

ANALISI DEL RISCHIO DA FRANA IN AMBIENTE GIS: IL CASO DEL SUB-APPENNINO DAUNO (PUGLIA)

Francesco MANCINI (*), Claudia CEPPI (*), Giuliano RITROVATO (*)

(*) Dipartimento di Architettura ed Urbanistica, Politecnico di Bari, Via Orabona, 4 – 70125 Bari
fmancini@racine.ra.it, giuritro@gmail.com, ceppicla@gmail.com

Riassunto

Il presente studio è il frutto del lavoro svolto all'interno di un progetto di ricerca finanziato dalla Regione Puglia e dall'unione Europea, che ha come principale scopo quello di valutare il rischio da frana nei centri minori del Sub-appennino Dauno. Allo scopo di produrre, a scala regionale, una carta della suscettività da frana, intesa come propensione al dissesto, nel lavoro vengono discusse le tecniche utilizzate per il trattamento dei dati e le modalità di inserimento delle variabili in ambiente GIS. Tali variabili hanno costituito un affidabile campione di partenza per l'analisi statistica della suscettività basata su un modello di regressione logistica. Si mostrano i risultati ottenuti attraverso l'analisi applicata a due campioni di dati, il primo comprendente la totalità delle informazioni ed il secondo suddiviso in due porzioni, la prima utilizzata per l'analisi e la seconda nella verifica dell'accuratezza nella determinazione della suscettività da frana. Tale valore, in entrambi i casi, si attesta sul 75% circa dimostrando anche che il risultato è indipendente dalla numerosità del campione iniziale.

Abstract

The paper is within a research project supported by Regione Puglia and European Community and aims to the landslide risk assessment on minor centre of the Sub-appennino Dauno (Puglia). Main goal is the creation of a landslide susceptibility map to be intended as proneness to risks. In particular, the statistical methodologies adopted in the data analysis and the selected variables to be used in the GIS environment will be addressed. A logistic regression model was used in the susceptibility assessment from the initial dataset. Finally, in order to validate the results provided by the analysis, a double solution was computed, the first based on the whole dataset and the second based on a subset of data. By this approach the reliability and robustness of the initial dataset could be verified. Both tests have produced a final accuracy at the 75% level of the susceptibility analysis (ability to recognize unstable area).

Introduzione

L'analisi di suscettibilità, o suscettività, riguardo ai movimenti franosi, richiede una base di conoscenza di fenomeni complessi, in grado da relazionare i movimenti ed i suoi fattori scatenanti. L'affidabilità delle mappe di suscettibilità dipende in grande misura dalla qualità dei dati utilizzati, dalla scala di lavoro e dalla scelta di una appropriata metodologia di analisi (Ayalew e Yamagishi, 2005). I metodi quantitativi per l'analisi di suscettibilità si basano sulla ricerca di una espressione numerica delle relazioni che legano i fattori predisponenti e il fenomeno franoso. Tra i metodi si distinguono principalmente due approcci, il primo deterministico ed il secondo di tipo statistico (Aleotti e Chowdhury, 1999). Tra i vari metodi utilizzati quello statistico cerca di superare gli ovvi limiti degli approcci qualitativi, dove l'attribuzione di un peso ai singoli fattori che governano l'instabilità, oltre alla scelta dei fattori stessi, avviene sulla base di osservazioni di tipo soggettivo o su conoscenze incomplete. Ovviamente nell'applicazione del metodo statistico la maggior difficoltà risiede nell'individuazione dei quei fattori che regolano l'instabilità e nella valutazione della loro relazione con il fenomeno in esame.

