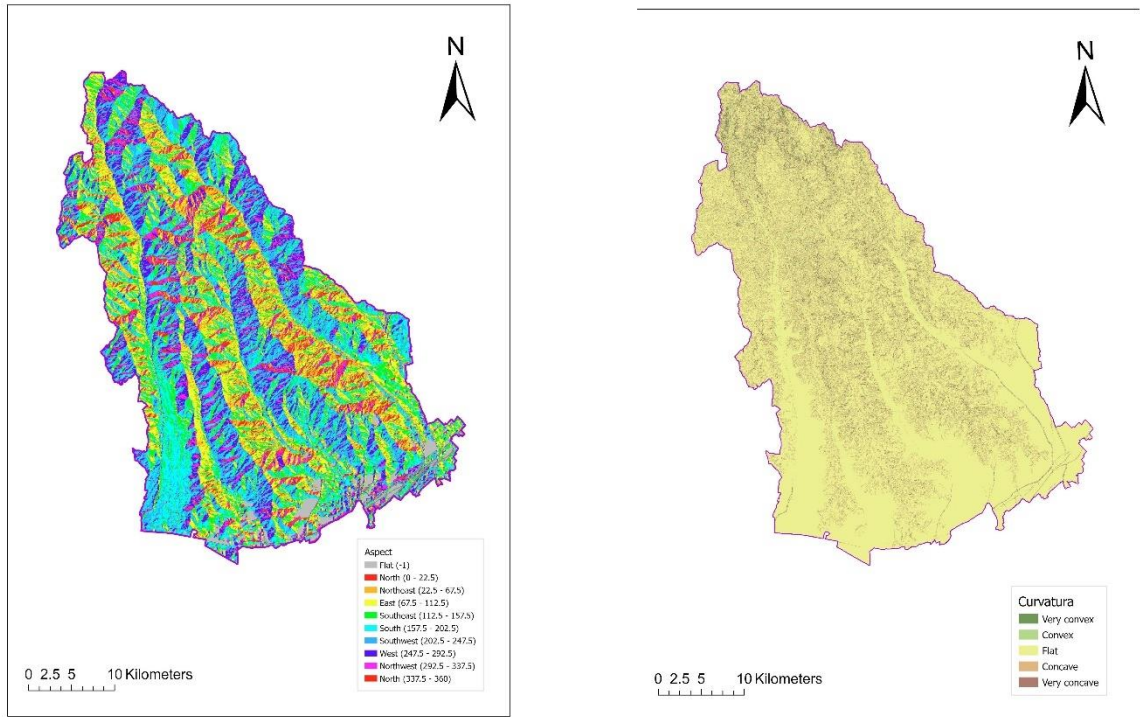


Figura 10: Sulla sinistra la mappa della variabile predisponente continua Aspect; è divisa in 10 classi in cui sono visibili tutti gli angoli di esposizione del pendio. Sulla destra la mappa della variabile predisponente continua Curvatura che è divisa in 5 classi.

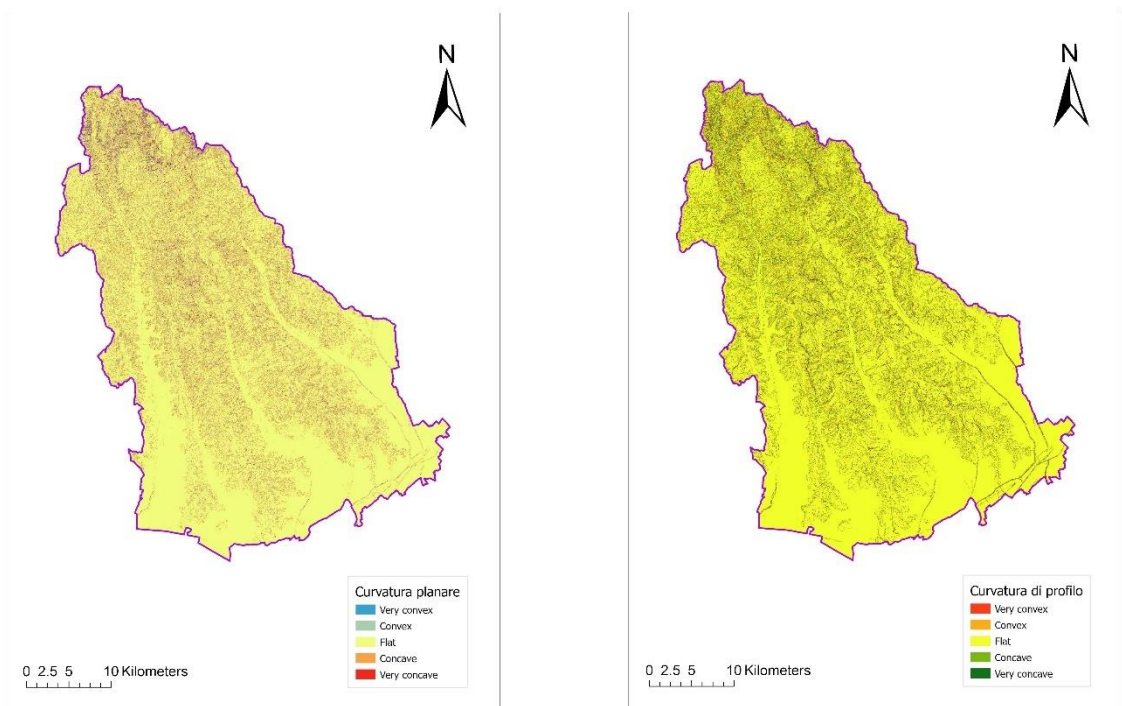


Figura 11: Sulla sinistra la mappa della variabile predisponente continua Curvatura planare; è divisa in 5 classi il cui valore minimo di convergenza è -39, mentre il valore massimo di divergenza è +47. Sulla destra la mappa della variabile predisponente continua Curvatura di profilo; il cui valore minimo è -47, mentre il valore massimo è +45).

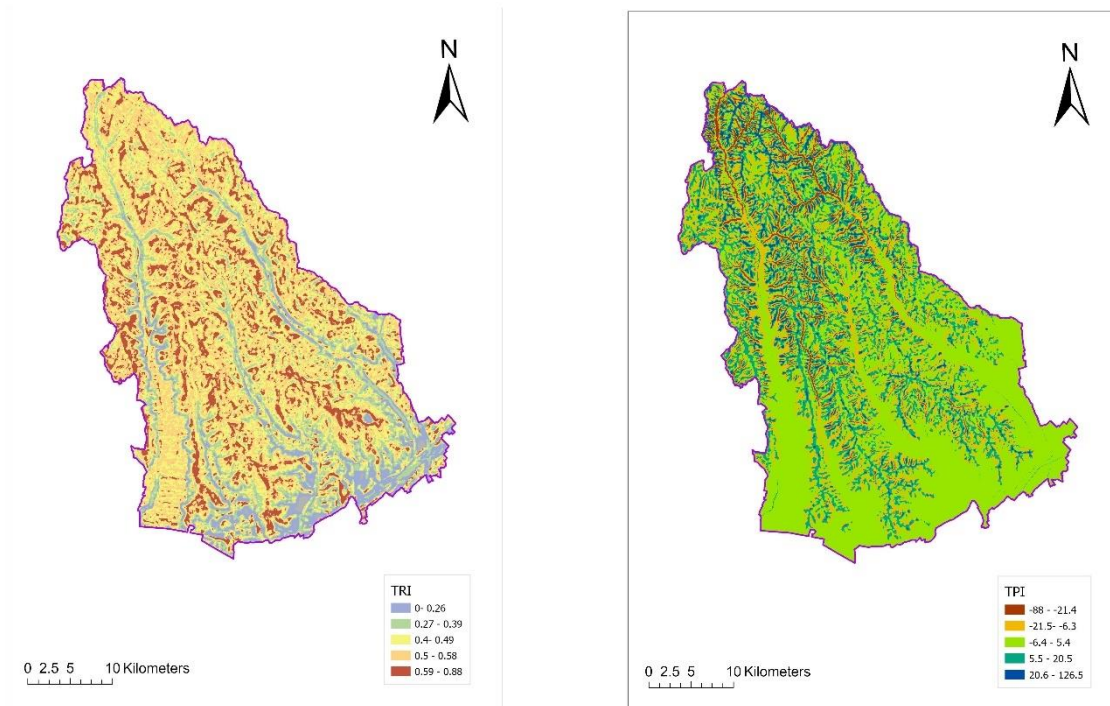


Figura 12: Sulla sinistra la mappa della variabile predisponente continua TRI; è divisa in 5 classi il cui valore minimo è 0 m, mentre il valore massimo è 0.8 m. Sulla destra la mappa della variabile predisponente continua Curvatura di profilo; è divisa in 5 classi il cui valore minimo è -87, mentre il valore massimo di accelerazione è +126.

I fattori idrologici sono:

Distanza dai fiumi (river): è un importante variabile predisponente alle frane per diversi motivi legati all'interazione tra l'erosione fluviale, la pendenza del terreno e il grado di saturazione. Sono tutti fattori che possono compromettere la stabilità del terreno e favorire il verificarsi delle frane (Yalcin et al., 2011) I valori di distanza vanno da un minimo di 0 m, in corrispondenza del corso fluviale, alla massima lontananza di 800 m (Figura 13)

Densità di drenaggio: questa proprietà fondamentale del terreno è stata descritta da (Horton, 1945) come la lunghezza totale dei corsi d'acqua divisa per l'area che occupano. Il drenaggio trasporta l'acqua, il che induce la saturazione del materiale, culminando in frane nelle valli (Shahabi & Hashim, 2015). Nella Figura 13 si osserva la mappa ricavata da questo fattore.

Topographic Wetness Index (TWI): è l'indice di umidità topografica. Permette di stimare l'apporto idrico e la distribuzione dell'umidità nel suolo di una specifica località, integrando il contenuto idrico proveniente dall'area di drenaggio a monte e il

deflusso dell'acqua a valle, approssimato dal gradiente di pendenza, per ciascuna cella di un DEM (Kopecky et al., 2021). Valori elevati di TWI indicano una maggiore possibilità di accumulo idrico e umidità del suolo, correlati a un rischio maggiore di eventi franosi (Meena et al., 2022). In questo lavoro l'intervallo di valori registrato va da un minimo di 0.8 ad un massimo di 29 (Figura 14).

Pioggie: le caratteristiche delle precipitazioni sono influenzate dalle condizioni climatiche e geografiche, con conseguenti variazioni significative sia temporali che spaziali nella quantità e nell'intensità delle piogge stesse. Questo può portare ad un verificarsi di frane non solo su ampie aree, ma anche in specifiche zone più ristrette (Meena et al., 2022). Per questa tesi di laurea sono stati utilizzati i dati delle precipitazioni annuali provenienti da 10 stazioni pluviometriche, prendendo in considerazione solo i periodi per i quali erano disponibili le informazioni per tutte le stazioni. I comuni in cui sono collocati i pluviometri sono: Recoaro Mille, Passo Santa Caterina (Val d'Agno), Cresparolo, San Bortoldo, Bosco Chiesanuova, San Giovanni Ilarione, Chiampo, Illasi e Colognola ai Colli (Figura 14).

Questa analisi consente di suddividere il territorio in fasce con precipitazione uniformi, basata sulla quantità di millimetri di pioggia caduti annualmente. Mediamente le precipitazioni maggiore si sono verificate nell'area montana della sezione, con una tendenza a diminuire spostandosi verso le aree pianeggianti.

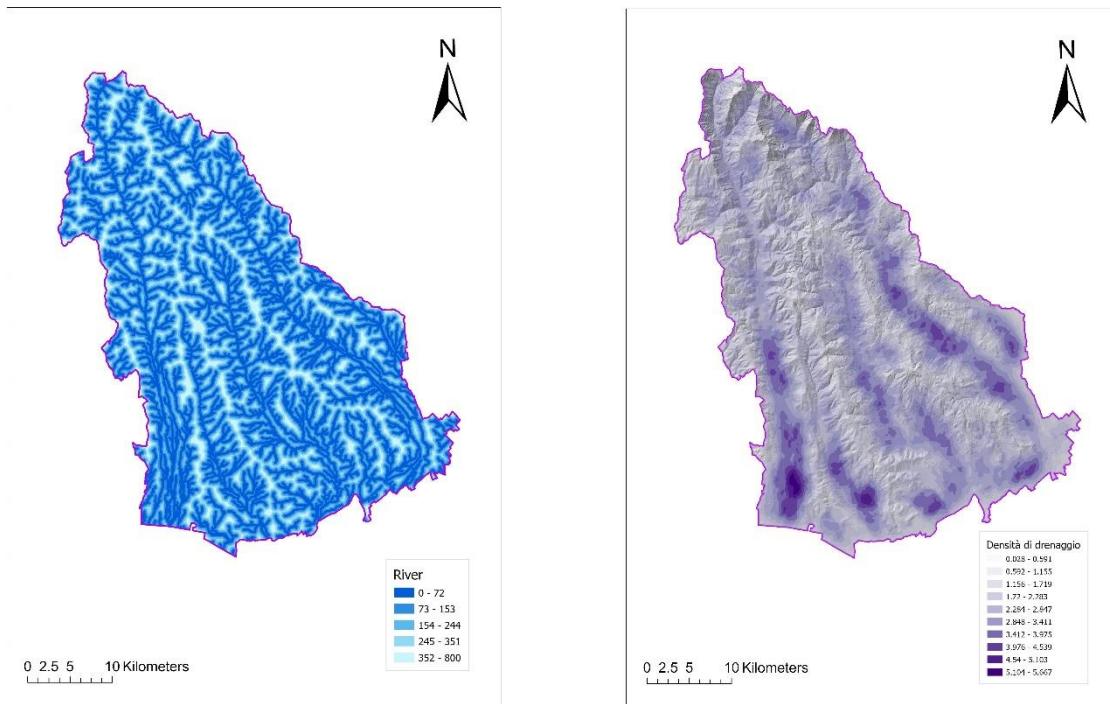


Figura 13: Sulla sinistra la mappa della variabile predisponente continua distanza dai fiumi; è divisa in 5 classi il cui valore minimo è 0 m, mentre il valore massimo è 800 m. Sulla destra la mappa della variabile predisponente continua densità di drenaggio; è divisa in 10 classi il cui valore minimo è 0 m, mentre il valore massimo è 5,6 m.

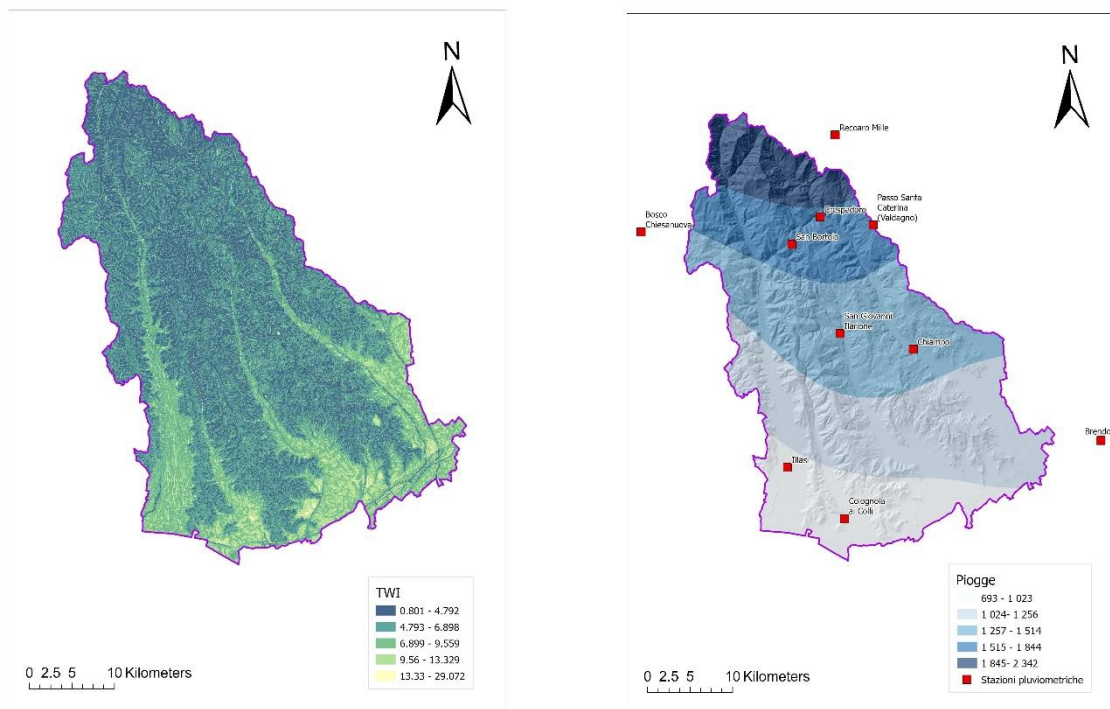


Figura 14: Sulla sinistra la mappa della variabile predisponente continua TWI; è divisa in 5 classi il cui valore minimo è 0,8 m, mentre il valore massimo è 29,1 m. Sulla destra la mappa della variabile predisponente continua pioffre; è divisa in 5 classi il cui valore minimo è 693 mm, mentre il valore massimo è 2342 mm.

I fattori non morfometrici sono:

Land use: questo fattore rappresenta l'insieme degli elementi naturali o artificiali che ricoprono la superficie terrestre. Condiziona il verificarsi delle frane perché ha la capacità di influenzare il funzionamento idrologico dei versanti, la ripartizione delle precipitazioni, le proprietà di infiltrazione e il deflusso, così come la resistenza al taglio del suolo (Meena et al., 2022). È presente un'abbondante letteratura che descrive come le diverse coperture del suolo possono influenzare i fenomeni franosi. Per questa tesi di laurea sono state selezionate sei categorie di suolo (Figura 15), ognuna delle quali è associata a un numero: 1-aree urbanizzate, 2-terreni agricoli, 3-culture permanenti, 4-foreste, 5-vegetazione sparsa, 6-corpi idrici.

Normal Difference Vegetation Index (NDVI): secondo il sito ufficiale dell'USGS, l'NDVI è il parametro utilizzato per quantificare il grado di verde della vegetazione, quindi la presenza di clorofilla; è utile per comprendere la densità della vegetazione e per valutare i cambiamenti della salute nelle piante. È un indice spettrale derivato dalla misurazione delle lunghezze d'onda riflesse dalla vegetazione nelle bande spaziali del rosso e dell'infrarosso.

La copertura vegetale influenza in modo significativo la stabilità di un versante, poiché una bassa copertura vegetale può creare condizioni favorevoli all'insorgere di frane (Viet Long et al., 2023).

Distanza dalle strade (road): i tagli stradali possono ridurre il carico alla base del pendio, creando condizioni favorevoli all'innescò di frane (Yalcin et al., 2011).

Litologia: questo fattore gioca un ruolo fondamentale nello studio delle frane, in quanto le diverse unità litologiche presentano differenti gradi di suscettibilità al cedimento (Yalcin et al., 2011), a causa delle specifiche caratteristiche geomecchaniche dei singoli tipi di roccia.

Per questo studio sono state selezionate 7 unità litologiche: vulcaniti, calcari e calcari dolomitici, calcari argillosi, sabbie, ghiaie e argille. Poiché si tratta di variabili categoriche, a ognuna di esse è stato associato un valore numerico da 1 a 7 (Tabella 2).

Material Strenght Ranking (MSR): è l'indice che valuta il grado di franosità di ogni litologia; si tratta di una classifica realizzata calcolando la densità di frana per ogni classe litologica. Considerando che i cedimenti si verificano su pendii con bassa resistenza al taglio (Segoni et al., 2020), i terreni più suscettibili ai fenomeni franosi sono quelli caratterizzati da scarsa coesione e elevata permeabilità.

L'utilizzo di questo parametro è innovativo nell'ambito dello studio della suscettibilità da frana. Essendo un'analisi basata su dati reali e concreti (densità dei punti in frana) fornisce una misura quantitativa e oggettiva della franosità delle litologie. Questo parametro, inoltre, essendo adattato alle caratteristiche locali, migliora la precisione e l'affidabilità del LSM, soprattutto se combinato con un modello avanzato come Random Forest.

In seguito (Tabella 2) è riportata la classificazione delle litologie, dalla più franosa alla meno franosa; come si diceva in precedenza, ad ogni litologia è associato un valore numerico che la identifichi.

Classifica	Litologia	Categoria
1	Vulcaniti	1
2	Calcari e Calcari Dolomitici	2
3	Depositi alluvionali misti	6
4	Calcari argillosi	3
5	Ghiaie	5
6	Argille	7
7	Sabbie	4

Tabella 2: Classificazione delle litologie più franose presenti nell'area studio, con i relativi valori categorici assegnati.

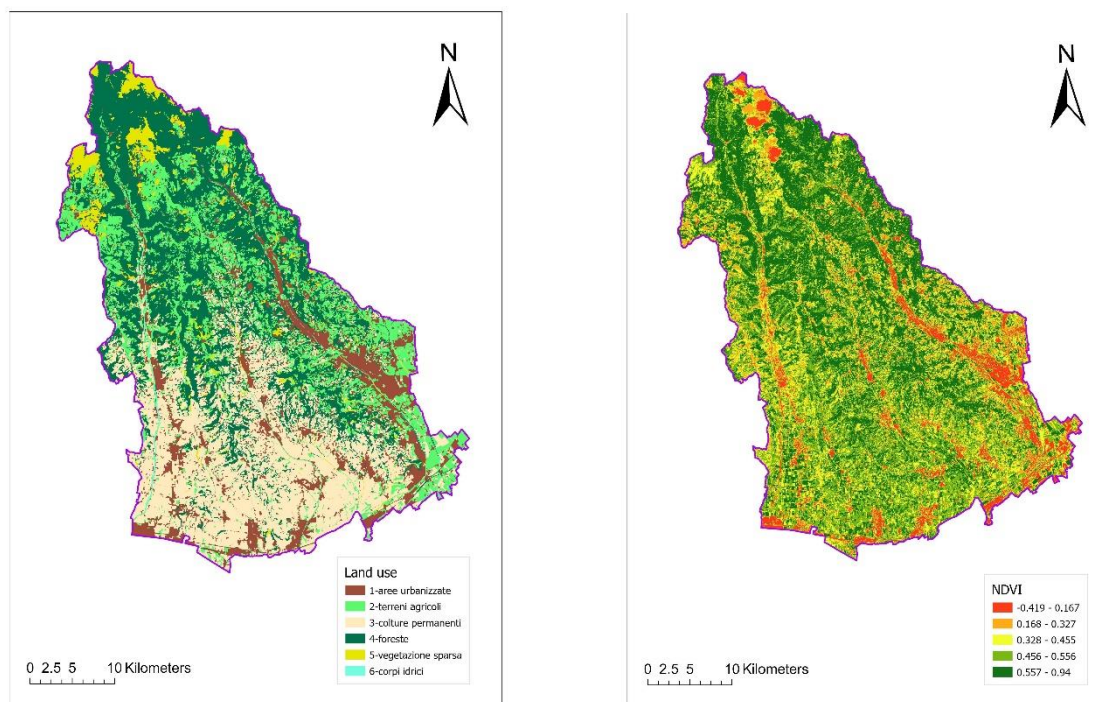


Figura 15: Sulla sinistra la mappa della variabile predisponente categorica land use; è divisa in 6 classi rappresentate da un valore categorico da 1 a 6. Sulla destra la mappa della variabile predisponente continua NDVI

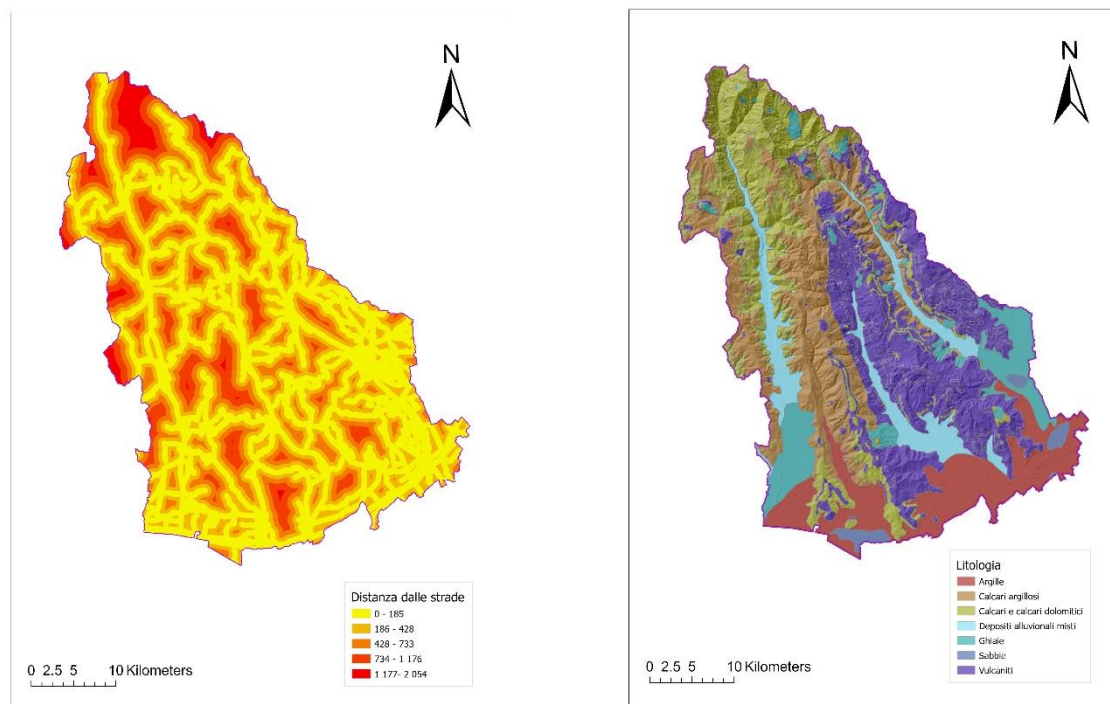


Figura 16: Sulla sinistra la mappa della variabile predisponente continua distanza dalle strade; è divisa in 5 classi il cui valore minimo è 0 m, mentre il valore massimo è 2054 m. Sulla destra la mappa della variabile predisponente categorica litologia; è divisa in 7 classi rappresentate da un valore categorico da 1 a 7.

3.2 Metodi

3.2.1 Modelli per la valutazione della suscettibilità da frana

Non esiste un metodo universalmente riconosciuto per il calcolo della suscettibilità e per la produzione della LSM. Nel corso dei decenni sono state proposte diverse tecniche per il raggiungimento di questo scopo, ma a causa dell'estrema variabilità dei fenomeni franosi, della diversità dei fattori predisponenti e innescanti, e a causa dei numerosi possibili effetti, non è stato ancora possibile individuare un metodo universalmente riconosciuto come affidabile. Per questo motivo ogni singolo caso studio va valutato e analizzato attentamente, al fine di selezionare il metodo migliore per il contesto.

Tutti i criteri proposti, però, si fondano su dei principi condivisi, definiti da Reichdnbach et al. (2018): tutti gli eventi lasciano dei segni sul terreno, che possono essere identificati, mappati e catalogati, pertanto le frane del presente e del passato

sono la chiave per comprendere e prevedere le frane del futuro.

Un altro concetto comune è l'esistenza di leggi fisiche che controllano il verificarsi dei fenomeni gravitativi, quindi esistono delle cause specifiche per ogni evento franoso, che influenzano direttamente o indirettamente il cedimento del pendio; l'ultima situazione condivisa è il prodotto finale dell'analisi di suscettibilità, infatti, il risultato di tutti i metodi è una mappa che suddivide il terreno in aree, sulla base della probabilità di accadimento di una frana in quella zona. Ad ogni settore è associato un colore e un valore di suscettibilità, più o meno alto, in relazione al livello di pericolo. Esiste una letteratura specialistica che si occupa di catalogare e classificare i metodi, derivati da differenti approcci, che sono stati utilizzati nel tempo, tra cui i lavori di (Shano et al., 2020) e di (Reichenbach et al, 2018). Le tecniche possono essere suddivise in qualitative e quantitative: le prime si fondano su osservazioni e valutazioni soggettive; mentre le seconde sono basate su un'interpretazione realistica dei dati, che produce una stima numerica della suscettibilità.

I metodi qualitativi principalmente utilizzati per la produzione di una LSM sono:

Analisi geomorfologica: consiste nella mappatura dell'area studio attraverso telerilevamento oppure studio diretto delle frane in campagna. I risultati di questa analisi sono un inventario dei fenomeni gravitativi presenti, a cui sono associate le rispettive caratteristiche geomorfologiche, quindi il singolo tipo di movimento, l'estensione, il volume di materiale coinvolto, i fattori predisponenti e innescanti, che permettono di studiare la meccanica della frana. Per lo sviluppo della LSM, vengono utilizzati generalmente fattori come la geologia, la geomorfologia, l'uso del suolo e la presenza o assenza di frane.

Essendo un lavoro di tipo qualitativo, il buon livello del prodotto finale dipende dalle capacità dei tecnici coinvolti e dall'estensione e dalla complessità dell'area investigata.

Approccio euristico: anche questa è una tecnica basata sulla soggettività dell'esperto, che individua dei fattori influenti sulla stabilità del pendio, acquisiti in campagna o con il telerilevamento. I fattori vengono suddivisi in classi a cui viene associato un peso, quindi un grado di incidenza sul verificarsi del dissesto. La mappa di suscettibilità totale si ottiene sommando le mappe di ogni fattore.

I metodi quantitativi più utilizzati in letteratura sono:

Approccio deterministico: si fonda sul calcolo del fattore di sicurezza (FS) per tutti i pendii dell'area studio. Il fattore di sicurezza è il rapporto tra le forze stabilizzanti, quindi tutte le forze che resistono allo scivolamento del versante, e le forze destabilizzanti, ovvero quelle che contribuiscono al movimento del versante. Si tratta perciò di valutare la predisposizione alla frana sulla base delle proprietà fisiche e meccaniche che regolano l'equilibrio del versante.

Questo tipo di approccio non è adatta per un'analisi su larga scala, in quanto è necessario utilizzare parametri di input estremamente precisi e dettagliati; pertanto, l'utilizzo è limitato a frane singole o a piccole aree studio (Tofani et al., 2017).

Approccio statistico: si basa sulla correlazione tra i parametri di input, ossia i fattori predisponenti, e i dissesti già verificati su un territorio, localizzati e catalogati in un inventario ben definito, che può essere l'IFFI oppure un inventario costruito ad hoc per lo studio di un'area, anche se questo metodo è generalmente utilizzato per l'analisi di zone molto ampie.

Questo approccio è suddiviso in due sottogruppi principali:

- L'approccio bivariato utilizza il concetto di rapporto di densità per sviluppare un modello di suscettibilità da frana. In questo metodo l'influenza dei fattori predisponenti sul verificarsi dei fenomeni franosi viene valutata singolarmente. Questo significa che per ciascun parametro vengono create delle classi, basate sui valori assunti da ogni variabile all'interno dell'area studio. In seguito, per tutte le classi, si osservano quante frane si sono verificate o non si sono verificate, in modo da calcolare il rapporto di densità. Dopo aver ottenuto questi valori, per tutti gli elementi e le classi, che sono stati analizzati separatamente, si possono ricombinare per ottenere la LSM complessiva.

Secondo Thiery et al. (2007) è un metodo robusto e flessibile, ma presenta delle limitazioni, tra cui la perdita di qualità e precisione dei dati con la semplificazione eccessiva dei dati tematici di ingresso, e la perdita di sensibilità dei dati nell'analisi forzata dei singoli fattori causali.

Esistono vari metodi statistici che utilizzano l'approccio bivariato, tra cui il

modello dell'analisi di frequenza, il modello del peso dell'evidenza, il modello di sovrapposizione ponderata, il modello del valore informativo e la logica Fuzzy.

