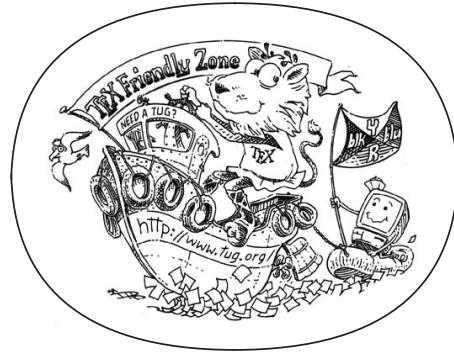


A CLASSIC THESIS STYLE

ANDRÉ MIEDE



An Homage to The Elements of Typographic Style

September 2015 – version 4.2

ABSTRACT

Short summary of the contents... a great guide by Kent Beck how to write good abstracts can be found here:

<https://plg.uwaterloo.ca/~migod/research/beck00PSLA.html>

CONTENTS

I	SOME KIND OF MANUAL	1
1	INTRODUZIONE	3
II	THE SHOWCASE	7
2	NOTAZIONI E DEFINIZIONI	9
3	METODO PER IL RANKING DELLE STAZIONI FERROVIARIE	15
4	PSEUDO-CODICE DEL METODO	19
5	CASO DI STUDIO	23
6	VALIDAZIONE E VALUTAZIONE DEL METODO	25
7	ESTENSIONE DEL METODO ALLE LINEE FERROVIARIE	39
8	PROGETTAZIONE DELLA BASE DI DATI SPAZIALE	41
III	APPENDIX	45
A	COSTRUZIONE DEL DATASET	47

LIST OF FIGURES

Figure 1	Frana di crollo. I massi che si staccano dalla cima rotolano lungo il pendio e arrivano a valle.	4
Figure 2	Frana di ribaltamento. Un'intera parete rocciosa si distacca dal pendio.	4
Figure 3	Frana da scivolamento planare. Un'area di terreno intera scivola verso la valle.	4
Figure 4	Frana da colata. I detriti scivolano verso la valle formando un corridoio.	4
Figure 5	Rappresentazione di una frana di crollo. Il masso si stacca sulla cima del pendio e scende già aumentando la sua energia meccanica. Finito il pendio l'energia meccanica accumulata diminuisce fino ad esaurirsi. Il punto esatto dove si ferma il masso non è dato saperlo, ma si può stimare una distanza d oltre la quale ragionevolmente il masso non potrà spingersi.	5
Figure 6	Frana di Colonia Las Colinas, El Salvador, 2001. E' possibile osservare come, in questo caso, la frana ha seguito una traiettoria ben delimitata tipica delle frane di colata. In questo caso specifico la distanza percorsa dalla frana è stata pari a circa 700 metri (URL!!!!!!).	5
Figure 7	Rappresentazione di una frana di colamento. E' possibile definire un raggio di azione r inteso come l'ampiezza del "corridoio" della frana.	6
Figure 8	Confini della provincia di Roma.	9
Figure 9	Confini della regione Lazio.	9
Figure 10	Confini del territorio italiano.	9
Figure 11	Rappresentazione delle Zones di una GeoArea. La geometria del confine delle porzioni di terreno è puramente indicativa. Le diverse tonalità di arancione rappresentano i valori di Szk delle Zones.	10
Figure 12	Il pallino rosso rappresenta il generico edificio b_i	10
Figure 13	In verde chiaro è rappresentato il buffer di raggio r centrato sull'edificio b_i	10
Figure 14	In rosa sono rappresentate le $nz_{i,j}$ con $i = 1$ che si trovano all'interno della HazardArea $_i$. E' possibile notare come $nz_{1,10}$ sia una Zone intera mentre $nz_{1,3}$ è parziale.	11
Figure 15	Con i pallini verdi sono rappresentati i centri di massa $cnz_{i,j}$ degli $nz_{i,j}$	11
Figure 16	Le linee in blu tratteggiate rappresentano le nearest isoipse $ni_{1,o}$	12
Figure 17	I quattro poligoni in blu con i contorni gialli rappresentano le $zf_{1,1,t}$ relative alla nearest zone $nz_{1,1}$	12
Figure 18	I pallini bianchi rappresentano i $czf_{1,1,t}$ ovvero i centri di massa delle zone fragments $zf_{1,1,t}$ della (Figura 17).	12
Figure 19	Nell'immagine a sinistra è possibile definire visivamente l'equazione 2. La retta di regressione lineare $lr_{i,j}$ è di colore celeste. Essa minimizza la distanza (rappresentata con la linea verde chiaro) tra i punti (in questo caso gli $czf_{i,j,t}$) e la retta stessa.	13

Figure 20	Due assets che si trovano nei pressi di un pendio. Entrambi hanno la propria <i>HazardArea</i> di raggio r , di cui quella dell'edificio in rosso comprende il punto di distacco del masso. Quindi l'edificio si trova in una zona pericolosa, al contrario di quello di colore verde.	15
Figure 21	Da scrivere	16
Figure 22	Da scrivere	16
Figure 23	La Buffered Linear Regression $blr_{i,j}$ ottenuta dalla <i>LineBuffering</i> della retta di regressione lineare $lr_{i,j}$. Essa rappresenta l'area attorno la retta entro una distanza d dalla stessa.	16
Figure 24	L'area di colore in giallo intorno all'edificio rappresenta il <i>BuildingBuffer_i</i> di raggio d	17
Figure 25	Il rettangolo verde rappresenta una <i>LandSlide</i> che non costituisce un pericolo per l'edificio. In rosso invece è stata raffigurata una <i>LandSlide</i> ad altissimo rischio.	17
Figure 26	A sinistra il caso di impatto totale a cui corrisponderà una <i>Impact Factor</i> pari a 1. A destra il caso di impatto parziale a cui corrisponderà una <i>Impact Factor</i> < 1	18
Figure 27	Da Scrivere	18
Figure 28	Da Scrivere	18
Figure 29	Il software QGis utilizzato per la visualizzazione del dataset.	24
Figure 30	Visualizzazione 3D del dataset attraverso il plugin Qgis2threejs di QGIS	24
Figure 31	Da scrivere	33
Figure 32	Da Scrivere	33
Figure 33	Da scrivere.	33
Figure 34	Da Scrivere	34
Figure 35	Da scrivere	34
Figure 36	In figura vengono brevemente mostrati i passi concettuali nella progettazione di una base di dati. In blu vengono rappresentate le fasi concettuali. In rosso vengono rappresentati i vari schemi e descrizioni che si acquisiscono in output alla fine di ogni fase progettuale	41
Figure 37	In figura è rappresentato il diagramma Er della base di dati. sono presenti 3 associazioni ed 8 entità.	42
Figure 38	47
Figure 39	47
Figure 40	In Figura si può osservare l'immagine raster del territorio abruzzese ottenuta dalla missione SRTM	48
Figure 41	In figura si possono osservare le curve di livello di una porzione di territorio della regione Abruzzo	49
Figure 42	In figura si può osservare il dataset ottenuto tenendo conto delle curve di livello	49
Figure 43	In figura si può osservare il dataset ottenuto tenendo conto delle curve di livello ed applicando la scomposizione di Voronoi	50

LIST OF TABLES

Table 1	Le frane censite verificatesi in Abruzzo.	23
Table 2	In Tabella sono rappresentate le classi di pericolosità che possono essere associate ad un edificio	25
Table 3	In Figura è rappresentata una generica matrice di contingenza con tre classi di appartenenza. Con $N(i,j)$ viene indicato il valore in posizione i,j . $T_{R1} = N(1,1) + N(1,2) + N(1,3)$ $T_{R2} = N(2,1) + N(2,2) + N(2,3)$ $T_{R3} = N(3,1) + N(3,2) + N(3,3)$ $T_{C1} = N(1,1) + N(2,1) + N(3,1)$ $T_{C2} = N(1,2) + N(2,2) + N(3,2)$ $T_{C3} = N(1,3) + N(2,3) + N(3,3)$ $TOT = T_{R1} + T_{R2} + T_{R3} = T_{C1} + T_{C2} + T_{C3}$	26
Table 4	In tabella è rappresentata la matrice di contingenza binaria della Classe 1. $N_1 = N(1,1)$; $N_2 = N(1,2) + N(1,3)$; $N_3 = N(2,1) + N(3,1)$; $N_4 = N(2,2) + N(2,3) + N(3,2) + N(3,3)$; $T_{C1} = N_1 + N_3$; $T_{C2} = N_2 + N_4$; $T_{R1} = N_1 + N_2$; $T_{R2} = N_3 + N_4$; $TOT = T_{C1} + T_{C2} = T_{R1} + T_{R2}$	28
Table 5	In tabella è rappresentata la matrice di contingenza binaria della Classe 2. $N_1 = N(2,2)$; $N_2 = N(1,1) + N(1,3)$; $N_3 = N(1,2) + N(3,2)$; $N_4 = N(1,1) + N(1,3) + N(3,1) + N(3,3)$; $T_{C1} = N_1 + N_3$; $T_{C2} = N_2 + N_4$; $T_{R1} = N_1 + N_2$; $T_{R2} = N_3 + N_4$; $TOT = T_{C1} + T_{C2} = T_{R1} + T_{R2}$	28
Table 6	In tabella è rappresentata la matrice di contingenza binaria della Classe 3. $N_1 = N(3,3)$; $N_2 = N(1,1) + N(1,2)$; $N_3 = N(3,1) + N(3,2)$; $N_4 = N(1,1) + N(1,2) + N(2,1) + N(2,2)$; $T_{C1} = N_1 + N_3$; $T_{C2} = N_2 + N_4$; $T_{R1} = N_1 + N_2$; $T_{R2} = N_3 + N_4$; $TOT = T_{C1} + T_{C2} = T_{R1} + T_{R2}$	29
Table 7	Nella tabella sono rappresentate le quantità sopra definite e le posizioni che occupano all'interno della matrice di contingenza binaria	29
Table 8	In Tabella sono rappresentati gli intervalli di pericolosità	30
Table 9	Nella tabella si possono osservare le discrepanze tra la classe di appartenenza reale, definita dal colore nella colonna delle stazioni, e quella calcolata attraverso l'algoritmo.	35
Table 10	Nella tabella viene mostrata la matrice di contingenza ricavata a partire dai risultati ottenuti dal metodo di calcolo	36
Table 11	In tabella viene mostrata la matrice di contingenza binaria della classe a bassa pericolosità ricavata a partire dalla tabella di contingenza non binaria	36
Table 12	In tabella viene mostrata la matrice di contingenza binaria della classe a media pericolosità ricavata a partire dalla tabella di contingenza non binaria	36
Table 13	In tabella viene mostrata la matrice di contingenza binaria della classe ad alta pericolosità ricavata a partire dalla tabella di contingenza non binaria	37
Table 14	In tabella, per ogni fascia di pericolo, sono esposti i risultati delle metriche introdotte in precedenza (True Positive Rate, True Negative Rate, Precision, Accuracy).	37

Table 15	In tabella sono descritte le entità in modo dettagliato	43
Table 16	tab:associazioni	44

LISTINGS

ACRONYMS

DRY Don't Repeat Yourself

API Application Programming Interface

UML Unified Modeling Language

Part I
SOME KIND OF MANUAL

INTRODUZIONE

Con il termine *frana* si indicano tutti i fenomeni di movimento di terreno, come ad esempio caduta di rocce e scivolamento di detriti superficiali, lungo un pendio. La forza di gravità, a cui è soggetta la massa di terreno, è sicuramente la causa scatenante più importante ma non è l'unica. Una frana avviene nel momento in cui un pendio passa da una condizione di stabilità ad una di instabilità. Tale passaggio di stato è causato da diversi fattori i quali possono contribuire contemporaneamente o singolarmente. Tra le cause naturali sono evidenziati le seguenti:

1. pressione di acqua sotterranea (acqua piana) che agisce per destabilizzare la pendenza;
2. assenza di vegetazione (dovuta ad esempio ad un incendio);
3. erosione della punta di un pendio;
4. indebolimento di un pendio a causa di piogge pesanti o scioglimento di ghiacci/neve;
5. terremoti che aggiungono carichi a pendenze già poco stabili.

Il fenomeno delle frane è aggravato inoltre da fattori riconducibili alle attività umane come:

1. deforestazioni e costruzioni che destabilizzano i pendii;
2. sterri che alterano la forma dei pendii e che quindi posso aggiungere nuovi carichi.

Una frana generalmente è caratterizzata dal tipo di materiale movimentato (ad esempio materiale roccioso o semplicemente detriti superficiali) e dal tipo di movimento che il materiale segue. Queste due caratteristiche principali vengono poi integrate considerando anche altri aspetti come la velocità e l'entità del movimento. In base al comportamento della frana possiamo quindi avere la seguente classificazione:

1. **Frane di crollo:** massa che si stacca da un versante molto acclive e che, successivamente, si muove per caduta libera, rimbalzo, rotolamento. Il distacco può avvenire per rottura di taglio o di trazione della roccia (Figura 1).
2. **Frane da ribaltamento:** rotazione in avanti, verso l'esterno del versante, di una massa di terra o roccia, intorno ad un punto situato al di sotto del baricentro della massa spostata (Figura 2).
3. **Frana da scivolamento planare:** movimento verso la base del versante di una massa di terra o roccia che avviene per deformazioni di taglio lungo una o più superfici o entro uno spessore limitato di materiale (Figura 3).
4. **Frana da colata o colamento:** movimento distribuito in maniera continua all'interno della massa spostata. Le superfici di taglio, se presenti, sono multiple e spesso temporanee. La distribuzione delle velocità nella massa spostata è analoga a quella all'interno di un fluido viscoso (Figura 4).

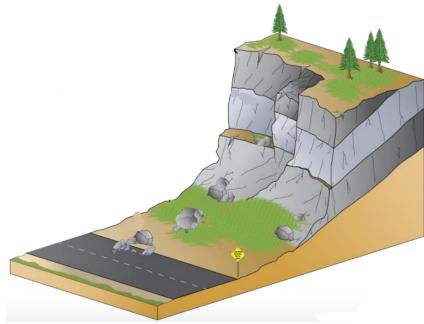


Figure 1: Frana di crollo. I massi che si staccano dalla cima rotolano lungo il pendio e arrivano a valle.

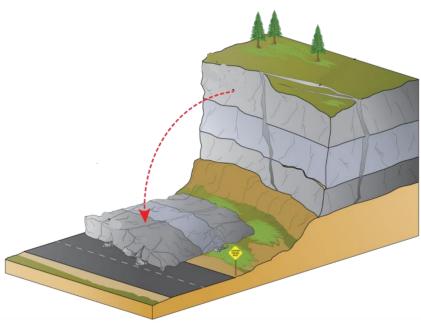


Figure 2: Frana di ribaltamento. Un'intera parete rocciosa si distacca dal pendio.

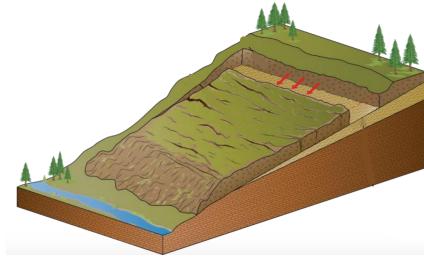


Figure 3: Frana da scivolamento planare. Un'area di terreno intera scivola verso la valle.

