

ALMA MATER STUDIORUM – UNIVERSITÀ DI BOLOGNA

SCUOLA DI SCIENZE

Corso di Laurea Magistrale in

FISICA DEL SISTEMA TERRA

CARATTERIZZAZIONE DELLE
ETEROGENEITÀ
DELLO SLIP SU UNA FAGLIA
MEDIANTE TECNICHE
DI FACE DETECTION

RELATORE:

Prof. Alberto Armigliato

PRESENTATA DA:

Elena Paoli

SESSIONE III

ANNO ACCADEMICO 2017-2018

Abstract

Il problema della caratterizzazione delle eterogeneità dello scorrimento cosismico su una faglia in occasione di un terremoto è affrontato mediante l'applicazione di una tecnica di "machine learning" nota come "face detection". In assenza di un metodo univoco per prevedere quale sarà la struttura spaziale dello slip in funzione della magnitudo, del meccanismo focale e della localizzazione geografica di un terremoto, questa tesi propone uno studio di fattibilità sull'utilizzo di algoritmi che normalmente sono utilizzati per il riconoscimento facciale, per lo studio delle suddette relazioni. Il database di riferimento è SRCMOD, una collezione di 180 modelli cosiddetti "a faglia finita", consistenti in altrettante rappresentazioni matriciali della distribuzione di slip sulle faglie responsabili di terremoti avvenuti a livello mondiale fra il 1906 e il 2017. Il software OpenCV è utilizzato in tutti i principali passaggi in cui l'approccio di face detection si articola: definizione di diversi classificatori basati sull'aspect ratio delle asperità sul piano di faglia, loro "allenamento", valutazione della performance dei classificatori sull'intero database SRCMOD, applicazione ad un sottoinsieme di modelli a singola asperità per la ricerca di eventuali correlazioni fra l'aspect ratio dell'asperità e la magnitudo, il meccanismo focale, la localizzazione geografica. Il principale risultato ottenuto conferma una conclusione già proposta in letteratura (Goda et al., 2016): non sembra esistere alcuna correlazione fra l'aspect ratio delle asperità e la magnitudo del terremoto, nemmeno se l'analisi è trattata separatamente per diversi meccanismi focali. E' invece possibile che alcune regolarità esistano fra l'area geografica in cui un terremoto avviene e l'aspect ratio dell'asperità dominante: i dati a disposizione sono però troppo poco numerosi per trarre conclusioni certe.

Indice

1.	Introduzione.....	1
2.	Databases di modelli “a faglia finita”.....	7
2.1	Il Database SRCMOD	7
2.2	Distribuzione geografica degli eventi contenuti in SRCMOD	10
2.3	Distribuzione dei meccanismi focali.....	12
2.4	Breve riepilogo delle leggi di scala pubblicate in letteratura.....	15
2.5	Relazione fra lo slip massimo e la profondità.....	16
2.6	Rappresentazione degli eventi di SRCMOD per l’applicazione degli algoritmi di face detection	18
3.	Algoritmi di “face detection”: il classificatore “Haar cascade”	21
3.1	Il classificatore “Haar Cascade”	21
3.1.1	Haar-like feature	22
3.1.2	L’algoritmo Adaboost.....	24
3.1.3	Classificatore Haar come cascata di classificatori	25
3.2	L’“allenamento” del classificatore.....	27
3.3	L’utilizzo del classificatore allenato per il riconoscimento	29
4.	Allenamento di classificatori HAAR sugli eventi di SRCMOD.....	33
4.1	Scelta delle caratteristiche da riconoscere in una data distribuzione di slip	33
4.2	Allenamento dei 7 classificatori con OpenCV.....	34
4.3	Applicazione dei classificatori ai modelli di slip a singola asperità	41
4.3.1	Effetto dei parametri <i>ScaleFactor</i> , <i>MinNeighbor</i> e <i>MinSize</i>	49
5.	Conclusioni.....	59
	APPENDICE A	64
	APPENDICE B.....	91
	Bibliografia.....	123

1. Introduzione

Il terremoto è un fenomeno fisico estremamente complesso. In termini molto generali, possiamo pensare ad un terremoto come ad un rilascio molto rapido (pochi secondi o poche decine di secondi) di energia come conseguenza del superamento di una data soglia di sforzo in un certo volume della crosta terrestre. Il rilascio di energia avviene attraverso un processo di frattura all'interno della crosta stessa, che inizia in un punto, chiamato ipocentro, e che si propaga lungo una superficie chiamata "faglia", che può pre-esistere nella crosta oppure può essere creata ex-novo dal terremoto. Il processo di frattura è accompagnato dalla generazione e propagazione di onde elastiche, cioè delle onde sismiche, e dalla deformazione permanente della crosta ed eventualmente della superficie terrestre (la cosiddetta "deformazione cosismica"). La Sismologia ha fatto progressi molto significativi negli ultimi 50-60 anni nella comprensione dei processi fisici che sono alla base della generazione di un terremoto, sia attraverso ricerche sperimentali in laboratorio sia, soprattutto, mediante lo sviluppo di un apparato teorico in grado di descrivere in termini di equazioni, e quindi di quantificare, la propagazione delle onde sismiche e i processi deformativi e di sforzo in modelli sempre più realistici della Terra. Ciononostante, i problemi aperti in sismologia sono ancora numerosi. Tralasciando qui il problema centrale della previsione di dove, quando e con quali caratteristiche avverrà il prossimo terremoto, un tema che ha ricevuto crescente interesse da una trentina d'anni a questa parte è quello della descrizione dettagliata del processo di frattura sulla faglia. A livello globale, la crosta terrestre viene descritta come il "guscio" più esterno e sottile del nostro pianeta, che esibisce un comportamento elastico (tende a tornare nel suo stato iniziale quando le forze che agiscono su di essa vengono rimosse) e fragile (si frattura quando è soggetta a sforzi molto "intensi" che agiscono su tempi molto brevi). Quando la si analizza nel dettaglio, si scopre che la crosta è un sistema fisico molto eterogeneo dal punto di vista della composizione e dello stato di sforzo. Non è quindi sorprendente che il modo in cui una faglia si rompe durante un terremoto sia tutt'altro che uniforme. Una faglia, che possiamo immaginare come una superficie di discontinuità lungo la quale due blocchi di crosta terrestre scorrono uno rispetto all'altro durante un terremoto, tipicamente non è planare e non ha una forma geometrica semplice. Allo stesso modo, non è uniforme lo scorrimento

relativo (“slip”) fra i due blocchi crostali lungo la superficie di faglia. Sulla base dell’analisi di un numero elevato di terremoti avvenuti a livello globale si è osservato che lo slip tende a concentrarsi in zone, dette “asperità”, il cui numero e la cui distribuzione sul piano di faglia sono il risultato di una combinazione non quantificabile a priori dello stato di sforzo locale e delle eterogeneità locali della crosta. Lo slip sulla faglia può essere rappresentato da un vettore, e come tale ha un modulo (l’ “entità” dello scorrimento, che dimensionalmente è una lunghezza), una direzione e un verso, che sono quelli del movimento relativo fra i due blocchi crostali. L’eterogeneità dello slip, che possiamo considerare una caratteristica universale di tutti i terremoti anche se risulta più evidente al crescere della magnitudo, si manifesta quindi come una variazione spaziale sia del suo modulo che della sua direzione e verso. Questi ultimi non variano del tutto casualmente, ma sono generalmente vincolati entro intervalli che sono strettamente legati allo stato di sforzo (“regime tettonico”) di una data porzione di crosta. Lo scopo principale di questa tesi è di proporre un metodo di caratterizzazione del modulo dello slip sulle faglie che permetta di individuare eventuali caratteristiche ricorrenti in funzione di un certo numero di parametri caratteristici di un terremoto, quali la sua magnitudo e la zona geografica (e quindi lo stato di sforzo) in cui un terremoto avviene.

La possibilità di definire caratteristiche ricorrenti della distribuzione di slip in funzione della magnitudo e del meccanismo focale non è un problema di interesse esclusivamente teorico, ma ha importanti conseguenze sulle stime di pericolosità sia in campo sismico che per quanto riguarda i fenomeni che ad un terremoto sono spesso associati, come frane e maremoti. Limitandoci al secondo tema, i terremoti di magnitudo più elevata che sono stati registrati negli ultimi 15-20 anni sono stati seguiti da maremoti che, in alcuni casi, sono stati catastrofici. E’ sufficiente pensare al terremoto del 26 dicembre 2004 al largo di Sumatra, che generò un maremoto capace di propagarsi attraverso l’intero Oceano Indiano e di provocare morte e distruzione non solo nel campo vicino (Sumatra) ma anche nel campo medio (Thailandia, India, Sri Lanka) e in quello lontano (coste africane orientali). Il terremoto ebbe magnitudo 9.1 e consistette nella rottura della crosta terrestre lungo una faglia lunga circa 1000 km, estesa dalla zona al largo dell’isola di Sumatra a sud fino al largo della Thailandia e alle isole Andamane a nord. Lo slip si concentrò su un certo numero di asperità (in numero

leggermente variabile a seconda degli studi) e l'effetto fu quello di concentrare gli effetti di maremoto su zone geografiche specifiche, soprattutto nel campo vicino (Banda Aceh nel nord di Sumatra, le coste thailandesi occidentali). Un terremoto con magnitudo molto simile ma con caratteristiche del tutto diverse a livello di eterogeneità dello slip e di effetti di maremoto è stato quello di Tohoku (Giappone) dell' 11 marzo 2011. In questo caso il terremoto rompe un'unica asperità principale, concentrata su una lunghezza totale di 400-500 km. Lo slip massimo sulla faglia ricavato da modelli di inversione fu quindi più alto che nel caso del terremoto del 2004: la combinazione di un'asperità "concentrata" spazialmente e quindi di uno slip massimo maggiore fu responsabile di effetti di maremoto devastanti specialmente lungo alcune zone delle coste giapponesi, come la piana di Sendai. Al di là dello studio di singoli eventi, modelli probabilistici o deterministici della distribuzione delle asperità su diverse geometrie di faglia potrebbero essere utilizzati da modelli per la stima della pericolosità sia in campo sismico che nel campo dei maremoti, con importanti ripercussioni sulla pianificazione territoriale e sui temi di protezione civile (si veda per esempio Goda e De Risi, 2018).

Al tema della pericolosità si aggiunge quello dell'early warning. E' evidente che avere a disposizione strategie efficaci per poter formulare ipotesi ragionevoli in tempi molto brevi riguardo alla possibile distribuzione di slip di un terremoto di magnitudo moderata o grande può rivelarsi una risorsa molto efficace nella stima in tempo reale degli effetti del terremoto stesso, ma soprattutto di possibili fenomeni ad esso associati, come i già citati maremoti.

In letteratura sono stati pubblicati numerosissimi studi che si occupano della definizione di strategie per la caratterizzazione delle eterogeneità dello slip su una faglia. Molti di questi studi si basano sullo sviluppo di modelli di inversione di diversi tipi di dati (sismologici, geodetici, satellitari, di maremoto), considerati singolarmente o in combinazione. Lo scopo finale è sempre quello di determinare il valore di slip e di eventuali altre caratteristiche della rottura (angolo di rake, collegato al meccanismo di rottura, rise time, ...) su una griglia di celle (sotto-faglie) in cui il piano di faglia viene discretizzato. Riportiamo di seguito solo alcuni fra i numerosissimi esempi possibili. Lo studio di Piatanesi et al. (1996) per il calcolo della distribuzione di slip sulla faglia del terremoto tsunamigenico del 2 settembre 1992 in Nicaragua, insieme al lavoro di Satake (1995) rappresenta uno dei primi esempi in cui i dati di run-up di maremoto sono stati

utilizzati per invertire la distribuzione di slip sulla faglia. Ammon et al. (2005), Rhie et al. (2007) e Fuji e Satake (2007) sono alcuni esempi di studi aventi lo scopo di ricavare la distribuzione dello slip per il terremoto di Sumatra del 2004 sulla base dell'inversione rispettivamente di dati sismici, della combinazione di dati sismici e geodetici, di dati di maremoto e satellitari.

Altri studi si occupano invece dell'indagine dei processi fisici che portano alla realizzazione di determinati "pattern" di slip. Un esempio è rappresentato dal lavoro di Lavallée et al. (2006), che dall'analisi di quattro terremoti (Imperial Valley 1979, Loma Prieta 1989, Northridge 1994, Kobe 1995) hanno ricavato che la legge probabilistica di Lévy riproduce in modo soddisfacente le principali caratteristiche della variabilità spaziale dello slip e a partire da questa hanno sviluppato un modello stocastico per modellare la distribuzione delle asperità sul piano di faglia. Lekshmy e Raghukanth (2019) hanno applicato il modello di Lavallée et al. (2006) per la simulazione di scenari di scuotimento realistici da utilizzare per scopi di ingegneria sismica. Aagaard e Heaton (2008) hanno utilizzato simulazioni dinamiche del processo di rottura con diverse combinazioni di condizioni iniziali per riprodurre diversi casi di slip eterogeneo ed evoluzioni temporali dello slip su faglie planari, verticali e con meccanismo strike-slip nel caso particolare di terremoti di magnitudo 7. L'obiettivo era principalmente quello di determinare le condizioni di stress drop dinamico e dell'energia di frattura necessarie per determinare un certo tipo di distribuzione finale di slip. Candela et al. (2011) hanno studiato la relazione che lega la distribuzione dello slip e la "roughness" della superficie di faglia. Lo studio di Aochi e Ide (2014) rappresenta infine un esempio in cui il problema è affrontato a partire da considerazioni sulla eterogeneità dell'energia di fratturazione e non dello stato di stress. La dimensione delle asperità è trovata essere in relazione di scala con l'energia di frattura, e il campo di radiazione delle onde generate da un terremoto è influenzato principalmente dalle patches più grandi e solo in misura minore da quelle più piccole.

In questa tesi è stato utilizzato un approccio completamente diverso: si è adottato il metodo di "face detection" per il riconoscimento delle eterogeneità dello slip. L'impiego di tecniche di riconoscimento basate su paradigmi di "machine learning" non è una novità nel campo della Geofisica. Gli esempi di applicazione sono numerosi e spaziano attraverso diversi campi di ricerca. Per citarne qualcuno, Sandri et al. (2005)

hanno utilizzato algoritmi di “pattern recognition” per studiare i cluster sismici che avvengono intorno all’Etna, per trovare possibili collegamenti tra le eruzioni e gli eventi sismici regionali. Ehret (2010) l’ha applicato a problemi di geofisica applicata, in particolare all’analisi di dati GPR (ground-penetrating radar) per la caratterizzazione delle interfacce fra strati rocciosi differenti in assenza di dati provenienti da pozzi. Questo secondo esempio illustra bene il procedimento generale, in cui diversi modelli artificiali che rispecchiano le caratteristiche del fenomeno fisico o delle proprietà fisiche di un materiale che si vogliono caratterizzare, vengono utilizzati per allenare opportune reti (algoritmi) di addestramento. Le reti così allenate vengono poi applicate ai dati sperimentali per cercare di ricavarne strutture ricorrenti o comunque proprietà difficilmente ricavabili attraverso altri approcci. Numerosi sono gli esempi di applicazione di tecniche di machine learning in sismologia: un’ottima sintesi è stata recentemente pubblicata da Kong et al. (2019).

Fra le tecniche di machine learning, quella di “face detection” è forse quella meno utilizzata in sismologia. Uno dei pochi studi che è stato possibile trovare in letteratura è quello di Mulia et al. (2018) lo hanno impiegato per il calcolo di mappe di inondazione da parte di uno tsunami in tempo reale, con applicazione al Giappone. Di conseguenza, almeno a conoscenza dell’autrice questo studio è fra i primi, se non il primo, che propone di applicare tecniche di face detection al problema della caratterizzazione delle eterogeneità dello slip cosismico su una faglia.

Come descritto nel lavoro già citato di Kong et al. (2019), in generale un algoritmo di “machine learning” si basa sui seguenti passaggi:

- 1) raccolta di dati;
- 2) pre-processing dei dati;
- 3) allenamento dell’algoritmo e analisi dei parametri specifici per il tipo di problema analizzato;
- 4) applicazione e valutazione dell’algoritmo su dati di interesse;
- 5) applicazione dell’algoritmo a nuovi dati.

La tesi è strutturata a grandi linee seguendo lo schema precedente. Il Capitolo 2 riguarda i punti 1) e 2): si presenta il database di modelli a faglia finita SRCMOD (<http://equake-rc.info/SRCMOD/>) e si discutono le strategie seguite per il pre-processing. Il Capitolo 3 riguarda la teoria generale relativa agli algoritmi di face

detection e include una discussione sulle strategie di allenamento mediante l'utilizzo del software OpenCV (Open Source Computer Vision Library, <https://opencv.org/>). Il Capitolo 4 riprende il tema dell'allenamento e lo discute con applicazione alle caratteristiche delle distribuzioni eterogenee di slip che si è deciso di riconoscere. L'algoritmo allenato è applicato al database SRCMOD e valutato tramite opportuni indici. Non viene trattato il punto 5) dello schema di Kong et al. (2019). Il Capitolo 5, infine, contiene un'estesa discussione e le conclusioni principali a cui si è giunti.

2. Databases di modelli “a faglia finita”

L'applicazione di strategie di machine learning in generale, e di face detection nel nostro caso particolare, richiede come primo tassello la disponibilità di un opportuno database. Gruppi di ricerca appartenenti ad enti mondiali diversi hanno cominciato a curare la compilazione di database di modelli a faglia finita negli ultimi quindici-venti anni. Un notevole impulso è stato dato dal terremoto di Sumatra del 26 dicembre 2004, caratterizzato da una rottura estremamente eterogenea sia spazialmente che temporalmente. La grande mole di dati sismologici, geodetici, satellitari, geologici e di maremoto che sono stati resi disponibili immediatamente dopo l'evento ha rappresentato un'occasione per accelerare lo sviluppo di algoritmi di inversione per ricavare la dinamica della rottura della faglia e la distribuzione finale di slip. L'approccio è stato esteso anche a terremoti di magnitudo significativa (tipicamente > 6) avvenuti sia precedentemente che successivamente al 2004, permettendo quindi di compilare database di modelli a faglia finita per un numero sempre crescente di eventi a livello globale. Fra i diversi enti che rendono disponibili online i propri database citiamo lo U.S. Geological Survey, Earthquake Hazards Program, (<https://www.sciencebase.gov/catalog/item/58da9d37e4b0543bf7fdaab3>), e il California Institute of Technology, Tectonics Observatory (http://www.tectonics.caltech.edu/slip_history/index.html). Nel primo caso l'evento più recente per cui è disponibile una soluzione a faglia finita è del 2017, mentre nel secondo caso il database non è più stato aggiornato dopo il 2014.

Il database usato come riferimento in questo lavoro è il “Finite-Source Rupture Model Database” SRCMOD (<http://equake-rc.info/SRCMOD/>), curato dal team del Prof. Martin Mai a partire dal 2004 e che attualmente include 180 modelli a faglia finita per terremoti avvenuti fra il 1906 e il 2017. Nel prossimo paragrafo descriviamo questo database con qualche dettaglio.

2.1 Il Database SRCMOD

A differenza dei databases curati dai singoli enti di ricerca, che per propria natura includono solo i risultati ottenuti mediante le specifiche strategie di inversione sviluppate dai team di ricerca che operano negli enti stessi, SRCMOD si propone come

un raccolta di soluzioni pubblicate da diversi autori a livello mondiale e con diverse tecniche. Non sorprende quindi che l'intervallo temporale coperto da SRCMOD sia molto più esteso rispetto ad altri databases, nè che per uno stesso terremoto siano presenti più soluzioni ricavate da diversi autori. Evidentemente, un'iniziativa di questo genere richiede innanzitutto un grande sforzo di standardizzazione, che al momento si traduce principalmente nella creazione di un certo numero di formati comuni per i files che descrivono i dettagli del processo di rottura e i parametri usati nell'inversione, e per la rappresentazione grafica della distribuzione finale di slip.

Nella versione online del database, ogni modello è rappresentato da una pagina come quella mostrata in Fig. 2.1, relativa al terremoto del Messico dell'8 settembre 2017.

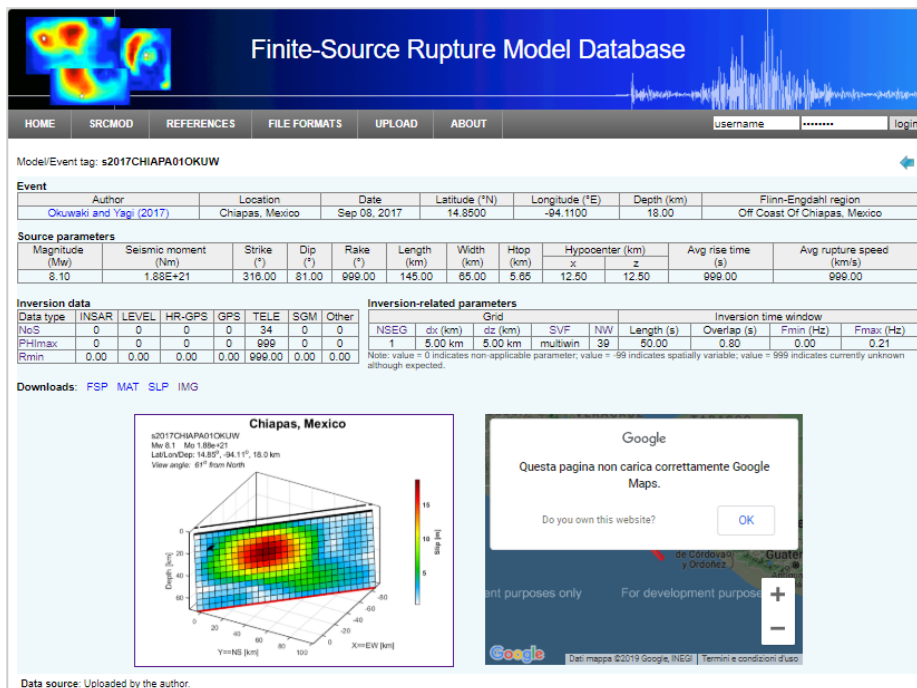


Figure 2.1

Ogni modello è caratterizzato da una stringa di caratteri che comprende l'anno dell'evento, i primi sei caratteri del luogo in cui esso è avvenuto, un contatore per specificare il numero del modello preso in esame (per distinguere modelli pubblicati da autori diversi per uno stesso terremoto) e infine il nome dell'autore.

Un tipico modello a finite-fault è formato da uno o più segmenti di faglia, rappresentati come superfici rettangolari, ciascuno dei quali viene discretizzato in un

dato numero di sotto-faglie, sempre rettangolari. La risoluzione spaziale del modello è definita dalla dimensione di queste sotto-faglie, ed è legata ai dati disponibili ed al metodo d'inversione impiegato.

Per ogni evento SRCMOD mette a disposizione diversi file di dati, che descrivono con diversi livelli di dettaglio le caratteristiche della discretizzazione e le proprietà del modello. Qui accenniamo solo ai seguenti due files, che sono stati utilizzati per le analisi descritte successivamente:

- SLP (Slip)

E' un file di testo diviso in due parti. La prima contiene in forma sintetica le informazioni sulla localizzazione del terremoto, sul meccanismo focale e la magnitudo, sui parametri cinematici di rottura sulle caratteristiche geometriche della discretizzazione (numero di sotto-faglie nelle direzioni dello strike e del dip). La seconda parte contiene il valore di slip su ciascuna sotto-faglia in forma matriciale:

- FSP (Finite Source Parameter)

E' un file molto più dettagliato rispetto al file in formato SLP. Contiene un elenco completo dei parametri di modellazione e d'inversione, il modello crostale utilizzato, dettagli sulla geometria della sorgente e il modello di rottura dipendente dallo spazio e dal tempo. Ciascuna sotto-faglia è individuata dalla posizione del punto centrale del lato superiore del rettangolo che la rappresenta, fornita sia in termini di latitudine e longitudine, sia rispetto ad un sistema cartesiano locale con origine nell'epicentro. Se disponibili, per ogni sotto-faglia possono essere riportati anche i valori dell'angolo di rake, del rise-time, del tempo di rottura.

Nella trattazione che segue, per ogni evento studiato è stato utilizzato il file FSP, da cui si sono estratte le informazioni relative alla posizione delle sotto-faglie e al valore di slip su ciascuna di esse.

Il database è riportato integralmente in Appendice A: rispetto alla versione originale, qui è stata introdotta una colonna aggiuntiva, dove viene riportato se un dato terremoto è stato seguito o meno da uno tsunami. Questa verifica è stata eseguita manualmente incrociando le informazioni contenute in SRCMOD e con i dati del

“Global Historical Tsunami Database” compilato dal National Center for Environmental Information (NGDC) della National Oceanographic and Atmospheric Administration (NOAA) (https://www.ngdc.noaa.gov/hazard/tsu_db.shtml).

2.2 Distribuzione geografica degli eventi contenuti in SRCMOD

La Fig. 2.2 mostra la localizzazione di tutti i terremoti contenuti in SRCMOD. La mappa riporta anche una versione semplificata del modello dei margini delle placche tettoniche pubblicato da Bird (2002).

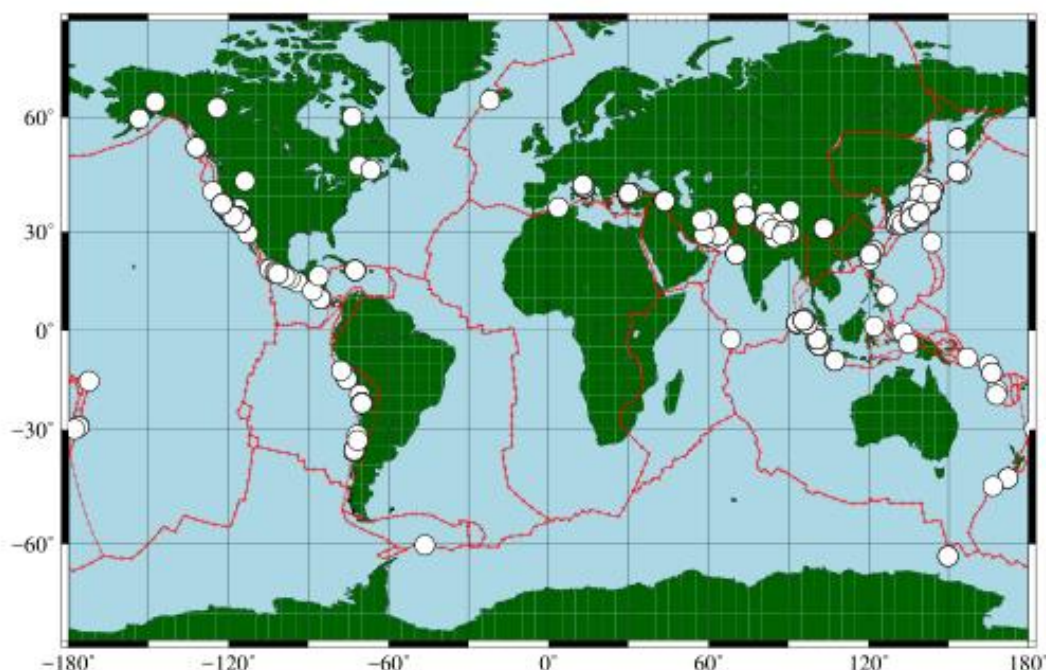


Figure 2.2 Distribuzione geografica dei 180 terremoti contenuti in SRCMOD. Le linee rosse rappresentano una versione semplificata dei contorni delle placche tettoniche secondo Bird (2002).

I terremoti contenuti nel catalogo si concentrano lungo le principali strutture tettoniche attive. Cominciando dalle coste orientali dell’oceano Pacifico, gli epicentri cadono lungo le coste dell’Alaska e nella zona delle isole Aleutine, lungo la faglia di San Andreas, le zone di subduzione centro-americana e sud-americana (Perù-Cile); a ovest, da nord a sud le strutture più attive sono le zone di subduzione al largo delle isole Kurili, del Giappone e delle Filippine, fino all’arco che comprende le isole Salomone, Vanuatu, Samoa ed infine la Nuova Zelanda. Nel bacino dell’oceano indiano, la zona di subduzione che va da nord dell’arcipelago delle isole Andamane passando per le isole

indonesiane di Sumatra e Java è stata sede di alcuni fra i più forti terremoti degli ultimi quindici anni, primo fra tutti il già citato terremoto del 26 dicembre 2004 (M_w 9). Più a nord, una densità importante di eventi è visibile in corrispondenza della catena himalayana e, più a ovest, nell'area della catena dei monti Zagros (Iran – Pakistan) e della subduzione di Makran. Nell'area euro-asiatica, SRCMOD include alcuni eventi lungo la faglia nord-anatolica, i terremoti recenti (dal 2009) avvenuti in Italia nella zona dell'Appennino centrale e il terremoto del 2003 in Algeria. Citiamo infine un certo numero di eventi localizzati in corrispondenza dei cosiddetti punti caldi, come in Islanda e a Yellowstone (Stati Uniti occidentali).

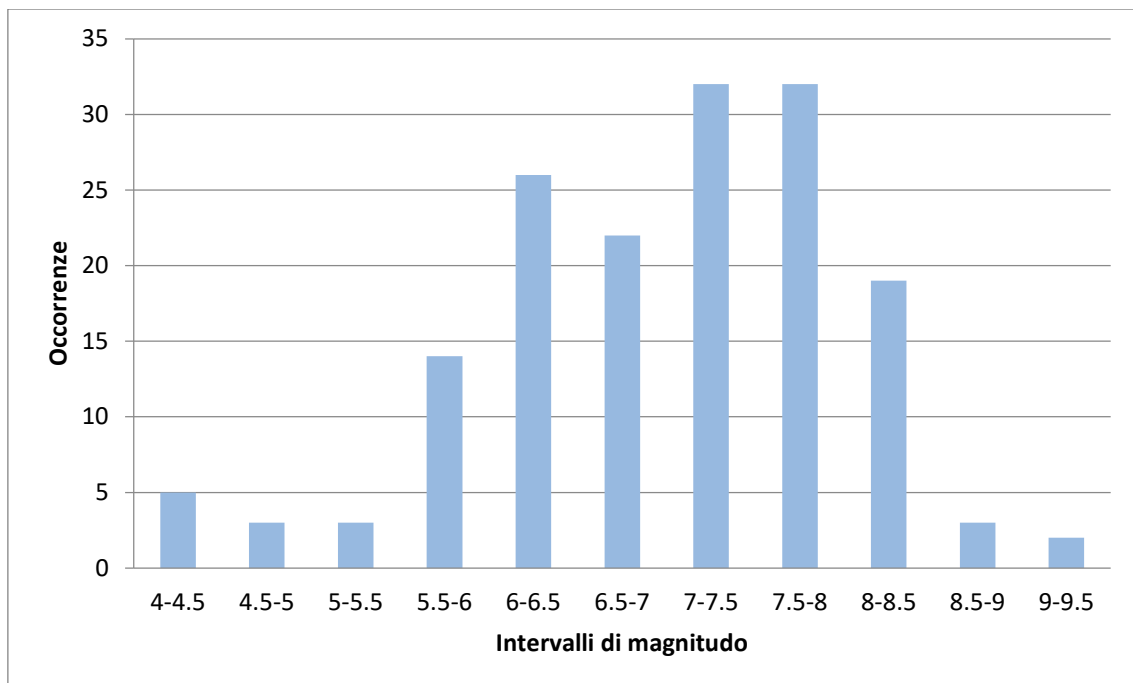


Figure 2.3 Distribuzione della magnitudo degli eventi nel database SRCMOD.

La Fig. 2.3 illustra la frequenza dei terremoti contenuti in SRCMOD per classi di magnitudo. Tralasciando la coda di eventi a magnitudo più bassa (< 5.5), che provengono da studi su sequenze sismiche specifiche (per esempio, compaiono 7 terremoti appartenenti allo sciame del monte Hida in Giappone, 12-22 agosto 1998), osserviamo che il database è largamente incompleto relativamente alle magnitudo moderate ($5.5 < M_w < 7$) rispetto alle magnitudo superiori ($M_w > 7$). In altri termini, il numero di terremoti contenuto in SRCMOD è un buon campione di tutti i terremoti avvenuti globalmente solo per le classi di magnitudo più alte. Tenendo conto che la

distribuzione globale delle occorrenze dei terremoti deve seguire l'andamento previsto dalla legge di Gutenberg-Richter, la frazione dei terremoti per cui è disponibile un modello a faglia finita rispetto al numero totale di terremoti registrati è nettamente più bassa per i terremoti con magnitudo fra 5.5 e 7.0. Questo può essere facilmente spiegato con il fatto che l'inversione per la determinazione del modello a faglia finita non è un'operazione di routine applicata a tutti i terremoti, ma viene effettuata quasi esclusivamente per i terremoti di magnitudo più alta, per i quali la mole e la qualità di dati è nettamente superiore rispetto al caso dei terremoti di magnitudo moderata o piccola.

2.3 Distribuzione dei meccanismi focali

Il meccanismo focale di un terremoto è l'insieme delle caratteristiche geometriche e cinematiche che descrivono il processo di rottura sul piano di faglia. In particolare, l'angolo di "rake" fornisce la direzione lungo cui avviene lo scorrimento relativo fra le due "facce" della superficie di faglia. Secondo una delle convenzioni più comuni in sismologia, l'angolo di rake viene misurato in senso antiorario a partire dalla direzione dello strike della faglia e può quindi variare fra 0° e 360° .

Si distinguono quattro tipi fondamentali di meccanismo focali:

- "Strike-slip" per angoli di rake 0° e 180° : in questo caso si dice che la faglia è di tipo "trascorrente", perché le due facce scorrono l'una rispetto all'altra quasi esclusivamente in direzione orizzontale.
- "Normal" per angoli di rake = 270° : in questo caso si parla di faglie "normali", caratteristiche di regimi di sforzo distensivo. Per faglie con inclinazione diversa dalla verticale, il cosiddetto "footwall" (il "tetto" in italiano) scende mentre il cosiddetto "hanging wall" (il "letto" in italiano) sale.
- Reverse per angoli di rake = 90° : in questo caso la faglia viene definita "inversa", perché il footwall sale e l'hanging wall scende. È un meccanismo tipo dei regimi tettonici compressivi.
- Per angoli di rake diversi dai precedenti, si parla genericamente di faglie "oblique".

Di seguito sono riportati i risultati ottenuti applicando una tolleranza di $\pm 15^\circ$ alle precedenti classificazioni. Nella Tabella 2.1 è riportato il numero di eventi presenti in

ognuno dei poligoni rappresentativi delle placche tettoniche disegnate in Figura 2.2 e la distribuzione dei meccanismi focali al loro interno (sono stati esclusi i poligoni che non avevano eventi al loro interno). La stessa informazione è fornita graficamente in Fig. 2.3.

Tabella 2.1. Distribuzione degli epicentri e dei meccanismi focali dei terremoti contenuti in SRCMOD fra le varie placche tettoniche, come definite in Bird (2002). La colonna intitolata "N° eventi a segmento singolo" indica quanti dei modelli a faglia finita sono rappresentati da una singola faglia.

	Nome poligono	N° eventi all'interno	N° a segmento o singolo	MECCANISMI FOCALI			
				Strike slip	Normal	Reverse	Oblique
AF	Africa	0	0				
AM	Amur	25	25	12	1	3	9
AN	Antarctica	2	2	1	0	0	1
AP	Altiplano	1	1	0	0	0	1
AR	Arabia	0	0				
AS	Aegean Sea	0	0				
AT	Anatolia	0	0				
AU	Australia	1	1	0	0	0	1
BH	Birds Head	2	2	1	0	0	1
BR	Balmoral Reef	0	0				
BS	Banda Sea	0	0				
BU	Burma	3	3	0	3	0	0
CA	Caribbean	2	2	0	0	0	2
CL	Caroline	0	0				
CO	Cocos	0	0				
CR	Conway Reef	0	0				
EA	Easter	0	0				
EU	Eurasia	31	30	5	11	2	12
FT	Futuna	0	0				
GP	Galapagos	0	0				
IN	India	4	4	3	0	1	0
JF	Juan de Fuca	1	1	0	0	1	0
JZ	Juan Fernandez	0	0				
KE	Kermadec	0	0				
MA	Mariana	0	0				
MN	Manus	0	0				

MO	Maoke	0	0				
MS	Molucca Sea	1	1	0	0	1	0
NA	North America	31	30	10	4	9	7
NB	North Bismarck	0	0				
ND	North Andes	0	0				
NH	New Hebrides	1	1	0	0	0	1
NI	Niuafou'ou	0	0				
NZ	Nazca	0	0				
OK	Okhotsk	20	19	1	3	11	4
ON	Okinawa	6	6	4	0	2	0
PA	Pacific	5	5	0	0	1	4
PM	Panama	1	1	0	0	1	0
PS	Philippine Sea	9	8	2	0	2	4
RI	Rivera	0	0				
SA	South America	3	3	0	0	1	2
SB	South Bismarck	0	0				
SC	Scotia	1	1	1	0	0	0
SL	Shetland	0	0				
SO	Somalia	0	0				
SS	Solomon Sea	0	0				
SU	Sunda	7	7	0	0	5	2
SW	Sandwich	0	0				
TI	Timo	0	0				
TO	Tonga	1	1	0	0	0	1
WL	Woodlark	0	0				
YA	Yangtze	1	1	1	0	0	0

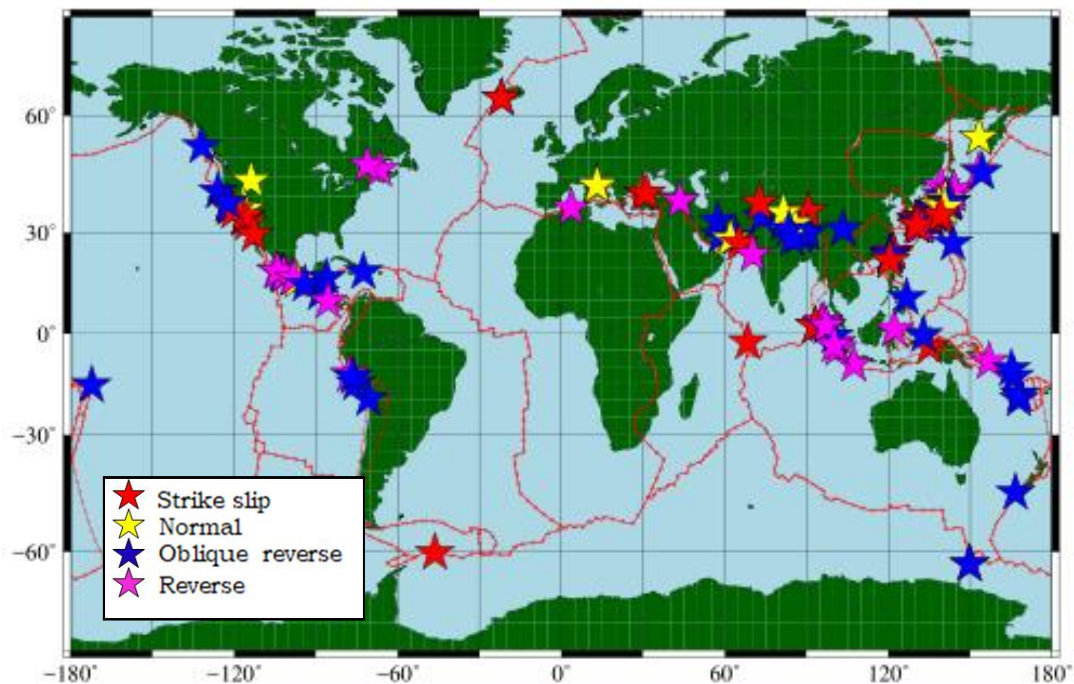


Figure 2.3 Mappa degli eventi presenti in SRCMOD con relativo meccanismo focale.

2.4 Breve riepilogo delle leggi di scala pubblicate in letteratura

In assenza di un modello deterministico “esatto” che, nota la magnitudo e la localizzazione di un terremoto, consenta di predire i parametri geometrici e cinematici fondamentali del processo di rottura, un problema classico in sismologia è quello di definire relazioni empiriche (regressioni) fra coppie alternative di parametri/proprietà caratteristiche del terremoto stesso. Un lavoro che viene citato come esempio di riferimento in tal senso è quello di Wells e Coppersmith (1994), in cui venivano presentate leggi di regressione fra la magnitudo momento e area, lunghezza, larghezza, slip medio, slip massimo delle relative faglie, specializzate per diversi meccanismi focali, a partire da un database di terremoti mondiali. Il lavoro di Wells e Coppersmith è servito come riferimento per numerosi studi successivi (ad esempio Blaser et al., 2010; Strasser et al., 2010; Leonard et al., 2010; Murotani et al., 2013), che hanno raffinato le leggi di regressione adottando database sempre più ricchi di terremoti, filtrandoli opportunamente per regimi tettonici, oppure, nel caso di specifici meccanismi focali, per caratteristiche della rottura (ad esempio intraplate o interplate earthquakes nel caso di terremoti di subduzione, come in Strasser et al., 2010). Nessuno degli studi citati si è però spinto oltre modelli in cui lo slip sulla faglia è fornito come un valore medio. E’

stato solo molto recentemente che Goda et al. (2016) hanno pubblicato il primo studio di dettaglio su leggi di regressione calcolate a partire da database di modelli a faglia finita, in particolare il database SRCMOD. Lo studio di Goda et al. propone anzitutto una possibile strategia per il pre-processing degli eventi contenuti in SRCMOD (selezione dei modelli a singolo segmento, “pulizia” dei modelli di faglia finita dalle celle con slip nullo in eccesso per ricavare la dimensioni effettive della faglia, trattamento dei terremoti che hanno più soluzioni a faglia finita proposti da diversi autori). Successivamente, propone nuove versioni per le leggi di regressione classiche fra magnitudo momento e le principali caratteristiche geometriche delle faglie (area, lunghezza, larghezza), oppure i valori caratteristici dello slip (medio e massimo). Due novità importanti riguardano 1) il calcolo della legge di regressione fra il rapporto fra area delle asperità e l’area totale della faglia, e la magnitudo momento, 2) la trattazione separata dei terremoti che sono stati seguiti da maremoto e quelli che non lo sono stati. Riguardo al primo punto, Goda et al. giungono alla conclusione che il rapporto fra l’area delle asperità (definita come l’area totale delle celle in cui lo slip è almeno 1.5 volte superiore allo slip medio, sulla base di Murotani et al., 2013) e l’area totale della faglia è indipendente dalla magnitudo momento e dalla separazione fra terremoti tsunamigenici e non-tsunamigenici.

La definizione di leggi di scala può essere visto come un metodo per definire caratteristiche ricorrenti nella variabilità dei parametri di faglia e delle distribuzioni di slip per diversi terremoti. In questa tesi, la ricerca di eventuali ricorrenze per quanto riguarda la distribuzione di slip sulle faglie sarà cercata in modo del tutto diverso, ovvero attraverso tecniche di “face detection” che descriveremo in qualche dettaglio nel prossimo capitolo.

2.5 Relazione fra lo slip massimo e la profondità

Una caratteristica che è stata trascurata in letteratura è la dipendenza dei valori caratteristici dello slip (medio, massimo) per i diversi modelli contenuti in SRCMOD dalla profondità. Ciò può essere in parte dovuto al fatto che la profondità è sempre fra i parametri peggio vincolati nel processo di inversione. D’altra parte, è certamente lecito chiedersi se meccanismi focali diversi siano caratterizzati da distribuzioni diverse di slip con la profondità. Inoltre, un problema tuttora aperto, ma che non affronteremo in

questa tesi, riguarda l'eventuale relazione che esiste fra posizione del massimo dello slip e posizione dell'ipocentro del terremoto.

Mediante le informazioni contenute nei files SLP e FSP per ogni modello a faglia finita è stata studiata la dipendenza del massimo valore dello slip dalla sua profondità espressa relativamente alla profondità dell'ipocentro. Il risultato è mostrato in Fig. 2.4.

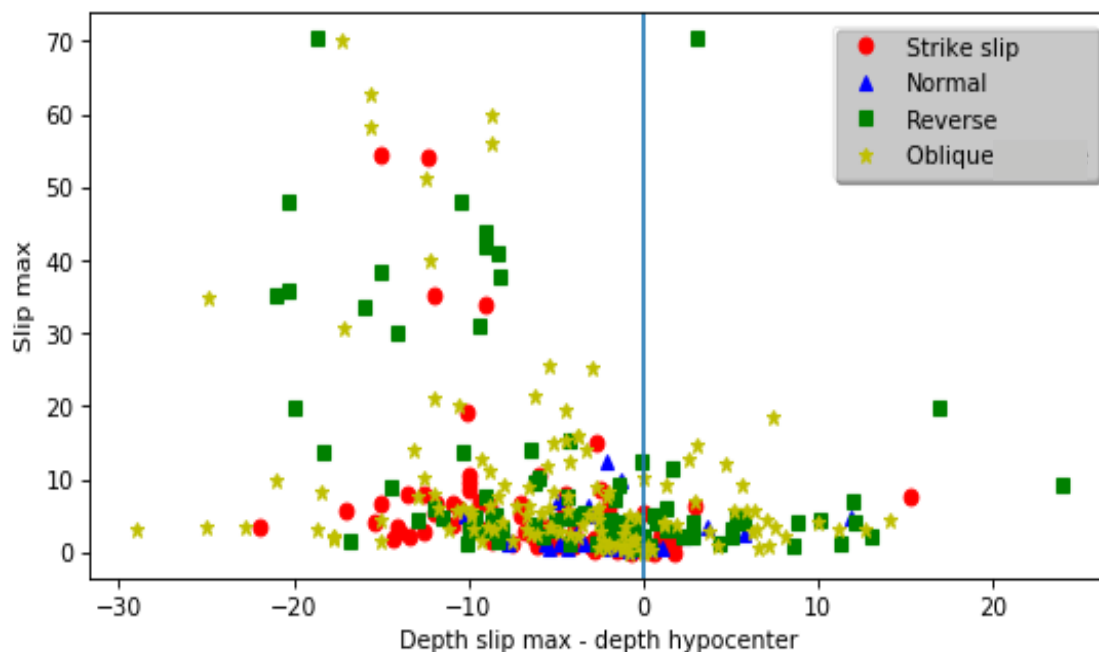


Figura 2.4 Andamento del massimo dello slip di ogni evento in SRCMOD in funzione della profondità data dalla differenza tra la profondità del punto in cui lo slip è massimo e quella dell'ipocentro. Ogni punto è rappresentato da un simbolo diverso a seconda del meccanismo focale.