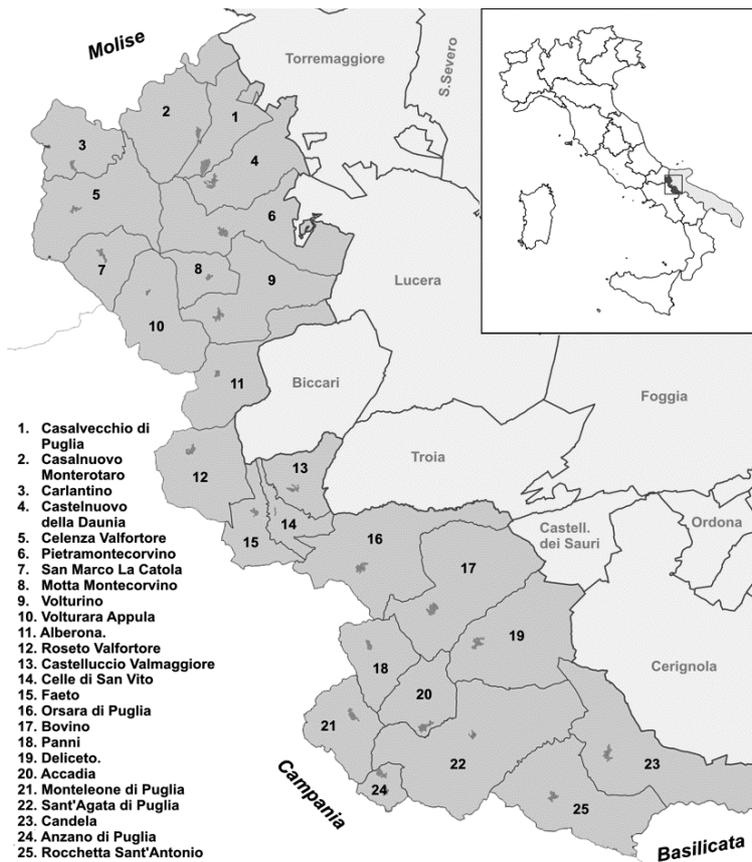


Figura 1 – Inquadratura dell'area di studio (Sub-Appennino Dauno) e localizzazione dei Comuni interessati dal calcolo della suscettibilità da frana.

D'altro canto esso consente di quantificare l'importanza di ogni fattore e di stabilire, quindi, quale sia il *dataset* iniziale più idoneo ai fini della produzione delle mappe di suscettibilità. L'implementazione delle informazioni iniziali (variabili geografiche) in ambiente GIS facilita enormemente molte delle operazioni di analisi e questo spiega, in larga misura, la crescente "popolarità" delle tecniche di analisi statistiche nel mondo dell'informazione geografica. Il lavoro descritto si inserisce proprio in questo contesto e si propone di mettere a punto per l'area di studio un metodo di analisi della suscettibilità affidabile e verificabile tramite il confronto con informazioni raccolte nel tempo per l'area di studio. L'area di studio comprende in particolare i 25 Comuni del Sub-appennino Dauno elencati e localizzati nella figura 1. I dati utilizzati nell'analisi proposta, e discussi in seguito, sono interamente disponibili per ognuno di essi.

Metodi statistici per l'analisi della suscettibilità: la regressione logistica

I metodi statistici multivariati studiano l'interazione e la dipendenza tra un numero di fattori, nella loro combinazione simultanea, e l'occorrenza di frana, al fine di stabilire l'influenza che ognuno di essi esercita nell'innescare del fenomeno. Le tecniche statistiche maggiormente utilizzate allo scopo sono la regressione multipla e l'analisi discriminante. Entrambi i metodi forniscono come risultato delle funzioni basate su combinazioni lineari dei fattori di maggior significato statistico per definire la condizione di instabilità basata sull'assenza o presenza di frana. La regressione multipla risulta particolarmente utile quando si cercano le relazioni esplicative tra le diverse variabili. Il metodo determina l'influenza di un numero di variabili indipendenti in relazione ad una o più variabili dipendenti. Tra i vari metodi statistici esistenti, ed utilizzabili per gli scopi del lavoro, è stato scelto quello della regressione logistica, appartenente alle tecniche di analisi multivariate, dove la relazione tra l'occorrenza del fenomeno franoso (in termini di probabilità) e la sua dipendenza con ciascuna variabile può essere espressa tramite la forma

$$Pr = \frac{1}{(1 + e^{-Z})}$$

dove Pr è la probabilità che un evento si verifichi. Z varia tra -1 e +1, la probabilità varia tra 0 e 1. Z deriva da una combinazione lineare del tipo $Z = B_0 + B_1X_1 + B_2X_2 + \dots + B_nX_n$ dove B_0 rappresenta l'intercetta della curva di regressione, B_i ($i=1, \dots, n$) il coefficiente angolare della curva di regressione logistica stimato dai dati e relativo ad ogni variabile indipendente ed n è il numero di variabili indipendenti. Nel nostro caso Pr rappresenta la stima della probabilità di occorrenza di

frana e le X_n rappresentano i fattori predisponenti (Dai e Lee, 2002; Lee e Sambath, 2006). Esaminando poi il segno dei coefficienti stimati per ogni variabile si ottengono delle considerazioni circa l'influenza di ogni variabile sulla probabilità di occorrenza. Un coefficiente positivo indica che tale variabile indipendente accresce la probabilità di frana mentre un coefficiente negativo, al contrario, indica che tale fattore agisce a vantaggio della stabilità (Chen e Wang, 2007).