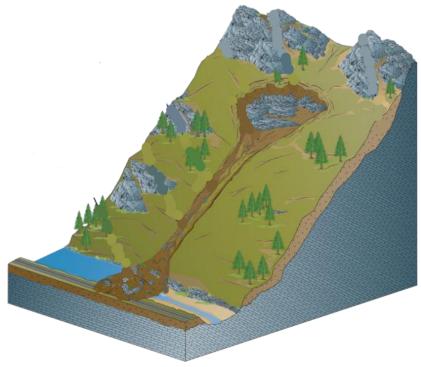


Figure 4: Frana da colata. I detriti scivolano verso la valle formando un corridoio.

Rispetto a questa classificazione è possibile trarre alcune conclusioni riguardo l'andamento spaziale di una frana. Innanzitutto è complicato stimare lo spazio percorso dalla frana. Prendendo, ad esempio, in considerazione una frana di crollo la distanza percorsa dal masso dipende da un numero elevato di variabili fisiche tra loro correlate (accelerazione, velocità, urti, massa ecc ecc). E' ragionevole però, con la dovuta prudenza che il caso richiede, stabilire un raggio di azione della frana, oltre il quale il fenomeno non può proseguire. Del resto un masso, dopo aver rotolato lungo un pendio, possiede un quantitativo limitato di energia meccanica la quale verrà a poco a poco dissipata lungo il tratto non in pendenza fino ad esaurirsi completamente (Figura 5).

Un'altra osservazione la si può dedurre analizzando le fotografie scattate nei luoghi dove si sono verificate delle frane. In molti casi, soprattutto nei casi di frana da colamento, il terreno movimentato segue una traiettoria ben precisa (Figura 6). Quest'ultima dipende dalla morfologia del pendio, ovvero da come sono distribuite le masse di terreno lungo la discesa. Anche in questo caso è possibile stabilire un raggio d'azione inteso come la larghezza del "corridoio" dentro il quale la frana scivola (Figura 7). Infine bisogna sempre tenere a mente che la forza di gravità è il fattore scatenante per eccellenza della frana, motivo per cui la pendenza del terreno giocherà sempre un ruolo fondamentale.

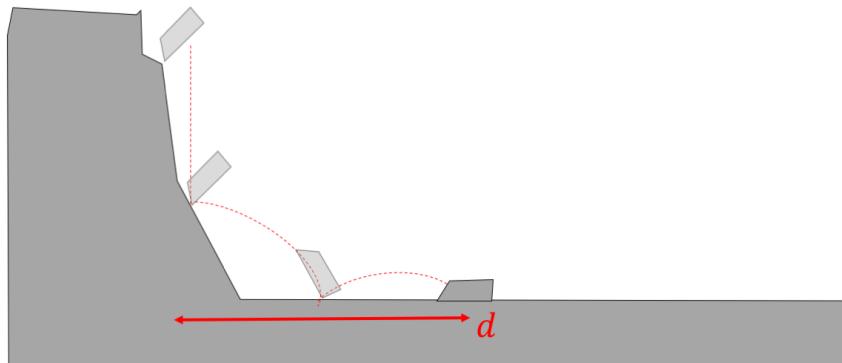


Figure 5: Rappresentazione di una frana di crollo. Il masso si stacca sulla cima del pendio e scende già aumentando la sua energia meccanica. Finito il pendio l'energia meccanica accumulata diminuisce fino ad esaurirsi. Il punto esatto dove si ferma il masso non è dato saperlo, ma si può stimare una distanza d oltre la quale ragionevolmente il masso non potrà spingersi.



Figure 6: Frana di Colonia Las Colinas, El Salvador, 2001. E' possibile osservare come, in questo caso, la frana ha seguito una traiettoria ben delimitata tipica delle frane di colata. In questo caso specifico la distanza percorsa dalla frana è stata pari a circa 700 metri ([URL!!!!!!!!!](#)).

La caratterizzazione di una frana è fondamentale e propedeutica ad enucleare una efficace strategia di prevenzione. Poter disporre di una strategia di questo tipo è particolarmente rilevante nel caso in cui l'ente che la mette in pratica ha un numero di *assets strategici* molto elevato. Prendiamo in considerazione, solo per citare un esempio, Ferrovie dello Stato Italiane Spa. L'azienda possiede, tra i suoi assets, una rete ferroviaria di 16734 Km e più di 3,000 stazioni dislocate su un territorio (L'Italia) la cui superficie è di 301,340 Km². E' impensabile che i vertici delle Ferrovie dello Stato approvino un piano aziendale di prevenzione che preveda il controllo di ciascuna stazione e ciascun Km della rete ferroviaria! Sarebbe impossibile sia in termini di tempo necessario che di sforzo economico che l'Ente dovrebbe sostenere.

La soluzione ideale sarebbe quella di avere un metodo che prenda in input un insieme di assets strategici e restituisca come output un lista ordinata in base alla loro esposizione al rischio frana. In questo modo sarebbe possibile occuparsi degli assets che sono veramente a rischio e quindi, oltre a risparmiare denaro, essere più rapidi negli interventi.

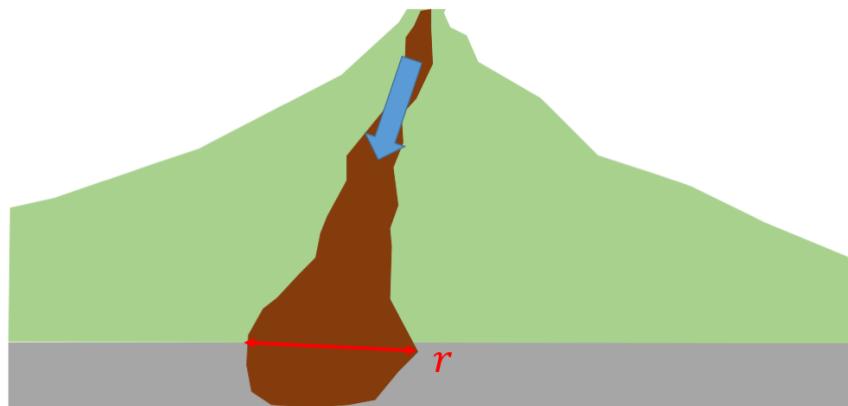


Figure 7: Rappresentazione di una frana di colamento. E' possibile definire un raggio di azione r inteso come l'ampiezza del "corridoio" della frana.

Il metodo da noi proposto verrà illustrato nelle prossime sezioni del documento. In particolare nella sezione 2 vengono fornite le definizioni dei concetti base, fondamentale per la comprensione del metodo, attraverso una formulazione matematica. Nella sezione 3 verrà spiegato in che modo le considerazioni fatte sulla pendenza del terreno e le traiettorie delle frane sono state affrontate e tradotte in equazioni matematiche. Infine nella sezione 4 verrà esposto come tali equazioni vengono utilizzate nello pseudo-codice dell'algoritmo.

Part II
THE SHOWCASE

NOTAZIONI E DEFINIZIONI

1. **GEOAREA** : Definisce un territorio di interesse per il calcolo dell'esposizione a rischio frana. La *GeoArea* può rappresentare un comune, una provincia (Figura 8), una regione (Figura 9), una nazione (Figura 10) fino ad arrivare ad un intero continente. *GeoArea* è descritta dalla tupla $\langle ID, description, boundary \rangle$.

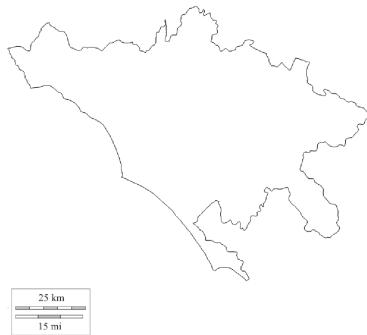


Figure 8: Confini della provincia di Roma.

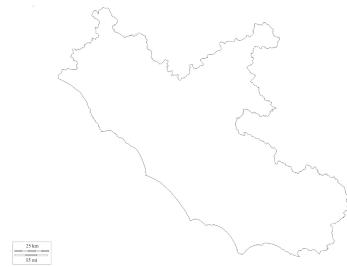


Figure 9: Confini della regione Lazio.



Figure 10: Confini del territorio italiano.

2. **Z (Zones)** = $\{z_k | k = 1, 2, \dots\} | z_k \text{ è una Zone di GeoArea}\}$. Una *Zone* definisce una porzione di terreno all'interno della *GeoArea* ed è descritta dalla tupla $\langle ID, boundary, Sz_k \rangle$. Il campo *ID* identifica univocamente la *k*-esima *Zone*, *boundary* la geometria del confine e *Sz_k* è un valore decimale che rappresenta la probabilità di frana di *z_k* (Figura 11). Si noti che *z_k* è una notazione sovraccaricata in quanto rappresenta sia l'*ID* di una *Zone* che la sua geometria.
3. **B (Buildings)** = $\{b_i | i = 1, 2, \dots\} | b_i \text{ è un building ubicato all'interno dei confini della GeoArea}\}$. Per *building* si intende un generico edificio (come ad esempio una scuola, un grattacielo, un hotel e strutture simili) descritto dalla tupla $\langle ID, description, position \rangle$. Il campo *ID* identifica univocamente l'edificio; *description* una sua descrizione testuale e infine *position* rappresenta la posizione geografica di *b_i* espressa da una coppia di coordinate.

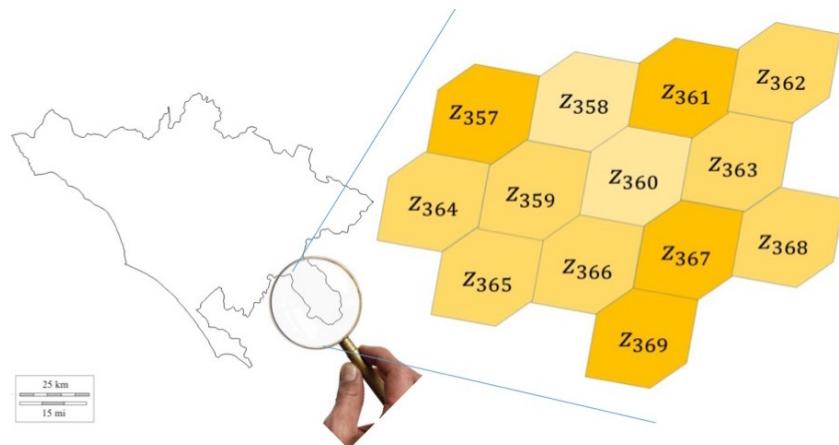


Figure 11: Rappresentazione delle Zones di una *GeoArea*. La geometria del confine delle porzioni di terreno è puramente indicativa. Le diverse tonalità di arancione rappresentano i valori di S_{Zk} delle Zones.

coordinate geografiche (Figura 12). Definiamo con $\text{card}(\mathcal{B})$ la cardinalità dell'insieme B , ovvero il numero di edifici presenti in *GeoArea*. Dettagli come la planimetria della struttura potrebbero fornire ulteriori dati sensibili ma sono molto difficili da reperire e quindi non verranno presi in considerazione. In seguito verrà usato sempre l'indice i per riferirsi al building i -esimo. Nelle notazioni introdotte di seguito l'occorrenza del pedice i viene usata sempre per richiamare che ci si riferisce all'edificio i – *mo*, ovvero all'elemento b_i di \mathcal{B} .

4. HazardArea_i rappresenta l'area a possibile rischio smottamento circostante l'edificio b_i , definita come la circonferenza centrata su b_i di raggio r (Figura 13).

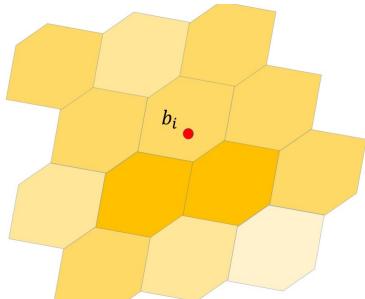


Figure 12: Il pallino rosso rappresenta il generico edificio b_i .

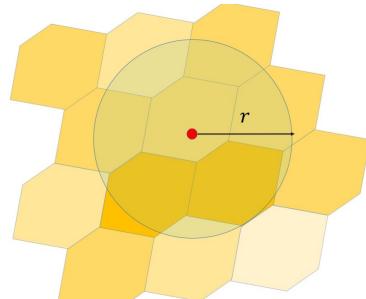


Figure 13: In verde chiaro è rappresentato il buffer di raggio r centrato sull'edificio b_i .

5. \mathcal{NZ}_i (**Nearest Zones**) = $\{nz_{i,j} | i = 1, \dots, \text{card}(\mathcal{B}), j = 1, \dots, \text{card}(\mathcal{NZ}_i)\} | nz_{i,j} \in \mathcal{Z} \cap \text{HazardArea}_i\}$ è una *Nearest Zone*, ovvero una *Zone* che si trova all'interno della HazardArea_i di b_i . Gli $nz_{i,j}$ possono rappresentare le *Zones* nella loro interezza, oppure parte di esse derivanti dall'intersezione con la HazardArea_i . (Figura 14). In seguito verrà usato sempre l'indice j accostato all'indice i per riferirsi al nearest zone j -esimo del building i .

esimo. Nelle notazioni introdotte di seguito l'occorrenza del pedice $\{i, j\}$ viene usata sempre per richiamare che ci si riferisce al nearest zone $j - mo$ dell'edificio $i - mo$, ovvero all'elemento $nz_{i,j}$ di \mathcal{NZ}_i . Si noti che $nz_{i,j}$ è una notazione sovraccaricata in quanto rappresenta sia l'*ID* di una *Nearest Zone* che la sua geometria.

6. $cnz_{i,j}$ (**Centroid of** $nz_{i,j}$) corrisponde al centro di massa di $nz_{i,j}$ la cui posizione è descritta dalle sue coordinate geografiche (Figura 15).
7. $\mathcal{I} = \{i_p (p = 1, 2, ..) | i_p \text{ è una } isoipse \text{ della GeoArea}\}$. Per isoipse si intende quella curva che unisce punti con uguale quota, descritta dalla tupla $< ID, elevation, geometry >$. Il campo *ID* identifica univocamente la curva di livello, *elevation* la quota e infine *geometry* una geometria che descrive l'andamento dell'isoipse.
8. \mathcal{NI}_i (**Nearest Isoipses**) = $\{ni_{i,o} (i = 1, .., \text{card}(\mathcal{B}), o = 1, .., \text{card}(\mathcal{NI}_i)) | ni_{i,o} \in \mathcal{I} \cap \text{HazardArea}_i\}$ è la o -esima *Nearest Isoipse* che si trova all'interno della *HazardArea* $_i$ di b_i . Le $ni_{i,o}$ possono rappresentare le *isoipse* nella loro interezza, oppure una parte di esse derivanti dall'intersezione con la *HazardArea* $_i$ (Figura 16). Si noti che $ni_{i,o}$ è una notazione sovraccaricata in quanto rappresenta sia l'*ID* di una *Nearest Isoipse* che la sua geometria.

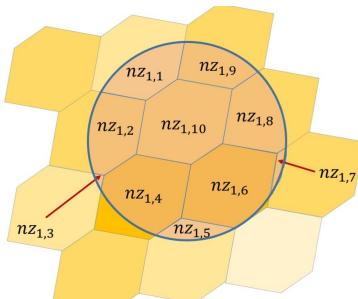


Figure 14: In rosa sono rappresentate le $nz_{i,j}$ con $i = 1$ che si trovano all'interno della *HazardArea* $_i$. È possibile notare come $nz_1,10$ sia una *Zone* intera mentre $nz_1,3$ è parziale.

