Consiglio Nazionale delle Ricerche
Istituto di Ricerca per la Protezione Idrogeologica

PROGRAMMA SHORT-TERM MOBILITY 2006
del CONSIGLIO NAZIONALE DELLE RICERCHE

RELAZIONE SCIENTIFICA
SULL'ATTIVITÀ DI RICERCA SVOLTA DALLA DOTT.SSA SILVIA PERUCCACCI

PROGRAMMA DI RICERCA
"DEFINIZIONE DI SOGLIE PLUVIOMETRICHE PER L'INNESCO DI MOVIMENTI FRANOSI"

Periodo: Nell'ambito del programma Short-Term Mobility del Consiglio Nazionale delle Ricerche, la dott.ssa Silvia Peruccacci, assegnista presso l'Istituto di Ricerca per la Protezione Idrogeologica del Consiglio Nazionale delle Ricerche sede di Perugia, ha trascorso il periodo di studio accordato presso il *Lamont-Doherty Earth Observatory (LDEO) of Columbia University, New York* dal 6 novembre al 27 novembre 2006.

Introduzione: Le frane possono essere prodotte da diversi fenomeni naturali e da numerose azioni antropiche. Fra i fenomeni naturali, i fenomeni meteorologici, ed in particolare le piogge intense e prolungate, sono certamente quelli che innescano frane con maggiore frequenza ed abbondanza.. Già da tempo alcuni ricercatori hanno messo a punto soglie di pioggia per la previsione dell'innesco di movimenti franosi. Le soglie disponibili si basano sull'analisi di eventi pluviometrici dei quali si conoscano le caratteristiche di durata, intensità e pioggia totale (cumulata), e che abbiano prodotto frane.

Obiettivo: Scopo del breve soggiorno di studio della dott.ssa Silvia Peruccacci presso il *Lamont-Doherty Earth Observatory* della Columbia University, era quello di sperimentare l'applicazione di tecniche d'analisi statistica alla definizione di soglie pluviometriche per l'innesco di fenomeni franosi al fine di comprendere meglio le relazioni fra le caratteristiche della precipitazione e l'innesco delle frane e garantire una definizione "oggettiva" delle stesse soglie.

Attività svolta: Durante il soggiorno presso il Dipartimento *Lamont-Doherty Earth Observatory* della Columbia University di New York, la dott.ssa Peruccacci ha collaborato con il Dr. Colin Stark, esperto nella modellazione matematica di fenomeni naturali. L'attività svolta è consistita essenzialmente:

a) nel raffinamento dei codici di calcolo sviluppati dal Dr. Colin Stark per la definizione di soglie di intensità-durata (ID) e di intensità-pioggia cumulata (IC) per l'innesco di frane;

b) nell'utilizzo di un nuovo strumento di analisi definito copula o funzione di dipendenza per unire o "accoppiare" funzioni di distribuzione multivariate alle rispettive distribuzioni marginali unidimensionali, in questo caso le distribuzioni di durata ed intensità media di eventi pluviometrici che hanno o non hanno prodotto frane;

via della Madonna Alta 126 • 06128 PERUGIA • Italy
Tel. +39 075 5014.421 • Fax. +39 075 5014.420 • http://www.irpi.cnr.it/

c) nella valutazione della qualità dei dati raccolti nell'archivio di informazioni relativo ad eventi pluviometrici che hanno o non hanno prodotto frane; e

d) nell'utilizzo della regressione logistica per stimare la probabilità che si verifichi un evento, in questo caso il verificarsi di un fenomeno franoso.