I passi preliminari per l'applicazione del metodo sono a) suddivisione dell'area di studio in unità discrete (Carrara et al., 1991), rappresentate nel nostro caso da un grigliato regolare generato con passo di 40x40m; b) identificazione dei fattori significativi ai fini dell'analisi e creazione dei dati di input. In questa fase è necessario convertire parametri nominali in forma numerici, ad esempio per dati quali la classificazione litologica o l'uso del suolo; c) costruzione della mappa di franosità in un database appositamente creato con l'aiuto della conoscenza esperta; d) suddivisione del campione in aree stabili ed instabili. La decisione in merito ad una classificazione in solo due classi stabili ed instabili facilita di molto l'applicazione del metodo statistico anche se ostacola l'identificazione delle diverse combinazioni di fattori per i vari livelli di pericolosità; e) combinazione dei parametri mappati con la matrice delle unità di base, data da ogni classe di ogni parametro, in funzione dell'assenza o presenza di frana; f) applicazione dell'analisi multivariata; g) riclassificazione delle unità in funzione dei risultati archiviati nella precedente fase in funzione della classi di suscettibilità (Aleotti e Chowdhury, 1999).

Variabili utilizzate nel calcolo della suscettività

Dunque la prima fase prevede l'individuazione delle variabili utili allo studio, la loro verifica in termini di qualità ed accuratezza e la loro predisposizione (formattazione) all'analisi. I fattori che devono essere considerati nel calcolo della suscettività possono essere raggruppati in 5 categorie: fattori orografici, fattori idrologici, fattori geomorfologici, fattori geotecnici e fattori d'uso (copertura). Nell'esecuzione di questo lavoro non tutti i fattori citati erano disponibili (anche per la difficoltà di strutturare in ambiente GIS informazioni di natura così differente) e solo quelli elencati di seguito sono stati utilizzati al fine di testare il metodo prescelto e verificarne poi la capacità predittiva.

1) *Fattori orografici*. A tale categoria appartengono i parametri descrittivi della forma del terreno: pendenza, esposizione, rugosità o curvatura. I parametri morfometrici vengono ricavati da prodotti come i DEM di derivazione aerofotogrammetria che, ovviamente possono essere disponibili con diversa accuratezza e risoluzione spaziale. Nel caso specifico è stato utilizzato un DEM prodotto nell'ambito del progetto Terra Italy - IT2000NR (CGR, Parma) disponibile per tutta la Regione Puglia. I Comuni interessati ricadono in 59 quadrangoli con taglio degli elementi al 10.000 del reticolato IGM ed il passo è di 40 m (accuratezza planimetrica 2m e accuratezza altimetrica di 5m). Tali dati sono stati importati in ambiente GIS, in forma raster, senza altre operazioni di correzione come, ad esempio, interpolazioni o riempimenti. In totale nell'area ricadono 2.856.411 pixel con quote tra i 47,73 m. e 1.142,98 m s.l.m. La *pendenza media* è stata calcolata con il metodo della variazione dei gradienti nelle direzioni Est-Ovest (x) e Nord-Sud (y). Viene calcolata la media tra le pendici delle linee che corrono da nord a sud e diagonalmente attraverso il centro della cella, analogamente per la direzione est ovest. Infine viene calcolato il valore $tg \theta = (x^2 + y^2)^{1/2}$ ove θ rappresenta la pendenza media, in gradi sessagesimali, della cella (Zevenbergen e Thorne, 1987). L'esposizione indica l'inclinazione della direzione di massima pendenza della cella rispetto alle direzioni cardinali. L'esposizione è stata suddivisa in 9 classi assegnando i seguenti valori: 1=E, 2=SE, 4=S, 8=SW, 16=W, 32=NW, 64=N, 128=NE, 255=pendenza media nulla. La curvatura rappresenta la variazione della pendenza nello spazio e si esprime come gradiente della pendenza. Sono state effettuate due valutazioni: curvatura planare e curvatura di profilo. L'analisi della curvatura planare permette di identificare e caratterizzare le aree di flusso: le curve di livello che in pianta si presentano convesse indicano un flusso convergente, concave di tipo divergente. L'analisi della curvatura di profilo indica se un qualsiasi punto lungo il profilo appartiene ad un'area concava

o convessa: la curvatura di profilo è convessa quando in pianta le curve di livello si avvicinano e concava quando le curve di livello si allontanano oppure, in sezione longitudinale, convessa se la pendenza aumenta verso valle e concava nel caso opposto. I metodi sono quelli descritti da Zevenbergen e Thorne, 1987.