Gli eventi con meccanismo focale normale sono tutti concentrati in basso e sullo zero, hanno quindi il valore di $slip_{max}$ più basso rispetto agli altri meccanismi e la sua profondità è confinata in una fascia di ± 10 km attorno a quella dell'ipocentro.

I meccanismi di strike slip, tranne che per alcune eccezioni, hanno $slip_{max}$ compreso tra 0 e 15-20 metri, e si trovano nella parte negativa delle ascisse, quindi la loro profondità è minore di quella dell'ipocentro. In altri termini, per questo meccanismo focale il massimo dello scorrimento si verifica quasi sempre a profondità più superficiali rispetto al punto in cui la rottura ha inizio.

Per quanto riguarda gli ultimi due meccanismi focali, “reverse” e “oblique”, la variabilità è molto più marcata. In generale, sono questi i meccanismi per i quali si

ricavano i valori più alti di $slip_{max}$. I terremoti con $slip_{max}$ più profondo dell'ipocentro hanno valori con $slip_{max}$ stesso confinato entro i 20 m, con la notevole eccezione del terremoto di Tohoku dell'11 marzo 2011, che ha raggiunto $slip_{max}$ di circa 70 m ad una profondità comparabile con quella dell'ipocentro. Nel caso dei terremoti caratterizzati da posizioni più superficiali delle celle con lo scorrimento massimo, quest'ultimo è molto più variabile sia come valore che come distanza verticale dall'ipocentro.

2.6 Rappresentazione degli eventi di SRCMOD per l'applicazione degli algoritmi di face detection

Per poter rendere i modelli contenuti in SRCMOD adatti ad essere trattati mediante gli algoritmi di “face detection” che saranno discussi nel capitolo successivo, è stato necessario adottare una rappresentazione puramente grafica degli eventi. In altre parole, in tutta la trattazione che seguirà, ogni evento sarà un'immagine in cui è rappresentato il piano di faglia e la distribuzione di slip su di esso. La strategia da seguire per questo pre-processing, apparentemente banale, costituisce uno dei punti chiave di tutta la procedura. Innanzitutto, si è scelto di lavorare in una geometria in cui tutti i piani di faglia sono verticali e gli assi rappresentano la lunghezza (lungo lo strike) e la larghezza (lungo il dip), con l'origine degli assi posta in alto a sinistra del rettangolo che rappresenta il piano di faglia. Ciò equivale a dire che viene eliminata ogni dipendenza dall'inclinazione del piano di faglia e dalla sua orientazione nello spazio, che al più possono essere salvati come metadati a corredo dell'immagine. La rappresentazione matriciale dello slip fornita da SRCMOD può essere mantenuta anche nella rappresentazione finale dell'immagine, come mostrato nel pannello superiore di Fig. 2.6. Le varie celle hanno dimensioni che ricalcano la discretizzazione spaziale definita in SRCMOD per ciascun evento e il valore dello slip è rappresentato in una scala di colori arbitraria, ma in cui la sequenza dei colori è mantenuta uguale per tutti gli eventi. Intuitivamente, algoritmi di “face detection” necessitano di immagini in cui la definizione sia di qualità “sufficiente”. In un caso come quello mostrato nel pannello superiore di Fig. 2.5, l'immagine è troppo “sgranata” e questo impedisce agli algoritmi di essere efficaci. Data la variabilità del passo di discretizzazione spaziale utilizzato per i diversi eventi di SRCMOD, si è scelto di adottare una rappresentazione “standard” per tutti gli eventi, ottenuta interpolando l'immagine iniziale mediante l'algoritmo “sinc” di

Matplotlib di Python, basato sulla funzione “seno cardinale”. La scelta, basata su criteri puramente qualitativi, è stata fatta dopo aver paragonato i risultati di altri metodi di interpolazione forniti dallo stesso software. Il risultato dell’interpolazione è mostrato nel pannello inferiore di Fig. 2.5.

Un problema che non è stato possibile affrontare in dettaglio nell’ambito di questa tesi è l’effetto sui risultati finali della scelta della particolare palette di colori per rappresentare lo slip. E’ ragionevole pensare che, poiché sequenze diverse di colori possono fare risaltare in modo diverso determinate caratteristiche della distribuzione di slip, gli algoritmi di riconoscimento possano essere influenzati da questa scelta. Il tema sarà oggetto di studio per eventuali sviluppi futuri.

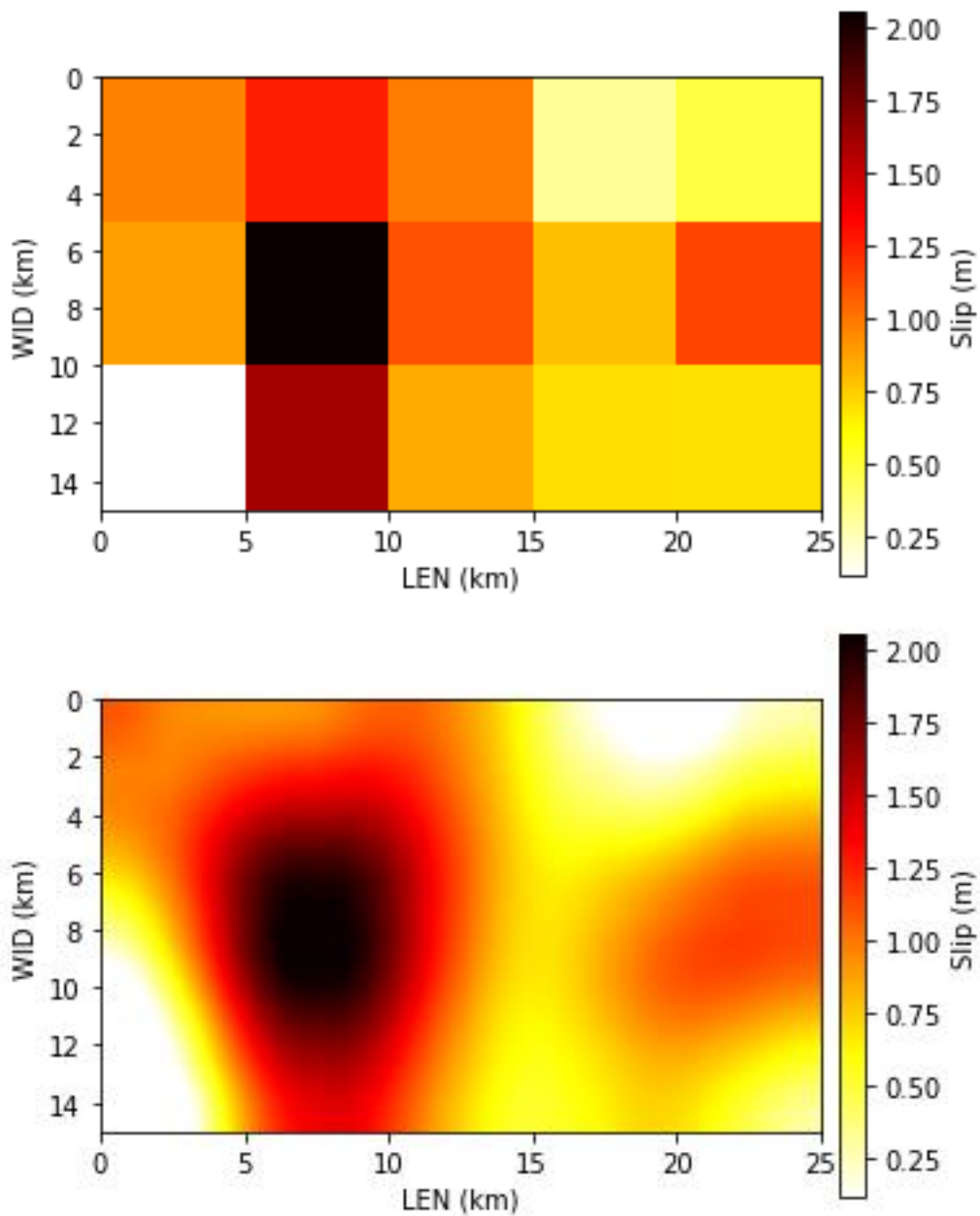


Figure 2.5 Esempio delle due tipologie di immagini tracciate. In alto: immagine ottenuta senza interpolazione quindi con le celle ben evidenziate. In basso: invece l'immagine interpolata. La divisione non è più presente, e l'immagine risulta molto più "nitida". La palette è stata scelta in modo tale che a tonalità di rosso più scuro corrisponda un valore di slip più elevato (evento s1945MIKAWA01KIKU).

3. Algoritmi di “face detection”: il classificatore “Haar cascade”

La “face detection”, e in generale l’utilizzo di classificatori per riconoscere oggetti contenuti in immagini, è uno dei tanti metodi che costituiscono la *Machine Learning*, che identifica l’abilità dei computer di apprendere senza essere state preventivamente programmate. La *Machine Learning* permette di evitare di scrivere il codice di programmazione attraverso il quale si illustra passaggio per passaggio le operazioni da eseguire. Fornendogli solo dei dati in input e informazioni relative ai risultati desiderati, sarà il sistema a indentificare una regola generale che colleghi i dati in ingresso con quelli in uscita, utilizzando la teoria delle probabilità, in modo da poter poi riutilizzare tale regola per nuovi dati fornitigli.

Il suo utilizzo negli ultimi anni in campo sismologico è in continuo aumento, soprattutto da quando l’intensità di dati che si riesce a raccogliere è incrementata e l’impiego di altre fonti (come le tecniche geodetiche tra cui GPS e SAR) consentono un continuo aggiornamento di questi dati, rendendo la loro analisi e elaborazione manuale molto complicata, oltre a tempi di lavorazione sempre più lunghi. La velocità di elaborazione è un fattore molto importante se ad esempio si vogliono monitorare terremoti in tempo reale o per sviluppare i pre allarmi.

L’utilizzo del *Machine Learning* è un approccio totalmente diverso dei metodi utilizzati finora, “imparando dall’esperienza” richiede molti dati, quindi più ne abbiamo a disposizione migliore sarà il suo apprendimento. L’aspetto però più interessante, in questo ambito, data molteplice varietà dei dati che si raccolgono è la possibilità che la machina possa individuare relazioni e modelli a cui finora non si era pensato.

3.1 Il classificatore “Haar Cascade”

Un *classificatore* è un programma che permette di distinguere un’immagine positiva (cioè contenente l’oggetto desiderato) da una negativa e sono impiegati principalmente per il riconoscimento facciale.

Affinché un classificatore sia in grado di riconoscere un determinato oggetto, è necessario “allenarlo” (come mostrato nel cap.3.2) cioè bisogna “insegnargli” a distinguere cosa deve individuare e tutto ciò che, invece, deve ignorare.

Per essere precisi come sarà spiegato in seguito non si parla di un singolo classificatore, ma di una “cascata” di classificatori che seguono le “Haar-like features”, immagini digitali utilizzate nel riconoscimento di oggetti. Il classificatore risultante sarà quindi costituito da più fasi (*stages*) formate da più classificatori più semplici, che sono applicati in modo iterativo. Questo permette di individuare subito le regioni dove è probabile trovare l’oggetto, così da riservare i calcoli più complessi solo su di esse, tramite l’utilizzo dell’algoritmo *AdaBoost*.

Per cercare nell’intera immagine, il classificatore lavora utilizzando una finestra scorrevole, situata in alto a sinistra, che andrà poi a scorrere su tutta l’immagine (è anche possibile ridimensionare la finestra in base alle proprie esigenze, in modo da poter individuare l’oggetto a varie dimensioni).

L’idea di applicare la finestra scorrevole deriva dal lavoro di Viola e Jones (2001), che si concentrò soprattutto sul riconoscimento di volti frontali. In questo caso il classificatore non agisce direttamente sull’immagine, ma su una sua rappresentazione, chiamata *Integral Image*, che come verrà illustrato in seguito permette valutazioni molto più veloci.

3.1.1 Haar-like feature

Le immagini digitali, Haar-like features, sono l’input del nostro classificatore; nella figura 3.1 ne vengono riportati alcuni esempi.

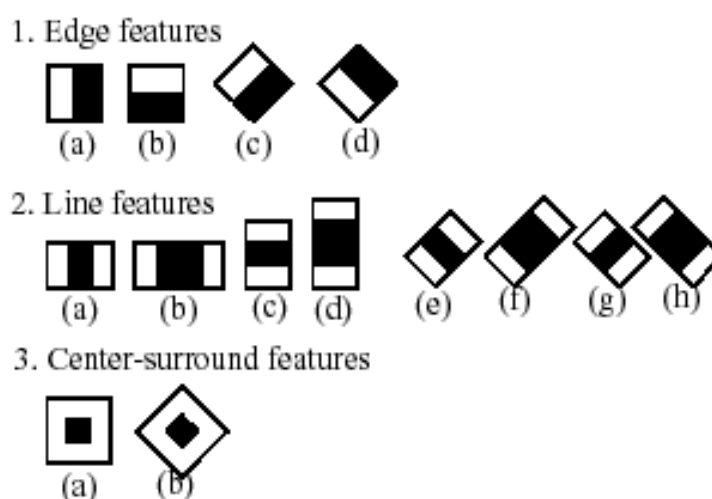


Figura 3.1 Principali filtri di Haar (immagine presa da https://docs.opencv.org/3.3.0/d7/d8b/tutorial_py_face_detection.html).

Ogni filtro è un singolo valore, ottenuto sottraendo la somma dei pixel sotto il rettangolo bianco dalla somma di quelli sotto il rettangolo nero.

In particolare, per il classificatore Haar, si usano tre tipi di features:

- “A due rettangoli”, dove il valore è la differenza tra la somma dei pixel all’interno delle due regioni rettangolari, le quali hanno la stessa taglia e forma e possono essere orizzontali o verticali;
- Quella “a tre rettangoli”, dove la somma dei due rettangoli esterni è sottratta dalla somma dei pixel di quella centrale;
- Quella “a quattro rettangoli” dove il valore è dato dalla differenza tra le coppie di rettangoli opposti.

La Fig. 3.2 mostra un esempio di come le diverse features vengono utilizzate all’interno di una sezione per effettuare il riconoscimento.

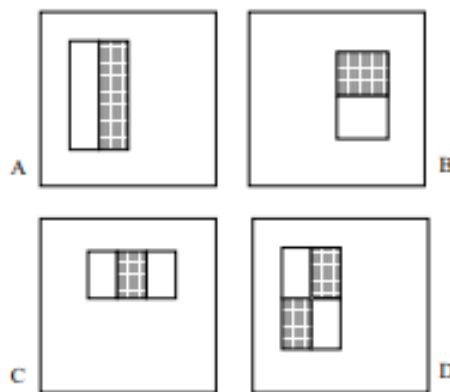


Figura 3.2 Esempi di features rettangolari all’interno della finestra di riconoscimento. La somma dei pixel all’interno della sezione bianca è sottratta dalla somma di quelli neri.

Quindi prendendo in considerazione tutte le varie dimensioni e le possibili posizioni che possono assumere le varie features, si avrebbe bisogno di un gran numero di quest’ultime (per un immagine 24 x 24 risultano più di 160000 features). Esse possono essere calcolate velocemente usando la *rappresentazione integrata dell’immagine*, cioè quella in cui il valore del pixel in una qualsiasi posizione (x,y) è la somma dei valori di tutti i pixel presenti prima di esso: se $I(x,y)$ è l’immagine iniziale e $II(x,y)$ è l’immagine integrata, allora

$$II(x,y) = \sum_{x' \leq x, y' \leq y} I(x',y')$$

A questo punto si può introdurre per ogni pixel la somma cumulativa per riga fino a quel pixel, $s(x, y)$, come

$$s(x, y) = s(x, y - 1) + I(x, y)$$

In modo tale che

$$II(x, y) = II(x - 1, y) + s(x, y)$$

Utilizzando questo metodo qualsiasi somma rettangolare può essere trovata riferendosi a quattro array, come riportato in Fig. 3.3.

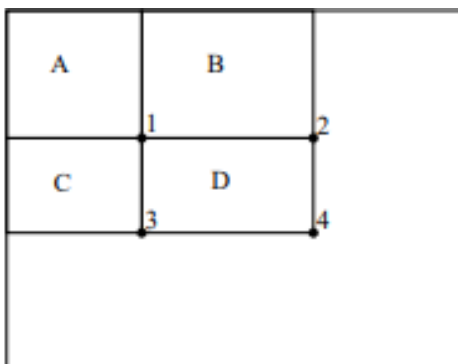


Figura 3.3 La somma dei pixel all'interno del rettangolo D può essere calcolata considerando 4 array. Il valore dell'immagine integrata nel punto 1 è la somma dei pixel in A, il valore in 2 è A+B, in 3 A+C, mentre in 4 è A+B+C+D. Così la somma all'interno di D si può ottenere facendo $4+1-(2+3)$.

3.1.2 L'algoritmo Adaboost

Adaboost è un algoritmo di apprendimento automatico, che ci permette immediatamente di valutare se in una determinata regione dell'immagine è presente l'oggetto ricercato da una regione che sicuramente non contiene tale elemento.

Questo è molto utile in termini di velocità dell'intero processo, infatti utilizzando l'algoritmo, come vedremo nel capitolo successivo, si evita di applicare le features in settori in cui non sarebbe necessario farlo, in quanto porterebbe a informazioni irrilevanti. L'algoritmo iniziale, il più debole, serve proprio a selezionare la singola features che meglio separa le immagini positive da quelle negative e dunque riduce il problema appena citato.

Se consideriamo una serie di immagini $(x_1, y_1), \dots, (x_n, y_n)$, con y_i pari a 0 oppure 1 a seconda che l'immagine sia positiva o negativa. I pesi iniziali sono

$$\omega_{1,i} = \frac{1}{2m}, \frac{1}{2l} \quad \text{per } y_i = 0,1$$

con m numero di negativi e l numero di positivi.

Per $t = 1, \dots, T$ si normalizzano i pesi così che ω_t sarà una distribuzione di probabilità

$$\omega_{t,i} = \frac{\omega_{t,i}}{\sum_{j=1}^n \omega_{t,j}}$$

Per ogni feature j si allena un classificatore h_j , vincolato a utilizzare una singola feature, l'errore associato sarà valutato rispetto ω_t :

$$\epsilon_j = \sum_i \omega_i |h_j(x_i) - y_i|$$

Il classificatore finale h_t scelto sarà quello associato all'errore minore ϵ_t . Aggiorniamo ora i pesi:

$$\omega_{t+1,i} = \omega_{t,i} \beta_t^{1-e_i} \quad \text{con } \beta_t = \frac{\epsilon_t}{1 - \epsilon_t}$$

$e_i = 0$ se x_i è classificato correttamente altrimenti $e_i = 1$. Si può scrivere quindi il classificatore finale come:

$$h(x) = \begin{cases} 1 & \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0 & \text{altrimenti} \end{cases} \quad \text{con } \alpha_t = \log \frac{1}{\beta_t}$$

Permette quindi a ogni iterazione di selezionare una sola caratteristica tramite la quale possiamo direttamente distinguere una sezione negativa da una potenzialmente positiva.

3.1.3 Classificatore Haar come cascata di classificatori

Tra le potenziali sezioni che potrebbero contenere un volto, c'è ne saranno molte in cui non è presente (falsi positivi) e quindi applicare anche qui tutte le features non è necessario. Occorre quindi avere metodo più veloce per verificare se all'interno di una determinata finestra è presente o meno l'oggetto da rilevare, in modo tale da escludere subito tali finestre e concentrarsi solo nelle altre regioni.

Si è quindi sviluppato il metodo “*Cascade of Classifiers*” che raggruppa le features in diversi stadi (normalmente i primi stadi contengono molte meno features), in modo tale da non dover applicare ogni volta tutte le features in ogni finestra. Ogni stadio della cascata è costituito dai classificatori addestrati al riconoscimento delle immagini positive, utilizzando l'algoritmo AdaBoost citato in precedenza.

Il classificatore finale sarà quindi la somma ponderata dei vari classificatori “deboli”, così chiamati perché da soli non sono in grado di classificare l’oggetto.

Se una finestra fallisce il primo stadio, la possiamo quindi scartare immediatamente, in caso contrario le viene applicato un secondo stadio di features e così via. Le finestre a cui saranno applicati tutti gli stadi conterranno un volto e saranno quindi quelle che determineranno il riconoscimento positivo.

In Fig. 3.4 è illustrato il meccanismo della finestra scorrevole, applicando il metodo *Cascade of Classifiers*, attraverso i classificatori più deboli, che utilizzano solo poche features, possiamo immediatamente scartare la parte superiore dell’immagine e applicare gli stadi superiori nella parte restante, questo permette una notevole diminuzione del tempo di classificazione.

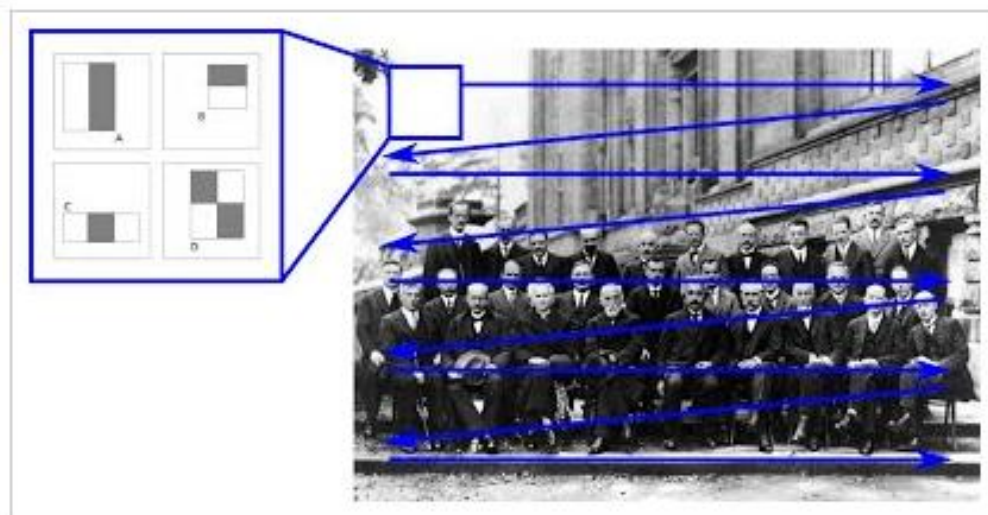


Figura 3.4 Esempio di come le diverse features vengono applicate all’intera immagine tramite una finestra a scorrimento. (immagine presa da http://en.wikipedia.org/wiki/Viola-Jones_object_detection_framework)

La parte negativa di quest’approccio è che se uno stadio classifica erroneamente un falso positivo, è possibile correggere l’errore nelle fasi successive, ma al contrario un falso positivo interrompe la classificazione e non è possibile correggere l’errore.

3.2 L'“allenamento” del classificatore

Finora abbiamo posto l'accento su come il classificatore riesce ad individuare un volto, questo perché, come già citato in precedenza, è impiegato principalmente nella “face detection” e esistono già numerosi classificatori allenati proprio per questo scopo.

Molti di essi sono disponibili tramite il software OpenCV, una libreria software multiplatforma per la visione artificiale che permette l'elaborazione di immagini, object detection e riconoscimento facciale.

Tra le librerie disponibili di OpenCV troviamo quella dedicata al “machine learning” che come già evidenziato permette a un programma di apprendere informazioni direttamente dai dati, senza equazioni o modelli predeterminati, migliorando le proprie prestazioni in modo “adattivo” a mano a mano che gli “esempi” da cui apprendere aumentano.

Questo sarà esattamente lo scopo di questo paragrafo, spiegare come poter allenare un classificatore, fornendogli esempi di immagini dell'oggetto che desideriamo riconoscere. Questa operazione risulta utile quando si vuole identificare oggetti che non sono volti e che quindi non hanno classificatori già allenati in precedenza disponibili in OpenCV.

Per poter essere allenato un classificatore ha bisogno di “imparare” cosa deve riconoscere da quello che invece è lo sfondo, cioè tutto ciò che non è l'oggetto richiesto, occorre quindi fornirgli due categorie di immagini: campioni positivi (contenente l'oggetto da rilevare) e una serie di immagini negative (rappresentanti ogni cosa che non sia quella che vogliamo rilevare).

Le immagini negative (dette anche *immagini di sfondo*) che provengono quindi da immagini arbitrarie possono avere dimensioni diverse, ma ognuna di esse deve essere almeno uguale alla dimensione dell'immagine positiva e serviranno, nella fase di apprendimento automatico, per aiutare il nostro classificatore a capire cosa non deve cercare.

Riportiamo ora brevemente i passaggi eseguiti per allenare un classificatore utilizzando le due tipologie di immagini appena citate.

Una volta scelte le immagini negative e quella positiva occorre convertirle in scala di grigio, poiché è in questo contesto che opera OpenCV, anche il numero di

immagini è importante, maggiore è il numero migliore sarà l'accuratezza del nostro classificatore.

Combinando l'immagine positiva con quelle negative otteniamo i cosiddetti "campioni positivi", che saranno il risultato di diverse combinazioni di distorsioni ridimensionamenti e collocamenti dell'immagine positiva in sovrapposizione ad ogni immagine negativa impiegata (la dimensione consigliata per questi campioni è 20x20 pixel).

I campioni positivi e le immagini negative sono i due elementi fondamentali per poter allenare il classificatore. Tramite il seguente comando, che riportiamo di seguito facendo alcune osservazioni: il numero di campioni positivi deve essere maggiore delle immagini negative per far sì che il classificatore "apprenda" al meglio quello che deve riconoscere, inoltre il numero di immagini negative da indicare deve essere sempre minore di quelle a disposizione in modo tale che se il classificatore dopo aver utilizzato il numero indicato non ha raggiunto la precisione richiesta può autonomamente prenderne altre di quelle non utilizzate per completare al meglio il suo addestramento.

```
opencv_traincascade -data data -vec positives.vec -bg bg.txt -numPos 1800 -numNeg 900 -  
numStages 10 -w 20 -h 20
```

Mentre il processo è in corso, sullo schermo viene visualizzata una schermata simile a quella riportata in Fig 3.5, la quale permette di controllare lo stato di avanzamento dell'allenamento e il numero di immagini utilizzate per quello stadio rapportato al numero totale indicatogli nel comando (indicati rispettivamente con POS e NEG). Mentre nelle tabelle sempre di Fig 3.5 la prima colonna indica il numero di features utilizzate, *HR*, *Hit Rate*, è la percentuale di campioni positivi che sono stati correttamente individuati come tali e *FA*, *False Alarm*, percentuale di falsi positivi trovati.

L'allenamento sarà completo quando verranno raggiunte due soglie che, di default, valgono *-minHitRate 0.995 -maxFalseAlarmRate 0.5*, dove *minHitRate* soglia percentuale che definisce quanti dei campioni positivi possono essere classificati erroneamente come negativi in ogni fase. Quindi su ad esempio 1000 campioni solo 5 possono essere classificati erroneamente su un intero stadio, mentre *maxFalseAlarmRate* soglia che definisce la percentuale massima di falsi positivi permessa in ogni stadio.

```

===== TRAINING 0-stage =====
<BEGIN
POS count : consumed    250 : 250
NEG count : acceptanceRatio    200 : 1
Precalculation time: 4.281
+-----+
| N |   HR |   FA |
+-----+
| 1 |     1 |     1 |
+-----+
| 2 |     1 | 0.025 |
+-----+
END>
Training until now has taken 0 days 0 hours 0 minutes 8 seconds.

===== TRAINING 1-stage =====
<BEGIN
POS count : consumed    250 : 250
NEG count : acceptanceRatio    200 : 0.0487211
Precalculation time: 3.376
+-----+
| N |   HR |   FA |
+-----+
| 1 | 0.996 | 0.02 |
+-----+
END>
Training until now has taken 0 days 0 hours 0 minutes 13 seconds.

===== TRAINING 2-stage =====
<BEGIN
POS count : consumed    250 : 251
NEG count : acceptanceRatio    0 : 0
Required leaf false alarm rate achieved. Branch training terminated.

```

Figura 3.5 Esempio della schermata che appare durante l'allenamento del classificatore.

3.3 L'utilizzo del classificatore allenato per il riconoscimento

Il classificatore allenato, come descritto nel capitolo precedente tramite il software OpenCV, può essere implementato per il riconoscimento tramite il linguaggio di compilazione Python. Come già evidenziato in precedenza, OpenCV opera con immagini in scala di grigi, quindi ogni immagine in cui vogliamo effettuare il riconoscimento deve essere opportunamente convertita in scala di grigi.

Tramite il seguente comando possiamo utilizzare il classificatore allenato a riconoscere l'oggetto da noi cercato. A questo punto come descritto nel capitolo 3.3 una finestra scorrevole delle dimensioni indicate dal parametro *minSize* (in pixel) inizierà a percorrere l'immagine in scala di grigi, analizzando la porzione dell'immagine attraverso le diverse features. Una volta finito la finestra scorrevole ricomincerà

aumentando la sua dimensione di un fattore indicato tramite il parametro *ScaleFactor*, finché non raggiungerà la dimensione dell'immagine da analizzare. Infine il parametro *MinNeighbour*, come spiegheremo in seguito, permette di limitare il riconoscimento di falsi positivi.

```
.detectMultiScale(gray, ScaleFactor=1.3, MinNeighbors=3, MinSize= (50,50))
```

Il risultato della classificazione è l'immagine scelta a colori in cui tramite dei rettangoli saranno individuati gli oggetti riconosciuti. In aggiunta verrà restituita una lista contenente la posizione del pixel in alto a sinistra (x,y) e la larghezza e l'altezza (w,h) del riquadro di delimitazione che circonda l'oggetto. Questi parametri possono quindi essere utilizzati per determinare la posizione dell'oggetto all'interno dell'immagine.

È evidente che la scelta di questi tre parametri influenzano fortemente la qualità e le prestazioni del riconoscimento tramite il classificatore.

Come appena citato lo *ScaleFactor* indica di quanto l'oggetto da riconoscere verrà ridimensionato ogni qual volta la finestra scorrevole ricomincia il suo percorso per analizzare l'immagine. Un valore piccolo quindi farà sì che si abbiano più livelli di ridimensionamento con la conseguenza che si ottengono riconoscimenti multipli per lo stesso oggetto e quindi ad una sovrapposizione dei rettangoli riconosciuti. Possiamo notare questa caratteristica in Fig. 3.6.

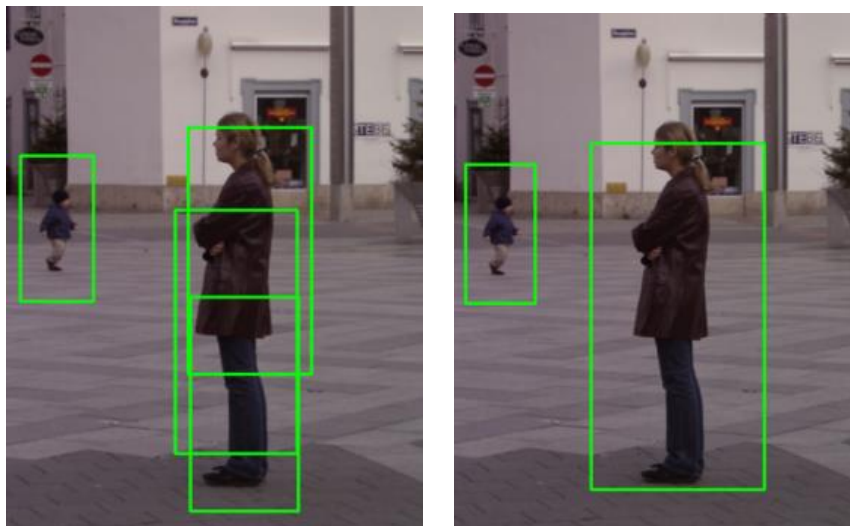


Figure 3.6 Confronto dei riconoscimenti ottenuti con *ScaleFactor* 1.1 (a sinistra) e 1.3 (a destra). Aumentando il valore non abbiamo più sovrapposizioni delle finestre per lo stesso oggetto.

Il parametro *MinSize* invece definisce la dimensione minima che deve avere l'oggetto per essere riconosciuto dal classificatore. La scelta di questo parametro di una dimensione troppo grande, come visibile in Fig.3.7, comporta ad una diminuzione dei riconoscimenti positivi che si avrebbero per parametri inferiori. Infatti un oggetto che verrebbe riconosciuto come positivo ma più piccolo del valore indicato viene automaticamente scartato dal classificatore.

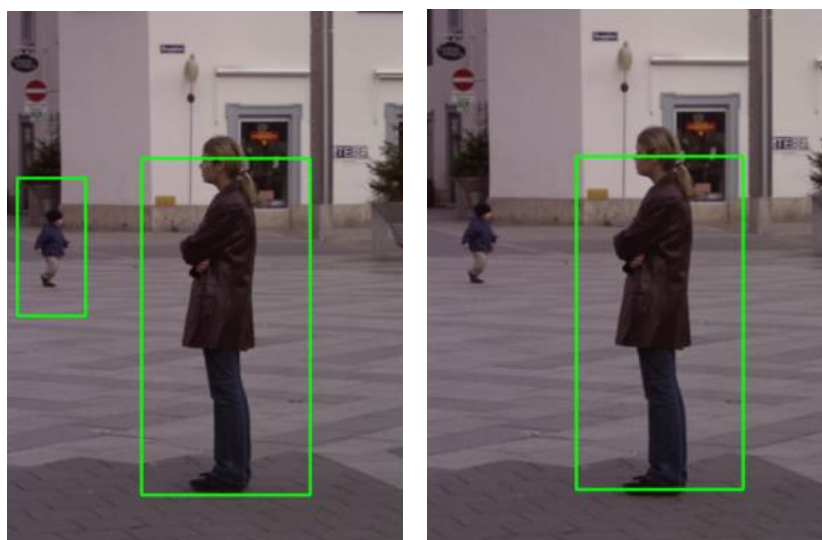


Figura 3.7 Confronto dei riconoscimenti ottenuti con Minsize (4,4) (a sinistra) e (16,16)(a destra). Aumentando le dimensioni del parametro, il bambino viene ignorato dal programma e quindi non più riconosciuto. (immagini prese da <https://www.pyimagesearch.com/2015/11/16/hog-detectmultiscale-parameters-explained/>)

L'ultimo parametro fondamentale per la qualità dei riconoscimenti è *MinNeighbors*, il quale permette, come già citato, di eliminare i falsi positivi rilevati dal classificatore. Questa situazione si verifica principalmente quando si utilizzano valori dei due parametri precedenti piccoli, poichè valori di *MinSize* piccoli comportano un maggior numero di riconoscimenti, mentre per *ScaleFactor* molto bassi si hanno riconoscimenti multipli del medesimo oggetto. Il parametro *MinNeighbors* serve proprio per cercare di risolvere questa situazione, impostando questo parametro con un valore diverso da zero fungerà da soglia inferiore, cioè un oggetto verrà considerato valido se è stato riconosciuto almeno un numero di volte pari al valore assegnato al *MinNeighbors*. L'importanza di questo parametro è evidente in Fig.3.8, nel pannello superiore con *MinNeighbors=1* osserviamo molti riconoscimenti, di cui molti sono falsi positivi, poichè il classificatore considera positivo tutte le porzioni di immagini che

almeno una volta hanno dato esito positivo. Impostando questo valore a 3 (pannello inferiore di Fig3.8) notiamo come il riconoscimento sia notevolmente migliorato, poiché in questo modo viene identificato come volto solo la porzione che è stata riconosciuta almeno 3 volte al passaggio della finestra scorrevole.

Occorre precisare che una scelta sbagliata di questo parametro comporterebbe anche la perdita di riconoscimenti positivi, infatti con valori troppo alti correremo il rischio che oggetti individuati un numero minore di volte rispetto ad altri verrebbero ignorati dal classificatore a causa di questo parametro.



Figura 3.8 Confronto dei riconoscimenti ottenuti con MinNeighbor 1 (sopra) e 3 (sotto). In alto vediamo molti falsi positivi riconosciuti come volti. Impostando il valore a 3, questi spariscono, infatti tutti quegli oggetti che non sono stati riconosciuti almeno 3 volte non vengono più considerati. (immagini prese da <https://johnallen.github.io/opencv-object-detection-tutorial/>)

4. Allenamento di classificatori HAAR sugli eventi di SRCMOD

In questo capitolo discuteremo l'adattamento e l'applicazione della teoria illustrata nel capitolo precedente al riconoscimento di specifiche caratteristiche degli eventi contenuti nel database SRCMOD.

Discuteremo anzitutto quali caratteristiche si è scelto di riconoscere in una data distribuzione di slip e come questa scelta ha guidato la definizione dei classificatori HAAR. Passeremo quindi a descrivere il processo di allenamento per i vari classificatori individuati, effettuato tramite il software OpenCV. Come si è visto nel capitolo precedente, la performance del processo di allenamento dipende in modo critico dalla scelta opportuna di una serie di parametri, che sarà qui discussa nell'applicazione a SRCMOD. Si discuteranno infine alcuni possibili criteri per stabilire quali configurazioni di parametri siano da preferire.

4.1 Scelta delle caratteristiche da riconoscere in una data distribuzione di slip

Come ricordato in precedenza, l'eterogeneità dello slip su una faglia è caratterizzata dalla presenza di asperità, che possiamo considerare come delle “zone” sulla superficie di faglia in cui lo slip tende a concentrarsi. In generale, il numero e la forma di queste asperità non sono prevedibili a priori per un dato terremoto. Il nostro scopo è verificare a posteriori se esistono caratteristiche ricorrenti per il numero e la forma delle asperità in funzione della magnitudo di un dato terremoto, del suo meccanismo focale, della zona geografica in cui è avvenuto.

Quando parliamo di “forma” di una asperità, intendiamo sia una sua rappresentazione mediante una data figura geometrica bi-dimensionale, sia la definizione delle sue dimensioni. Una misura che spesso viene adottata come rappresentativa è l' “aspect ratio”, che possiamo definire come il rapporto fra le due dimensioni caratteristiche della figura geometrica scelta: nel caso di un rettangolo è il rapporto fra la sua altezza e la sua base, per un'ellisse il rapporto fra le lunghezze del semi-asse minore e del semi-asse maggiore.

In questa tesi, utilizzeremo i risultati di Baglione et al. (2017), secondo cui la forma delle asperità e il valore dello slip al suo interno possono essere riprodotti in

modo soddisfacente tramite funzioni gaussiane a sezione ellittica, con semi-assi minore e maggiore descritti tramite leggi di regressione in funzione della magnitudo. Di conseguenza, la forma che è stata scelta in questo studio come base per l'immagine positiva da utilizzare nel processo di allenamento dei classificatori è un'ellisse.

Nel seguito indicheremo con a l'asse maggiore di una data ellisse e con b il suo asse minore. L'aspect ratio per le nostre ellissi sarà quindi $ar = a/b$. Nel processo di allenamento descritto nel seguito sono stati utilizzati sette diversi valori di ar (1, 2, 3, 4, $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$): ciascuna definisce una diversa immagine positiva, su è stato allenato uno specifico classificatore. Evidentemente, $ar = 1$ identifica un cerchio, mentre gli altri valori rappresentano ellissi via via più allungate lungo uno degli assi.

4.2 Allenamento dei 7 classificatori con OpenCV

Come è stato brevemente discusso nel paragrafo 2.6, il successo del processo di allenamento, e quindi di riconoscimento, è fortemente influenzato dalla risoluzione delle immagini su cui il riconoscimento viene effettuato. Lo stesso vale per le immagini positive utilizzate nell'allenamento. Un primo problema che si è posto è stato quindi capire se fosse opportuno interpolare o meno i pixel che definiscono le sette immagini positive corrispondenti ad altrettanti valori di ar .

Un primo tentativo ha coinvolto immagini non interpolate sia per le distribuzioni di slip (come quella in Fig. 4.1) sia per le immagini positive (Fig. 4.2) rappresenta il caso $ar = 1$). Si noti che le immagini positive sono state generate con palette di scale di grigi, poiché, come già ricordato nel capitolo precedente, il software OpenCV può accettare in input immagini a colori, ma opera su immagini convertite in scala di grigi.

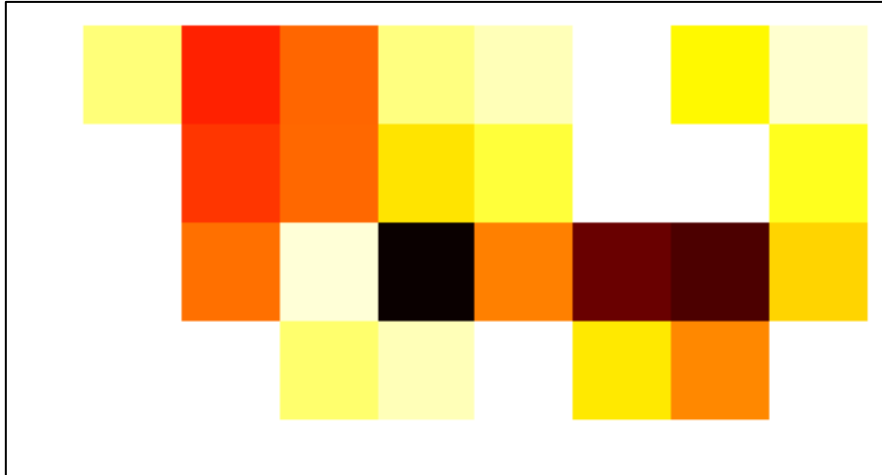


Figure 4.1 Esempio di immagine della distribuzione slip con pixel non interpolati (evento s1946NANKAI01TANI).

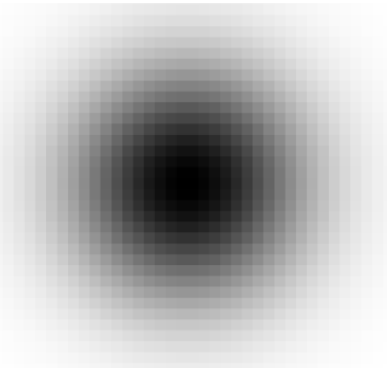


Figure 4.2 Prima scelta di immagine positiva utilizzata

Le immagini negative per l'allenamento sono state create come matrici con numeri random, così da simulare il "rumore" presente nelle distribuzioni di slip degli eventi di SRCMOD. Per garantire una sufficiente variabilità, sono state generate 500 immagini negative: un esempio è fornito in Fig. 4.3.

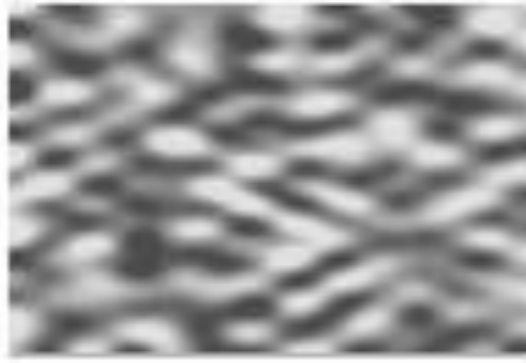


Figure 4.3 Esempio di immagine negativa utilizzata per l'allenamento.

Il tentativo che ha coinvolto immagini non interpolate sia per la distribuzione di slip che per l'immagine positiva si è rivelato fallimentare. Similmente, non si sono ottenuti risultati soddisfacenti in tutti i casi in cui anche solo una fra l'immagine da riconoscere e l'immagine positiva veniva fornita per pixel. Si è quindi scelto di lavorare solo con immagini interpolate, tanto per le immagini da riconoscere quanto per le immagini positive. La Fig. 4.4 illustra la rappresentazione della distribuzione di slip per lo stesso evento riportato in Fig. 4.1, dopo l'interpolazione. Analogamente, la Fig. 4.5 mostra diverse possibili scelte per l'immagine positiva (caso $ar = 1$).

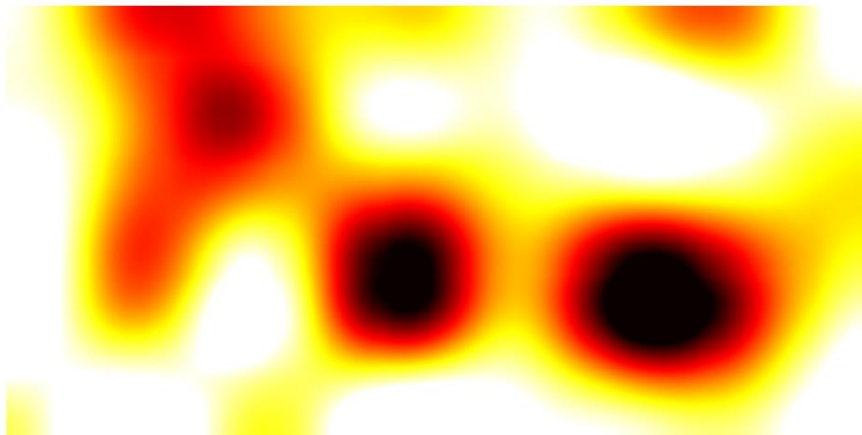


Figure 4.4 Esempio d'interpolazione dell'immagine rappresentante l'andamento dello slip. Saranno quelle su cui poi sarà eseguito il riconoscimento.