Descrizione dei dati utilizzati: Durante lo svolgimento del programma di ricerca si è utilizzato il più vasto archivio di informazioni oggi disponibile al mondo relativo ad eventi pluviometrici che hanno o non hanno prodotto frane. L'archivio raccoglie informazioni relative a 7419 eventi pluviometrici di cui 3501 che hanno generato dissesti e 3918 che non hanno generato movimenti franosi. Per ogni evento sono note: la durata e l'intensità della pioggia, il totale cumulato di precipitazione, il numero e la tipologia delle frane innescate dall'evento, e le principali caratteristiche litologiche e climatiche dell'area interessata dall'evento. L'attività di ricerca è stata focalizzata sull'identificazione di soglie di pioggia per l'innesco di movimenti franosi di tipo superficiale. A tal proposito si sono utilizzate 2626 informazioni pluviometriche relative a tali eventi franosi e 3918, poi successivamente ridotte a 1185, relative ad eventi che non hanno generato frane.

Risultati ottenuti: I principali risultati ottenuti durante il breve soggiorno di studio possono essere così brevemente riassunti:

a) Nel codice di calcolo per la definizione delle soglie pluviometriche per l'innesco di frane sono stati diminuiti il numero di updates che il modello bayesiano sviluppato deve eseguire. Tale modifica consente di visualizzare meglio l'andamento delle catene di Markov utilizzate nel modello (indicatore della bontà del modello stesso) e di snellire e rendere più rapida la convergenza del modello.

b) La tecnica di regressione tramite copula tra due variabili consiste nel definire la distribuzione congiunta delle stesse, associando una funzione di dipendenza delle distribuzioni marginali. In questo modo è possibile determinare la probabilità di accadimento di due variabili in maniera congiunta. In via del tutto sperimentale si è deciso di provare ad utilizzare una tale tecnica per la stima della probabilità congiunta delle variabili durata, D, ed intensità media, I, di eventi pluviometrici che hanno o non hanno prodotto frane. L'equazione 1 mostra la relazione matematica alla base:

$$P_{I,D}(I,D) = P_I(I) \times P_D(D) \times \frac{\partial^2 C}{\partial I \partial D}(F_I(I), F_D(D)) \qquad \text{(Eq. 1)}$$

La densità bivariata delle variabili I e D, $P_{I,D}(I,D)$, può essere pertanto calcolata moltiplicando le densità marginali di I e D (rispettivamente $P_I(I)$ e $P_D(D)$) per la densità di copula (ultimo termine nell'eq. 1). Una sfida è stata quella di stimare la densità di copula utilizzando il semplice ed ampliamente utilizzato software Excel. Il primo passo è consistito nel determinare la probabilità cumulata P della variabile D e la probabilità cumulata P della variabile I per tutti gli eventi pluviometrici che hanno generato frane, rispettivamente $P_D(D_j)$ e $P_I(I_j)$ per il generico evento j di durata D_j ed intensità I_j. Le coppie di valori $P_D(D_j)$ e $P_I(I_j)$ sono state graficate ottenendo la Figura 1a. E' stata poi individuata una griglia di 20 × 20 celle, e per ciascuna cella è stato calcolato il numero di punti di coordinate $(P_D(D_j);P_I(I_j))$ che ricadevano all'interno della stessa. Tali valori, normalizzati per il numero totale di eventi pluviometrici e per l'area di ciascuna cella, sono stati opportunamente graficati ottenendo la densità di copula per le condizioni pluviometriche che hanno prodotto dissesti (si veda Figura 1b). La stessa procedura è stata seguita per la determinazione della densità di copula per le condizioni pluviometriche che hanno generato movimenti franosi (si veda Figura 2a e 2b).