- L'approccio multivariato tratta simultaneamente tutti i fattori predisponenti di instabilità. Per ogni parametro viene calcolato un peso relativo, in relazione all'influenza che ha sull'instabilità del pendio, e attraverso la loro interazione è possibile calcolare la mappa di suscettibilità da frana.

Anche per questo approccio esistono diversi metodi statistici, tra cui il modello della regressione logistica, quello della regressione multipla e il modello dell'analisi discriminante.

L'intelligenza artificiale (AI): è una tecnologia estremamente avanzata che ha l'obiettivo di automatizzare compiti tipicamente svolti dagli esseri umani. Nel corso dei decenni, è stata utilizzata in diversi ambiti, tra cui la gestione della natura e della società. È uno strumento potente, quindi è perfetto per comprendere e prevedere fenomeni complessi come le frane; permette di elaborare grandi quantità di dati, grazie alla sua capacità di identificare e analizzare le relazioni non lineari (Baghbani et al., 2022).

Esistono diversi metodi di AI che possono essere utilizzati per l'elaborazione della LSM: il Machine Learning (ML), il Deep Learning (DL) e i modelli ibridi (HMs).

- Il Machine Learning è una tecnica basata sulla creazione di algoritmi che permettono al computer di imparare attraverso l'esperienza (Nasteski, 2017). Viene utilizzata quando non si ha una conoscenza teorica completa del problema. Il ML, infatti, estrae informazioni direttamente dai dati per addestrare gli algoritmi, senza dipendere da un modello o da equazioni. All'aumentare del numero di campioni, migliorano le loro prestazioni. In base al risultato che si vuole ottenere esistono due diversi modi di addestramento: l'apprendimento supervisionato insegna al modello a fare previsioni accurate su dati nuovi e non etichettati, dopo essere stato addestrato utilizzando un set di dati contenente sia i valori di input che gli output associati. Questo metodo svolge sia compiti di regressione, quindi la previsione di un

valore numerico continuo, sia compiti di classificazione, quindi la previsione di una classe o categoria per ogni dato di input. Questa tecnica usata per la regressione o la classificazione di sistemi non lineari multivariati, permette di determinare l'associazione tra i fattori predisponenti e il verificarsi delle frane (Lary et al., 2016) e (Merghadi et al., 2020).

L'apprendimento non supervisionato, invece, permette di fare previsioni accurate senza essere stato addestrato utilizzando un set di dati contenente sia i valori di input che gli output associati. In questo caso i dati di input non possono essere confrontati alla serie di valori corretti e risposte predeterminate, le etichette di classificazione devono essere sviluppate automaticamente durante l'addestramento, attraverso la ricerca di similitudini tra i dati, in modo da categorizzarli e creare dei gruppi (Nasteski, 2017).

Tra gli algoritmi più utilizzati dal ML ci sono la regressione lineare, la regressione logistica, il Support Vector Machine (SVM), il clustering, il Decisional Tree (DT), il eXtreme Gradient Boosting (XGBoost) e il Random Forest (RF).

Sebbene inizialmente Machine Learning e il metodo statistico fossero considerati dei concetti distinti, oggi è possibile osservare come diversi algoritmi utilizzati nel ML abbiano in realtà una un'origine statistica, come il Bootstrap, applicato nell'algoritmo di Random Forest, o la regressione lineare.

- Il Deep Learning è un approccio avanzato di ML che richiede una maggiore quantità di dati (LeCun et al., 2015) e utilizza modelli composti da reti neurali profonde per lavorare su dati molto complessi. Queste reti artificiali sono costituite da layer di nodi (neuroni) collegati tra loro, ognuno con un proprio peso e ruolo all'interno del modello.
- I modelli ibridi sono costituiti dall'integrazione di più modelli di ML, al fine di combinare i vantaggi di ognuno di essi, per migliorare l'accuratezza della previsione dei fenomeni franosi (He et al., 2024).

L'algoritmo di Machine Learning utilizzato per lo sviluppo della mappa di suscettibilità dei Monti Lessini, oggetto di questa tesi di laurea, è il Random Forest (RF) implementato sul software MATLAB. È un algoritmo di classificazione

supervisionato introdotto per la prima volta da Breiman (2001), basato sulla creazione e combinazione di molteplici alberi decisionali profondi per migliorare il modello. L'utilizzo dell'algoritmo è vincolato alla suddivisione del dataset in due gruppi: i dati di training vengono utilizzati per addestrare il modello a individuare le relazioni presenti tra i parametri indipendenti, quindi i fattori predisponenti, e i parametri indipendenti o target; l'obiettivo è rendere il modello capace di fare previsioni quando gli vengono sottoposti nuovi dati. I dati di test, invece, servono per valutare le prestazioni previsionali del modello, quando, appunto incontra nuovi dati. Attraverso la tecnica del bootstrapping (Breiman et al. 2003) vengono selezionati casualmente, con sostituzione, dei campioni di valori dal gruppo di training, che servono per addestrare ogni singolo albero decisionale. I dati scartati durante questa fase vengono definiti da Breiman (2001) Out-Of-Bag (OOB). Possono essere utilizzati per determinare l'importanza delle variabili nell'innescare delle frane attraverso il calcolo dell'errore di classificazione (OOB-error), man mano che gli alberi decisionali vengono aggiunti alla foresta.

L'algoritmo RF permette di gestire simultaneamente sia variabili categoriche che continue, in modo da catturare interazioni complesse tra le variabili e modellare relazioni non lineari (Catani et al., 2013). È più vantaggioso rispetto ad altre tecniche in termini di riduzione dell'overfitting; la costruzione di una foresta di alberi impedisce che il modello si adatti perfettamente ai dati di training, performando in modo scarso quando viene applicato ai dati di test.

3.2.2 Workflow

3.2.2.1 Preparazione del dataset

Il dataset utilizzato dal modello RF, è una tabella degli attributi costituita da 491 righe (corrisponde al numero di SU), escluse le intestazioni. Sono contenute tutte le Slope Unit selezionate all'interno dell'area studio, ciascuna identificata con un object, una categoria (cat), una serie di valori e un'etichetta, corrispondente alla variabile dipendente "frana", che identifica la presenza o assenza di fenomeni gravitativi all'interno della SU. La variabile assume valore 1 se la Slope Units è soggetta a frana e 0 se non lo è.

Le altre colonne sono relative ai parametri statistici dei fattori predisponenti descritti nel paragrafo 3.1.3. Utilizzando il tool di ArcGIS Pro “Statistic as Table”, per ogni variabile continua, in ogni SU, sono stati definiti: valore massimo e minimo, range, media, deviazione standard (STD), somma, mediana e il novantesimo percentile (PCT90) ovvero il valore che indica il limite sotto il quale cade il 90% dei dati. Per le variabili categoriche state calcolare anche la majority (classe maggiormente rappresentata) e la minority (classe con minore frequenza).

Sono state calcolate anche la distribuzione di frequenza delle Slope Units, in frana o non in frana, rispetto alla media o alla majority di alcuni fattori predisponenti: media dello slope, majority del land use e majority della litologia più franosa. La distribuzione di frequenza mostra come i dati di un insieme (SU in frana o non in frana) siano distribuiti tra i vari intervalli di valori (media e majority dei parametri). Per ottenere questo risultato i valori statistici dei parametri predisponenti sono stati organizzati in classi, la cui ampiezza è stata determinata con la regola di Sturges. Le ultime due colonne della tabella, denominate X e Y, sono le coordinate spaziali di tutte le Slope Units, nel sistema WGS 1984 UTM Zone 33N.

3.2.2.2 Analisi preliminare

I dati preparati nella tabella vengono caricati nel codice Random Forest su MATLAB, per essere sottoposti ad una prima fase di analisi preliminare, utile per valutare la stabilità del modello e l’influenza delle singole variabili sull’innesco delle frane.

Solo in questa fase viene aggiunta alla tabella una nuova colonna che rappresenta la variabile random, con valori compresi tra 1 e 0; si tratta di una variabile fittizia, scelta casualmente dal modello per ogni interazione. Serve per valutare se l’algoritmo interpreta bene i dati, distinguendo tra le informazioni rilevanti e il rumore.

In questa fase il database è analizzato attraverso tre differenti configurazioni:

- Prima configurazione: considera tutti i valori della media o della majority dei fattori predisponenti, per un totale di 17 variabili, esclusa la variabile random.

- Seconda configurazione: comprende tutti i valori della STD e della majority dei fattori predisponenti; anche qui il totale è di 17 variabili, sempre escludendo la variabile random.
- Terza configurazione: analizza insieme i valori di STD, media e majority di tutti i fattori predisponenti; il totale è di 31 variabili esclusa la variabile random. Questo assetto considera quindi due parametri statistici per ogni variabile continua (media e STD) e uno solo per le tre variabili categoriche (majority).

Per tutte le configurazioni i dati vengono suddivisi casualmente nei gruppi di training e test; le frane utilizzate per la fase di test sono l'80%, mentre il 20% viene usato per la validazione. Sui campioni di addestramento viene utilizzata la tecnica di selezione bootstrapping o bagging, attraverso il comando "bag_predictor". L'algoritmo, in questo modo calcola gli alberi decisionali e l'OOB. I risultati permettono di elaborare il grafico dell'OOB error e il diagramma dell'importanza delle variabili.

Vengono poi calcolati la matrice di confusione e la curva ROC, con la associata AUC, separatamente per i dati di training e test di ogni configurazione.

La matrice di confusione è lo strumento utilizzato per valutare la qualità dei risultati, attraverso l'identificazione gli errori di classificazione commessi dell'algoritmo.

Appare composta da questi elementi:

- True Positive (TP): gli elementi che sono stati correttamente classificati nella classe positiva. In questo caso si fa riferimento alle SU catalogate come frane, che sono effettivamente in frana.
- False Positive (FP): gli elementi che sono stati erroneamente classificati nella classe positiva. In questo caso si fa riferimento alle SU catalogate come frane, che in realtà non sono in frana.
- True Negative (TN): gli elementi che sono stati correttamente classificati nella classe negativa. In questo caso si fa riferimento alle SU catalogate come non frane, che effettivamente non sono in frana.
- False Negative (FN): gli elementi che sono stati erroneamente classificati nella classe negativa. In questo caso si fa riferimento alle SU catalogate come non frane, quando nella realtà dalla frana.

La curva Receiver Operating Characteristic (ROC) è un altro strumento grafico per valutare le prestazioni del modello, evidenziando come vengono distinte la classe dei veri positivi e quella dei falsi positivi, al variare della soglia decisionale; ovvero il valore limite che separa l'area in cui i valori vengono assegnati alla classe positiva e l'area a cui vengono assegnati i valori negativi.

L'Area Under Curve (AUC) è letteralmente l'area al di sotto della curva ROC, che misura l'efficacia del modello:

- $AUC = 1$ indica un modello perfetto
- $AUC = 0,5$ indica un modello senza nessun potere discriminante

Il modello è stato eseguito 5 volte per ciascuna configurazione, con la costruzione di 2000 alberi per ogni esecuzione.

3.2.2.3 Analisi finale e Partial Dipendence Plots

L'analisi finale viene condotta seguendo una serie di passaggi simili a quelli dell'analisi preliminare, considerando alcune modifiche.

Il dataset inserito nel codice non subisce variazioni e viene analizzato utilizzando le tre configurazioni descritte in precedenza, insieme a una nuova configurazione, definita di "pruning". I parametri statistici utilizzati sono stati selezionati partendo dal grafico dell'importanza delle variabili della terza configurazione. La scelta è stata fatta scartando tutti i dati poco influenti, oppure eliminando quelli con un significato simile o sovrapponibile a parametri di un'importanza maggiore, come nel caso della litologia e del MSR. Pertanto, i parametri mantenuti sono i seguenti 11:

- STD dell'elevazione
- Media dello slope
- Media della curvatura di profilo
- STD della curvatura planare
- media delle piogge
- STD della densità di drenaggio
- Media del Topographic Wetness Index (TWI)
- STD del Topographic Position Index (TPI)

- Majority dell'uso del suolo
- Media del Topographic Roughness Index (TRI)
- Majority del Material Strength Ranking (MSR).

Anche nell'analisi finale il database viene suddiviso nei dati di training e di test, dove i primi vengono ulteriormente campionati casualmente con la tecnica bootstrapping, che genera gli alberi decisionali. Le osservazioni escluse sono chiamate OOB, da cui si può calcolare l'errore dell'Out-Of-Bag e il grafico dell'importanza delle variabili. In seguito, si calcolano la matrice di confusione e la curva ROC, con la associata AUC, separatamente per i dati di training e test di ogni configurazione.

Il modello è stato eseguito 10 volte per ciascuna configurazione, con la costruzione di 500 alberi per ogni esecuzione. Il numero di alberi è stato ridotto rispetto all'analisi preliminare, infatti osservando il grafico dell'OOB error, si nota dove l'errore del modello si stabilizza; quindi, dove possono selezionare gli alberi con la migliore capacità predittiva.

Soltanto per la quarta configurazione dell'analisi finale, che è risultata quella con i risultati migliori, sono stati calcolati i Partial Dependence Plots (PDPs), ovvero una rappresentazione grafica che aiuta a interpretare e comprendere come una variabile predisponente influisca sulle previsioni del modello. Attraverso una riapplicazione del modello, vengono calcolati i valori più importanti relativamente ad ogni variabile. Quindi, ogni PDP, mostra la relazione che esiste tra ogni classe della variabile indipendente e la variabile dipendente (target del modello), quindi mostra come si modifica la previsione in risposta alle variazioni dei fattori predisponenti (Nocentini et al., 2023).

3.2.2.4 Mappa di suscettibilità

A questo punto dell'indagine, si dispone di tutti gli elementi necessari per l'elaborazione delle mappe di suscettibilità, che vengono calcolate per le quattro le configurazioni presentate in precedenza. Ogni schema viene inserito singolarmente nella sezione dedicata all'applicazione del codice ed eseguita 10 volte. Per ogni run si generano la matrice di confusione, la curva ROC e l'AUC, che valutano il rendimento del modello.

L'algoritmo calcola il landslide susceptibility index (LSI) per ogni SU, considerando tutte le configurazioni delle variabili, per ciascuna esecuzione del modello, producendo, quindi, 40 mappe di suscettibilità. Il risultato finale è costituito da 4 mappe, ottenute mediando ciascuna delle 10 simulazioni fatte per ogni configurazione. Si avrà, quindi, una mappa di sintesi per ogni configurazione. Ogni albero decisionale, prodotto dal modello, fornisce la probabilità di appartenenza di una SU alla classe frana o non frana, che viene poi mediata tra tutti gli alberi. Questo indice è un valore numerico che rappresenta la variazione della probabilità di accadimento. I valori variano da:

- LSI = 0 significa che secondo molti alberi la probabilità di accadimento di una frana per la SU è nullo.
- LSI = 1 significa che secondo molti alberi decisionali la probabilità di accadimento di una frana per la SU è massimo.

Lo shape-file della suscettibilità, prodotto da RF su MATLAB, viene importato in ArcGIS Pro, che visualizza la disposizione e il posizionamento delle zone caratterizzate da diversa pericolosità spaziale. Questo software permette di suddividere i valori di LSI in classi basate sulla probabilità di accadimento dei fenomeni; od ognuna delle quali è associato un colore.

4 Risultati

In questo capitolo sono illustrati i risultati ottenuti attraverso le modellazioni descritte nel capitolo precedente. Gli esiti di ciascuna analisi svolta, vengono presentati organizzati in paragrafi che seguono l'ordine utilizzato nel capitolo precedente. Questo approccio permette di raggiungere progressivamente l'esposizione del risultato finale, ovvero l'obiettivo di questa tesi di laurea: l'elaborazione della mappa di suscettibilità da frana dei Monti Lessini.

4.1 Analisi del dataset

Come spiegato in precedenza, sui fattori predisponenti selezionati per l'indagine sono state eseguite delle analisi statistiche preliminari con ArcGIS Pro, poi organizzati in una tabella, che è stata importata in MATLAB e utilizzata dall'algoritmo RF.

Separatamente, utilizzando il software Excel, è stata calcolata la distribuzione di frequenza delle Slope Units, in frana e non in frana, rispetto alla media dello Slope e alla majority del Land Use. Lo scopo è fare una stima iniziale dell'importanza di queste variabili nel verificarsi di un evento franoso.

I valori medi dello Slope sono stati suddivisi in 10 classi, con un'ampiezza di 4 valori consecutivi (es. 0-4, 4-8,..). Il valore medio minimo dello Slope per le SU non in frana è $0,6^\circ$ mentre il massimo è $37,4^\circ$; il valore medio minimo per le SU in frana è $7,6^\circ$, mentre il massimo è $39,3^\circ$. L'istogramma seguente (Figura 17) rappresenta la distribuzione di frequenza delle pendenze delle Slope Units in frana o non in frana.

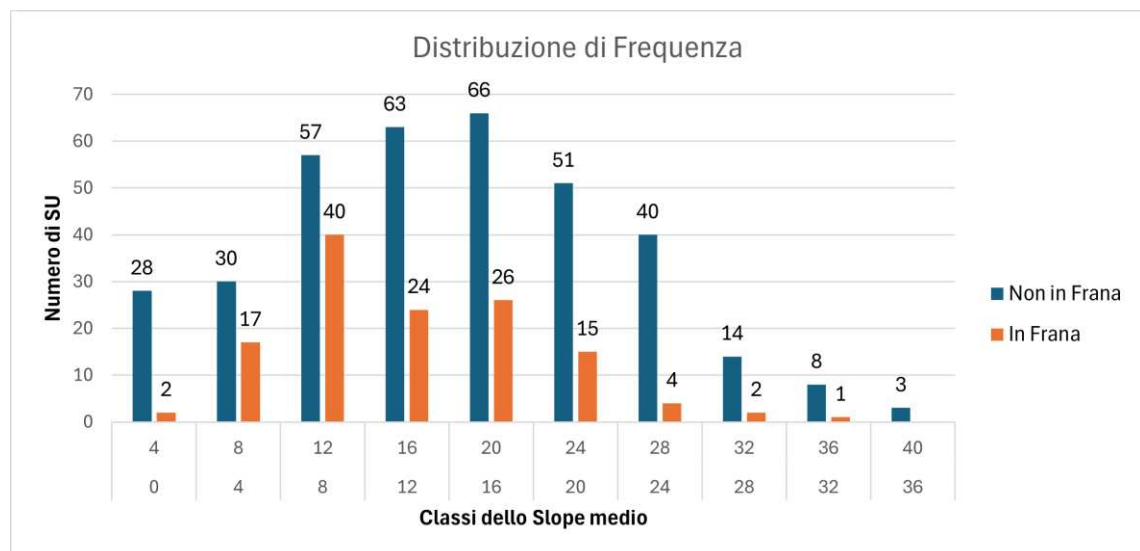


Figura 17: Distribuzione di frequenza delle SU in frana e non in frana, rispetto alle classi dello Slope medio

Questo grafico illustra come la maggior parte delle SU in frana e non in frana abbiano una pendenza media compresa tra 8° e 24° . Corrisponde circa al 65% delle Slope Units non in frana e all'80% delle SU in frana. Questa statistica mostra, in prima analisi, come non ci sia una netta differenza tra le distribuzioni delle pendenze per le SU in frana e le SU non in frana.

La majority del Land Use calcola quale categoria di uso del suolo appare più frequentemente in ogni SU. Per la distribuzione di frequenza delle Slope Units, in frana o non in frana, rispetto alle classi di uso del suolo, le categorie sono stati suddivise in 4 classi, con ampiezza di 1 valore consecutivo:

- La classe 1 rappresenta la categoria di uso del suolo 1, ovvero le aree urbanizzate.
- La classe 2 rappresenta la categoria uso del suolo 2, ovvero i terreni agricoli.
- La classe 3 rappresenta la categoria di uso del suolo 3, ovvero le colture permanenti
- La classe 4 rappresenta la categoria di uso del suolo 4, ovvero le foreste
- La classe 5 rappresenta la categoria di uso del suolo 5, ovvero la vegetazione sparsa

Quindi il valore di majority minimo è 1, mentre il massimo è 5. L'istogramma rappresenta il numero di Slope Units, in frana e non in frana, all'interno di ogni intervallo (Figura 18).

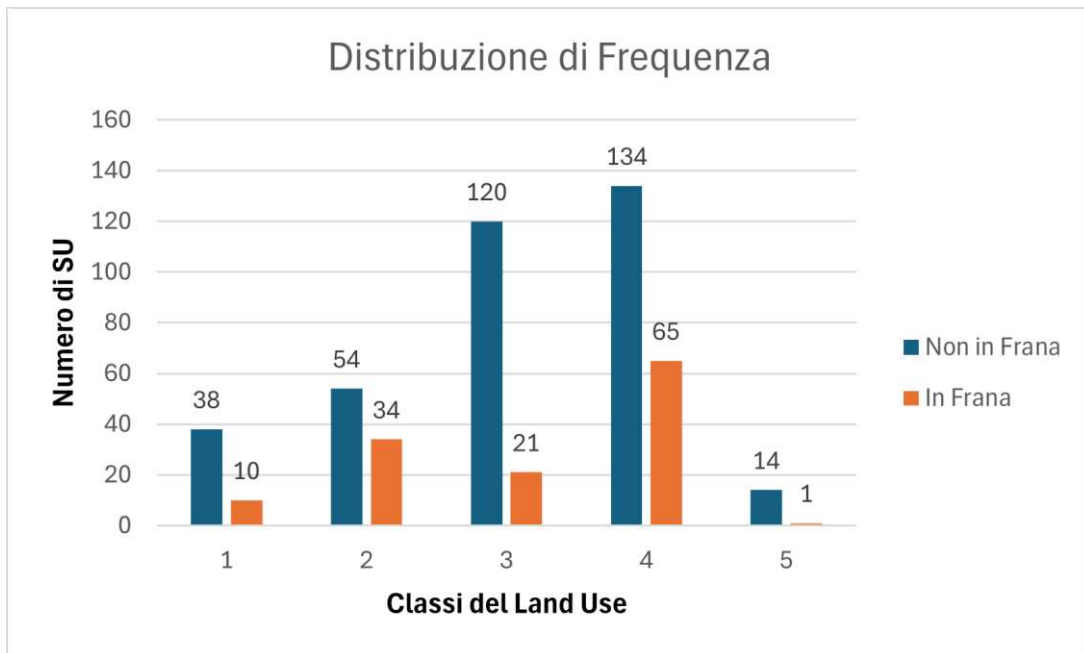


Figura 18: Distribuzione di frequenza delle SU in frana e non in frana, rispetto alle classi della majority del Land Use

Questo grafico mostra come l'uso del suolo più comune per le SU non in frana siano le colture permanenti e le foreste, così come per 86 Slope Units in frana. Sono rispettivamente il 70% e il 65% del totale.

La majority calcola quale litologia appare più frequentemente in ogni SU. Per la distribuzione di frequenza delle Slope Units, in frana o non in frana, rispetto alle classi litologiche, le categorie sono stati suddivise in 7 classi, con ampiezza di 1 valore consecutivo:

- La classe 1 rappresenta la categoria di litologia 1, ovvero le vulcaniti.
- La classe 2 rappresenta la categoria di litologia del suolo 2, ovvero i calcari e i calcari dolomitici.
- La classe 3 rappresenta la categoria di litologia 3, ovvero i calcari argillosi.
- La classe 4 rappresenta la categoria di litologia 4, ovvero le sabbie.
- La classe 5 rappresenta la categoria di litologia 5, ovvero le ghiaie.

- La classe 6 rappresenta la categoria di litologia 6, ovvero i depositi alluvionali misti.
- La classe 7 rappresenta la categoria di litologia 7, ovvero le argille.

L'istogramma seguente il numero di Slope Units, in frana e non in frana, all'interno di ogni intervallo (Figura 19).

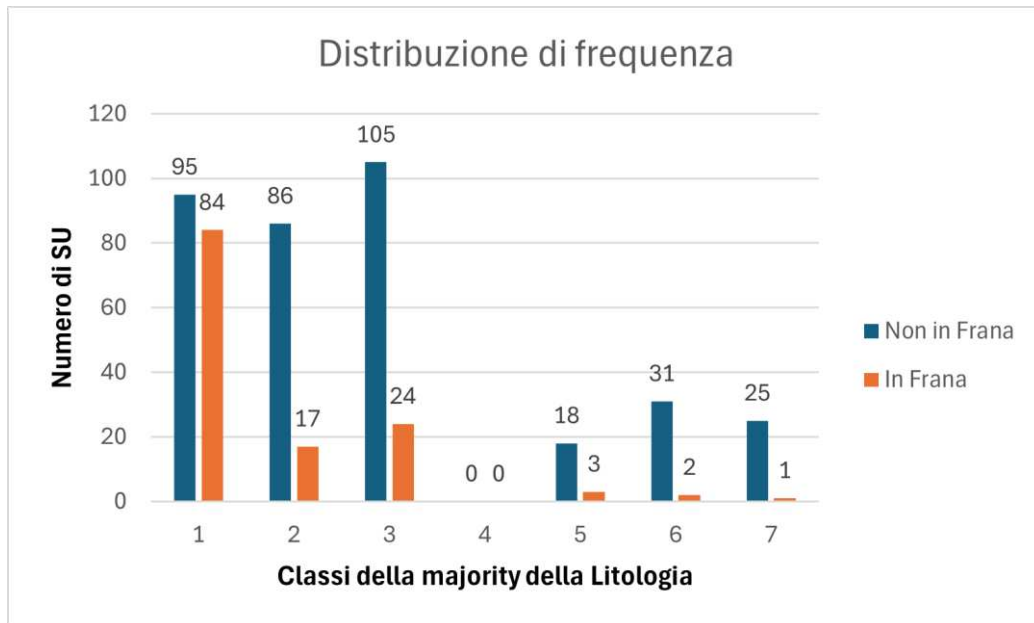


Figura 19: Distribuzione di frequenza delle SU in frana e non in frana, rispetto alle classi della majority della Litologia.

Il grafico mostra che la litologia più frequente nelle SU in frana è quella delle vulcaniti (84), mentre nelle SU non in frana è quella dei calcari argillosi (105). Si osserva che non sono presenti Slope Units nella classe delle sabbie, e che le prime tre classi sono quelle che hanno le litologie presenti più frequentemente nelle SU. Questo dato indica che le tre litologie in questione sono le più abbondanti all'interno dell'area studio, pertanto contengono più SU, a prescindere dalla loro franosità. Questa valutazione non corrisponde con la classifica riportata nel capitolo 3.1.3, in cui le litologie sono ordinate dalla più franosa alla meno franosa in base a quanti punti in frana sono presenti in ogni litologia.

Partendo da questa classificazione, è stato condotto l'ultimo studio della distribuzione di frequenza del MRS. La majority è stata calcolata sulle SU che contengono, in piccola o grande parte, la litologia più franosa: le vulcaniti. In totale

sono 310 le unità su cui è stata calcolata la majority per identificare la categoria litologica più frequente al loro interno. Le classi sono le seguenti:

- La classe 1 contiene la categoria delle vulcaniti
- La classe 2 contiene la categoria di litologia dei calcari e dei calcari dolomitici.
- La classe 3 contiene la categoria di litologia dei calcari argillosi.
- La classe 4 contiene la categoria di litologia delle sabbie.
- La classe 5 contiene la categoria di litologia delle ghiaie.
- La classe 6 dei depositi alluvionali misti.
- La classe 7 delle argille.

L'istogramma in (Figura 20) rappresenta il numero di Slope Units, in frana e non in frana, all'interno di ogni intervallo.

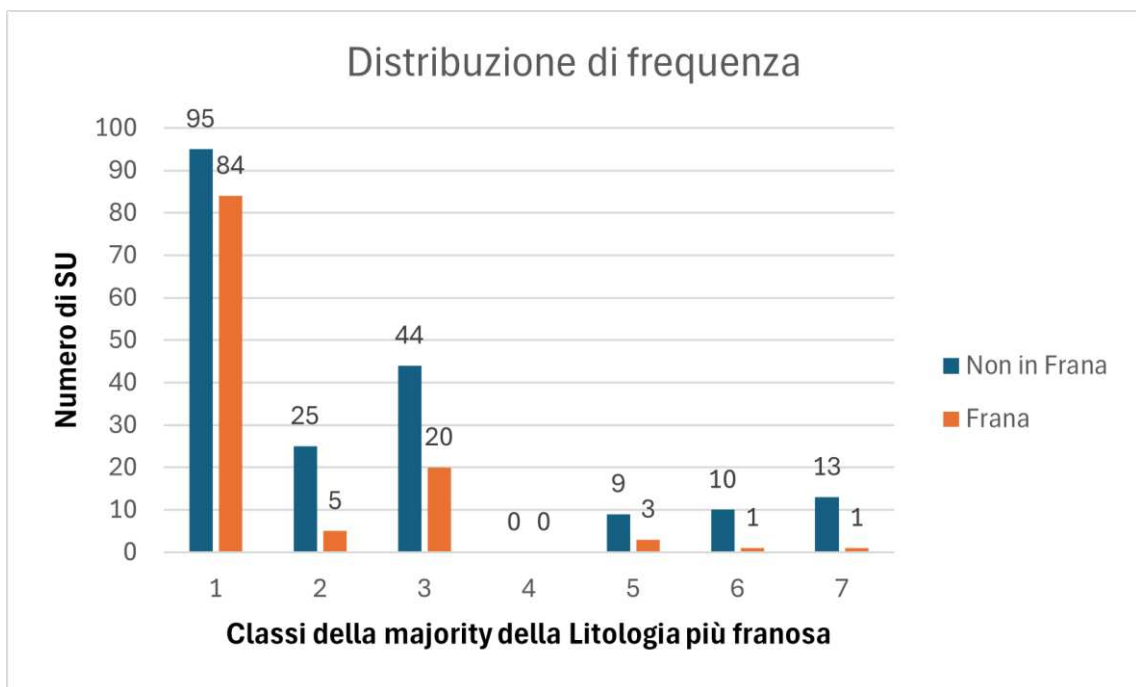


Figura 20: distribuzione di frequenza delle SU in frana e non in frana, rispetto alle classi della majority della Litologia più franosa.

Il grafico mostra che in 95 SU non in frana e in 84 SU in frana la litologia più frequente è quella delle vulcaniti, come si osservava nel caso precedente; così come la totale assenza di SU con litologia più frequente delle sabbie.

4.2 Risultati analisi preliminare

Nei paragrafi successivi vengono riportati ed analizzati i risultati dell'analisi preliminare, per le varie configurazioni dei fattori predisponenti utilizzati. Saranno quindi esaminati: l'errore di Out-Of-Bag, l'istogramma con la stima dell'importanza delle variabili, la matrice di confusione per i dati di training e test, e la curva ROC sempre sia per i dati di training che di test.

Come già descritto in precedenza, durante la fase preliminare il modello è stato applicato 5 volte per ogni configurazione e, per ciascuna esecuzione, ha costruito 2000 alberi decisionali.

La prima fase di elaborazione consiste nella suddivisione del dataset nel gruppo di training e di test, e il successivo utilizzo della tecnica di bootstrapping per costruire gli alberi decisionali. I dati scartati da questo lavoro di bagging, chiamati Out-Of-Bag, non vengono visti durante l'addestramento degli alberi, ma vengono utilizzati per testare la performance dell'albero stesso. L'errore di previsione del modello (OOB error) viene calcolato mediando gli errori previsionali ottenuti per ogni albero, utilizzando gli OOB.

Per ogni esecuzione dell'algoritmo, si genera un grafico dell'OOB error. Sull'asse delle ascisse è riportato il numero degli alberi, che va da 0 a 2000; su quello delle ordinate l'OOB classification error, che varia da 0.46 a 0.26. La linea blu mostra l'andamento dell'errore al crescere degli alberi.