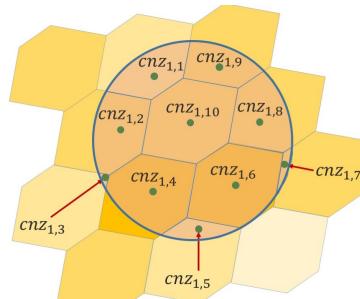


Figure 15: Con i pallini verdi sono rappresentati i centri di massa $cnz_{i,j}$ degli $nz_{i,j}$.

9. \mathcal{ZF}_i (**Zone Fragments**) = $\{zf_{i,j,t} (i = 1, .., \text{card}(\mathcal{B}), j = 1, .., \text{card}(\mathcal{NZ}_i), t = 1, ..) | zf_{i,j,t}$ è la t -esima *Zone Fragment* interna alla nearest zone $nz_{i,j}$ restituita dall'operazione di splitting di detta nearest zone con le sole nearest isoipse in \mathcal{NZ}_i che l'attraversa. I quattro poligoni in blu con i contorni gialli della Figura 17 rappresentano gli altrettanti *Zone Fragments* nei quali la Nearest Zone $nz_{i,j}$ viene frammentata dall'operazione di split con le tre isoipse che la attraversano. \mathcal{ZF}_i è l'insieme delle zone fragments prodotte dallo split tra gli $nz_{i,j} \in \mathcal{NZ}_i$ e le $ni_{i,o} \in \mathcal{NI}_i$ (Figura 17).
10. $czf_{i,j,t}$ (**Centroid of** $zf_{i,j,t}$) corrisponde al centro di massa dello zone fragment $zf_{i,j,t}$ la cui posizione è descritta dalle sue coordinate geografiche date dalla coppia $(czfx_{i,j,t}, czfy_{i,j,t})$ (Figura 18).

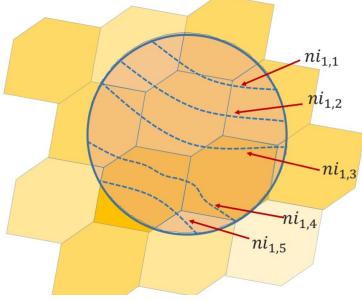


Figure 16: Le linee in blu tratteggiate rappresentano le nearest isoipse $ni_{1,0}$.

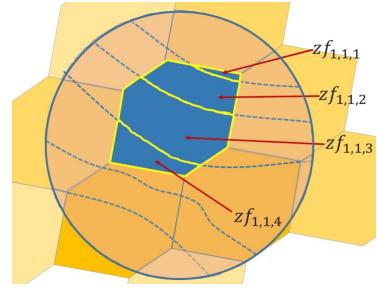


Figure 17: I quattro poligoni in blu con i contorni gialli rappresentano le $zf_{1,1,t}$ relative alla nearest zone $nz_{1,1}$.

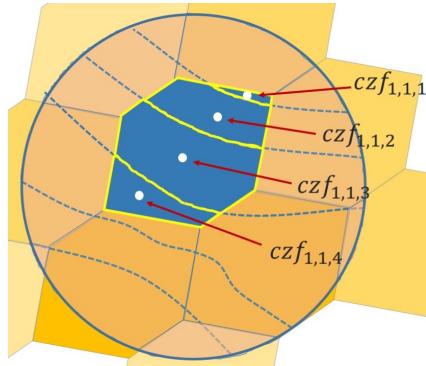


Figure 18: I pallini bianchi rappresentano i $czf_{1,1,t}$ ovvero i centri di massa delle zone fragments $zf_{1,1,t}$ della (Figura 17).

11. \mathcal{LR}_i (**LinearRegressions**) = { $lr_{i,j}$ ($i = 1, \dots, \text{card}(\mathcal{B})$), ($j = 1, \dots, \text{card}(\mathcal{NZ}_i)$)} | $lr_{i,j}$ rappresenta la retta di regressione lineare della Nearest Zone $nz_{i,j}$. Nel caso specifico rappresenta la retta passante per il centro di massa $cnz_{i,j}$ del nearest zone $nz_{i,j}$ con coefficiente angolare m tale che, dati i centri di massa $(czfx_{i,j,t}, czfy_{i,j,t})$ delle zone fragment $zf_{i,j,t}$, sia minima la somma dei quadrati dei loro scarti (Figura 19). Matematicamente essa è descritta come segue:

l'equazione (1) rappresenta la retta passante per il centro di massa $cnz_{i,j}$ rappresentata in Figura 19

$$y - cnzy_{i,j} = m(x - cnzx_{i,j}) \quad (1)$$

per la quale dati t punti $(czfx_{i,j,t}, czfy_{i,j,t})$ risulta minima la somma dei quadrati degli scarti:

$$S(m, q) = \sum_{i=0}^n (m * czfx_{j,t} + q - czfy_{j,t})^2 \quad (2)$$

$$\text{ove } q = cnzy_{i,j} - m * cnzx_{i,j} \quad (3)$$

Questo concetto matematico sarà utilizzato per stimare la direzione di caduta della frana relativa al $nz_{i,j}$. Per una migliore comprensione è possibile fare riferimento alla Figura 19.

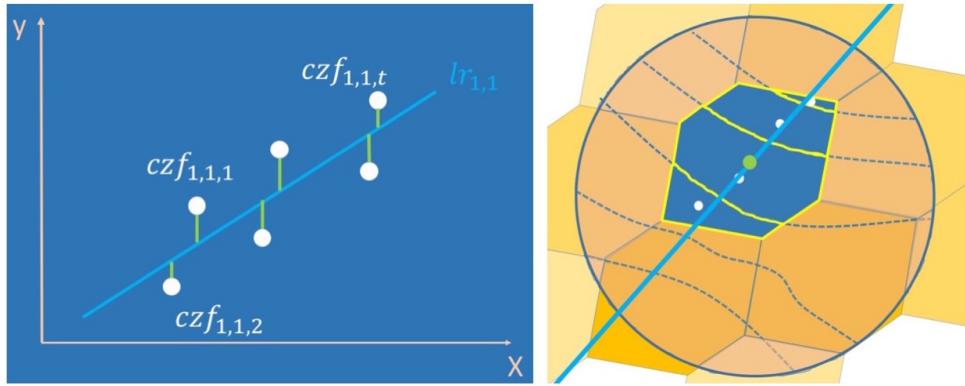


Figure 19: Nell'immagine a sinistra è possibile definire visivamente l'equazione 2. La retta di regressione lineare $lr_{i,j}$ è di colore celeste. Essa minimizza la distanza (rappresentata con la linea verde chiaro) tra i punti (in questo caso gli $czf_{i,j,t}$) e la retta stessa.

12. $BuildingBuffer_i$ rappresenta l'area circostante l'edificio b_i e approssima la sua estensione. È definita come la circonferenza centrata su b_i di raggio l dove $l < r$ con r raggio del HazardArea_i (Figura 24).
13. EXP (Exposure) = { $\exp_i | i = 1, \dots, \text{card}(\mathcal{B})$ } | \exp_i è il valore di *exposure* del building *i*-esimo. Per *exposure* si intende il valore numerico che indica quanto il building è esposto al rischio frana, rappresentato dalla tupla $< ID, description, position, exposure >$

METODO PER IL RANKING DELLE STAZIONI FERROVIARIE

Il fine del metodo che ci si accinge a descrivere è valutare il livello di rischio frana (Exposure) al quale potrebbero essere soggetto un insieme di assets. Con tale valutazione si potrà creare una graduatoria degli assets più a rischio di una determinata zona (GeoArea) e potrà essere di supporto agli enti responsabili di manutenzione o controllo della zona stessa. Il metodo che verrà sviluppato tiene a mente tutte le considerazioni presentate nell'introduzione, da cui è emerso che oltre alla ripidità del pendio e allo spazio percorso dalla frana è fondamentale anche la traiettoria seguita. Quest'ultima è alla base del metodo ed è stato il principale elemento di studio. Per un'analisi approfondita di questi aspetti sono state introdotte le curve di livello (\mathcal{J}). Inoltre il territorio preso in considerazione (GeoArea), dove sono dislocati gli assets strategici (\mathcal{B}), è stato suddiviso in piccole particelle di terreno (\mathcal{Z}) dalle quali può aver origine una frana se si verificano determinate condizioni che andremo a definire.

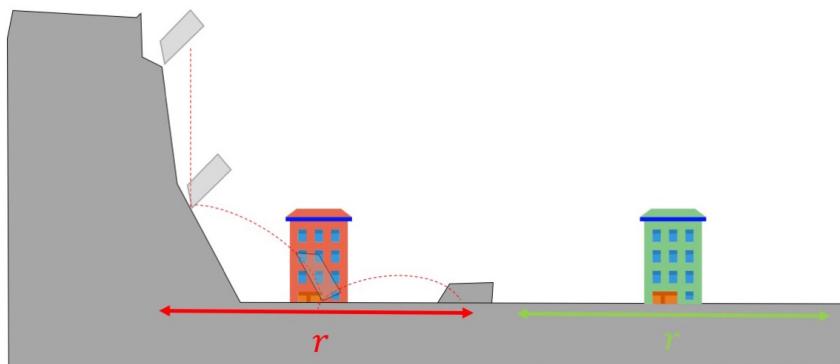


Figure 20: Due assets che si trovano nei pressi di un pendio. Entrambi hanno la propria *HazardArea* di raggio r , di cui quella dell'edificio in rosso comprende il punto di distacco del masso. Quindi l'edificio si trova in una zona pericolosa, al contrario di quello di colore verde.

Il metodo calcola per ogni asset b_i , il proprio valore di exposure. Ricordando dall'introduzione che una frana può estendersi per una distanza notevole ma pur sempre limitata (Figure 20), sembra logico limitare l'area di studio alla sola (HazardArea_i). L'analisi che segue quindi interesserà le sole Zones Z all'interno di tale area, ovvero le particelle appartenenti all'insieme NZ_i (Nearest Zones di b_i) (Figura 21). Come detto in precedenza, elemento fondamentale alla base del metodo è la traiettoria di frana. Il metodo infatti si occupa di stimare tale traiettoria per ognuna delle Nearest Zones $nz_{i,j} \in NZ_i$ dunque è indispensabile considerare le curve di livello (Figura 22).

Naturalmente le sole curve di livello di interesse alla $nz_{i,j}$ sono quelle che intersecano la (HazardArea_i), ovvero le (Nearest Isoipses) NI_i . Queste ultime vengono utilizzate per determinare la distribuzione delle masse all'interno della particella di terreno $nz_{i,j}$. Pertanto esse verranno utilizzate per frazionare la $nz_{i,j}$ in frammenti $zf_{i,j,t}$. I centri di massa ($czf_{i,j,t}$) di tali frammenti sono alla base del calcolo della regressione lineare. Dal calcolo della regressione lineare otteniamo la retta ($lr_{i,j}$) che stima appunto la traiettoria di frana della nostra particella $nz_{i,j}$.

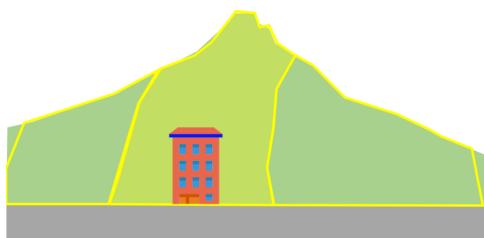


Figure 21: Da scrivere

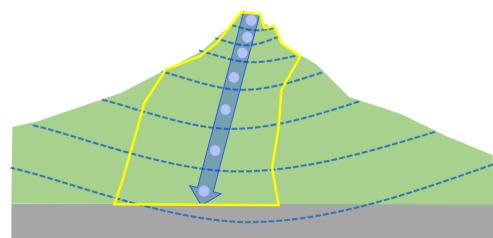


Figure 22: Da scrivere

A questo punto, ricavata la retta che stima la traiettoria di frana $lr_{i,j}$ della particella $nz_{i,j}$, bisogna determinare se la presunta frana di tale particella "impatta" o meno sull'edificio b_i oggetto di analisi. Per poter comprendere appieno cosa si intende per "impatto" è necessario definire le seguenti notazioni:

14. **\mathcal{BLR}_i (Buffered Linear Regressions)** = { $blr_{i,j} (i = 1,..,\text{card}(\mathcal{B})), (j = 1,..,\text{card}(\mathcal{NZ}_i))$ } | $\forall lr_{i,j} \in \mathcal{LR}_i$ si ha che $blr_{i,j}$ è la geometria risultante dalla *LineBuffering* di $lr_{i,j}$ Figura 23. $blr_{i,j}$ rappresenta quella che potrebbe essere la scia della frana della particella $nz_{i,j}$ alla quale ci si riferisce. Ovvero identifica l'area che tale frana mette a rischio. Per una migliore comprensione riferirsi alla (Figura 23)

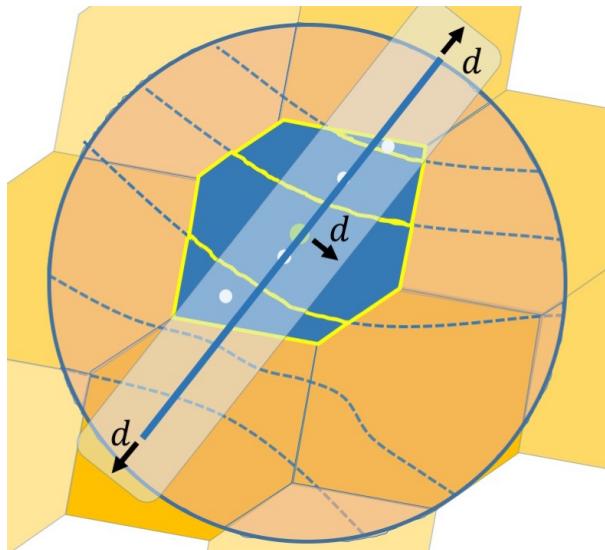


Figure 23: La **Buffered Linear Regression** $blr_{i,j}$ ottenuta dalla *LineBuffering* della retta di regressione lineare $lr_{i,j}$. Essa rappresenta l'area attorno la retta entro una distanza d dalla stessa.

15. **\mathcal{LS}_i (LandSlides)** = { $ls_{i,j} (i = 1,..,\text{card}(\mathcal{B})), (j = 1,..,\text{card}(\mathcal{LS}_i))$ } con $\mathcal{LS}_i \subseteq \mathcal{BLR}_i$ sotointersezione | $\mathcal{LS}_i = (\mathcal{BLR}_i \cap \text{BuildingBuffer}_i) \neq \emptyset$. Con parole esplicite, gli elementi dell'insieme \mathcal{LS}_i sono le sole $blr_{i,j}$ che intersecano il *BuildingBuffer* $_i$. Nello specifico, le sole frane che "impattano" l'edificio b_i (Figura 25).

Avendo ben definite le frane che "impattano" l'edificio b_i , viene da tener conto di un altro fattore determinante, ossia l'entità di tale "impatto". Una *LandSlide* di fatto può "impattare" su un edificio in modo totale o parziale e di conseguenza influire diversamente sul valore di exposure della stessa. Introduciamo pertanto la notazione di *Impact Factor*

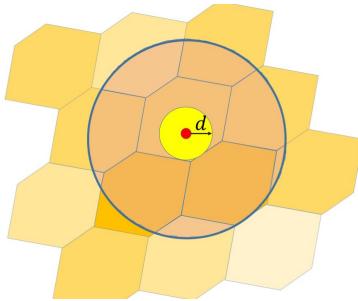


Figure 24: L'area di colore in giallo intorno all'edificio rappresenta il BuildingBuffer_i di raggio d.