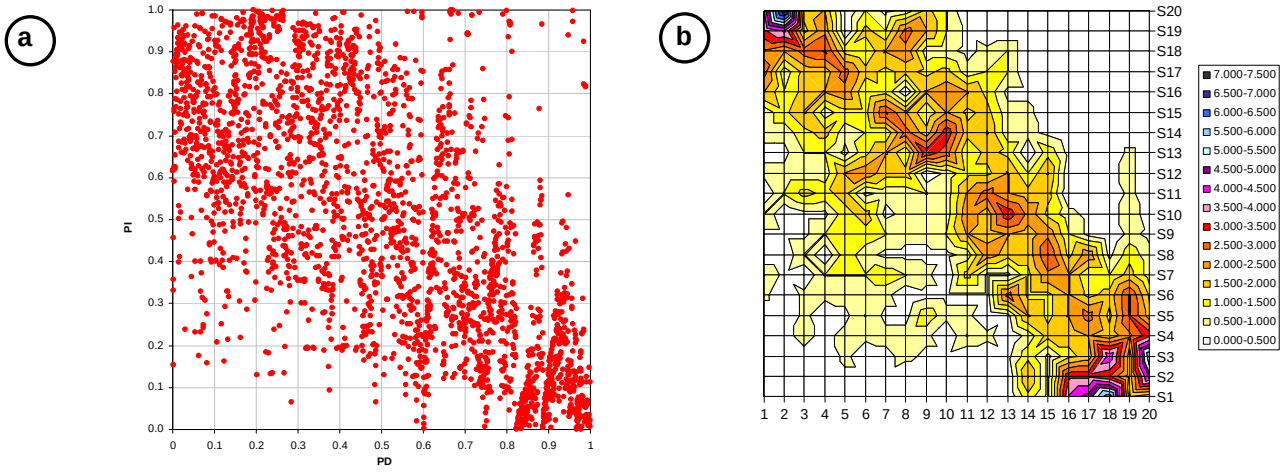


Figura 1. In a) sono riportate le coppie di valori $P_D(D_j)$ e $P_I(I_j)$ degli eventi pluviometrici che hanno generato frane. In b) è riportata la densità di copula per condizioni di D e I che hanno generato frane.

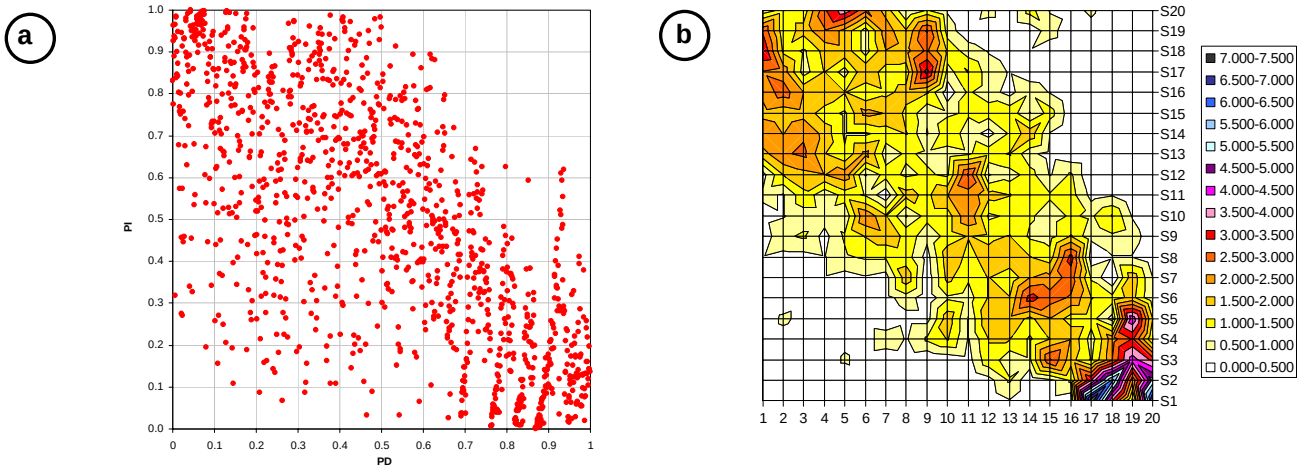


Figura 2. In a) sono riportate le coppie di valori $P_D(D_j)$ e $P_I(I_j)$ degli eventi pluviometrici che non hanno generato frane. In b) è riportata la densità di copula per condizioni di D e I che non hanno generato frane.

L'analisi è stata poi ripetuta utilizzando al posto della variabile intensità media, I, la variabile trasformata z definita come:

$$z = 1 - \left(\frac{2.22}{I \cdot D^{0.442}} \right) \qquad \text{(Eq. 2)}$$

La figura 3 mostra i grafici della densità di copula rispettivamente per le condizioni di D e z che hanno e non hanno generato frane. Come si evince dall'osservazione della suddetta Figura, l'effetto di una tale trasformazione è quello di condensare lungo la bisettrice i dati sia per gli eventi pluviometrici che hanno prodotto frane che non.