2) *Fattori idrologici*. Tra questi rientra l'area drenata, che rappresenta l'estensione (in pixel o m²) del sottobacino drenato. Il calcolo richiede la conoscenza delle direzioni di flusso, procedura iterativa che verifica quali celle inviano il flusso verso il pixel considerato entro l'area del bacino. In questo lavoro è stata utilizzata la procedura proposta da Tarboton (1997), che non si descrive nel dettaglio, la quale richiede la rimozione delle depressioni utilizzando un criterio di riempimento delle stesse e conservando la direzione prevalente dell'intorno. Nel caso di studio le correzioni apportate sono risultate di piccola entità ed utilizzate solo in questa occasione. Infine va sottolineato che i valori ottenuti si riferiscono ad una piovosità costante su tutta l'area di studio (1 litro/cella) e permeabilità nulla.

3) *Fattori geomorfologici*. I fattori geomorfologici utilizzati sono quelli raccolti nell'ambito del Progetto Strategico citato e fanno riferimento all'*inventario dei fenomeni franosi* per i siti oggetto di studio. Tali dati sono ottenuti da foto-interpretazione, da indagini speditive di campagna e dall'integrazione con informazioni storiche e di base sulla tipologia di frana. Tale inventario è stato strutturato in un GeoDatabase. Anche questi dati, ai fini dell'applicazione dell'analisi di regressione, sono stati rasterizzati riducendo l'informazione a quella di area in frana (1) o non in frana (VOID) per ogni cella dell'area di studio.

4) *Fattori geotecnici*. Alcune informazioni derivano dalla *carta litologica* della Regione Puglia redatta in scala 1:100.000. Le classi che interessano l'area di studio sono 36, successivamente accorpate in 13, in funzione del comportamento meccanico generale, con un codice identificativo nella struttura raster dell'informazione. Grandi miglioramenti su questa voce sono attesi con la nuova carta litologica alla scala 1:50'000 in produzione ad opera dell'APAT. Per ragioni di spazio si omettono ulteriori discussioni su questo punto.

5) *Fattori d'uso*. L'*uso del suolo*, per definizione, costituisce un elaborato dove confluiscono molteplici fattori, naturali che antropici, di carattere ambientali geomorfologico, pedologico e vegetazionale, in interazione con le modificazioni dovute all'urbanizzazione, all'agricoltura e all'attività zootecnica. La carta di copertura del territorio utilizzata è quella prodotta a metà anni '90 nel progetto CORINE Land Cover (Heyman et al., 1994) e successivamente aggiornata (per la Puglia all'anno 2000). Ai fini del calcolo della suscettibilità, il numero delle classi è stato ridotto a 10 in relazione all'influenza che viene associata ad ognuna di queste.

L'applicazione del metodo

La variabili significative e disponibili (in totale ne sono state selezionate 9) sono state rappresentate in grigliati dove in corrispondenza dei nodi sono disponibili coordinate e valore della variabile. Tali dati sono stati poi importati nell'ambiente MIS (SPSS) per l'analisi statistica. Una delle operazioni preliminari all'applicazione del metodo è stata l'elaborazione dei diversi fattori in variabili che fossero costituite da classi omogenee. In particolare le variabili di tipo nominale (uso del suolo, orientazione, litologia) sono state convertite in variabili binarie, tale operazione è necessaria sia per creare delle variabili che abbiano con la variabile dipendente (assenza/presenza di frana) una relazione prossima a quella lineare sia per poter esaminare quali valori delle variabili predispongono il suolo all'instabilità, o, in altre parole, siano maggiormente correlate con la variabile dipendente. Il campione è stato poi depurato dai valori nulli, ed è stato suddiviso in due sottocampioni di egual numerosità. Infatti, essendo il metodo applicato a tutta l'area di studio, vi erano un elevato numero di celle non in frana, e quindi tramite estrazione *random* di campioni è stato derivato un sottocampione numericamente equilibrato in modo da avere una variabile dicotomica come variabile dipendente. Per la validazione del metodo, e quindi al fine di determinare la sua capacità predittiva, sono stati preparati due set di dati. Il primo con la totalità delle informazioni ed il secondo con esclusione dall'analisi di regressione del 25% delle aree in

frana. Questo al fine di verificare l'effetto della numerosità del campione (e quindi della sua rappresentatività del fenomeno) sui risultati ottenuti.