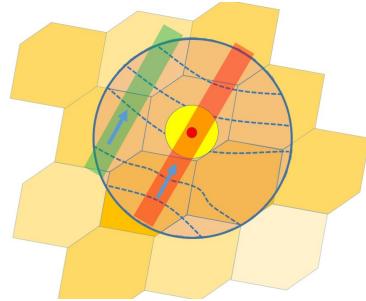


Figure 25: Il rettangolo verde rappresenta una LandSlide che non costituisce un pericolo per l'edificio. In rosso invece è stata raffigurata una LandSlide ad altissimo rischio.

16. IF_{ls_{i,j}} (**Impact Factor**) è un valore percentuale $\in (0, 1]$ che esprime l'entità dell'impatto della ls_{i,j} sull'edificio b_i. Essa corrisponde al rapporto tra l'area della ls_{i,j} ∩ BuildingBuffer_i e l'area del BuildingBuffer_i. Un valore di IF_{ls_{i,j}} = 1 equivale a dire che la frana impatta in modo totale sull'edificio. Da tenere conto che un valore pari a 0 non è possibile essendo già l'insieme LS_i quella delle sole frane che impattano su b_i. Matematicamente la *Impact Factor* è descritta come segue:

$$\text{ImpactFactor} = \frac{\text{Area}(\text{BuildingBuffer}_i \cap \text{ls}_{i,j})}{\text{Area}(\text{BuildingBuffer}_i)} \quad (4)$$

In (Figura 26) vediamo da due diverse prospettive i casi di impatto di una *LandSlide*.

$$\text{Exp}_{-b_i} = \sum_{k=1}^n \text{Exp}_{-b_{i,k}} \quad (5)$$

$$\text{Exp}_{-b_{i,k}} = \frac{\text{Size} * S_{z_k}}{d} * \text{ImpactFactor} \quad (6)$$

$$\text{Size} = \frac{\text{Area}(z_k)}{\text{the average values of all the areas of zones } Z} \quad (7)$$

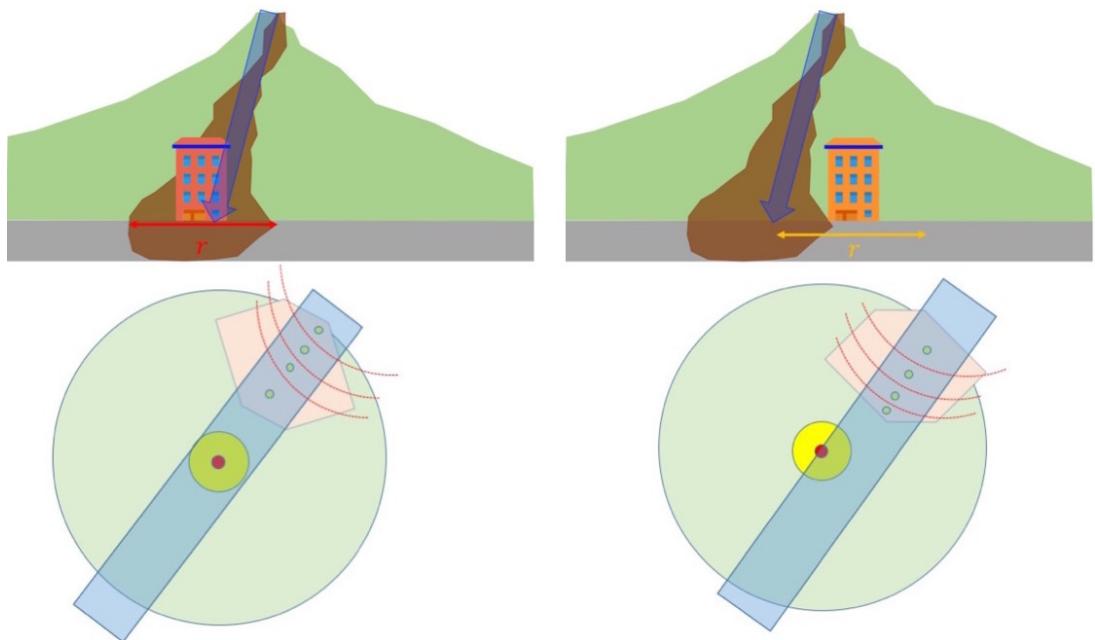


Figure 26: A sinistra il caso di impatto totale a cui corrisponderà una *Impact Factor* pari a 1. A destra il caso di impatto parziale a cui corrisponderà una *Impact Factor* < 1

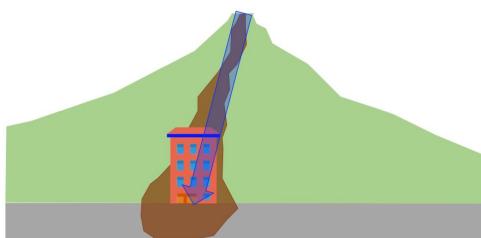


Figure 27: Da Scrivere

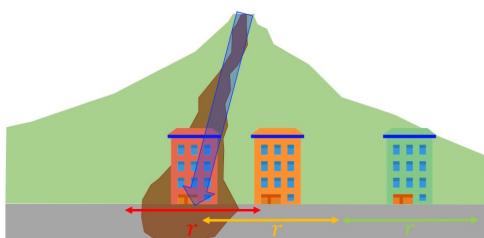


Figure 28: Da Scrivere

PSEUDO-CODICE DEL METODO

Di seguito verrà proposto lo pseudo-codice dell'algoritmo per il calcolo dell'esposizione al rischio di frana di un *Building*. L'algoritmo qui proposto utilizza un approccio funzionale alla risoluzione del problema: Permette infatti di separare la problematica del calcolo dell'exposure in sotto-problemi atomici. Inoltre ciò semplifica notevolmente la scrittura e la comprensione dello stesso.

Algorithmus 1 : Exposure (b_i)

Input : b_i
Output : exp_i

- 1 $\mathcal{NZ}_i \leftarrow \text{NearestZonesFinder}(\mathcal{Z}, b_i, r);$
- 2 $\mathcal{NI}_i \leftarrow \text{NearestIsoipseFinder}(\mathcal{J}, b_i, d);$
- 3 $\mathcal{ZF}_i \leftarrow \text{ZoneFragmentFinder}(b_i, \mathcal{NI}_i, \mathcal{NZ}_i);$
- 4 $\mathcal{BLR}_i \leftarrow \text{LinearRegressionFinder}(b_i, \mathcal{NZ}_i, \mathcal{ZF}_i, \mathcal{NZ}_i, \text{BufferSize});$
- 5 $\mathcal{LS}_i \leftarrow \text{LandSlideFinder}(b_i, \mathcal{LR}_i);$
- 6 $\text{exp}_i \leftarrow \text{ContributionOfLandSlide}(b_i, \mathcal{LS}_i);$
- 7 return exp_i;

La funzione Exposure (b_i) elenca i passi fondamentali per poter calcolare il valore d'exposure dell'edificio b_i. Di seguito verranno esposti i dettagli di tutte le funzioni richiamate.

Algorithmus 2 : NearestZonesFinder(\mathcal{Z} , b_i, r)

Input : b_i, r
Output : \mathcal{NZ}_i

- 1 HazardArea_i $\leftarrow \text{ST_Buffer}(b_i, r);$
- 2 $\mathcal{NZ}_i \leftarrow \text{HazardZone}_i \cap \mathcal{Z};$
- 3 return $\mathcal{NZ}_i;$

1. Crea un buffer (HazardArea_i) di raggio r intorno all'edificio b_i
 2. Interseca tale buffer con le zone nell'insieme \mathcal{Z} . Il risultato è l'insieme contenente le sole zone con intersezione non vuota con il buffer costruito, ovvero l'insieme \mathcal{NZ}_i
-

Algorithmus 3 : NearestIsoipseFinder(\mathcal{J} , b_i, r)

Input : \mathcal{J} , b_i, d
Output : \mathcal{NI}_i

- 1 HazardArea_i $\leftarrow \text{ST_Buffer}(b_i, r);$
- 2 $\mathcal{NI}_i \leftarrow \text{HazardArea}_i \cap \mathcal{J};$
- 3 return $\mathcal{NI}_i;$

1. Crea un buffer (HazardArea_i) di raggio r intorno all'edificio b_i
2. Interseca tale buffer con le zone nell'insieme \mathcal{J} . Il risultato è l'insieme contenente le sole isoipse con intersezione non vuota con il buffer costruito, ovvero l'insieme \mathcal{NI}_i

Algorithmus 4 : ZoneFragmentFinder(b_i)

```

Input :  $b_i, \mathcal{NZ}_i, \mathcal{NJ}_i$ 
Output :  $\mathcal{ZF}_i$ 

1  $\mathcal{ZF}_i \leftarrow \emptyset$ 
2 for each  $nz_{i,j}$  in  $\mathcal{NZ}_i$  do
3    $Tf \leftarrow \emptyset;$ 
4    $Ti \leftarrow nz_{i,j} \cap \mathcal{NJ}_i;$ 
5    $Tf \leftarrow Tf + ST\_split(nz_{i,j}, Ti[1]);$ 
6    $Ti \leftarrow Ti - \{Ti[1]\};$ 
7   while  $Tf$  is not empty do
8      $Ci \leftarrow Tf[1] \cap Ti;$ 
9     if  $Ci$  is not empty then
10        $Tf \leftarrow Tf + ST\_split(Tf[1], Ci[1]);$ 
11        $Ti \leftarrow Ti - \{Ci[1]\};$ 
12        $Tf \leftarrow Tf - \{Tf[1]\};$ 
13     end
14   else
15      $\mathcal{ZF}_i \leftarrow \mathcal{ZF}_i + \{Tf[1]\};$ 
16      $Tf \leftarrow Tf - \{Tf[1]\};$ 
17   end
18 end
19 end
20 return  $\mathcal{ZF}_i;$ 

```

La funzione $\text{ZoneFragmentFinder}(b_i)$ esegue una multi-split tra \mathcal{NZ}_i e l'insieme delle isoipse contenute in \mathcal{NJ}_i . Per multi-split si intende il risultato descritto in figura 17.

1. Si inizializza l'insieme \mathcal{ZF}_i .
2. Per ogni nz_i contenuto nell'insieme \mathcal{NZ}_i entro nel ciclo. L'indice i è bloccato sul building i -esimo mentre la j si muove nell'intervalo $(1, \dots, \text{card}(\mathcal{NZ}_i))$.
3. Si inizializza l'insieme Tf (Temp_fragment) che conterrà i frammenti parziali su cui è possibile eseguire delle split.
4. Si inizializza l'insieme Ti (Temp_isoipse) come il risultato dell'intersezione tra nz_i corrente con gli elementi \mathcal{NJ}_i .
5. Si esegue la split con la prima isoipse contenuta nell'insieme Ti . L'isoipse una volta usata per la split non ne produrrà nessun'altra.
6. Elimino l'isoipse $Ti[1]$ usata nella split precedente.
7. Finché l'insieme Tf non è vuoto significherà che gli elementi sono ulteriormente splittabili.
8. Interseco $Tf[1]$ con Ti . Il risultato genera un nuovo insieme chiamato Ci (Current_isoipses).
9. Se l'insieme Ci non è vuoto vorrà dire che l'elemento $Tf[1]$ è ulteriormente splittabile in quanto intersecato da almeno una isoipse.

10. Eseguo la split ed il risultato lo aggiungo all'insieme T_f .
11. Elimino dall'insieme T_i l'isoipse che ha prodotto lo split.
12. Elimino da T_f il frammento che è stato appena spartito in sotto frammenti.
13. Se C_i è un insieme vuoto significherà che il frammento non è attraversato da nessuna isoipse, per cui non sarà più splittabile.
14. Aggiungo all'insieme $Z\mathcal{F}_i$ l'elemento T_f [1] non più slittabile.
15. Rimuovo dall'insieme T_f il primo elemento che ho appena aggiunto all'insieme dei frammenti atomici non più splittabili.
16. infine restituisco l'insieme $Z\mathcal{F}_i$

Algorithmus 5 : LinearRegressionFinder($b_i, Z\mathcal{F}_i, NZ_i, BufferRadius$)

Input : $b_i, Z\mathcal{F}_i, NZ_i, BufferRadius$

Output : \mathcal{BLR}_i

```

1  $\mathcal{BLR}_i \leftarrow \emptyset;$ 
2 for  $nz_{i,j}$  in  $NZ_i$  do
3    $c nz_{i,j} \leftarrow ST\_centroid(nz_{i,j})$  ;
4   for  $(zf_{i,j,t} \in Z\mathcal{F}_i) \subset nz_{i,j}$  do
5      $C_p \leftarrow C_p + \{ST\_centroid(zf_{i,j,t})\}$  ;
6   end
7    $slope \leftarrow regr\_slope(C_p)$  ;
8    $lr_{i,j} \leftarrow ST\_MakeLine(slope, cnz_{i,j})$  ;
9    $\mathcal{BLR}_i \leftarrow \mathcal{BLR}_i + \{(ST\_Buffer(lr_{i,j}, BufferRadius))\}$  ;
10 end
11 return  $\mathcal{BLR}_i$ ;

```

1. Inizializzo l'insieme \mathcal{BLR}_i .
2. Per ogni elemento all'interno dell'insieme NZ_i . L'indice i è bloccato sul building i -esimo mentre la j si muove nell'intervallo $(1,..,\text{card}(NZ_i))$.
3. Calcolo il centroide $cnz_{i,j}$.
4. Per ogni $zf_{i,j,t}$ contenuto in nz_i (figura 17)
5. Calcolo i centroidi di $zf_{i,j,t}$ e inserisco i punti risultati in un insieme chiamato C_p (Centroid_point) (Figura 18)
6. Calcolo la pendenza della retta di regressione lineare sul insieme dei punti C_p con la primitiva $regr_slope$. Inserisco il risultato in una variabile chiamata $slope$
7. Calcolo la retta $lr_{i,j}$ passante per il punto $cnz_{i,j}$ e coefficiente angolare $slope$.
8. Inserisco in \mathcal{BLR}_i il buffer della retta.
9. Restituisco l'insieme \mathcal{BLR}_i

Algorithmus 6 : LandSlideFinder(b_i, \mathcal{LR}_i)

Input : b_i, \mathcal{LR}_i **Output** : \mathcal{LS}_i

```

1 BuildingBufferi ← ST_Buffer( $b_i, d$ ) ;
2  $\mathcal{LS}_i$  ← ST_intersection(BuildingBufferi,  $\mathcal{LR}_i$ ) ;
3 return  $\mathcal{LS}_i$ ;
```

1. Crea un buffer di raggio d intorno all'edificio b_i
2. Interseca il buffer appena creato con l'insieme \mathcal{LR}_i
3. Ne Risulta un insieme contenente le sole \mathcal{LR}_i intersecate al buffer.(figura 25)

Algorithmus 7 : ContributionOfLandSlide(b_i, \mathcal{LS}_i)

Input : $b_i, \mathcal{LR}_i, \mathcal{NZ}_i$ **Output** : exp_i

```

1  $exp_i \leftarrow 0$  ;
2 BuildingBufferi ← ST_Buffer( $b_i, d$ ) ;
3 for each  $ls_{i,j} \in \mathcal{LS}_i$  do
4     if ST_intersects( $b_i, nz_{i,j}$ ) then
5          $exp_i \leftarrow$  compute equation 5 ;
6     end
7     else
8          $exp_i \leftarrow$  compute equation 6 ;
9     end
10 end
11 return  $exp_i$ ;
```

La funzione ContributionOfLandSlide(b_i, \mathcal{LS}_i) somma tutti i contributi delle **Landslide** che impattano sull'edificio.