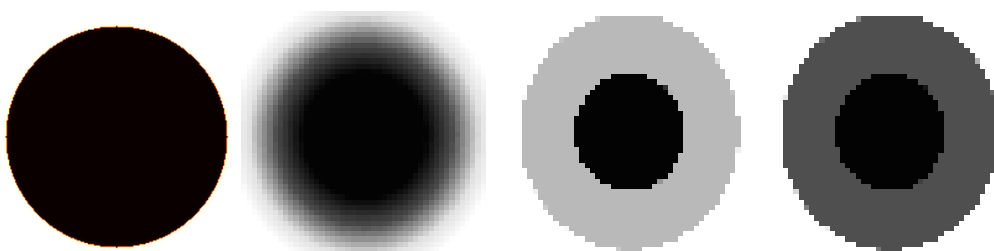


Figura 4.5 Immagini positive utilizzate per l'allenamento del classificatore.

A seguito di vari test che hanno coinvolto ciascuna delle immagini positive illustrate in Fig. 4.5, è risultato che l'immagine positiva che offre il risultato migliore è data da un semplice cerchio nero (immagine di sinistra in Fig. 4.5). Pur essendo un cerchio la forma geometrica rappresentata dall'immagine, l'immagine stessa è una matrice quadrata (rettangolare nel caso delle immagini relative ad $ar \neq 1$), in cui la porzione di area esterna al cerchio deve essere trattata come trasparente: questo si ottiene mediante opportune direttive fornite al software OpenCV.

La Fig. 4.6 mostra tipici esempi di campioni positivi (combinazioni di immagini positive e negative) utilizzati nel processo di allenamento per il caso $ar = 1$.

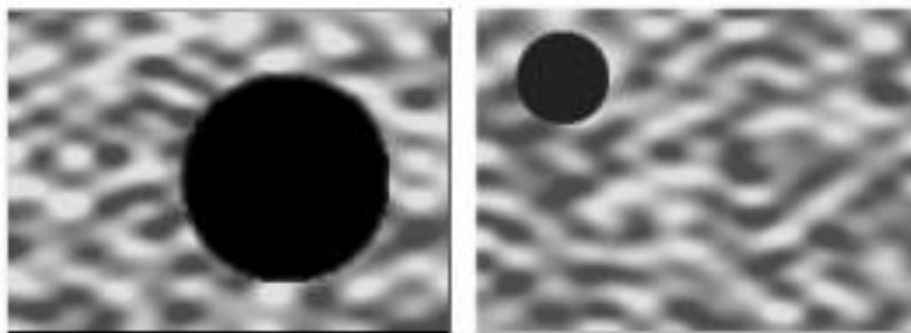


Figura 4.6 Risultati ottenuti dal comando `opencv_createsamples`, combinazioni dell'immagine positiva con quelle negative.

Analogamente, la Fig. 4.7 contiene un esempio di campioni positivi utilizzati per l'allenamento nel caso $ar \neq 1$. In questo caso, oltre alla posizione dell'immagine positiva è stato necessario variare anche il suo angolo di rotazione attorno all'asse perpendicolare al piano della faglia. Si è scelto di consentire al software di operare

rotazioni fino ad un angolo massimo pari a 0.7 radianti (circa 40°), in quanto valori maggiori portavano al troncamento parziale delle immagini e alla loro distorsione.

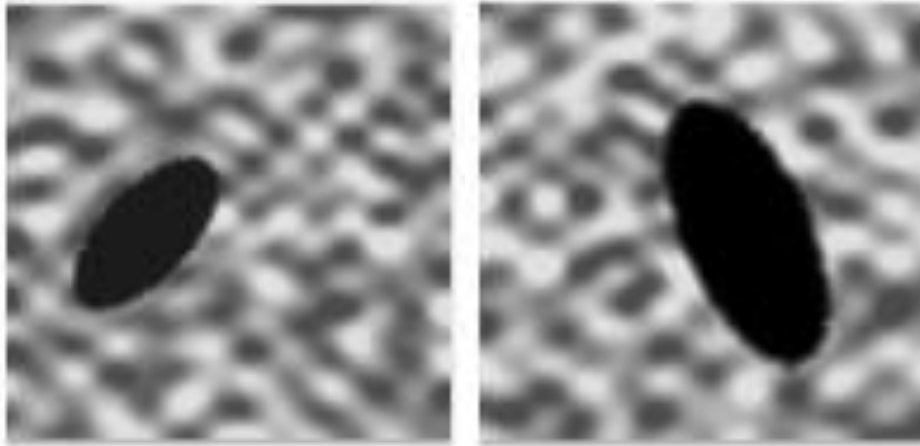


Figura 4.7 Esempi di immagini combinate con immagini positive non circolari ($ar \neq 1$).

La Fig. 4.8 mostra un tipico output di OpenCV relativo al processo di allenamento per uno dei classificatori con $ar \neq 1$. Per il significato dei diversi parametri che compaiono nell'elaborazione si rimanda alle spiegazioni fornite nel Capitolo 3.

```

C:\Users\Elena\misto\Desktop\TESI\OPENCV\OVALE>
+-----+
|  9| 0.995556| 0.106667|
+-----+
END>
Training until now has taken 0 days 0 hours 0 minutes 43 seconds.

===== TRAINING 4-stage =====
<BEGIN
POS count : consumed   450 : 455
NEG count : acceptanceRatio   300 : 0.000654417
Precalculation time: 2.59
+-----+
|  N |   HR |   FA |
+-----+
|  1 |     1|     1|
+-----+
|  2 |     1|     1|
+-----+
|  3 |     1|     1|
+-----+
|  4 |     1| 0.693333|
+-----+
|  5 | 0.995556| 0.466667|
+-----+
|  6 | 0.995556| 0.466667|
+-----+
|  7 | 0.995556| 0.406667|
+-----+
|  8 | 0.995556|    0.11|
+-----+
END>
Training until now has taken 0 days 0 hours 1 minutes 9 seconds.

C:\Users\Elena\misto\Desktop\TESI\OPENCV\OVALE>

```

Figure 4.8 Esempio di come il classificatore si allena. Prima colonna: numero di features utilizzate. HR: percentuale di campioni positivi che sono stati correttamente individuati come tali. FA: False Alarms, numero di falsi positivi trovati

I risultati del processo di allenamento applicato a tutti i classificatori possono essere riassunti nei seguenti punti:

- Rispetto a quanto accade nella “face detection” vera e propria, in cui i classificatori vengono allenati per riconoscere volti, il numero di stadi necessari per i casi di nostro interesse è molto più limitato. Il processo giunge a termine dopo appena 3 stadi nel caso $ar = 1$ e dopo 5 stadi nel caso $ar \neq 1$. Questo è un risultato ragionevole data la relativa semplicità delle nostre immagini positive rispetto ad un qualunque volto.

- Sempre riferendosi ai casi “classici” in cui OpenCV viene applicato, nelle nostre applicazioni per migliorare la precisione del classificatore è stato necessario utilizzare una percentuale massima di falsi positivi permessi in ogni stadio inferiore alla metà del valore predefinito in OpenCV, consigliato per il riconoscimento di volti.
- In termini di tempi di esecuzione, per il singolo classificatore l’allenamento richiede non più di qualche minuto.

Quando i classificatori sono stati allenati, essi possono essere forniti ad OpenCV per effettuare il riconoscimento sugli eventi del database SRCMOD. La Fig. 4.9 mostra un esempio di output grafico fornito dal software nel caso che abbiamo adottato fin qua come esemplificativo: gli “oggetti” trovati sono segnati tramite rettangoli.

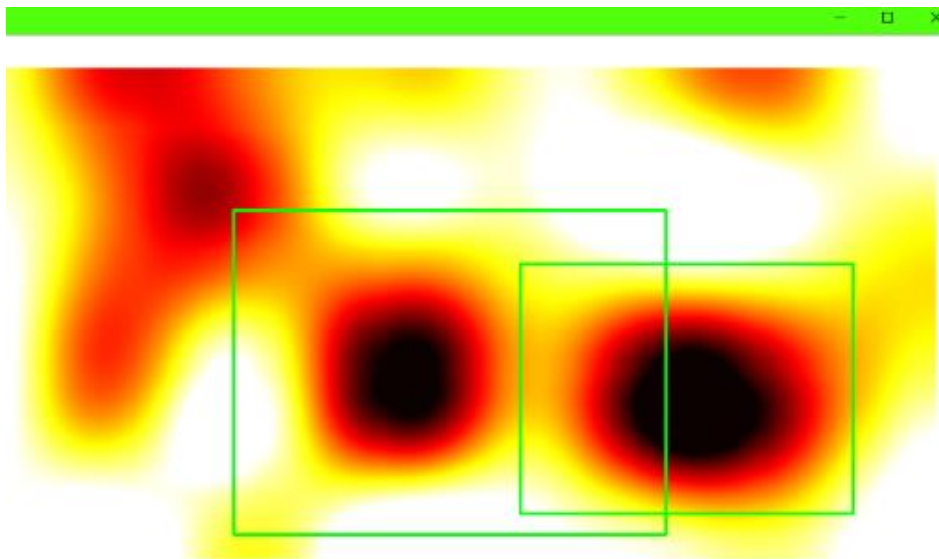


Figura 4.9 Esempio di come lavora il classificatore appena allenato. I rettangoli verdi mostrano gli oggetti riconosciuti.

La Fig. 4.9 mette in evidenza un problema che si è riproposto per numerosi eventi del database SRCMOD. Il software è in grado di riconoscere in modo soddisfacente due asperità, ma è evidente che sono presenti altre due zone di concentrazione significativa di slip (in alto a sinistra e a destra in prossimità del bordo

superiore della faglia) che non vengono riconosciute. Questo può essere imputato ad almeno due cause.

Come abbiamo visto nel capitolo precedente, fissato il numero di classificatori, mediante i parametri *MinSize* e *MinNeighbor* di OpenCV è possibile specificare la dimensione minima che un oggetto deve avere per poter essere riconosciuto, e quante volte deve essere riconosciuto un oggetto nella stessa posizione perchè il riconoscimento vada a buon fine. Tipicamente, al diminuire di *MinSize* aumenta il numero di falsi positivi: questo imporrebbe l'aumento del valore di *MinNeighbor*, che porterebbe però il rischio di perdere i riconoscimenti corretti. Con riferimento alla Fig. 4.9, diminuendo *MinSize* saremmo in grado di riconoscere l'asperità in alto a sinistra, ma rischieremmo di non riconoscere più almeno una delle due asperità principali. In generale, dati i nostri 7 classificatori, non è stato possibile trovare una combinazione di *MinSize* e di *MinNeighbor* che producesse risultati soddisfacenti per tutti gli eventi di SRCMOD.

La seconda causa potrebbe essere legata al tipo di classificatori che abbiamo scelto. E' possibile che allenando ulteriori classificatori con valori di *ar* compresi fra 1 e 2 (e fra ½ e 1) il problema illustrato precedentemente diventi meno severo. L'aumento del numero di classificatori potrebbe essere un tema per futuri sviluppi di questa ricerca.

Per le considerazioni svolte sopra, si è scelto quindi di limitare le analisi successive ai soli eventi di SRCMOD con asperità singola, che sono stati estratti manualmente dal database.

4.3 Applicazione dei classificatori ai modelli di slip a singola asperità

Dei 180 modelli contenuti in SRCMOD, 70 sono stati catalogati "visivamente" (cioè non tramite l'applicazione del riconoscimento con OpenCV) come caratterizzati da una singola asperità. Per completezza, la distribuzione geografica dei 70 modelli è mostrata in Fig. 4.10, mentre in Fig. 4.11 viene fornita la versione aggiornata della Fig. 2.4, relativa alla distribuzione degli slip massimi in funzione della profondità dell'ipocentro.

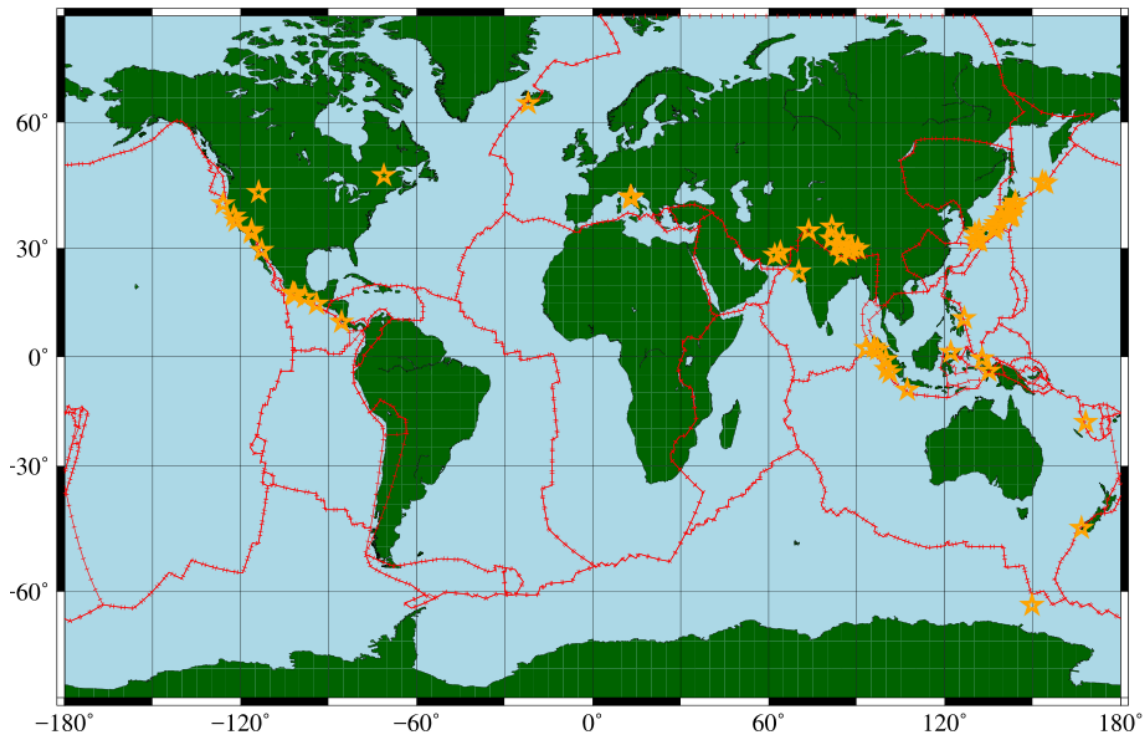


Figura 4.10. Mappa dei 70 modelli a singola asperità presi in esame.

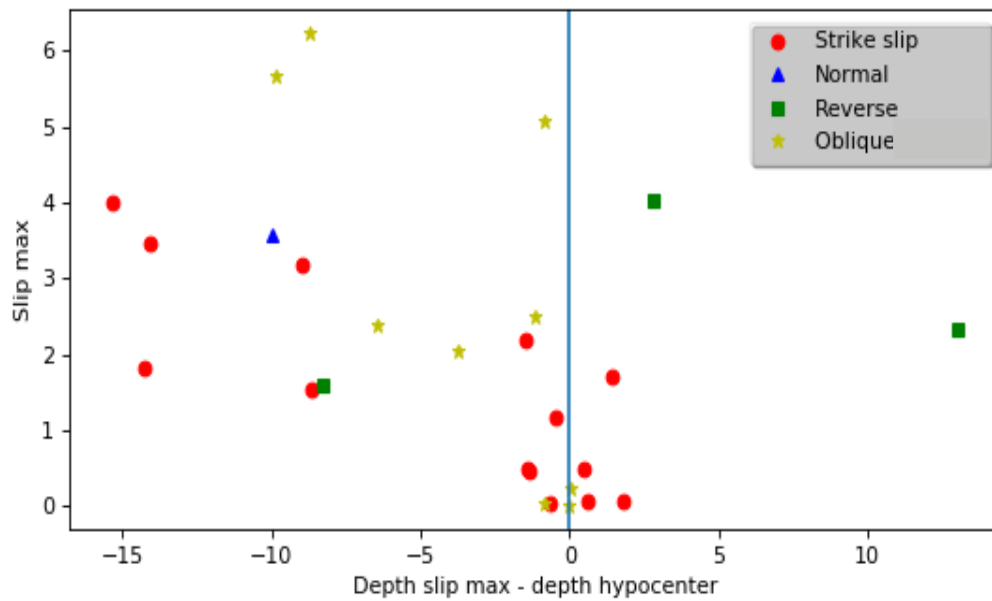


Figura 4.11 Andamento del massimo dello slip di ogni evento in funzione della profondità data dalla differenza tra la profondità del punto in cui lo slip è massimo e quella dell'ipocentro. Ogni punto è basato sul meccanismo focale che lo caratterizza.

Applicando i diversi classificatori di Haar ai 70 eventi selezionati si ottengono i risultati riportati in Tabella 4.1. Il risultato del riconoscimento, che risulta stabile al variare dei parametri *ScaleFactor*, *MinNeighbor* e *MinSize*, è illustrato graficamente per

ognuno dei 70 modelli in Appendice B. I risultati sono forniti in termini del classificatore utilizzato da OpenCV per effettuare il riconoscimento, e quindi in termini di aspect ratio ar riconosciuta per la singola asperità. Il risultato è incrociato con l'informazione sul meccanismo focale predominante di ciascun evento.

Tabella 4.1. Risultato del processo di riconoscimento applicato ai 70 modelli SRCMOD a singola asperità.

	Meccanismo focale	ar
s1945MIKAWA01KIKU	Oblique	1
s1968TOKACH01NAGA	Reverse	2
s1969GIFUxK01TAKE	Strike slip	3
s1978MIYAGI01YAMA	Reverse	1
s1979COYOTE01LIUx	Strike slip	1
s1979PETATL01MEND	Reverse	1/3
s1981PLAYAA01MEND	Reverse	1
s1983BORAHP01MEND	Normal	1
s1985ZIHUAT01MEND	Reverse	1
s1988SAGUEN01HART	Reverse	1
s1992JOSHUA01BENN	Strike slip	1
s1993PUMQUx01WANG	Oblique	2
s1994SANRIK01NAGA	Reverse	3
s1996HYUGAx01YAGI	Reverse	1/2
s1996HYUGAx02YAGI	Reverse	1/2
s1996PUMQUx01WANG	Oblique	1
s1997COLFIO02HERN	Normal	1
s1997KAGOSH01HORI	Strike slip	1
s1997YAMAGU01IDEx	Strike slip	1
s1998ANTARC02ANTO	Oblique	1
s1998HIDASW07IDEx	Strike slip	1
s1998HIDASW10IDEx	Oblique	1/2
s1998IWATEJ01NAKA	Oblique	2
s1998PUMQUx01WANG	Oblique	2
s1999HECTOR01JONS	Strike slip	1
s2000KLEIFA01SUDH	Strike slip	1
s2001BHUIJIN01YAGI	Reverse	1
s2003MIYAGI01MIUR	Strike slip	1
s2003TOKACH01KOKE	Reverse	1
s2004IRIANx01WEIx	Strike slip	1
s2004NIIGAT01ASAN	Reverse	1
s2004ZHONGB01ELLI	Normal	1
s2005FUKUOK02ASAN	Strike slip	1

s2005HONSHU01SHAO	Reverse	1
s2005KASHMI01SHAO	Oblique	2
s2005NORTHE01SHAO	Oblique	2
s2005SUMATR01SHAO	Oblique	1
s2005ZHONGB01ELLI	Oblique	3
s2006JAVAIN01YAGI	Reverse	1
s2006KURILIO1JIxx	Reverse	1
s2007BENGKU02GUSM	Reverse	3
s2007KURILIO1JIxx	Oblique	3
s2008GERZET01ELLI	Oblique	1
s2008GERZET02ELLI	Normal	1
s2008SIMEUL01SLAD	Reverse	1
s2008SULawe01SLAD	Reverse	2
s2008YUTIAN01ELLI	Normal	1
s2008ZHONGB01ELLI	Oblique	1
s2009FIORDL01HAYE	Oblique	1
s2009GULFOF01HAYE	Strike slip	3
s2009PADANG02SLAD	Oblique	1
s2009PAPUAx01HAYE	Oblique	1
s2010SUMATR01HAYE	Reverse	2
s2011HONSHU01SHAO	Oblique	4
s2011OFFSHO01HAYE	Reverse	1
s2011PAKIST01HAYE	Oblique	1
s2011TOHOKU01AMMO	Reverse	3
s2011VANUAT01HAYE	Oblique	1
s2012COSTAR01LIUx	Reverse	1
s2012EASTOF01HAYE	Oblique	1
s2012OAXACA01WEIx	Reverse	1
s2012SUMATR03HAYE	Strike slip	3
s2013KHASHI01WEIx	Normal	1
s2015GORKHA01YAGI	Oblique	2
s2016AMATRIO1PIZZ	Normal	3
s2016NORCIA01PIZZ	Normal	2
s2016USSITA01PIZZ	Normal	1
s2017CHIAPA01OKUW	Oblique	1
s20142014SO01GALL	Oblique	2

La Fig. 4.12 riassume i risultati di Tabella 4.1 in forma di istogramma. Circa il 64% delle asperità viene classificato con $ar = 1$. Percentuali non trascurabili sono ottenute anche per $ar = 2$ e $ar = 3$. I risultati cambiano in modo trascurabile se vengono raggruppate le classi $ar = 2$ e $ar = 1/2$, così come le classi $ar = 3$ e

$ar = 1/3$. La dipendenza dal meccanismo focale sembra essere evidente solo nel caso $ar = 2$ (e $ar = 1/2$), in cui si osserva che gli eventi hanno quasi esclusivamente meccanismo misto (“oblique”) o “reverse” (e sono del tutto assenti meccanismi strike-slip). Osserviamo infine che il singolo caso caratterizzato da $ar = 4$ è quello del terremoto di Tohoku dell’11/03/2011.

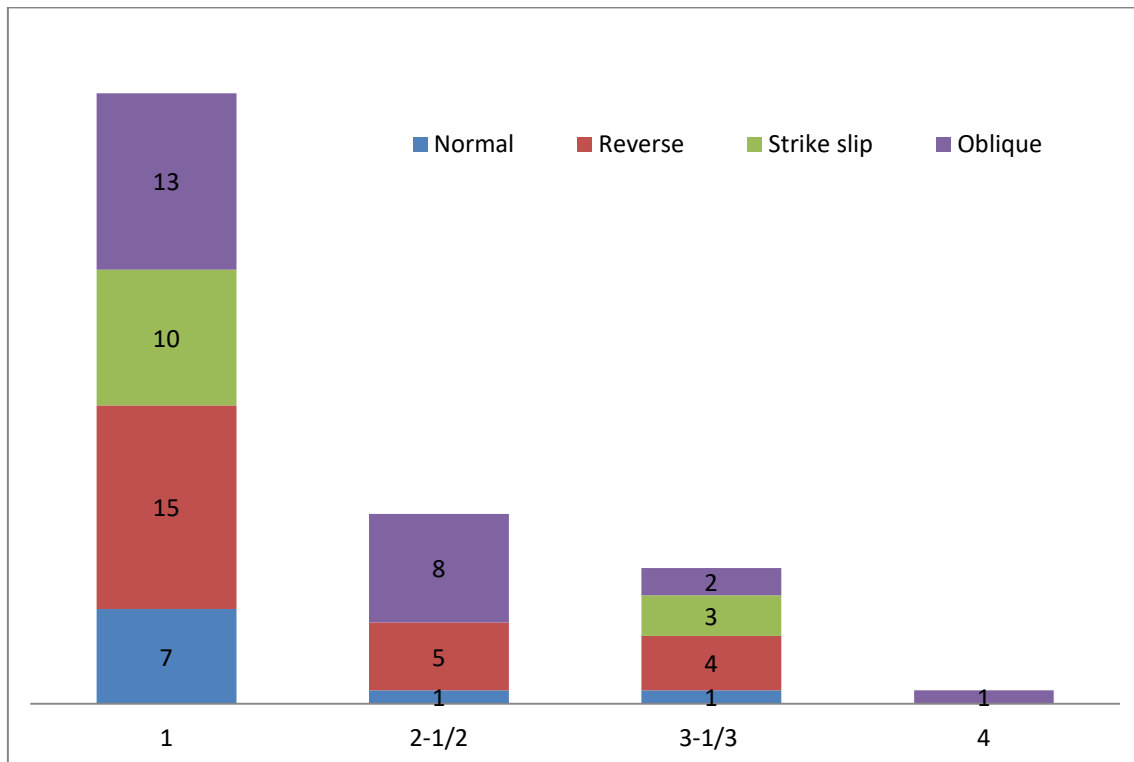


Figure 4.12. Istogramma del numero di riconoscimenti per ogni configurazione, con suddivisione per meccanismo focale predominante. In ascissa è riportato il valore dell’aspect ratio.

La Tabella 4.2 permette di indagare l’eventuale correlazione fra l’aspect ratio che definisce i vari classificatori e l’intervallo di magnitudo dei terremoti riconosciuti da ciascun classificatore. L’informazione viene fornita distinguendo fra i vari meccanismi focali. Non sembra esistere alcuna correlazione evidente fra gli intervalli di magnitudo e il valore dell’aspect ratio: la distinzione per meccanismo focale, a sua volta, non aiuta a riconoscere alcun tipo di caratteristica ricorrente. Questa conclusione è coerente con quanto già trovato da Goda et al. (2016) sulla base dello studio della legge di regressione fra l’aspect ratio e la magnitudo (anche se in quel caso, come già ricordato,

l'aspect ratio era definita sulla base dell'area totale delle celle in cui lo slip è almeno 1.5 volte superiore allo slip medio).

Table 4.2. Magnitudo e meccanismo focale dei 70 terremoti in base all'aspect ratio.

Classificatore (<i>ar</i>)	Valori della magnitudo in base al meccanismo focale			
	Normal	Reverse	Strike-slip	Oblique
1	5.9	5.81	4.41	6.08
	5.97	6.62	5.57	6.4
	6.09	7.25	5.81	6.67
	6.2	7.3	5.87	6.7
	6.82	7.4	5.92	7.2
	7.1	7.4	6.08	7.3
	7.8	7.42	6.1	7.59
		7.5	6.25	7.6
		7.6	7.16	7.6
		7.61	7.2	7.6
		7.66		7.76
		7.82		8.1
		8.21		8.68
	8.3			
2, 1/2	6.5	6.68		4.45
		6.81		6.07
		7.3		6.16
		7.7		6.29
		8.35		6.3
				7.2
			7.6	
			7.9	
3, 1/3	6.21	7.39	6.43	6.2
		7.7	6.9	8.1
		8.5	7.2	
		9		
4				9.1

Un ultimo tentativo per cercare eventuali caratteristiche ricorrenti a partire dai risultati della classificazione riguarda l'analisi della distribuzione geografica dei terremoti riconosciuti dai diversi classificatori. Le Figg. 4.13 – 4.15 mostrano su scala globale i dati di Tabella 4.1. A questa scala, le conclusioni che si possono trarre sono solo qualitative e riguardano solo alcune specifiche aree geografiche. Ad esempio, dal confronto fra le tre figure sembra emergere che i terremoti a singola asperità che

avvengono lungo la zona di subduzione centro-americana siano tutti classificabili secondo $ar = 1$, con l'unica eccezione di un terremoto vicino alla terminazione nord-occidentale dell'area presa in esame. In modo simile, i terremoti con meccanismo normale localizzati nella catena Himalayana e nelle catena dei monti Zagros cadono nella categoria $ar = 1$, così come i terremoti con meccanismo "oblique" nelle Filippine e a Papua Nuova Guinea, in Nuova Zelanda, a Vanuatu. Al contrario, può succedere che terremoti di magnitudo e meccanismo focale comparabili che avvengono in zone geografiche molto limitate, e addirittura su intervalli temporali molto brevi, siano classificati con valori diversi di ar : è il caso dell'Appennino centrale, in cui i tre eventi del 24 agosto 2016, 26 e 30 ottobre 2016 sono classificati dall'algoritmo di riconoscimento con $ar = 3,2,1$ rispettivamente.

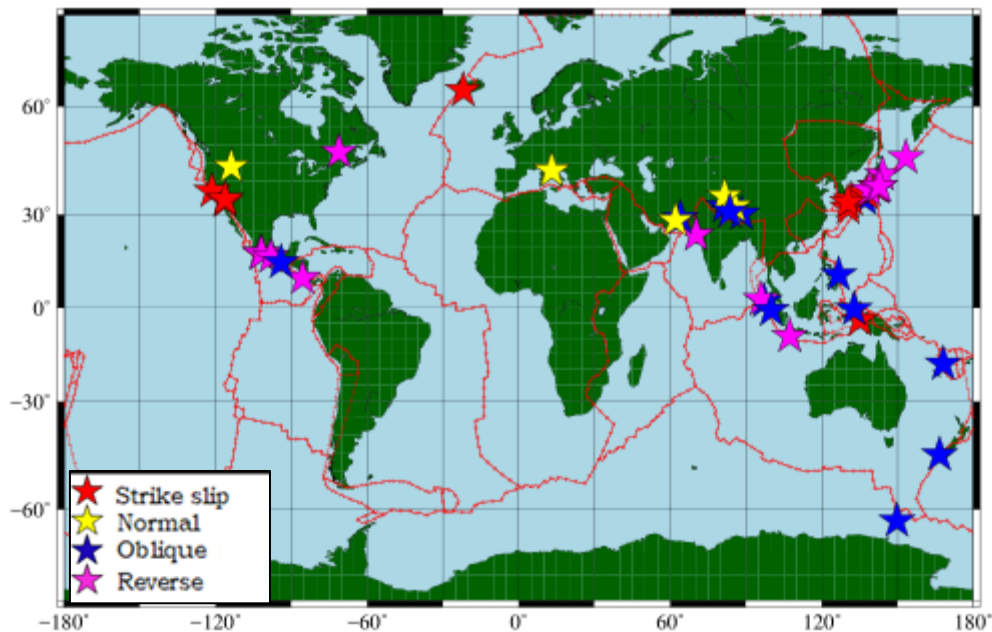


Figure 4.13 Mappa degli eventi classificati con $ar=1$.

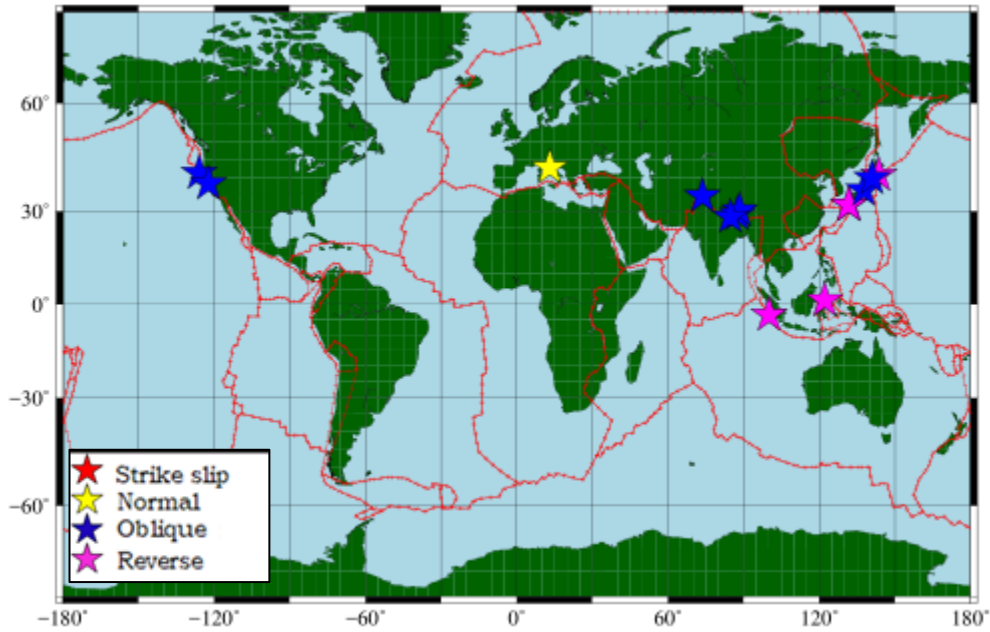


Figure 4.14 Mappa degli eventi classificati con $ar=2$ e $a=1/2$.

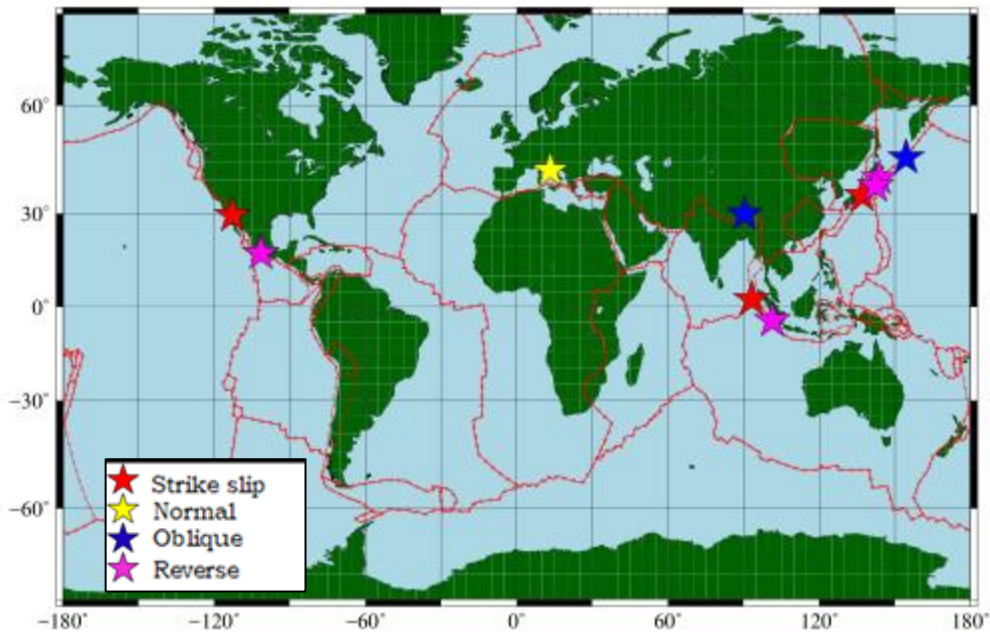


Figure 4.15 Mappa degli eventi classificati con $ar=3$ e $ar=1/3$.

Esistono, infine, alcune zone, come il Giappone, in cui non sembra esistere una correlazione evidente fra aspect ratio e meccanismo focale. La Fig. 4.16 mostra uno “zoom” sull’area dell’estremo oriente che comprende il Giappone e le isole Kurili. Se quasi tutti i terremoti con meccanismo strike-slip hanno epicentro sulla, o vicino alla, terraferma e cadono nella classe $ar = 1$, per i terremoti di subduzione (“reverse” e

“oblique”) non sembra essere possibile una distinzione chiara. E’ possibile che segmenti diversi di una stessa zona subduzione, oppure terremoti generati lungo lo stesso segmento ma a profondità diverse, possano condividere valori simili di aspect ratio. Dato lo scarso numero di eventi a disposizione e la dispersione osservata nei valori di ar, crediamo corretto non trarre conclusioni e limitare le considerazioni fatte a pure ipotesi, eventualmente da verificare più in profondità in futuro.

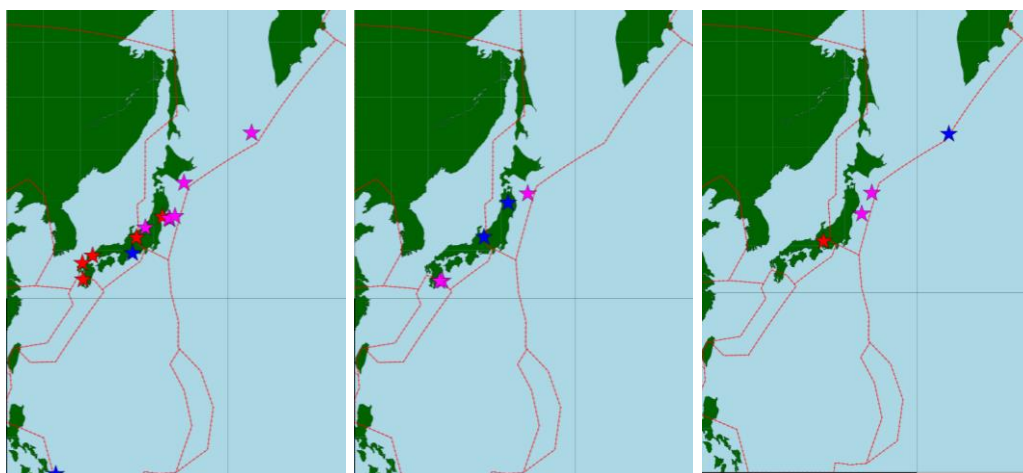


Figure 4.16. Zoom delle Figg. 4.13-4.15 sull’estremo oriente. Da sinistra a destra, i tre pannelli corrispondono a $ar = 1$, $ar = 2$, $ar = 3$.

4.3.1 Effetto dei parametri *ScaleFactor*, *MinNeighbor* e *MinSize*

Nonostante i risultati illustrati in Tabella 4.2 e in Fig. 4.12 per la classificazione dei modelli SRCMOD a singola asperità in base ad ar risultino stabili al variare dei parametri *ScaleFactor*, *MinNeighbor* e *MinSize*, ha certamente senso chiedersi quali effetti la scelta di questi parametri abbia sulla “qualità” del riconoscimento. La Fig. 4.17 illustra tre esempi di altrettanti modelli a singola asperità, con evidenziati gli “oggetti” riconosciuti al variare dei tre parametri. Non sembra esistere un modo univoco per determinare quali valori dei parametri utilizzare per avere il risultato migliore, né il programma è costruito in modo tale da fornirci autonomamente la risposta più adatta. L’unica soluzione, confermata anche dalla consultazione di diversi siti e forum sull’argomento, è effettuare questa scelta sulla base di un approccio “trial and error”.

Per studiare il problema in modo sistematico, si è esplorato lo spazio dei parametri negli intervalli $ScaleFactor \in [1.1, 1.7]$ a passi di 0.1, $minNeighbor \in [1, 9]$ a passi di 1 e $minSize$ pari a 10, 50 o 100. Per ogni combinazione dei parametri si è

eseguito il riconoscimento per i 70 modelli a singola asperità e si è salvato il risultato del riconoscimento stesso. Come mostrato in Fig. 4.17, per ogni configurazione il risultato è rappresentato da un rettangolo circoscritto al particolare classificatore utilizzato dal programma per il riconoscimento. Di ogni rettangolo è possibile quantificare posizione e dimensioni nel sistema di riferimento del piano di faglia.

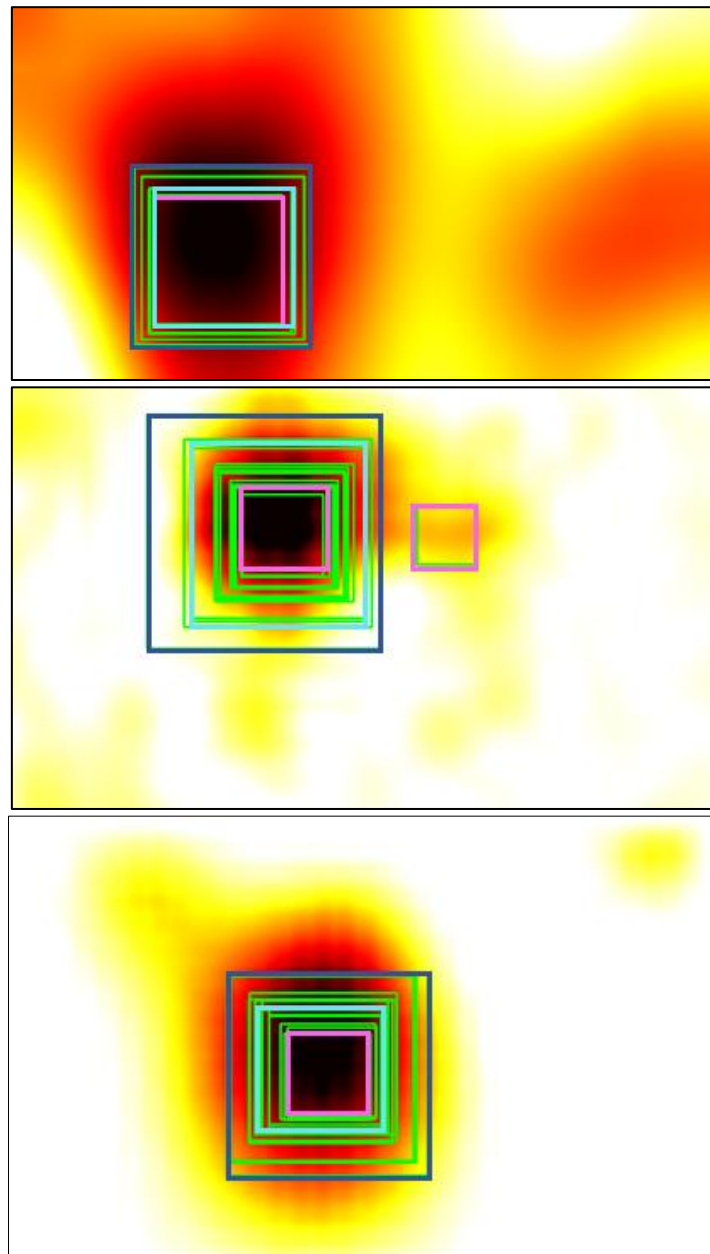


Figura 4.17 Alcuni esempi dei diversi riconoscimenti ottenuti al variare delle configurazioni. Si sono evidenziate alcune configurazioni mantenendo il $MinNeighbor=1$, mentre lo $ScaleFactor$ assume i valori 1.1 (colore rosa), 1.3(azzurro) e 1.4(blu).

Per analizzare criticamente i risultati della “sensitivity analysis” si è cercato di definire alcuni indici, legati a caratteristiche rilevanti della distribuzione di slip, che permettano di quantificare la “bontà” di ciascuna delle configurazioni ottenute. Proponiamo qui due indici, che indicheremo con I_S e con I_D . Il primo è il rapporto fra la sommatoria dei valori di slip contenuti nelle celle del piano di faglia coperte dal rettangolo di riconoscimento ed il valore totale dello slip sulla faglia. Se n_R è il numero di celle coperte dal rettangolo, N è il numero totale di celle nel modello originale e s_j il valore di slip della cella j -esima, allora:

$$I_S = \frac{\sum_{j=1}^{n_R} s_j}{\sum_{j=1}^N s_j}$$

Il secondo indice è invece dato dalla distanza (in km) fra il baricentro del rettangolo di riconoscimento e il centro della cella che nel modello originale è caratterizzata dallo slip massimo. Poichè i diversi modelli di faglia hanno valori di passi di campionamento dx e dz per la discretizzazione in sotto-faglie molto variabili, si è deciso di pesare la distanza per un’opportuna combinazione di questi parametri:

$$I_D = \frac{d}{\sqrt{dx^2 + dz^2}}$$

In altri termini, la distanza è pesata sulla lunghezza della diagonale della singola sotto-faglia per ciascun modello a faglia finita. In questo modo, modelli con discretizzazione molto fine, e quindi con errori “piccoli” sulla posizione dello slip massimo, pesano di più sul valore di I_D dei modelli con discretizzazione più grossolana, in cui la posizione del massimo ha un errore più grande.

Le configurazioni generate nell’ambito della “sensitivity analysis” sono state ordinate separatamente per valori di I_S e di I_D . Nel primo caso, le configurazioni che possiamo definire “migliori” sono quelle per cui I_S è più vicino a 1, mentre nel secondo caso i risultati “migliori” sono quelli per cui I_D è minimo. I risultati dell’analisi sono discussi separatamente per i due indici.

4.3.1.1 Analisi delle configurazioni rispetto a I_S

In questo caso, il criterio per la ricerca delle configurazioni “migliori” privilegia l’informazione relativa alla “quantità di slip” che ciascuna configurazione attribuisce all’asperità analizzata. Per ogni modello di faglia, e quindi a parità di area totale e di

rigidità del mezzo, questo equivale a stimare la frazione di momento sismico rispetto al momento sismico totale che ciascuna configurazione è in grado di riprodurre.

La Fig. 4.18 illustra l'andamento di I_S al variare delle configurazioni per due diversi modelli di faglia. Si tenga conto che la numerazione delle configurazioni è basata sulla sequenza in cui i parametri sono stati fatti variare. In tutto ci sono 189 configurazioni, che possiamo suddividere in 7 blocchi, corrispondenti ai 7 valori presi in considerazione per il parametro *ScaleFactor*, contenenti 27 configurazioni ciascuno. I due grafici di Fig. 4.18 mettono in evidenza almeno due aspetti, oltre a quello ovvio che l'intervallo dei valori di I_S varia a seconda dell'evento che si sta studiando. Il primo aspetto è che molte configurazioni si equivalgono in termini di valori di I_S . Il secondo, ben visibile nel pannello inferiore della figura, è che possono esistere più riconoscimenti, e quindi diversi valori di I_S , per una stessa configurazione. Questo avviene in particolare per valori bassi di *MinNeighbor*: ricordando le osservazioni fatte nel Capitolo 3, la spiegazione è che il riconoscimento va a buon fine per diverse posizioni dell'immagine positiva sull'asperità, ciascuna delle quali può corrispondere a valori diversi di I_S .