Per ogni configurazione risultano 5 grafici dell'Out-Of-Bag error; nella (Figura 21), (Figura 22) e (Figura 23) si possono osservare quelli relativi al primo run, gli altri sono riportati nell' (Allegato 1), (Allegato 2) e (Allegato 3).

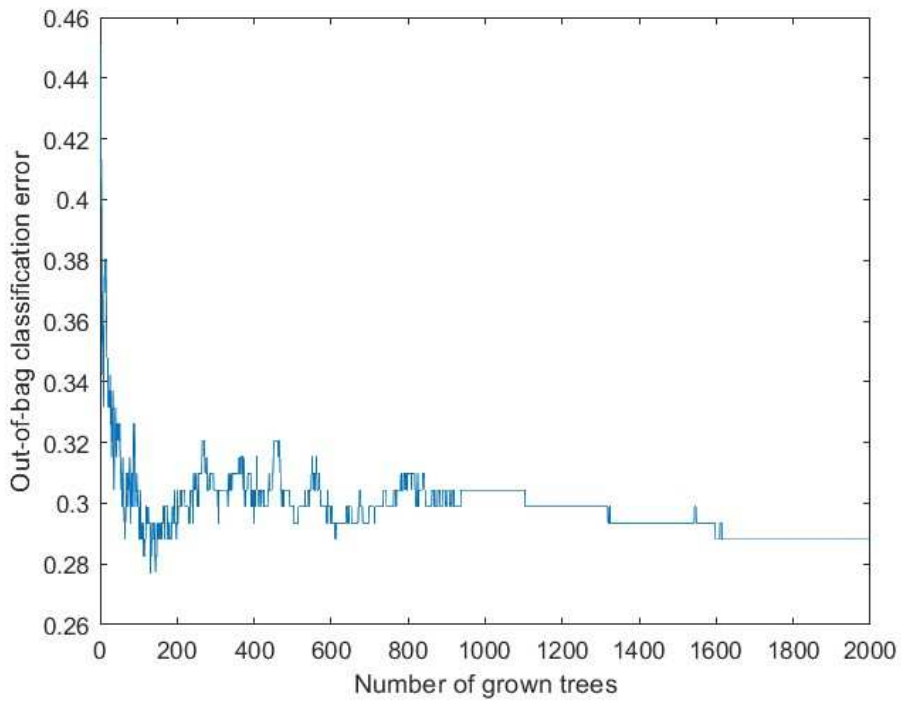


Figura 21: Grafico dell'OOB error del primo run della configurazione con media e majority (prima configurazione).

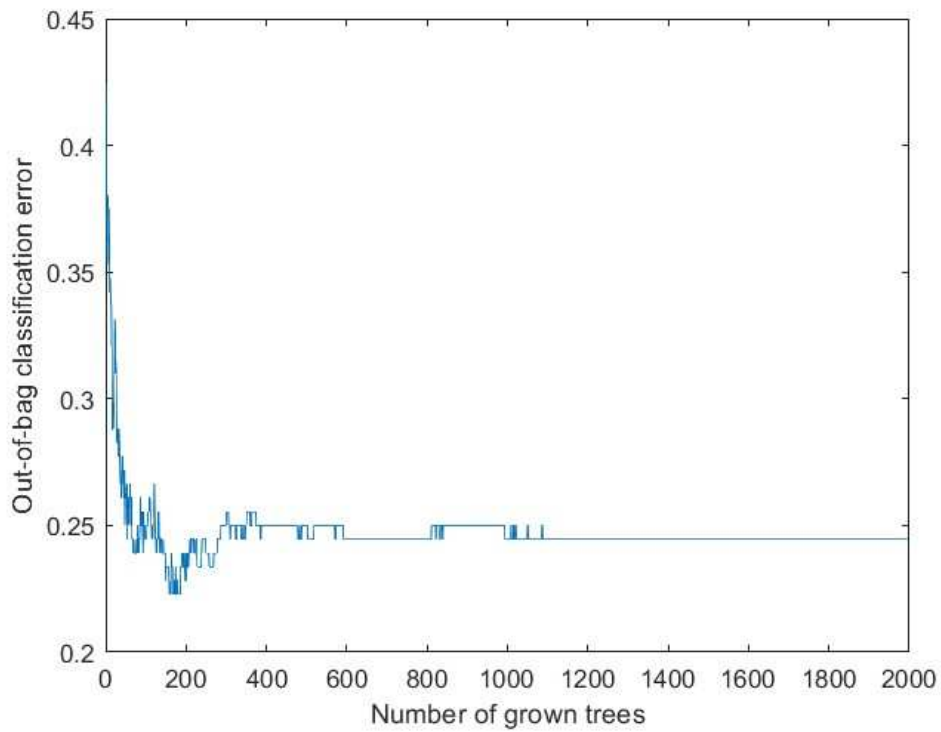


Figura 22: Grafico dell'OOB error del primo run della configurazione con STD e majority (seconda configurazione)

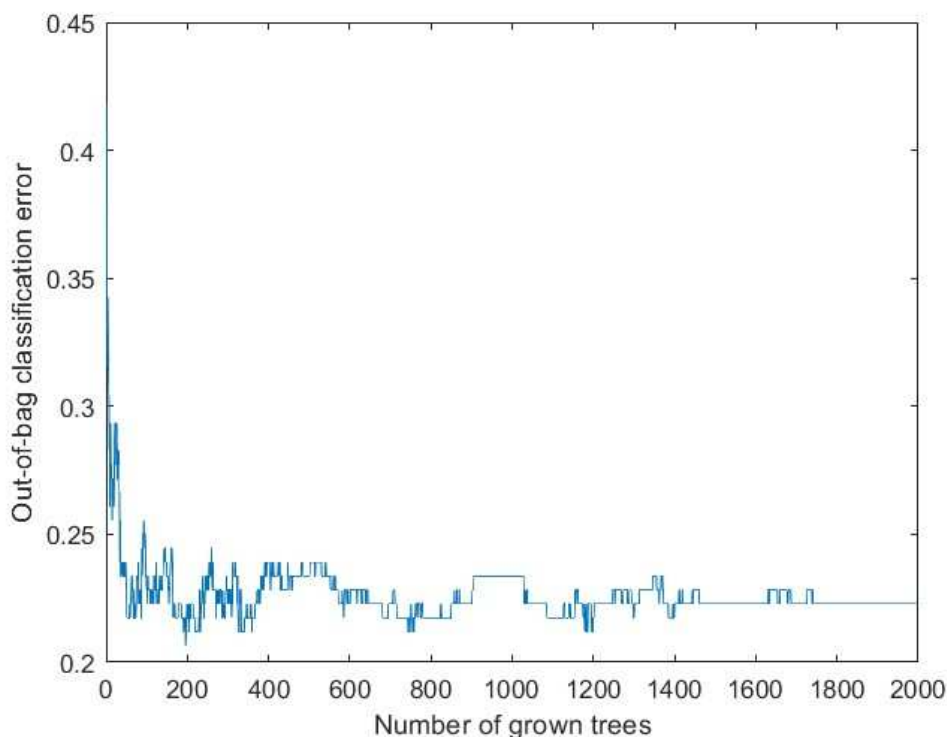


Figura 23: Grafico dell'OOB error del primo run della configurazione con media, STD e majority di tutti i fattori predisponenti (terza configurazione)

Nell'area iniziale dei grafici, dove ci sono ancora pochi alberi, si ha un elevato errore dell'OOB. In questa fase il modello non è ancora in grado di stabilire correttamente le relazioni esistenti tra i dati, quindi non sta ancora facendo delle previsioni accurate. Man mano che aumentano gli alberi si stabilizza anche l'errore, fino a diventare costante, circa in corrispondenza di 500. Questo valore, che è stato ottenuto mediando i risultati di tutti i run delle tre configurazioni, è stato scelto per eseguire l'algoritmo nella fase di analisi finale. Quando l'errore si stabilizza, non ci sono ulteriori possibilità di miglioramento, nel senso che aggiungendo nuovi alberi non si correggerebbero in modo significativo gli errori precedenti, ma si aggiungerebbe solamente una ulteriore, e inutile, complessità al modello.

Il calcolo dell'OOB error è utile anche per stabilire l'importanza delle variabili utilizzate nell'algoritmo. Il modello, per ogni albero decisionale, procede con la perturbazione di una variabile predisponente. Per questa nuova configurazione viene calcolato un OOB error, che confrontato con quello originale, permette di fare la valutazione sull'incidenza della variabile. I valori positivi dell'OOB error indicano

che la variabile incide molto sul modello, quindi, se questi parametri venissero esclusi dall'elaborazione, l'errore aumenterebbe peggiorando le capacità previsionali del modello. Viceversa, i valori negativi dell'OOB error, indicano incidenza negativa delle variabili: se questi parametri venissero esclusi dall'elaborazione, l'errore potrebbe diminuire, migliorando le capacità previsionali del modello.

In seguito sono riportati i tre istogrammi dell'importanza delle variabili (OOB Permutated Predictor Importance), proposti per ogni configurazione (Figura 24), (Figura 25) e (Figura 26). Sono stati ottenuti mediando i risultati dei singoli grafici prodotti per ogni esecuzione del modello.

Sull'ascissa sono riportati i nomi dei predittori, con il prefisso che indica il tipo di analisi statistica che è stata fatta su quella variabile. L'asse delle ordinate, invece, mostra il valore dell'importanza del predittore; valori maggiori indicano un'importanza maggiore sul modello. Si può osservare anche la barra di errore per ogni colonna, che mostra l'intervallo di confidenza dei valori, indicando il massimo e il minimo calcolati.

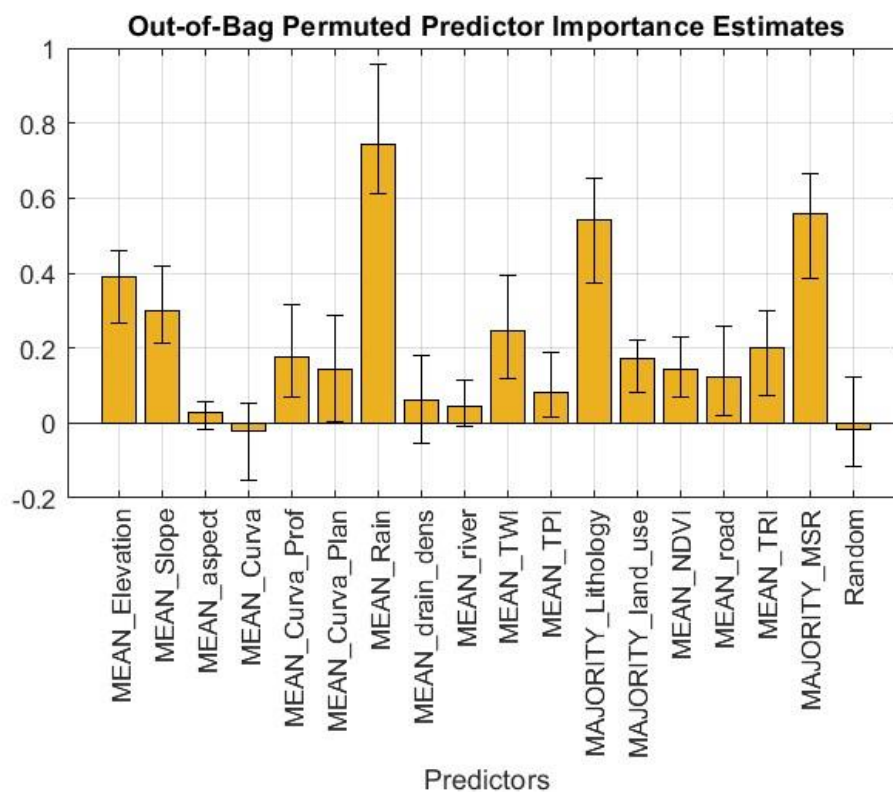


Figura 24: Istogramma dell'importanza delle variabili per la prima configurazione

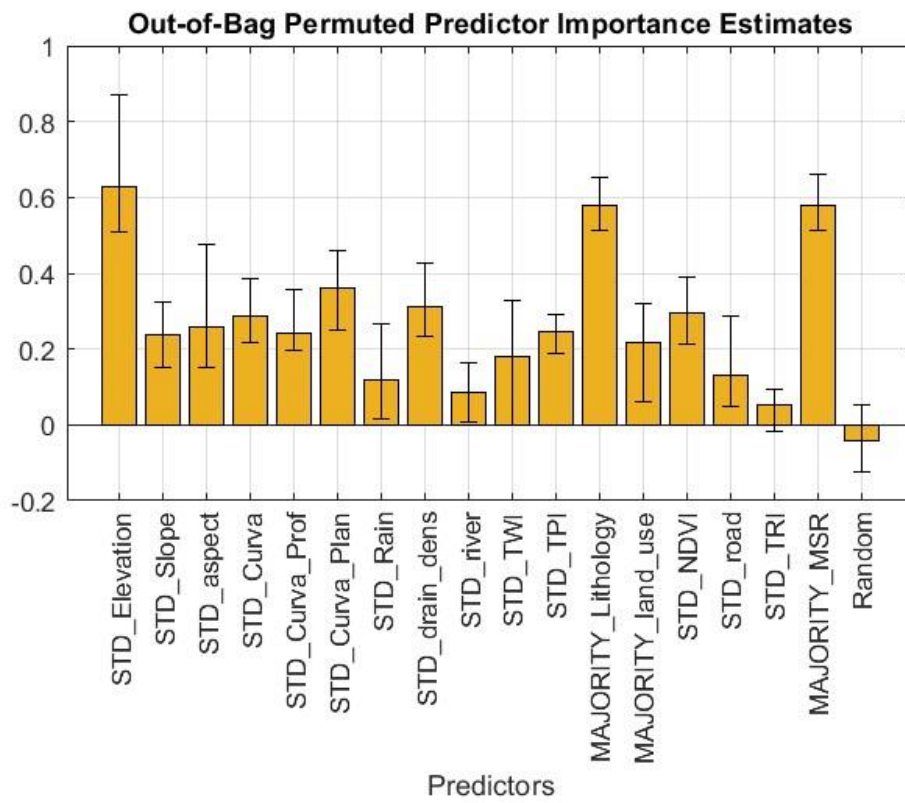


Figura 25: Istogramma dell'importanza delle variabili per la seconda configurazione

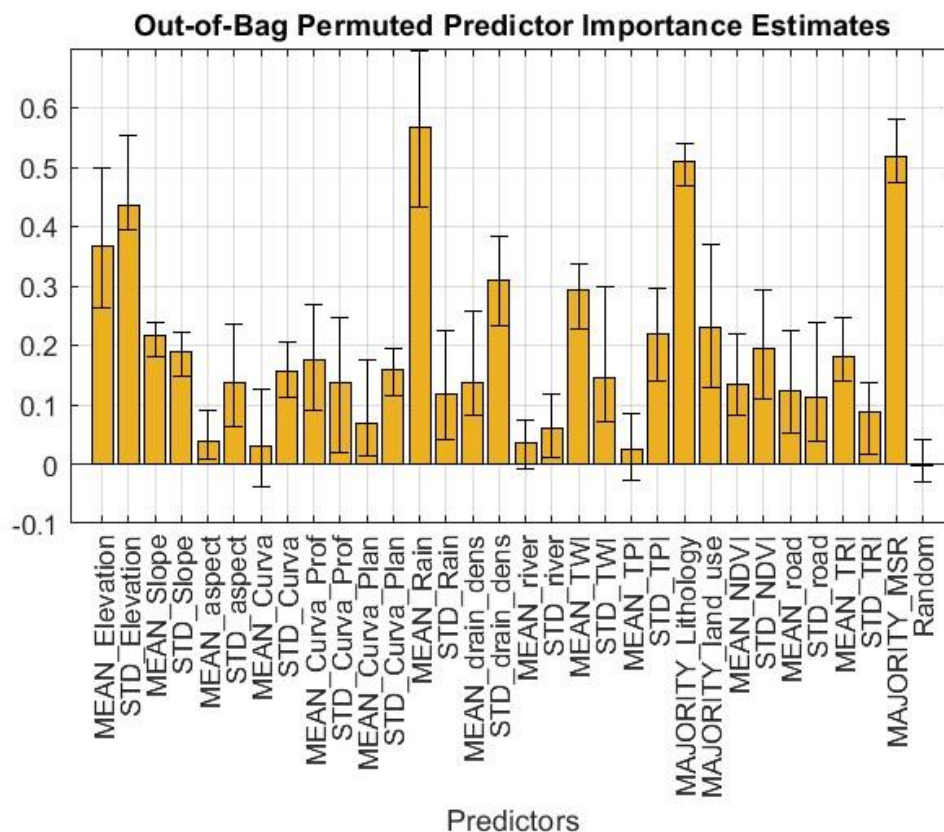


Figura 26: Istogramma dell'importanza delle variabili per la terza configurazione

Come prima cosa si può osservare la presenza della variabile random, che in tutti i grafici mostra un'importanza bassa e trascurabile; infatti, non verrà più considerata per l'analisi ufficiale. La sua scarsa rilevanza viene usata come metro di paragone per l'esclusione delle altre variabili.

Nella Figura 24 sono presenti tre fattori particolarmente incidenti: le piogge, la litologia e l'MSR. I parametri meno significativi sono sicuramente l'aspect, la curvatura, la densità di drenaggio, la distanza dai fiumi e la TPI.

Nella seconda configurazione (Figura 25) questa classifica di rilevanza viene parzialmente mantenuta; infatti, l'elevazione ha un'importanza maggiore rispetto al caso precedente, mentre si mantengono in vetta la litologia e l'MSR. L'incidenza della variabile che diminuisce più significativamente è quella delle piogge. Questo potrebbe essere spiegato in base al significato dei parametri statistici media e deviazione standard. La media è il valore medio delle precipitazioni in un lasso di tempo, quindi descrive l'accumulo; mentre la deviazione standard descrive quanto le

piogge si discostano da un valore medio, relazionandosi ai picchi di precipitazioni che si possono verificare nell'area, come nel caso di un evento estremo.

L'importanza maggiore della MEAN_Rain, rispetto alla STD_Rain, potrebbe indicare che l'innescò di fenomeni franosi, in questa area studio, è determinato dall'accumulo prolungato di acqua.

In generale l'importanza delle variabili che nella prima configurazione risultavano molto basse, aumentano nella seconda, dove la STD mostra la variabilità dei parametri rispetto alla loro media.

Nella terza configurazione (Figura 26) vengono utilizzati la media, la majority e la deviazione standard di tutti i predittori; pertanto, il risultato ottenuto ricalca l'andamento della combinazione delle due configurazioni precedenti.

Il modello produce anche la matrice di confusione dei dati di training e dei dati test per ogni configurazione considerata. Ovvero lo strumento utilizzato per valutare la qualità dei risultati, attraverso l'identificazione gli errori di classificazione commessi dell'algoritmo.

La matrice di confusione si presenta come un rettangolo composto da 4 celle, che contengono i valori di TN, FP, FN e TP, calcolati dal modello. In "ascissa" sono presenti le classi predette, che possono essere pari a 0 (non in frana) o 1 (in frana); mentre in "ordinata" sono presenti le classi reali, anche queste pari a 0 e 1.

Associata alla matrice di confusione si hanno la curva ROC e l'area sottesa alla curva (AUC), anche queste calcolate sia per i dati di training che di test, per tutte le esecuzioni delle tre relazioni. La prima è utilizzata per valutare le prestazioni del modello, evidenziando come vengono distinte la classe dei veri positivi e quella dei falsi positivi, al variare della soglia decisionale; ovvero il valore limite che separa l'area in cui i valori vengono assegnati alla classe positiva e l'area a cui vengono assegnati i valori negativi.

La seconda è l'area sottesa alla curva ROC, a cui viene assegnato un valore che misura l'efficacia del modello.

Sull'ordinata del grafico sono riportati i valori di TP rate, ovvero la percentuale di veri positivi correttamente identificati dal modello; mentre in ascissa sono riportati i valori di FP rate, ovvero la probabilità che il modello identifichi erroneamente un campione negativo (non in frana) come positivo (in frana).

Il risultato della matrice di confusione, costruita valutando il modello sui dati di training, è uguale per tutti i run di tutte le configurazioni, come si può osservare in Figura 27.

- Sono presenti 92 veri negativi. Il modello ha giustamente predetto la stabilità.
- Sono presenti 0 falsi negativi. Il modello non ha predetto stabilità dove era presente instabilità.
- Sono presenti 92 veri positivi. Il modello ha giustamente predetto instabilità.
- Sono presenti 0 falsi positivi. Il modello non ha predetto instabilità dov'era presente stabilità.

Questo risultato indica una buona performance del modello, prevedibile considerando che si tratta della fase di addestramento.

Per verificare effettivamente il rendimento della previsione bisogna esaminare i risultati anche della fase di test.

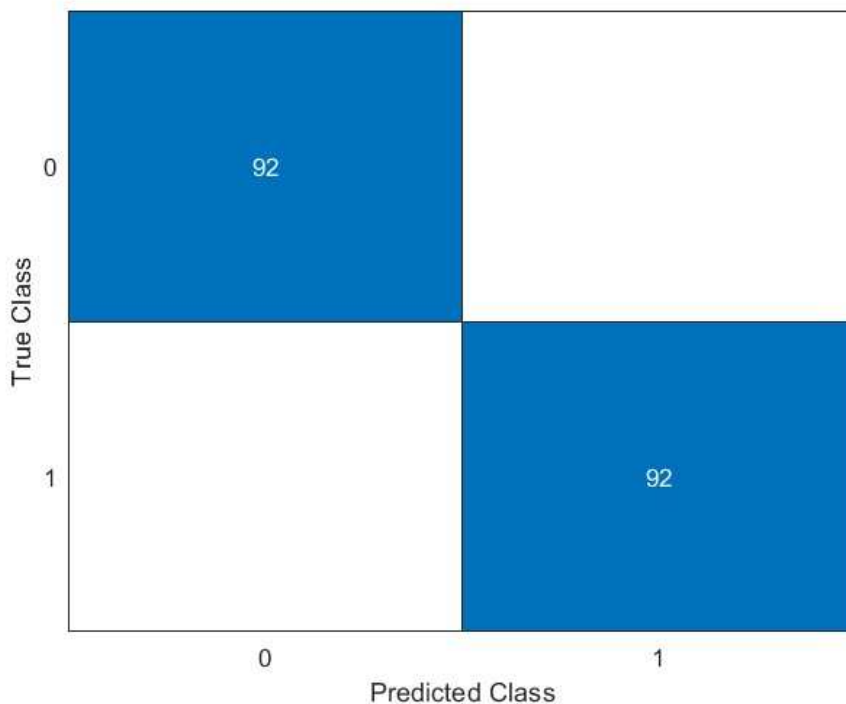


Figura 27: Matrice di confusione della fase di training, uguale per tutte le applicazioni delle tre configurazioni

Anche la curva ROC e il valore di AUC si mantengono costanti per tutte le

applicazioni delle tre configurazioni (Figura 28). Si tratta di un modello perfetto, in cui la curva sale lungo l'asse del TP rate, fino all'angolo superiore sinistro, dove TPR è pari a 1 e FPR è pari a zero. In queste condizioni il valore di AUC è 1, che indica la perfetta capacità del modello di separare le classi positive da quelle negative. Questo risultato indica che il modello è riuscito ad “addestrarsi” ottimamente con i dati disponibili.

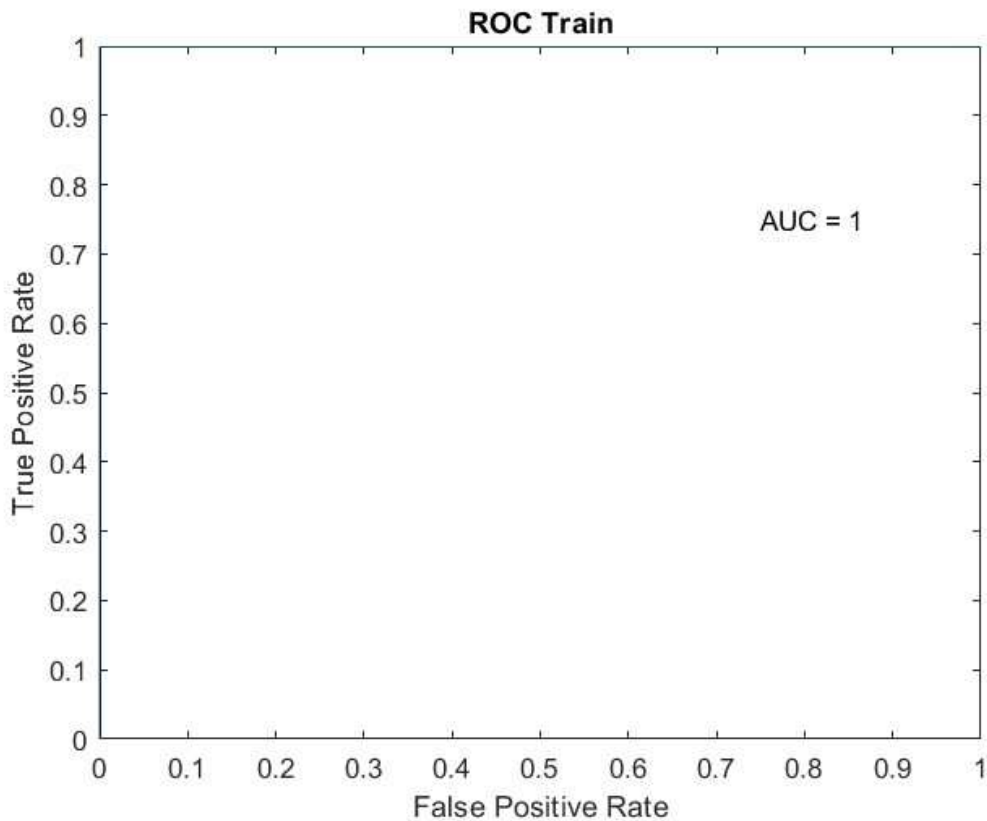


Figura 28: grafico della curva ROC e valore di AUC costante per tutti i run delle tre configurazioni.

I risultati prodotti dalla fase di training e di test per l'analisi preliminare delle varie configurazioni si possono riassumere nelle seguenti tabelle.

Il numero di TN, FN, TP e FP per la fase di test della prima configurazione risulta leggermente diverso per ogni applicazione dell'algoritmo. I valori medi riscontrati sono analizzati in seguito:

- Sono presenti mediamente 29 veri negativi, significa che il modello ha giustamente predetto la stabilità in 29 casi.

- Sono presenti mediamente 10,2 falsi negativi, significa che il modello ha predetto erroneamente stabilità in 10,2 casi, quando nella realtà era presente instabilità.
- Sono presenti mediamente 28.8 veri positivi, significa che il modello ha giustamente predetto l'instabilità in 28 casi.
- Sono presenti mediamente 10 falsi positivi, significa che il modello ha predetto erroneamente instabilità in 10 casi, quando nella realtà era presente stabilità.

Questa suddivisione media ha un'accuratezza circa del 74%, ovvero il modello ha classificato correttamente il 74% dei campioni presenti nel dataset di test. Questo valore non è sufficiente per valutare correttamente la performance del modello; perciò, vanno consideranti anche la curva ROC e il valore di AUC; anche in questo caso si differenziano leggermente per ogni run del codice, bisogna considerare quindi il valore medio di AUC, pari a circa 0,83. Nel grafico si può osservare distintamente la curva (Allegato 4) della ROC. Questa misura di AUC indica una buona capacità discriminatoria del modello: c'è la possibilità dell'83 % che il modello classifichi correttamente un campione positivo come instabile, prima di classificare erroneamente il campione negativo come instabile.

TN_Train	FP_Train	FN_Train	TP_Train	AUC_Train	TN_Test	FP_Test	FN_Test	TP_Test	AUC_Test	
92	0	0	92	1	27	12	11	28	0.820513	
92	0	0	92	1	28	11	6	33	0.856016	
92	0	0	92	1	26	13	14	25	0.748849	
92	0	0	92	1	30	9	10	29	0.851414	
92	0	0	92	1	34	5	10	29	0.888231	
92	0	0	92	1	29	10	10.2	28.8	0.833005	Media
0	0	0	0	0	3.16227766	3.16227766	2.863564	2.863564	0.052812	STD

Tabella 3: tabella riassuntivi dei valori di training e test per le 5 applicazioni della prima configurazione

Anche nel caso della seconda configurazione (Tabella 4) il numero di TN, FN, TP e FP risulta leggermente diverso per ognuna delle 5 applicazioni dell'algoritmo. Per questo motivo vengono analizzati i valori medi:

- Sono presenti mediamente 26 ,4 veri negativi, significa che il modello ha giustamente predetto la stabilità in 26,4 casi.

- Sono presenti mediamente 7,4 falsi negativi, significa che il modello ha predetto erroneamente stabilità in 7,4 casi, quando nella realtà era presente instabilità.
- Sono presenti mediamente 31,6 veri positivi, significa che il modello ha giustamente predetto l'instabilità in 31,6 casi.
- Sono presenti mediamente 12,6 falsi positivi, significa che il modello ha predetto erroneamente instabilità in 12,6 casi, quando nella realtà era presente stabilità.

Nonostante i valori siano differenti rispetto alla prima configurazione, l'accuratezza è ancora pari al 74%, ovvero il modello ha classificato correttamente il 74% dei campioni presenti nel dataset. I valori di AUC variano leggermente a ogni esecuzione del codice, con una media che si attesta intorno a 0,82. Nel grafico (Allegato 5) si può osservare distintamente la curva (Allegato 5) della ROC. Questa misura di AUC indica una buona capacità discriminatoria del modello: c'è la possibilità dell'82 % che il modello classifichi correttamente un campione positivo come instabile, prima di classificare erroneamente il campione negativo come instabile.

TN_Train	FP_Train	FN_Train	TP_Train	AUC_Train	TN_Test	FP_Test	FN_Test	TP_Test	AUC_Test	
92	0	0	92	1	27	12	5	34	0.864563	
92	0	0	92	1	31	8	6	33	0.864563	
92	0	0	92	1	25	14	6	33	0.86259	
92	0	0	92	1	24	15	12	27	0.761999	
92	0	0	92	1	25	14	8	31	0.757396	
92	0	0	92	1	26.4	12.6	7.4	31.6	0.822222	Media
0	0	0	0	0	2.792848	2.792848	2.792848	2.792848	0.057106	STD

Tabella 4: tabella riassuntivi dei valori di training e test per le 5 applicazioni della seconda configurazione

Il numero di TN, FN, TP e FP della terza configurazione risulta leggermente diverso per ognuna delle 5 applicazioni dell'algorithm, come si osserva nei due casi precedenti. Per questo motivo vengono analizzati i valori medi:

- Sono presenti mediamente 30,6 veri negativi, significa che il modello ha giustamente predetto la stabilità in 30,6 casi.

- Sono presenti mediamente 6,4 falsi negativi, significa che il modello ha predetto erroneamente stabilità in 6,4 casi, quando nella realtà era presente instabilità.
- Sono presenti mediamente 32,6 veri positivi, significa che il modello ha giustamente predetto l'instabilità in 32,6 casi.
- Sono presenti mediamente 8,4 falsi positivi, significa che il modello ha predetto erroneamente instabilità in 8,4 casi, quando nella realtà era presente stabilità.

Questa configurazione media ha un'accuratezza media circa dell'81%, ovvero il modello ha classificato correttamente l'81% dei campioni presenti nel dataset. Il valore medio di AUC è pari a 0,87. Nel grafico (Allegato 6) si può osservare distintamente la curva della ROC. Questa misura di AUC indica una capacità discriminatoria del modello migliore delle configurazioni precedenti: c'è la possibilità dell'87 % che il modello classifichi correttamente un campione positivo come instabile, prima di classificare erroneamente il campione negativo come instabile.

