An Alternative Pulse Classification Algorithm Based on Multiple Wavelet Analysis

Deniz Ertuncay & Giovanni Costa

University of Trieste
Department of Mathematics and Geosciences
SeisRaM Working Group

deniz.ertuncay@phd.units.it

29/05/2019

Overview

Introduction

- Velocity waveform with a unexpected patterns are called pulse shaped signals.
- Pulse-shape signals are important to analyze since they can create high demands on structures around the period of the pulse (Kalkan & Kunnath 2006) .
- Because of their scarcity, velocity pulses are not taken into account in most of the ground motion prediction equations (GMPE) (Abrahamson et, al. 2016; Boore et al. 2014) .
- Proper identification should give the position and the period of the impulsive signal

Reason

- Forward directivity (Somerville et al. 1997)
- Fling step effect (Mavroeidis & Papageorgiou 2002)
- When rupture velocity is close to the velocity of the bed rock of the site of interest

Indicators

- Signals with long and large amplitudes (Somerville et al. 1997)
- High PGV/PGA ratio (Bray & Rodriguez-Marek 2004)
- Earthquake energy concentrated on one (or a few) pulse(s) (Somerville et al. 1997)
- Unexpectedly high response values at the pulse period on spectral response (Yang & Wang 2012)

Data

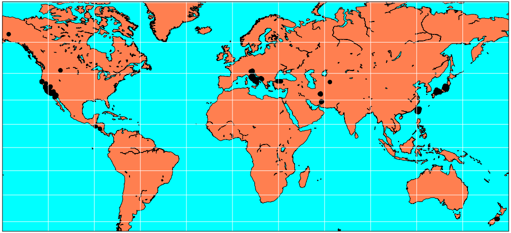


Figure: Stations in the database

2738 earthquake waveform

$5.5 \leq M_w \leq 7.9$

Epicentral distance $\leq$ 115 km

1168 strike slip, 296 normal and 1274 reverse faulting

Wavelet Analysis

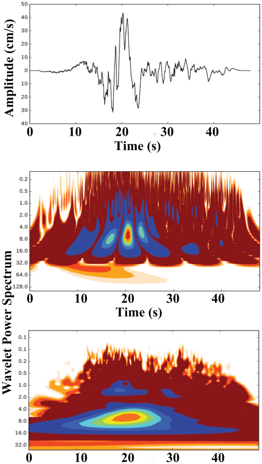


Figure: 1992 Landers earthquake ($M_w = 7.3$), Yermo Fire Station (upper), Ricker wavelet power spectrum (center) and Morlet wavelet spectrum (lower)

- Wavelet analysis package of Torrence & Compo (1998) is used for signal analysis
- Ricker and Morlet Wavelet are implemented to the wavelet analysis process
- The maximum power spectrum values at PGV and the biggest power spectrum value of the signal (if it does not occur at PGV) are used in the pulse identification.

Velocity Pulse at PGV

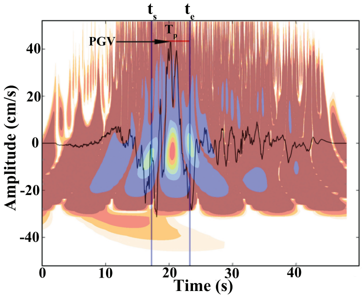


Figure: 1992 Landers earthquake, Yermo Fire Station

1. $PGV \geq 30cm/s$
2. $$\frac{\left(\frac{\int_{t_s}^{te} v^2(\tau)d\tau}{\int_0^\infty v^2(\tau)d\tau} + \frac{\int_{t_s}^{te} WPS(\tau)d\tau}{\int_0^\infty WPS(\tau)d\tau}\right)}{2} \geq 0.30 \qquad (1)$$