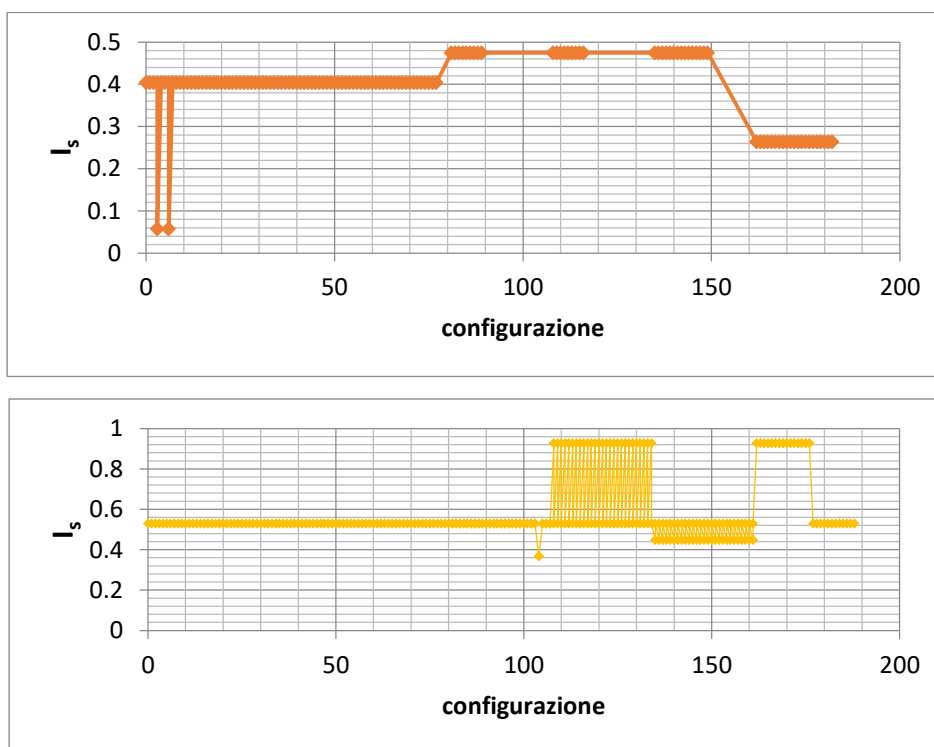


Figure 4.18 Esempi dell'andamento di I_S in base al numero della configurazione per due dei 70 eventi SRCMOD presi in esame (sopra evento s1945MIKAWA01KIKU, sotto s1969GIFUxK01TAKE)

Per ovviare al problema appena discusso, si è scelto di prendere in esame per ogni evento solo le configurazioni con il valore massimo di I_S . Il risultato è illustrato in Fig. 4.19. Esiste un'ampia variabilità del numero di configurazioni per un dato evento a cui corrisponde il massimo di I_S . Le configurazioni sono colorate in modo diverso a seconda del valore massimo di I_S , così da poter “pesare visivamente” i diversi numeri di occorrenza per uno stesso evento. A parte due casi “patologici” (eventi 25 e 27), per cui si ottiene lo stesso valore praticamente per tutte le configurazioni (ma il valore del massimo di I_S è inferiore a 0.5), in tutti gli altri casi il numero di configurazioni equivalenti è al più uguale al corrispettivo di due blocchi da 27 configurazioni, relative a due valori di ScaleFactor diversi. Ci sono anche numerosi casi (come quello dell'evento 63) in cui la variabilità è estremamente limitata e il valore di I_S è prossimo a 1. Infine, è possibile notare come ci sia una massima concentrazione di eventi per cui esiste almeno un riconoscimento in corrispondenza dell'81-esima configurazione.

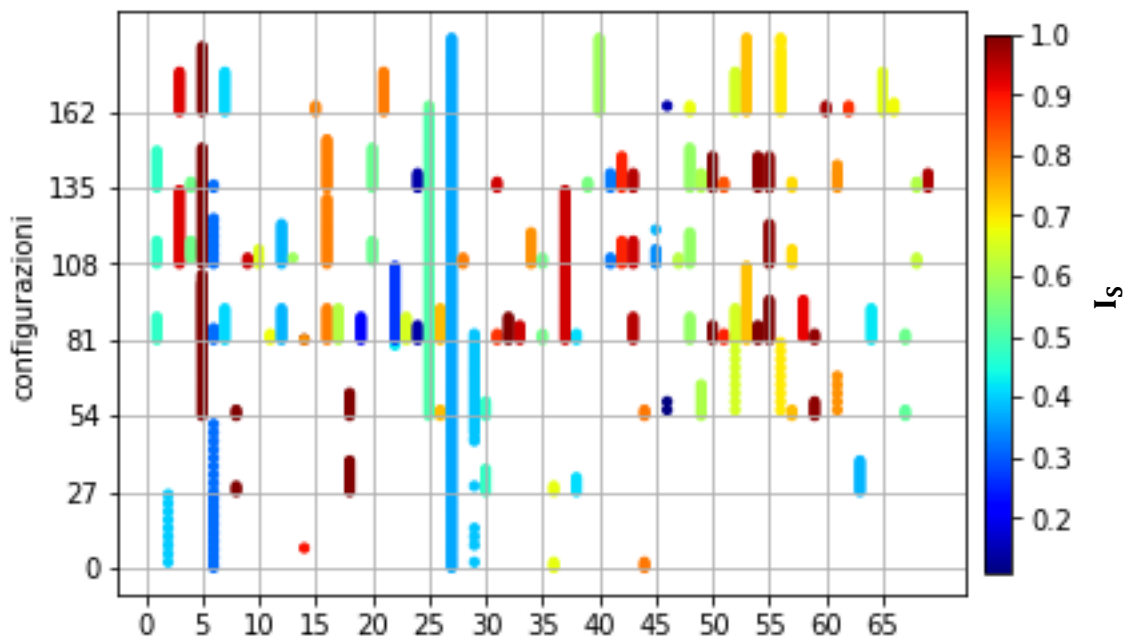


Figure 4.19 Configurazioni corrispondenti al valore massimo di I_S . In ascissa a ogni valore corrisponde un evento (stesso ordine seguito in Tabella 2.1).

Se integriamo le informazioni appena discusse con l'analisi delle immagini corrispondenti per ogni evento (elencate in appendice B), notiamo comunque che esiste una certa variabilità nella qualità dei riconoscimenti anche nel caso in cui il massimo di

I_S appartiene alla classe più alta ($[0.8,1.0]$). Le due conclusioni che si possono trarre sono 1) che, in base al criterio utilizzato, i riconoscimenti migliori avvengono quando la dispersione dello slip intorno all'asperità è molto limitata, 2) che molto probabilmente l'infittimento dei classificatori negli intervalli $ar \in [1,2]$ e $ar \in [1/2,1]$ è necessario per ottenere riconoscimenti più precisi.

La Fig. 4.20 permette di eseguire un'analisi sull'intero insieme dei 70 modelli a singola asperità. Vi è illustrato l'istogramma del numero totale di volte in cui una data configurazione fornisce il massimo di I_S (indipendentemente dal suo valore), che esibisce un picco in corrispondenza delle configurazioni caratterizzate da $ScaleFactor = 1.4$ e valori minimi del parametro $MinNeighbor$. L'andamento decrescente al crescere del valore di $MinNeighbor$ è evidente anche per gli altri valori di $ScaleFactor$. Il valore 1.4 per quest'ultimo parametro potrebbe quindi essere visto come la migliore scelta per l'applicazione del processo di riconoscimento mediante i classificatori che abbiamo definito. Questa conclusione, come vedremo fra breve, è almeno parzialmente confermata anche dai risultati ottenuti nel caso dell'indice I_D .

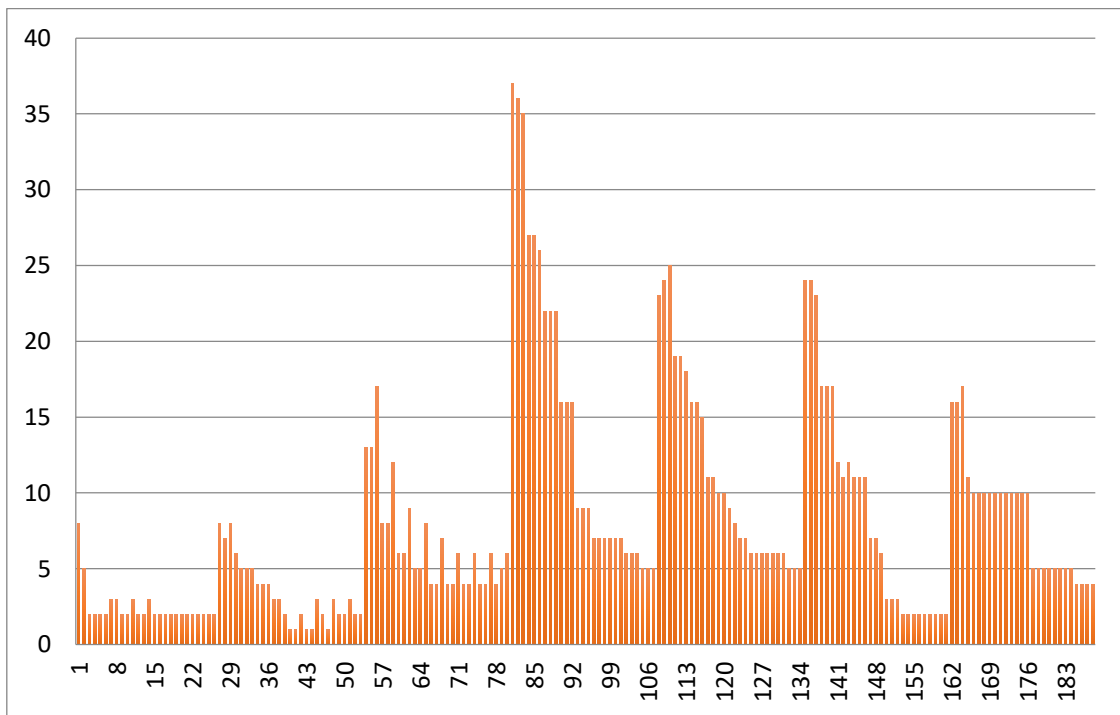


Figure 4.20 Istogramma del numero di occorrenze (in ordinata) in cui I_S è massimo per ogni configurazione (in ascissa), per l'intera lista di eventi.

4.3.1.2 *Analisi delle configurazioni rispetto a I_D*

Passiamo ora al secondo criterio, basato sulla distanza pesata tra il centro del rettangolo riconosciuto e il centro della cella a cui corrisponde il valore massimo di slip nel modello di faglia originale. In questo caso, il criterio per la ricerca delle configurazioni “migliori” privilegia l’informazione relativa alla posizione del rettangolo riconosciuto rispetto alla posizione dell’asperità, cioè la capacità dell’algoritmo di riconoscimento di posizionare correttamente l’asperità sul piano di faglia.

La discussione è condotta sulla falsariga di quella svolta nel paragrafo precedente. La Fig. 4.21 illustra l’andamento di I_D al variare delle configurazioni per due diversi modelli a singola asperità. Evidentemente, in questo caso siamo interessati alle configurazioni che forniscono i valori più bassi di I_D . La Fig. 4.21 mostra una variabilità molto limitata di I_D in entrambi i casi, anche se il valore dell’indice è sensibilmente diverso. Si noti che valori di I_D intorno a 1 indicano distanze fra il centro del rettangolo di riconoscimento e il centro della cella paragonabili alla dimensione caratteristica della cella stessa. Nel caso illustrato nel pannello superiore di Fig. 4.21, a parte pochissime configurazioni il valore di I_D è stabilmente intorno a 0.24, che indica una “centratura” del rettangolo di riconoscimento soddisfacente. Al contrario, per l’evento studiato nel pannello inferiore il valore di I_D si mantiene prossimo a 1 per la grande maggioranza delle configurazioni e scende al massimo fino a 0.84 per un sottoinsieme limitato di casi. Al di là dei valori, la persistenza del valore di I_D al variare delle configurazioni può essere interpretato come un indizio del fatto che I_D è un indice meno “potente” di I_S per la ricerca delle configurazioni che possano essere considerate “migliori” per il processo di riconoscimento.

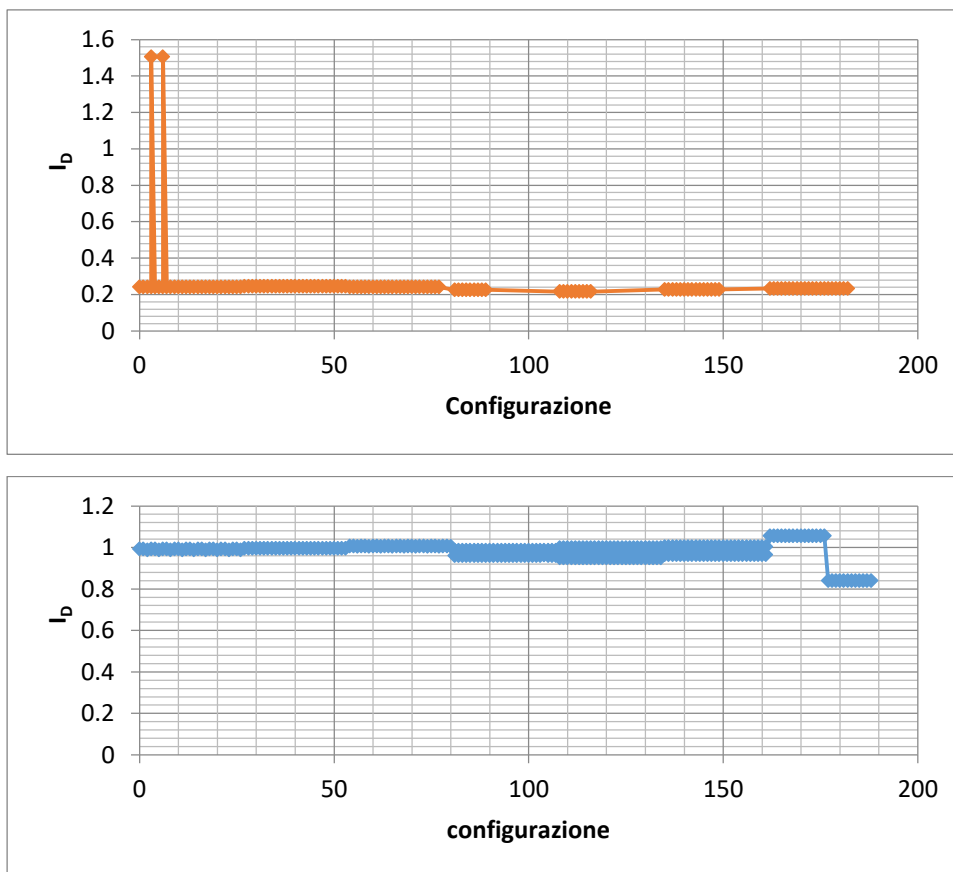


Figure 4.21 Esempi di andamento di I_D al variare della configurazione (sopra evento s1945MIKAWA01KIKU, sotto s1969GIFUxK0ITAKE)

La Fig. 4.22 mostra per ogni evento solo le configurazioni con il valore minimo di I_D . Osserviamo innanzitutto che, con l'eccezione dell'evento 25, tutti gli eventi sono caratterizzati da minimi di I_D inferiori a 2. Analizzando il grafico per righe (cioè per configurazioni), non si apprezzano clusters di eventi sufficientemente lunghi per poter individuare una data configurazione come la più frequente. In tal senso ci viene invece in aiuto (almeno parzialmente) la Fig. 4.23, che contiene l'istogramma del numero totale di volte in cui una data configurazione fornisce il minimo di I_D (indipendentemente dal suo valore) per l'intero insieme dei 70 modelli a singola asperità. Si osservano picchi evidenti di occorrenze in corrispondenza dei gruppi di configurazioni corrispondenti a *ScaleFactor* 1.4, 1.5 e 1.6, con valori decrescenti al crescere di *MinNeighbor* e *MinSize*. Fa eccezione il gruppo con *ScaleFactor* 1.7, in cui la variazione in funzione di *MinNeighbor* e *MinSize* non è univoca.

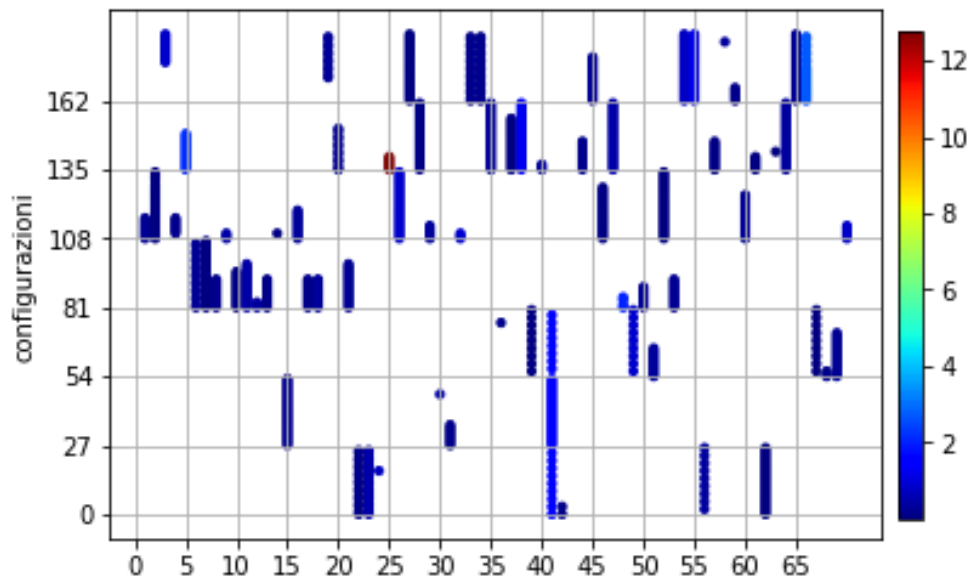


Figure 4.22 Configurazioni in cui è I_D è minimo.

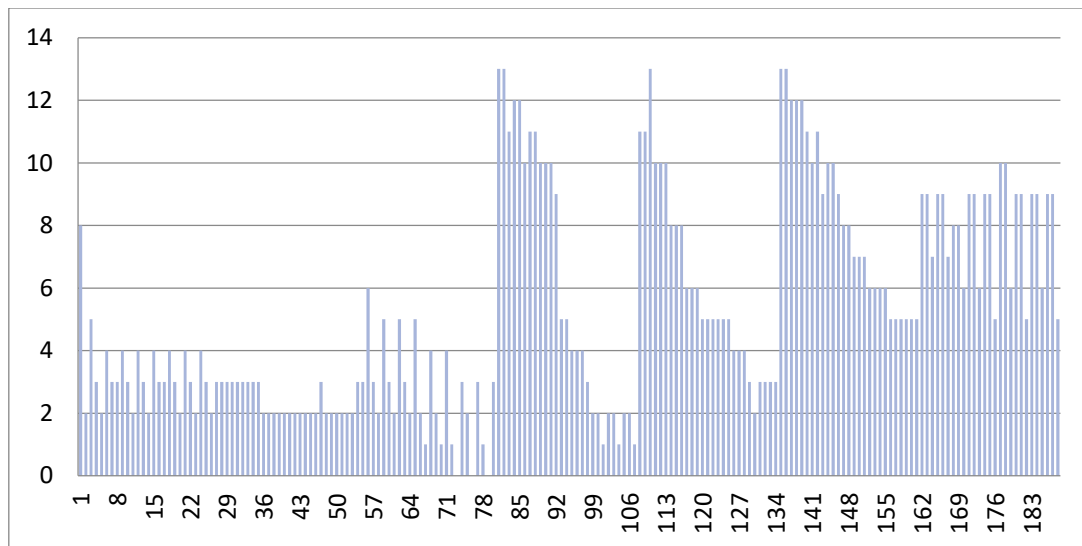


Figure 4.23 Rappresentazione delle occorrenze totali in cui è I_D è minimo.

Incrociando l'informazione desumibile dalla Fig. 4.23 con il risultato discusso in base alla Fig. 4.20, possiamo concludere che il valore di 1.4 per il parametro *ScaleFactor* è quello per cui l'algoritmo di riconoscimento produce "stabilmente" i risultati migliori. Per inciso, questo risultato conferma quando risulta dall'applicazione di OpenCV a campi completamente diversi dalla sismologia, così come desumibile da forums e siti web dedicati. E' più complicato trarre conclusioni definitive sulla scelta dei valori ottimali per *MinNeighbor* e *MinSize*, come già era stato anticipato nel Capitolo 3.

5. Discussione e conclusioni

L'obiettivo che ci si è prefissi in questa tesi è stata la descrizione e la classificazione delle eterogeneità della distribuzione dello slip cosismico sulle faglie mediante tecniche di riconoscimento di volti ("face detection"). A conoscenza dell'autrice, questo è il primo caso in cui questa particolare tecnica viene applicata in sismologia, mentre altre tecniche che possono essere raggruppate sotto l'etichetta generale di "machine learning" hanno già trovato e continuano a trovare applicazioni in sismologia e, più in generale, in geofisica.

L'idea generale è nata immaginando che il possibile riconoscimento di caratteristiche ricorrenti nella eterogeneità dello slip cosismico abbia delle affinità con il problema del rilevamento di un volto. Supponiamo di dover riconoscere in una determinata fotografia, o in un'immagine presa da un video, una persona di cui non abbiamo alcuna immagine oppure non abbiamo immagini "recenti": se abbiamo qualche informazione (avuta per esempio tramite descrizioni di testimoni) riguardante le caratteristiche salienti del volto di quella persona, si può partire da un database di volti, cercare fra questi quelli che hanno caratteristiche simili a quelle della persona cercata e formalizzare quindi un'ipotesi sulla presenza o meno di quella persona nella foto di partenza.

Il nostro approccio verso le distribuzioni eterogenee di slip è stato simile. Siamo partiti da un database di modelli a faglia finita (database SRCMOD), corrispondenti ad un certo numero di terremoti su scala globale, di magnitudo variabile da 4.5 a 9.1, con meccanismi focali diversi fra loro. Piuttosto che cercare se, per ogni terremoto, esiste nel database qualche altro terremoto che sia "identico" a quello scelto in quanto a eterogeneità dello slip, abbiamo definito delle caratteristiche "standard" che tutte le distribuzioni di slip possono avere e abbiamo cercato di catalogare i diversi modelli eterogenei in base alla presenza o meno di quelle caratteristiche.

Le caratteristiche fondamentali delle distribuzioni eterogenee sono state individuate nelle asperità, cioè in quelle zone in cui lo scorrimento relativo fra le facce di una superficie di faglia si concentra, e nella loro forma. Basandosi sulle conclusioni di Baglione et al. (2017), secondo cui le asperità hanno forme che possono essere riprodotte da distribuzioni gaussiane a sezione ellittica, abbiamo scelto un'ellisse come

forma caratteristica da cercare nei modelli a faglia finita di SRCMOD. In particolare, abbiamo posto l'attenzione sull'aspect ratio delle asperità e quindi sul rapporto fra gli assi minore e maggiore delle ellissi utilizzate per rappresentarle.

Nel linguaggio della "face detection", l'ellisse con una data aspect ratio rappresenta un' "immagine positiva", cioè una caratteristica la cui presenza deve essere riconosciuta in immagini all'interno di un database, che può essere più o meno eterogeneo. Una fase fondamentale del processo consiste nell' "allenamento" di un "classificatore", cioè di un algoritmo capace di trovare l'immagine positiva in un'immagine qualsiasi. L' "allenamento" viene svolto su un numero molto elevato di "campioni" ottenuti sovrapponendo l'immagine positiva desiderata ad un numero elevato di "immagini negative", cioè di immagini in cui la caratteristica cercata certamente è assente.

Mediante il software OpenCV, sono stati allenati 7 classificatori diversi, corrispondenti ad altrettanti valori dell'aspect ratio delle asperità ($ar = 1, 2, 3, 4, 1/2, 1/3, 1/4$). I sette classificatori così allenati sono stati applicati inizialmente all'intero database SRCMOD, studiando la qualità del riconoscimento del numero e del tipo delle asperità al variare di alcuni parametri caratteristici del software, chiamati *ScaleFactor*, *MinNeighbor*, *MinSize*: purtroppo non è stato possibile determinare una configurazione dei tre parametri che permettesse un riconoscimento ottimale su tutti i membri del database. Si è quindi scelto di eliminare il problema del conteggio delle asperità e di limitarsi ai soli eventi contenuti in SRCMOD caratterizzati da una distribuzione di slip con una singola asperità predominante. Dei 180 membri di SRCMOD se ne sono quindi selezionati 70 e su questi sono stati effettuati due principali tipi di analisi.

Il primo tipo di analisi è consistito nella classificazione dei vari modelli a faglia finita in base al valore di aspect ratio e nella ricerca di eventuali correlazioni con la magnitudo, il meccanismo focale, la localizzazione geografica. Si è trovato che non esiste una correlazione chiara fra aspect ratio e magnitudo. Questa conclusione è coerente con quanto già discusso in Goda et al. (2016), che era giunto allo stesso risultato calcolando rette di regressione fra il rapporto dell'area dell'asperità (definita come l'area totale delle celle in cui lo slip è almeno 1.5 volte superiore allo slip medio, sulla base di Murotani et al., 2013) e l'area totale della faglia, e la magnitudo momento.

Non si giunge a conclusioni diverse se lo stesso tipo di analisi viene svolta separatamente per i diversi meccanismi focali. Spostando l'attenzione sulla correlazione fra classi di aspect ratio e localizzazione geografica, si possono formulare ipotesi che, dato il numero limitato di eventi e quindi la scarsa significatività statistica delle osservazioni, non possono tradursi in conclusioni. Fra queste ipotesi, citiamo che i terremoti a singola asperità che avvengono lungo la zona di subduzione centro-americana siano tutti classificabili secondo $ar = 1$, cioè che siano caratterizzati da asperità sub-circolari. Un'osservazione simile vale per i terremoti con meccanismo normale localizzati nella catena Himalayana e nelle catena dei monti Zagros, e per quelli con meccanismo misto (definito "oblique" in questa tesi) nelle Filippine e a Papua Nuova Guinea, in Nuova Zelanda, a Vanuatu. Si è anche messo in evidenza che si possono ottenere classificazioni diverse per terremoti di magnitudo e meccanismo focale comparabili che avvengono in zone geografiche molto limitate e su intervalli temporali molto brevi, per quali quindi ci si aspetterebbe una classificazione omogenea: è il caso dei tre terremoti del 24 agosto 2016, 26 e 30 ottobre 2016 in Appennino centrale. Infine, non si può escludere che la classificazione porti a risultati diversi per segmenti diversi di una stessa area tettonica (per esempio un'area di subduzione come quella giapponese e quella del Perù-Cile) o per profondità diverse di nucleazione lungo uno stesso segmento.

Il secondo tipo di analisi è consistito nella ricerca di indicatori che possano aiutare a quantificare la bontà di una dato riconoscimento. Sempre limitandosi al caso dei 70 modelli SRCMOD a singola asperità, si è svolta una "sensitivity analysis" sull'effetto dei parametri *ScaleFactor*, *MinNeighbor*, *MinSize* di OpenCV sul risultato del riconoscimento. Sono stati scelti due indici (numeri) basati rispettivamente sulla percentuale di slip contenuto nell'area di riconoscimento rispetto allo slip totale sulla faglia, e sulla distanza fra il centro dell'area di riconoscimento e il centro dell'asperità nel modello a faglia finita originale. Per entrambi gli indici, indicati come I_S e I_D , l'analisi è stata svolta sia per evento sia globalmente su tutto il database ristretto; inoltre, per aggirare il problema di riconoscimenti multipli per alcune combinazioni dei parametri, si è ristretto lo studio alle sole configurazioni per cui I_S è massimo e I_D è minimo. In base all'analisi svolta su I_S , risulta che il valore ottimale per *ScaleFactor* è 1.4, in perfetto accordo con quanto consigliato dagli sviluppatori e dagli utenti di

OpenCV per applicazioni in campi totalmente diversi. E' più difficile vincolare *MinNeighbor* e *MinSize*, anche se i risultati migliori sono stati ottenuti più frequentemente per i valori più bassi fra quelli presi in esame per entrambi. I risultati ottenuti sulla base di I_D vanno nella stessa direzione, anche se le conclusioni sono molto meno nette che nel caso precedente.

Le analisi descritte in questa tesi portano a concludere che l'approccio alla caratterizzazione delle eterogeneità dello slip cosismico mediante tecniche di face detection, oltre a rappresentare una novità nel campo sismologico, può portare a risultati interessanti, soprattutto se inserito in un approccio integrato con altre strategie.

Evidentemente, il lavoro presentato qui rappresenta solo un punto di inizio. Numerose sono le direzioni di ricerca lungo cui orientarsi per gli sviluppi futuri dell'applicazione della metodologia. Tralasciando la necessità di poter operare in futuro su database sempre più ricchi di eventi, più completi sia a livello geografico che per quanto riguarda la copertura delle magnitudo moderate (fino a 7), citiamo qui tre possibili spunti. Il primo riguarda il raffinamento del numero di classificatori in OpenCV, che equivale ad un aumento dei valori di aspect ratio per le asperità da riconoscere. La classificazione che è stata adottata può essere resa più dettagliata almeno nell'intervallo $ar \in [1,2]$ (e quindi anche $ar \in [1/2,1]$). Chiaramente, va trovato un compromesso fra tale raffinamento e l'esigenza di non far perdere di significato alla classificazione. Si ritiene che un tentativo in tal senso vada fatto e che esso possa portare ad un miglioramento dei riconoscimenti.

Un secondo spunto è legato al modo in cui è stato effettuato il pre-processing sulle immagini dei modelli a faglia finita. OpenCV opera su immagini interpolate a scala di grigi. La conversione in questa scala è svolta automaticamente dal software sulla base dell'immagine che viene fornita in input, che può essere anche a colori. E' ragionevole pensare che il modo di rappresentare la variazione dello slip sul piano da faglia a livello puramente cromatico (la "palette") possa incidere sulla qualità del riconoscimento. Data la scelta di ellissi di diverse aspect ratio come caratteristiche da riconoscere, si può pensare di preparare le immagini in modo tale che le asperità siano evidenziate in partenza tramite l'adozione di opportune mappe di colore. Ad esempio, seguendo Murotani et al. (2013), si potrebbero saturare le palette al di sopra di un valore di slip che rappresenti una certo multiplo dello slip medio sulla faglia. Questo

permetterebbe di stabilire anche una più diretta connessione fra la qualità del riconoscimento e il momento sismico associato al riconoscimento stesso.

Infine, in parte collegandosi a quanto appena osservato, si potrebbero cercare ulteriori indici da utilizzare per l'ottimizzazione del processo di riconoscimento e quindi per migliorarne la qualità.

APPENDICE A – II database SRCMOD

Earthquake ID	Region	Date	Flinn-Engdahl Region	Mag (Mw)	Lat (°N)	Lon (°E)	Dep (km)	Tsunami (yes/no)	Author
2017080901	Chiapas, Mexico	08/09/2017	Off Coast Of Chiapas, Mexico	8.1	14.85	-94.11	18	Y	<u>Okuwaki and Yagi (2017)</u>
2016301001	Norcia	30/10/2016	Central Italy	6.5	42.84	13.08	9.5	N	<u>Pizzi et al, (2017)</u>
2016261001	Ussita	26/10/2016	Central Italy	6.09	42.88	13.1	7.5	N	<u>Pizzi et al, (2017)</u>
2016240801	Amatrice	24/08/2016	Central Italy	6.21	42.7063	13.2532	5	N	<u>Pizzi et al, (2017)</u>
2016150401	Kumamoto, Japan	15/04/2016	Kyushu, Japan	7.04	32.7545	130.763	12.5	N	<u>Asano and Iwata (2016)</u>
				7.1	32.755	130.763	9		<u>Yagi et al, (2016)</u>
2016140401	Kumamoto, Japan	14/04/2016	Kyushu, Japan	6.14	32.7417	130.8087	11.4	N	<u>Asano and Iwata (2016)</u>
2016240101	Alaska (Southern)	24/01/2016	Southern Alaska	7.12	59.6585	-153.4521	127.8	N	<u>Hayes (NEIC, Alaska South 2014)</u>
2015071201	Murghab, Tajikistan	07/12/2015	Tajikistan	7.19	38.211	72.78	22	N	<u>Sangha et al, (2017)</u>
2015160901	Illapel, Chile	16/09/2015	Near Coast Of Central Chile	8.3	-31.637	-71.741	25	Y	<u>Okuwaki et al, (2016)</u>
2015250401	Gorkha, Nepal	25/04/2015	Nepal	7.9	28.147	84.708	15	N	<u>Yagi and Okuwaki (2015)</u>
2014240801	2014 South Napa	24/08/2014	Northern California	6.07	38.22	-122.313	10	N	<u>Gallovic (2016)</u>

2014010401	Iquique, Chile	01/04/2014	Near Coast Of Northern Chile	8.1	-19.6423	-70.817	21.6	Y	Wei (Caltech, Iquique 2014)
2013171101	Scotia Sea	17/11/2013	Scotia Sea	7.7	-60.3	-46.4	10.7	Y	Hayes (USGS, Scotia Sea 2013)
2013240901	Balochistan, Pakistan	24/09/2013	Pakistan	7.7	26.87	65.325	-12.2	Y	Avouac et al, (2014)
2013060201	Santa Cruz Islands	02/06/2013	Santa Cruz Islands	8.06	-10.738	165.138	12.7	N	Lay et al, (2013)
2013240501	Okhotsk Sea	24/05/2013	Sea Of Okhotsk	8.3	54.874	153.281	608	N	Ye at al, (2013)
				8.39	54.874	153.281	620		Wei et al, (2013b)
2013160401	Khash, Iran	16/04/2013	Southwestern Pakistan	7.8	28.1134	62.048	83.4	N	Wei (Caltech, Khash 2013)
2012281001	Masset, Canada	28/10/2012	Queen Charlotte Islands Region	7.83	52.769	-131.927	17	Y	Wei (Caltech, Masset 2012)
				7.72	52.769	-131.927	17		Shao and Ji (UCSB, Masset 2012)
				7.82	52.7881	-132.1011	15.9		Lay et al, (2013)a
2012050901	Costa Rica	05/09/2012	Costa Rica	7.57	10.1017	-85.3446	39.3	Y	Hayes (NEIC, Costa Rica 2012)
				7.6	9.76	-85.56	13		Yue et al, (2013)
				7.6	9.76	-85.56	13.1		Liu et al, (2015)
2012310801	East of Sulangan, Philippines	31/08/2012	Philippine Islands Region	7.59	10.816	126.6725	34	Y	Hayes (USGS, Philippines 2012)

2012270801	Offshore El Salvador	27/08/2012	Off Coast Of Central America	7.3	12.2596	-88.5224	19.9	Y	<u>Hayes (NEIC, Offshore El Salvador 2012)</u>
2012260801	Brawley Swarm	26/08/2012	Southern California	5.45	33.019	-115.537	6.4	N	<u>Wei et al, (2013)</u>
2012260802	Brawley Swarm	26/08/2012	Southern California	5.38	33.0152	-115.548	4.9	N	<u>Wei et al, (2013)</u>
2012110401	Sumatra	11/04/2012	Off W Coast Of Northern Sumatra	8.57	2.311	93.063	22	Y	<u>Wei (Caltech, Sumatra 2012)</u>
				8.64	2.31	93.06	22.9		<u>Shao et al, (UCSB, sumatra 2012)</u>
				8.56	2.3658	93.0812	21.4		<u>Hayes (NEIC, Sumatra 2012)a</u>
				8.56	2.3554	93.0545	21		<u>Hayes (NEIC, Sumatra 2012)b</u>
				8.7	2.31	93.06	30		<u>Yue et al, (2012)</u>
2012200301	Oaxaca, Mexico	20/03/2012	Near Coast Of Guerrero, Mexico	7.4	16.662	-98.187	20	Y	<u>Wei (Caltech, Oaxaca 2012)</u>
				7.4	16.6478	-98.1829	19.7		<u>Hayes (NEIC, Oaxaca 2012)</u>
2012100101	sumatra	10/01/2012	Off W Coast Of Northern Sumatra	7.2	2.4007	93.165	18.4	N	<u>Hayes (NEIC, Sumatra 2012)</u>
2011231001	Van, Turkey	23/10/2011	Turkey	7.1	38.7511	43.4698	15.2	N	<u>Hayes (NEIC, Turkey 2011)</u>

				7.13	38.628	43.486	20		<u>Shao and Ji (UCSB, Turkey 2011)</u>
				7.2	38.7511	43.4698	15.2		<u>Altiner et al, (2013)</u>
				7.1	38.73	43.43	16		<u>Utkucu 2013</u>
				7.1	38.708	43.366	15.3		<u>Elliott et al, (2013)</u>
				7.1	38.691	43.497	19		<u>Konca (2015)</u>
2011211001	Kermadec Islands	21/10/2011	Kermadec Islands Region	7.4	-28.9326	-176.0346	32.1	Y	<u>Hayes (NEIC, Kermadec Islands 2011)</u>
2011200801	Vanuatu	20/08/2011	Vanuatu Islands	7.3	-18.2714	168.0602	31.4	N	<u>Hayes (NEIC, Vanuatu 2011)</u>
2011060701	Kermadec Islands	06/07/2011	Kermadec Islands Region	7.3	-29.2967	-176.1995	19.1	Y	<u>Hayes (USGS, Kermadec Islands 2011)</u>
2011110401	Fukushima Hamadori, Japan	11/04/2011	Near East Coast Of Honshu, Japan	6.68	36.952	140.686	7	N	<u>Tanaka et al, (2014)</u>
2011110301	Tohoku-Oki, Japan	11/03/2011	Near East Coast Of Honshu, Japan	9	38.1	142.85	25	Y	<u>Ide et al, (2011)</u>
				9	38.0373	142.842	20.4		<u>Wei and Sladen (Caltech, Tohoku 2011)</u>
				9.11	38.103	142.86	22		<u>Yagi and Fukahata (2011)a</u>

				9.12	38.1	142.8	21	<u>Wei et al, (Caltech: Tohoku 2011)</u>
				9.09	38.19	142.68	21	<u>Wei et al, (2012)</u>
				9.1	38.322	142.369	24	<u>Shao et al, (2011)</u>
				9.09	38.322	142.369	24	<u>Shao et al, (2011)</u>
				9.1	38.1	142.86	23	<u>Shao et al, (2011)</u>
				9.1	38.1	142.86	23	<u>Shao et al, (2011)</u>
				9	38.3	142.34	28	<u>Ammon et al, (2011)</u>
				9.05	38.3509	142.4562	28.8	<u>Hayes (2011)</u>
				9	38.147	142.915	17	<u>Lay et al, (2011)</u>
				9	38.107	142.916	18.3	<u>Yamazaki et al, (2011)</u>
				9	38.104	142.861	13.5	<u>Gusman et al, (2012)</u>
				9	38.104	142.861	13.5	<u>Gusman et al, (2012)</u>
				9	38.1035	142.861	20.3	<u>Fujii et al, (2011)</u>
				9	38.1035	142.861	20.3	<u>Fujii et al, (2011)</u>
				9	38.1035	142.861	20.3	<u>Satake et al, (2013)</u>

				9	38.1035	142.861	12.8		<u>Satake et al, (2013)</u>
				9	38.1035	142.861	12.8		<u>Satake et al, (2013)</u>
				9	38.107	142.916	20		<u>Yue and Lay (2013)</u>
2011090301	Offshore Honshu, Japan	09/03/2011	Near East Coast Of Honshu, Japan	7.3	38.4487	142.8574	20.7	Y	<u>Hayes (NEIC, Offshore Honshu 2011)</u>
2011180101	Pakistan	18/01/2011	Southwestern Pakistan	7.2	28.8453	63.9338	67.3	N	<u>Hayes (NEIC, Pakistan 2011)a</u>
				7.2	28.8423	63.9578	66.8		<u>Hayes (NEIC, Pakistan 2011)a</u>
2010251201	Vanuatu	25/12/2010	Vanuatu Islands Region	7.3	-19.7685	167.892	15.2	Y	<u>Hayes (NEIC, Vanuatu 2010)</u>
2010211201	Bonin Islands	21/12/2010	Bonin Islands, Japan Region	7.4	26.8537	143.6651	17.8	Y	<u>Hayes (NEIC, Bonin Islands 2010)</u>
2010251001	Sumatra	25/10/2010	Southern Sumatra, Indonesia	7.7	-3.5003	100.0793	17.4	Y	<u>Hayes (NEIC, Southern Sumatra 2010)</u>
2010030901	Darfield, South Island New Zealand	03/09/2010	South Island, New Zealand	7.02	-43.6148	172.0386	10.8	N	<u>Hayes (NEIC, Darfield 2010)</u>
				7.1	-43.55	172.2	10		<u>Atzori et al, (2012)</u>

2010090501	Northern Sumatra	09/05/2010	Northern Sumatra, Indonesia	7.2	3.7137	96.0833	44.6	N	<u>Hayes (NEIC, Northern Sumatra 2010)</u>
2010060401	Northern Sumatra	06/04/2010	Northern Sumatra, Indonesia	7.8	2.3384	97.1331	30.6	Y	<u>Hayes (USGS, Northern Sumatra 2010)</u>
2010040401	El Mayor-Cucapah, Mexico	04/04/2010	Calif,-Baja Calif, Border Region	7.29	32.3	-115.267	5.5	N	<u>Wei et al, (2011)</u>
				7.35	32.3	-115.3	10		<u>Mendoza and Hartzell (2013)</u>
2010270201	Maule, Chile	27/02/2010	Near Coast Of Central Chile	8.76	-35.846	-72.719	35	Y	<u>Sladen (Caltech, Maule 2010)</u>
				8.77	-35.8765	-72.7746	33.8		<u>Hayes (NEIC, Maule 2010)</u>
				8.8	-36	-73	20.3		<u>Luttrell et al, (2011)</u>
				8.9	-35.85	-72.72	37		<u>Shao et al, (UCSB, Maule 2010)</u>
				8.8	-36.208	-72.963	31.6		<u>Delouis at al, (2010)</u>
				8.8	-36.12	-72.9	22.9		<u>Lorito et al, (2011)</u>
				8.8	-36.12	-72.9	22.9		<u>Lorito et al, (2011)</u>
				8.8	-35.917	-72.886	20.1		<u>Pollitz et al, (2011)</u>

2010120101	Haiti	12/01/2010	Haiti Region	7.1	18.5076	-72.6375	12.9	Y	<u>Calais et al, (2010)</u>
				7	18.5	-72.53	11		<u>Sladen (Caltech, Haiti 2010)</u>
				7	18.445	-72.535	10.2		<u>Hayes et al, (2010)</u>
				7	18.5064	-72.5323	12.2		<u>Hayes (NEIC, Haiti 2010)</u>
2009071001	Vanuatu	07/10/2009	Vanuatu Islands	7.6	-13.0642	166.19	35	Y	<u>Sladen (Caltech, Vanuatu 2009)</u>
				7.8	-13.0675	166.1794	39.2		<u>Hayes (NEIC, Vanuatu 2009)</u>
2009300901	Padang, Indonesia	30/09/2009	Southern Sumatra, Indonesia	7.6	-0.789	99.961	80	Y	<u>Sladen (Caltech, Padang 2009)</u>
				7.6	-0.789	99.961	80		<u>Sladen (Caltech, Padang 2009)</u>
				7.6	-0.7982	99.8899	77.3		<u>Hayes (NEIC, Padang 2009)a</u>
2009290901	Samoa	29/09/2009	Samoa Islands Region	8	-15.6163	-172.3019	16.9	Y	<u>Hayes (NEIC, Samoa 2009)</u>
2009030801	Gulf of California	03/08/2009	Gulf Of California	6.9	29.4094	-112.8066	9.2	N	<u>Hayes (NEIC, Gulf of California 2009)</u>
2009150701	Fiordland, New Zealand	15/07/2009	Off W, Coast Of S, Island, N,Z,	7.6	-45.7288	166.6181	24.2	Y	<u>Hayes (NEIC, New Zealand 2009)</u>

2009280501	Offshore Honduras	28/05/2009	North Of Honduras	7.3	16.7786	-86.1808	13.5	Y	<u>Hayes and Ji (Offshore Honduras 2009)</u>
2009060401	L Aquila, Italy	06/04/2009	Central Italy	6.3	42.348	13.38	8.2	N	<u>Poiata et al, (2012)</u>
				6.3	42.348	13.38	8.2		<u>Poiata et al, (2012)</u>
				6.3	42.3439	13.3784	8.6		<u>Gualandi et al, (2013)</u>
				6.1	42.35	13.38	9.5		<u>Cirella et al, (2012)</u>
				6.1	42.35	13.38	9.5		<u>Cirella et al, (2012)</u>
				6.1	42.35	13.38	9.5		<u>Cirella et al, (2012)</u>
				6.3	42.35	13.38	9.5		<u>Cirella et al, (2009)a</u>
				6.18	42.339	13.381	8.8		<u>Gallovič et al, (2015)</u>
2009030101	Papua	03/01/2009	Irian Jaya Region, Indonesia	7.6	-0.5021	132.773	34.6	Y	<u>Hayes (NEIC, Papua 2009)</u>
2008161101	Sulawesi, Indonesia	16/11/2008	Minahassa Peninsula, Sulawesi	7.3	1.3001	122.103	25.5	Y	<u>Sladen (Caltech, Sulawesi 2008)</u>