1. Inizializzo la variabile exp_i
2. Calcolo un buffer di raggio d intorno al building i -esimo
3. Per ogni elemento contenuto in \mathcal{LS}_i
4. Se il building è situato sull' $nz_{i,j}$ che ha generato $ls_{i,j}$ allora calcolo l'equazione 5 scritta in precedenza.
5. Altrimenti calcolo l'equazione 6 scritta in precedenza
6. Restituisco il valore di exposure del building i -esimo

CASO DI STUDIO

Il metodo proposto è valido per una grande varietà di assets. Tuttavia in questo lavoro è stato testato sulle stazioni e linee ferroviarie del territorio abruzzese. Quest'ultimo infatti è particolarmente soggetto al rischio frana, com'è possibile osservare nella tabella 1, dove i numeri delle frane verificatesi (accanto alla tipologia) sono davvero elevati. Inoltre tali frane sono quelle censite, quindi è molto probabile che il numero reale sia di gran lunga maggiore.

Table 1: Le frane censite verificatesi in Abruzzo.

TIPO DI MOVIMENTO	NUM FRANE	%
Crollo/ribaltamento	128	1,51
Scivolamento rotazionale/traslativo	3.401	40,05
Espansione	2	0,02
Colamento lento	2.364	27,84
Colamento rapido	704	8,29
Sprofondamento	1	0,01
Complesso	1.331	15,67
DGPV	92	1,08
Aree soggette a crolli/ribaltamenti diffusi	63	0,74
Aree soggette sprofondamenti diffusi	6	0,07
Aree soggette a frane superficiali diffuse	257	3,03
n.d. (tipo non determinato)	144	1,69

Per questo motivo il territorio abruzzese è sicuramente un caso di studio interessante. I dataset di partenza, presi in input per il caso di studio sono i seguenti:

1. geo_area: rappresenta la GEOAREA, ovvero la porzione di territorio di interesse;
2. railway_stations: contiene le stazioni appartenenti alla rete ferroviaria abruzzese;
3. railway_routes: contiene le linee ferroviarie abruzzesi;
4. abruzzo_raster: i dati raster del territorio abruzzese.

I files geo_area, railway_stations e railway_routes sono degli shapefile. Il file raster abruzzo_raster è stato utilizzato al fine di ricavarsi le curve di livello dell'intera geo_area. Nell'appendice A è possibile leggere una descrizione dettaglia delle procedura della creazione del dataset, ovvero il modo in cui gli shapefile sono stati importati all'interno della base di dati e poi utilizzati. Per tutta la fase di sperimentazione del metodo è stato utilizzato il software **PostgreSQL** con l'estensione spaziale **PostGIS**. In questo ambiente sono stati creati la base di dati, le viste le user defined function (UDF) e sono stati importati gli shapefile. Il software **QGIS** è stato

utilizzato per visualizzare i dati dalle tabelle della basi dati. Il programma è in grado di visualizzare contemporaneamente, su più layers, i dati provenienti da diverse tabelle. In figura 29 si può vedere il dataset completo utilizzato per il caso di studio. E' possibile riconoscere i confini dell'Abruzzo, le linee ferroviarie evidenziate con segmenti colore verde e infine le stazioni ferroviarie rappresentate con dei pallini di colore rosso.

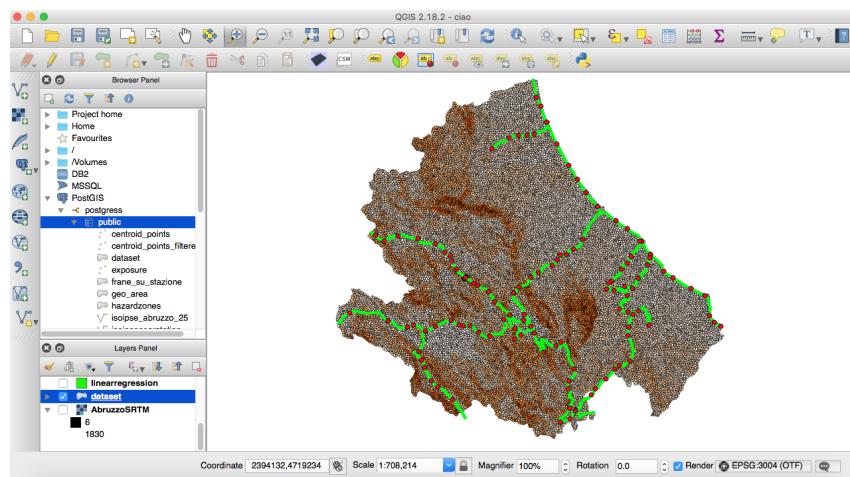


Figure 29: Il software QGis utilizzato per la visualizzazione del dataset.

I dati raster permettono inoltre una visualizzazione 3D del dataset all'interno del browser installato di default sulla macchina. Ciò è possibile attraverso il plugin **Qgis2threejs** di QGIS. Questa estensione è molto utile per la validazione e valutazione del metodo illustrate nella sezione successiva.

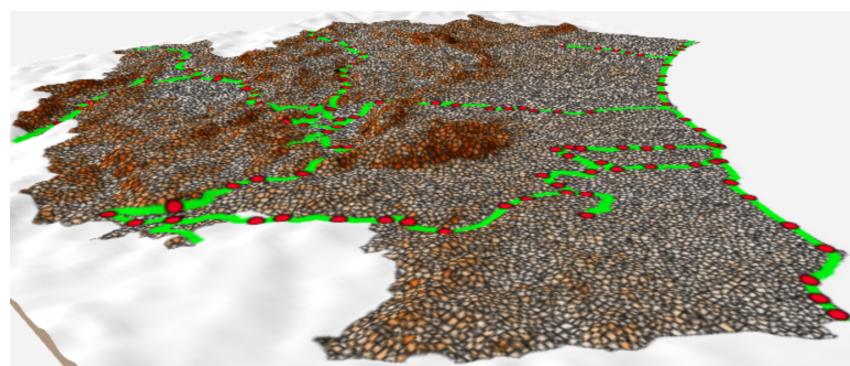


Figure 30: Visualizzazione 3D del dataset attraverso il plugin Qgis2threejs di QGIS

VALIDAZIONE E VALUTAZIONE DEL METODO

Per poter ritenere che il metodo proposto sia effettivamente valido, i risultati prodotti dovranno essere confrontati con dati noti ed attendibili. Questo confronto ci permetterà attraverso delle metriche ben precise di verificare la conformità dei valori calcolati da quelli preconosciuti. I principali metodi di validazione, sfruttano una classificazione dei risultati in classi di appartenenza. Nel caso specifico queste classi rappresentano il livello di pericolosità dei versanti montuosi/collinari rispetto all'edificio preso in esame. Maggiore sarà la pericolosità associata all'edificio, maggiore sarà la possibilità che esso sia coinvolto in un fenomeno franoso.

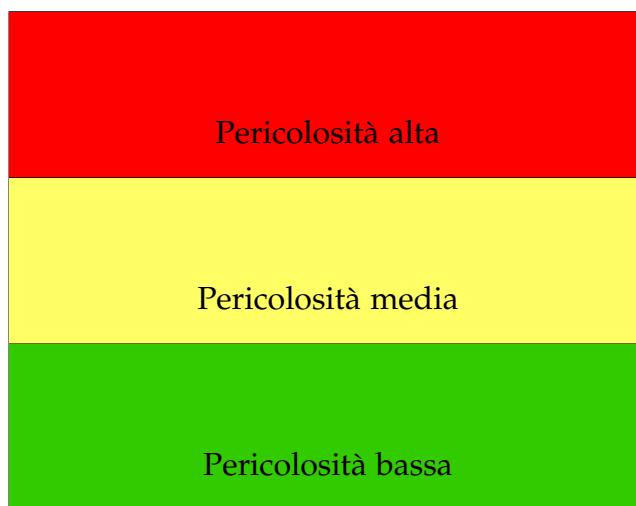


Table 2: In Tabella sono rappresentate le classi di pericolosità che possono essere associate ad un edificio

Essendo i risultati del metodo espressi in un intervallo di valori continui essi andranno mappati all'interno di classi. La problematica nello stabilire il numero di classi necessario a rendere i risultati espressivi non è da sottovalutare in quanto, un'eccessiva divisione in classi, porterebbe a frammentare i risultati in modo sproporzionato, commettendo errori non dovuti al metodo di calcolo ma alla dimensione ridotta degli intervalli su cui sono costruite le classi. Contrariamente, una divisione in poche classi abbasserebbe l'espressività del metodo stesso nel considerare le differenze territoriali importantissime per una valutazione corretta dell'esposizione al rischio. Si è deciso quindi che tre classi sarebbero state il giusto compromesso, consentendoci quindi di distinguere le stazioni a più alto rischio con quelle a rischio più basso non dimenticandoci di quelle intermedie. Il dataset su cui il metodo è stato validato è quello delle stazioni ferroviarie della regione Abruzzo. Ogni stazione presente nel dataset è stata classificata in modo tale da ricadere in una delle tre classi proposte. Tale classificazione "umana" è da considerarsi corretta per definizione in quanto stabilita da esperti. In definitiva, attraverso questo modo di procedere, si potrà verificare se il metodo di calcolo restituisce valori affidabili confrontando i risultati ottenuti con quelli stabiliti dagli esperti. Ottenendo dati statistici sui risultati ottenuti sul dataset delle stazioni ferroviarie, si potrà stimare come l'algoritmo si comporti con un dataset di input più grande. Avendo tale classi-

ficazione si possono usare le metodologie di analisi dei dati che sono fornite nell'ambito del machine learning. Introduciamo quindi una matrice detta matrice di contingenza. Tale strumento mette in relazione le classi di appartenenza reali, con quelle predette dall'algoritmo che si sta esaminando.

La Matrice di contingenza viene calcola nel seguente modo:

1. Si stabilisce un dataset di cui si conoscono le classi di appartenenza.
2. Attraverso l'algoritmo che si sta studiando, per ogni elemento presente nel dataset, si stima la classe di appartenenza.
3. Dai risultati ottenuti si conta:
 - a) Il numero di predizioni esatte per ogni classe.
 - b) Il numero di predizioni sbagliate per ogni classe organizzate secondo le classi di appartenenza stimate.

Questi numeri sono organizzati all'interno di una matrice.

1. Ogni riga rappresenta una classe di appartenenza reale.
2. Ogni colonna rappresenta una classe di appartenenza stimata.

		Classi di appartenenza Predette			Totale
		Classe 1	Classe 2	Classe 3	
Classi di Appartenenza Reali	Classe 1	$N(1,1)$	$N(1,2)$	$N(1,3)$	T_{R1}
	Classe 2	$N(2,1)$	$N(2,2)$	$N(2,3)$	T_{R2}
	Classe 3	$N(3,1)$	$N(3,2)$	$N(3,3)$	T_{R3}
Totale		T_{C1}	T_{C2}	T_{C3}	TOT

Table 3: In Figura è rappresentata una generica matrice di contingenza con tre classi di appartenenza.

Con $N(i,j)$ viene indicato il valore in posizione i,j .

$$T_{R1} = N(1,1) + N(1,2) + N(1,3)$$

$$T_{R2} = N(2,1) + N(2,2) + N(2,3)$$

$$T_{R3} = N(3,1) + N(3,2) + N(3,3)$$

$$T_{C1} = N(1,1) + N(2,1) + N(3,1)$$

$$T_{C2} = N(1,2) + N(2,2) + N(3,2)$$

$$T_{C3} = N(1,3) + N(2,3) + N(3,3)$$

$$TOT = T_{R1} + T_{R2} + T_{R3} = T_{C1} + T_{C2} + T_{C3}$$

Tale modo di organizzare la matrice permette di avere sulla diagonale principale il numero di elementi predetti in modo corretto. Sommando per righe i valori della matrice otteniamo il

numero totale di elementi afferenti ad una classe di appartenenza reale. La somma di tutti i totali corrisponderà al numero di elementi presenti nel dataset.

Esempio:

1. $N(1,1)$: Numero elementi predetti in Classe 1 afferenti in Classe 1 reale.
2. $N(1,2)$: Numero elementi predetti in Classe 2 afferenti in Classe 1 reale.
3. $N(1,3)$: Numero elementi predetti in Classe 3 afferenti in Classe 1 reale.

Da questa matrice possiamo ricavare altre matrici di contingenza dette binarie. Queste matrici per costruzione hanno dimensione 2×2 . Ogni classe reale avrà la sua matrice di contingenza binaria. Il perché di questo passaggio ad una rappresentazione binaria è da ricercare nel significato degli elementi della matrice, infatti essi hanno una carica espressiva maggiore rispetto alla tabella di contingenza non binaria, inoltre permettono di analizzare le classi in modo distinto. Tali matrici sono costruite nel seguente modo:

		Classi di appartenenza stimate		Totale
		Classe 1	No classe 1	
Classi di appartenenza reale	Classe 1	N_1	N_2	T_{R1}
	No classe 1	N_3	N_4	T_{R2}
Totale		T_{C1}	T_{C1}	TOT

Table 4: In tabella è rappresentata la matrice di contingenza binaria della Classe 1.

$$N_1 = N(1,1) ; N_2 = N(1,2) + N(1,3) ; N_3 = N(2,1) + N(3,1) ; N_4 = N(2,2) + N(2,3) + N(3,2) + N(3,3) ;$$

$$T_{C1} = N_1 + N_3 ; T_{C2} = N_2 + N_4 ; T_{R1} = N_1 + N_2 ; T_{R2} = N_3 + N_4 ; TOT = T_{C1} + T_{C2} = T_{R1} + T_{R2}$$

		Classi di appartenenza stimate		Totale
		Classe 2	No classe 2	
Classi di appartenenza reale	Classe 2	N_1	N_2	T_{R1}
	No classe 2	N_3	N_4	T_{R2}
Totale		T_{C1}	T_{C1}	TOT

Table 5: In tabella è rappresentata la matrice di contingenza binaria della Classe 2.

$$N_1 = N(2,2) ; N_2 = N(1,1) + N(1,3) ; N_3 = N(1,2) + N(3,2) ; N_4 = N(1,1) + N(1,3) + N(3,1) + N(3,3) ;$$

$$T_{C1} = N_1 + N_3 ; T_{C2} = N_2 + N_4 ; T_{R1} = N_1 + N_2 ; T_{R2} = N_3 + N_4 ; TOT = T_{C1} + T_{C2} = T_{R1} + T_{R2}$$

Questo modo di rappresentare i dati sarà molto utile per ricavare delle misure assolute su parametri utili a stabilire l'effettiva bontà dell'algoritmo nel trovare risultati esatti. Definiamo quindi 4 quantità:

1. Veri Positivi (**VP**). Tale quantità denota i casi nei quali l'algoritmo ha riconosciuto correttamente la classe di appartenenza.
2. Falsi Positivi (**FP**). Tale quantità denota i casi nei quali l'algoritmo ha riconosciuto erroneamente la classe di appartenenza.
3. Veri Negativi (**VN**). Tale quantità denota i casi nei quali l'algoritmo ha riconosciuto correttamente che un elemento non appartiene ad una certa classe.
4. Falsi Negativi (**FN**). Tale quantità denota i casi nei quali l'algoritmo non ha riconosciuto correttamente la classe di appartenenza di un elemento scambiandola con un'altra.