Analisi dei risultati

L'analisi di regressione restituisce come prima informazione il confronto fra due modelli: il modello con solo l'intercetta (regressione lineare) ed il modello con tutte le variabili presenti (regressione multipla). Nel caso del campione comprendente la totalità dei dati tale confronto presenta il 50.4% di casi correttamente classificati nel primo modello (quindi un modello che indipendentemente dalle covariate disponibili considera tutte le unità come suscettibili a frana) contro il 75,6% fornito dal secondo modello, la cui predizione dipende dalle variabili disponibili. Nel secondo campione tali valori erano rispettivamente 50.1% e 75.4%. L'analisi fornisce anche la stima dei parametri del modello logistico con s.q.m. associato e la significatività del parametro. Nonostante alcuni parametri avevano un basso livello di significatività si è scelto di non escluderli dall'analisi per non avere una pur minima perdita di informazione. Alla fine delle due analisi abbiamo potuto verificare che nonostante la perdita di informazione (provocata dalla riduzione del campione del 25%) la stima delle aree correttamente classificate era molto simile.

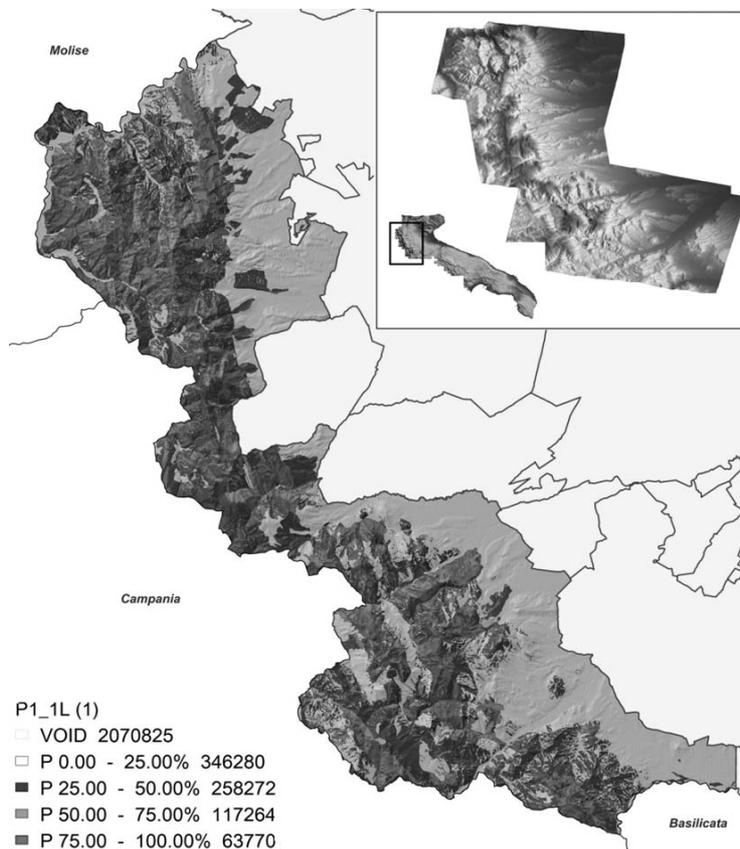


Figura 2 – Mappa di suscettività ottenuta con la totalità del campione

Quindi tramite l'analisi di regressione logistica è stato calcolato ed assegnato un valore della probabilità ad ogni cella ed ottenere le rispettive mappe di suscettività, riportate nelle figure 2 e 3, rispettivamente calcolate sulla totalità del campione e con il campione ridotto del 25%.

I valori di probabilità sono stati raggruppati in 4 classi di egual range. Successivamente è possibile confrontare per ognuno dei pixel la corrispondenza tra predizione e comportamento reale. L'analisi dei coefficienti poi mette in evidenza come alcuni fattori risultano fortemente significativi, vedi ad esempio le classi di uso del suolo, ed alcune classi di orientazione.

Molto interessante è notare come alcune classi litologiche presentano dei fattori significativamente negativi, ad indicare come tali affioramenti siano meno suscettibili al processo franoso.