2008290901	Kermadec Islands, New Zealand	29/09/2008	Kermadec Islands, New Zealand	7	-29.8937	-177.7883	39.5	N	<u>Hayes (NEIC, New Zealand 2008)</u>
2008250801	Zhongba, Tibet	25/08/2008	Xizang	6.7	30.7747	83.4825	7.6	N	<u>Elliott et al, (2010)</u>
2008130601	Iwate-Miyagi Nairiku, Japan	13/06/2008	Eastern Honshu, Japan	6.89	39.0298	140.8807	7.8	N	<u>Asano and Iwata (2011b)</u>
				6.8	39.1095	140.6775	9.5		<u>Hayes and Ji (Honshu 2008)</u>
				7	39.027	140.878	6.5		<u>Cultrera et al, (2013)</u>
2008120501	Wenchuan (China)	12/05/2008	Sichuan, China	7.92	30.858	103.24	10.8	Y	<u>Fielding et al, (2013)</u>
				7.9	30.858	103.24	10.8		<u>Fielding et al, (2013)</u>
				7.92	30.858	103.24	10.8		<u>Fielding et al, (2013)</u>
				8.03	30.986	103.364	16		<u>Yagi et al, (2012)</u>
				7.9	30.9925	103.4579	9.3		<u>Sladen (Caltech, Wenchuan 2008)</u>
				7.9	31.1024	103.3166	9.5		<u>Ji and Hayes (Wenchuan 2008)</u>
				7.9	30.87	103.2	13.2		<u>Qi et al, (2011)</u>
2008200301	Yutian, Tibet	20/03/2008	Southern Xinjiang, China	7.1	35.4824	81.5146	4.1	N	<u>Elliott et al, (2010)</u>

2008200201	Simeulue, Indonesia	20/02/2008	Off W Coast Of Northern Sumatra	7.4	2.7313	95.9707	24.8	N	<u>Sladen (Caltech, Simeulue 2008)</u>
				7.4	2.7074	95.9661	23.6		<u>Hayes and Ji (Simeulue 2008)</u>
2008160101	Gerze, Tibet	16/01/2008	Xizang	5.9	32.45	85.3	4	N	<u>Elliott et al, (2010)</u>
2008090101	Gerze, Tibet	09/01/2008	Xizang	6.4	32.41	81.35	7.5	N	<u>Elliott et al, (2010)</u>
2007141101	Tocopilla, Chile	14/11/2007	Northern Chile	7.7	-22.1972	-69.853	49.7	Y	<u>Sladen (Caltech, Tocopilla 2007)</u>
				7.7	-22.2194	-69.8169	39.4		<u>Zeng et al, (2007; USGS, Online Model)</u>
				7.81	-22.2271	-69.8972	37		<u>Ji (UCSB, Tocopilla 2007)</u>
				7.8	-22.33	-70.11	47.9		<u>Motagh et al, (2010)</u>
				7.7	-22.34	-70.06	51.8		<u>Bejar-Pizarro et al, (2010)</u>
				7.7	-22.34	-70.06	51.8		<u>Bejar-Pizarro et al, (2010)</u>
				7.7	-22.34	-70.06	51.8		<u>Bejar-Pizarro et al, (2010)</u>
2007120901	Benkulu, Indonesia	12/09/2007	Southern Sumatra,	8.4	-4.5321	101.3918	30	Y	<u>Konca (Caltech, Benkulu 2007)</u>

			Indonesia	8.4	-4.5453	101.378	29.4		<u>Konca et al, (2008)</u>
				8.4	-4.5469	101.3795	29.6		<u>Ji (UCSB, Benkulu 2007)</u>
				8.5	-4.52	101.374	21.2		<u>Gusman et al, (2010)</u>
				8.5	-4.52	101.374	21.2		<u>Gusman et al, (2010)</u>
2007120902	Pagai, Indonesia	12/09/2007	Southern Sumatra, Indonesia	7.9	-2.5151	100.9139	45		<u>Sladen and Konca (Caltech, Pagai 2007)</u>
				7.9	-2.7637	100.4982	34.7		<u>Konca et al, (2008)</u>
				7.9	-2.7156	100.8011	44.6		<u>Ji and Zeng (Pagai 2007))</u>
2007170701	Niigata-ken Chuetsu-oki	17/08/2007	Near West Coast Of Honshu, Japan	6.6	37.54	138.61	8.9	N	<u>Cirella et al, (2008)</u>
2007150801	Pisco, Peru	15/08/2007	Near Coast Of Peru	8	-13.0483	-76.5427	39	Y	<u>Konca (Caltech, Peru 2007)</u>
				8	-13.3247	-76.5154	38.6		<u>Sladen (Caltech, Peru 2007)</u>
				8	-13.3758	-76.5213	29.4		<u>Ji and Zeng (Peru 2007)</u>
2007010401	Solomon Islands	01/04/2007	Solomon Islands	8.1	-8.4878	156.9636	11.6	Y	<u>Ji (UCSB, Solomon Islands 2007)</u>

2007250301	Noto Hanto, Japan	25/03/2007	Near West Coast Of Honshu, Japan	6.73	37.217	136.7019	9.6	Y	<u>Asano and Iwata (2011)</u>
2007130101	Kuril Islands	13/01/2007	East Of Kuril Islands	8.1	46.2542	154.432	10	Y	<u>Sladen (Caltech, Kuril 2007)</u>
				8.1	46.2856	154.4348	18.2		<u>Ji (UCSB, Kuril 2007)</u>
2006261202	Pingtung evt2 (Taiwan)	26/12/2006	Taiwan Region	6.8	21.97	120.42	50.2	Y	<u>Yen et al, (2008)</u>
2006261201	Pingtung evt1 (Taiwan)	26/12/2006	Taiwan Region	6.9	21.69	120.56	44.1	Y	<u>Yen et al, (2008)</u>
2006151101	Kuril Islands	15/11/2006	Kuril Islands	8.3	46.6216	153.2723	25.9	Y	<u>Ji (UCSB, Kuril 2006)</u>
				8.3	46.607	153.23	20		<u>Sladen (Caltech, Kuril 2006)</u>
				8.4	46.592	153.266	12		<u>Lay et al, (2009)</u>
2006170701	Java, Indonesia	17/07/2006	South Of Java, Indonesia	7.82	-9.31	107.284	15	Y	<u>Yagi and Fukahata (2011)</u>
				7.9	-9.295	107.345	9.8		<u>Konca (Caltech, Java 2006)</u>
				7.7	-9.308	107.3598	14.7		<u>Ji (UCSB, JAVA 2006)</u>
2005081001	Kashmir, Pakistan	08/10/2005	Pakistan	7.6	34.4905	73.6253	10.9	N	<u>Konca (Caltech, Kashmir 2005)</u>
				7.6	34.4813	73.6259	10.5		<u>Shao and Ji (UCSB, Kashmir 2005)</u>

2005160801	Honshu , Japan	16/08/2005	Near East Coast Of Honshu, Japan	7.5	38.2398	142.0458	34.5	Y	<u>Shao and Ji (UCSB, Honshu 2005)</u>
2005150601	Northern Californi	15/06/2005	Off Coast Of Northern California	7.2	41.2932	-125.9812	9	Y	<u>Shao and Ji (UCSB, Northern California 2005)</u>
2005200401	Fukuoka (Japan)	20/04/2005	Kyushu, Japan	5.57	33.67	130.29	13	N	<u>Asano and Iwata (2006)</u>
2005070401	Zhongba, Tibet	07/04/2005	Xizang	6.2	29.76	90.43	6	N	<u>Elliott et al, (2010)</u>
2005280301	Sumatra	28/03/2005	Northern Sumatra, Indonesia	8.7	2.01	97.01	40	Y	<u>Ji (Caltech, Sumatra 2005)</u>
				8.5	2.0455	97.0082	29.6		<u>Konca et al, (2007)</u>
				8.68	2.0623	97.1048	25.7		<u>Shao and Ji (UCSB, Sumatra 2005)</u>
2005200301	Fukuoka (Japan)	20/03/2005	Kyushu, Japan	6.64	33.75	130.16	14	N	<u>Asano and Iwata (2006)</u>
2004261201	Sumatra (Indonesia)	26/12/2004	Off west coast of northern Sumatera, Indonesia	9.1	3.3	95.78	35	N	<u>Ammon et al, (2005)</u>
				9.17	3.12	95.49	27		<u>Rhie et al, (2007)</u>
				8.89	3.3	95.78	35		<u>Ji et al, (2005)</u>
				9.19	3.12	95.49	27		<u>Rhie et al, (2007)</u>
2004231001	Niigata-Ken Chuetsu, Japan	23/10/2004	Near West Coast Of Honshu, Japan	6.62	37.307	138.839	10.6	N	<u>Asano and Iwata (2009)</u>

2004280901	Parkfield (Calif.)	28/09/2004	Central California, United States	6	35.8185	-120.3706	8.3	N	<u>Dreger et al, (2005)</u>
				6.06	35.8185	-120.3706	8.3		<u>Custodio et al, (2005)</u>
				5.9	35.8236	-120.367	8		<u>Ji (Caltech, Parkfield 2004)</u>
2004110701	Zhongba, Tibet	11/07/2004	Xizang	6.2	30.73	83.83	10	N	<u>Elliott et al, (2010)</u>
2004070201	Irian-Jaya, Indonesia	07/02/2004	Irian Jaya Region, Indonesia	7.2	-3.9906	135.051	11.2	N	<u>Wei (Caltech, Irian-Jaya 2004)</u>
2003261201	Bam, Iran	26/12/2003	Southern Iran	6.5	29.052	58.365	8	N	<u>Poiata et al, (2012a)</u>
2003250901	Tokachi-oki (Japan)	25/09/2003	Hokkaido, Japan, region	8.03	41.78	144.08	25	Y	<u>Yamanaka and Kikuchi (2003)</u>
				8.16	41.78	144.08	27		<u>Yagi (2004)</u>
				7.96	41.78	144.08	25		<u>Tanioka et al, (2004)</u>
				8.21	41.7797	144.0785	25		<u>Koketsu et al, (2004)</u>
2003250701	Miyagi-hokubu (Japan)	25/07/2003	Near east coast of eastern Honshu, Japan	6.08	38.403	141.174	6.5	N	<u>Miura et al, (2004)</u>
				6.1	38.403	141.174	6.5		<u>Hikima and Koketsu (2004)</u>
2003150701	Carlsberg Ridge	15/07/2003	Carlsberg Ridge	7.6	-2.6116	68.3704	11.3	N	<u>Wei (Caltech, Carlsberg 2003)</u>

2003210501	Boumerdes (Algeria)	21/05/2003	Northern Algeria	7.25	36.83	3.65	16	Y	<u>Semmane et al, (2005)</u>
2003220101	Colima (Mexico)	22/01/2003	Near coast of Jalisco, Mexico	7.5	18.71	-104.13	20	Y	<u>Yagi et al, (2004)</u>
2002031101	Denali (Alaska)	03/11/2002	Central Alaska, United States	7.91	63.52	-147.53	7.5	N	<u>Oglesby et al, (2004)</u>
				7.87	63.5141	-147.4529	6		<u>Asano et al, (2005)</u>
				7.97	63.5175	-147.444	8		<u>Hayes (NEIC, Alaska Central 2015)</u>
2001141101	Qinghai, China	14/11/2001	Qinghai, China	7.84	35.95	90.54	10	N	<u>Hayes (NEIC, Qinghai 2015)</u>
2001240301	Geiyo (Japan)	24/03/2001	Western Honshu, Japan	6.79	34.1292	132.6962	46.5	N	<u>Sekiguchi and Iwata (2001)</u>
				6.68	34.1292	132.6962	46.5		<u>Takehi (2004)</u>
2001260101	Bhuj (India)	26/01/2001	Southern India	7.66	23.63	70.24	18	N	<u>Yagi (2003)</u>
				7.38	23.4	70.32	20		<u>Antolik and Dreger (2003)</u>
				7.6	23.4035	70.2835	22.7		<u>Copley et al, (2011)</u>
2000061001	Tottori (Japan)	06/10/2000	Western Honshu, Japan	6.73	35.269	133.357	14.5	N	<u>Semmane et al, (2005)a</u>
				6.83	35.2694	133.3567	7.8		<u>Sekiguchi et al, (2003)</u>
				6.86	35.269	133.357	14		<u>Iwata et al, (2000)</u>

				6.6	35.275	133.35	12.5		<u>Piatanesi et al, (2007)</u>
2000170601	Kleifarvatn (Iceland)	17/06/2000	Iceland region	5.87	63.9038	-21.9664	4	N	<u>Sudhaus and Jonsson (2008)</u>
1999121101	Duzce (Turkey)	12/11/1999	Turkey	7.18	40.818	31.198	18	N	<u>Delouis et al, (2004)</u>
				6.71	40.76	31.16	10		<u>Birgoren et al, (2004)</u>
1999161001	Hector Mine (Calif.)	16/10/1999	Southern California, United States	7.14	34.59	-116.27	7.5	N	<u>Salichon et al, (2004)</u>
				7.24	34.59	-116.27	6		<u>Kaverina et al, (2002)</u>
				7.16	34.59	-116.27	15		<u>Jonsson et al, (2002)</u>
				7.17	34.597	-116.27	15		<u>Ji et al, (2002)</u>
1999300901	Oaxaca (Mexico)	30/09/1999	Southeastcentral Pacific Ocean	7.47	16	-97.02	39.7	N	<u>Hernandez et al, (2001)</u>
1999200901	ChiChi (Taiwan)	20/09/1999	Taiwan	7.61	23.869	120.84	6.8	N	<u>Zeng and Chen (2001)</u>
				7.63	23.869	120.84	7		<u>Sekiguchi et al, (2002)</u>
				7.67	23.869	120.84	6.9		<u>Wu et al, (2001)</u>
				7.58	23.869	120.84	8.8		<u>Johnsson et al, (2002)</u>
				7.68	23.869	120.84	8		<u>Chi et al, (2001)</u>
				7.69	23.869	120.84	7		<u>Ma et al, (2000)</u>

				7.69	23.869	120.84	7		<u>Ma et al, (2001)</u>
1999170801	Izmit (Turkey)	17/08/1999	Turkey	7.4	40.7	29.91	16	Y	<u>Yagi and Kikuchi (2000)</u>
				7.44	40.71	29.91	17		<u>Sekiguchi and Iwata (2002)</u>
				7.42	40.76	29.97	17		<u>Reilinger et al, (2000)</u>
				7.56	40.76	29.97	17		<u>Delouis et al, (2002)</u>
				7.47	40.7	29.91	16		<u>Cakir et al, (2004)</u>
				7.59	40.725	29.987	17		<u>Bouchon et al, (2002)</u>
1998180901	Hida Swarm EV16 (Japan)	18/09/1998	Eastern Honshu, Japan	4.41	36.3337	137.6577	1.9	N	<u>Ide (2001)</u>
1998030901	Iwate (Japan)	03/09/1998	Eastern Honshu, Japan	6.3	39.8132	140.8966	3	N	<u>Nakahara et al, (2002)</u>
				6.27	39.8132	140.8966	7.5		<u>Miyakoshi et al, (2000)</u>
1998250801	Pumqu-Xainza, Tibet	25/08/1998	Xizang	6.16	30.1345	88.31	8.3	N	<u>Wang et al, (2014)</u>
1998220801	Hida Swarm EV11 (Japan)	22/08/1998	Eastern Honshu, Japan	4.33	36.2396	137.6529	2.6	N	<u>Ide (2001)</u>
1998170801	Hida Swarm EV10 (Japan)	17/08/1998	Eastern Honshu, Japan	4.45	36.3512	137.6228	3	N	<u>Ide (2001)</u>
1998160802	Hida Swarm EV09 (Japan)	16/08/1998	Eastern Honshu, Japan	5.13	36.3222	137.6327	3.4	N	<u>Ide (2001)</u>

1998160801	Hida Swarm EV08 (Japan)	16/08/1998	Eastern Honshu, Japan	4.11	36.318	137.641	2.6	N	<u>Ide (2001)</u>
1998140801	Hida Swarm EV07 (Japan)	14/08/1998	Eastern Honshu, Japan	4.41	36.3017	137.6307	3.7	N	<u>Ide (2001)</u>
1998120801	Hida Swarm EV05 (Japan)	12/08/1998	Eastern Honshu, Japan	4.54	36.2351	137.6256	3.3	N	<u>Ide (2001)</u>
1998170701	Rueyli (Taiwan)	17/07/1998	Taiwan	5.84	23.5	120.66	6	N	<u>Ma and Wu (2001)</u>
1998250302	Antarctica Normal-Faulting Segment	25/03/1998	Balleny Islands region	7.76	-62.86	149.74	12	Y	<u>Antolik et al, (2000)</u>
1998250301	Antarctica Strike-Slip Segment	25/03/1998	Balleny Islands region	7.98	-62.86	149.74	12	Y	<u>Antolik et al, (2000)</u>
1997141001	Colfiorito Sequence # 3 (Italy)	14/10/1997	Central Italy	5.86	42.919	12.926	5.6	N	<u>Hernandez et al, (2004)</u>
1997260901	Colfiorito Sequence # 1 (Italy)	26/09/1997	Central Italy	5.72	43.0225	12.8917	4	N	<u>Hernandez et al, (2004)</u>
1997260902	Colfiorito Sequence # 2 (Italy)	26/09/1997	Central Italy	5.97	43.0305	12.8622	5.1	N	<u>Hernandez et al, (2004)</u>
1997250601	Yamaguchi (Japan)	25/06/1997	Western Honshu, Japan	5.82	34.43	131.676	8.2	N	<u>Miyakoshi et al, (2000)</u>
				5.81	34.441	131.676	7.5		<u>Ide (1999)</u>
1997130501	Kagoshimaen-hoku-seibu (Japan)	13/05/1997	Kyushu, Japan	6.01	31.952	130.34	7.7	N	<u>Horikawa (2001)</u>
1997100501	Zirkuh (East Iran)	10/05/1997	Northern And Central Iran	7.2	33.82	59.8	10	N	<u>Sudhaus and Jonsson (2011)</u>
1997260301	Kagoshimaen-	26/03/1997	Kyushu, Japan	6.1	31.9681	130.3875	7.6	N	<u>Horikawa (2001)</u>

	hoku-seibu (Japan)			6.04	31.9681	130.3875	7.6		<u>Miyakoshi et al, (2000)</u>
1996021201	Hyuga-nada2 (Japan)	02/12/1996	Kyushu, Japan	6.68	31.803	131.588	20.4	Y	<u>Yagi et al, (1999)</u>
1996121101	Nazca Ridge (Peru)	12/11/1996	Near coast of Peru	8.06	-14.99	-75.63	21	Y	<u>Spence et al, (1999)</u>
				7.84	-14.99	-75.63	21		<u>Salichon et al, (2003)</u>
1996191001	Hyuga-nada1 (Japan)	19/10/1996	Southeast of Shikoku, Japan	6.81	31.835	131.958	11.6	N	<u>Yagi et al, (1999)</u>
1996030701	Pumqu-Xainza, Tibet	03/07/1996	Xizang	6.08	30.1456	88.2894	8.3	N	<u>Wang et al, (2014)</u>
1995311001	Taiwan (9510312227P000)	31/10/1995	Taiwan	4.93	23.291	120.359	10.7	N	<u>Ma and Wu (2001)</u>
				4.93	23.291	120.359	10.7		<u>Ma and Wu (2001)</u>
1995091001	Colima (Mexico)	09/10/1995	Southeastcentral Pacific Ocean	7.96	18.864	-104.579	16.6	Y	<u>Mendoza and Hartzell (1999)</u>
1995140901	Copala, Mexico	14/09/1995	Near Coast Of Guerrero, Mexico	7.3	16.48	-98.76	15.9	Y	<u>Courboulex et al, (1997)</u>
1995140701	Taiwan (9507141652)	14/07/1995	Taiwan	5.07	24.368	121.743	9.8	N	<u>Ma and Wu (2001)</u>
1995170101	Kobe (Japan)	17/01/1995	Near south coast of western Honshu, Japan	6.9	34.595	135.012	14.3		<u>Zeng and Anderson (2000)</u>
				6.86	34.595	135.012	14.3		<u>Yoshida et al, (1996)</u>

				6.92	34.595	135.012	14.3		<u>Wald (1996)</u>
				6.87	34.595	135.012	14.3		<u>Koketsu et al, (1998)</u>
				6.89	34.595	135.012	14.3		<u>Ide et al, (1996)</u>
				7.01	34.607	135.043	14		<u>Horikawa et al, (1996)</u>
				6.8	34.6	135.04	14		<u>Cho and Nakanishi (2000)</u>
				6.99	34.598	135.044	16.4		<u>Sekiguchi et al, (2000)</u>
				7.02	34.598	135.044	16.4		<u>Sekiguchi et al, (2002)a</u>
1994281201	Sanrikuki (Japan)	28/12/1994	Off east coast of Honshu, Japan	7.7	40.43	143.75	10	Y	<u>Nagai et al, (2001)</u>
				7.88	40.43	143.75	10		<u>Nakayama and Takeo (1997)</u>
1994050601	Nanao (Taiwan)	05/06/1994	Taiwan	6.28	24.468	121.787	5.1	N	<u>Ma and Wu (2001)</u>
1994060401	Taiwan (9404060112P160)	06/04/1994	Taiwan	4.54	23.533	120.421	13.4	N	<u>Ma and Wu (2001)</u>
				4.54	23.533	120.421	13.4		<u>Ma and Wu (2001)</u>
1994170101	Northridge (Calif,)	17/01/1994	Southern California, United States	6.66	34.213	-118.537	17.5	Y	<u>Dreger (1994)</u>
				6.73	34.213	-118.537	17.5		<u>Hartzell et al, (1996)</u>

				6.81	34.213	-118.537	17.5		<u>Hudnut et al, (1996)</u>
				6.84	34.213	-118.537	17.5		<u>Shen et al, (1996)</u>
				6.8	34.213	-118.537	17.5		<u>Wald et al, (1996)</u>
				6.71	34.213	-118.537	17.5		<u>Zeng and Anderson (2000)</u>
1993151201	Tapu (Taiwan)	15/12/1993	Taiwan	5.65	23.194	120.507	15.2	N	<u>Ma and Wu (2001)</u>
1993120701	Hokkaido-nansei-oki (Japan)	12/07/1993	Hokkaido, Japan, region	7.76	42.777	139.197	17	Y	<u>Tanioka et al, (1995)</u>
				7.6	42.78	139.183	20		<u>Mendoza and Fukuyama (1996)</u>
1993200301	Pumqu-Xainza, Tibet	20/03/1993	Xizang	6.29	29.0464	87.4905	8.3	N	<u>Wang et al, (2014)</u>
1992290601	Little Skull Mnts (Nevada)	29/06/1992	California-Nevada border region	5.63	36.4316	-116.1776	11.8	N	<u>Silva et al, (2003)</u>
1992280601	Landers (Calif,)	28/06/1992	Southern California, United States	7.2	34.2	-116.43	7	N	<u>Zeng and Anderson (2000)</u>
				7.28	34.2	-116.43	7		<u>Wald and Heaton (1994)</u>
				7.22	34.2	-116.43	7		<u>Hernandez et al, (1999)</u>
				7.29	34.2	-116.43	7		<u>Cotton and Campillo (1995)</u>

				7.08	34.2	-116.43	5		<u>Cohee and Beroza (1994)</u>
1992230401	Joshua Tree (Calif.)	23/04/1992	Southern California, United States	6.15	33.966	-116.317	12.5	N	<u>Hough and Dreger (1995)</u>
				6.25	33.966	-116.317	12.5		<u>Bennet et al, (1995)</u>
1991280601	Sierra Madre (Calif.)	28/06/1991	Southern California, United States	5.59	34.264	-117.99	20.9	N	<u>Wald (1992)</u>
1989251201	Ungava (Canada)	25/12/1989	Northern Quebec, Canada	6.02	60.12	-73.6	2.5	N	<u>Hartzell et al, (1994)</u>
1989181001	Loma Prieta (Calif.)	18/10/1989	Central California, United States	6.98	37.041	-121.883	17.6	Y	<u>Zeng and Anderson (2000)</u>
				6.94	37.041	-121.883	17.6		<u>Wald et al, (1991)</u>
				6.99	37.041	-121.883	17.6		<u>Steidl et al, (1991)</u>
				6.91	37.041	-121.883	17.6		<u>Emolo and Zollo (2005)</u>
				6.95	37.041	-121.883	17.6		<u>Beroza (1991)</u>
1988251101	Saguenay (Canada)	25/11/1988	Southern Quebec, Canada	5.81	48.117	-71.184	26	N	<u>Hartzell et al, (1994)</u>
1987241101	Superstition Hills (Calif.)	24/11/1987	Southern California, United States	6.51	33.0166	-115.85	10	N	<u>Wald et al, (1990)</u>
				6.6	33.0166	-115.85	10		<u>Larsen et al, (1992)</u>

1987241102	Elmore Ranch (Calif.)	24/11/1987	Southern California, United States	6.52	33.0833	-115.796	10	N	<u>Larsen et al, (1992)</u>
1987011001	Whittier Narrows (Calif.)	01/10/1987	Southern California, United States	5.89	34.05	-118.08	14.6	N	<u>Hartzell and Iida (1990)</u>
1986080701	North Palm Springs (Calif.)	08/07/1986	Southern California, United States	6.14	34	-116.57	11	N	<u>Mendoza and Hartzell (1988)</u>
				6.21	34	-116.57	11		<u>Hartzell (1989)</u>
1985231201	Nahanni2 (Canada)	23/12/1985	Northwest Territories, Canada	6.66	62.187	-124.24	8	N	<u>Hartzell et al, (1994)</u>
1985051001	Nahanni1 (Canada)	05/10/1985	Northwest Territories, Canada	6.66	62.208	-124.217	8	N	<u>Hartzell et al, (1994)</u>
1985210901	Zihuatanejo (Mexico)	21/09/1985	Near coast of Guerrero, Mexico	7.42	17.598	-101.819	20	Y	<u>Mendoza (1993)</u>
1985190901	Michoacan (Mexico)	19/09/1985	Michoacan, Mexico	8.01	18.18	-102.57	17	Y	<u>Mendoza and Hartzell (1989)</u>
1985030301	Central Chile (Chile)	03/03/1985	Near coast of central Chile	8.16	-33.125	-71.61	40	Y	<u>Mendoza et al, (1994)</u>
1984140901	Nagano-ken seibu (Japan)	14/09/1984	Eastern Honshu, Japan	6.29	35.82	137.49	3	N	<u>Takeo and Mikami (1987)</u>
1984240401	Morgan Hill (Calif.)	24/04/1984	Central California, United States	6.28	37.317	-121.682	8.5	N	<u>Beroza and Spudich (1988)</u>
				6.07	37.317	-121.682	8.5		<u>Hartzell and Heaton (1986)</u>

1983281001	Borah Peak (Idaho)	28/10/1983	Eastern Idaho, United States	6.82	44.06	-113.86	16	N	<u>Mendoza and Hartzell (1988)</u>
1983260501	Japanea-chubu (Japan)	26/05/1983	Near West Coast Of Honshu, Japan	7.58	40.357	139.077	8.1	Y	<u>Fukuyama and Irikura (1986)</u>
1982090101	New Brunswick (Canada)	09/01/1982	New Brunswick, Canada	5.47	46.984	-66.656	8	N	<u>Hartzell et al, (1994)</u>
1981251001	Playa Azul (Mexico)	25/10/1981	Near coast of Michoacan, Mexico	7.25	17.742	-102.238	15	Y	<u>Mendoza (1993)</u>
1980290601	Izu-hanto-toho-oki (Japan)	29/06/1980	Near south coast of eastern Honshu, Japan	6.61	34.9231	139.1855	8	Y	<u>Takeo (1988)</u>
1979151001	Imperial Valley (Calif,)	15/10/1979	California-Baja California border region	6.35	32.644	-115.309	8	N	<u>Zeng and Anderson (2000)</u>
				6.53	32.644	-115.309	8		<u>Olson and Apsel (1982)</u>
				6.58	32.644	-115.309	8		<u>Hartzell and Heaton (1983)</u>
				6.53	32.644	-115.309	8		<u>Archuleta (1984)</u>
1979060801	Coyote Lake (Calif,)	06/08/1979	Central California, United States	5.92	36.97	-121.51	8	N	<u>Liu and Helmberger (1983)</u>
1979140301	Petatlan (Mexico)	14/03/1979	Near coast of Guerrero, Mexico	7.39	17.46	-101.46	15	Y	<u>Mendoza (1995)</u>
1978160901	Tabas (Iran)	16/09/1978	Northern and central Iran	7.09	33.22	57.36	5.5	N	<u>Hartzell and Mendoza (1991)</u>

1978120601	Miyagiki (Japan)	12/06/1978	Near east coast of eastern Honshu, Japan	7.61	38.15	142.167	37	Y	<u>Yamanaka and Kikuchi (2004)</u>
1975210401	Oita-ken-chubu (Japan)	21/04/1975	Kyushu, Japan	6.36	33.15	131.3	8.5	N	<u>Takeo (1990)</u>
1974031001	Peru (Central Coast)	03/10/1974	Near coast of Peru	8.01	-12.39	-77.76	12	Y	<u>Hartzell and Langer (1993)</u>
1974090501	Izu-hanto-oki (Japan)	09/05/1974	Near south coast of eastern Honshu, Japan	6.56	34.4	138.8	8	Y	<u>Takeo (1990)</u>
1971090201	San Fernando (Calif,)	09/02/1971	Southern California, United States	6.82	34.43	-118.37	13	N	<u>Heaton (1982)</u>
1969090901	Gifu-ken-chubu (Japan)	09/09/1969	Eastern Honshu, Japan	6.43	35.47	137.05	2	N	<u>Takeo (1990)</u>
1968160501	Tokachi-oki (Japan)	16/05/1968	Off east coast of Honshu, Japan	8.35	40.73	143.58	9	Y	<u>Nagai et al, (2001)</u>
1968010401	Hyuga-nada (Japan)	01/04/1968	Shikoku, Japan	7.53	32.28	132.53	15	Y	<u>Yagi et al, (1998)</u>
1961190801	Kitamino (Japan)	19/08/1961	Near west coast of eastern Honshu, Japan	6.47	36.2	136.8	10	N	<u>Takeo (1990)</u>
1948280601	Fukui (Japan)	28/06/1948	Near west coast of eastern Honshu, Japan	6.65	36.247	136.188	10.1	N	<u>Ichinose et al, (2005)</u>
1946201201	Nankaido (Japan)	20/12/1946	Near south coast of western Honshu, Japan	8.4	33.03	135.62	30	Y	<u>Tanioka and Satake (2001)</u>
				8.31	33.03	135.62	30		<u>Satake (1993)</u>
				8.3	33.03	135.62	30		<u>Kato and Ando (1997)</u>

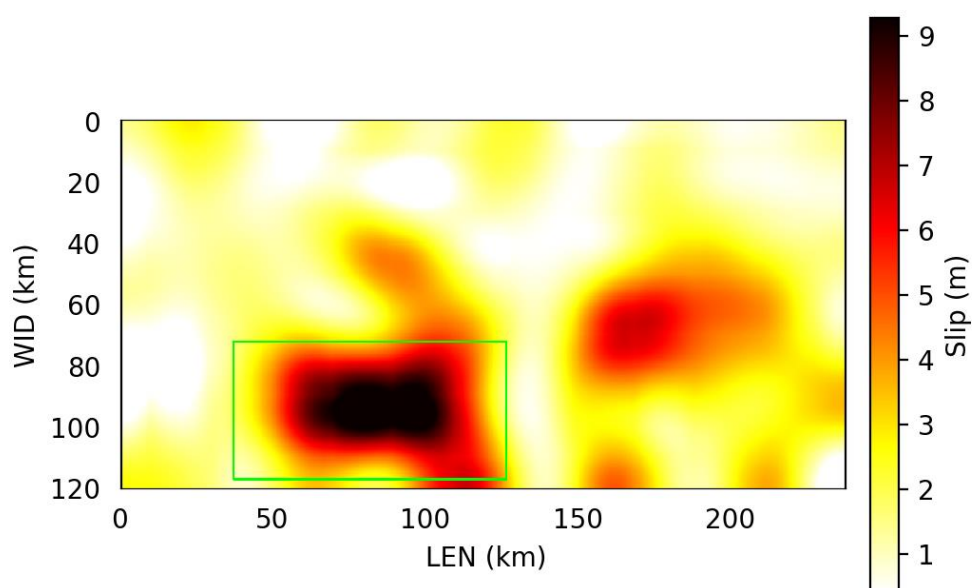
				8.4	33.03	135.62	30		<u>Baba et al, (2002)</u>
1945120101	Mikawa (Japan)	12/01/1945	Near south coast of eastern Honshu, Japan	6.67	34.7	137.1	10	Y	<u>Kikuchi et al, (2003)</u>
1944071201	Tonankai (Japan)	07/12/1944	Near south coast of western Honshu, Japan	8.1	33.8	136.62	30	Y	<u>Tanioka and Satake (2001)a</u>
				8.1	33.8	136.62	30		<u>Satake (1993)</u>
				7.99	33.7	136.05	30		<u>Kikuchi et al, (2003)</u>
				8.2	33.8	136.62	30		<u>Kato and Ando (1997)</u>
				8.04	33.77	135.96	30		<u>Ichinose et al, (2003)</u>
1923010901	Kanto (Japan)	01/09/1923	Near south coast of eastern Honshu, Japan	7.95	35.4	139.2	14.6	Y	<u>Wald and Somerville (1995)</u>
				8.08	35.4	139.2	14.6		<u>Kobayashi and Koketsu (2005)</u>
1906180401	San Francisco (Calif.)	18/04/1906	Central California, United States	7.64	37.78	-122.51	10	Y	<u>Wald et al, (1993)</u>
				7.91	37.78	-122.51	10		<u>Song et al, (2008)</u>
				7.91	37.78	-122.51	10		<u>Thatcher et al, (1997)</u>

APPENDICE B

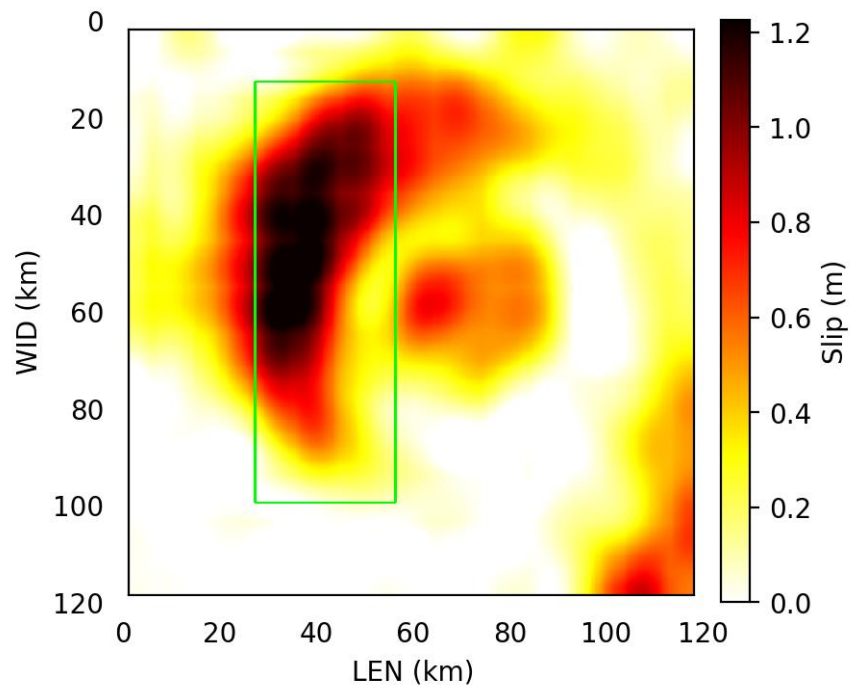
Per ognuno dei terremoti appartenenti al sottoinsieme di SRCMOD che abbiamo inizialmente riconosciuto come caratterizzati da una singola asperità, riportiamo di seguito il risultato del riconoscimento migliore sulla base dell'indice I_1 (rapporto fra lo slip contenuto nel rettangolo di riconoscimento e lo slip totale sulla faglia).

I risultati sono raggruppati in quattro classi distinte di I_1 : $I_1 \in [0,0.4[$, $I_1 \in [0.4,0.6[$, $I_1 \in [0.6,0.8[$, $I_1 \in [0.8,1.0]$.

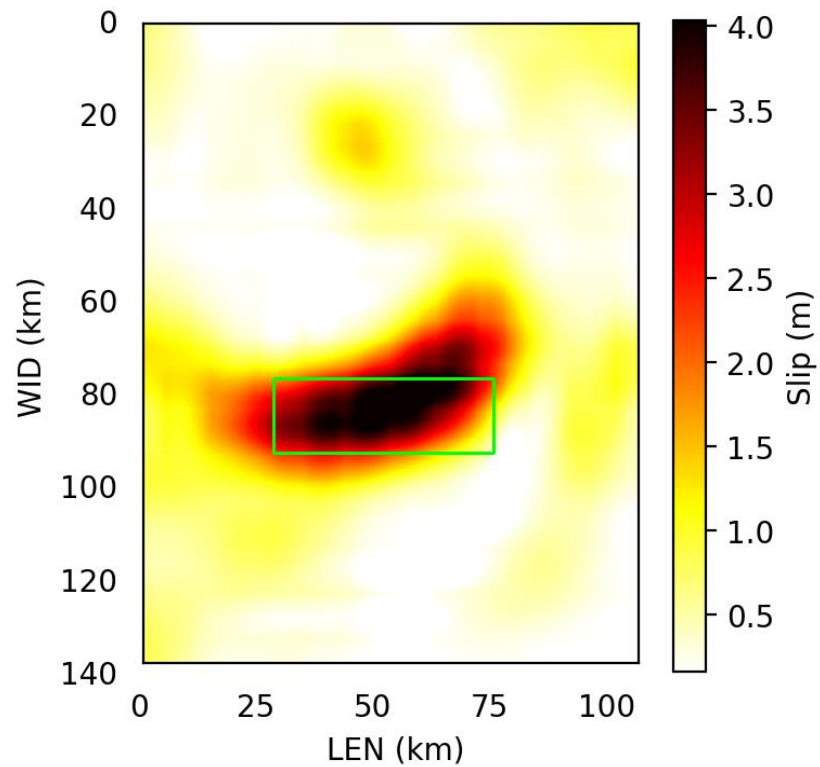
Intervallo $I_1 \in [0, 0.4[$



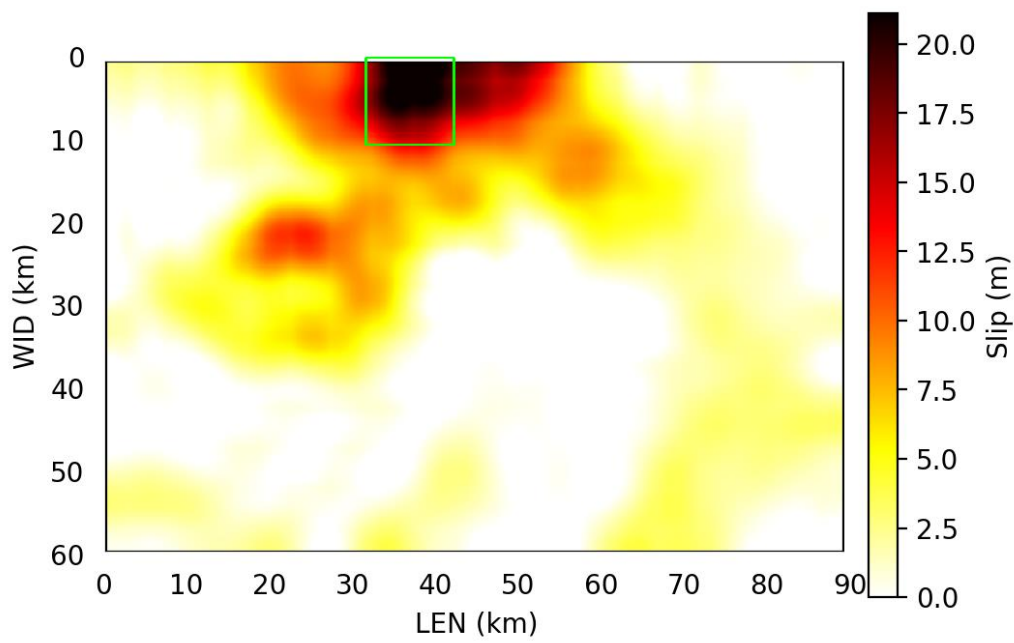
s1968TOKACH01NAG. Evento del 16/05/1968, Off east coast of Honshu, Japan, magnitudo 8.35



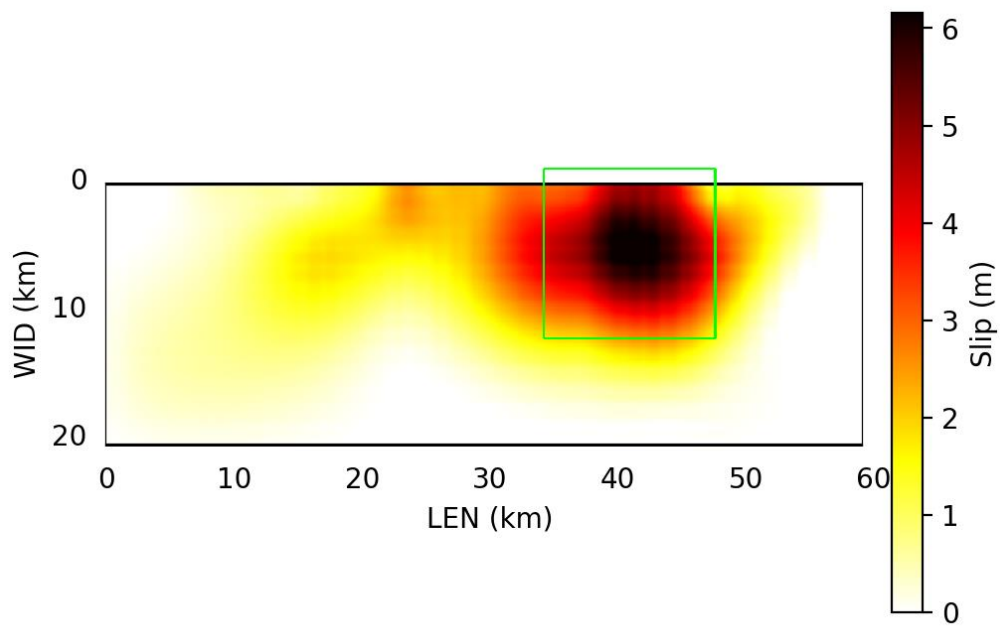
s1979PETATLO1MEND. Evento del 14/03/1979, Near coast of Guerrero, Mexico, magnitudo 7.39.



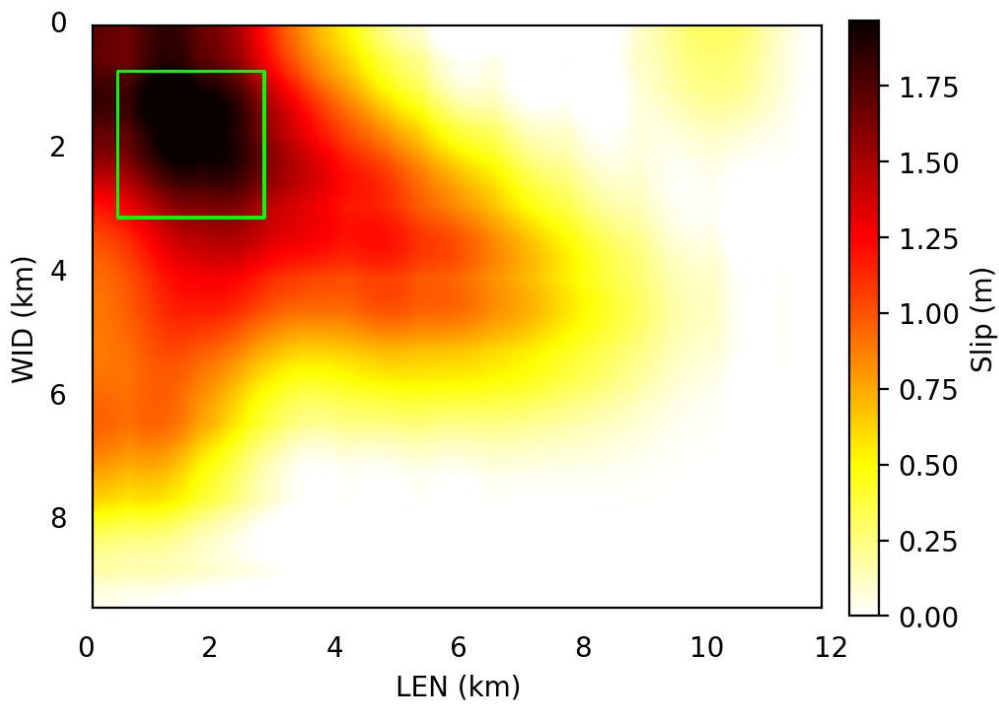
s1994SANRIK01NAGA. Evento del 28/12/1994, Off east coast of Honshu, Japan, magnitudo 7.70.