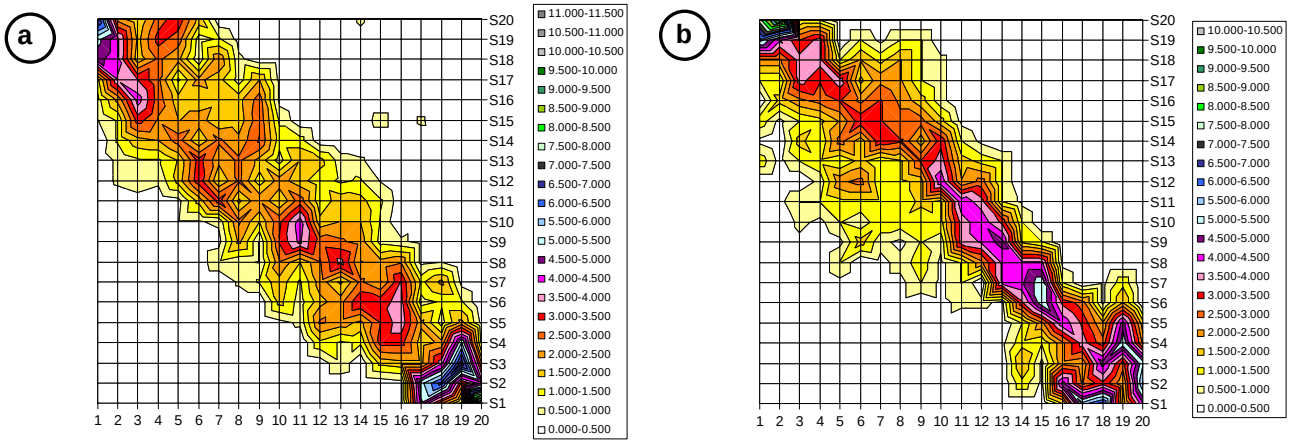


Figura 3. In a) è riportata la densità di copula per condizioni di D e z che hanno prodotto frane. In b) è riportata la densità di copula per condizioni di D e z che non hanno prodotto frane.

c) L'utilizzo della copula sulle condizioni di durata ed intensità media relative ad eventi pluviometrici che non hanno generato dissesti ha evidenziato la presenza di dati non congruenti raccolti nell'archivio (si veda Figura 4). Da una revisione approfondita degli articoli da cui si sono evinte le suddette condizioni pluviometriche, il numero delle informazioni pluviometriche relative ad eventi che non hanno generato frane è stato ridotto da 3918 a 1185 (si veda Figura 2a). L'uso della copula può pertanto essere un indice di buona qualità dei dati raccolti utili ad indagini successive evidenziando la presenza di un campione congruente.

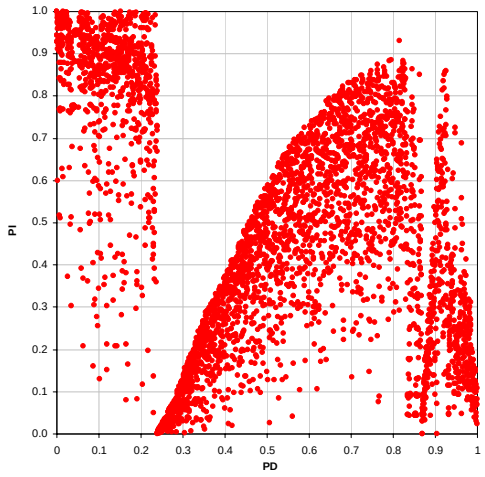


Figura 4. Coppie di valori $P_D(D_j)$ e $P_I(I_j)$ di tutti gli eventi pluviometrici presenti nell'archivio che non hanno generato frane.

d) Una tecnica di regressione logistica è stata utilizzata per stimare la probabilità che si verifichi o meno un evento franoso. Tale tecnica usa i fattori osservati, durata D di un evento pluviometrico ed intensità media di pioggia relativa allo stesso evento pluviometrico, assieme all'occorrenza o non occorrenza dell'evento, per stimare la probabilità che l'evento si verifichi in determinate circostanze.