		Classi di appartenenza stimate		Totale
		Classe 3	No classe 3	
Classi di appartenenza reale	Classe 3	N_1	N_2	T_{R1}
	No classe 3	N_3	N_4	T_{R2}
Totale		T_{C1}	T_{C1}	TOT

Table 6: In tabella è rappresentata la matrice di contingenza binaria della Classe 3.

$$N_1 = N(3,3) ; N_2 = N(1,1) + N(1,2) ; N_3 = N(3,1) + N(3,2) ; N_4 = N(1,1) + N(1,2) + N(2,1) + N(2,2) ;$$

$$T_{C1} = N_1 + N_3 ; T_{C2} = N_2 + N_4 ; T_{R1} = N_1 + N_2 ; T_{R2} = N_3 + N_4 ; TOT = T_{C1} + T_{C2} = T_{R1} + T_{R2}$$

		Classi di appartenenza stimate		Totale
		Classe Generica	No classe Generica	
Classi di appartenenza reale	Classe Generica	VP	FN	$P = VP + FN$
	No classe Generica	FP	VN	$N = FP + VN$
Totale		VP + FP	FN + VN	$P + N$

i

Table 7: Nella tabella sono rappresentate le quantità sopra definite e le posizioni che occupano all'interno della matrice di contingenza binaria

Tali valori vanno ricercati nella matrice di contingenza binaria. Essi per costruzione della matrice saranno esattamente i quattro elementi di cui è composta.

Definiamo inoltre delle metriche di valutazione.

- True Positive Rate (**TPR**)

$$TPR = \frac{VP}{VP + FN} \quad (8)$$

- True Negative Rate (**TNR**)

$$TNR = \frac{VN}{FP + VN} \quad (9)$$

- Precision (**P**)

$$P = \frac{VP}{VP + FP} \quad (10)$$

- Accuracy (**Acc**)

$$Acc = \frac{VP + VN}{(VP + FN) + (VN + FN)} \quad (11)$$

Avendo introdotto tali strumenti matematici per fare la validazione, di seguito verranno illustrati i risultati ottenuti dall'esecuzione del metodo di calcolo sul dataset delle stazioni ferroviarie abruzzesi. Come prima cosa si dovranno definire gli intervalli a cui le classi sono associate. Il metodo restituisce valori continui di exposure per cui sono state definite i seguenti intervalli che corrisponderanno alle classi definite in precedenza. Il valore minimo restituito dal metodo è 0. Esso corrisponde ad una situazione per nulla pericolosa. Il valore massimo invece è 1,8. Tale valore corrisponde ad un' situazione ad altro rischio. Tali intervalli sono i seguenti:

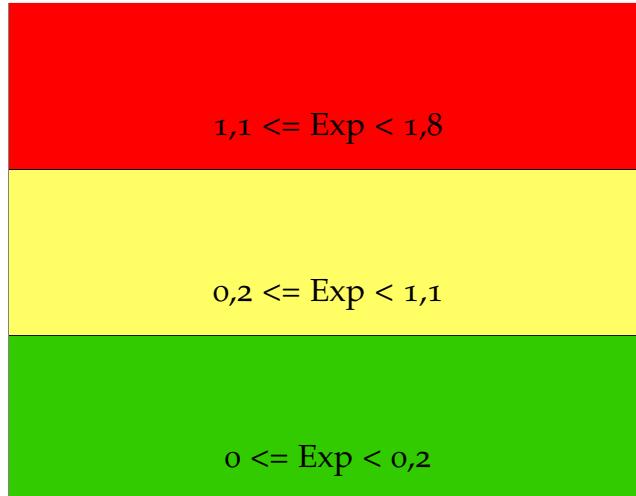


Table 8: In Tabella sono rappresentati gli intervalli di pericolosità

Di seguito verrà fornita una tabella che riassume i risultati ottenuti dall'esecuzione del metodo sul dataset delle stazioni. La tabella 9 contiene le seguenti informazioni:

1. La colonna "Stazione" rappresenta la lista delle stazioni ordinate secondo il valore di exposure fornito dagli esperti. I colori indicano le classi di appartenenza reali.
2. La colonna "Exp" rappresenta la lista dei valori di exposure che l'algoritmo ha calcolato. I colori indicano le classi di appartenenza stimate.

Partendo da questa rappresentazione tabulare si possono costruire le matrici di contingenza introdotte precedentemente.

Possiamo osservare dalla tabella 10 che i valori più alti sono concentrati lungo la diagonale principale della matrice di contingenza. Questo fenomeno si genera perché nella diagonale sono presenti il numero di valori stimati correttamente. Si può anche osservare che l'algoritmo non sbaglia mai assegnando una pericolosità alta ad una stazione a bassa pericolosità, stessa cosa per il viceversa infatti, dalla tabella, sia la classe bassa sia la classe alta, hanno valori nulli in posizione N(1,3) ed N(3,1). Da tale tabella si possono ricavare le matrici di contingenza binarie. Possiamo osservare dalla tabella 11 che l'algoritmo riesce a riconoscere 38 stazioni su 47 in fascia di pericolo bassa. Per quanto riguarda i falsi negativi esso restituisce 9 casi. L'algoritmo categorizza in modo scorretto 8 stazioni su 67 stimando la classe di appartenenza come più bassa rispetto alla realtà. Infine riesce a riconoscere 59 casi su 67 che la stazione è in fascia più alta.

Per la tabella 12 si possono trarre conclusioni simili a quelle viste per la matrice di contingenza binaria descritta in tabella 11. Si può notare come l'algoritmo categorizza in modo corretto la maggior parte delle stazioni, ben 52 su 61.

L'algoritmo, essendo stato pensato per individuare i casi pericolosi, dovrà restituire, a maggior ragione, dei risultati validi per la classi di pericolosità alta. Come possiamo vedere dalla tabella 13 vengono categorizzati correttamente i casi che effettivamente sono pericolosi. Si può notare come in un unico caso l'algoritmo confonde la classe stimata di appartenenza. Tale caso verrà analizzato più approfonditamente in seguito. Si può osservare dalla tabella che l'algoritmo non sbaglia, ad eccezione di un caso, di categorizzare come non altamente pericolosa una stazione (N(2,2)). Ci sono infatti 107 casi su 108 in cui l'algoritmo stima correttamente.

Di seguito viene proposta la tabella che riassume i risultati ottenuti dal calcolo delle metriche definite in precedenza.

Dai risultati ottenuti e riassunti in tabella 14 possiamo trarre alcune conclusioni. Il TRN misura la percentuale di positivi che sono correttamente identificati come tali, per cui, possiamo vedere che mediamente l'algoritmo si classificano correttamente per più dell'80% dei casi le classi di appartenenza corrette. Il TNR misura la percentuale di negativi che siano correttamente identificate come tali. Possiamo osservare come anche in questo caso l'algoritmo non scende mai al disotto del 83%. Sia per quanto riguarda la precisione che l'accuratezza l'algoritmo si comporta in linea agli altri valori. Da notare che la classe di pericolosità più alta a valori molto alti di TNR e Acc. Questo fenomeno è dovuto al fatto che nel dataset ci sono poche stazioni ad alto rischio rispetto al numero associato agli altri rischi di pericolosità.

I risultati appena discussi confermano la bontà del metodo proposto. Ciò nonostante è stata eseguita indagine approfondita sulle stazioni la cui classificazione risulta errata. L'obiettivo è quello di comprendere se ci sono ulteriori margini di miglioramento del metodo. I falsi negativi analizzati sono i seguenti: Popoli - Vittorito, Canistro e Fontecchio. Vengono considerati falsi negativi in quanto la loro classificazione non coincide con quella proposta dagli esperti. Inoltre queste 3 stazioni sono state scelte in quanto i rispettivi valori di exposure si discostano di più dal valore dell'estremo superiore della classe di appartenenza proposta dagli esperti. Ad esempio la stazione di Popoli - Vittorito risulta di classe 1 secondo il giudizio degli esperti. Il metodo però la classifica come classe 2. Per questo motivo è un falso negativo. Inoltre per assegnare alla classe 1 il valore di exposure deve essere inferiore a 0.2, mentre il valore restituito dal metodo è pari a 0.42. L'errore è quindi pari a 0.22 e risulta il più elevato tra tutti i falsi negativi della classe 1. Lo stesso ragionamento è stato applicato alla scelta delle altre 2 stazioni prese in considerazione.

1. Popoli - Vittorito: in figura 31 dalle prime due immagini è possibile osservare che la stazione non ha pendii importanti nelle immediate vicinanze, in quanto si trova in una zona prettamente pianeggiante. Nell'ultima immagine è possibile vedere, colorate in viola, le $nz_{i,j}$, con le rispettive landslides rappresentate con dei rettangoli colorati con i contorni evidenziati in nero. La diversa colorazione ha l'unico scopo di differenziare ulteriormente, esclusivamente a livello visivo, tra loro le landslides. Apparentemente il metodo sembra funzionare correttamente, ma in realtà la landslide di colore rosso ha una traiettoria inesatta. Come è possibile osservare in figura 32 le curve di livello all'interno del $nz_{i,j}$ non sono numerose e in più hanno un andamento molto particolare. Di conseguenza i centri di massa $czf_{i,j,t}$ degli zone fragments $zf_{i,j,t}$ non sono ben allineati ma sparsi nella $nz_{i,j}$. Questo comporta che l'interpolazione della retta è molto imprecisa e quindi la traiettoria finale non è quella giusta. In questo caso particolare la landslide interseca il BuildingBuffer_i e quindi contribuisce all'aumento dell'exposure della stazione. Ciò comporta il passaggio dalla classe 1 alla classe 2 e quindi ad una valutazione errata.
2. Canistro: dalle prime due immagini della figura 33 si nota che la stazione si trova in una vallata. Nell'ultima immagine si vede che le landslides che impattano sulla stazione sono 4, 2 delle quali però non sono corrette. La landslide di colore rosso è affetta dagli stessi problemi di traiettoria non corretta emersi nella stazione precedente. Inoltre la landslide di colore giallo è troppo larga. Infatti nella figura 34 è possibile vedere che la landslide è più larga dello $nz_{i,j}$ da cui parte. Ciò accade perché la larghezza della landslides è pro-

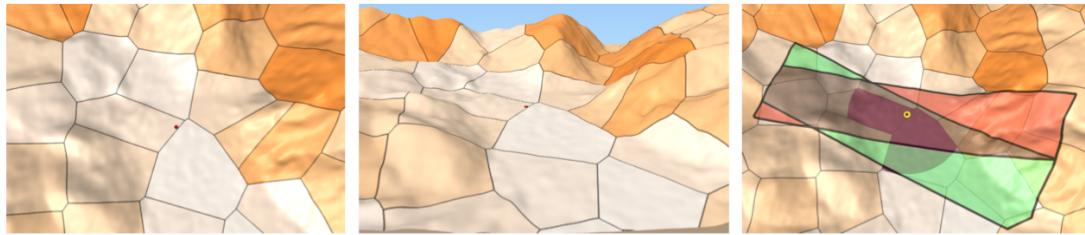


Figure 31: Da scrivere



Figure 32: Da Scrivere

porzionata al perimetro della $nz_{i,j}$. Quest'ultima non è però di forma quadrata, ovvero non ha i lati tutti della stessa lunghezza. Per questo motivo bisognerebbe fare attenzione all'orientamento della landslide rispetto alla $nz_{i,j}$. In questo caso la landslide è perpendicolare ai lati "corti" della $nz_{i,j}$. Ruotandola la landslide di 90 gradi in senso orario risulterebbe perpendicolare rispetto ai lati più lunghi e quindi la larghezza sarebbe corretta. Questo aspetto ha una notevole rilevanza sull'equazione 4 dell'impact factor. Infatti l'area dell'intersezione tra il $BuildingBuffer_i$ e la landslide $ls_{i,j}$ risulta maggiorata, contribuendo all'aumento dell'exposure e di conseguenza ad una classificazione più alta.

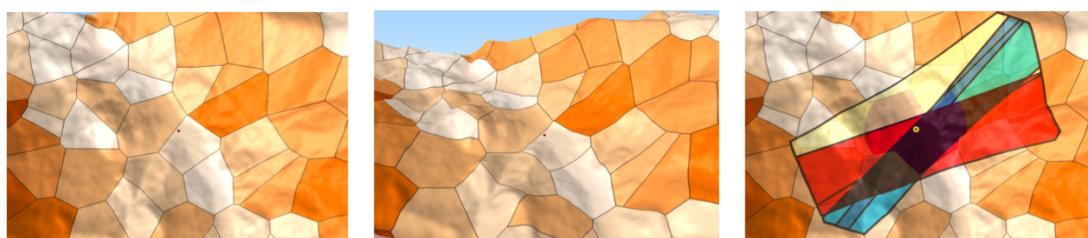


Figure 33: Da scrivere.

3. Fontecchio: dalle prime due immagini della figura 35 si nota che la stazione si trova vicinissima ai piedi di una montagna. Nell'ultima immagine è possibile osservare il numero elevato di landslides che impattano sulla stazione. A differenza di Popoli-Vittorito e Canistro, la stazione di Fontecchio è un falso negativo che il metodo classifica con

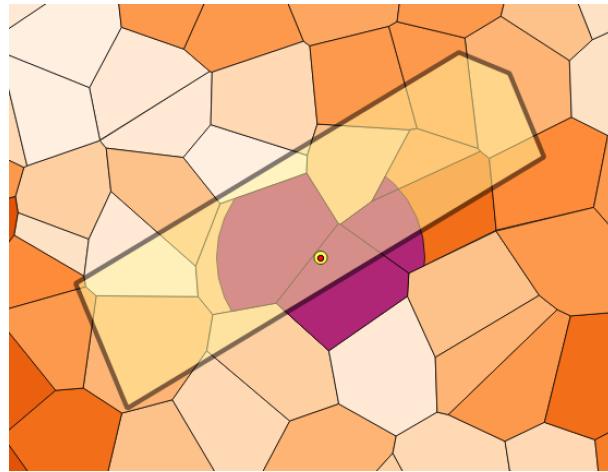


Figure 34: Da Scrivere

una classe più bassa rispetto a quella proposta dagli esperti. In questo caso le landslides hanno una traiettoria corretta e dimensione vicinissima a quella della $nz_{i,j}$ da cui partono. Per questo motivo la classificazione errata è legata al dataset, ovvero gli sz_k associati alle $nz_{i,j}$ hanno valori troppo bassi.