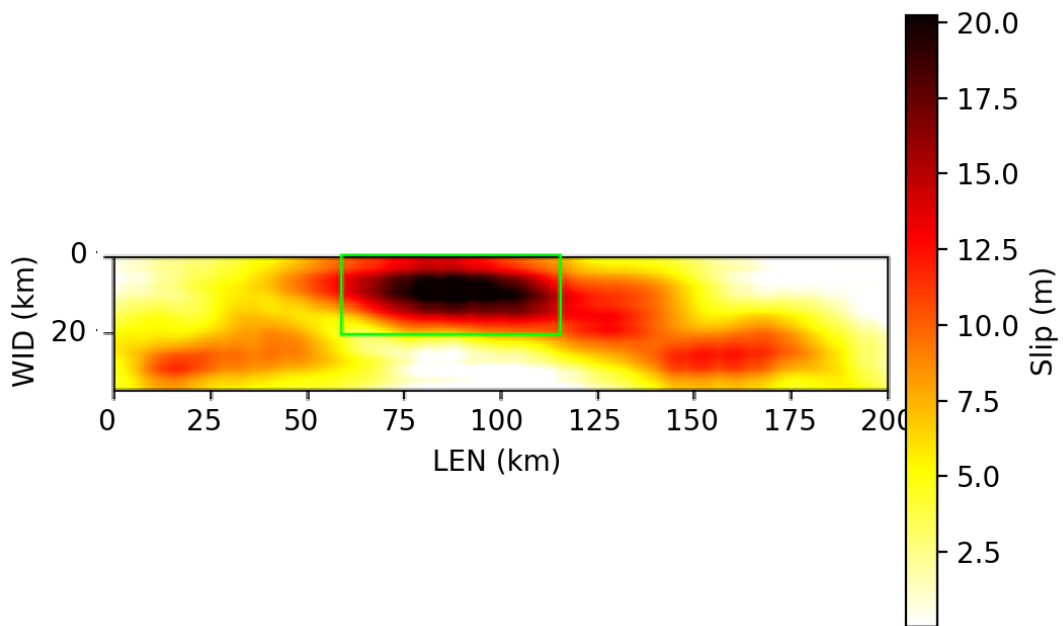
s1998ANTARC02ANTO. Evento del 25/03/1998, Balleny Islands region, magnitudo 7.76.



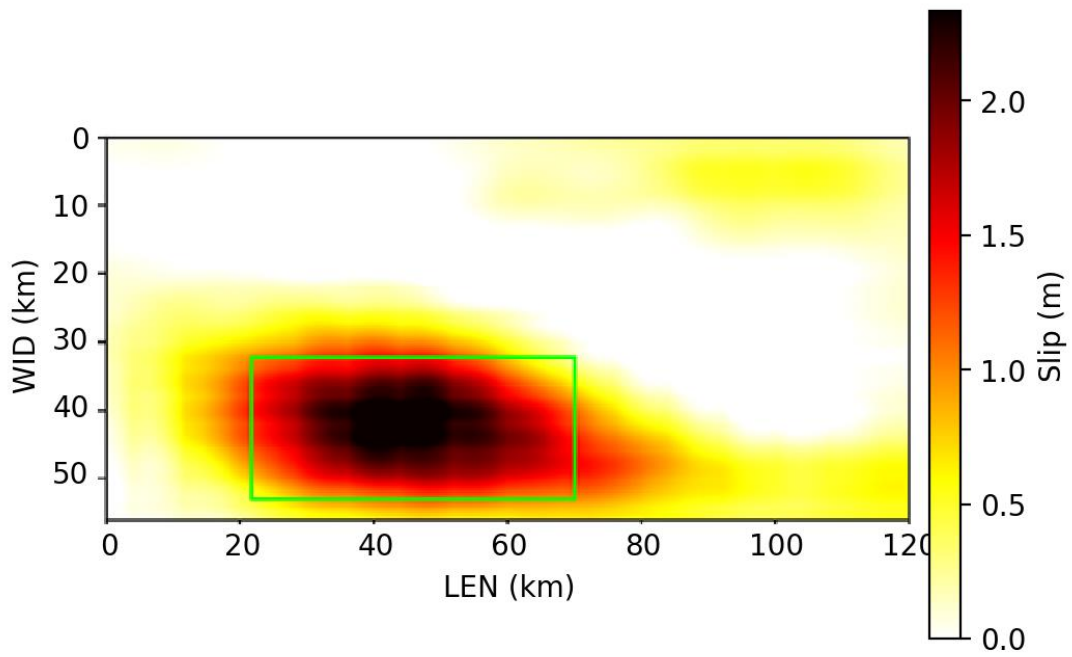
s1999HECTOR01JONS. Evento del 16/10/1999, Southern California, United States, magnitudo 7.16.



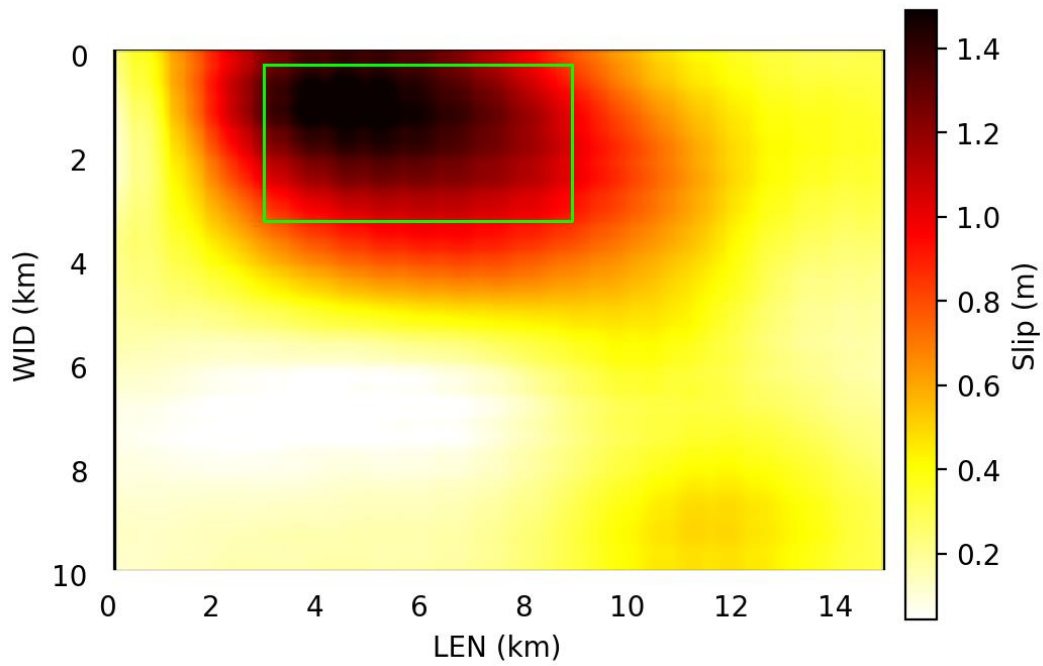
s2003MIYAGI01MIUR. Evento del 25/07/2003, Near east coast of eastern Honshu, Japan , magnitudo 6.08.



s2007KURILI01JIxx. Evento del 13/01/2007, East Of Kuril Islands, magnitudo 8.1.

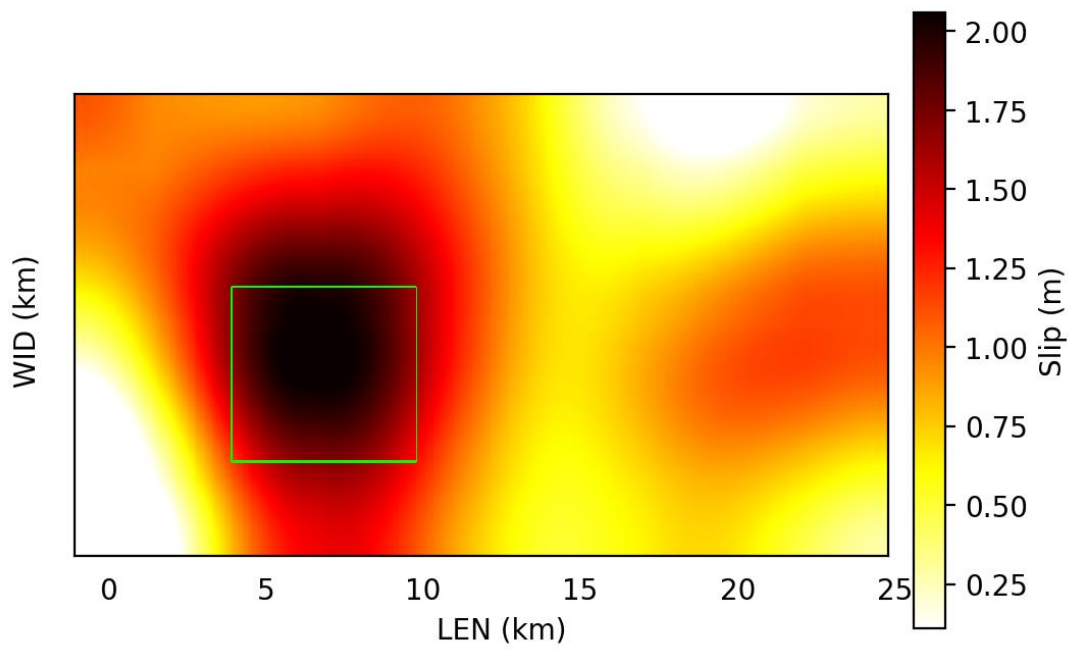


s2008SULAWE01SLAD. Evento del 16/11/2008, Minahassa Peninsula, Sulawesi, magnitudo 7.30.

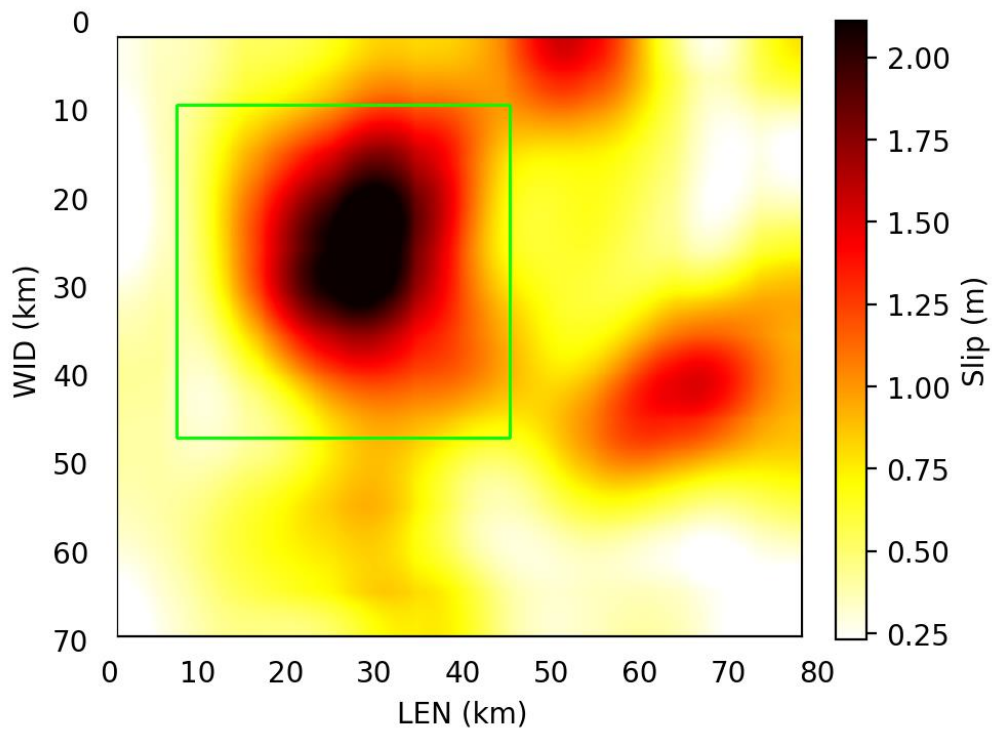


s20142014SO01GALL. Evento del 24/08/2014, Northern California, magnitudo 6.07.

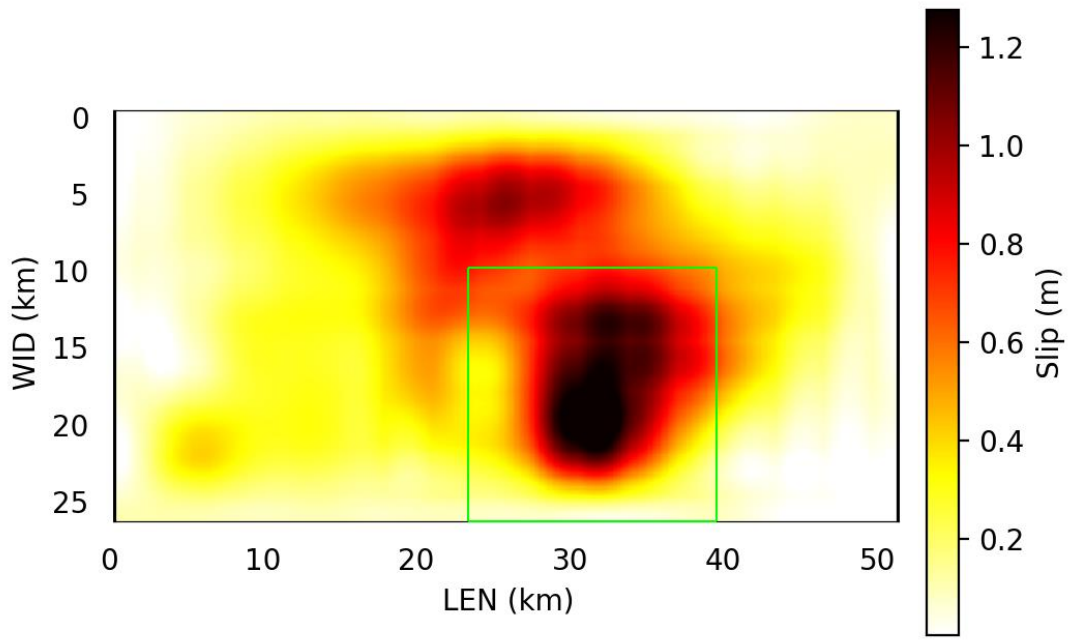
Intervallo $I_1 \in [0.4, 0.6[$



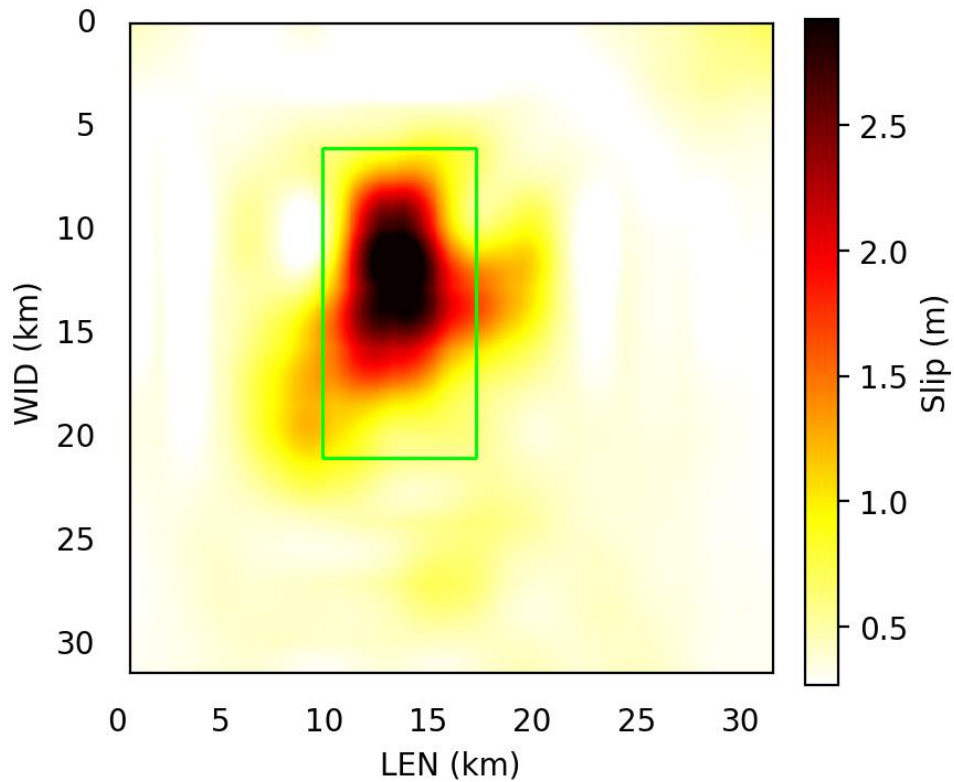
s1945MIKAWA01KIKU. Evento del 12/01/1945, Near south coast of eastern Honshu, Japan, magnitudo 6.67.



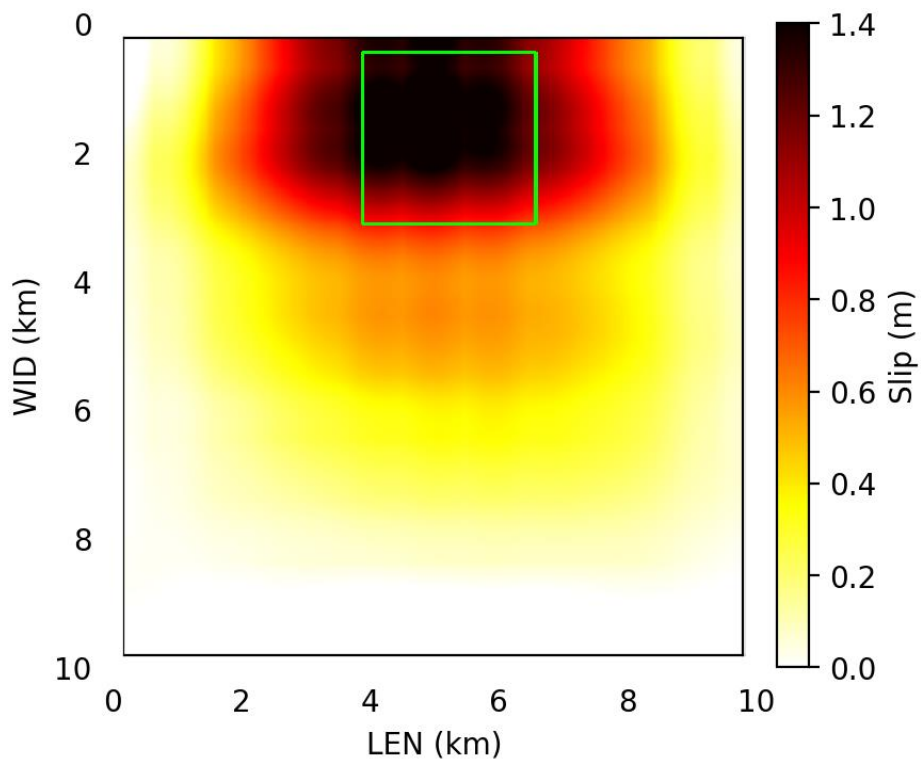
s1978MIYAGI01YAMA. Evento del 12/06/1978, Near east coast of eastern Honshu, Japan, magnitudo 7.61.



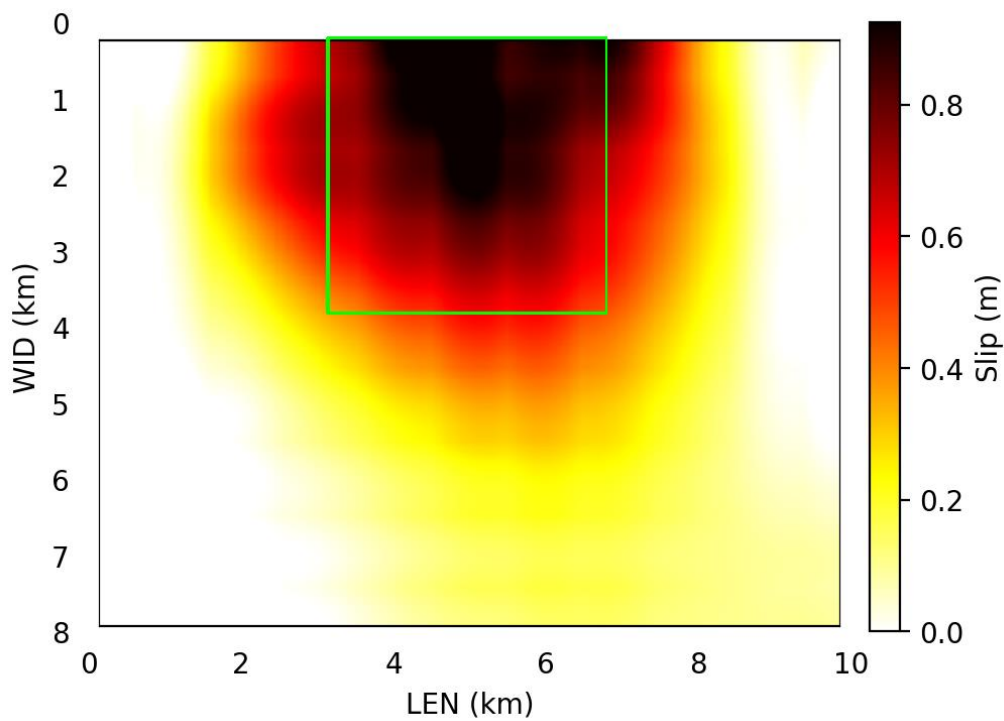
s1983BORAHP01MEND. Evento del 28/10/1983, Eastern Idaho, United States, magnitudo 6.82.



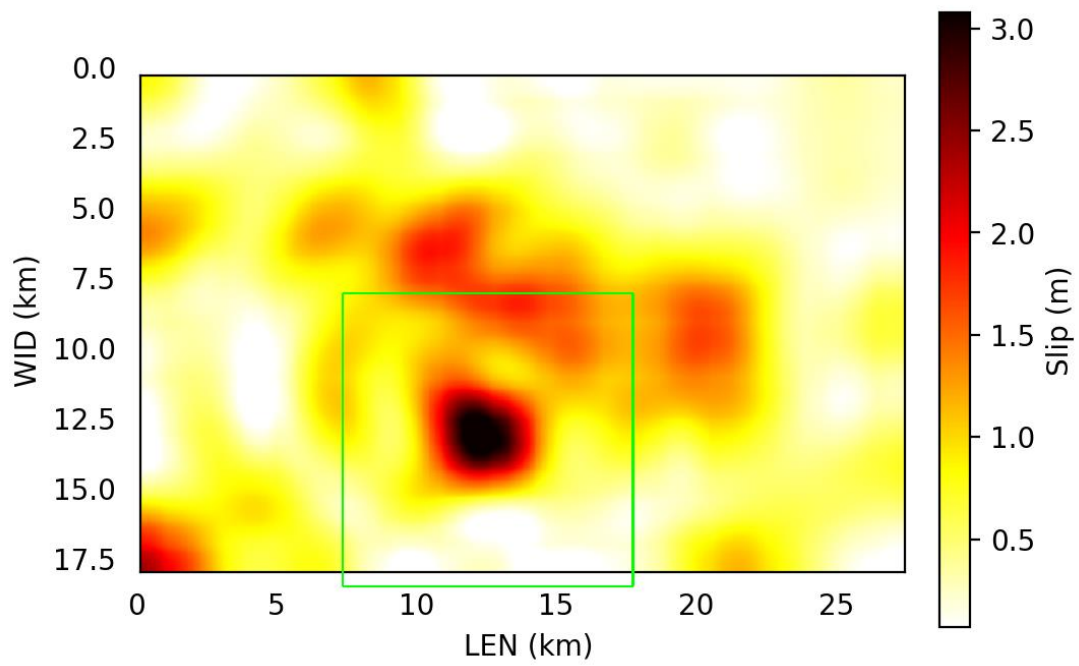
s1996HYUGAx01YAGI. Evento del 19/10/1996, Southeast of Shikoku, Japan, magnitudo 6.81.



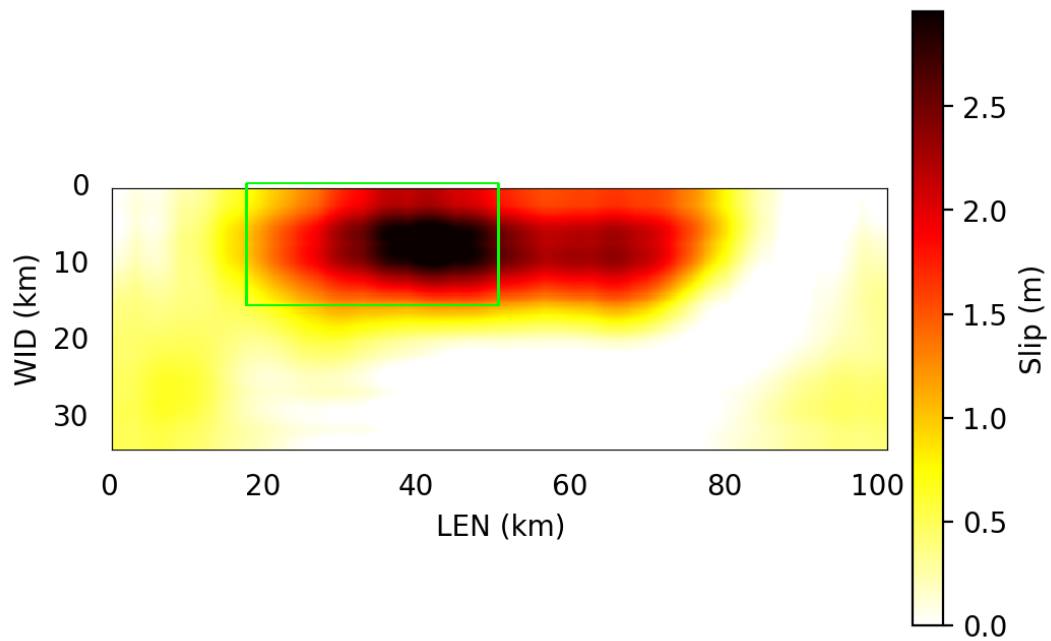
s1998IWATEJ01NAKA. Evento del 03/09/1998, Eastern Honshu, Japan, magnitudo 6.30.



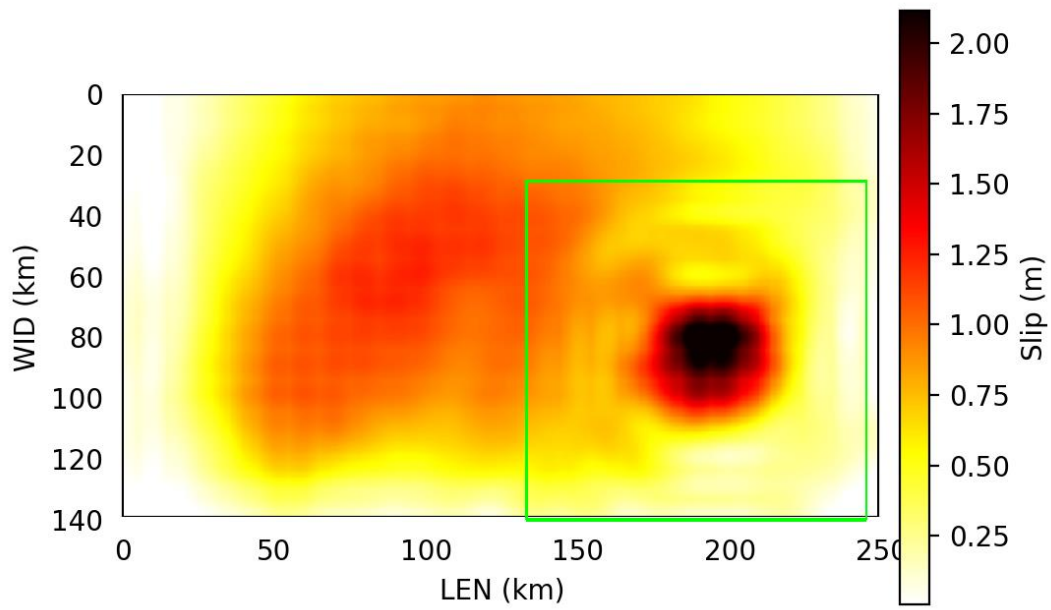
s2000KLEIFA01SUDH. Evento del 17/06/2000, Iceland region, magnitudo 5.87.



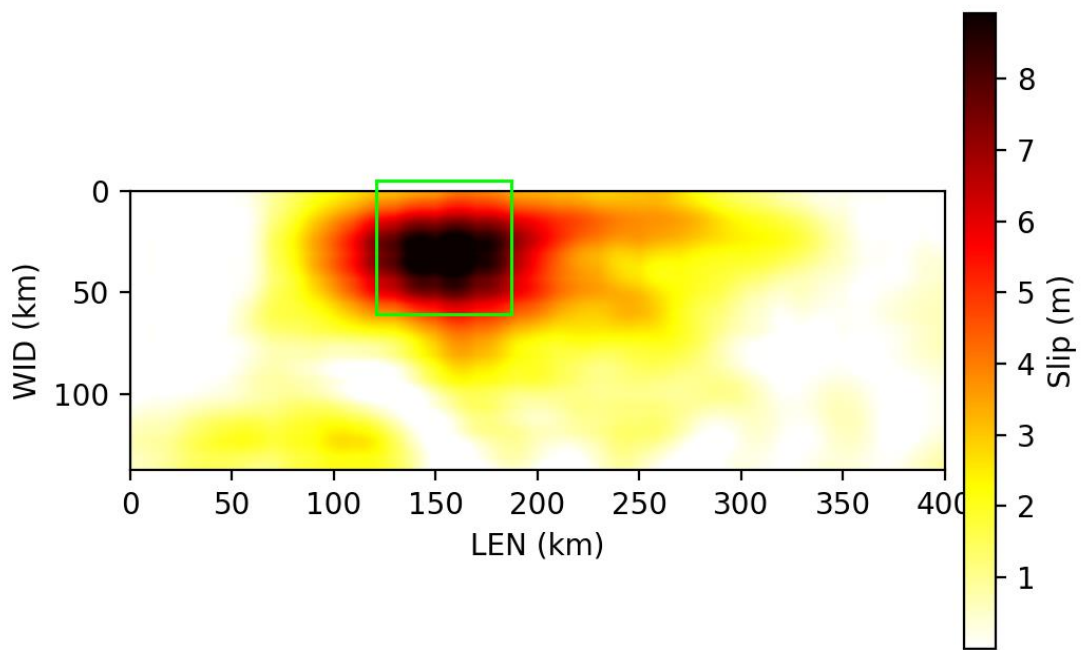
s2004NIIGAT01ASAN. Evento del 23/10/2004, Near West Coast Of Honshu, Japan, magnitudo 6.62.



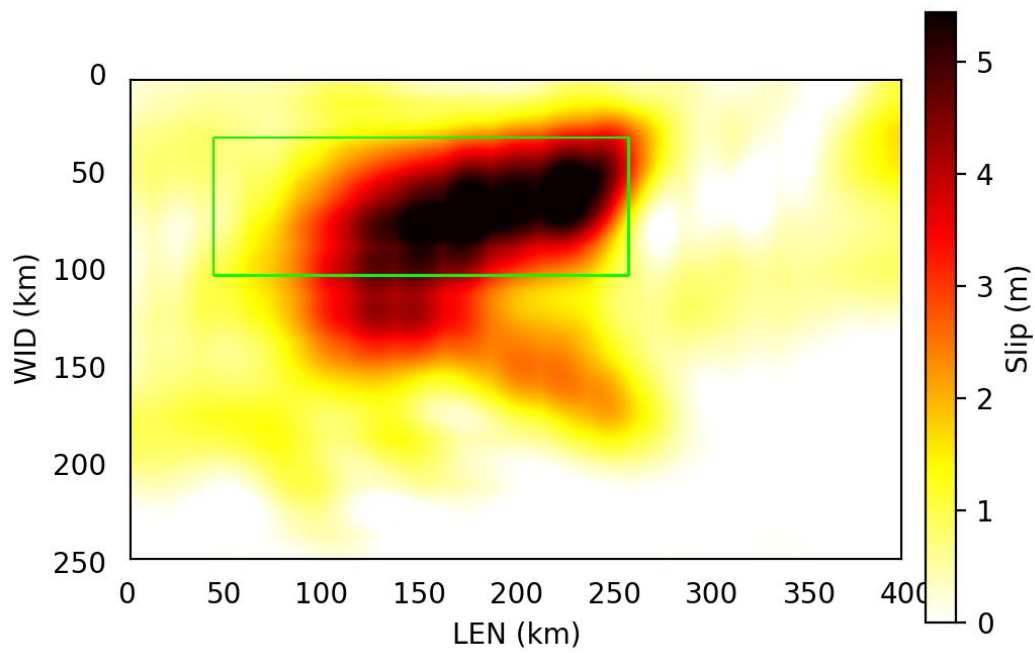
s2005NORTHE01SHAO. Evento del 15/06/2005, Off Coast Of Northern California, magnitudo 7.20.



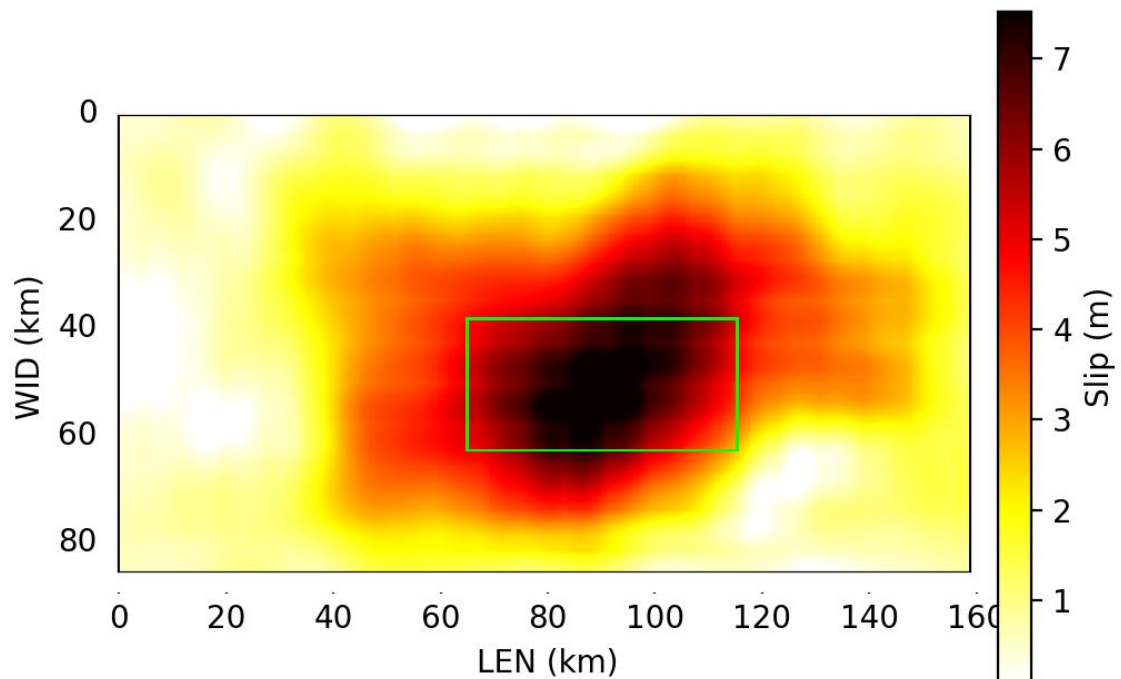
s2006JAVAIN01YAGI. Evento del 17/07/2006, South Of Java, Indonesia, magnitudo 7.82.



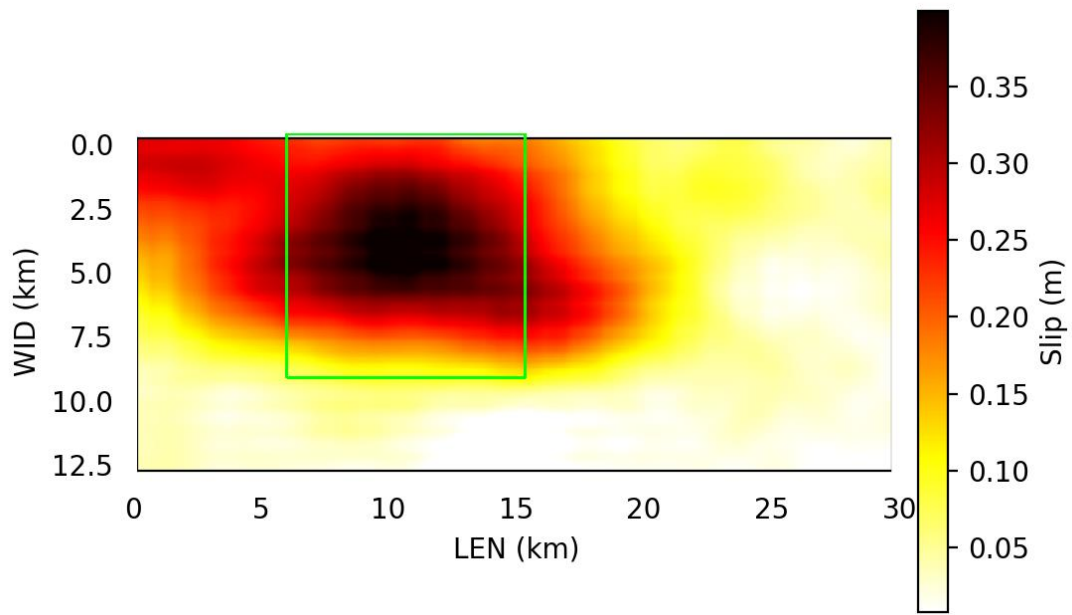
s2006KURILI01JIxx. Evento del 15/11/2006, Kuril Islands, magnitudo 8.30.



s2007BENGKU02GUSM. Evento del 12/09/2007, Southern Sumatra, Indonesia, magnitudo 8.40.

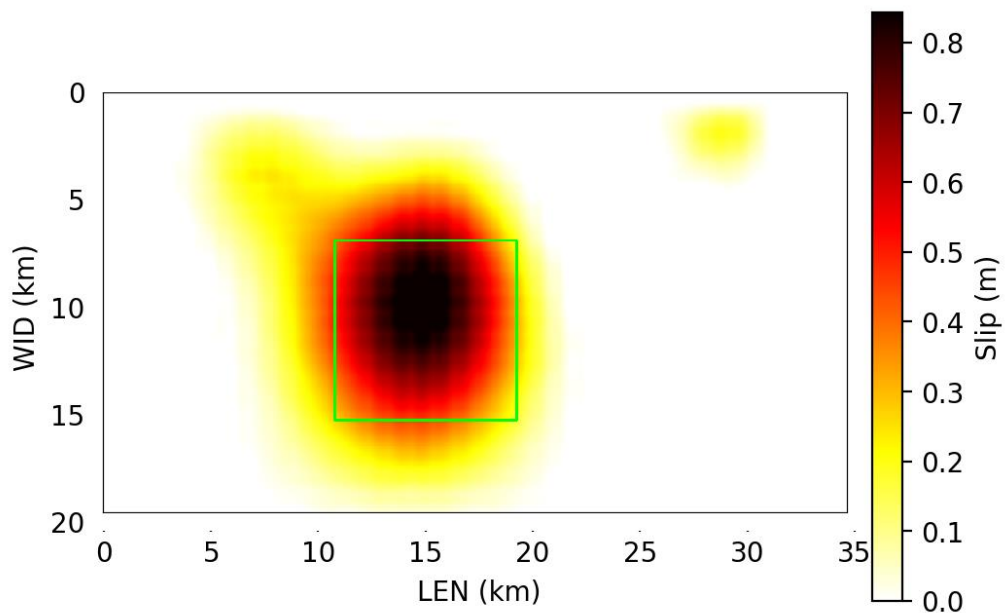


s2015GORKHA01YAGI. Evento del 25/04/2015, Nepal, magnitudo 7.90.

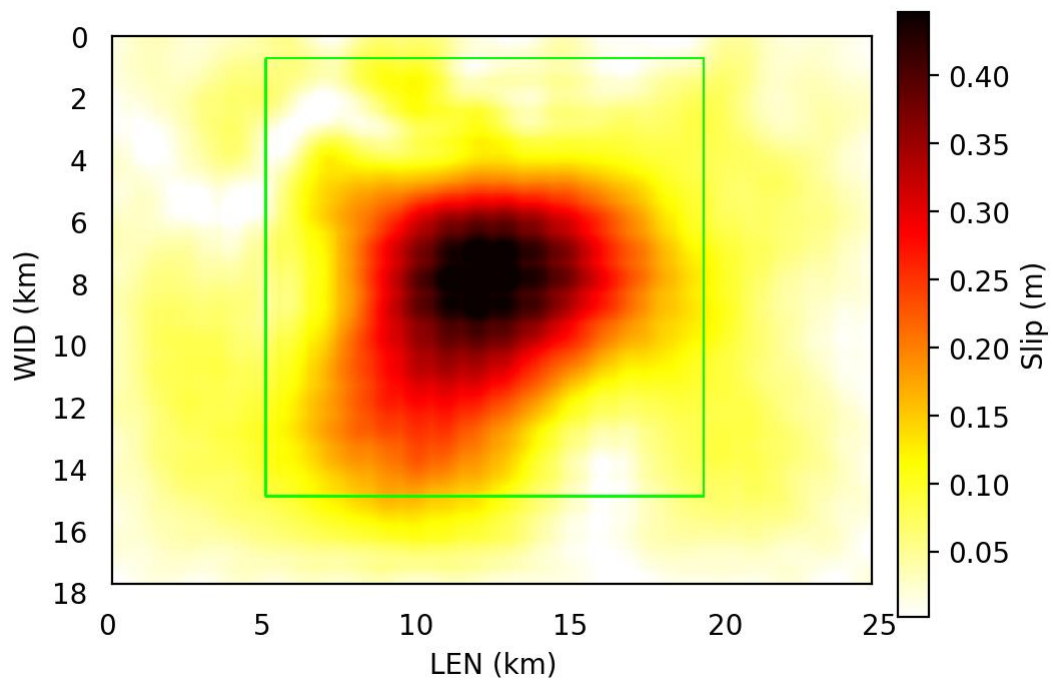


s2016USSITA01PIZZ. Evento del 26/10/2016, Central Italy, magnitudo 6.09.

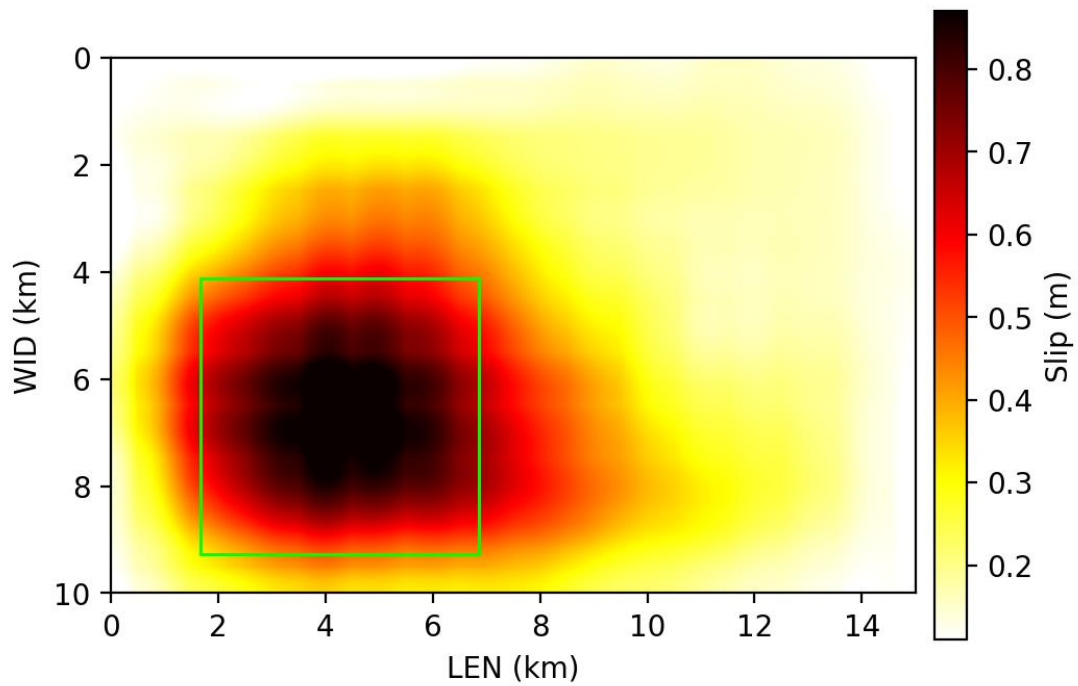
Intervallo $I_1 \in [0.6, 0.8[$



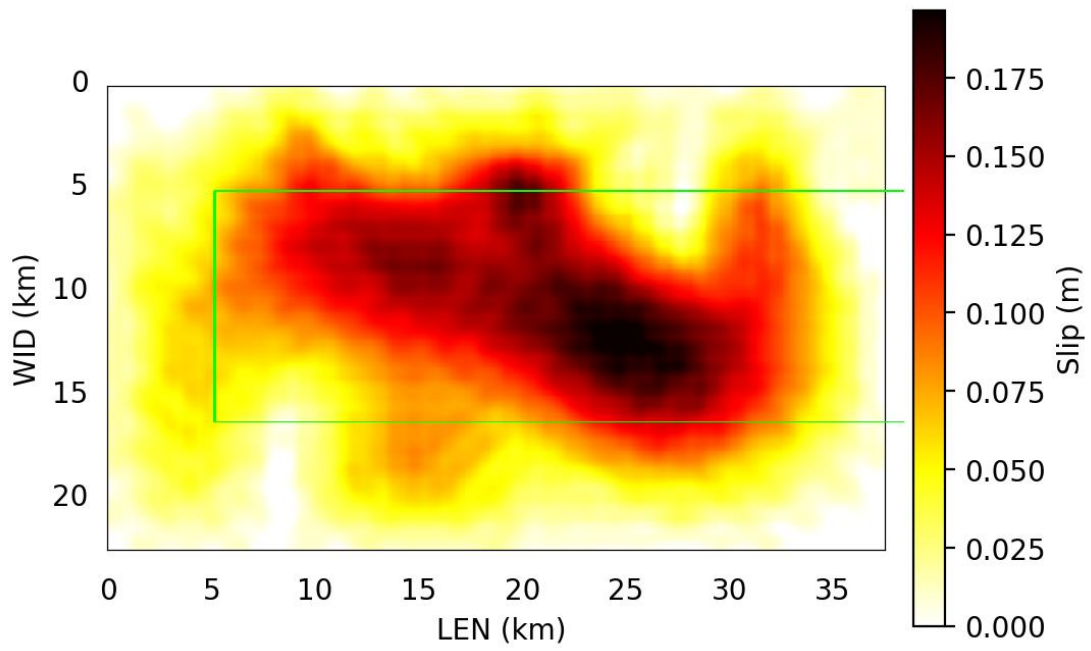
s1992JOSHUA01BENN. Evento del 23/04/1992, Southern California, United States, magnitudo 6.15.4



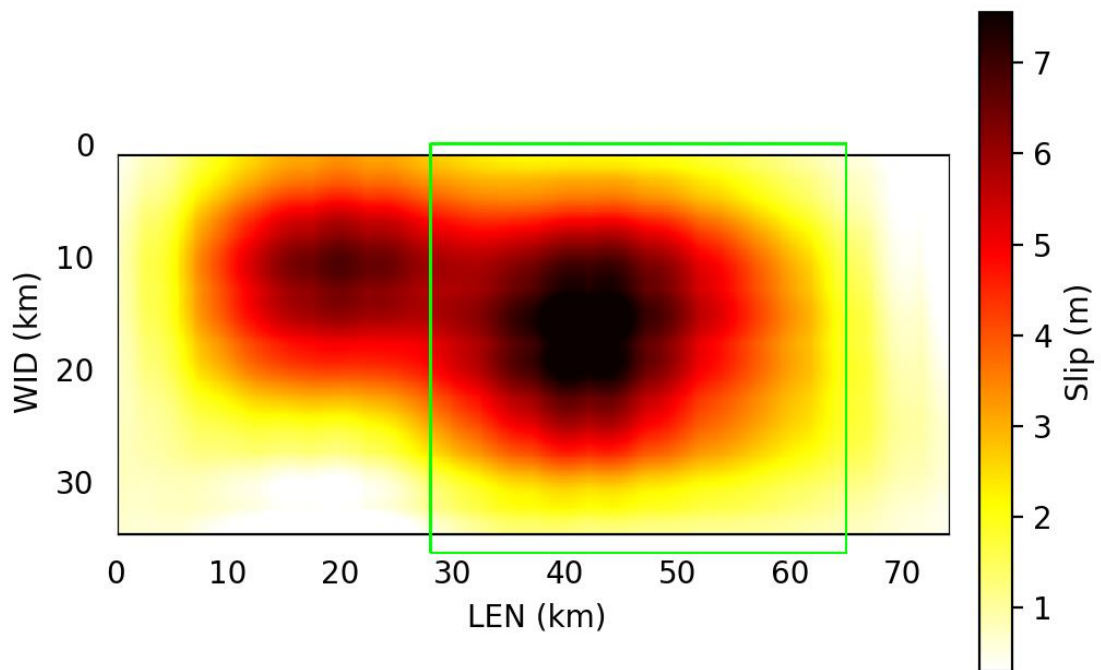
s1996PUMQUx01WANG. Evento del 03/07/1996, Xizang, magnitudo 6.08.