Si è scelto un modello di regressione logistica per variabile dipendente di tipo dicotomico (che può assumere due valori: 0,1). La variabile dipendente dicotomica r_i può essere codificata numericamente

$r_i = 1$ se si è in presenza di un fenomeno franoso

$r_i = 0$ se non si è in presenza di un fenomeno franoso

r_i indica dunque la presenza o l'assenza di una particolare caratteristica (in questo caso un fenomeno franoso). In termini di probabilità si ha

r_i	$P(r_i)$
1	p_i
0	$1 - p_i$

r_i si distribuisce come una variabile casuale Bernoulliana di parametro p_i, che descrive l'esito di un esperimento casuale che ha probabilità di risultare in "fenomeno franoso" con probabilità pari a p_i. Si ha quindi:

$$r_i \cong Bern(p_i) \qquad \text{(Eq. 3)}$$

$$P(r_i) = p_i^{r_i}(1-p_i)^{1-r_i} \qquad \text{(Eq. 4)}$$

Per l'interpretazione dei parametri è utile definire l'*odds*, ossia il rapporto fra la probabilità che r_i assuma valore 1 e quella che assuma valore 0:

$$odds(p_i) = \frac{p_i}{1-p_i} \qquad \text{(Eq. 5)}$$

Ogni evento pluviometrico di durata D ed intensità media I è stato considerato usando la variabile trasformata z definita come:

$$z = -\ln\left(\frac{\alpha}{I \cdot D^{\beta}}\right) \qquad \text{(Eq. 6)}$$

Essendo il *Logit*, ossia il logaritmo dell'*odds*, una funzione lineare dei parametri ed essendo la curva logistica parametrizzata attraverso un valore critico di soglia z_c ed un valore di smoothing σ si ha:

$$Logit(p_i) = \ln[odds(p_i)] = \ln\left(\frac{p_i}{1-p_i}\right) = \frac{z_i - z_c}{\sigma} + \beta_i \cong \frac{z_i - z_c}{\sigma} \qquad \text{(Eq. 7)}$$

con semplici passaggi matematici si ottiene:

$$\frac{p_i}{1-p_i} = \exp^{\frac{z_i - z_c}{\sigma}} \qquad \text{(Eq. 8)}$$

$$p_i = (1-p_i)\exp^{\frac{z_i - z_c}{\sigma}} \qquad \text{(Eq. 9)}$$

$$p_i = \exp^{\frac{z_i - z_c}{\sigma}} - p_i \cdot \exp^{\frac{z_i - z_c}{\sigma}} \qquad \text{(Eq. 10)}$$

$$p_i\left(1 + \exp^{\frac{z_i - z_c}{\sigma}}\right) = \exp^{\frac{z_i - z_c}{\sigma}} \qquad \text{(Eq. 11)}$$

$$p_i = \frac{\exp^{\frac{z_i - z_c}{\sigma}}}{1 + \exp^{\frac{z_i - z_c}{\sigma}}} \qquad \text{(Eq. 12)}$$

Per l'inferenza dei parametri z_c e σ, il Dr. Stark ha sviluppato un modello basato sull'assunzione che i dati di input seguano una distribuzione di tipo Bernoulliano con una probabilità data da una curva logistica in z. Come input sono stati utilizzati gli eventi pluviometrici che hanno indotto frane, z_{fi}, e gli eventi pluviometrici che non hanno generato frane, z_{nfi}, a cui sono stati associati rispettivamente i valori 1 e 0:

$$EventiPluviometrici : \left\{z_{f1}, z_{f2}, ..., z_{fi}, z_{nf1}, z_{nf2}, ..., z_{nfi}\right\} \qquad \text{(Eq. 13)}$$