TN_Train	FP_Train	FN_Train	TP_Train	AUC_Train	TN_Test	FP_Test	FN_Test	TP_Test	AUC_Test	
92	0	0	92	1	30	9	5	34	0.896778	
92	0	0	92	1	33	6	13	26	0.817226	
92	0	0	92	1	30	9	6	33	0.854043	
92	0	0	92	1	27	12	3	36	0.878369	
92	0	0	92	1	33	6	5	34	0.913872	
92	0	0	92	1	30.6	8.4	6.4	32.6	0.872058	Media
0	0	0	0	0	2.50998008	2.50998	3.847077	3.847077	0.03785	STD

Tabella 5: tabella riassuntivi dei valori di training e test per le 5 applicazioni della terza configurazione

Considerando tutte le configurazioni, si osserva che la terza è la migliore in termini di valore medio dell'AUC. La presenza di tutti i parametri statistici permette al modello di avere una migliore capacità discriminatoria, grazie alla possibilità di comprendere meglio il comportamento delle variabili.

4.3 Risultati analisi finale

L'analisi finale, come spiegato nel paragrafo 3.2.2.3, è stata condotta prendendo in considerazione le tre configurazioni già utilizzate per l'analisi preliminare, con l'aggiunta della quarta configurazione. Quest'ultima è composta esclusivamente dalle variabili significative per il modello, selezionate sulla base dai risultati ottenuti e riportati in Figura 26.

Nei paragrafi successivi vengono presentati e analizzati i risultati dell'analisi finale, per le quattro configurazioni utilizzate. Saranno quindi esaminati: l'errore di Out-Of-Bag, l'istogramma con la stima dell'importanza delle variabili, la matrice di confusione per i dati di training e test, e la curva ROC sempre sia per i dati di training che di test.

Durante questa fase il modello è stato applicato 10 volte per ogni configurazione e, per ciascuna esecuzione, ha costruito 500 alberi decisionali. Questo numero di alberi è stato scelto valutando la stabilizzazione dell'Out-Of-Bag error all'aumentare degli alberi decisionali, nell'analisi preliminare. Oltre questa soglia aumenta solo il tempo di calcolo, senza produrre miglioramenti nel modello; quindi 500 è il numero di alberi ottimale da utilizzare nella foresta.

Per ogni configurazione risultano 10 grafici dell'Out-Of-Bag error; nella Figura 29, Figura 30, Figura 31 e Figura 32 si possono osservare quelli relativi al primo run, gli altri sono riportati nell'Allegato 7, Allegato 8, Allegato 9 e Allegato 10.

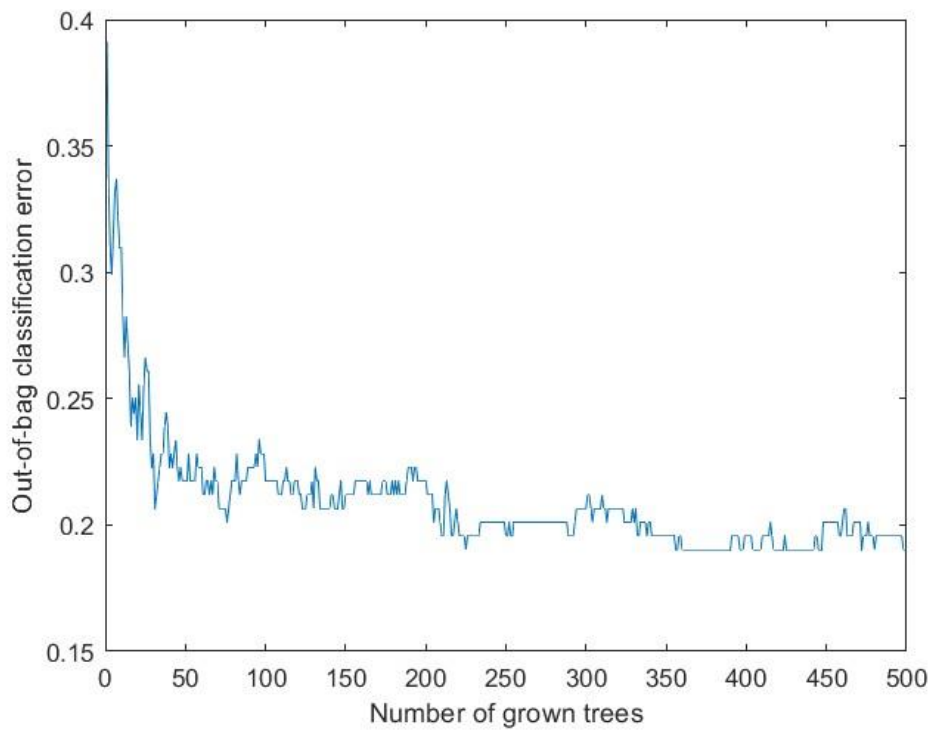


Figura 29: Grafico dell'OOB error del primo run della configurazione con media e majority (prima configurazione).

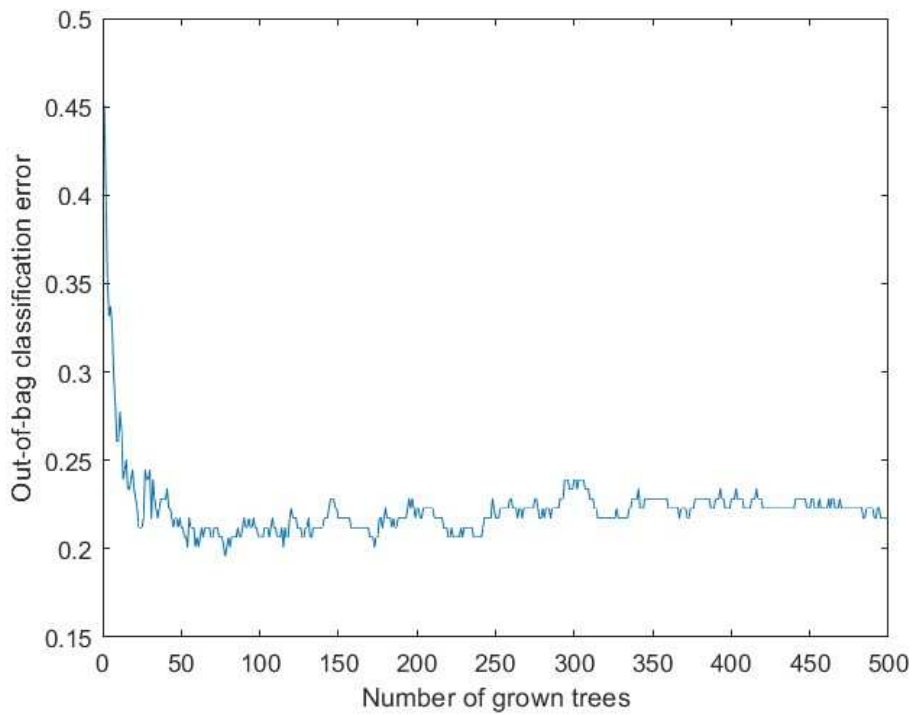


Figura 30: Grafico dell'OOB error del primo run della configurazione con STD e majority (seconda configurazione).

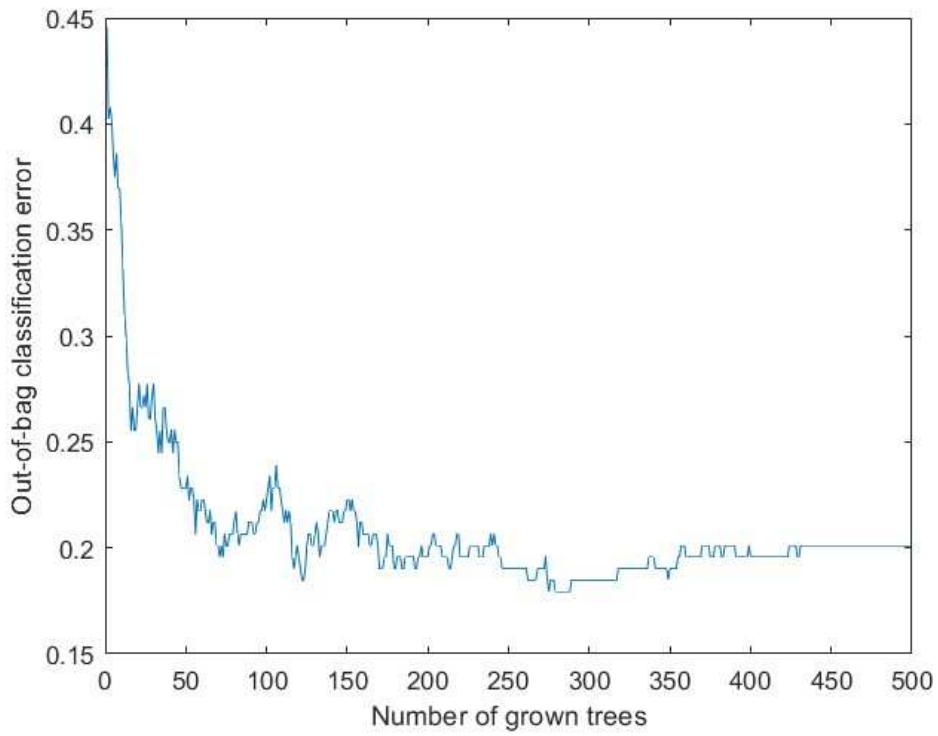


Figura 31 : Grafico dell'OOB error del primo run della configurazione con STD , media e majority delle 31 variabili (terza configurazione).

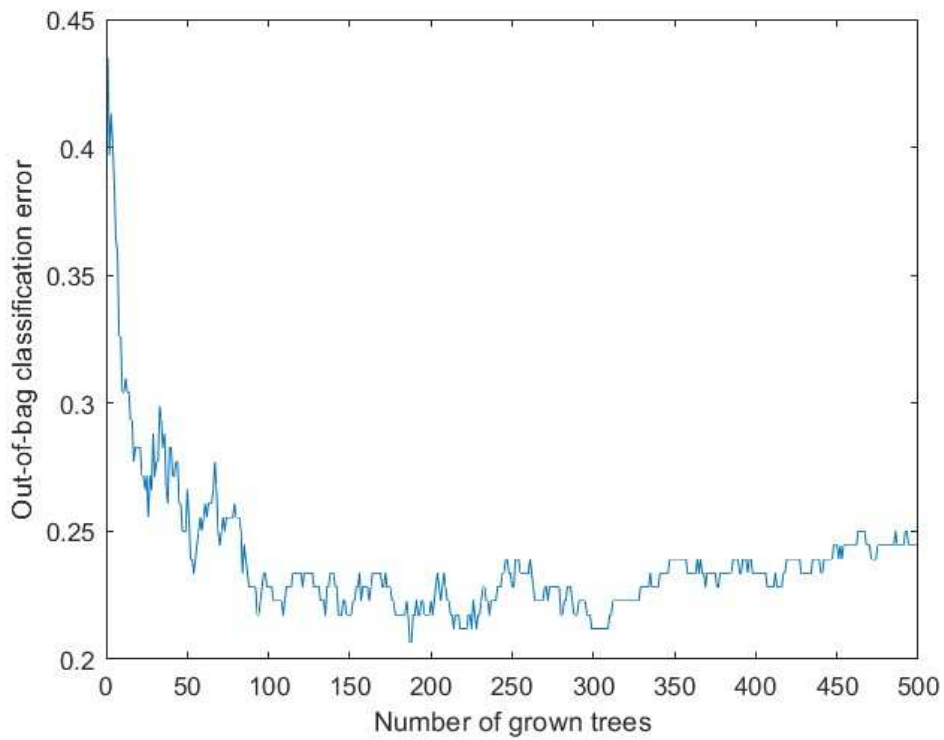


Figura 32: Grafico dell'OOB error del primo run della configurazione considerando solamente le variabili più rilevanti (quarta configurazione).

Sull'asse delle ascisse è riportato il numero degli alberi, che va da 0 a 500; su quello delle ordinate l'OOB classification error, che varia da 0,15 a 0,40 nella prima configurazione; da 0,15 a 0,50 nella seconda configurazione; da 0,15 a 0,45 nella terza configurazione e da 0,2 a 0,45 dell'ultima configurazione. La linea blu mostra l'andamento dell'errore al crescere degli alberi.

Mediamente i grafici riportati hanno un rapido calo dell'OOB errore nella zona iniziale, circa in corrispondenza dei primi 50 alberi. Il modello, con l'aggiunta dei primi alberi, impara dai dati migliorando le proprie capacità predittive.

La massima efficienza è attorno a 450, oltre questa soglia l'errore si stabilizza e l'aumento ulteriore di alberi da un contributo minimo o nullo all'ulteriore miglioramento del modello.

Attraverso il calcolo dell'OOB error, l'algoritmo elabora quattro grafici dell'OOB Permutated Predictor Importance; ovvero gli istogrammi di importanza delle variabili sulle capacità previsionali del modello.

Ogni applicazione dell'algoritmo sulle configurazioni indagate, deriva un istogramma. Successivamente, la media dei singoli risultati per esecuzione, produce il grafico totale con gli intervalli di confidenza (Figura 33), (Figura 34), (Figura 35) e (Figura 36). Sono Quest'ultimi sono rappresentati dalla barretta nera posizionata sulla testa di ogni colonna e indicano il range di valori di importanza, dal massimo al minimo, ottenuti da tutte le simulazioni effettuate. Sull'asse delle ascisse sono riportati i nomi dei predittori, con il prefisso che indica il tipo di analisi statistica che è stata fatta su quella variabile. L'asse delle ordinate, invece, mostra il valore dell'importanza del predittore; valori maggiori indicano un'importanza maggiore sul modello.

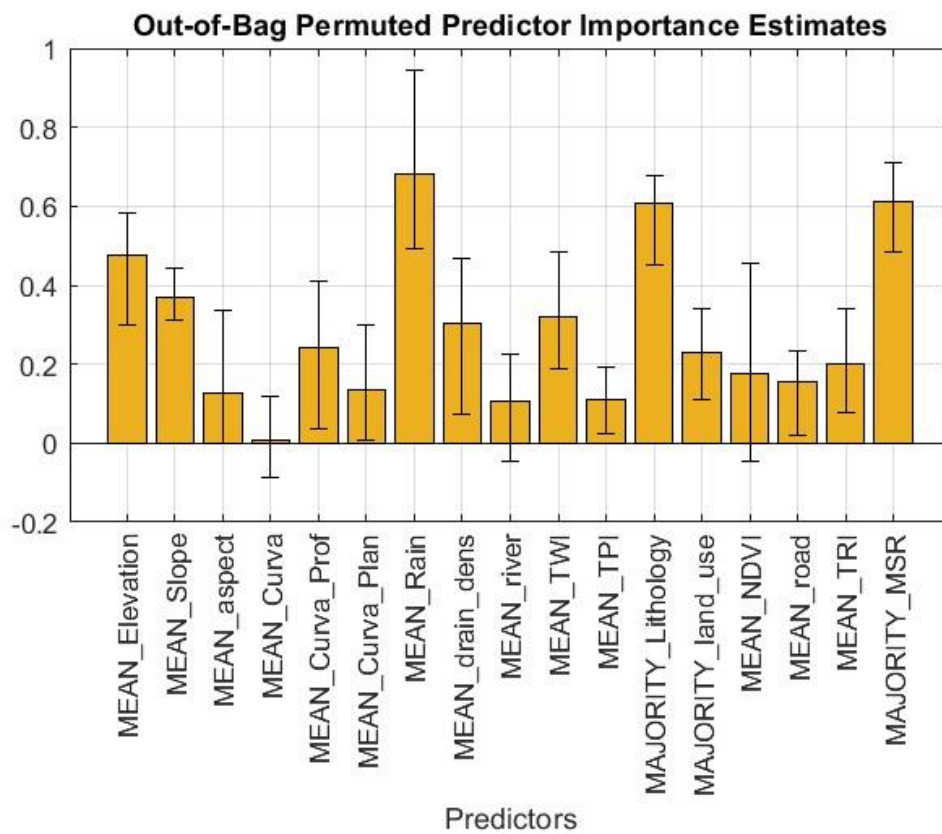


Figura 33: Istogramma dell'importanza delle variabili per la prima configurazione

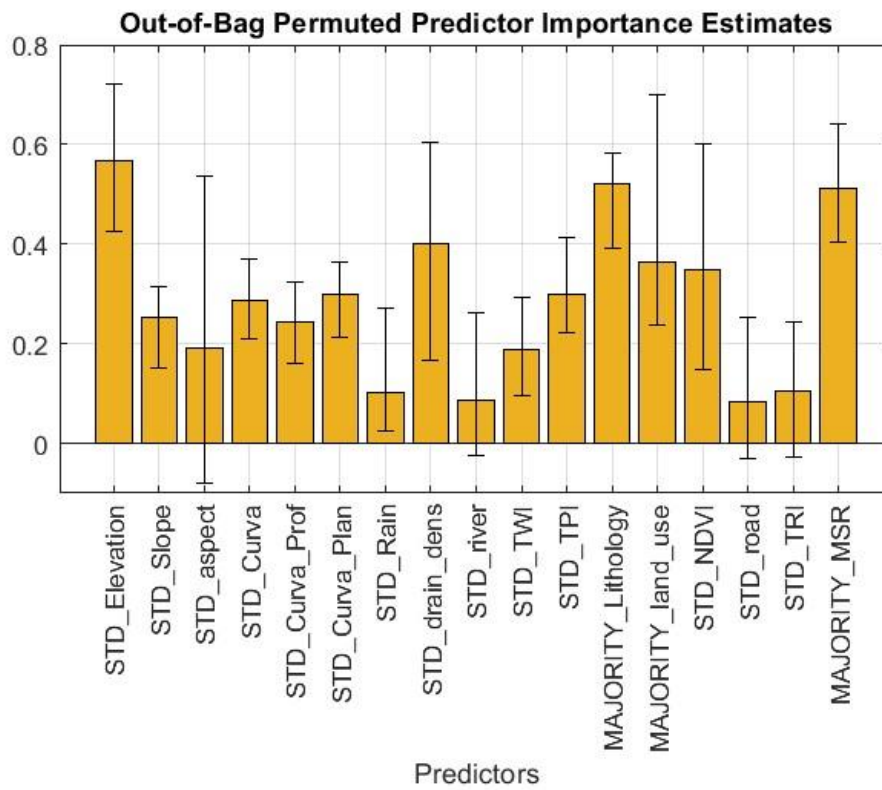


Figura 34: Istogramma dell'importanza delle variabili per la seconda configurazione

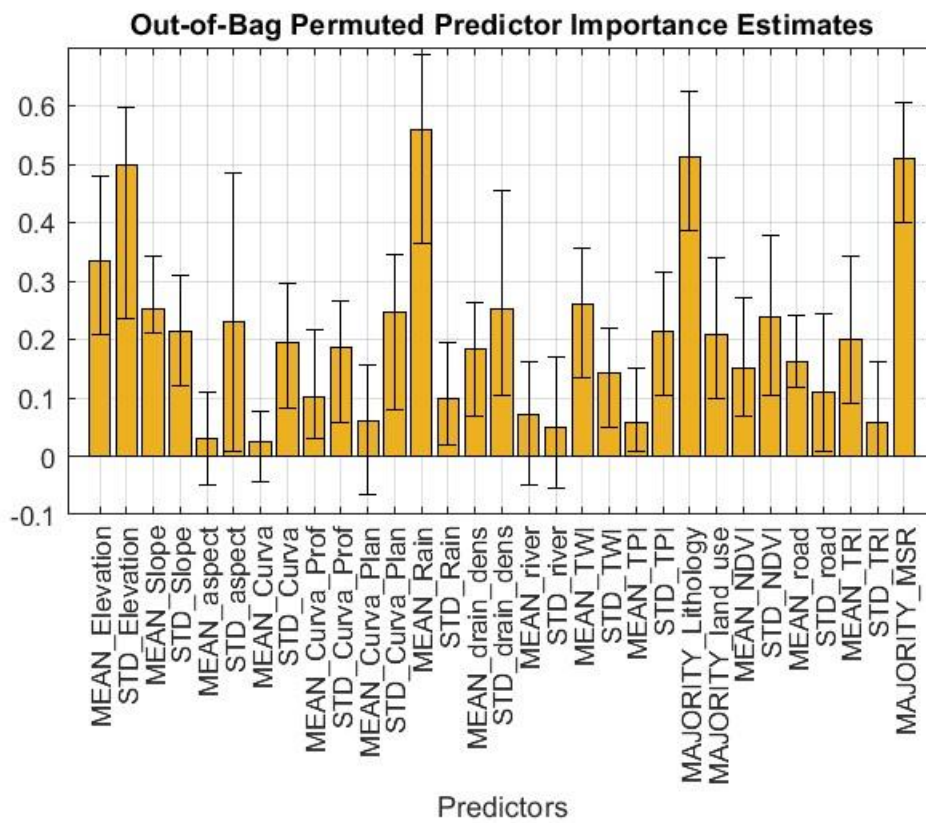


Figura 35: Istogramma dell'importanza delle variabili per la terza configurazione

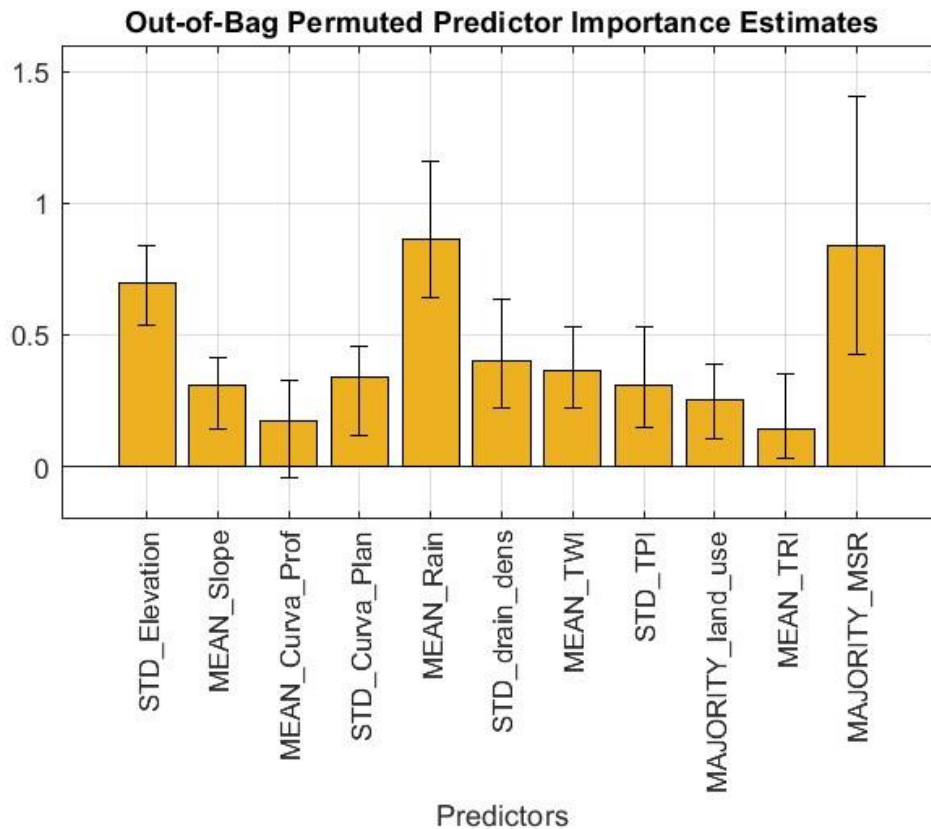


Figura 36: Istogramma dell'importanza delle variabili per la quarta configurazione

I quattro i grafici mostrano intervalli di confidenza, per l'importanza dei predittori, significativamente più ampi rispetto a quanto osservato nell'analisi preliminare. Questo maggiore grado di incertezza è probabilmente determinato dalla maggior sensibilità del modello con 500 alberi al campionamento casuale di dati, durante la fase di bootstrapping. Un numero più elevato di alberi dovrebbe compensare alle fluttuazioni individuali, producendo una stima dell'importanza della variabile più stabile.

L'istogramma della prima configurazione (Figura 33) rispecchia l'andamento osservato nello stesso grafico dell'analisi preliminare. Quindi le variabili più significative sono: le piogge medie, la litologia e l'MSR. I parametri meno influenti, invece, sono l'aspect, la curvatura, la curvatura planare, la distanza dai fiumi e la TPI.

In Figura 34 si osserva l'importanza delle variabili per la seconda configurazione, che è simile all'analisi preliminare: si confermano pesanti, sulle capacità predittive

del modello, la quota, la litologia e l'MSR. Inoltre, incrementano mediamente l'incidenza delle STD dell'uso del suolo, dell'NDVI e della densità di drenaggio. L'importanza delle piogge, come si osservava in precedenza, diminuisce significativamente dalla prima alla seconda configurazione. Questo fenomeno può essere spiegato alla luce dei parametri utilizzati per ogni configurazione. Nel primo caso sono state calcolate le piogge medie, che descrivono l'accumulo complessivo in un lasso di tempo. Nel secondo caso la deviazione standard descrive di quanto le precipitazioni si discostano dal valore medio, evidenziando i picchi e la variabilità delle precipitazioni, come quelli associati a eventi estremi. Essendo le piogge medie più importanti rispetto alla STD delle precipitazioni, si può ipotizzare che, in questa area di studio, l'innescò dei fenomeni franosi è associata all'accumulo prolungato di acqua, piuttosto che a eventi piovosi estremi.

Nella terza configurazione (Figura 35) vengono utilizzati la media, la majority e la deviazione standard di tutti i predittori; pertanto, il risultato ottenuto ricalca l'andamento della combinazione delle due configurazioni precedenti.

La quarta configurazione, mostrata in (Figura 36), propone tre picchi di importanza dei predittori: la deviazione standard della quota, la media delle piogge e la majority dell'MSR. Essendo composta dai parametri più importanti, selezionati sulla base delle indagini preliminari, l'andamento dell'istogramma è prevedibile.

Le successive fasi di analisi della matrice di confusione, della curva ROC e dei valori di AUC, consentono di comprendere come le diverse combinazioni di parametri hanno influito sulle capacità predittive e di classificazione del modello.

La matrice di confusione è stata generata per i dati di training e di test delle 4 configurazioni, con l'algoritmo applicato 10 volte per ciascuna configurazione. Si presenta come un rettangolo composto da 4 celle, che contengono i valori di TN, FP, FN e TP calcolati. Sull'ascissa sono individuate le classi predette, che possono essere pari a 0 (non in frana) e 1 (in frana); mentre in "ordinata" sono individuate le classi reali, a loro volta pari a 0 e 1.

La matrice di confusione, per i dati di training, anche per l'analisi ufficiale è uguale per tutti i run delle configurazioni, come si osserva in Figura 37.

- Sono presenti 92 veri negativi. Il modello ha giustamente predetto la stabilità.

- Sono presenti 0 falsi negativi. Il modello non ha predetto stabilità dove era presente instabilità.
- Sono presenti 92 veri positivi. Il modello ha giustamente predetto instabilità.
- Sono presenti 0 falsi positivi. Il modello non ha predetto instabilità dov'era presente stabilità.

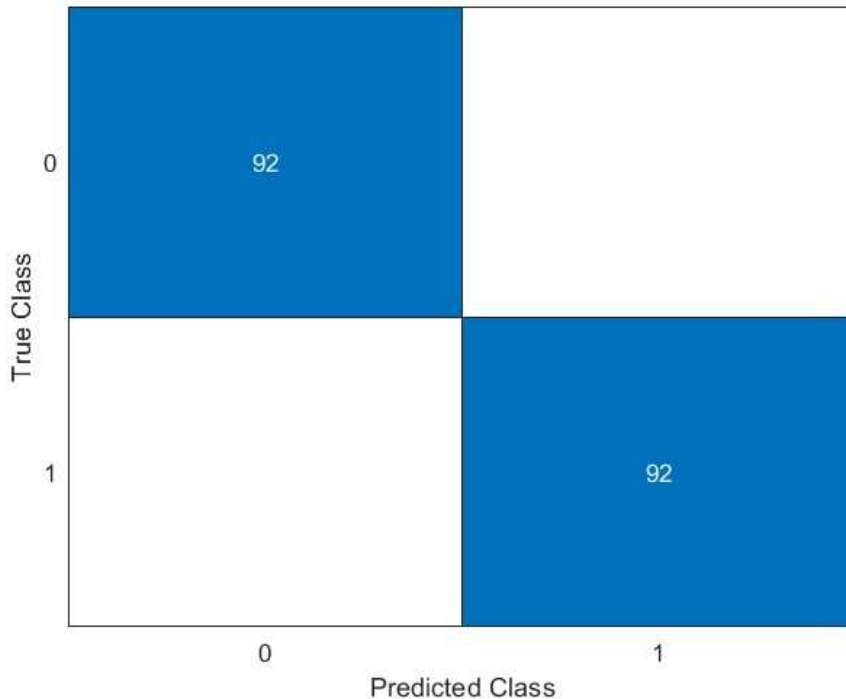


Figura 37: Matrice di confusione della fase di training, uguale per tutte le applicazioni delle quattro configurazioni

Questo tipo di risultato indica un'ottima performance del modello durante la fase di addestramento.

Associate alla matrice di confusione ci sono la curva ROC e l'area sottesa alla curva (AUC), sempre calcolate per i dati di training e di test.

In (Figura 38) si può osservare anche queste si mantengono costanti per tutte le applicazioni delle quattro configurazioni. L'AUC è pari a 1, quindi il modello è in grado di separare perfettamente le classi positive da quelle negative. La curva, invece, risale lungo l'asse delle ordinate, fino all'angolo superiore sinistro, dove il TPR è pari a 1 e il FPR è zero.

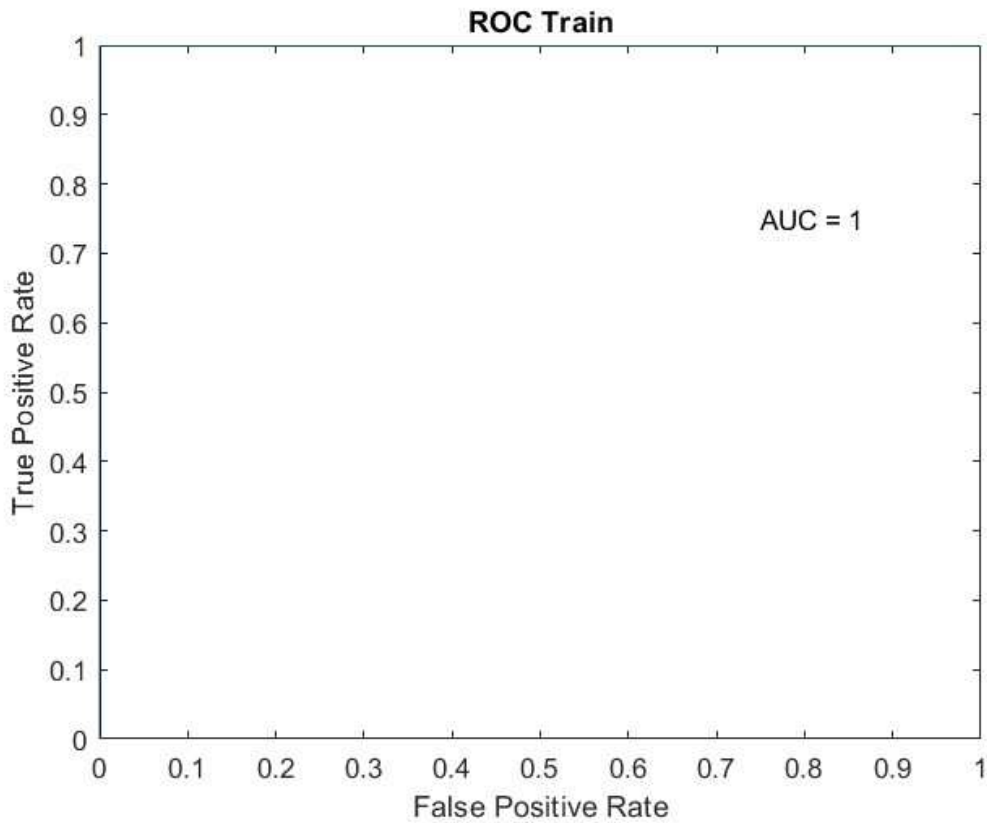


Figura 38: grafico della curva ROC e valore di AUC costanti per tutti i run delle tre configurazioni.