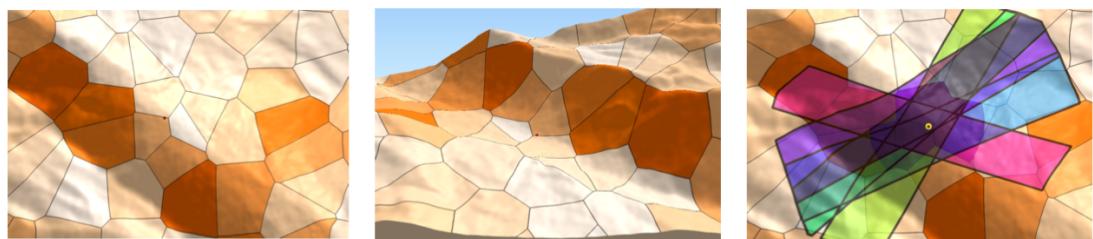


Figure 35: Da scrivere

Stazione	Exp	Stazione	Exp	Stazione
Alba Adriatica - Nereto - Controguerra	0,00	Roseto degli Abruzzi	0,00	Pescina
Tortoreto	0,00	San Pietro Avellana	0,00	Aielli
Giulianova	0,09	Archi	0,35	Celano - Ovindoli
Pineto - Atri	0,11	Perano	0,32	Paterno - San Pelino
Montesilvano	0,00	Lanciano	0,07	Cappelle - Maglianica
Pescara Centrale	0,00	Villa Caldari	0,02	Tagliacozzo
Casalbordino - Pollutri	0,01	Selceroli	0,17	San Vito - Lanciano
Porto di Vasto	0,01	Arielli	0,08	Carsoli
San Salvo	0,00	Vasto	0,01	Capistrello
Teramo	0,12	Balsorano	0,48	Canistro
San Nicolò a Tordino	0,04	Silvi	0,30	Civitella Roveto
Bellante - Ripattoni	0,20	Francavilla al Mare	0,31	Civita d'Antino - M.
Notaresco	0,07	Foro	0,11	Morreia - Castronovo
Mosciano Sant'Angelo	0,00	Ortona	0,29	Ateleta
Chieti	0,05	Ortona - Sangritana	0,33	Gamberale
Brecciarola	0,13	San Vito - Lanciano	0,19	Quadri
Rosciano	0,07	Fossacesia	0,08	Civiluparella
Alanno	0,25	Torino di Sangro - Paglieta	0,29	Villa Santa Maria
Scafa - San Valentino - Caramanico Terme	0,26	Vasto - San Salvo	0,19	Bomba
Popoli - Vittorito	0,42	Manoppello	0,12	Isca d'Archi
Pratola Peligna	0,14	Torre de'Passeri	0,37	Atessa
Sulmona	0,14	Tocco - Castiglione	0,48	Altino
Sulmona - Introdacqua	0,07	Bussi	0,98	Casoli
Rivisondoli - Pescocostanzo	0,16	Cansano	0,63	Sant'Eusonio del S.
Montenero Valcocchiara	0,02	Campo di Giove	0,58	Crocetta
Castel di Sangro 1	0,26	Palena	0,83	Castel Frentano
Castel di Sangro 2	0,18	Roccaraso	0,64	Treglio
Pratola Peligna Superiore	0,00	Alfedena - Scontrone	0,45	San Vito Chietino
Villa Sant'Angelo	0,01	Raiano	0,34	Orsogna
San Demetrio ne' Vestini	0,11	Molina Aterno	0,59	Filetto
Fossa	0,00	Beffi	0,70	Guardiagrele
Paganica	0,00	Tione degli Abruzzi	0,76	San Vincenzo
Sassa - Tornimparte	0,00	Fagnano - Campana	0,49	Vigliano d'Abruzzo

		Classi di appartenenza Predette			Totale
		Bassa	Media	Alta	
Classi di Appartenenza Reali	Bassa	39	8	0	47
	Media	7	53	1	61
	Alta	0	1	5	6
Totale		46	62	6	114

Table 10: Nella tabella viene mostrata la matrice di contingenza ricavata a partire dai risultati ottenuti dal metodo di calcolo

		Classi di appartenenza Predette		Totale
		Bassa	No bassa	
Classi di Appartenenza Reali	Bassa	39	8	47
	No bassa	7	60	67
Totale		46	68	114

Table 11: In tabella viene mostrata la matrice di contingenza binaria della classe a bassa pericolosità ricavata a partire dalla tabella di contingenza non binaria

		Classi di appartenenza Predette		Totale
		Media	No Media	
Classi di Appartenenza Reali	Media	53	8	61
	No media	9	44	53
Totale		62	52	114

Table 12: In tabella viene mostrata la matrice di contingenza binaria della classe a media pericolosità ricavata a partire dalla tabella di contingenza non binaria

		Classi di appartenenza Predette		Totale
		Alta	No alta	
Classi di Appartenenza Reali	Alta	5	1	6
	No alta	1	107	108
Totale		6	108	114

Table 13: In tabella viene mostrata la matrice di contingenza binaria della classe ad alta pericolosità ricavata a partire dalla tabella di contingenza non binaria

	Classe a rischio di pericolosità bassa	Classe a rischio di pericolosità medio	Classe a rischio di p
TPR	0,83	0,87	0,83
TNR	0,90	0,83	0,99
P	0,85	0,85	0,83
Acc	0,87	0,85	0,98

Table 14: In tabella, per ogni fascia di pericolo, sono esposti i risultati delle metriche introdotte in precedenza (True Positive Rate, True Negative Rate, Precision, Accuracy).

ESTENSIONE DEL METODO ALLE LINEE FERROVIARIE

Il metodo proposto è stato esteso sulle linee ferroviarie.

Definizioni

1. \mathcal{R} (**routes**) = $\{r_k | k = 1, \dots, \text{card}(\mathcal{R}) | r_k \text{ è una route ubicata all'interno dei confini della GeoArea}\}$. Per routes si intende una generica tratta (come ad esempio una ferrovia, un'autostrada) descritta dalla tupla $<ID, Name, geometry>$. Il campo *ID* identifica univocamente la tratta; *Name* una sua etichetta nominale *geometry* rappresenta una geometria che descrive sul territorio la tratta.
2. \mathcal{SR} (**Segmentize Routes**) = $\{sr_{k,p} | k = 1, \dots, \text{card}(\mathcal{R}), p = 1, 2, \dots | sr_{k,p} \text{ è una segmentize route ubicata all'interno dei confini della GeoArea}\}$. Per segmentize routes si intende una generica sotto-tratta di un elemento r_k dell'insieme \mathcal{R} . L'unione di tutti gli $sr_{k,p}$ restituisce l'elemento r_k .
3. **EXPSR (Exposure Segmentize Routes)** = $\{\text{expsr}_{k,p} | k = 1, \dots, \text{card}(\mathcal{R}), p = 1, 2, \dots | \text{expsr}_{k,p}$ è il valore di *exposure* del segmento di tratta della routes k -esima. Per *exposure* si intende il valore numerico che indica quanto il segmento di tratta è esposto al rischio frana.

PROGETTAZIONE DELLA BASE DI DATI SPAZIALE

I risultati ottenuti dal metodo devono poter essere consultati agevolmente per facilitare le decisioni riguardo la prevenzione del territorio. Questa esigenza si traduce nella creazione di una basi di dati geografica che riassume i risultati ottenuti dal metodo. Sfruttando le tecnologie messe a disposizione dai sistemi GIS (Geographical Information System), sarà possibile valutare i risultati presenti nella base di dati attraverso il rendering dei dati su una mappa geografica. Tale visualizzazione contestualizza meglio i risultati rendendoli più chiari. Questa motivazione esprime l'importanza di una base dati. Verrà quindi illustrata la sua progettazione attraverso le tecniche tipiche di questo ambito.

Si è progettata quindi la base di dati seguendo i seguenti passi progettuali.



Figure 36: In figura vengono brevemente mostrati i passi concettuali nella progettazione di una base di dati. In blu vengono rappresentate le fasi concettuali. In rosso vengono rappresentati i vari schemi e descrizioni che si acquisiscono in output alla fine di ogni fase progettuale

Come primo passo si devono trasporre le definizioni e le notazioni introdotte in concetti propri delle basi di dati. Definiamo quindi una relazione:

1. geoarea \leftarrow GEOAREA
2. dataset \leftarrow Z
3. isoipse_abruzzo_25 \leftarrow \mathcal{I}
4. raylway_station \leftarrow \mathcal{B}
5. station_exposure \leftarrow \mathcal{EXP}
6. railway_routes \leftarrow \mathcal{R}
7. railway_routes_segment \leftarrow \mathcal{SR}
8. segment_exposure \leftarrow $\mathcal{EXP}SR$

Tale relazione definisce le entità. Esse rappresentano concetti complessi e di rilievo che descrivono classi di oggetti con esistenza autonoma. Un'istanza di un' entità è un oggetto della classe rappresentata. Un'entità ha un nome univoco all'interno dello schema concettuale e viene rappresentata nel diagramma ER con un rettangolo con il nome dell'entità all'interno. Per ogni entità vengono mostrate

1. Gli attributi descritti da un pallino vuoto;
2. L'attributo identificante descritto da un pallino nero pieno;

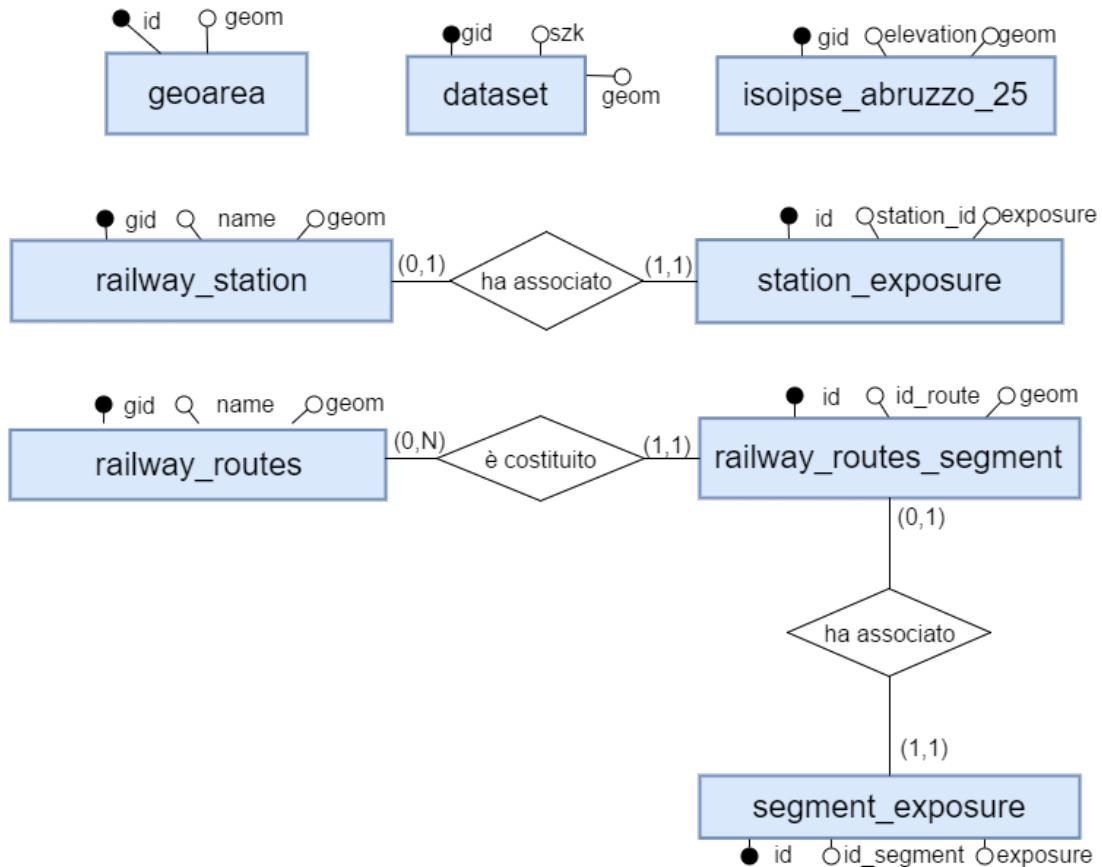


Figure 37: In figura è rappresentato il diagramma Er della base di dati. sono presenti 3 associazioni ed 8 entità.

3. I vincoli di partecipazione nelle relazioni, con notazione (min,max).

Dallo schema Concettuale si è progettato quello logico nel seguente modo.

1. geoarea(id,geom)
2. dataset(gid,szk,geom)
3. isoipse_abruzzo_25(gid,elevation,geom)
4. railway_station(gid,name,geom)
5. station_exposure(id,station_id,exposure)
 - a) FK: station_id REFERENCES railway_station
6. railway_routes(gid,name,geom)
7. railway_routes_segment(id,id_route,geom)
 - a) FK: id_route REFERENCES railway_routes
8. segment_exposure(id,id_segment,exposure)
 - a) FK: id_segment REFERENCES railway_routes_segment

Per quanto riguarda lo schema fisico esso verrà esposto in appendice.

Entità	Descrizione	Attributi	Attributi Identificanti
geoarea	Entità relativa al territorio di interesse, nel nostro caso la regione Abruzzo	id geom	id
dataset	Entità relativa al territorio partizionato della regione abruzzo	gid szk geom	gid
isoipse_abruzzo_25	Entità relativa alle isoipse contenute all'interno del territorio abruzzese.	gid elevation geom	gid
railway_station	Entità relativa alle stazioni presenti sul suolo della regione di interesse	gid name geom	gid
station_exposure	Entità relativa all'esposizione al rischio frana da parte di una stazione	id station_id exposure	id
railway_routes	Entità relativa alle linee ferroviarie della regione di interesse	gid name geom	gid
railway_routes_segment	Entità relativa ai vari segmenti che compongono le linee ferroviarie	id id_route geom	id
segment_exposure	Entità relativa all'esposizione al rischio di frana da parte di un segmento che compone una linea ferroviaria	id id_segment exposure	id

Table 15: In tabella sono descritte le entità in modo dettagliato

Associazione	Descrizione	Entità coinvolte
ha associato	La seguente relazione evidenzia che una stazione può avere associato un valore di exposure. Stessa considerazione si può fare per i segmenti	railway_station<>station_exposure railway_routes_segment<>segment_exposure
è costituito	La seguente relazione evidenzia che una route può essere associata con N segmenti. Un segmento può appartenere ad una sola route	railway_routes<>railway_routes_segment

Table 16: tab:associazioni

Part III
APPENDIX

COSTRUZIONE DEL DATASET

Nel ambito dei sistemi informatici geografici hanno una rilevanza importantissima i dataset di input. Su di essi infatti possono essere fatte le elaborazioni utili alla risoluzione e all'analisi di problematiche territoriali. Tali dati geografici purtroppo sono difficili da reperire e molto spesso non esistono. Questa difficoltà si traduce nel dover costruire il dataset. La costruzione può essere fatta in due modalità:

1. Studio sul campo. Esso prevede di andare effettivamente sul territorio a campionare i dati. Tale metodo risulta molto costoso e lento.
2. Costruzione sintetica del dataset. Questo modo di realizzazione usa dei dati preesistenti e attraverso elaborazioni restituisce un dataset conforme alle specifiche.

Nel caso in esame non esistendo un dataset contenente i dati relativi all'esposizione al pericolo frane della regione Abruzzo, si è deciso di costruirne uno. Tale costruzione è stata fatta tenendo conto di diverse specifiche:

1. Partizionamento completo. Tutto il territorio della regione Abruzzo deve essere partizionato in modo tale da non avere alcuna zona priva di un valore di esposizione al rischio. Ogni partizione verrà chiamata *Zone*.
2. Ogni *Zone* deve avere associato un valore che esprime la pericolosità di quella particella rispetto al pericolo che essa frani. Tale valore è chiamato *Szk*
3. Ogni *Zone* deve avere un area minima di 500 m^2

Si è quindi proposta una prima soluzione. Essa verrà esplicitata seguendo una lista di passi.