$$r_i : \{1,1,...,1,0,0,...,0\} \qquad \text{(Eq. 14)}$$

Il modello viene risolto attraverso il package denominato WinBUGS (http://www.mrc-bsu.cam.ac.uk/bugs/) e sfrutta il principio della massima verosimiglianza per l'inferenza dei parametri z_c e σ. Per i dati utilizzati il modello regressivo ha stimato $z_c = 1.61$ e $\sigma = 0.64$.
Assegnati dei valori arbitrari a z, z_k, attraverso l'eq. 12, è possibile stimare il valore di p_k:

$$p_k = \frac{\exp^{\frac{z_k - 1.61}{0.64}}}{1 + \exp^{\frac{z_k - 1.61}{0.64}}}$$

Per gli eventi pluviometrici che hanno e non hanno generato frane, rispettivamente z_{fi} e z_{nfi}, si sono tracciate le curve di probabilità cumulata, PC, e, attraverso un tool di interpolazione, si sono stimati i valori di probabilità in corrispondenza degli assegnati z_k. La figura 5 mostra gli eventi pluviometrici che hanno e non hanno generato frane, rispettivamente z_{fi} in rosso e z_{nfi} in verde, le curve di probabilità cumulata per gli eventi che hanno e non hanno indotto frane in corrispondenza dei valori z_k, rispettivamente in fucsia ed azzurro, ed infine la curva di probabilità ottenuta mediante il modello di regressione logistica p_k in corrispondenza di z_k, curva in blu. Dall'analisi della figura si evince come la curva del modello di regressione si posiziona tra le curve di probabilità cumulata degli eventi pluviometrici che hanno o non hanno prodotto fenomeni franosi evidenziando la bontà del modello sviluppato.

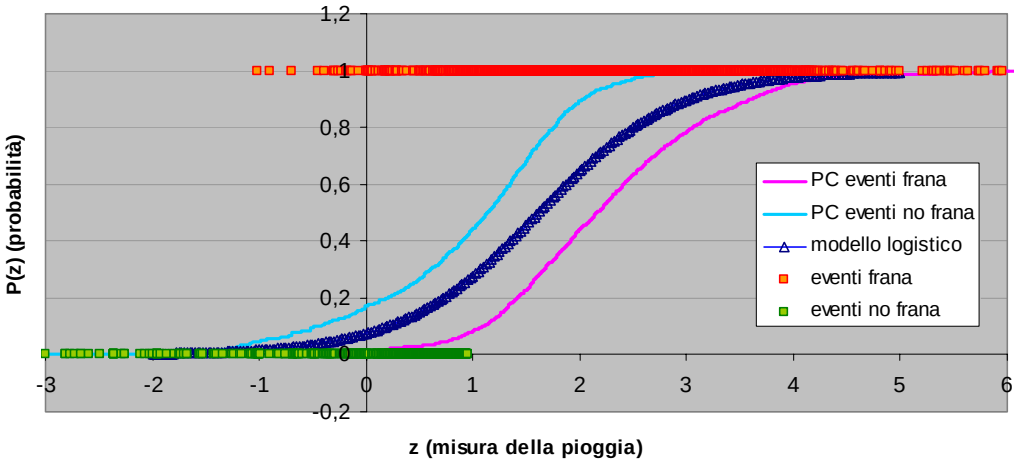
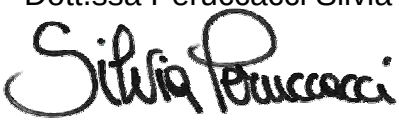


Figura 5. Eventi pluviometrici che non hanno indotto frane in verde, eventi pluviometrici che hanno indotto frane in rosso, curva di probabilità per gli eventi che non hanno generato frane in azzurro, curva di probabilità per gli eventi che hanno generato frane in fucsia, e curva di probabilità del modello di regressione logistica.

il proponente

Prof. Ing. Lucio Ubertini

il fruitore

Dott.ssa Peruccacci Silvia