I risultati della fase di training e test sono riassunti nelle seguenti tabelle.

Il numero di TN, FN, TP e FP per la prima configurazione (Tabella 6) varia leggermente ad ogni applicazione dell'algoritmo, pertanto sono stati considerati i valori medi:

- Sono presenti mediamente 29 veri negativi, significa che il modello ha giustamente predetto la stabilità in 29 casi.
- Sono presenti mediamente 8,9 falsi negativi, significa che il modello ha predetto erroneamente stabilità in 8,9 casi, quando nella realtà era presente instabilità.
- Sono presenti mediamente 30,1 veri positivi, significa che il modello ha giustamente predetto l'instabilità in 30,1 casi.
- Sono presenti mediamente 10 falsi positivi, significa che il modello ha predetto erroneamente instabilità in 10 casi, quando nella realtà era presente stabilità.

Questa suddivisione media ha un'accuratezza circa del 76%, ovvero il modello ha classificato correttamente il 76% dei campioni presenti nel dataset. Questo valore non è sufficiente per valutare correttamente la performance del modello; perciò, vanno consideranti anche la curva ROC e il valore di AUC. Anche in questo caso si differenziano leggermente per ogni run del codice: il valore medio di AUC è 0,84. Questa misura di AUC indica una buona capacità discriminatoria del modello: c'è la possibilità dell'84 % che il modello classifichi correttamente un campione positivo come instabile, prima di classificare erroneamente il campione negativo come instabile.

In Figura 39 si possono osservare i grafici della curva ROC per le 10 esecuzioni del modello.

Lo stesso caso, nella fase preliminare, presenta un'accuratezza del 74% e un valore di AUC pari a 0,83. Questi risultati indicano una prestazione leggermente migliore del modello con 500 alberi, che è stato testato su un maggior numero di iterazioni. Va inoltre osservata la bassa variabilità dei valori di AUC (espressa come deviazione standard), ad indicare una buona stabilità del modello tra le varie iterazioni.

TN_Train	FP_Train	FN_Train	TP_Train	AUC_Tra	TN_Test	FP_Test	FN_Test	TP_Test	AUC_Test	
92	0	0	92	1	27	12	14	25	0.767258	
92	0	0	92	1	29	10	8	31	0.838264	
92	0	0	92	1	28	11	9	30	0.815253	
92	0	0	92	1	29	10	5	34	0.892834	
92	0	0	92	1	30	9	13	26	0.844839	
92	0	0	92	1	27	12	8	31	0.847469	
92	0	0	92	1	28	11	10	29	0.798817	
92	0	0	92	1	26	13	6	33	0.800789	
92	0	0	92	1	33	6	7	32	0.923734	
92	0	0	92	1	33	6	9	30	0.895464	
92	0	0	92	1	29	10	8.9	30.1	0.842472	Media
0	0	0	0	0	2.4037	2.40370085	2.84605	2.84605	0.049483	STD

Tabella 6: tabella riassuntivi dei valori di training e test per le 10 applicazioni della prima configurazione

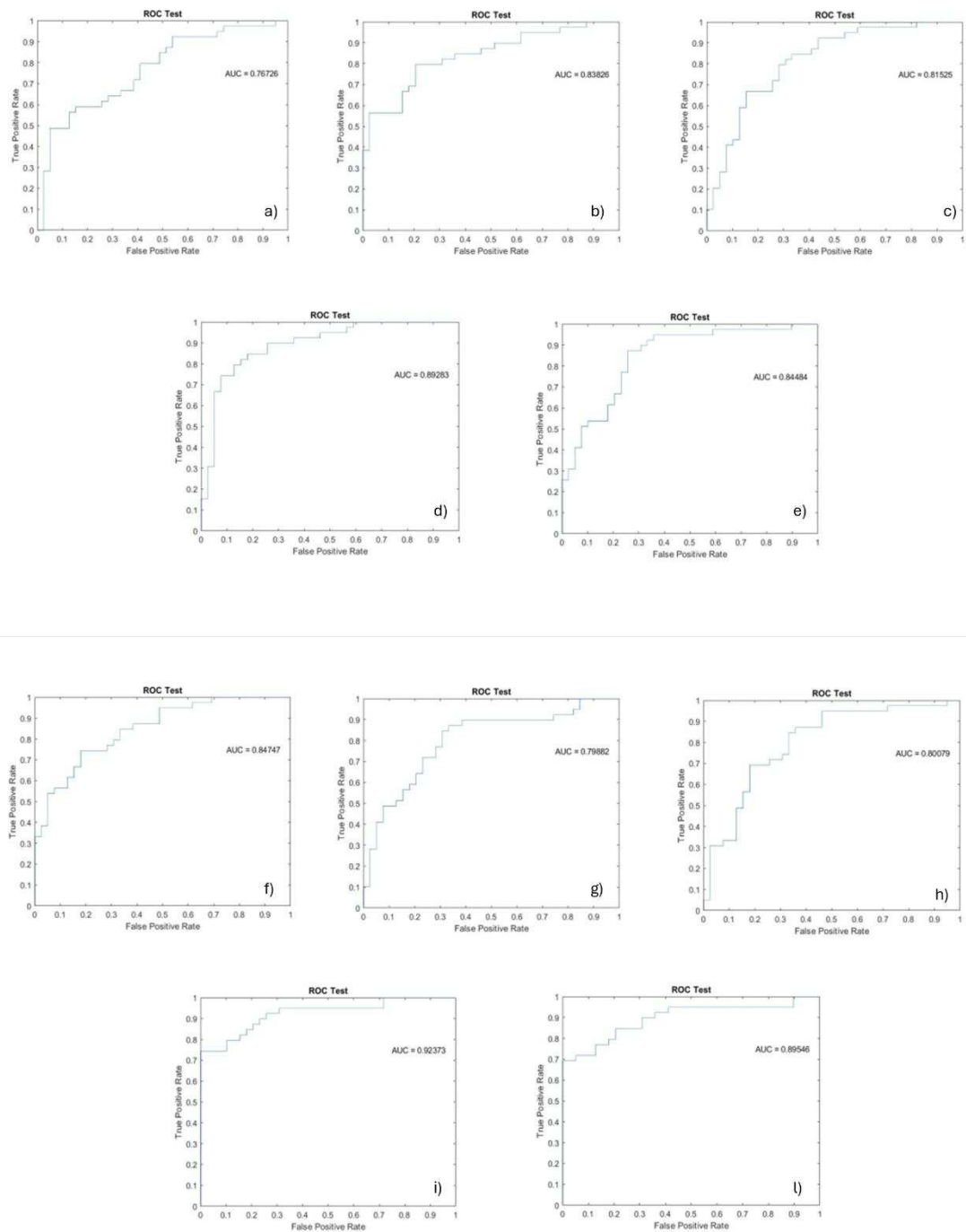


Figura 39: Curve ROC e valori di AUC delle 10 esecuzioni della prima configurazione dell'analisi finale. Nella prima riga, da sinistra verso destra, si osservano il run 1, run 2 e run 3. Nella seconda riga, da sinistra verso destra, si osservano il run 4 e il run 5. Nella terza riga, da sinistra verso destra, si osservano run 6, run 7, run 8. Nella quarta riga, da sinistra verso destra, si osservano run 9 e run 10 (analisi finale)

Per la seconda configurazione (Tabella 7) il numero di TN, FN, TP e FP risulta leggermente diverso per ognuna delle 10 applicazioni dell'algoritmo. I valori medi selezionati sono:

- Sono presenti mediamente 27,1 veri negativi, significa che il modello ha giustamente predetto la stabilità in 27,1 casi.
- Sono presenti mediamente 8,3 falsi negativi, significa che il modello ha predetto erroneamente stabilità in 8,3 casi, quando nella realtà era presente instabilità.
- Sono presenti mediamente 30,7 veri positivi, significa che il modello ha giustamente predetto l'instabilità in 30,7 casi.
- Sono presenti mediamente 11,9 falsi positivi, significa che il modello ha predetto erroneamente instabilità in 11,9 casi, quando nella realtà era presente stabilità.

Questa configurazione ha un'accuratezza media circa del 74%, ovvero il modello ha classificato correttamente il 74% dei campioni presenti nel dataset. Il valore medio di AUC è pari a 0,83. Questa misura di AUC indica una buona capacità discriminatoria del modello: c'è la possibilità dell'83% che il modello classifichi correttamente un campione positivo come instabile, prima di classificare erroneamente il campione negativo come instabile.

In Figura 40 si possono osservare i grafici delle curve ROC delle dieci iterazioni di questa configurazione.

Lo stesso caso presenta un'accuratezza del 74% anche nella fase preliminare; mentre il valore di AUC pari a 0,82. Anche questi risultati indicano una performance leggermente migliore del modello con 500 alberi.

TN_Train	FP_Train	FN_Train	TP_Train	AUC_Train	TN_Test	FP_Test	FN_Test	TP_Test	AUC_Test	
92	0	0	92	1	28	11	10	29	0.840237	
92	0	0	92	1	26	13	11	28	0.788297	
92	0	0	92	1	28	11	4	35	0.869822	
92	0	0	92	1	27	12	6	33	0.81854	
92	0	0	92	1	24	15	3	36	0.876397	
92	0	0	92	1	24	15	7	32	0.816568	
92	0	0	92	1	29	10	6	33	0.888231	
92	0	0	92	1	28	11	9	30	0.836949	
92	0	0	92	1	32	7	19	20	0.808021	
92	0	0	92	1	25	14	8	31	0.775148	
92	0	0	92	1	27.1	11.9	8.3	30.7	0.831821	Media
0	0	0	0	0	2.469817807	2.469818	4.522782	4.522782	0.037701	STD

Tabella 7: tabella riassuntivi dei valori di training e test per le 10 applicazioni della seconda configurazione

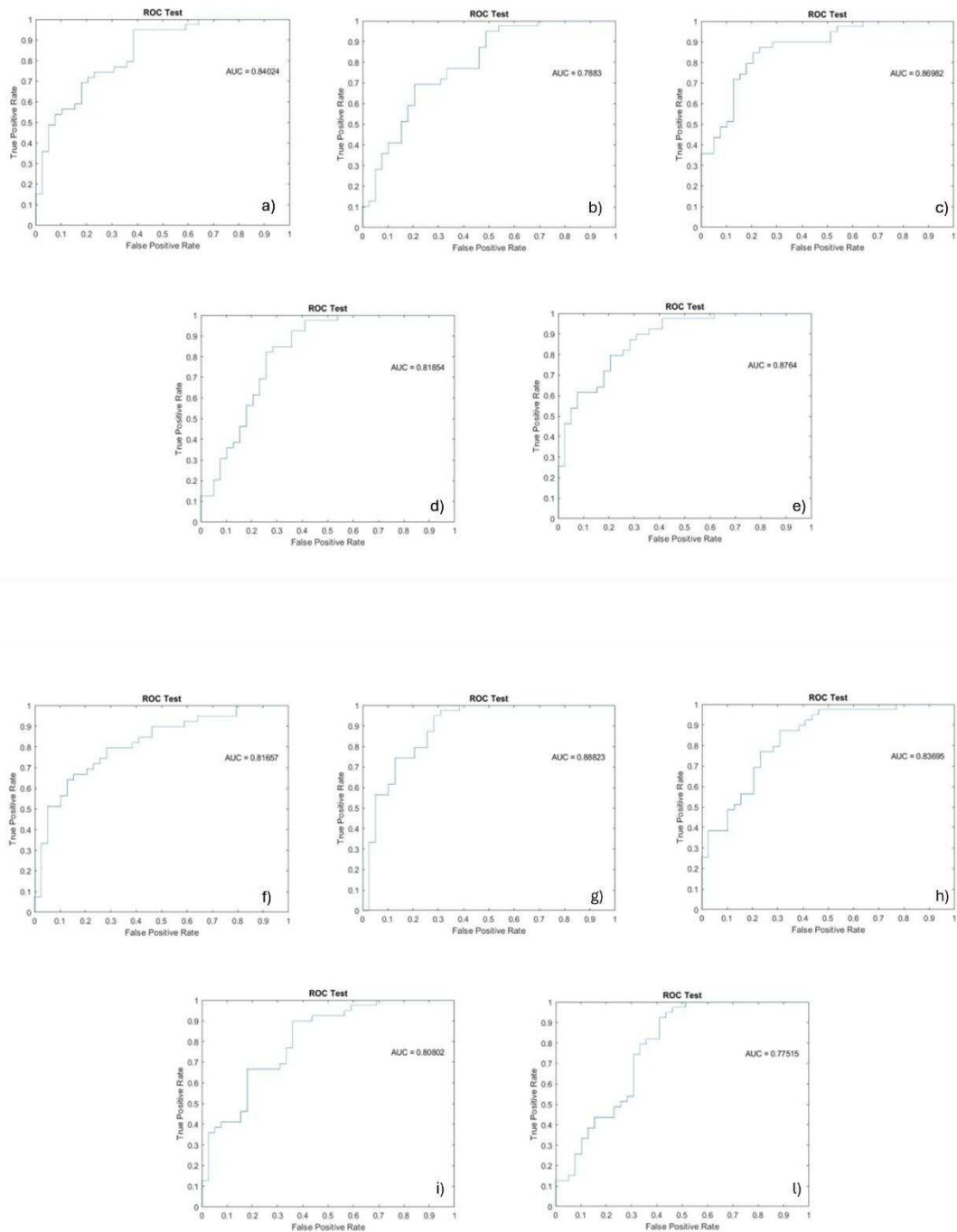


Figura 40: Curve ROC e valori di AUC delle 10 esecuzioni della seconda configurazione. Nella prima riga, da sinistra verso destra, si osservano il run 1, run 2 e run 3. Nella seconda riga, da sinistra verso destra, si osservano il run 4 e il run 5. Nella terza riga, da sinistra verso destra, si osservano run 6, run 7, run 8. Nella quarta riga, da sinistra verso destra, si osservano run 9 e run 10 (analisi finale)

Il numero di TN, FN, TP e FP per la terza configurazione (Tabella 8) risulta leggermente diverso per ognuna delle 10 applicazioni dell'algorithm, come si osserva nei due casi precedenti. Per questo motivo vengono analizzati i dati medi:

- Sono presenti mediamente 28,1 veri negativi, significa che il modello ha giustamente predetto la stabilità in 28,1 casi.
- Sono presenti mediamente 8,6 falsi negativi, significa che il modello ha predetto erroneamente stabilità in 8,6 casi, quando nella realtà era presente instabilità.
- Sono presenti mediamente 30,4 veri positivi, significa che il modello ha giustamente predetto l'instabilità in 30,4 casi.
- Sono presenti mediamente 10,9 falsi positivi, significa che il modello ha predetto erroneamente instabilità in 10,9 casi, quando nella realtà era presente stabilità.

Questa suddivisione media ha un'accuratezza circa del 75%, ovvero il modello ha classificato correttamente il 75% dei campioni presenti nel dataset.

Per questa configurazione il valore medio di AUC è pari 0,82, che indica una buona capacità discriminativa del modello. I risultati della curva ROC per le 10 esecuzioni di questa configurazione si possono osservare in Figura 41.

Lo stesso assetto, nella fase preliminare, presenta un'accuratezza dell'81%, mentre il valore di AUC pari a 0,87. Questi risultati indicano un peggioramento della performance del modello con 500 alberi; anche se complessivamente rimane un rendimento molto buono. Per questo studio, infatti, tutti i valori di AUC superiori a 0,7 sono considerati validi e accettabili.

TN_Train	FP_Train	FN_Train	TP_Train	AUC_Train	TN_Test	FP_Test	FN_Test	TP_Test	AUC_Test	
92	0	0	92	1	23	16	11	28	0.756739	
92	0	0	92	1	28	11	8	31	0.840894	
92	0	0	92	1	28	11	7	32	0.82117	
92	0	0	92	1	30	9	10	29	0.850099	
92	0	0	92	1	27	12	4	35	0.811966	
92	0	0	92	1	29	10	10	29	0.813281	
92	0	0	92	1	28	11	7	32	0.849441	
92	0	0	92	1	31	8	7	32	0.876397	
92	0	0	92	1	27	12	10	29	0.773833	
92	0	0	92	1	30	9	12	27	0.793557	
92	0	0	92	1	28.1	10.9	8.6	30.4	0.818738	Media
0	0	0	0	0	2.233582076	2.233582	2.412928	2.412928	0.036995	STD

Tabella 7: tabella riassuntiva dei valori di training e test per le 10 applicazioni della terza configurazione

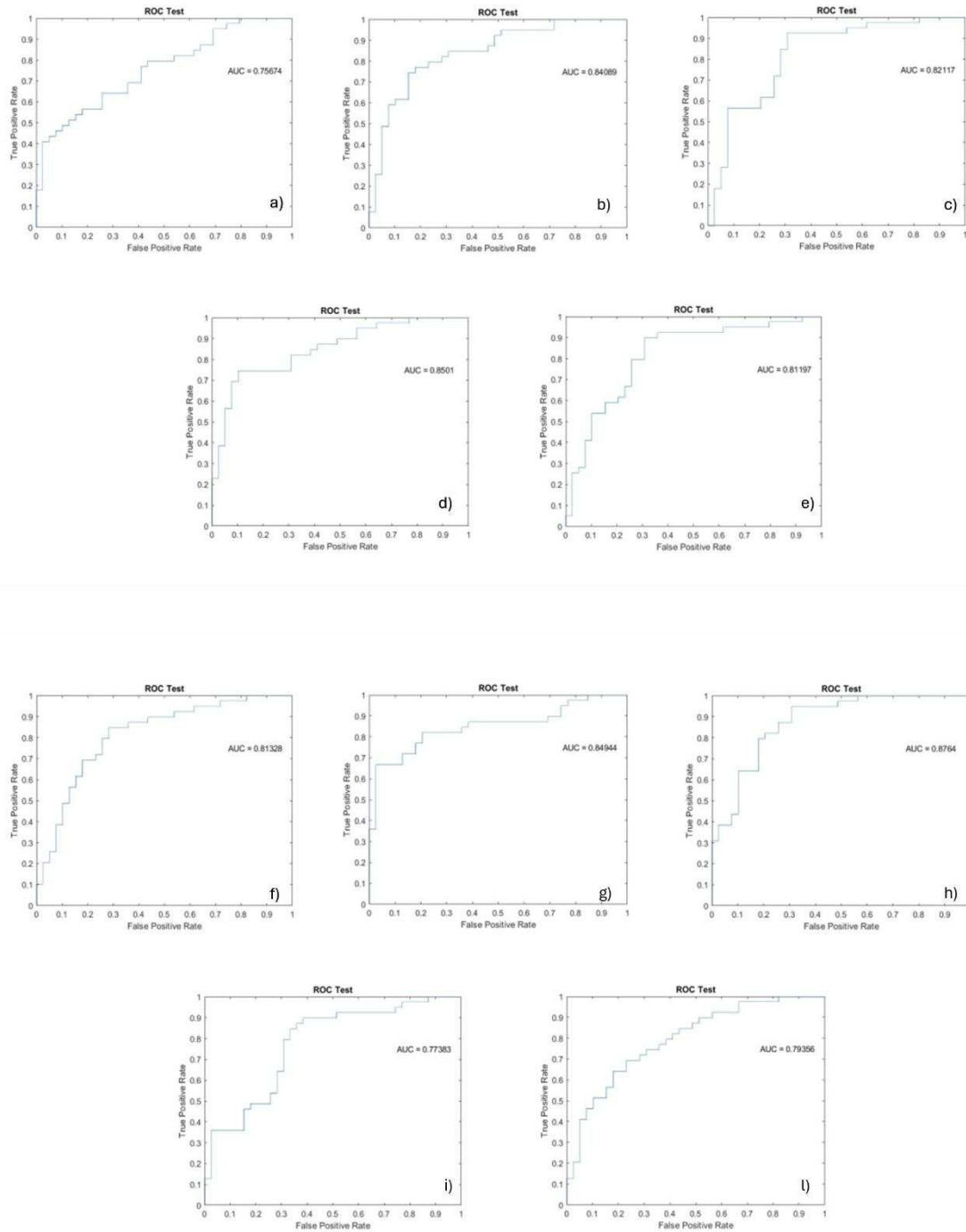


Figura 41: Curve ROC e valori di AUC delle 10 esecuzioni della terza configurazione. Nella prima riga, da sinistra verso destra, si osservano il run 1, run 2 e run 3. Nella seconda riga, da sinistra verso destra, si osservano il run 4 e il run 5. Nella terza riga, da sinistra verso destra, si osservano run 6, run 7, run 8. Nella quarta riga, da sinistra verso destra, si osservano run 9 e run 10 (analisi finale)

Anche per questa configurazione inedita (quarta configurazione, Tabella 9) il numero di TN, FN, TP e FP risulta leggermente diverso per ognuna delle 10 applicazioni dell'algorithm. Quindi i valori medi considerati sono:

- Sono presenti mediamente 29,1 veri negativi, significa che il modello ha giustamente predetto la stabilità in 29,1 casi.
- Sono presenti mediamente 8 falsi negativi, significa che il modello ha predetto erroneamente stabilità in 8 casi, quando nella realtà era presente instabilità.
- Sono presenti mediamente 31 veri positivi, significa che il modello ha giustamente predetto l'instabilità in 31 casi.
- Sono presenti mediamente 9,9 falsi positivi, significa che il modello ha predetto erroneamente instabilità in 9,9 casi, quando nella realtà era presente stabilità.

Questa suddivisione media ha un'accuratezza circa dell'87% e un AUC medio di 0,85. Questa misura di AUC indica una buona capacità discriminativa del modello: c'è la possibilità dell'85% che il modello classifichi correttamente un campione positivo come instabile, prima di classificare erroneamente il campione negativo come instabile.

In Figura 42 si possono osservare i relativi grafici della curva ROC per le dieci iterazioni.

TN_Train	FP_Train	FN_Train	TP_Train	AUC_Train	TN_Test	FP_Test	FN_Test	TP_Test	AUC_Test	
92	0	0	92	1	27	12	7	32	0.822485	
92	0	0	92	1	32	7	8	31	0.839579	
92	0	0	92	1	30	9	9	30	0.86259	
92	0	0	92	1	29	10	12	27	0.844839	
92	0	0	92	1	26	13	9	30	0.819855	
92	0	0	92	1	30	9	7	32	0.876397	
92	0	0	92	1	30	9	8	31	0.846154	
92	0	0	92	1	33	6	6	33	0.905983	
92	0	0	92	1	26	13	5	34	0.868508	
92	0	0	92	1	28	11	9	30	0.780408	
92	0	0	92	1	29.1	9.9	8	31	0.84668	Media
0	0	0	0	0	2.3781	2.378141198	1.943651	1.943651	0.03477	STD

Tabella 8: Tabella 9: tabella riassuntivi dei valori di training e test per le 10 applicazioni della quarta configurazione

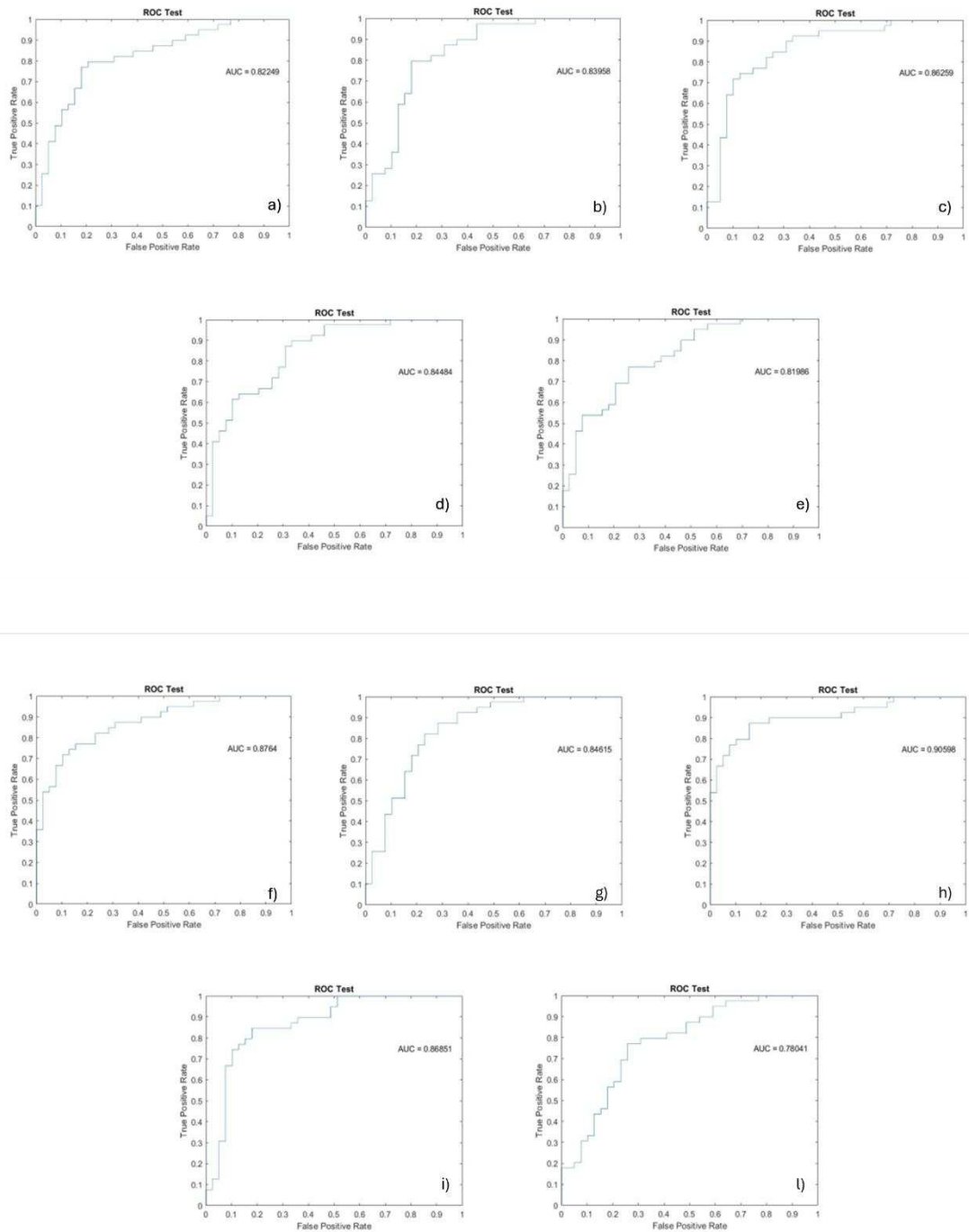


Figura 42: Curve ROC e valori di AUC delle 10 esecuzioni della quarta configurazione. Nella prima riga, da sinistra verso destra, si osservano il run 1, run 2 e run 3. Nella seconda riga, da sinistra verso destra, si osservano il run 4 e il run 5. Nella terza riga, da sinistra verso destra, si osservano run 6, run 7, run 8. Nella quarta riga, da sinistra verso destra, si osservano run 9 e run 10 (analisi finale).

Tutti i risultati ottenuti dalle analisi finali sono ampiamente soddisfacenti, in quanto illustrano dei modelli molto performanti; nonostante la rimozione di alcuni predittori poco influenti, il modello migliore risulta essere quello relativo all'ultima configurazione. Questo risultato suggerisce che la mappa di suscettibilità da frana può essere generata efficientemente utilizzando soltanto i fattori condizionanti più significativi.

4.3.1 Interpretazione dei risultati

I Partial Dependence Plots (PDPs) sono stati calcolati esclusivamente per la quarta configurazione, quella che si è dimostrata essere la più efficace. Il loro scopo è visualizzare come ogni classe, di ciascun predittore, influenzi il risultato finale del modello, quindi, sulla previsione della suscettibilità da frana.

Per tutte le variabili sono stati generati dieci PDPs, uno per ogni applicazione del modello. Nella Figura 43, i risultati sono riassunti in un unico grafico per fattore predisponente, in cui ogni linea rappresenta una diversa applicazione del modello. Sull'asse delle ascisse sono indicate le classi di valori dei diversi parametri, mentre sull'asse delle ordinate sono riportati i valori dello score, ovvero l'effetto medio della variabile sul risultato previsto.

Lo score della deviazione standard della quota mostra un iniziale picco di crescita per tutti i run, seguito da una stabilizzazione in corrispondenza del valore pari 200 m della STD. Questo significa che, all'aumentare della variabilità della quota delle Slope Units, aumenta anche l'influenza della variabile sul risultato finale. La rapida crescita iniziale suggerisce un impatto maggiore di questa variabile, rispetto ad altre considerate. I risultati ottenuti sono in linea con quanto osservato nell'istogramma dell'OOB, dove la quota risultava tra i predittori più importanti (Figura 36).

Lo score dello slope medio ha un andamento oscillante, con dei picchi tra 0° e 20° . Questo comportamento indica un'influenza sul modello non lineare.

La media della curvatura di profilo inizialmente ha uno score elevato; cala rapidamente intorno allo zero, andando ad impattare negativamente sulla capacità previsionale del modello. Si stabilizza a valori bassi.

La deviazione standard della curvatura planare ha un andamento praticamente opposto: lo score mostra una crescita iniziale, paragonabile a quella dell'elevazione e

delle piogge medie. Si osserva però una successiva fase di calo.

La curva della media delle piogge ha un andamento più simile a quello già descritto per la deviazione standard dell'elevazione. L'impatto di questo termine sul modello finale cresce linearmente con l'aumento delle precipitazioni, fino a valori di 1500 mm, dove la curva si stabilizza. Significa che precipitazioni media più basse influenzano meno il risultato (positivo) del modello, mentre valori maggiori sono più impattanti. Questo iniziale picco di crescita, come detto in precedenza, suggerisce una maggiore rilevanza delle piogge medie tra 1200 e 1500 mm/anno, rispetto agli altri predittori. Anche questo risultato è in linea con l'istogramma dell'importanza delle variabili (Figura 36).

L'andamento crescente dello score si può osservare anche per la deviazione standard della densità di drenaggio; in questo caso sono presenti anche delle fluttuazioni che rendono più complessa la comprensione dell'influenza che ogni classe ha sul risultato finale.

Lo score del TWI medio ha un picco tra 5 e 6, a cui segue una fase decrescente del valore. Questa rapida crescita indica l'elevata importanza del predittore in questo intervallo ristretto. L'indice di posizione topografica standardizzato (TPI), invece, ha effetti complessi e non lineari sul modello: presenta due picchi, con oscillazioni, iniziali, con valori che successivamente tendono a scendere e stabilizzarsi. Per questo parametro si osserva che i valori elevati hanno una minore influenza sulla suscettibilità di frana, indicando che versanti più regolari sono maggiormente associati all'innescio di frane nell'area studio.

Anche l'uso del suolo influenza il modello in modo discontinuo, legato all'apporto di ogni singola classe: i terreni agricoli (classe 2) e le foreste (classe 4) si sono dimostrate le classi più favorevoli alla discriminazione di aree in frana.

L'indice di rugosità medio (TRI) ha uno score che cala nettamente oltre il valore di 0,5: questo può indicare che terreni più rugosi influenzano meno nettamente la performance del modello.

L'MRS è la variabile chiave di questo modello. Questa osservazione può essere supportata dalle conoscenze pregresse sulla posizione delle frane catalogate nell'area e dai risultati ottenuti dall'istogramma dell'Out-Of-Bag, che riflettono l'elevata importanza di questa variabile nel modello.