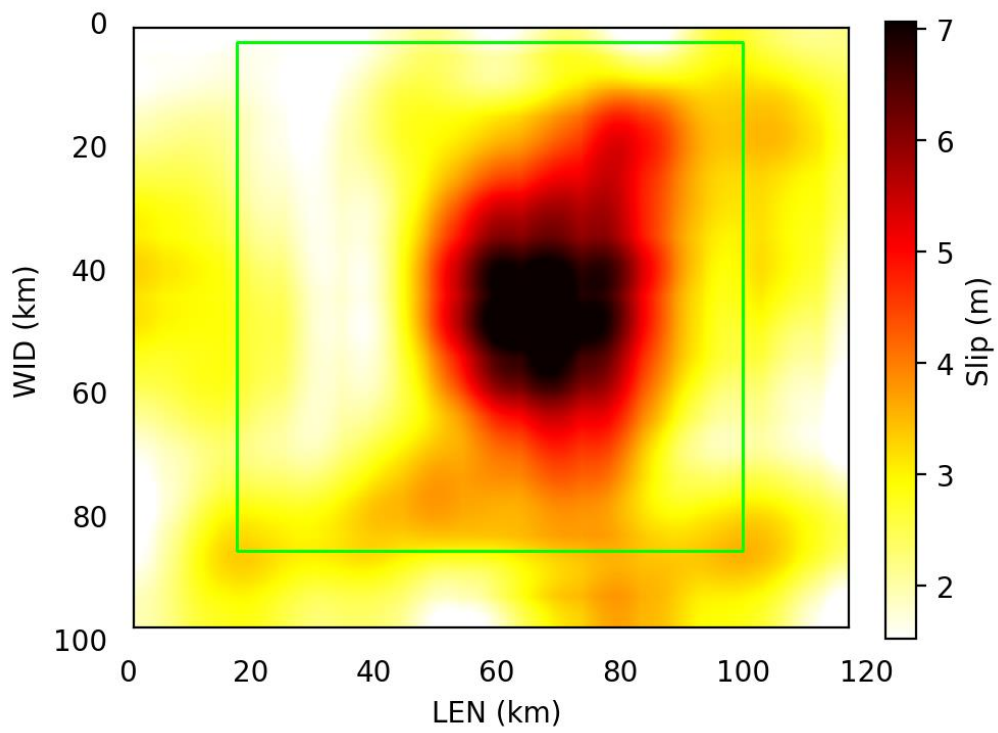
s1997KAGOSH01HORI. Evento del 26/03/1997, Kyushu, Japan , magnitudo 6.10.



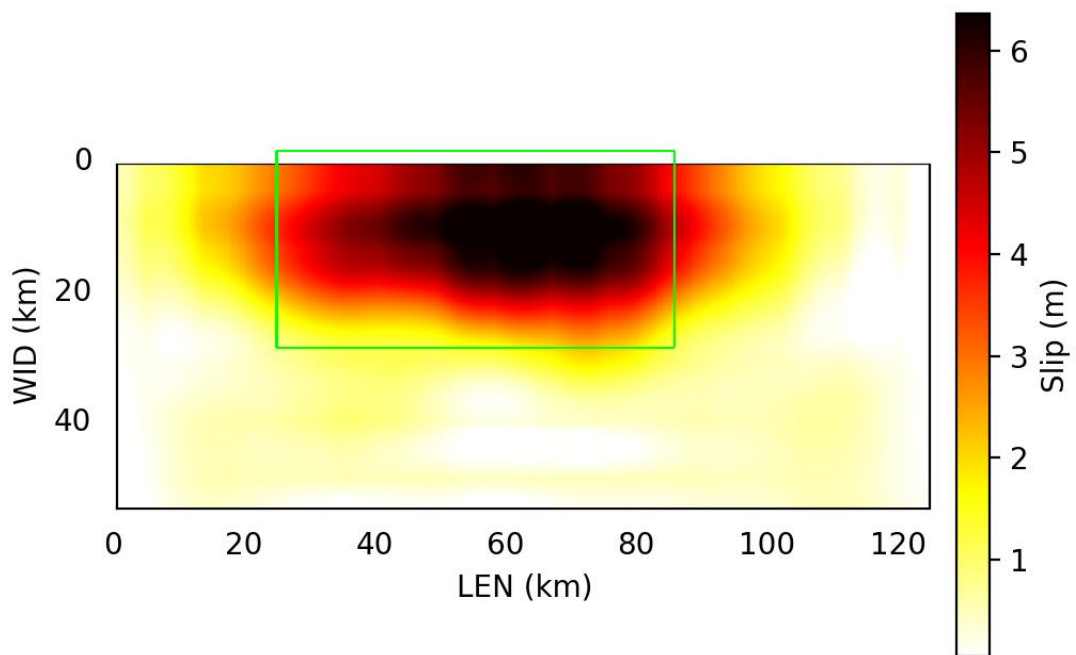
1998PUMQUx01WANG. Evento del 25/08/1998, Xizang, magnitudo 6.16.



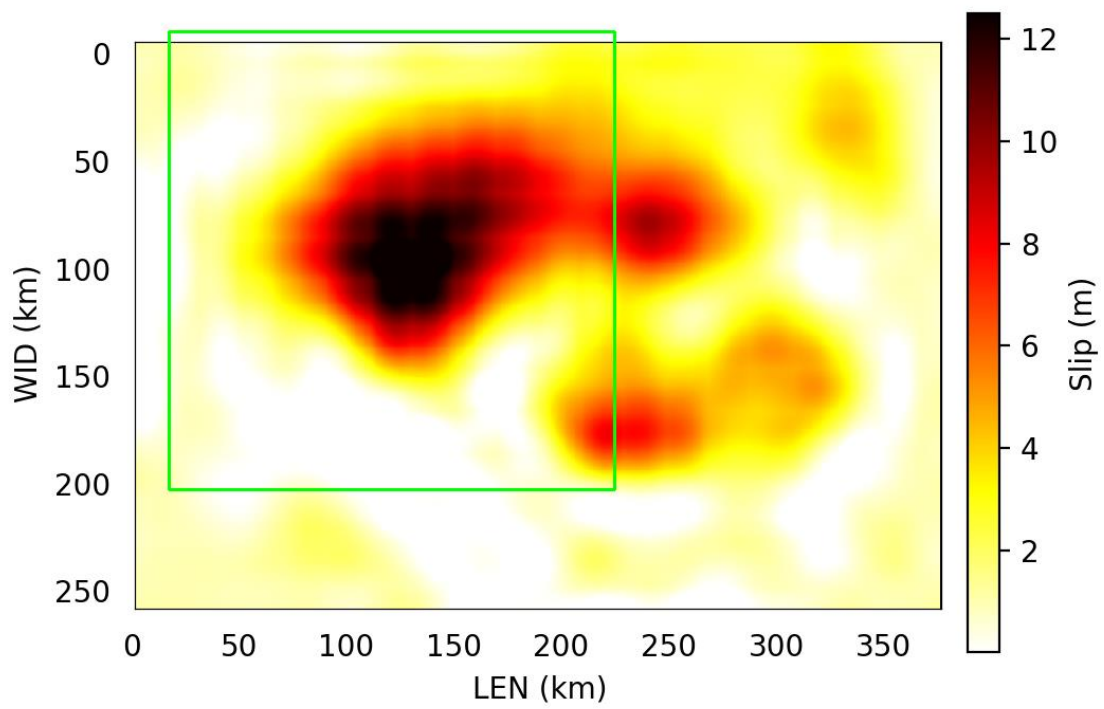
s2001BHUJIN01YAGI. Evento del 26/01/2001, Southern India, magnitudo 7.66.



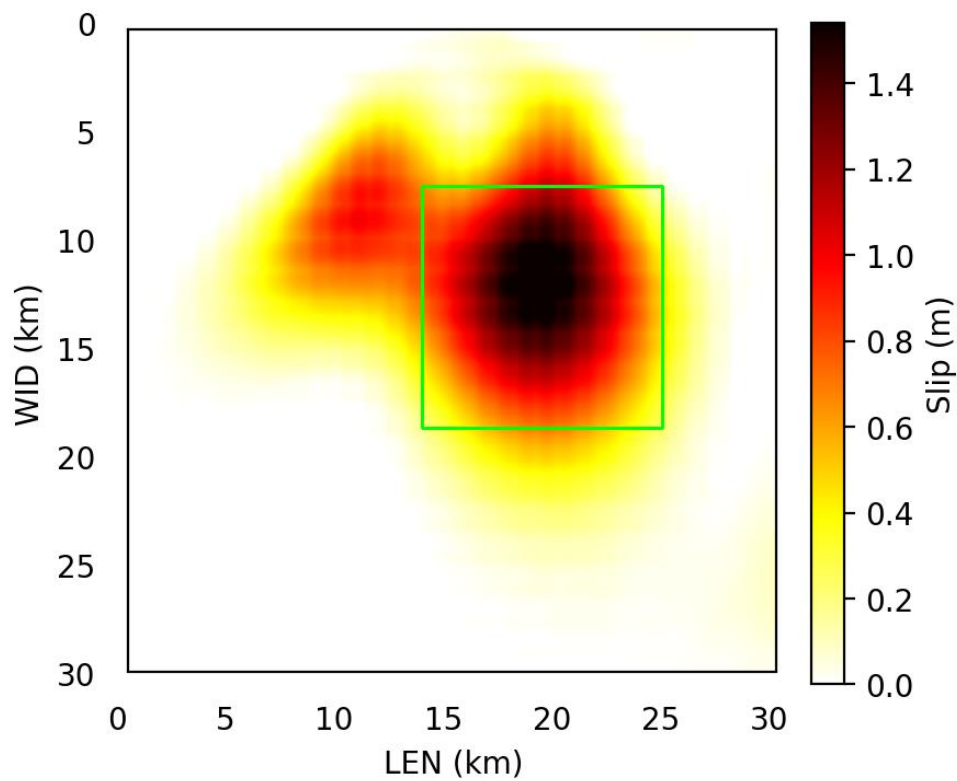
s2003TOKACHO1KOKE. Evento del 25/09/2003, Hokkaido, Japan, region, magnitudo 8.03.



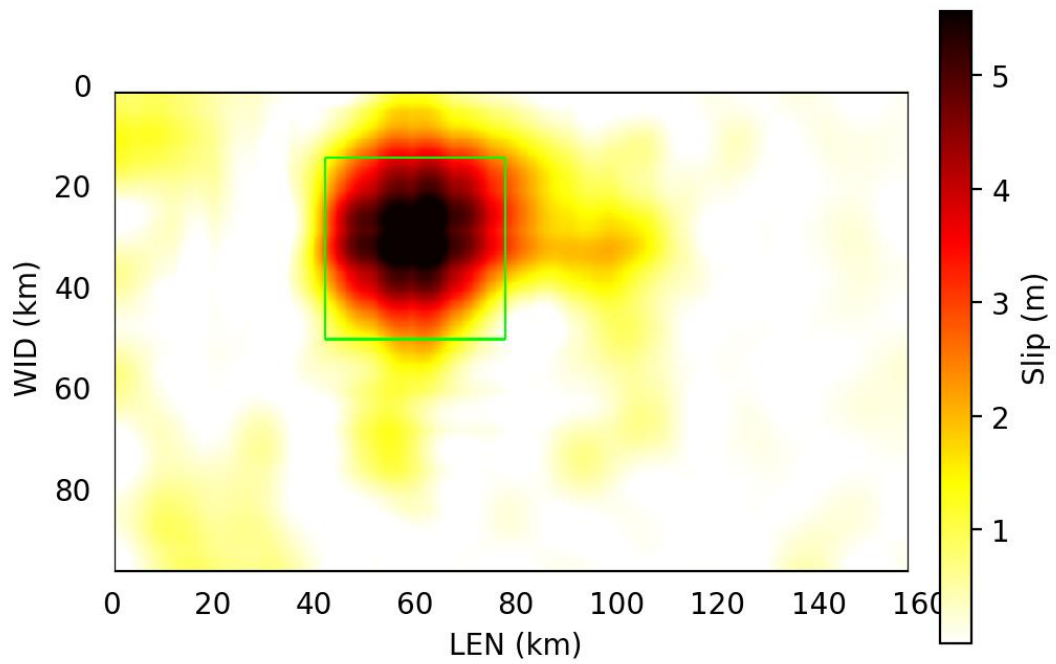
s2005KASHMI01SHAO. Evento del 08/10/2005, Pakistan, magnitudo 7.60.



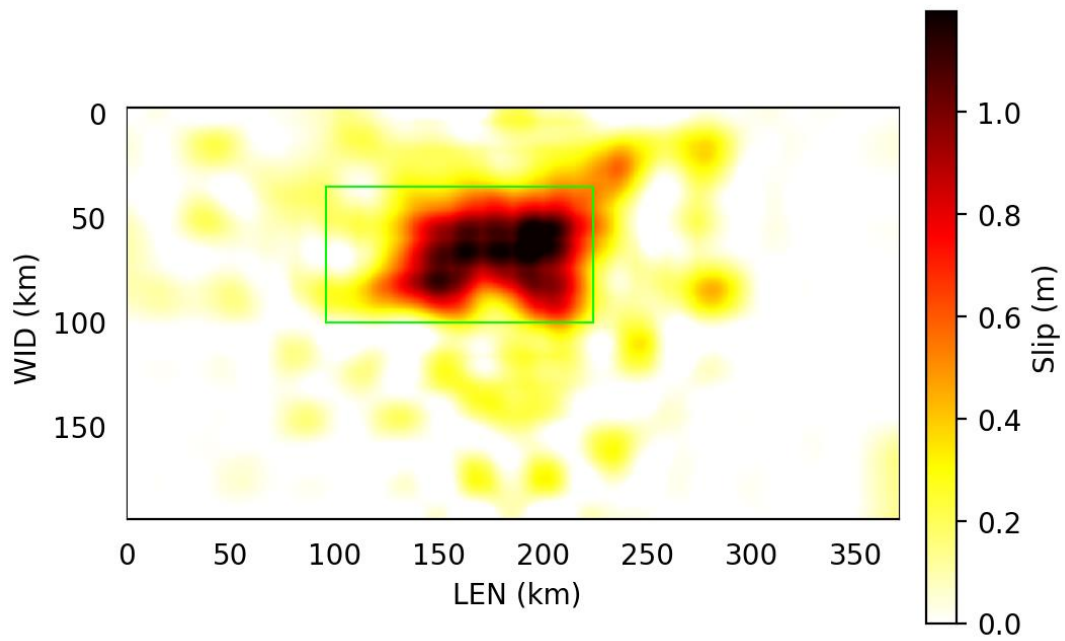
s2005SUMATR01SHA0. Evento del 28/03/2005, Northern Sumatra, Indonesia, magnitudo 8.70.



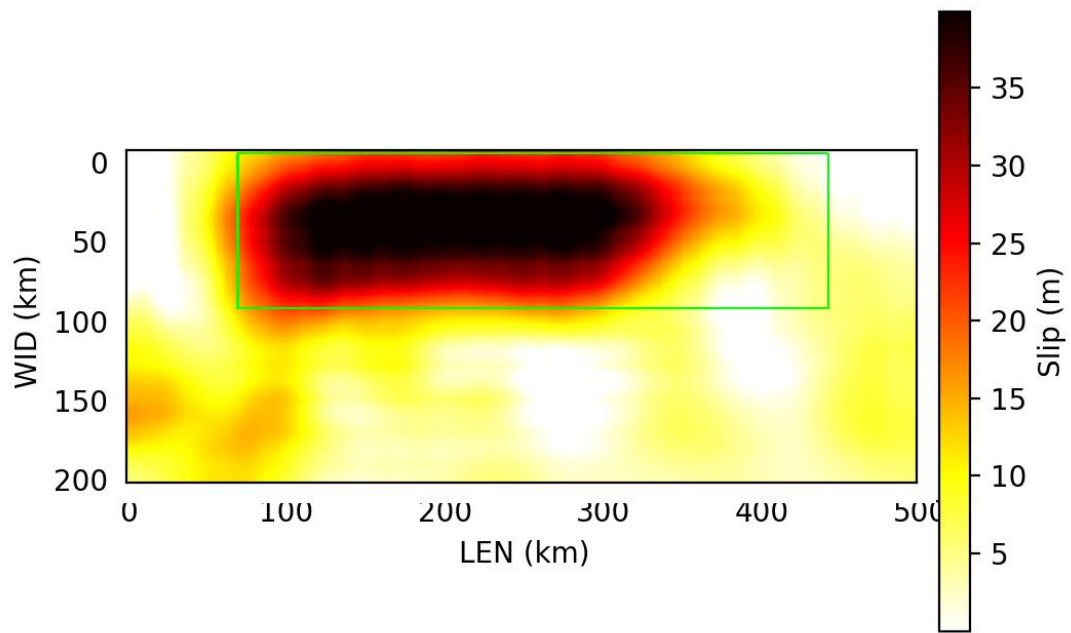
s2008ZHONGB01ELLI. Evento del 25/08/2008, Xizang , magnitudo 6.70.



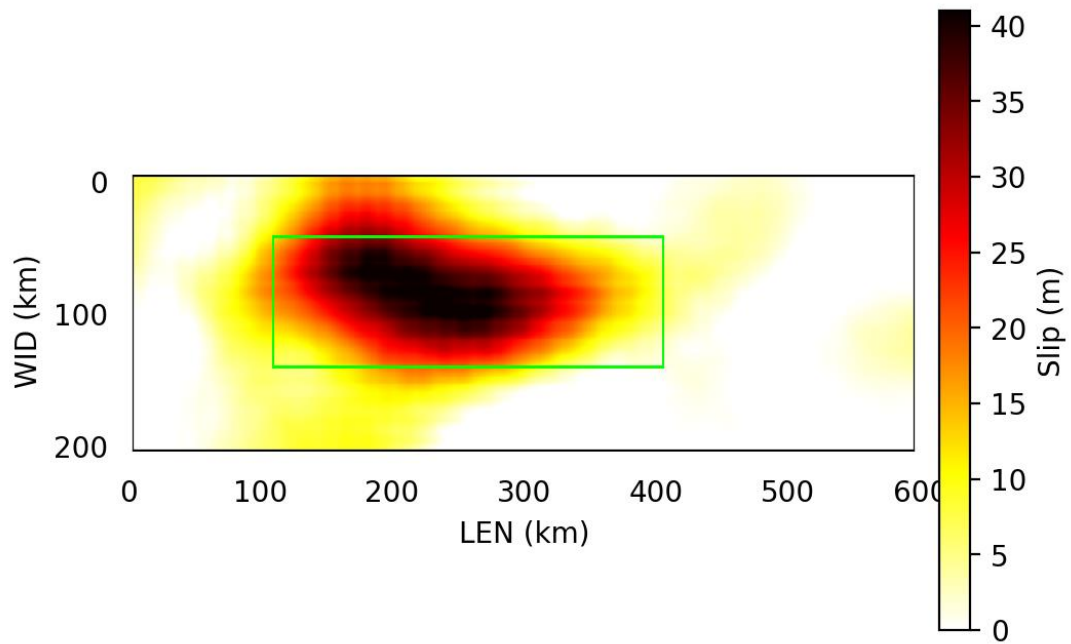
s2009FIORDL01HAYE. Evento del 15/07/2009, Off W. Coast Of S. Island, N.Z., magnitudo 7.60.



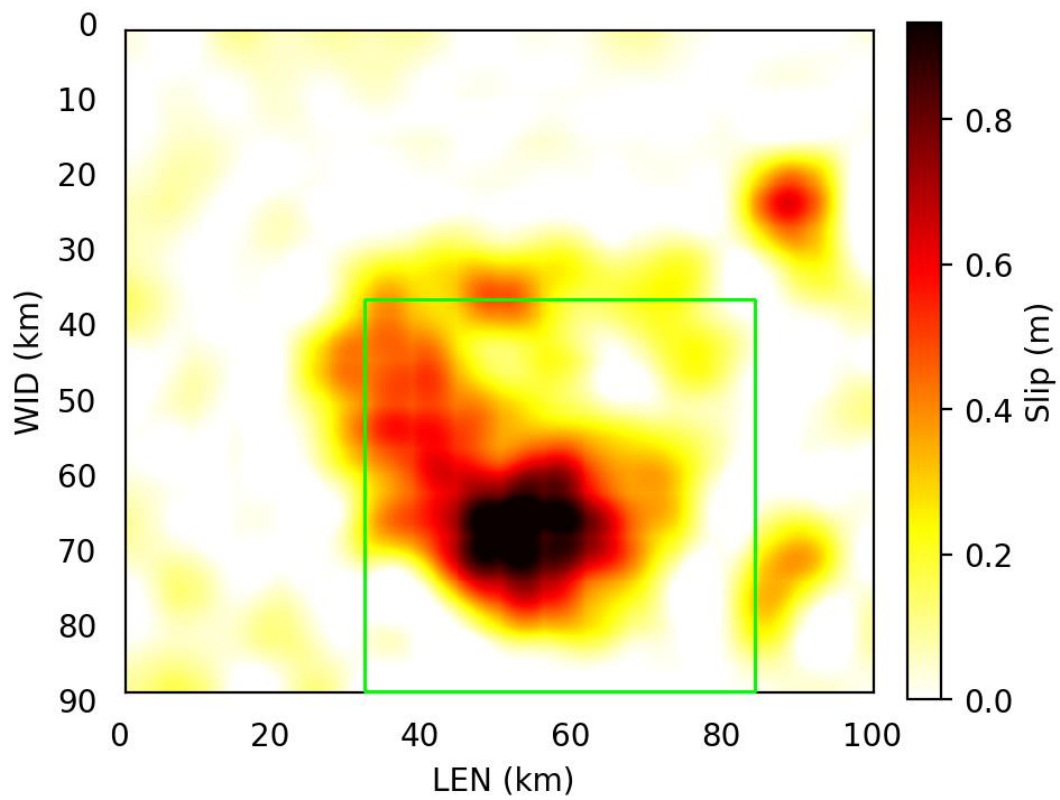
s2010SUMATR01HAYE. Evento del 25/10/2010, Southern Sumatra, Indonesia, magnitudo 7.70.



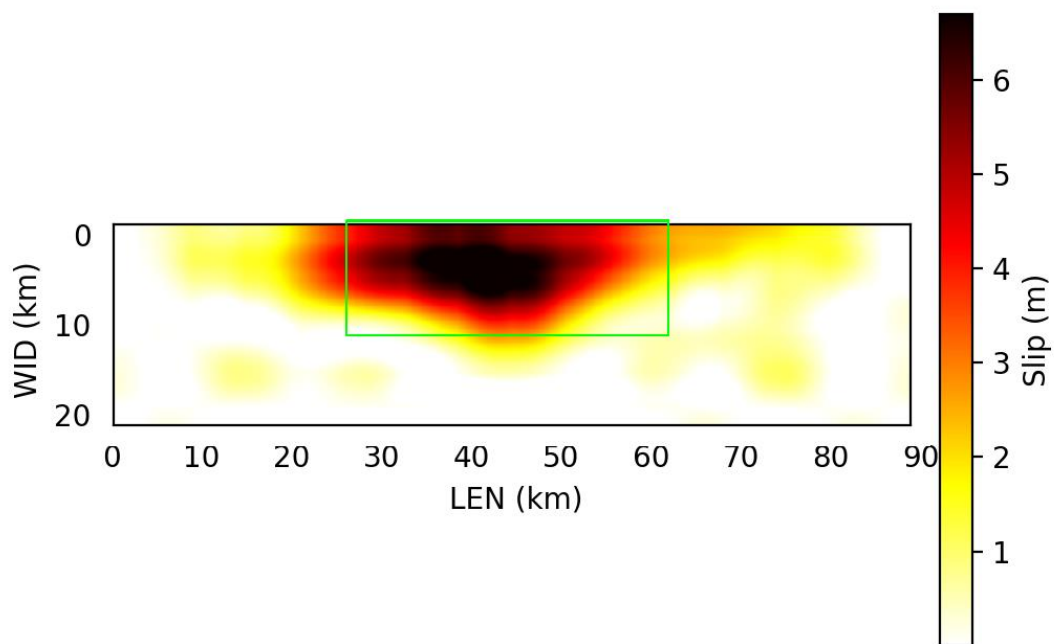
s2011HONSHU01SHAO. Evento del 09/03/2011, Near East Coast Of Honshu, Japan , magnitudo 7.30.



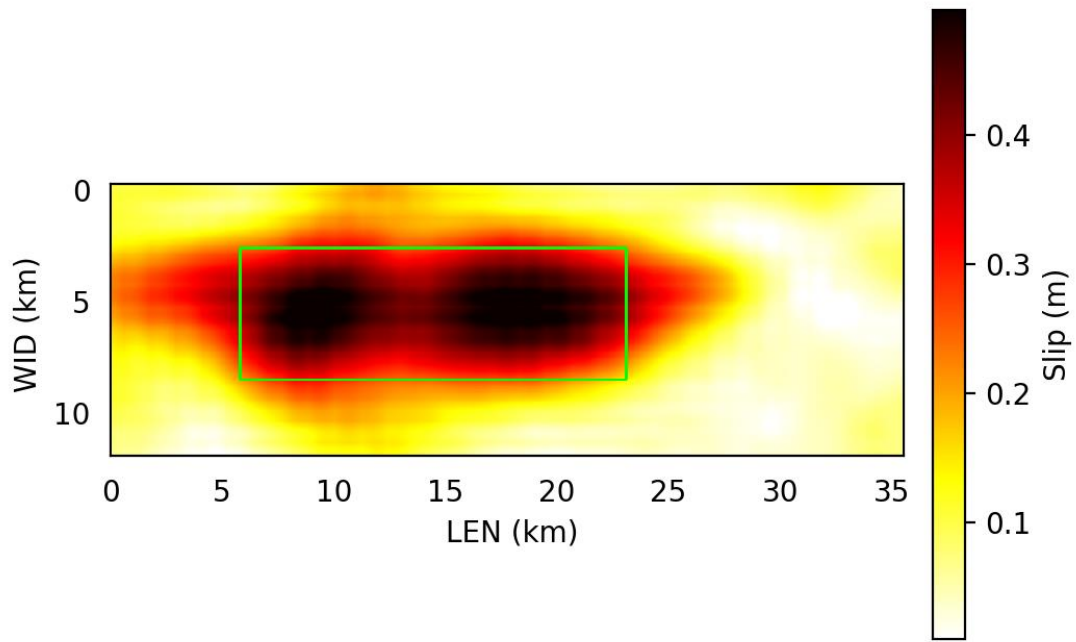
s2011TOHOKU01AMMO. Evento del 11/03/2011, Near East Coast Of Honshu, Japan , magnitudo 9.00.



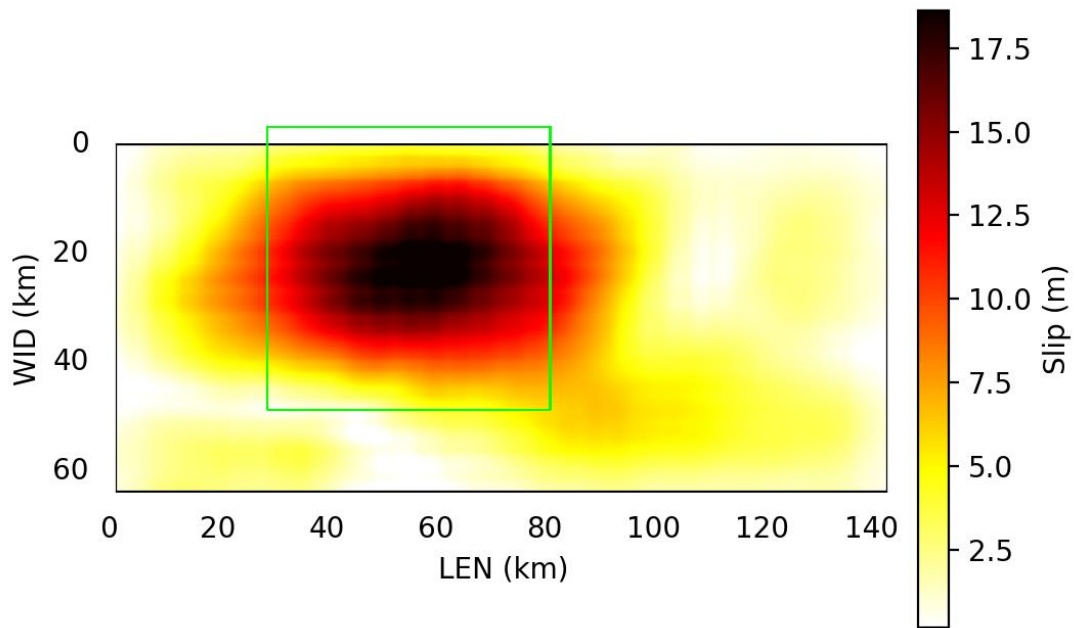
s2011VANUAT01HAYE. Evento del 20/08/2011, Vanuatu Islands, magnitud 7.30.



s2012SUMATR03HAYE. Evento del 11/04/2012, Off W Coast Of Northern Sumatra, magnitud 8.57.

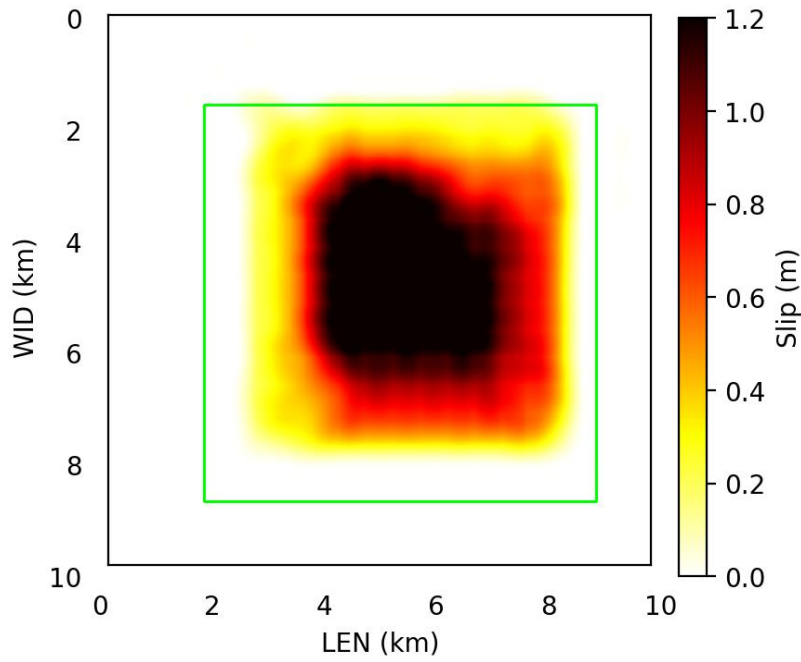


s2016AMATRI01PIZZ. Evento del 24/08/2016, Central Italy, magnitudo 6.21

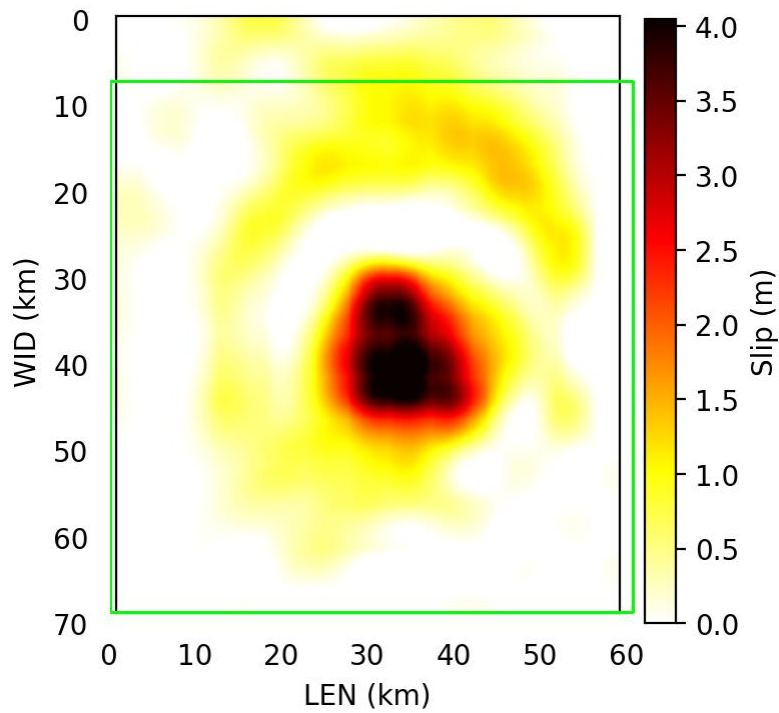


s2017CHIAPA01OKUW. Evento del 08/09/2017, Off Coast Of Chiapas, Mexico, magnitudo 8.10.

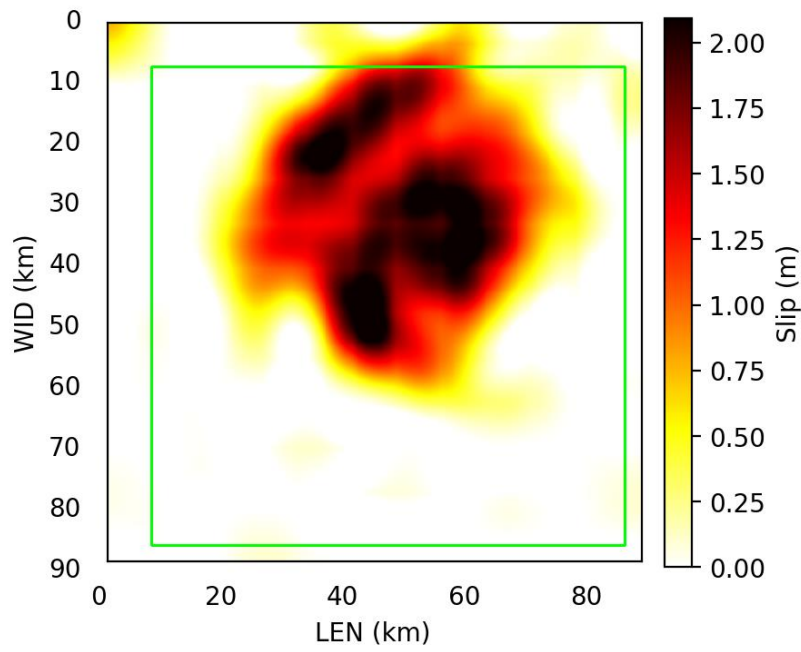
Intervallo $I_1 \in [0.8, 1.0[$



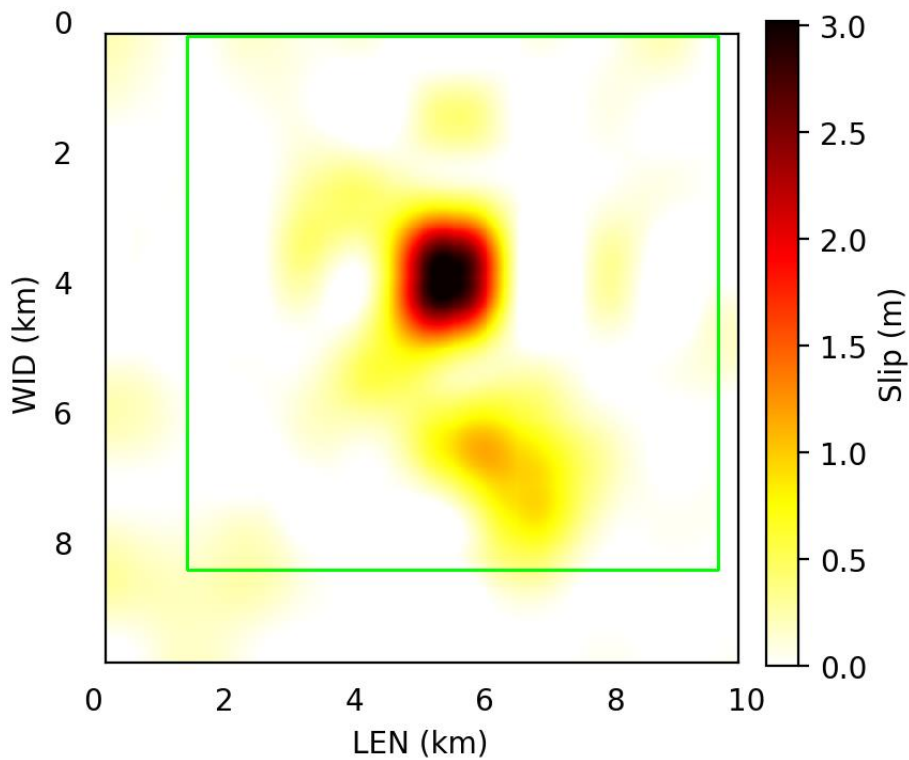
s1979COYOTE01LIUx. Evento del 06/08/1979, Central California, United States, magnitudo 5.92.



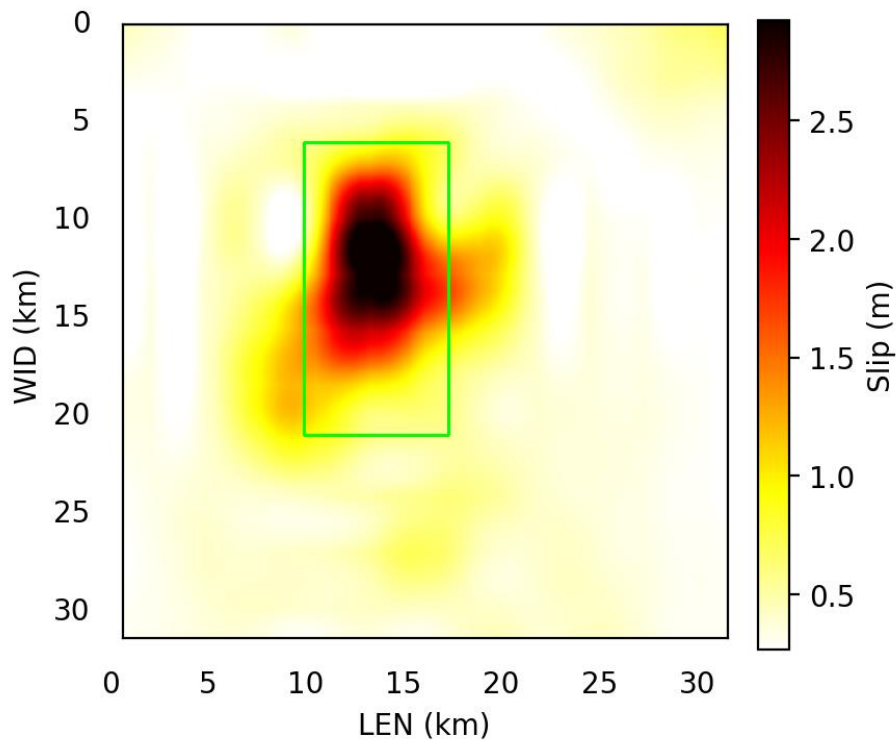
s1981PLAYAA01MEND. Evento del 25/10/1981, Near coast of Michoacan, Mexico, magnitudo 7.25.



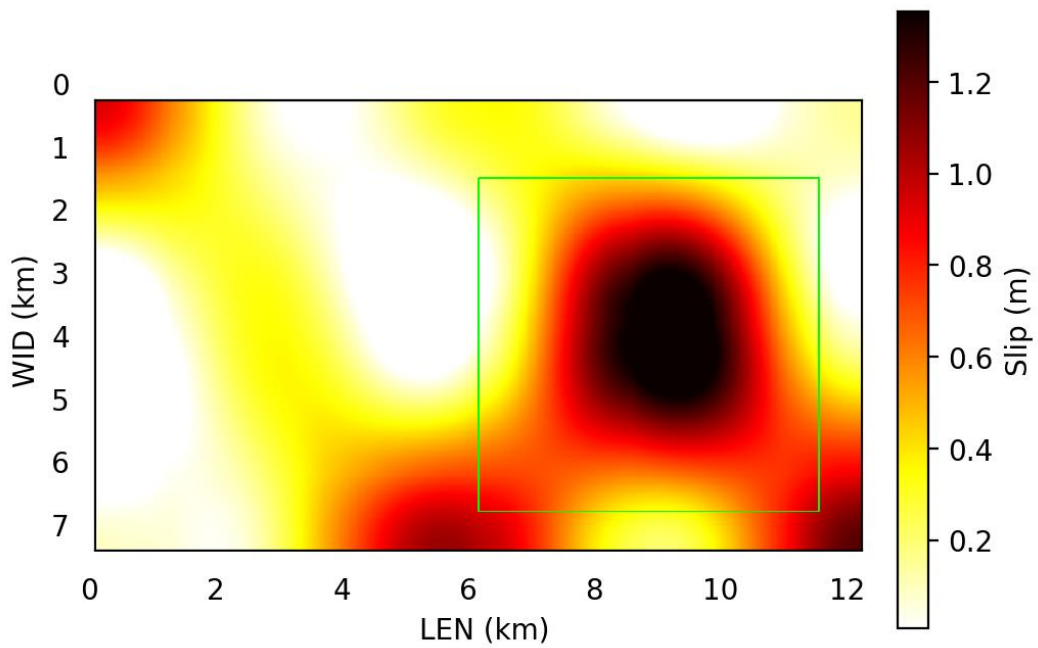
s1985ZIHUAT01MEND. Evento del 21/09/1985, Near coast of Guerrero, Mexico, magnitudo 7.42.



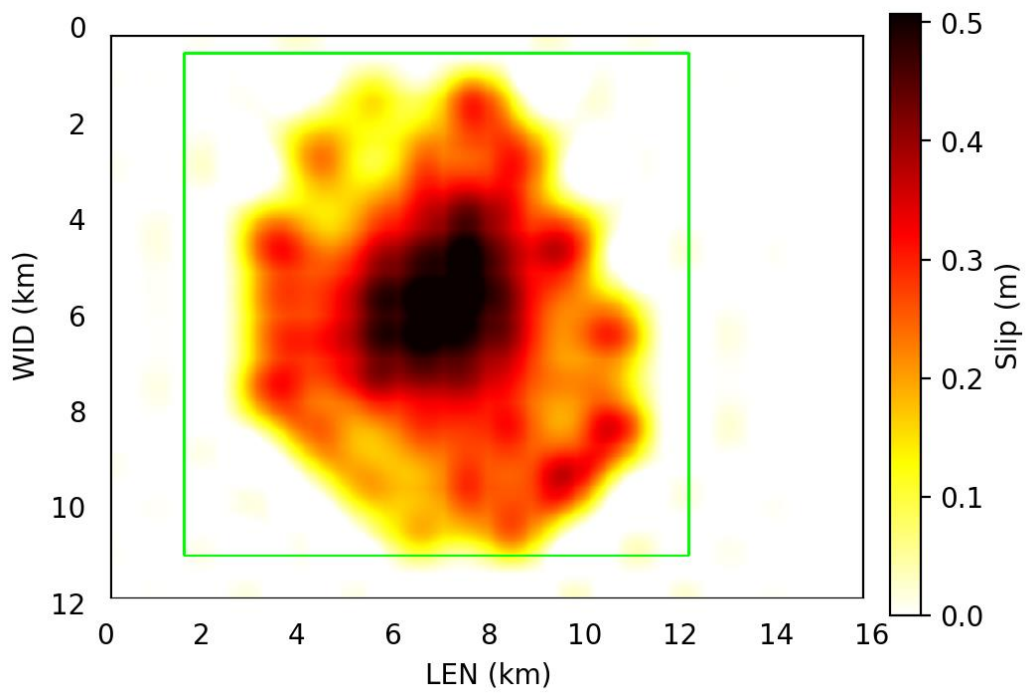
s1988SAGUEN01HART. Evento del 25/11/1988, Southern Quebec, Canada, magnitudo 5.81.