1. Creazione di una griglia di 155×155 elementi di lato 1 km^2 (fig.38)
2. Intersezione della griglia con la regione Abruzzo (fig.39)

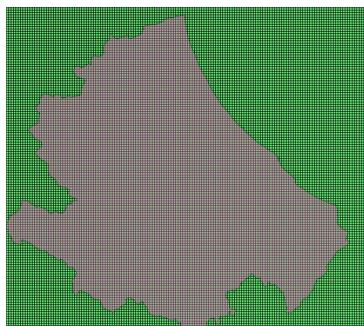


Figure 38



Figure 39

In primo momento si è proceduto ad assegnare un valore di *Szk* casuale. Tale valore però non avrebbe avuto una corrispondenza con la realtà e quindi qualsiasi risultato ottenuto con

qualsiasi metodo sarebbe stato impossibile da validare. Si è quindi deciso che tale valore doveva essere deciso secondo una logica ben precisa. Non avendo uno studio geologico completo di tutto il territorio abruzzese si è deciso di tenere in considerazione una sola variabile che contribuisce agli eventi franosi, ovvero la conformazione del suolo. Essa viene espressa attraverso delle curve di livello. Esse rappresentano tutti i punti a stessa quota sul territorio. La semplificazione ad una sola variabile permette di poter fare una validazione secondo una metrica ben precisa. Le curve di livello sono state ottenute a partire dai dati DEM del territorio abruzzese. Tali dati sono stati reperiti sul portale USGS (United States Geological Survey). L'USGS è la maggiore agenzia per la cartografia civile degli Stati Uniti. I dati prelevati da tale portale sono quelli della missione NASA SRTM. Lo Shuttle Radar Topography Mission (SRTM) è un'impresa internazionale che è riuscita ad ottenere un Modello digitale di elevazione su una scala quasi globale.

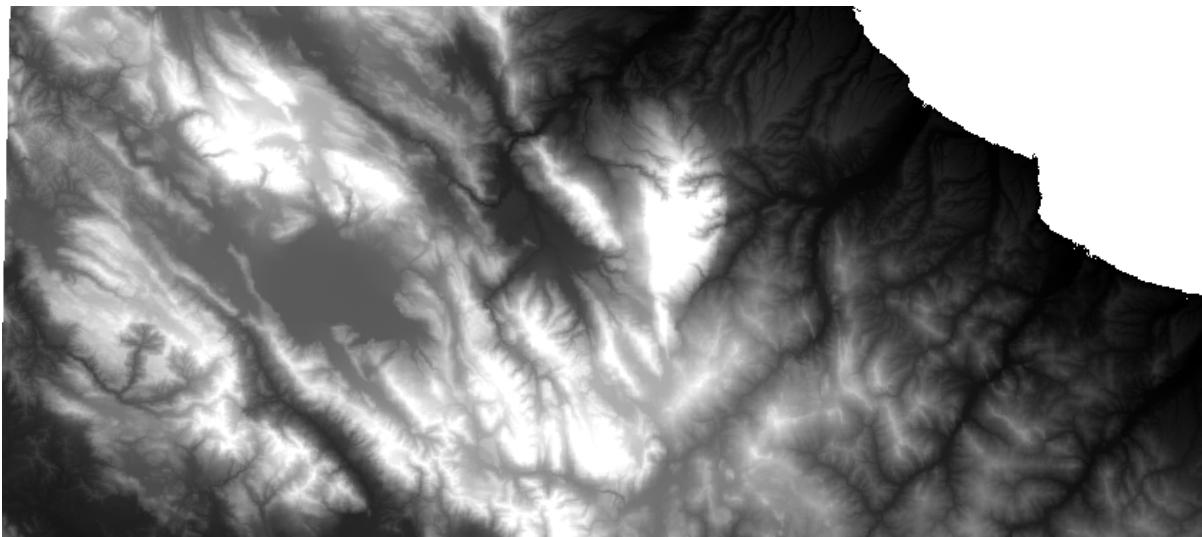


Figure 40: In Figura si può osservare l'immagine raster del territorio abruzzese ottenuta dalla missione SRTM

Lo SRTM consisteva in un sistema radar modificato che ha volato a bordo dello Space Shuttle Endeavour durante gli 11 giorni della missione STS-99 del febbraio 2000. Per acquisire i dati topografici dei dati di elevazione, il carico SRTM è stato equipaggiato con due antenne radar. Un'antenna era posizionata nello spazio di carico dello Shuttle, l'altra alla fine di un braccio di 60 metri che si estendeva dallo spazio di carico una volta che lo Shuttle era nello spazio. La tecnica impiegata è conosciuta come Interferometric Synthetic Aperture Radar. I modelli di elevazione ricavati dai dati dello SRTM vengono usati nei Geographic information system (GIS). Possono essere liberamente scaricati tramite internet ed il loro formato di file è supportato da parecchi programmi software. Sono stati prelevati i dati raster della regione Abruzzo e attraverso QGis (Software per la visualizzazione e l'elaborazione di dati geografici) sono state generate le curve di livello. Tali curve hanno una risoluzione di 25m. Infine sono state esportate come shape file. Avendo introdotto le curve di livello (Isoipse) si sono decisi i valori di Szk tenendo conto della morfologia effettiva del terreno. Per far ciò si è semplicemente contato il numero di curve di livello che intersecano ogni Zone. Maggiore sono il numero di intersezioni maggiore sarà la ripidità del suolo quindi maggiore sarà il pericolo frana. Come detto in precedenza questa è una semplificazione estrema in quanto nella realtà ci sono

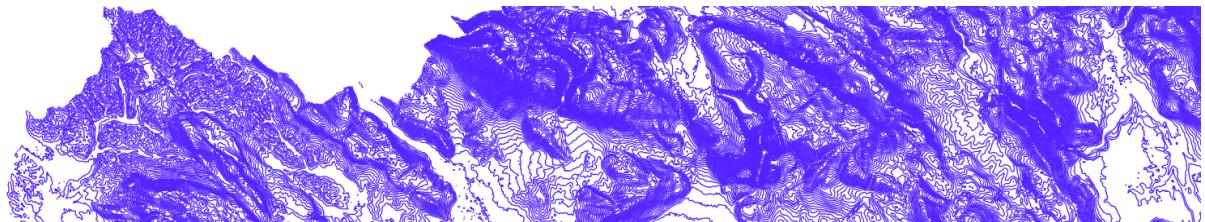


Figure 41: In figura si possono osservare le curve di livello di una porzione di territorio della regione Abruzzo

moltissimi altri fattori che scatenano una frana. Si è quindi ottenuto un dataset che rimarcasse attraverso gli *Szk* l’andamento del terreno. Tale dataset ha due grandi limiti.

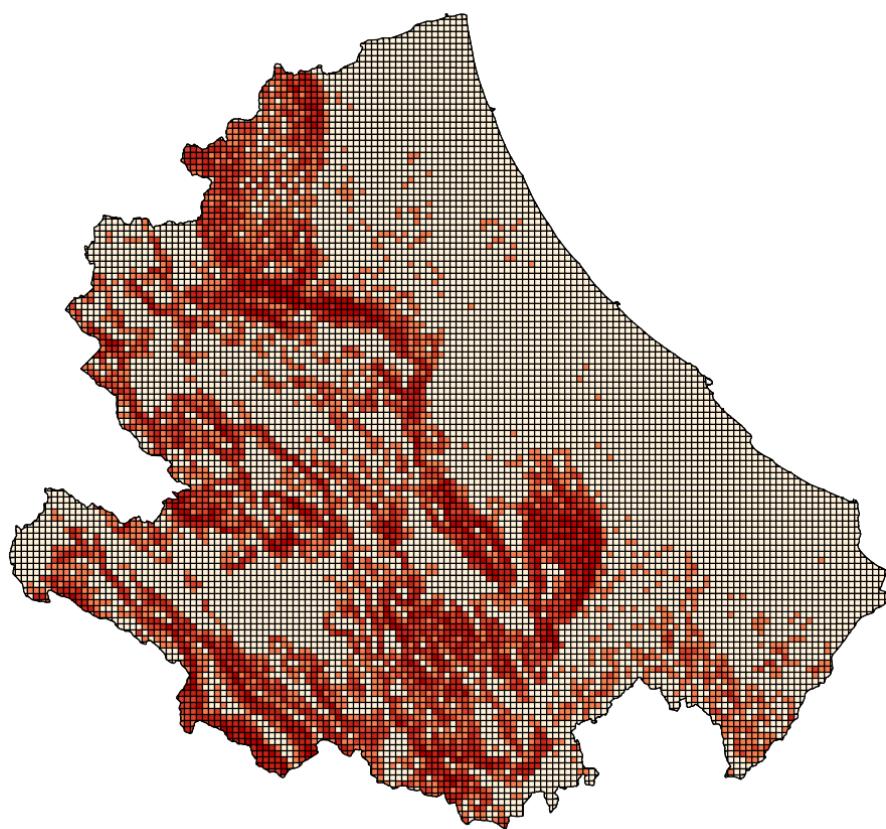


Figure 42: In figura si può osservare il dataset ottenuto tenendo conto delle curve di livello

1. Estrema regolarità delle *Zone*
2. Aree estremamente piccole lungo i bordi

Per quanto riguarda il secondo limite si è risolto il problema aggregando le aree piccole a quelle vicine a loro. Essendo il primo limite troppo vincolante si è deciso di cambiare modello di generazione delle *Zone*. Si è quindi provato ad usare il modello di Voronoi. In matematica, un diagramma di Voronoi (dal nome di Georgij Voronoi), anche detto tassellatura di Voronoi, decomposizione di Voronoi, o tassellatura di Dirichlet (dal nome di Lejeune Dirichlet) è un

particolare tipo di decomposizione di uno spazio metrico determinata dalle distanze rispetto ad un determinato insieme discreto di elementi dello spazio (ad esempio, un insieme finito di punti). Il risultato ottenuto dall'unione della decomposizione di Voronoi con le curve di livello ha prodotto un risultato davvero soddisfacente. Il limite delle aree regolari si è superato in quanto tale tassellatura genera aree tutte differenti tra di loro. Mentre il limite delle aree molto piccole sui bordi non si presenta affatto. Per aumentare la variazione S_{zk} in pianura si è deciso di aggiungere un valore casuale. In definitiva attraverso le curve di livello ricavate

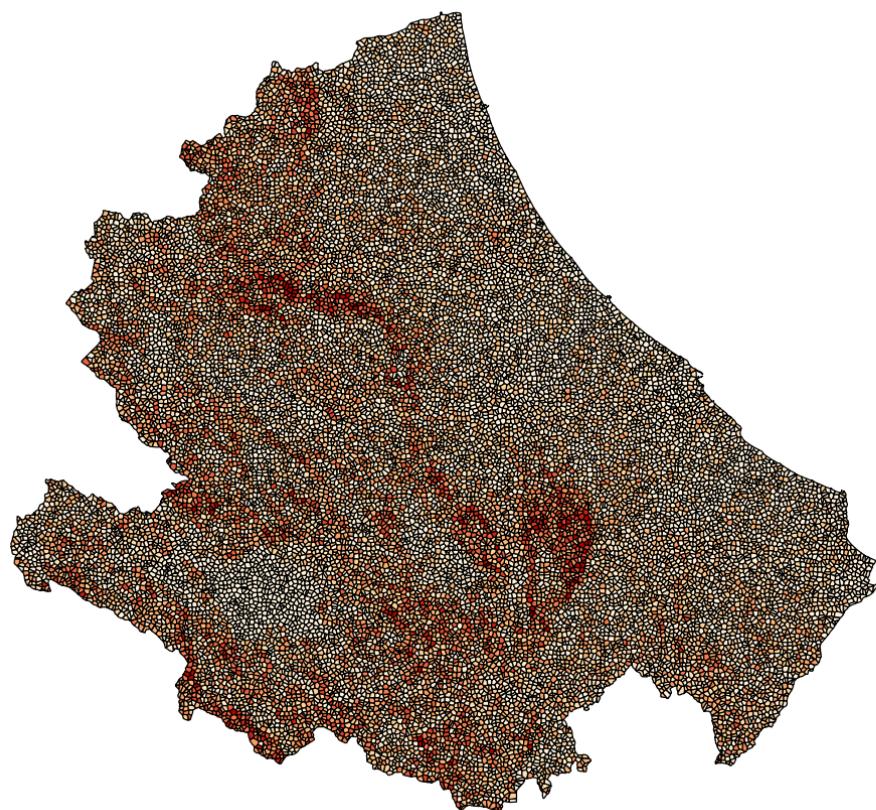


Figure 43: In figura si può osservare il dataset ottenuto tenendo conto delle curve di livello ed applicando la scomposizione di Voronoi

dai dati raster, la suddivisione in poligoni di Voronoi si è ottenuto un dataset sintetico che va a rispettare tutte le specifiche che ci eravamo imposti. Di seguito verranno proposte le UDF relative alla creazione del dataset.

```

1 CREATE TABLE voronoi AS
  SELECT (ST_DUMP(geom_voronoi)).geom AS geom_dump
  FROM (
    SELECT ST_collectionextract(
      ST_voronoi polygons(geom_point), 3) AS geom_voronoi
  6 FROM (
    SELECT ST_GeneratePoints( geom , 22000) AS geom_point
    FROM geo_area AS query3
  )AS query2
  )AS query;
11
CREATE TABLE dataset AS
SELECT ST_Intersection(abruzzo.geom,voronoi.geom_dump) AS geom
FROM voronoi, geo_area AS abruzzo;

CREATE OR REPLACE FUNCTION zkcalcolus (idc INTEGER) RETURNS void
LANGUAGE plpgsql
3 AS $$

DECLARE
  max INTEGER;
  min INTEGER;
  probabilita FLOAT;
  diff FLOAT;
  cursore RECORD;
  posneg FLOAT;
  variance FLOAT;
  BEGIN
13 FOR cursore in SELECT id,geom FROM dataset WHERE idc = id LOOP
  max := 0;
  min := 0;
  diff := 0;
  probabilita := 0;
18 posneg := 0;
  variance := 0;

  CREATE TEMP TABLE intersezioneQuadratino
  id SERIAL PRIMARY KEY , elevation float, geom geometry;
23
  INSERT INTO intersezioneQuadratino(elevation, geom)
  SELECT isoipseabruzzo25.elevation,st_collectionextract(
    st_intersection(dataset.geom,isoipseabruzzo25.geom),2)
  FROM isoipseabruzzo25,dataset
28 WHERE dataset.id = cursore.id;

  CREATE TEMP TABLE intersezioneQuadratinoPulita
  (id SERIAL PRIMARY KEY , elevation float, geom geometry);
  INSERT INTO intersezioneQuadratinoPulita
33 SELECT *
  FROM intersezioneQuadratino
  WHERE st_astext(geom) <> 'MULTILINESTRING EMPTY';

```

```
max :=( SELECT elevation
38 FROM intersezioneQuadratinoPulita
      ORDER BY elevation DESC LIMIT 1);
min :=( SELECT elevation
      FROM intersezioneQuadratinoPulita
      ORDER BY elevation ASC LIMIT 1);

43 DROP TABLE intersezioneQuadratino;
DROP TABLE intersezioneQuadratinoPulita;

posneg :=(SELECT random());
48 IF posneg > 0.5 THEN
  posneg = 1;
ELSE
  posneg = -1;
END IF;
53 variance := (SELECT (random()*0.3));
variance := variance * posneg;
diff := max - min;
IF diff <> 0 THEN
  IF diff > 800 THEN
58 probabilita := 1;
  ELSE
    probabilita := diff / 800 ;
    probabilita := probabilita + variance;
    IF probabilita > 1 THEN
63 probabilita := 1;
    END IF;
    IF probabilita < 0 THEN
      probabilita := 0;
    END IF;
  END IF;
68 END IF;
ELSE
  probabilita := 0;
END IF ;
UPDATE dataset SET zk = probabilita WHERE id = cursore.id;
73 END LOOP;
END;
$$
```