Nel PDP del MRS, è ragionevole aspettarsi che la litologia più franosa, ovvero le vulcaniti (MSR = 7), abbia uno score molto elevato, che indica un significativo impatto sulle capacità predittive del modello. Man mano che si scende lungo la scala della franosità delle litologie, lo score diminuisce.

Tuttavia, si osserva un aumento dello score per litologie meno franose; pertanto queste classi sono influenti sulle capacità predittive del modello: essendo litologie stabili, giocano un ruolo significativo nel determinare la previsione di “non frana”, pesando sul modello quando deve distinguere le aree che probabilmente non sono in frana.

I risultati dei PDPs, calcolati sulla quarta configurazione, si sono rivelati efficienti per confermare la variazione di quota, le piogge medie e l'MSR come i predittori che contribuiscono maggiormente alle performance del codice. Come già detto, questo rendimento ricalca quello espresso dai diagrammi di importanza delle variabili, non solo della quarta configurazione, ma anche delle precedenti. Il valore aggiunto dei PDPs ricade nella capacità di identificare i valori di ogni singola variabile che maggiormente influenzano la suscettibilità da frana.

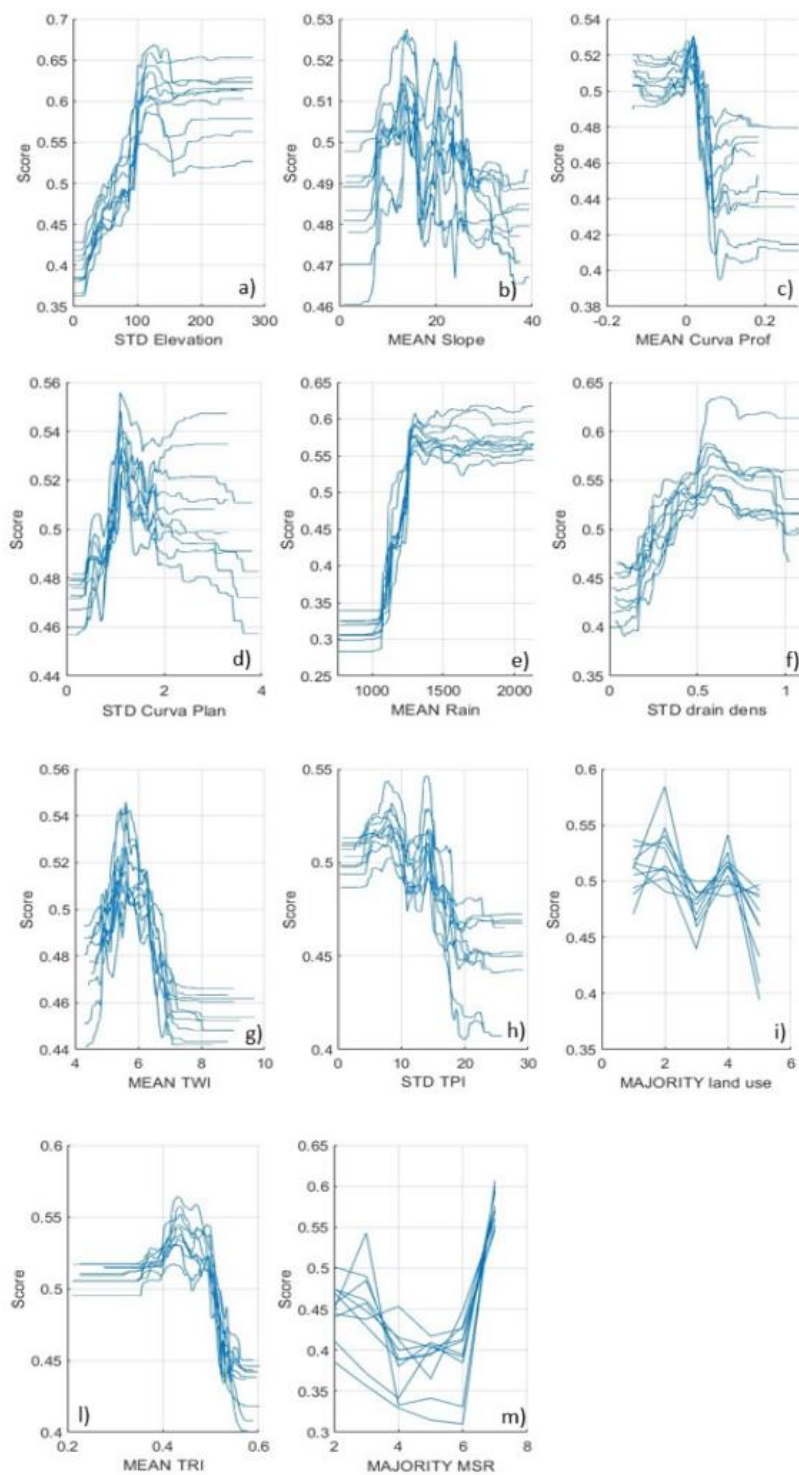


Figura 43: PDPs di tutte le variabili predisponenti della quarta configurazione.

4.4 Mappa di suscettibilità da frana

Lo scopo delle fasi descritte in precedenza era l'addestramento del modello di Random Forest per prevedere la pericolosità spaziale di un evento franoso.

L'algoritmo ha imparato a sfruttare i dati di input per calcolare la possibilità che una frana si verifichi all'interno di ogni SU.

Sono state generate quaranta mappe di suscettibilità da frana, una per ogni esecuzione delle 4 configurazioni adottate. Successivamente, sono riportate solo le quattro mappe finali, ottenute mediando i valori delle dieci simulazioni fatte per ogni configurazione.

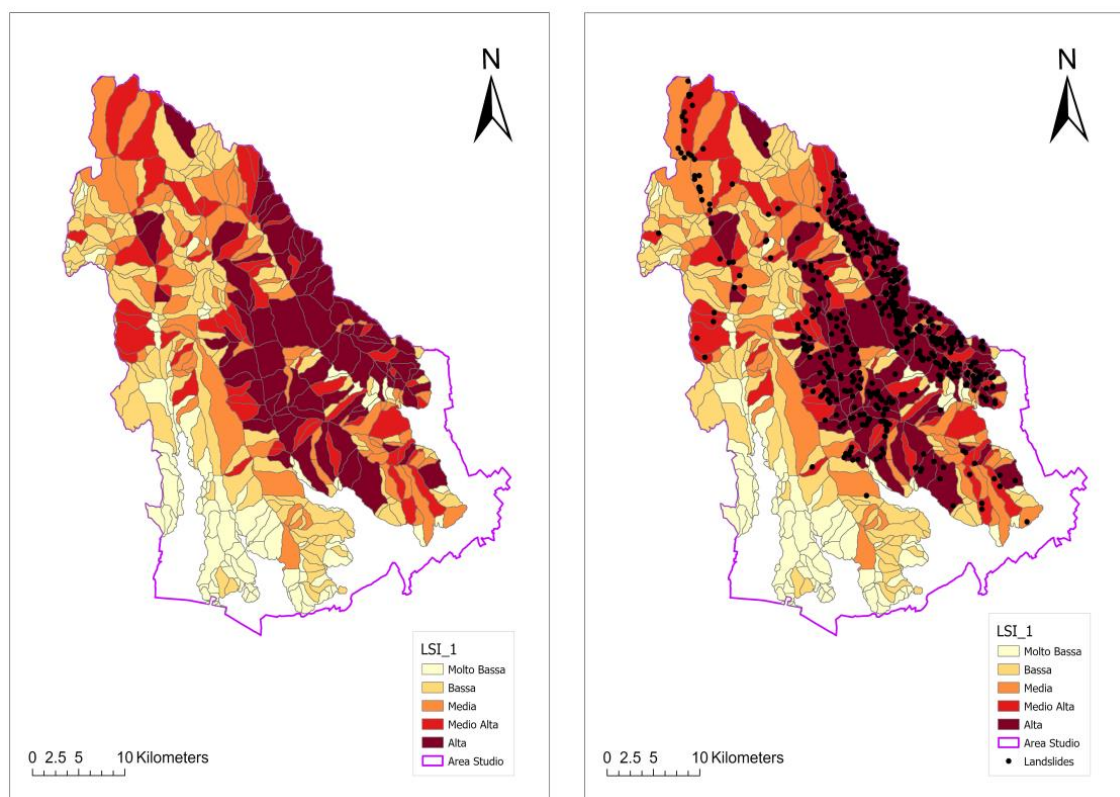


Figura 44: A sinistra è rappresentata la mappa di suscettibilità da frana media relativa alla prima configurazione, mentre a destra è mostrata la stessa mappa con la sovrapposizione dei punti di frana ricavati dall'Inventario dei Fenomeni Franosi Italiani (IFFI).

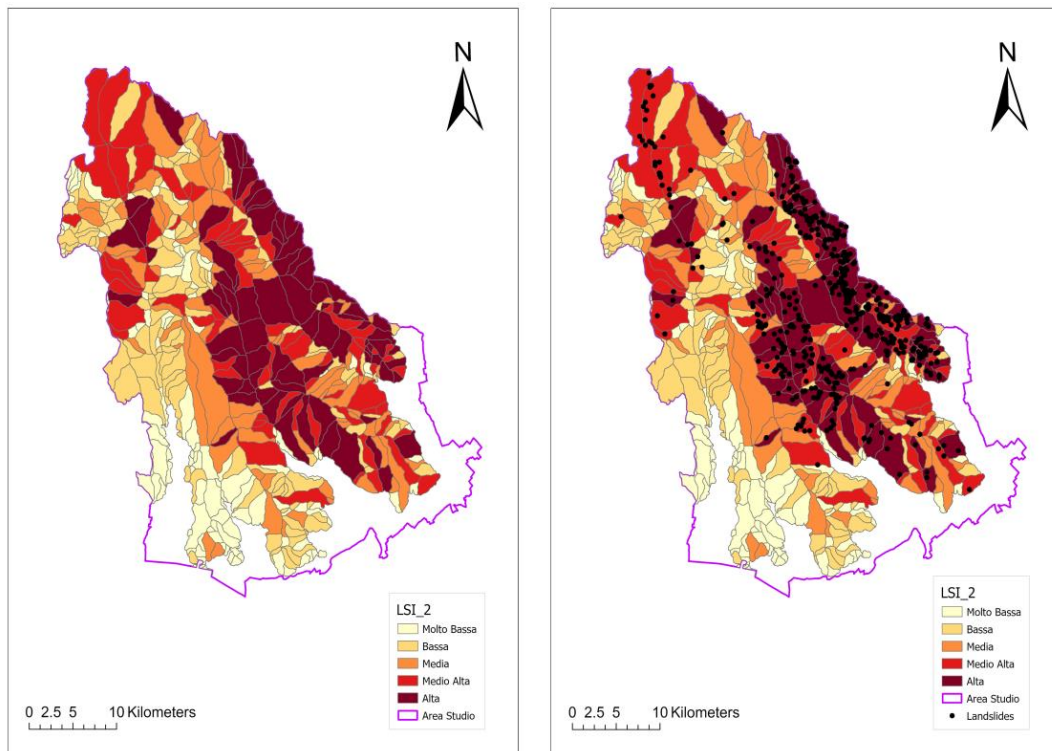


Figura 45: A sinistra è rappresentata la mappa di suscettibilità da frana media relativa alla seconda configurazione, mentre a destra è mostrata la stessa mappa con la sovrapposizione dei punti di frana ricavati dall'Inventario dei Fenomeni Franosi Italiani (IFFI).

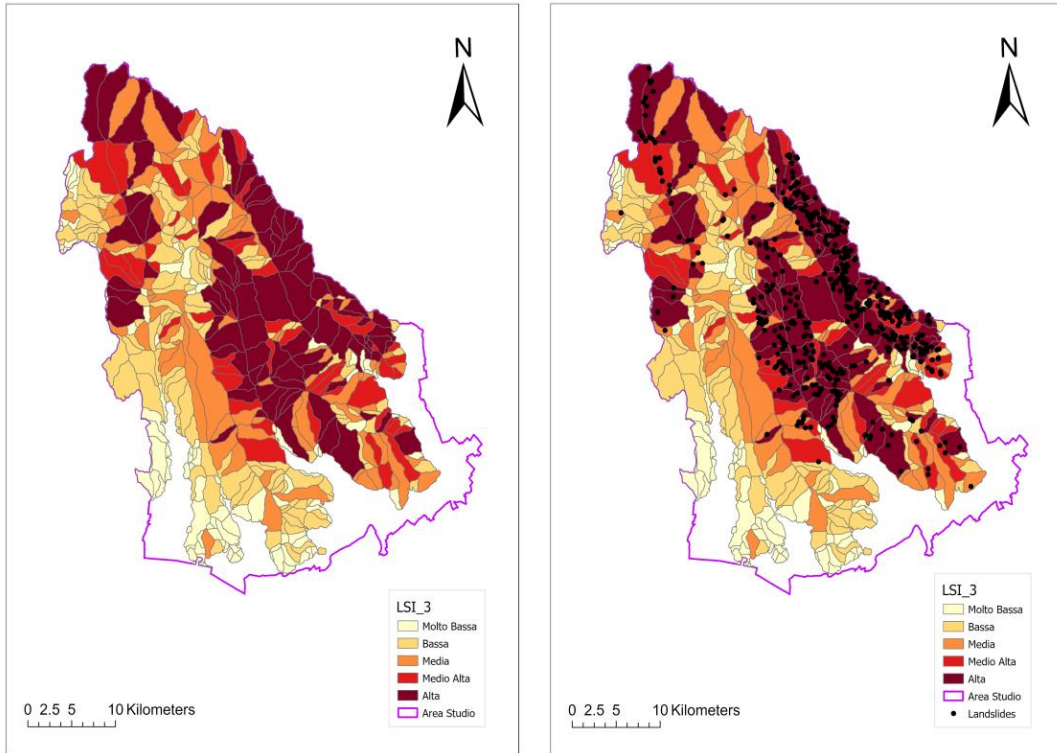


Figura 46: A sinistra è rappresentata la mappa di suscettibilità da frana media relativa alla terza configurazione, mentre a destra è mostrata la stessa mappa con la sovrapposizione dei punti di frana ricavati dall'Inventario dei Fenomeni Franosi Italiani (IFFI).

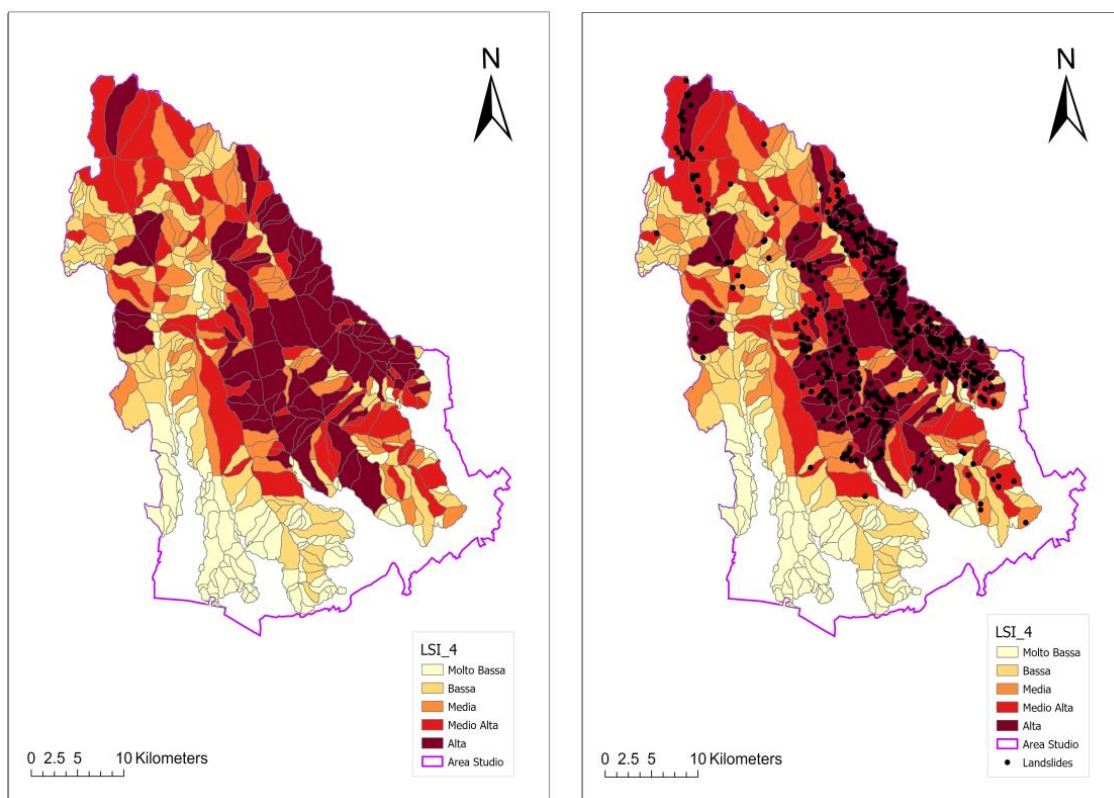


Figura 47: A sinistra è rappresentata la mappa di suscettibilità da frana media relativa alla quarta configurazione, mentre a destra è mostrata la stessa mappa con la sovrapposizione dei punti di frana ricavati dall'Inventario dei Fenomeni Franosi Italiani (IFFI).

Gli indici di suscettibilità da frana (LSI) calcolato dal modello sono stati suddivisi in 5 classi, come segue:

- $0 < LSI \leq 0,2 \rightarrow$ Suscettibilità da frana molto bassa
- $0,2 < LSI \leq 0.4 \rightarrow$ Suscettibilità da frana bassa
- $0,4 < LSI \leq 0.6 \rightarrow$ Suscettibilità da frana media
- $0,6 < LSI \leq 0.8 \rightarrow$ Suscettibilità da frana medio alta
- $0,8 < LSI \leq 1 \rightarrow$ Suscettibilità da frana alta

Si può osservare come, in tutte le configurazioni, la maggior parte dei punti in frana cadono all'interno delle SU classificate ad alta suscettibilità. Conseguentemente le SU caratterizzate in questo modo hanno la stessa collocazione geografica in tutte e quattro le mappe. Un discorso analogo si può fare anche per le aree con suscettibilità molto bassa: generalmente non contengono punti in frana e sono più o meno le stesse in tutte le configurazioni. È evidente che si riscontrano delle variazioni e discrepanze

tra le quattro mappe proposte e tra quelle presenti negli allegati, soprattutto nelle fasce di LSI non estreme. Le SU classificate con livelli di suscettibilità bassa, media e medio alta tendono a variare maggiormente tra le varie mappe generate. Nonostante questo, si può osservare un significato coerente e omogeneo dal punto di vista interpretativo.

Nelle (Figura 48) proposta in seguito, si può osservare il numero di Slope Units presenti in ogni classe di suscettibilità. Il conteggio è stato fatto per tutte le applicazioni delle quattro configurazioni.

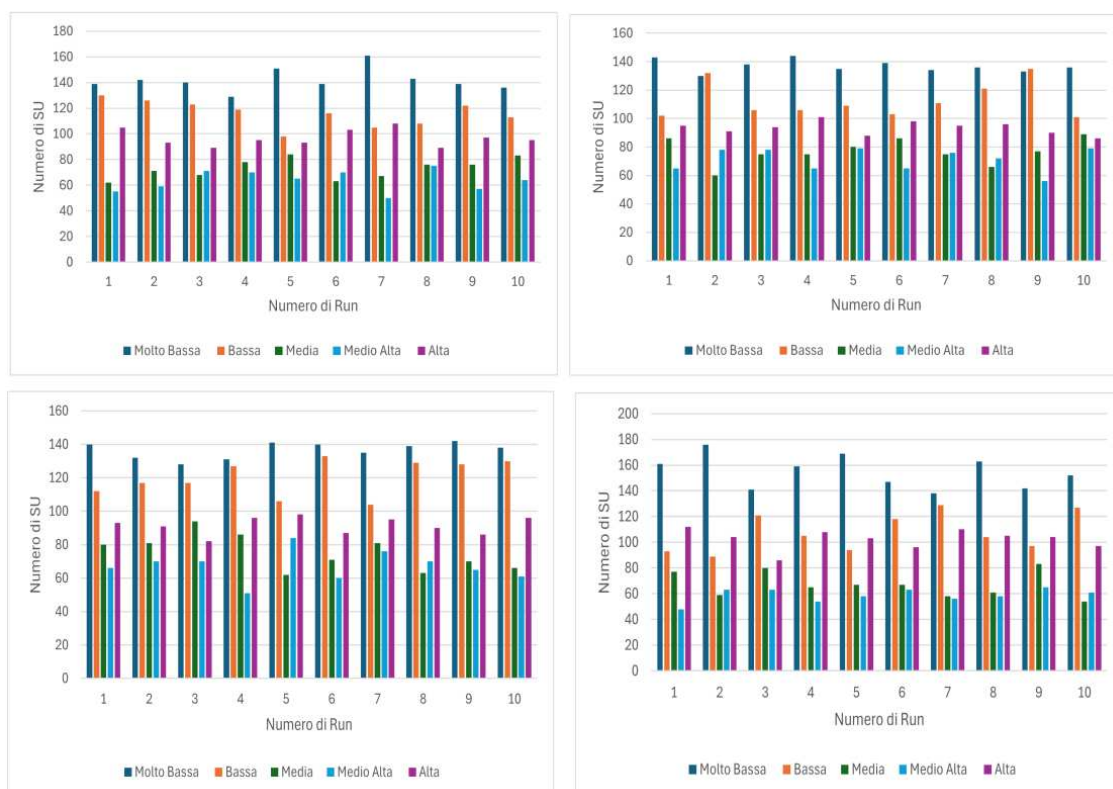


Figura 48: Mostra il numero di SU in ogni classe del LSI per ogni interazione del modello nelle 4 configurazioni. In alto a sinistra si può osservare il grafico relativo alla prima configurazione; in alto a destra, invece, si osserva il grafico della seconda configurazione; in basso a sinistra il grafico della terza e, in basso a destra, il grafico quella quarta configurazione. Ogni istogramma mostra la distribuzione delle SU per i 10 i run del modello.

Nella tabella seguente (Tabella 10) è riassunto il numero medio di SU appartenenti ad ogni classe, per le 4 configurazioni:

	I configurazione	II configurazione	III configuraione	IV configurazione	Media
LSI molto basso	142	137	137	155	142.75
LSI basso	116	112	120	108	114
LSI medio	73	77	76	67	73.25
LSI medio alto	63	71	67	59	65
LSI alto	97	94	91	102	96

Tabella 10: Valori medi del numero di SU contenute in ogni classe di LSI, per tutte le configurazioni

Come si diceva in precedenza, si osserva un range di variabilità minimo tra il numero di Slope Units appartenenti alla medesima classe di suscettibilità, per le diverse combinazioni. Si nota che il numero maggiore di Slope Units appartengono alle classi di suscettibilità bassa e molto bassa.

In seguito (Figura 49) è stato fatto un conteggio più specifico, considerando solo le SU già inizialmente catalogate come “in-frana”.

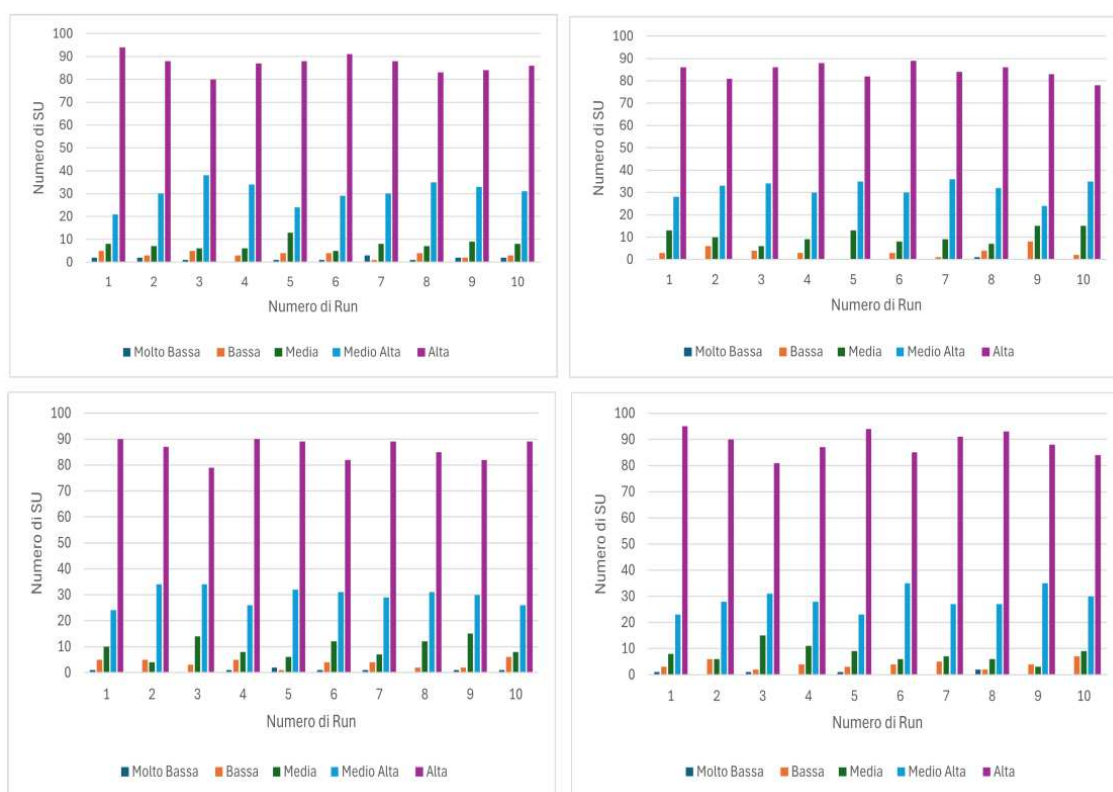


Figura 49: Mostra il numero di SU in frana presenti in ogni classe del LSI. In alto a sinistra si può osservare il grafico relativo alla prima configurazione; in alto a destra, invece, si osserva il grafico della seconda configurazione; in basso a sinistra il grafico della terza e, in basso a destra, il grafico quella quarta configurazione. Ogni istogramma mostra la distribuzione delle SU per i 10 i run del modello.

In tutti i grafici proposti, la maggior parte delle SU presenta un indice di suscettibilità alto, che risulta completamente in linea con le aspettative. Si osservano anche un

numero limitato di unità con suscettibilità medio-alta, mentre le altre classi sono rappresentate da un numero decisamente più ridotto di elementi. Come si può osservare in Tabella 11, mediamente solo una SU ha LSI molto basso e solo 4 hanno LSI basso. Questo andamento è opposto rispetto a quello descritto in precedenza, ma coerente con le aspettative. È anche un indicatore della buona performance del modello: la maggior parte delle SU assunte in frana sono state classificate con una suscettibilità alta o medio-alta, mentre un numero trascurabile di queste ha un grado di suscettibilità inferiore.

	I configurazione	II configurazione	III configuraizone	IV configurazione	Media
LSI molto basso	2	0	1	1	1
LSI basso	4	4	4	4	4
LSI medio	8	11	10	8	9.25
LSI medio alto	30	32	30	29	30.25
LSI alto	87	84	86	89	86.5

Tabella 11: Valori medi del numero di SU in frana contenute in ogni classe di LSI, per tutte le configurazioni

L’approccio adottato, si è quindi rivelato efficace nel prevedere la suscettibilità da frana, con risultati coerenti con le aspettative e indicatori solidi delle performance del modello. Questi risultati supportano l’uso di Random Forest come strumento affidabile per le valutazioni della pericolosità spaziale delle frane, fornendo un’importante base scientifica per interventi di mitigazione del rischio.

5 Discussioni

Per poter utilizzare un qualsiasi modello di data-driven, tra cui ricade il modello di Random Forest utilizzato in questo studio, è fondamentale selezionare in modo accurato i materiali di input (unità di base del terreno, catalogo dei fenomeni franosi e fattori predisponenti), in quanto essi influenzano in modo decisivo i risultati delle analisi. In questo lavoro le Slope Units sono state scelte come unità di base per suddividere il terreno in poligoni delimitati, con caratteristiche interne omogenee, e caratteristiche disomogenee rispetto a quelli adiacenti (Guzzetti et al., 1999). Nel

modello sono state utilizzate 491 SU, con dimensioni diversificate tra loro. Non essendo un numero molto elevato, l'algoritmo viene addestrato sulla base di un dataset piccolo, con una variabilità dei dati ridotta. Il risultato è un modello altamente specializzato, ma non applicabile globalmente. Tuttavia, essendo il RF nativamente poco affetto a problemi di overfitting, e non richiedendo assunzioni preliminari sulla distribuzione dei dati, risulta adatto all'utilizzo con campioni ridotti di dati.

Per avere un campione più ampio, il modello potrebbe essere testato variando le dimensioni delle Slope Units tramite il software `r.slopeunits`, al fine di verificare anche come la sua efficienza cambi in relazione a un numero diverso di SU. Alvioli et al. (2016) hanno evidenziato che, quando le Slope Units sono relativamente piccole, hanno un'omogeneità interna estremamente elevata, ma non rispettano completamente l'eterogeneità richiesta tra le unità adiacenti. Quando le SU sono grandi, hanno un comportamento opposto rispetto a quello descritto. È quindi fondamentale determinare le dimensioni e la quantità ottimale delle Slope Units per addestrare un modello altamente efficiente, capace di mantenere buone prestazioni anche in condizioni di applicazione diverse da quelle di training.

Oltre alle modifiche sulle SU, il modello potrebbe anche essere testato utilizzando unità di base diverse. Nonostante presentino a loro volta delle criticità, le celle di griglia (pixel) potrebbero essere una buona alternativa per suddivisione del territorio. Secondo Reichdnebach et al. (2018), sono le unità di mappatura utilizzate più comunemente nella letteratura scientifica: circa l'86 % degli studi è effettuato utilizzando i pixel. La loro popolarità è dovuta all'elevata compatibilità tra le celle di griglia e i modelli digitali di elevazione (DEM); inoltre hanno una risoluzione personalizzabile, che li rende adattabili alle diverse scale di analisi e alle risorse disponibili. Sono unità regolari e uniformi, quindi più facilmente manipolabili rispetto alle SU che, essendo irregolari e disomogenee, introducono complessità nei calcoli del modello.

Un altro limite del modello riguarda l'incertezza legata alle SU considerate non soggette a frana. Infatti, al loro interno potrebbe essere presenti frane sconosciute, non registrata nell'Inventario dei Fenomeni Franosi (IFFI), che è la fonte dei punti di

frana utilizzati in questa tesi di laurea. L'inventario in questione, per quanto in continuo aggiornamento, non è considerabile totalmente completo, in quanto non vi è certezza che al suo interno siano catalogati tutti gli eventi franosi presenti sul territorio nazionale. Per esempio, non contiene le frane più recenti o quelle collocate in aree remote, che non interessano in modo diretto infrastrutture e centri abitati, che, quindi, vengono individuate e segnalate meno frequentemente. Per ridurre questa incertezza l'inventario IFFI potrebbe essere integrato con dati appartenenti a cataloghi specialistici, redatti da Comuni o Provincie; con dati ottenuti da lavori di professionisti su dissesti specifici; oppure con dati interferometrici satellitari.

Analizzare dettagliatamente i fattori predisponenti che controllano la pericolosità spaziale delle frane è tra gli obiettivi principali di questa tesi di laurea. L'algoritmo Random Forest, attraverso il calcolo dell'errore di classificazione dell'Out-Of-Bag, ha permesso di stabilire l'importanza delle variabili sulle capacità previsionali del modello, determinando così la loro influenza sul verificarsi delle frane. I risultati prodotti da questa analisi sono una serie di istogrammi (tre per l'analisi preliminare e quattro per quella finale), che mostrano graficamente i livelli di incidenza delle singole variabili selezionate per ogni configurazione.