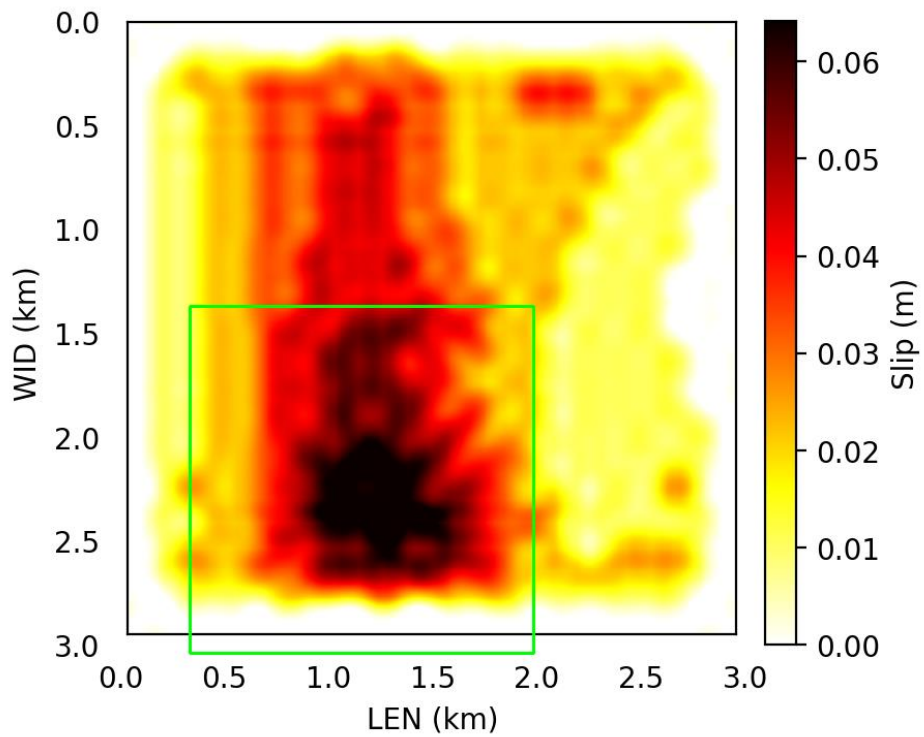
s1996HYUGAx01YAGI. Evento del 19/10/1996, Southeast of Shikoku, Japan, magnitudo 6.81.



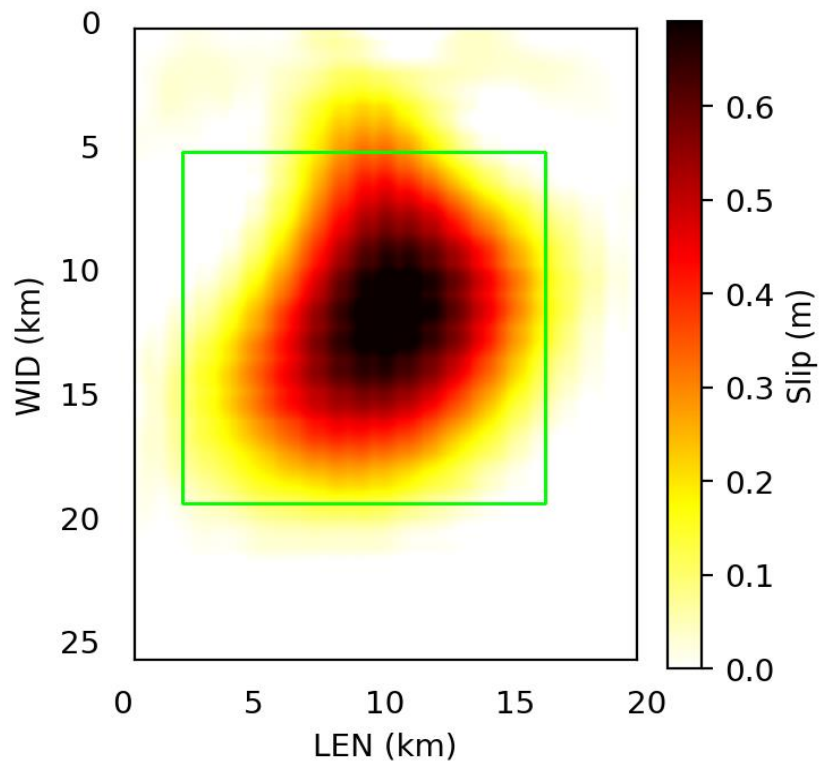
s1997COLFIO02HERN. Evento del 26/09/1997, Central Italy, magnitudo 5.97.



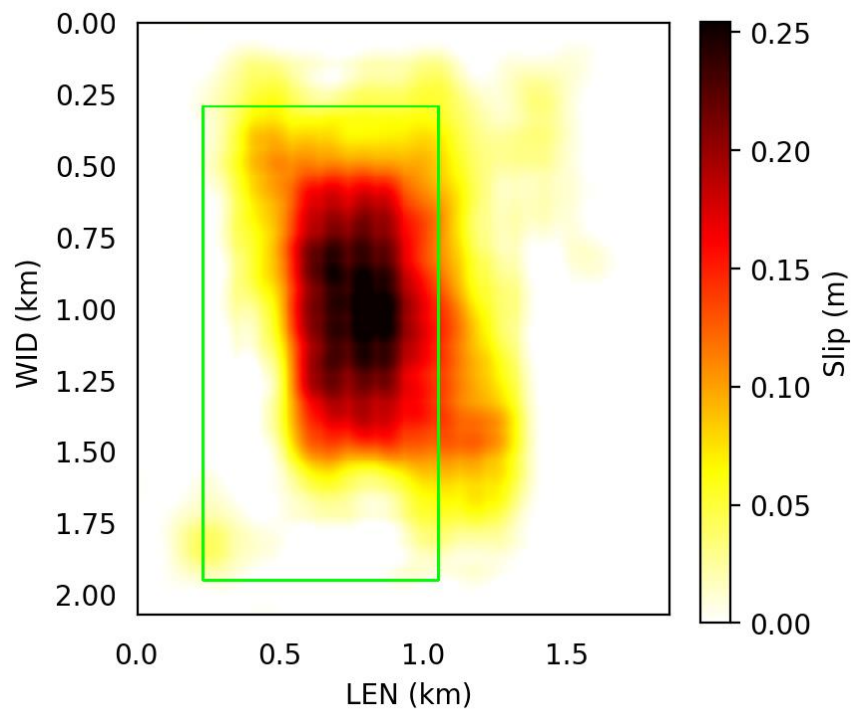
s1997YAMAGU01IDEx. Evento del 25/06/1997, Western Honshu, Japan, magnitudo 5.82.



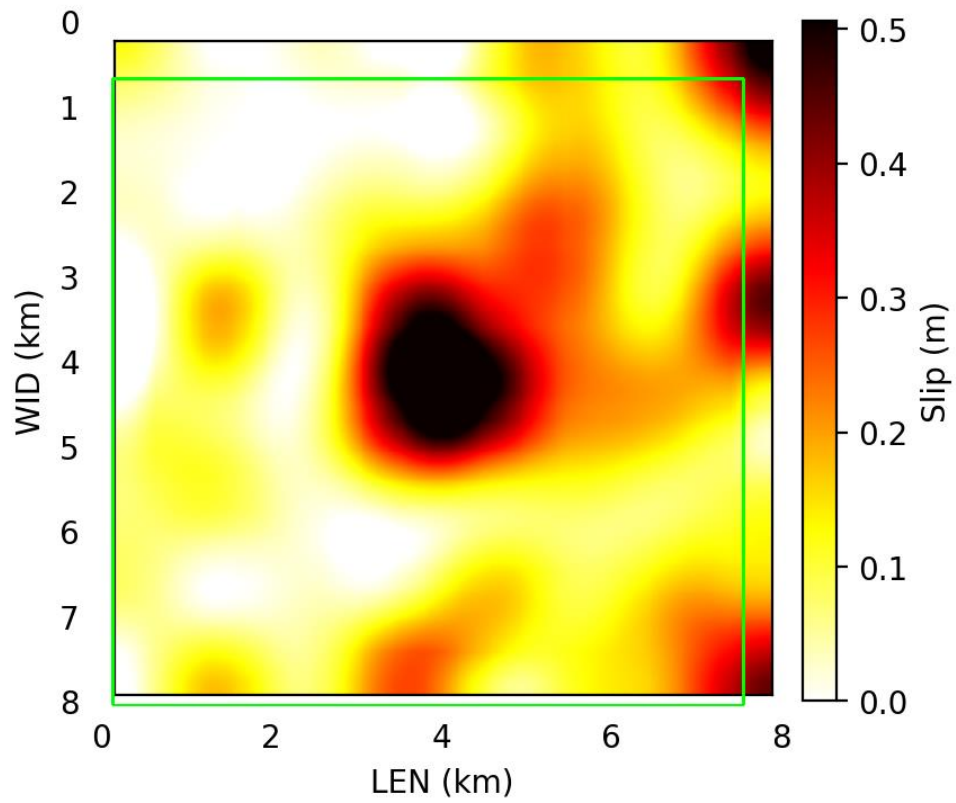
s1998HIDASW07IDEx. Evento del 22/08/1998, Eaastern Honshu, Japan, magnitudo 4.33.



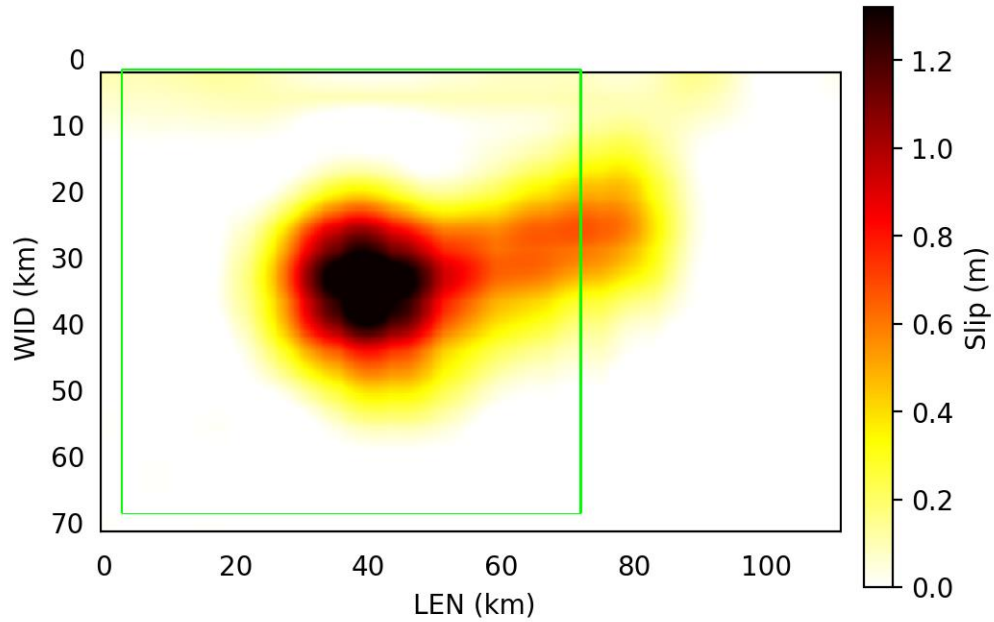
s2004ZHONGB01ELLI. Evento del 11/07/2004, Xizang, magnitudo 6.20.



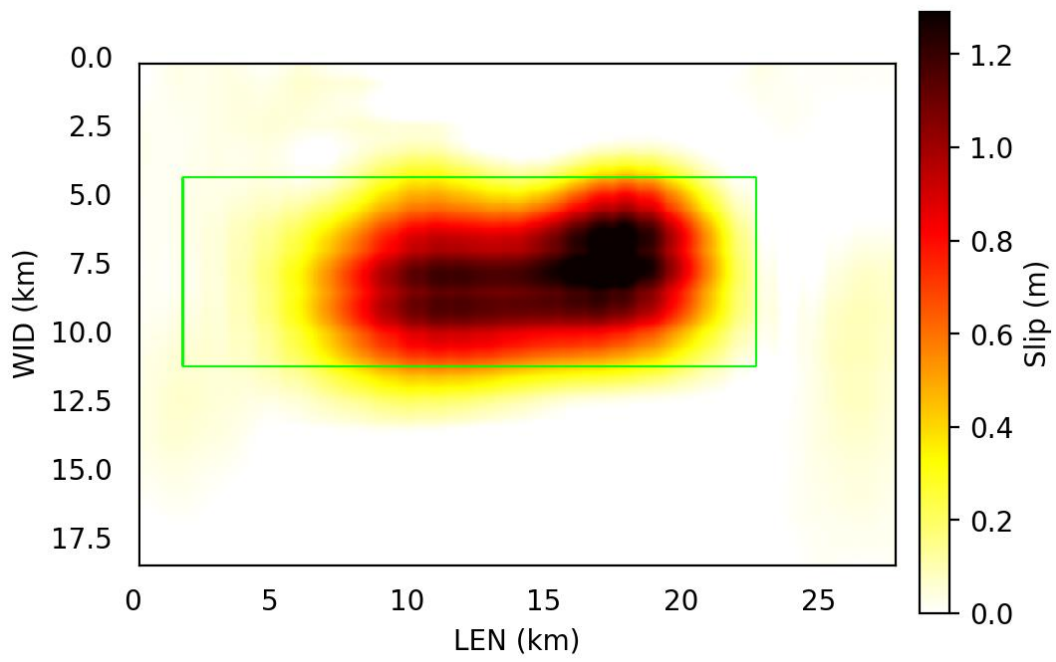
s1998HIDASW10IDEx. Evento del 18/09/1998 Eastern Honshu, Japan, magnitudo 4.41.



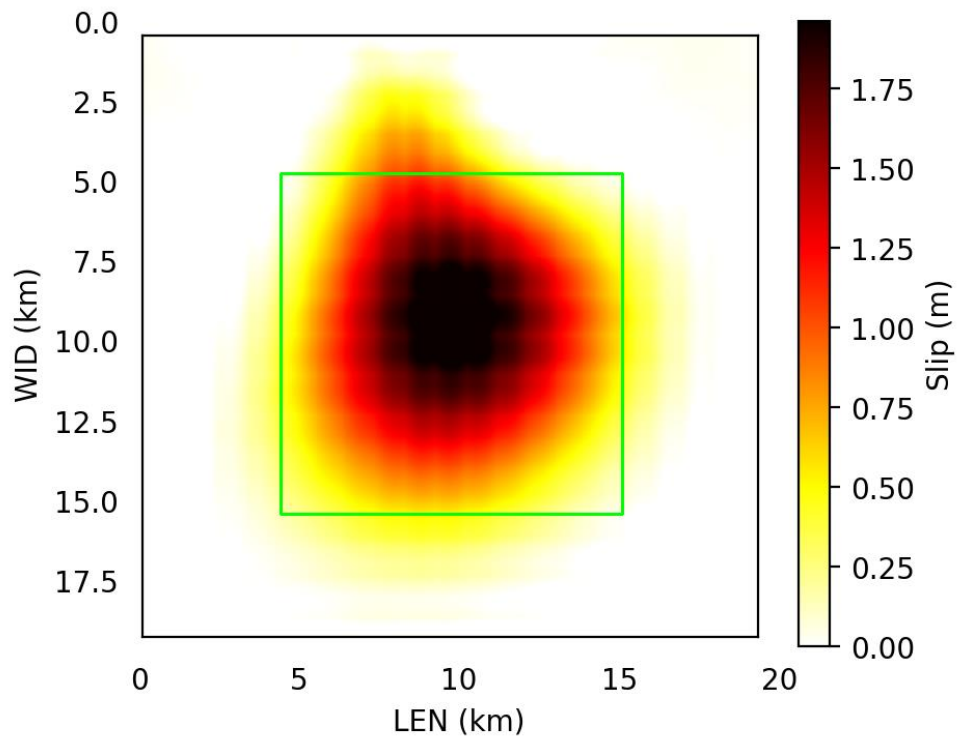
s2005FUKUOK02ASAN. Evento del 20/04/2005, Kyushu, Japan, magnitudo 5.57.



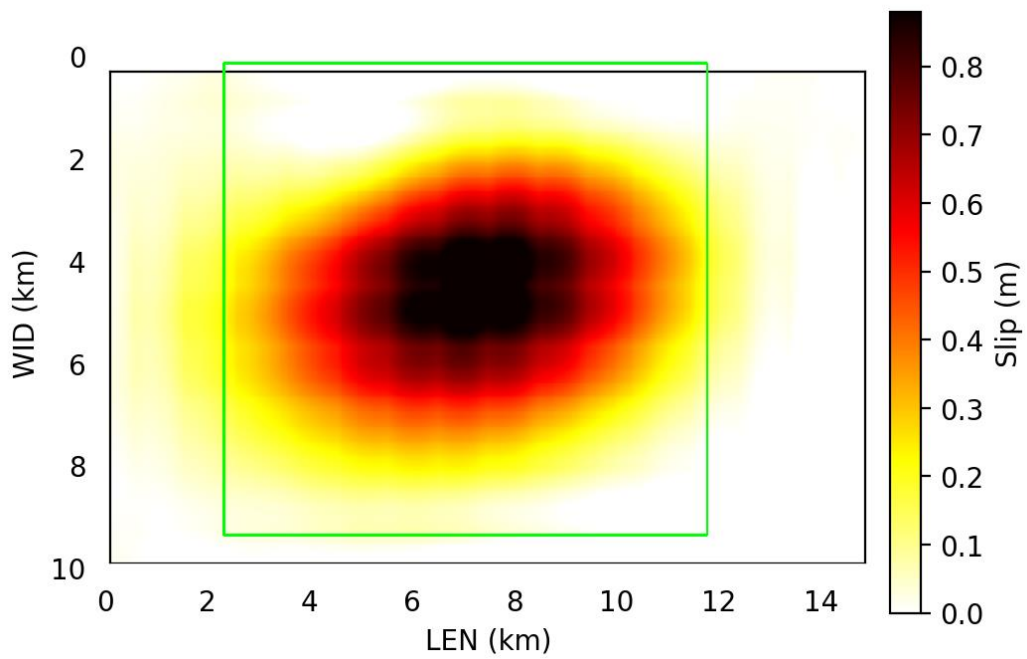
s2005HONSHU01SHAO. Evento del 16/08/2005, Near East Coast Of Honshu, Japan , magnitudo 7.50.



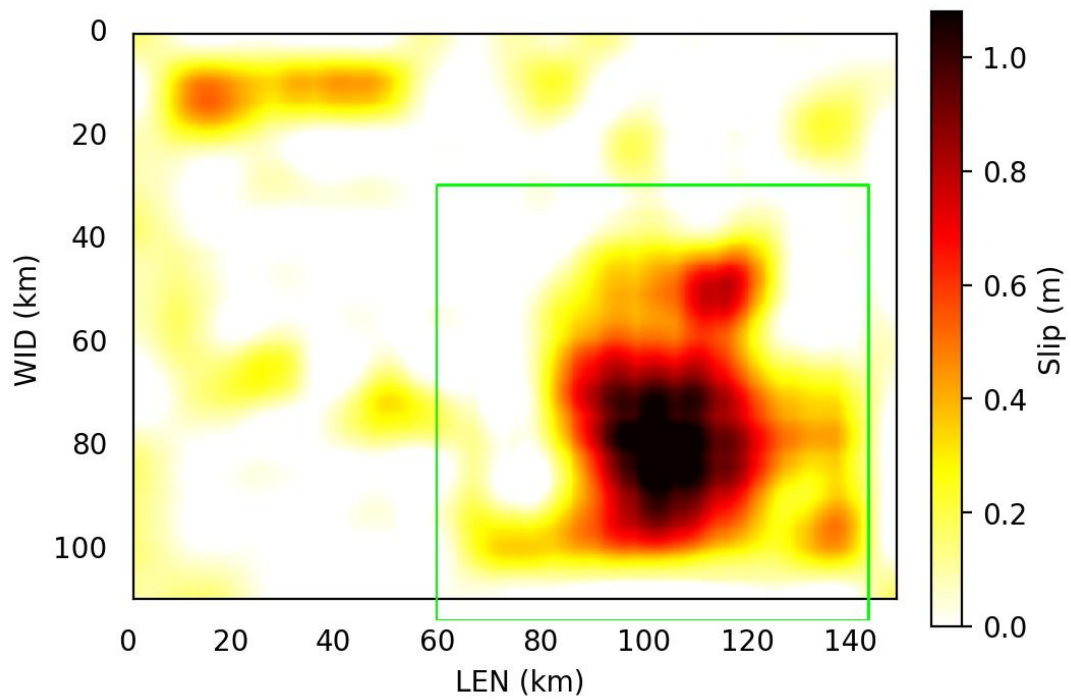
s2005ZHONGB01ELLI, Evento del 07/04/2005, Xizang, magnitudo 6.20



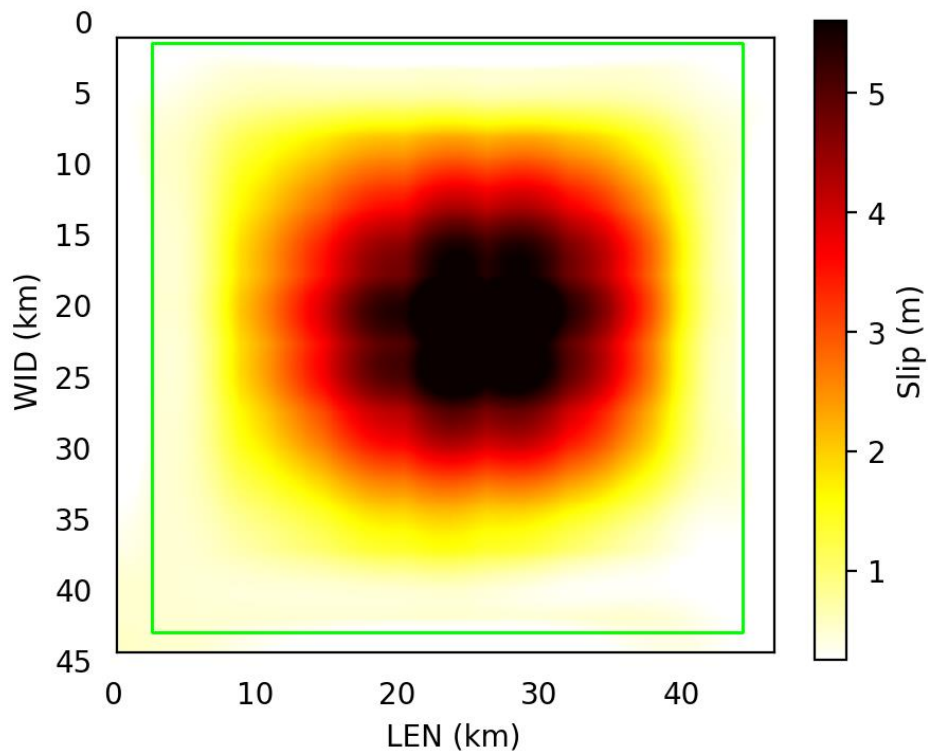
s2008GERZET01ELLI .Evento del 09/01/2008, Xizang, magnitudo 6.40.



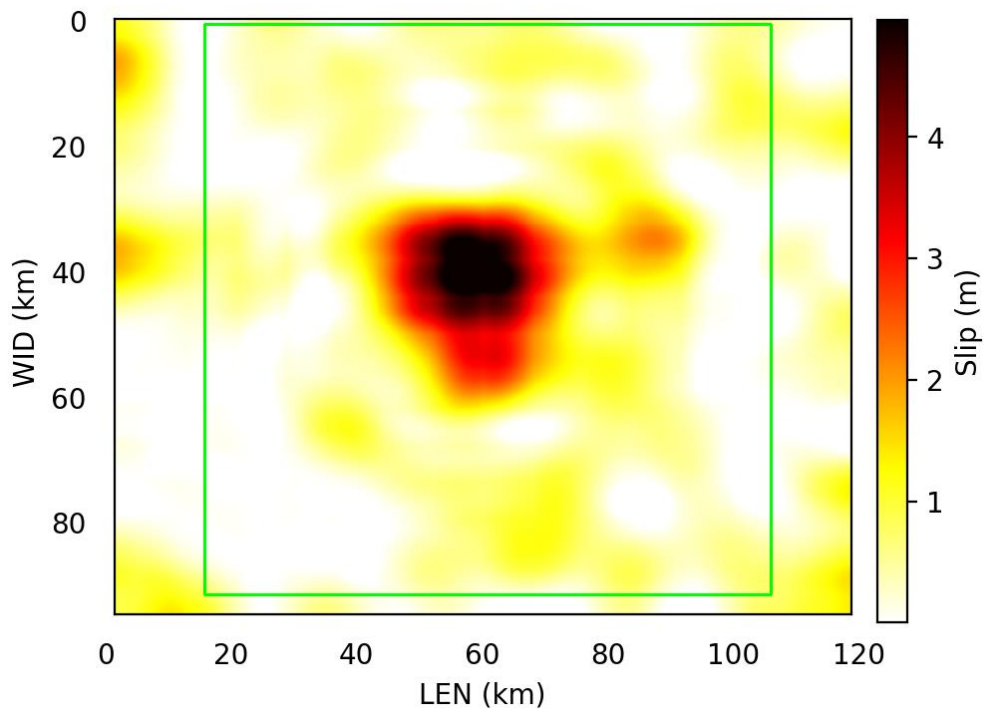
s2008GERZET02ELLI. Evento del 16/01/2008, Xizang, magnitudo 5.90.



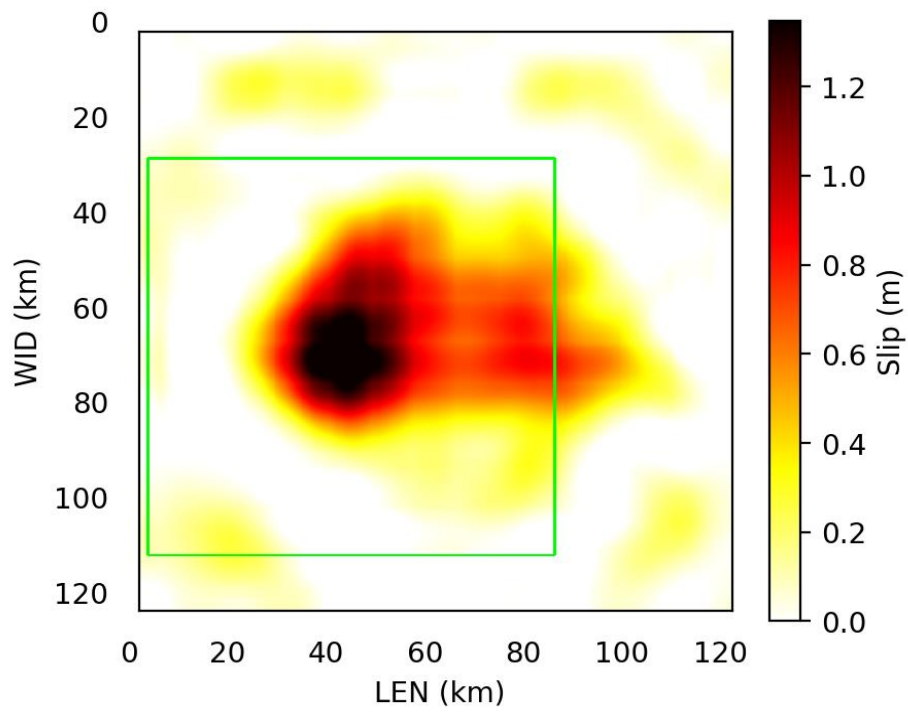
s2008SIMEUL01SLAD. Evento del 20/02/2008, Off W Coast Of Northern Sumatra, magnitudo 7.40.



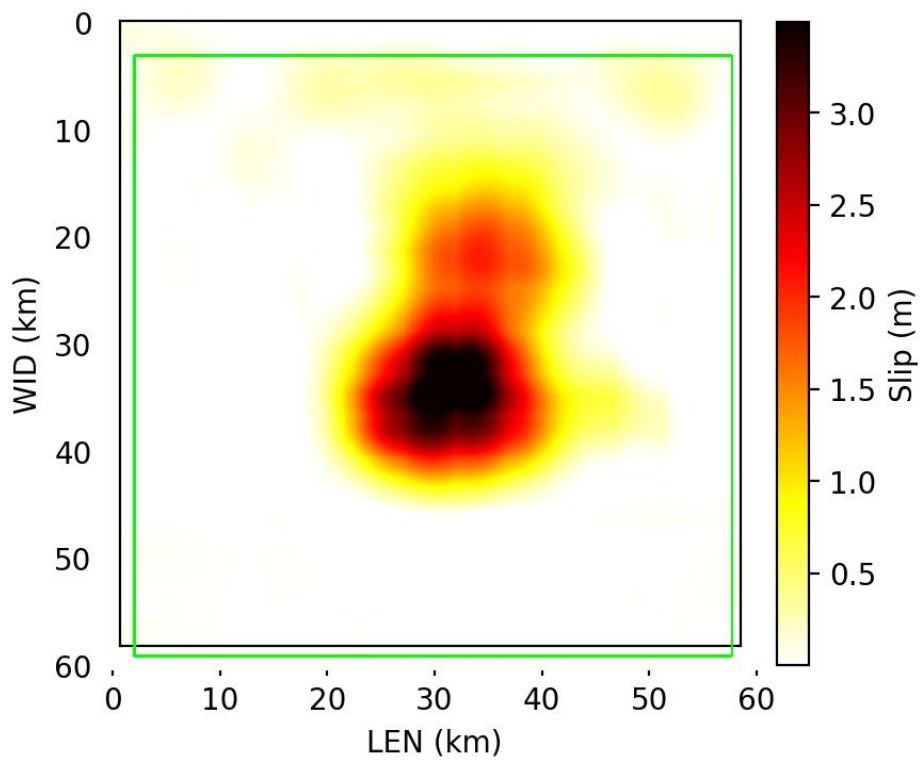
s2009PADANG02SLAD. Evento del 30/09/2009, Southern Sumatra, Indonesia, magnitudo 7.60.



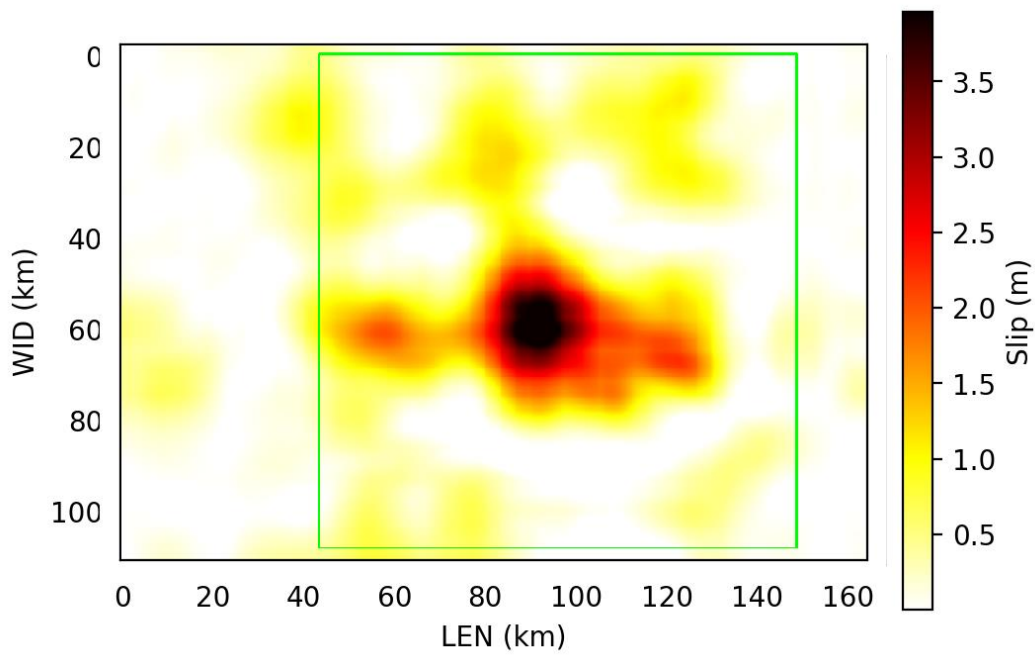
s2009PAPUAx01HAYE. Evento del 03/01/2009, Irian Jaya Region, Indonesia, magnitudo 7.60.



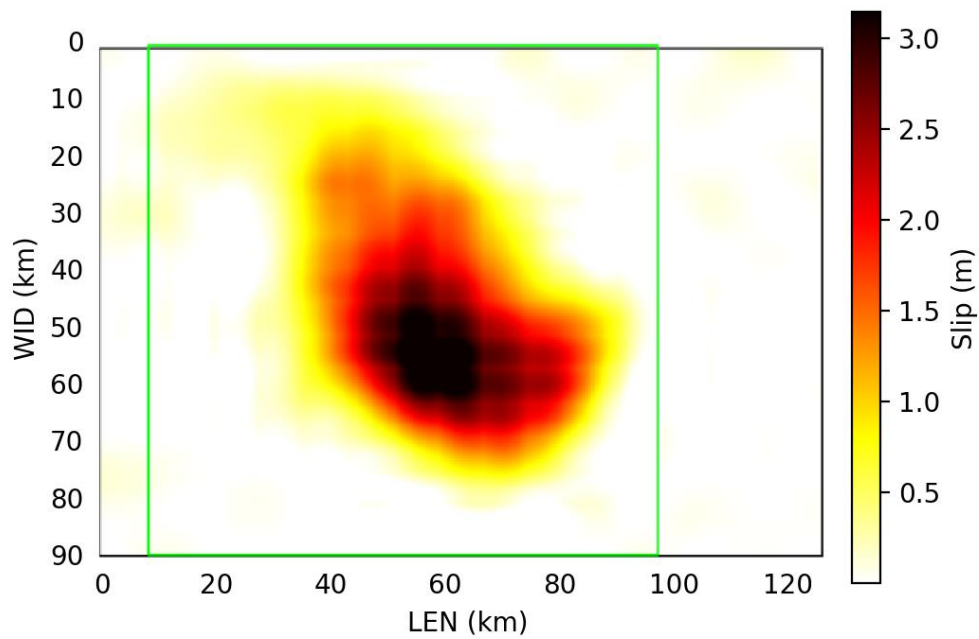
s2011OFFSHO01HAYE. Evento del 09/03/2011, Near East Coast Of Honshu, Japan, magnitudo 7.30.



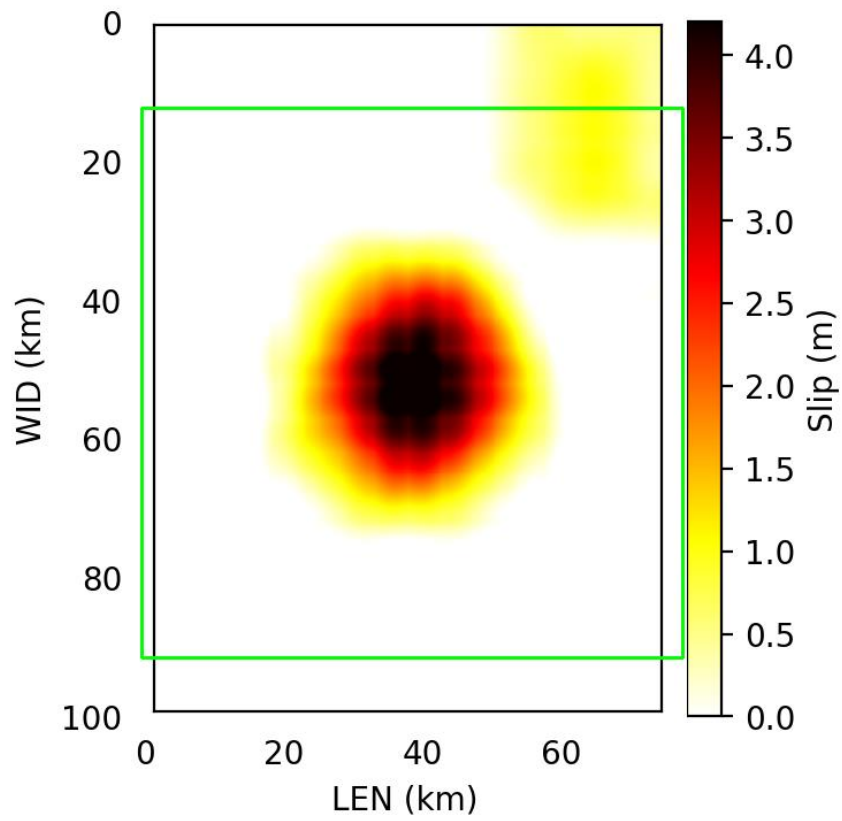
s2011PAKIST01HAYE. Evento del 18/01/2011, Southwestern Pakistan, magnitudo 7.20.



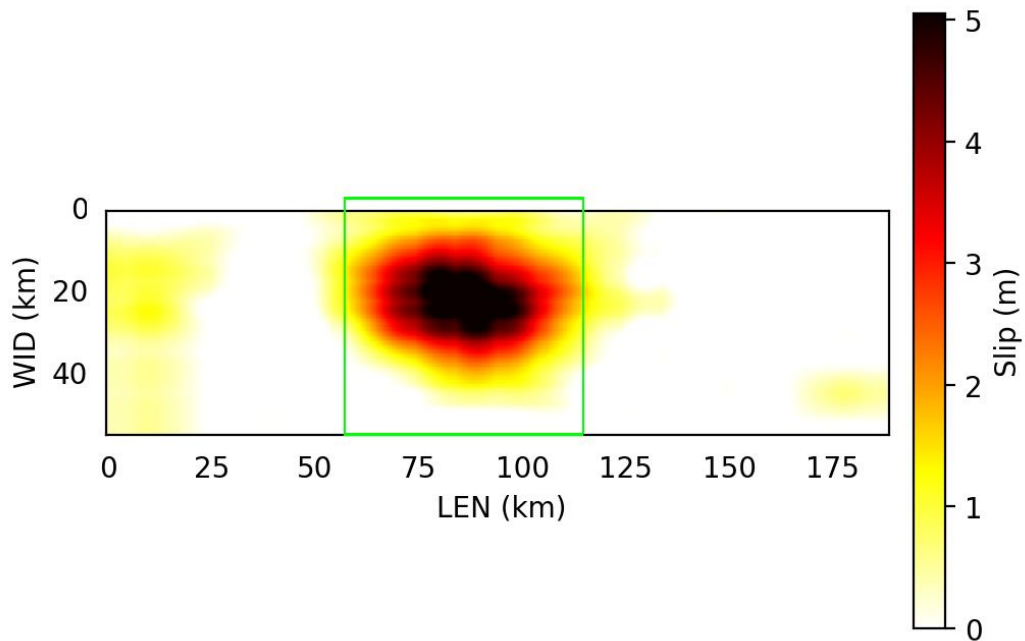
s2012COSTAR01LIUx. Evento del 05/09/2012, Costa Rica, magnitudo 7.57.



s2012EASTOF01HAYE. Evento del 31/08/2012, Philippine Islands Region, magnitudo 7.59



s2012OAXACA01WEI. Evento del 24/05/2013, Sea Of Okhotsk, magnitudo 8.30.



s2013KHASHI01WEI. Evento del 16/04/2013,, Southwestern Pakistan, magnitudo 7.80.

Bibliografia

Aagaard B. T. and Heaton T. H., *Constraining fault constitutive behavior with slip and stress heterogeneity*, *Journal Of Geophysical Research*, VOL. 113, B04301, doi:10.1029/2006JB004793, 2008.

Ajna R. and Hernes T., *Training Haar Cascades*, 2016.

Ammon C. J., et al., *Rupture process of the 2004 Sumatra-Andaman earthquake*. *Science*, 308, 1133-1139, 2005.

Baglione E., Armigliato A., Tinti S.: *Heterogeneous slip distribution on faults responsible for large earthquakes: characterization and implications for tsunami modelling*. European Geosciences Union General Assembly 2017, Vienna, Austria, 23-28 April 2017, abstract n. EGU2017-7929.

Berger W., *Deep Learning Haar Cascade Explained*, 2017.

Bird P., *An updated digital model of plate boundaries*, *Geochemistry*, Published by AGU and the Geochemical Society, Article Volume 4, Number 3, 14 March 2003

Blaser, L., Kruger F., Ohrnberger M. and Scherbaum, F., [2010], *Scaling relations of earthquake source estimates with special focus on subduction environment*, *Bull. Seismol. Soc. Am.* 106, 2914–2926.

Candela T., Renard F., Schmittbuhl J., Bouchon M. and Brodsky E., *Fault slip distribution and fault roughness*, *Geophys. J. Int.* (2011) 187, 959–968.

Creating your own Haar Cascade OpenCV Python Tutorial, pythonprogramming.net.

Docs, OpenCV. *Face Detection Using Haar Cascades*, OpenCV: Face Detection Using Haar Cascades, 2017.

Ehret B., *Pattern recognition of geophysical data*, *Geoderma* 160 (2010)

Fuji Y., Satake K., *Tsunami source of the 2004 Sumatra-Andaman earthquake inferred from tide-gauge and satellite data*. *Bull. Seism. Soc. Am.*, 97, S192-S207, 2007.

Goda et al, *New Scaling Relationships of Earthquake Source Parameters for Stochastic Tsunami Simulation*, Coastal Engineering Journal, Vol. 58, No. 3 (2016).

Goda K., De Risi R., *Multi-hazard loss estimation for shaking and tsunami using stochastic rupture sources*, International Journal of Disaster Risk Reduction 28 (2018) 539–554.

Kong Q., Trugman D.T., Ross Z.E., Bianco M.J., Meade B.J. and Gerstoft P., *Machine Learning in Seismology: Turning Data into Insights*, Seismological Research Letters Volume 90, Number 1 January/February 2019

Lang J.P., *GMT*, 2006-20134.

Lavallée D., Liu P. and Archuleta R.J., *Stochastic model of heterogeneity in earthquake slip spatial distributions*, Geophys. J. Int. (2006) 165, 622–640

Lekshmy P.R. and Raghukanth S.T.G., *Stochastic earthquake source model for ground motion simulation*, Earthq Eng & Eng Vib (2019) 18: 1-34

Leonard M., [2010] *Earthquake fault scaling: Self-consistent relating of rupture length, width, average displacement, and moment release*, Bull. Seismol. Soc. Am. 100, 1971–1988.

Mai P.M. and Thingbaijam K.K.S., *SRCMOD: An Online Database of Finite-Fault Rupture Models*, Seismological Research Letters, November 2014.

Mai P.M. and Beroza G.C., *A spatial random field model to characterize complexity in earthquake slip*, Journal Of Geophysical Research, Vol. 107, No. B11, 2308

Mehner R., *Train your own Opencv Haar Classifier*, 2013.

Minini A., *Gli alberi di decisione*, online personal knowledge base.

Mulia I.E., Gusman A.R. and Satake K., *Alternative to non-linear model for simulating tsunami inundation in real-time*, Geophys. J. Int. (2018) 214, 2002–2013.

Murotani S., Satake K. and Fujii Y. [2013] *Scaling relations of seismic moment, rupture area, average slip, and asperity size for M9 subduction-zone earthquakes*, Geophys. Res. Lett. 40, 5070–5074.

Piatanesi A., Tinti S., Gavagni I., *The slip distribution of the 1992 Nicaragua earthquake from tsunami run-up data.* Geophys. Res. Lett., <https://doi.org/10.1029/95GL03606>

Raya R., *Face detection using opencv and python: a beginner's guide*, 2017.

Rezaei M., *Tutorial- Creating a Cascade of Haar-like Classifiers: Step by Step*, 2013.

Rhie J., Dreger D., Burgmann R., Romanowicz B., *Slip of the 2004 Sumatra–Andaman Earthquake from Joint Inversion of Long-Period Global Seismic Waveforms and GPS Static Offsets.* Bull. Seism. Soc. Am., 97, S115-S127, 2007.

Sandri L., Marzocchi W. and Gasperini P., *Some insights on the occurrence of recent volcanic eruptions of Mount Etna volcano (Sicily, Italy)*, Geophys. J. Int. (2005) 163, 1203–1218.

Satake K., *Linear and nonlinear computations of the 1992 Nicaragua earthquake tsunami.* Pure Appl. Geophys., 144, 455-470, 1995.

Singh G., *Training Haar-Cascade*, 2012.

Strasser F., Arango M. and Bommer J. J., [2010], *Scaling of the source dimensions of interface and intraslab subduction-zone earthquakes with moment magnitude*, Seismol. Res. Lett. 81, 941–950.

Thompson J., *Training Cascade Files*, 2014.

Viola P. and Jones M., *Rapid Object Detection using a Boosted Cascade of Simple Features*, Accepted Conference On Computer Vision And Pattern Recognition, 2001.

Wells D. L. and Coppersmith, K. J. [1994], *New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement*, Bull. Seismol. Soc. Am. 84,974–1002.