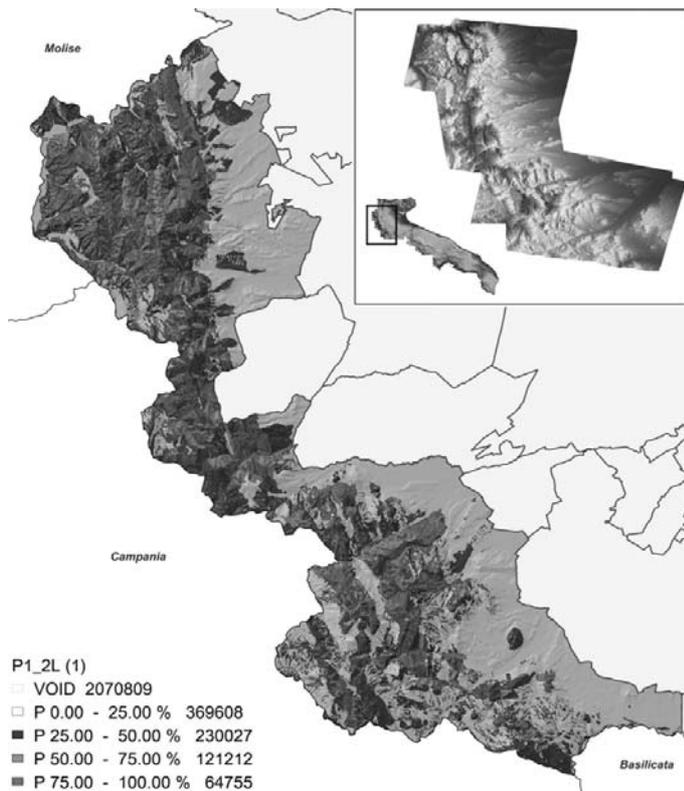


Figura 3 – Mappa di suscettività ottenuta con un campione ridotto del 25%.

Tali ricoprimenti vanno quindi a vantaggio della stabilità per i meccanismi che interessano l'area.

Conclusioni

L'applicazione del metodo ha portato ad una ottima percentuale di aree correttamente classificate come franose ed in entrambi i campioni utilizzati, testimoniando una analisi non troppo legata alle ipotesi iniziali. Tale percentuale si attesta in entrambi i casi sul livello del 75% circa. Risultano quindi suscettibili sia aree in cui si sono già verificati fenomeni come aree dove essi non si sono ancora verificati. Peraltro i fattori da considerare come scatenanti sulla suscettività trova riscontro in base a considerazioni di tipo geomorfologico e geotecnico suggerendo nuovi spunti di riflessioni sulla natura del fenomeno nell'area di studio.

Ringraziamenti

Il lavoro svolto nell'ambito delle attività del Progetto Strategico "Valutazione del rischio da frana per la pianificazione di centri urbani minori in zone di catena: il caso della Daunia" (coordinatrice: prof.ssa Federica Cotecchia – Politecnico di Bari. Si ringrazia la dott.ssa Francesca Santaloia (CNR-IRPI Bari) per la collaborazione nella fase di implementazione in ambiente GIS dei dati geomorfologici. Il software Geomedia 6.0 (Intergraph) è fornito nell'ambito del progetto "Registered Research Laboratory (RRL) Program".

Bibliografia

- Aleotti P., Chowdhury R. (1999), "Landslide hazard assessment: summary review and new perspectives", *Bulletin of Engineering Geology and the Environment*, 58(1), 21-44.
- Ayalew L. Yamagishi H. (2005), "The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan", *Geomorphology*, 65(1-2), 15-31.
- Carrara A., Cardinali M., Detti R., Guzzetti F., Reichenbach P. (1991), "GIS techniques and statistical models in evaluation landslide hazard", *Earth surface processes and landforms*, 16, 427-445.
- Chen Z., Wang J. (2007), "Landslide hazard mapping using logistic regression model in Mackenzie Valley, Canada", *Natural Hazards*, 42(1), 75-89.
- Dai F.C., Lee C. F. (2002), "Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong", *Geomorphology*, 42(3-4), 213-228.
- Lee S., Sambath T. (2006), "Landslide susceptibility mapping in the Damrei Romel area, Cambodia using frequency ratio and logistic regression models", *Environmental Geology*, 50(6), 847-855.
- Tarboton D. G. (1997), "A New Method for the Determination of Flow Directions and Contributing Areas in Grid Digital Elevation Models", *Water Resources Research*, 33(2), 309-319.
- Zevenbergen L.W., Thorne C.R. (1987), "Quantitative Analysis of Land Surface Topography", *Earth Surface Processes and Landforms*, 12, 1987.
- Heymann Y., Steenmans C., Croisille G., Bossard M. (1994) "CORINE land cover project -Technical guide", ECSC-EEC-EAEC, Brussels- Luxembourg, 136 pp.