I quattro predittori che si confermano essere i più influenti per tutte le configurazioni proposte sono: le piogge medie, la deviazione standard della quota, la litologia e l'MSR.

La media delle piogge annuali, registrate nell'area studio, descrive il volume totale delle precipitazioni che avvengono mediamente in quel determinato periodo di tempo. È un indicatore del bilancio idrico medio e del potenziale di accumulo di acqua nel sottosuolo. La deviazione standard delle piogge, invece, descrive come queste variano rispetto alla media, individuando gli eventi frequenti, sia in termini di picchi di piovosità, sia in termini di periodi secchi. Il modello considera la media come parametro più significativo, rispetto alla deviazione standard; questo risultato indica che, nell'area di studio, è più probabile che i fenomeni gravitativi si inneschino in seguito alla saturazione del terreno determinata da un accumulo graduale dell'acqua nel suolo, piuttosto che in seguito a eventi estremi e isolati di precipitazioni. Nel primo caso la saturazione prolungata riduce la resistenza al taglio del suolo, favorendo condizioni di precarietà, in cui anche un evento piovoso, di

intensità media, potrebbe essere decisivo per l'innescò della frana. Nel secondo caso i terreni non sono saturi, quindi, affinché si generi un movimento del versante, in seguito al raggiungimento delle condizioni critiche, è necessario un elevato apporto di acqua in tempi brevi.

La deviazione standard della quota, che è un altro predittore molto importante secondo il modello, misura le variazioni di altimetria delle varie SU. Questo fattore può essere correlato con la pendenza media, infatti pendii caratterizzati da un'elevata variabilità della quota sono tendenzialmente più ripidi, quindi più vulnerabili a frane; mentre versanti con bassa deviazione standard, hanno generalmente una pendenza minore, che può essere associata a una suscettibilità da frana più bassa.

L'elevata importanza della Litologia e del MSR nel calcolo della suscettibilità da frana conferma le osservazioni effettuate inizialmente. Infatti, si era ipotizzato che ci fosse un forte controllo della litologia sul verificarsi delle frane, poiché la maggior parte dei punti di frana, individuati nell'Inventario dei Fenomeni Franosi in Italia (IFFI), si trovano in aree caratterizzate dalla stessa litologia (vulcaniti). Per questa ragione la litologia vulcanica è stata indicata come la più franosa nella classifica del Material Strangth Ranking. Pertanto, l'osservazione iniziale sulla densità di frane, trova un riscontro concreto nelle capacità previsionali del modello.

I risultati di importanza delle variabili ottenuti con gli istogrammi sono stati specificati e analizzati nel dettaglio grazie ai Partial Dipendence Plots (PDPs), calcolati solo per la quarta configurazione, nell'analisi finale. Come spiegato in precedenza, essi permettono di visualizzare graficamente come i valori di ogni singola variabile influenzano la suscettibilità da frana. Sono dunque uno strumento molto potente, che rende più accessibile e comprensibile il comportamento del modello.

I risultati ottenuti dell'analisi dei PDPs si sono rivelati coerenti con l'importanza dei fattori predisponenti ricavati dall'istogramma dell'OOB e descritti in precedenza. In particolare, i PDPs delle piogge medie e della deviazione standard della quota, mostrano un'iniziale crescita dello score: nel primo caso continua fino a 1500 mm/anno, mentre nel secondo si osserva fino a 200 m. Entrambe le curve si stabilizzano oltre queste soglie, indicando che la loro influenza sulla suscettibilità da frana non cresce ulteriormente in modo significativo. L'iniziale picco dello score

suggerisce, quindi, l'aumento repentino dell'importanza associata all'aumento delle precipitazioni medie e della variabilità della quota.

Nel PDP dell'MRS, la litologia più franosa, le vulcaniti ($MSR = 7$), mostra uno score elevato, indicando un impatto significativo sul modello. Lo score diminuisce man mano che si scende lungo la scala della franosità delle litologie. Tuttavia, si osserva un aumento dello score per litologie meno franose, che, pur essendo stabili, influenzano il modello nella previsione di "non frana", aiutando a distinguere le aree non a rischio.

Per analizzare la qualità dei risultati prodotti da Random Forest, attraverso l'identificazione degli errori di classificazione commessi, sono state calcolate la matrice di confusione, la curva ROC e il valore dell'area sotto alla curva (AUC), per tutte le applicazioni di ogni configurazione, della fase preliminare e finale, sia per i dati di training che di test.

I risultati ottenuti per i dati di addestramento indicano un modello che ha lavorato in modo perfetto: ha giustamente predetto zero FP e zero FN, la curva ROC indica un TPR pari a 1 e un FPR pari a zero e, per finire, il valore di AUC è 1.

I risultati ottenuti per i dati della fase di test evidenziano un'elevata capacità predittiva del modello Random Forest, come evidenziano i valori medi di AUC, che si aggirano attorno a 0,82.

Il confronto tra le diverse configurazioni ha permesso di identificare l'opzione più adatta all'analisi, riducendo il margine di incertezza che si sarebbe creato concentrandosi su un singolo modello specifico; in questo modo è migliorata la conoscenza delle singole variabili, per come influenzano il modello. La configurazione che è risultata essere più efficiente, in termini di performance del modello, si è rivelata essere quella composta dagli 11 parametri più influenti (quarta configurazione dell'analisi finale), che sono stati selezionati partendo dell'istogramma in Figura 26. Mostra un'accuratezza circa dell'87% e un AUC medio di 0,85.

Eliminando le variabili meno influenti si riduce la complessità del modello, senza comprometterne l'efficienza. Quindi si può produrre una mappa di suscettibilità credibile utilizzando soltanto i fattori condizionanti più significativi; in questo modo si risparmiano tempo e risorse, mantenendo una buona precisione.

Per questa tesi di laurea sono comunque state prodotte delle LSM per ognuna delle quattro configurazioni indagate. I risultati ottenuti, anche in questo caso, sono risultati coerenti con le aspettative, che supportano l'affidabilità di Random Forest come strumento efficace per la valutazione della suscettibilità, fornendo una solida base scientifica per pianificare interventi di mitigazione del rischio e gestione del territorio.

Il modello si è rivelato molto efficiente e performante per questa specifica area studio, ma, nella configurazione attuale, la sua generalizzabilità e applicabilità ad altri contesti territoriali non è immediata, ma necessita di un approfondimento specifico, basato sul nuovo contesto territoriale in analisi. L'algoritmo di machine learning viene addestrato utilizzando una serie di dati di training, selezionati sulla base delle condizioni ambientali della zona; quindi, se il modello dovesse essere applicato in un contesto nuovo e diverso, in cui sono presenti dei valori non rappresentati nel database di training, potrebbe non essere in grado di interpretarli correttamente, generando dei risultati non attendibili o, nel peggiore dei casi, generare un messaggio di errore.

6 Conclusione

Questo lavoro di tesi è stato sviluppato nella zona dei Monti Lessini, che comprende un'area molto estesa (circa 476 km^2) tra le province di Verona e Vicenza. Si tratta di un territorio che presenta una grande variabilità dal punto di vista paesaggistico e litologico, con zone in cui si registrano numerosi fenomeni franosi. Buona parte di questi dissesti sono stati identificati e catalogati nell'Inventario dei Fenomeni Franosi in Italia (IFFI). Al fine di fornire una base utile per la gestione del territorio e la previsione del rischio geologico sono state elaborate quattro mappe di suscettibilità da frana. Per raggiungere questo obiettivo è stata utilizzata l'intelligenza artificiale, nello specifico l'algoritmo Random Forest della tecnica del machine learning. Questo modello si basa su un processo di apprendimento automatico, in cui il sistema si addestra autonomamente a partire dai dati forniti, migliorando le proprie prestazioni

attraverso l'esperienza.

L'applicazione del modello RF per la generazione della LSM è subordinata all'utilizzo di unità di base che suddividono la mappa in porzioni di territorio con caratteristiche omogenee all'interno, ma eterogenee rispetto alle unità adiacenti (Guzzetti et al., 1999). In letteratura esistono unità di base diverse (Reichenbach et al., 2018), ma per questa tesi si è scelto di utilizzare le Slope Units (SU), ovvero delle unità di terreno che vengono delimitate da linee di drenaggio e spartiacque, utilizzando esclusivamente le informazioni ottenute da un modello digitale di elevazione (DEM) (Alvioli et al., 2016).

Siccome il verificarsi dei processi gravitativi è strettamente correlato con una serie di fattori che agiscono sul versante, per poter fare una previsione spaziale delle frane è necessario definire una serie di fattori predisponenti. Per questo specifico caso ne sono stati utilizzati 17: la quota, la pendenza del versante (slope angle), l'orientazione del versante (aspect), la curvatura totale, la curvatura di profilo e planare, la distribuzione delle piogge, la densità di drenaggio, la distanza dai fiumi, il Topographic Wetness Index (TWI), il Topographic Position Index (TPI), il Normal Difference Vegetation Index (NDVI), la distanza dalle strade, il Topographic Roughness Index (TRI), l'uso del suolo (land use), la litologia e il Material Strength Ranking (MSR). L'obiettivo di questa tesi di laurea, oltre all'elaborazione della LSM, è anche quello di individuare le variabili più influenti, che contribuiscono maggiormente all'instabilità dei versanti.

A ciascuna SU è associato un valore per ogni variabile considerata. Su questi valori sono stati eseguiti calcoli statistici, in particolare la media e la deviazione standard per le variabili continue, e la majority per le variabili categoriche, che vengono utilizzati per costruire la tabella dei dati di input, successivamente importata in MATLAB e utilizzata dall' algoritmo RF. L'altro parametro molto importante, che fa parte di questa tabella, è l'indice che identifica quali SU sono "in frana", contrassegnate con 1, e quali sono "non in frana", contrassegnate con 0.

L'analisi dei dati con Random Forest è stata eseguita in due step successivi: per l'analisi preliminare sono state considerate 3 diverse configurazioni di parametri, per ognuna il modello è stato applicato 5 volte, con la creazione di 2000 alberi per ciascuna esecuzione. L'analisi finale si è svolta considerando le stesse 3

configurazioni del caso precedente, con l'aggiunta di una quarta realizzata considerando solo i fattori predisponenti con maggiore importanza, selezionati partendo dall'istogramma dell'OOB in Figura 26. In questo caso il modello è stato applicato 10 volte, con la creazione di 500 alberi per ciascuna esecuzione; questo valore è stato scelto in base all'errore di classificazione dell'Out-Of-Bag osservato nella fase preliminare, che si stabilizza attorno a questo numero di alberi. Per tutte le configurazioni, sia della fase preliminare che della fase finale, i risultati prodotti dal modello sono: il grafico dell'errore di classificazione dell'Out-of-Bag e il grafico di importanza delle variabili, stabilito a partire dal OOB error. I fattori predisponenti più significativi sono risultati essere: la deviazione standard della quota, le precipitazioni medie e la majority del MSR. Un'analisi più specifica, relativamente all'importanza delle variabili, è stata fatta, solo nella fase finale, tramite l'utilizzo dei Partial Dipendenze Plots (PDPs), i quali permettono di identificare i valori che influenzano maggiormente la suscettibilità da frana, per ogni singola variabile.

Per poter valutare la qualità dei risultati del modello, in modo da definire la configurazione migliore, vengono calcolati anche la matrice di confusione, la curva ROC e il relativo valore di AUC, sia per i dati di training che per i dati di test. Con un'accuratezza dell'87% e un valore di AUC medio pari a 0,85, la configurazione più efficace risulta essere la quarta. Nonostante ciò, tutti i risultati ottenuti rivelano un modello performante ed efficiente per questa area studio.

Sono state prodotte quattro mappe di suscettibilità, una per ogni configurazione, ottenute mediando i valori degli indici di suscettibilità da frana (LSI) derivanti dalle dieci simulazioni effettuate per ogni configurazione dell'analisi finale. I risultati ottenuti confermano l'affidabilità del modello di Random Forest come strumento efficace per la valutazione della pericolosità spaziale da frana.

Un potenziale sviluppo futuro del presente lavoro potrebbe essere il passaggio da un'analisi "statica" della suscettibilità ad una "dinamica", in cui le condizioni meteorologiche e climatiche future vengono considerate per il potenziale innesco di frane.

"Il riscaldamento climatico è inequivocabile" (IPCC, 2023) e sono evidenti i suoi effetti sul manifestarsi di eventi meteorologici estremi che influenzano

significativamente la stabilità dei versanti, anche se queste dinamiche risultano complesse e incerte (Gariano & Guzzetti, 2016). Benché questa dipendenza sia evidente, gli autori sottolineano alcuni limiti nelle modellazioni frana-clima, dovuti alla mancata considerazione della variabilità non stazionaria del clima e dalla dipendenza temporale della suscettibilità da frana. Evidenziano inoltre che i dettagli delle previsioni sono relazionati alla gravità degli scenari di emissione, elaborati nel sesto rapporto di valutazione sui cambiamenti climatici dell'Intergovernmental Panel on Climate Change (IPCC). È stato pubblicato nel 2023 e sintetizza i risultati dei tre gruppi di lavoro: le basi fisico-scientifiche (2021), in cui sono contenuti i possibili scenari, catalogati in base al livello di emissioni; impatti, adattamento e vulnerabilità (2022) e mitigazione dei cambiamenti climatici (2022).

Alla luce di queste considerazioni, il modello di Random Forest sviluppato in questa tesi di laurea potrebbe essere applicato per valutare la suscettibilità da frana, integrando variabili predisponenti correlate a eventi meteo-climatici futuri derivati da scenari di cambiamento climatico. Si potrebbero utilizzare i dati di temperatura e precipitazione proposti per ogni scenario descritto nell' IPCC (2023).

7 Bibliografia

ALVIOLI M., GUIZZETTI F., MARCHESINI I., (2020) – *Parameter-free delineation of slope units and terrain subdivision of Italy*. *Geomorphology* 358 (2020), 107124

ALVIOLI M., MARCHESINI M., REICHENBACH P., ROSSI M., ARDIZZONE F., FIORUCCI F., GUZZETTI F., (2016) – *Automatic delineation of geomorphological slope units with r.slopeunits v1.0 and their optimization for landslide susceptibility modeling*, *Geoscientific Model Development*, 9, 3975-3991

ANTONELLI R., BARBIERI G., DAL PIAZ G.V., DAL PRA A., DE ZANCHE V., GRANDESSO P., MIETTO P., SEDEA R., ZANFERRARI A., (1990) - *Carta geologica del Veneto. Scala 1:250.000. Una storia di cinquecento milioni di anni*, Regione del Veneto, Venezia.

- BAGHBANI A., CHOUDHURY T., COSTA S., REINER J., (2022) - *Application of artificial intelligence in geotechnical engineering: a state-of-the-art review*, Earth-Science Review, 228, 103991
- BAI S., LU P., THIEBES B., (2020) – *Comparing characteristics of rainfall- and earthquake- triggered landslides in the Upper Minjiang catchment, China*, Eng. Geol., 268, article 105518
- BARBIERI G., DAL PRÀ A., DE ZANCHE V., SACCARDI P., SEDEA R., ZANFERRARI A., (1973) – *Studio geologico sulla franosità nel bacino del torrente Chiampo (Prealpi venete)*, Memorie degli Istituti di Geologia e Mineralogia dell'università di Padova, Vol. XXIX
- BOSELLINI A., (1989) - *Dynamics of Tethyan carbonate platform. Controls on Carbonate Platform and Basin Development*, Crevello P.D, James L.W, Sarg J.F, Read J.F, Soc. Econ. Paleont. Miner., Spec. Publ. 44: 3-13.
- BOSELLINI A., CARRARO F., CORSI M., DE VECCHI G.P., GATTO G.O., MALARODO R., STURANI C., UNGARO S., ZANETTIN B., (1967) – *Note illustrative della carta geologica d'Italia alla scala 1:100.000 Foglio 49 Verona 61*, Nuova Tecnica Grafica, Roma
- BOSELLINI A., HARDIE L.A., (1973) - *Depositional theme of a marginal marine evaporite*. Sedimentology, v.20, 5-27.
- BOSELLINI A., HSU K.J., (1973) - *Mediterranean Plate Tectonics and Triassic Palaeogeography*. Nature, v. 249, n.5412, 144-146.
- BREIMAN L., (2001) – *Random Forest*, Machine Learning, 45, 5-32
- BREIMAN L., LAST M., RICE J., (2003) - *Random Forest: Finding Quasars*, Statistical Challenges in Astronomy, Springer-Verlag, pp-243-254
- CALLIGARIS C., (2016) – *Algoritmi geomorfometrici*, D.M.G Università degli Studi di Trieste

CANTELLI L. & CASTELLARIN A., (1994) - *Analisi e inquadramento strutturale del sistema "Schio-Vicenza"*. Atti Tic. Sc. Della Terra, Serie Speciale, 1: 231-245, Pavia.

CANTON A. & CASTALDINI D., (1985) – *Approfondimento di morfoneotettonica tra il Lago di Garda ed il Torrente Alpone (Provincia di Verona)*. Boll. Mus. Civ. St. nat. Verona, 12, 1985, pp. 461-491

CARRARA A., (1988) – *Drainage and divide networks derived from high-fidelity digital terrain models*, Quantitative Analysis of Mineral and Energy Resources, Vol 223, pp. 581-597

CASATI P., (a cura di) (1996) – *Scienze della Terra: elementi di Geologia Generale, Volume 1*. CittàStudiEdizioni, 3° edizione, pgg 391-426. Torino.

CASTELLARIN A. & VAI G.B., (1982) – *Guida alla geologia del Sudalpino centro orientale*. Guide Geol. Reg. S.G.I., pp. 1-22.

CATANI F., LAGOMARSINO D., SEGONI S., TOFANI V., (2013) – *Landslide susceptibility estimation by random forests techniques: sensitivity and scaling issues*, Natural Hazards Earth Syst. Sci., 13, 2815-2831

CHEN X. & CHEN W., (2021) – *Gis-based landslide susceptibility assessment using optimized hybrid machine learning methods*, Catena, 196, 104833

COLLESELLI F., DAL PRÀ A., PREVIATELLO P., SEDEA R., (1976) – *Indagine geologica e geotecnica di un'ampia area di frana nella valle del T. Chiampo*. Rig 1976, numero 4, pp. 206

COSTANTINI M., MINATI F., TRILLO F., FERRETTI A., NOVALI F., PASSERA E., DEHLS J., LARSEN Y., MARINKOVIC P., EINEDER M., BRCIC R., SIEGMUND R., KOTZERKE P., PROBECK M., KEYERES A., PROIETTI S., SOLARI L., ANDERSON H.S., (2021) – *European Ground Motion Service (EGMS)*. International Geoscience and Remote Sensing Symposium (IGARSS), 3293-3296.

CRUDEN, D. M., (1991). *A Simple Definition of a Landslide*. Bulletin of the International Association of Engineering Geology, 43, 27-29.

- DE REU J., BOURGEOIS J., BATS M., ZWERTVAEGHER A., GELORINI V., DE SMEDT P., CHU W., ANTROP M., DE MAEYER P., FINKE P., VAN MEIRVENNE M., VERNIERS J., CROMBE P., (2013) – *Application of the topographic position index to heterogeneous landscape*, *Geomorphology*, Vol 186, pp. 39-49
- DE VECCHI G.P., GREGNANIN A., PICCIRILLO E.M., (1976) – *Aspetti petrogenetici del vulcanesimo terziario veneto*. Mem. Ist. Geol. Mineral. Univ. di Padova, 30, pp. 1-33
- DOGLIONI C., BOSELLINI A., (1997) - *Eoalpine and mesoalpine tectonics in the Southern Alps*, *Geol. Rundsch.*, 76, 735-754, 1987.
- FAILLA O., SCIENZA A., FIORINI P., MINELLI R., (1998) - *La zonazione della valle d'Illasi*. IVES conference Series, Terroir 1998
- FRATTINI P., CROSTA G.B., (2006) - *Valutazione dell'accettabilità del rischio da frana e analisi costi-benefici*, *Giornale di geologia applicata* 4 (2006), pp. 49-56
- GARIANO S.L., GUIZZETTI F., (2016) – *Landslides in a changing climate* Earth-Science reviews, Vol. 162, November 2016, pp. 227-252
- GIUSBERTI L., ROGHI G., ZORZIN R., (2020) – *Geologia dei Monti Lessini Veronesi*. – Le terre gialle e rosse della Lessinia, Cavallo G. e Museo di Storia Naturale, pp. 32-47
- GUZZETTI F., REINCHENBACH P., ARDIZZONE F., CARDINALI M., GALLI M., (2006) – *Estimating the quality of landslide susceptibility models*. *Geomorphology*, Vol 81, pp 166-184
- GUZZETTI F., CARRARA AL., CARDINALI M., REICHENBACH P., (1999) - *Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy*, *Geomorphology*, Vol. 31, Issues 1-4-, pp. 181-216
- HANSEN A., (1984) – *Landslide hazard analysis*, D. Brunssden, D.B. Prior (Eds.), *Slope Instability*, Wiley, pp. 523-602

HE R., ZHANG W., DOU J., JIANG N., XIAO H., ZHOU J., (2024) – *Application of artificial intelligence in three aspects of landslide risk assessment: A comprehensive review*, Rock Mechanics Bulletin, Vol. 3, Issue 4, 100144

HORTON R.E., (1945) – *Erosional development of the streams and their drainage basins; hydrophysical approach to quantitative morphology*, Geol. Soc. Am. Bull., 56, pp. 275-370

HUNTER K., SVEDSEN B., RICKNMANN D., (1994) – *Debris flow modeling: A review*, Continuum Mech. Thermodyn 8,1-35 (1994)

IADANZA C., TRIGLIA A., STACE P., DRAGONI A., BIONDO T., ROCCISANO M., (2021) - *IdroGEO: A collaborative web mapping application based on REST API services and open data on landslides and floods in Italy*. ISPRS International Journal of Geo-Information, 10(2).

IPCC, (2023) - *Summary for Policymakers. In: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero(eds.)]*.

JACOBS L., KERVYN M., REINCHENBACH P., ROSSI M., MARCHESINI I., ALVIOLI M., DEWITTE O., (2020) – *Regional susceptibility assessment with heterogeneous landslide information: Slope Unit – vs. pixel based approach*, Geomorphology, Vol. 356, 107084.

KOPECKY M., MACEK M., WILD J., (2021) – *Topographic Wetness Index calculation guidelines based on measured soil moisture and plant species composition*, Science of Total Environment, Vol 757, 143785

LARY D.J., ALAVI A.H., GANDOMI A.H, WALKER A.L., (2016) - *Machine learning in geoscience and remote sensing*, Geoscience Frontiers, Vol. 7, Issue 1, pp. 3-10

LECUN Y., BENGIO Y., HINTON G., (2015) – *Deep Learning*, Nature, 521, 436-444

- LEE M.L., NG K.Y., HUANG Y.F., LI W.C., (2014) – *Rainfall-induced landslides in Hulu Kelang area, Malaysia*, Nat. Hazards, 70, pp. 353-375
- MALAPERDAS G.M. & PANAGIOTIS V.V., (2018) – *The aspect of Aspect: Understanding land exposure and its part in geographic information systems analysis*, Energy & Environment 0(0), 1-16.
- MARTINELLO C., CAPPADONIA C., CONOSCENTI C., AGNESI V, ROTIGLIANO E., (2021) – *Optimal slope units partitioning in landslide susceptibility mapping*, Journal of maps, 17(3), pp. 152-162
- MASETTI D., FANTONI R., ROMANO R., SARTORIO D., TREVISANI E., (2012) – *Tectonostratigraphic evolution of the Jurassic extensional basins of eastern Southern Alps and Adriatic foreland based on an integrated study of surface and subsurface data*. AAPG Bulletin, v. 96, NO. 11, PP. 2065-2089
- MEENA S.R., PULIERO S., BHUYAN K., FLORIS M., CATANI F., (2022) – *Assessing the importance of conditioning factor selection in landslide susceptibility for the province of Belluno (region of Veneto, northeastern Italy)*, Natural Hazards and earth System Science, 22(4), 1395-1417
- MERGHADI A., YUNUS A.P., DOU J., WHITELEY J., THAIPHAM B., BUI D.T., AVTAR R., ABDERRAHMANE B., (2020) – *Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance*, Earth-Science Reviews, 207, 103225
- MINGYONG L., HAIJAIA W., YANG L., (2022) – *Identifying the essential conditioning factors of landslide susceptibility models under different grid resolutions using hybrid machine learning: A case of Wushan and Wuxi countries, China*, Catena, Vol 217, 106428
- MULLER W., PROSSER G., MANCKTELOW N., VILLA I.M., KELLEY P.S., VIOLA G., OBERLI F., (2001) - *Geochronological constraints on the evolution of the 102 Periadriatic fault system (Alps)*. International Journal of Earth Sciences 90 (3), 623-653.

- NASTESKI V., (2017) - *An overview of the supervised machine learning methods*, Faculty of Information and Communication Technologies, Partizanska
- NOCENTINI N., ROSI A., SEGONI S., FANTI R., (2023) – *Towards landslide space-time forecasting through machine learning: the influence of rainfall parameters and model setting*, *Frontiers in Earth Science*, 11, 1-20
- O'LOUGHLIN E.M., (1986) – *Prediction of surface saturation zones in natural catchments by topographic analysis*, *Water Resources Research*, 22(5), pp. 794-804
- P.A.T.I dell'Alpone (2011) – *Valutazione ambientale strategica*. Regione Veneto Provincia di Verona
- PETLEY D., (2012) – *Global patterns of loss of life from landslides*, *Geology* (2012), 40(10), 927-930
- PICCOLI G., (1966) – *Subaqueous and subaerial basic volcano eruption in the Paleogens of the Lessini Alps (Southern Alps, NE-Italy)* *Bulletin of Volcanology*, v.29, 253-266
- REICHENBACH P., ROSSI M., MALAMUD B.D., MIHIR M., GUIZZETTI F., (2018) - *A review of statistically-based landslide susceptibility models*. *Earth-Science Reviews*, 180 (November 2017), 60–91.
- RILEY S.J, DEGLORIA S.D., ELLIOT R., (1999) – *Index that quantifies topographic heterogeneity*, *Intermountain Journal of Science*, 5, 23-27.
- SALGATO S., (2019) - *La Val d'Alpone: un territorio alla ricerca della propria identità, tra utopia e realtà*. *Territori Spezzati – Spopolamento e abbandono nelle aree interne dell'Italia contemporanea* a cura di Giancarlo Macchi Jánica e Alessandro Palumbo, CISGE – Centro Italiano per gli Studi Storico-Geografici, Roma 2019
- SAURO U., (1973) – *Il paesaggio degli Alti Lessini. Studio geomorfologico*. *Museo Civico di storia naturale di Verona*. Memorie fuori serie, numero 6.
- SAURO U., (2010) – *Lessinea*. Verona, Cierre edizioni.

- SCHICKER R., MOON V., (2012) – *Comparison of bivariate and multivariate statistical approaches in landslide susceptibility mapping at a regional scale*, *Geomorphology*, Vol. 161-162, pp. 40-57
- SEGONI S., PAPPAFICO G., LUTI T., CATANI F., (2020) - *Landslide susceptibility assessment in complex geological settings: Sensitivity to geological information and insights on its parameterization*, *Landslides*, 17, 2443-2453
- SHANO L., RAGHVANSHI T.K., METEN M., (2020) - *Landslide susceptibility evaluation and hazard zonation techniques*, *Geoenviron Disaster* 7, 18 (2020).
- THIERY Y., MALET J.P., STERLACCHINI S., PUISSANT A., MAQUAIRE O., (2007) – *Landslide susceptibility assessment by bivariate methods at large scales: application to a complex mountainous environment*, *Geomorphology*, 92 (2007), pp. 38-59
- TOFANI V., BIOCCHI G., ROSSI G., SEGONI S., D’AMBROSIO M., CASAGLI N., CATANI F., (2017) – *Soil characterization for shallow landslides modeling: A case study in the northern Apennines (Central Italy)*, *Landslide* 14, 755-770
- TROMBETTA G., (2010) – *Anatomia comparata fra due sistemi di piattaforma carbonatica del Triassico del Sudalpino: il Norico del Bacino Lombardo (Dosso delle Saette/Magasa, Prealpi Bresciane) ed il Ladinico superiore delle Dolomiti occidentali (Sciliar – Alpe di Siusi)* Ann. Mus. Civ. St. nat. Ferrara, Vol. 13 pp. 3-31
- UBYDUL H., BLUM P., DA SILVIA P.F., ANDERSEN P., PILZ J., CHALOV S.C., MALET J.-P., AUFLIČ M.J., ANDRES N., POYIADJI E., LAMAS P.C., ZHANG W., PESHEVSKI I., PETURSSON H.G., KURT T., DOBREV N., GARCIA-DAVALILLO J.C., HALKIA M., FERRI S., GAPRINDASHVILI G., ENGSTRÖM J., KEELLINGS D., (2016) – *Fatal landslides in Europe*, *Landslides* 13(6), 1545-1554
- VARNES D.J., (1984), - *Landslide Hazard Zonation: a Review of Principles and Practice*, Unesco Press, Parigi, pp. 63
- VIETLONG D., BA-QUANG-VINH N., HUNG T. P., CHI CONG N., CUONG TIEN N., (2023) – *Effect of time-variant NDVI on landslide susceptibility: A case*

study in Quang Ngai province, Vietnam, Open Geoscience, vol. 15, no. 1, pp. 20220550

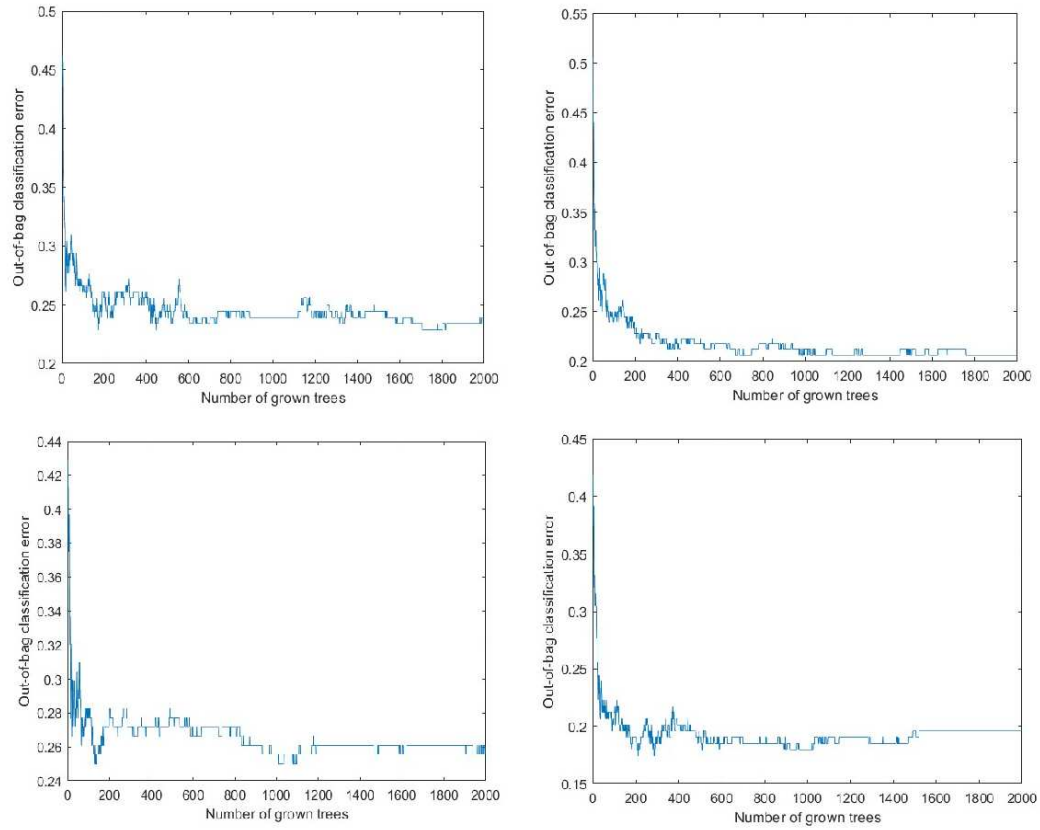
YALCIN A., REIS S., AYDINOGLU A.C., YOMRALIOGLU T., (2011) – *A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistic regression methods for landslide susceptibility mapping in Trabzon, NE Turkey*, Catena, Vol. 85, 3, pp. 274-287

ZAMPIERI D., (1995) – *Tertiary extension in the southern Trento Platform, Southern Alps, Italy*. Tectonics, 14,3, pp. 645-657.

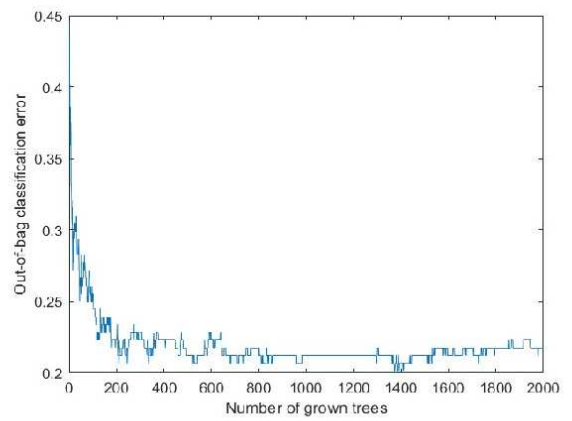
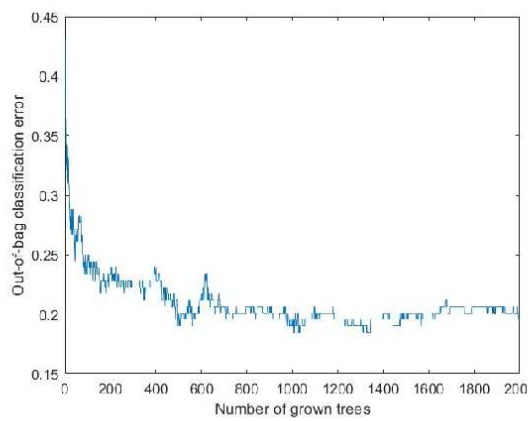
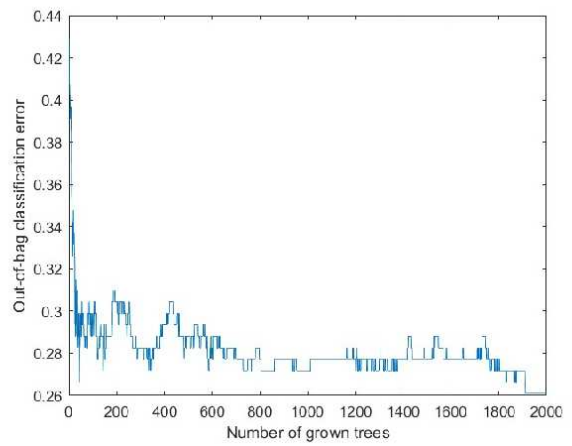
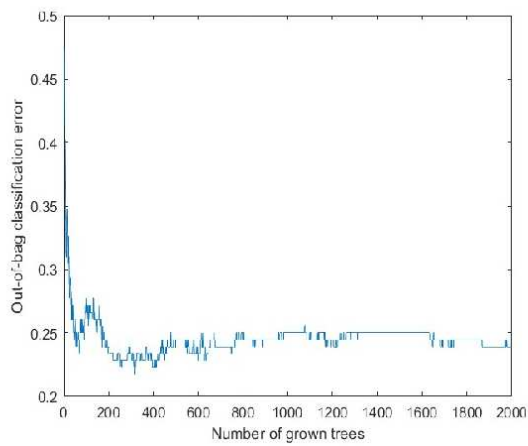
ZAMPIERI D., (2000) – *Segmentation and linkage of the Lessini Mountains normal faults, Southern Alps, Italy*. Tectonophysics 349, Elsevier, 19-31.

ZAMPIERI D., (2022) – *Modello strutturale della deformazione estensionale paleogenica nella ex-cava Main (Arzignano, Vicenza)* Studi e ricerche, Associazione Amici del Museo, Museo Civico “G. Zannato”, pp. 19-24.

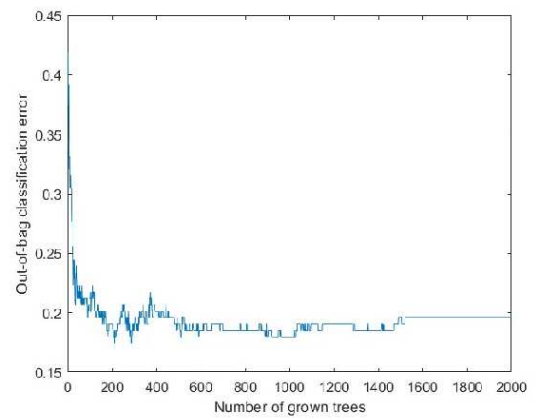
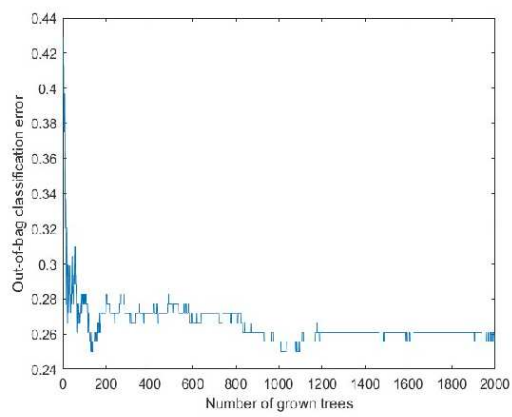
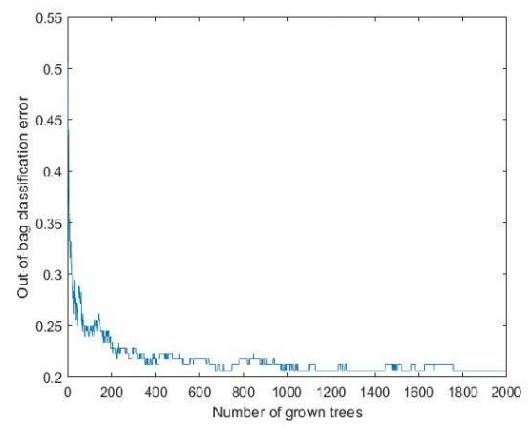
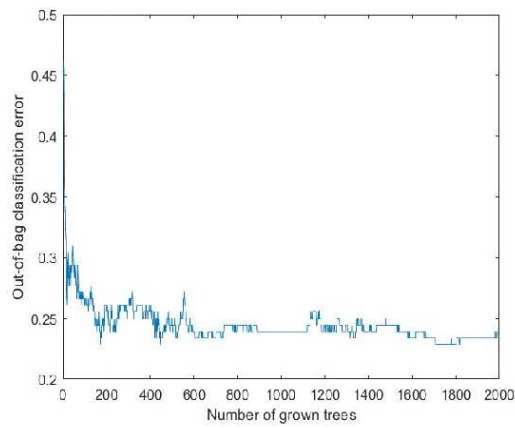
8 Allegati



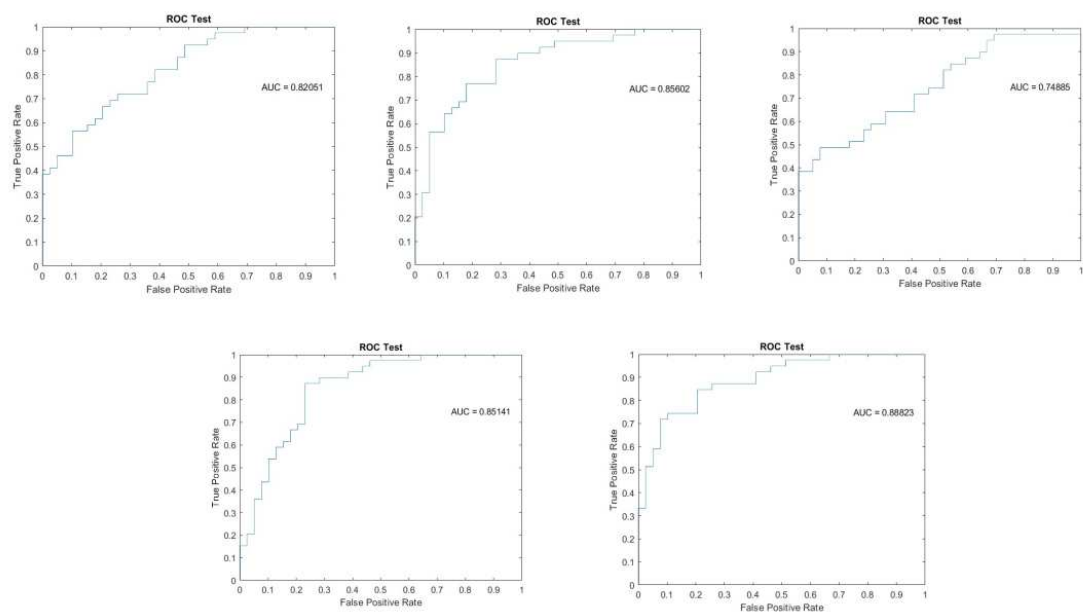
Allegato 21: Grafici dell'OOB della prima configurazione. Run 2 in alto a destra, run 3 in alto a sinistra; run 4 in basso a destra e run 5 in basso a sinistra (analisi preliminare)



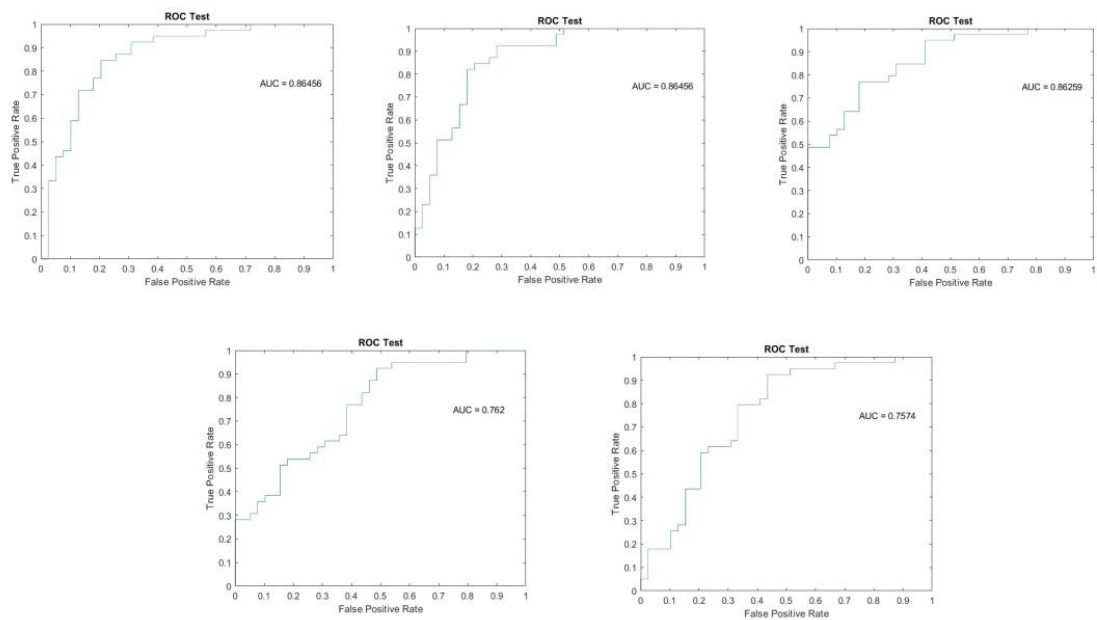
Allegato 22: Grafici dell'OOB della seconda configurazione. Run 2 in alto a destra, run 3 in alto a sinistra; run 4 in basso a destra e run 5 in basso a sinistra (analisi preliminare)



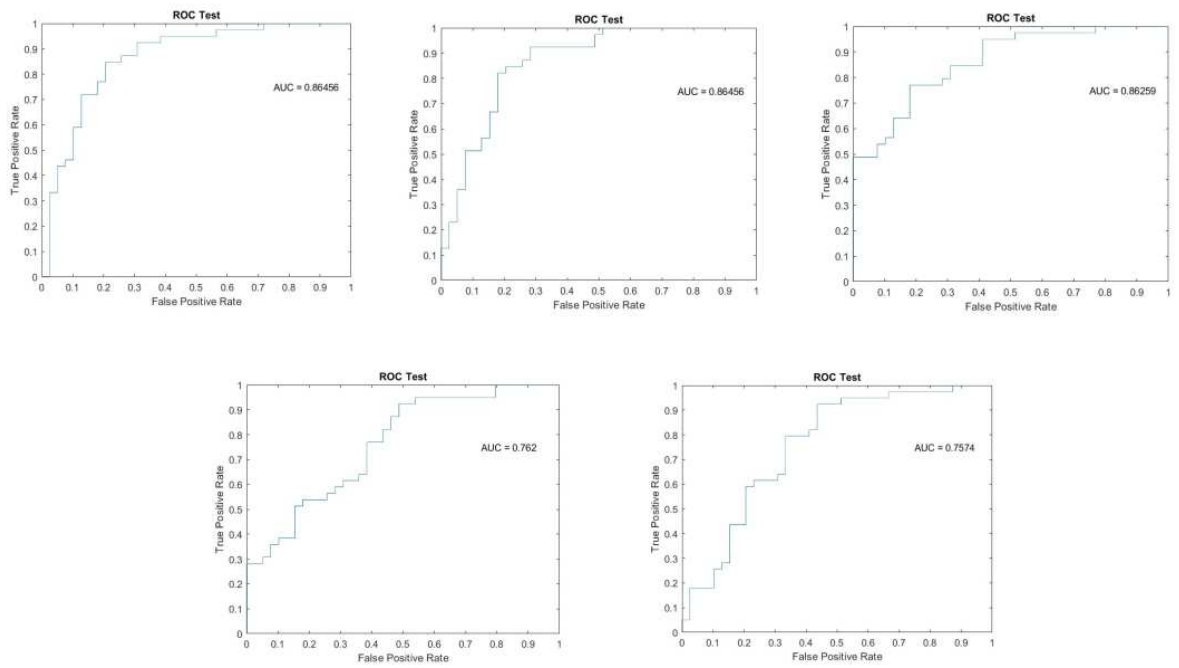
Allegato23: Grafici dell'OOB della terza configurazione. Run 2 in alto a destra, run 3 in alto a sinistra; run 4 in basso a destra e run 5 in basso a sinistra (analisi preliminare)



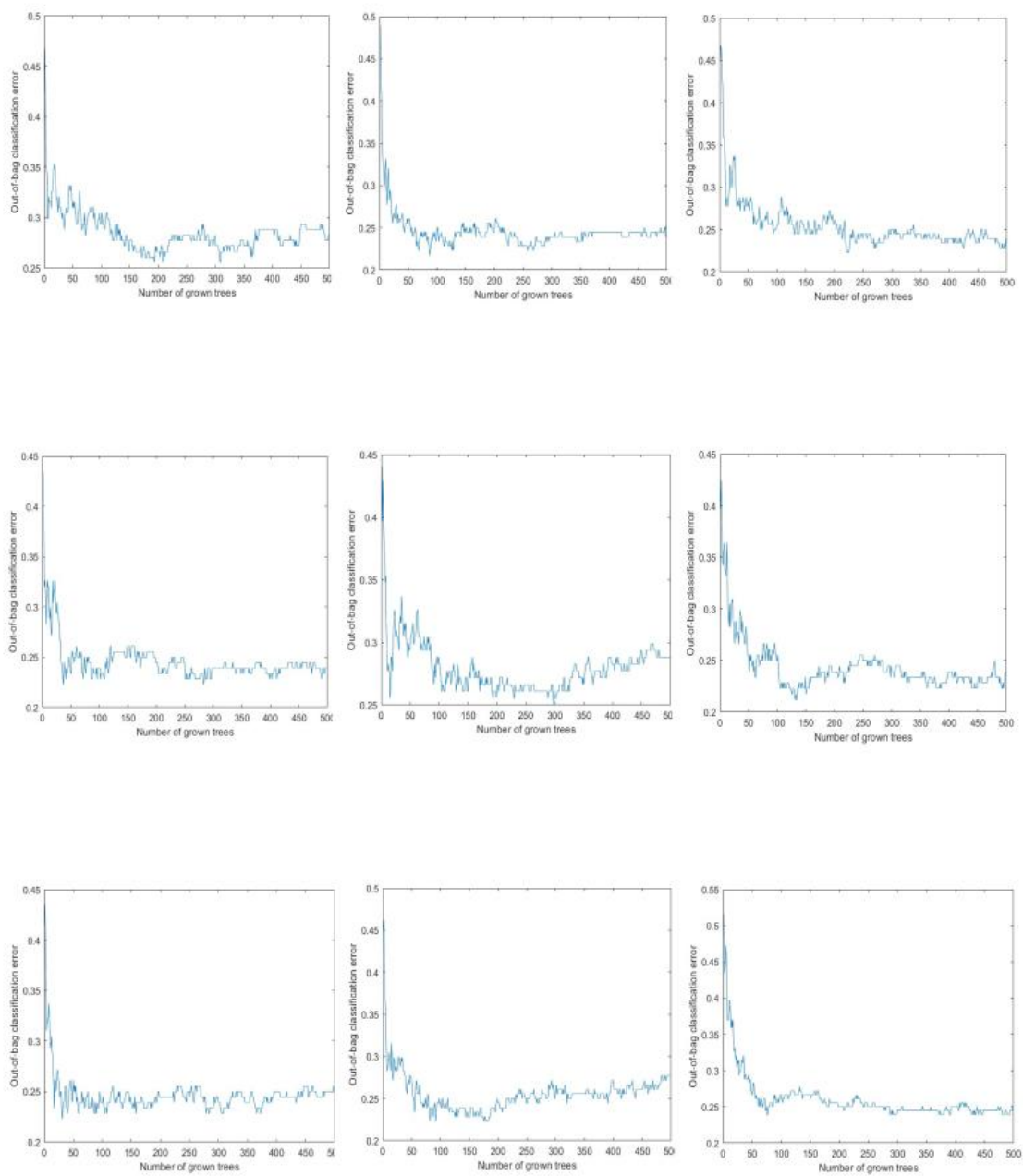
Allegato 24: Curve ROC e valori di AUC delle 5 esecuzioni della prima configurazione. Nella prima riga, da sinistra verso destra, si osservano il run 1, run 2 e run 3. Nella seconda riga, da sinistra verso destra, si osservano il run 4 e il run 5. (analisi preliminare)



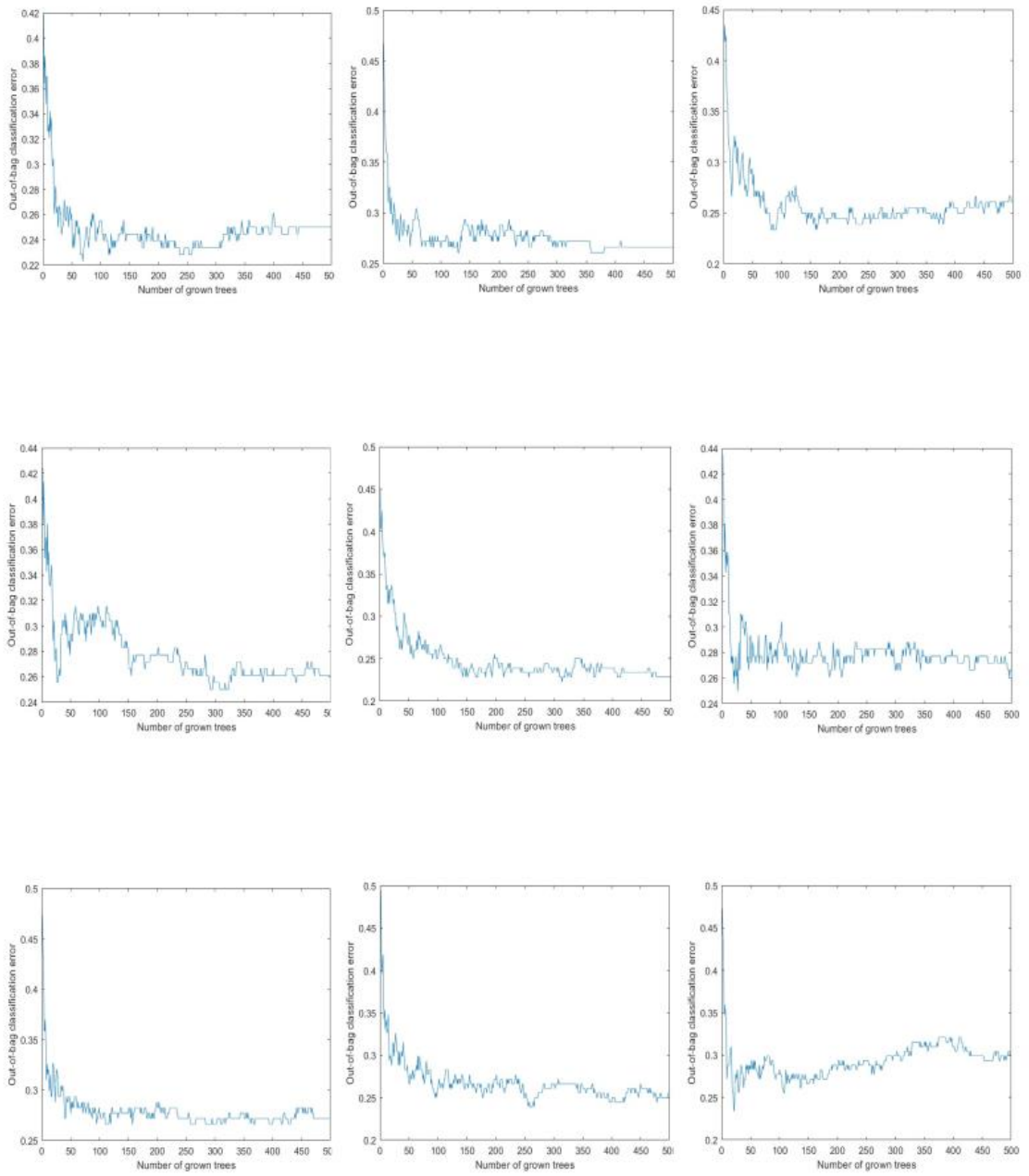
Allegato 25: Curve ROC e valori di AUC delle 5 esecuzioni della seconda configurazione. Nella prima riga, da sinistra verso destra, si osservano il run 1, run 2 e run 3. Nella seconda riga, da sinistra verso destra, si osservano il run 4 e il run 5. (analisi preliminare)



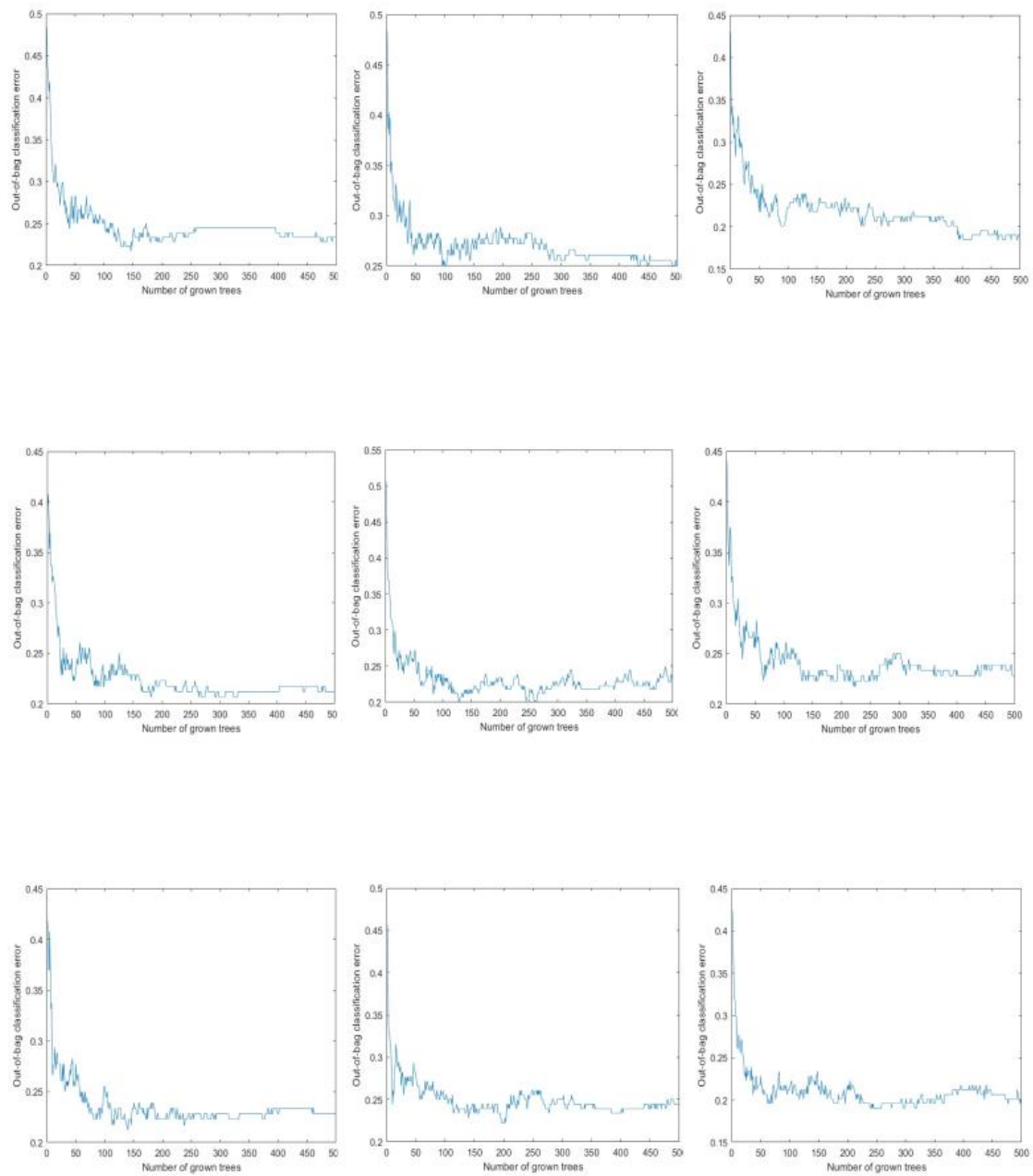
Allegato 26:: Curve ROC e valori di AUC delle 5 esecuzioni della terza configurazione. Nella prima riga, da sinistra verso destra, si osservano il run 1, run 2 e run 3. Nella seconda riga, da sinistra verso destra, si osservano il run 4 e il run 5. (analisi preliminare)



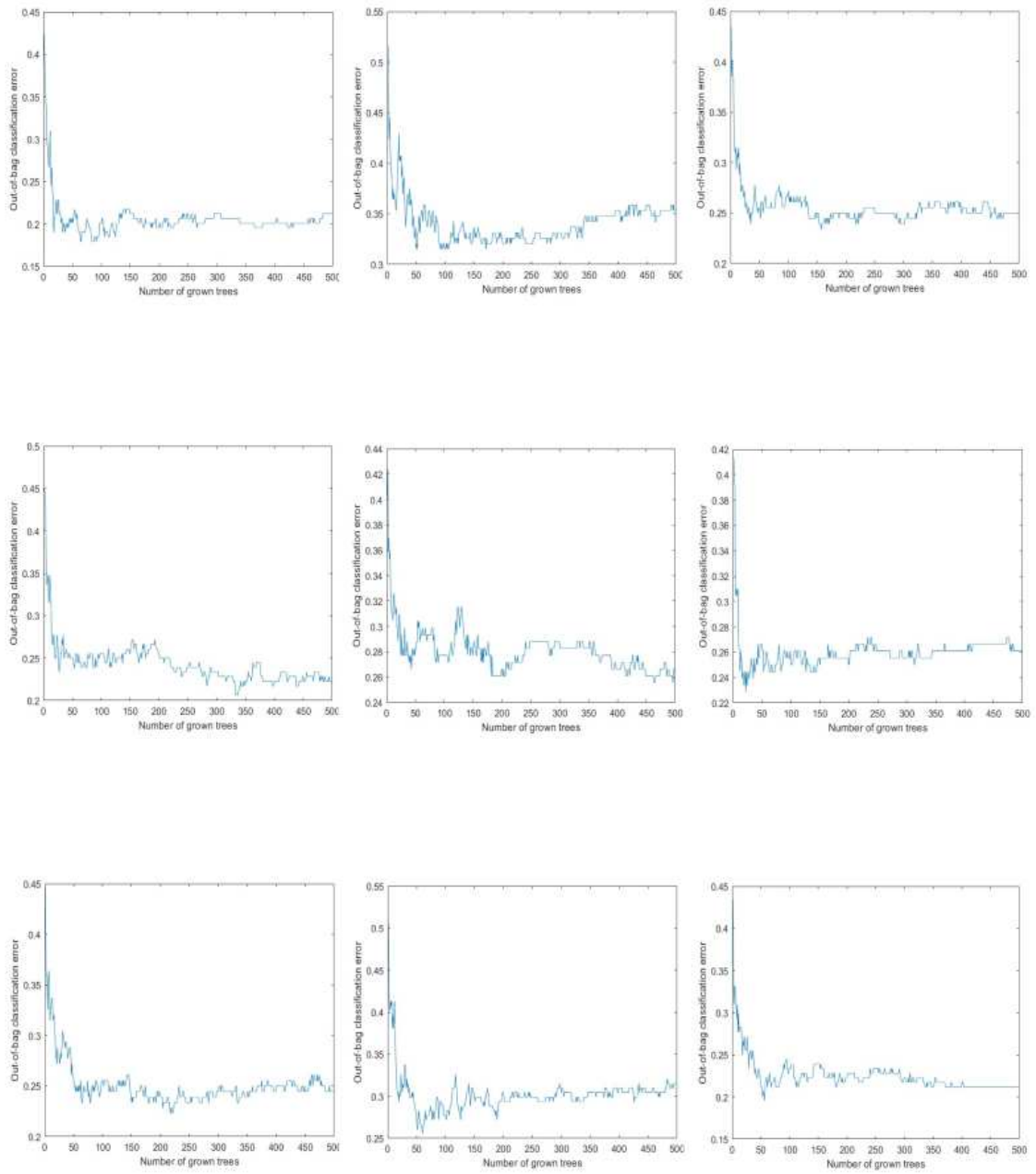
Allegato 27: Grafici dell'OOB della prima configurazione. Prima riga da sinistra a destra: run 2, run 3 e run 4. Seconda riga da sinistra a destra: run 5, run 6 e run 7. Terza riga da sinistra a destra: run 8, run 9 e run 10 (analisi finale)



Allegato 28: Grafici dell'OOB della seconda configurazione. Prima riga da sinistra a destra: run 2, run 3 e run 4. Seconda riga da sinistra a destra: run 5, run 6 e run 7. Terza riga da sinistra a destra: run 8, run 9 e run 10 (analisi finale)



Allegato 29: Grafici dell'OOB della terza configurazione configurazione. Prima riga da sinistra a destra: run 2, run 3 e run 4. Seconda riga da sinistra a destra: run 5, run 6 e run 7. Terza riga da sinistra a destra: run 8, run 9 e run 10 (analisi finale)



Allegato 30: Grafici dell'OOB della quarta configurazione. Prima riga da sinistra a destra: run 2, run 3 e run 4. Seconda riga da sinistra a destra: run 5, run 6 e run 7. Terza riga da sinistra a destra: run 8, run 9 e run 10 (analisi finale)