Velocity Pulse at Outside of PGV Region

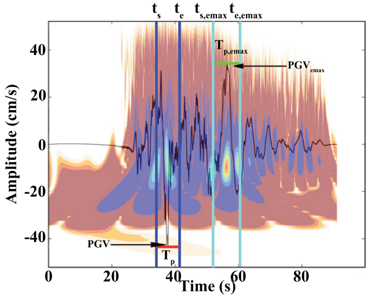


Figure: 1999 Chi-Chi Taiwan Earthquake ($M_w = 7.6$), TCU051 Station

1. $PGV_{E_{max}} \geq 25cm/s$
2. $t_{PGV} - t_{E_{max}} \geq T_p/4.$
3. $$\frac{\int_{ts\,emax}^{te\,emax} v^2(\tau)d\tau}{\int_{ts}^{te} v^2(\tau)d\tau} \geq 1.1 \qquad (2)$$
4. $$\frac{\int_{ts\,emax}^{te\,emax} WPS(\tau)d\tau}{\int_{ts}^{te} WPS(\tau)d\tau} \geq 1.1 \qquad (3)$$
5. $$\frac{\left(\frac{\int_{ts\,emax}^{te\,emax} v^2(\tau)d\tau}{\int_0^\infty v^2(\tau)d\tau} + \frac{\int_{ts\,emax}^{te\,emax} WPS(\tau)d\tau}{\int_0^\infty WPS(\tau)d\tau}\right)}{2} \geq 0.30 \qquad (4)$$

Results

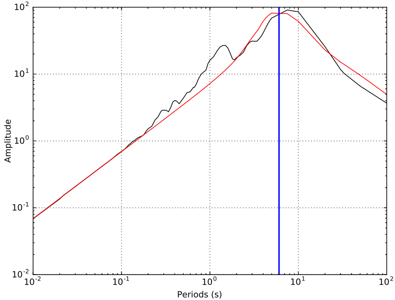


Figure: Pseudo spectral velocity graph of 1992 Landers Earthquake, Yermo Fire Station signal (black) and obtained Ricker wavelet signal (red)

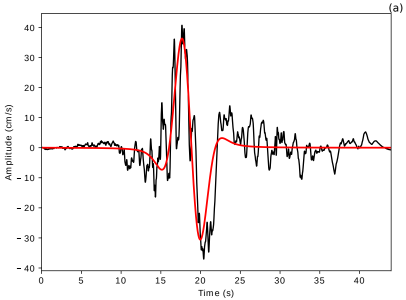


Figure: Waveform (black) and Ricker wavelet (red)

Results

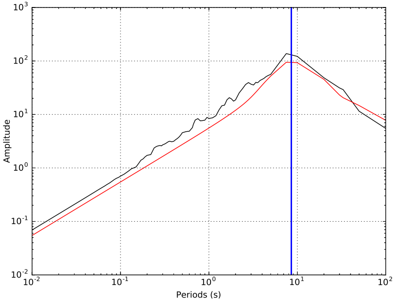


Figure: 1999 Chi-Chi Taiwan Earthquake, TCU051 Station signal (black) and obtained Ricker wavelet signal (red)

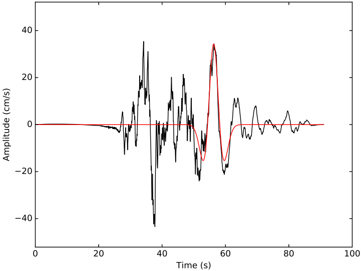


Figure: Waveform (black) and Ricker wavelet (red)

Results

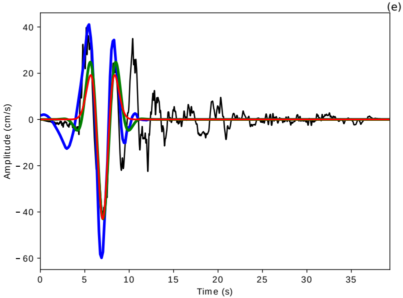


Figure: 1980 Irpinia Earthquake (M_w = 6.9), STN Station signal

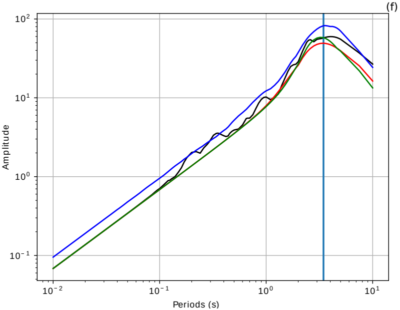


Figure: Spectral Response of waveform (black), Ricker wavelet (red), Chang et al. (2016) (green) and Shahi & Baker (2014) (blue).

Results

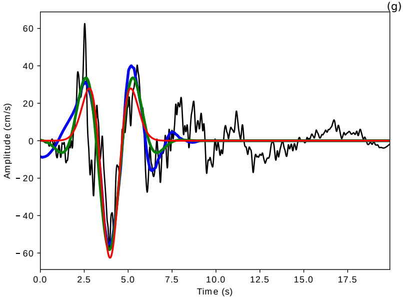


Figure: 1994 Northridge Earthquake ($M_w = 6.7$), Rinaldi Reveiving Station signal

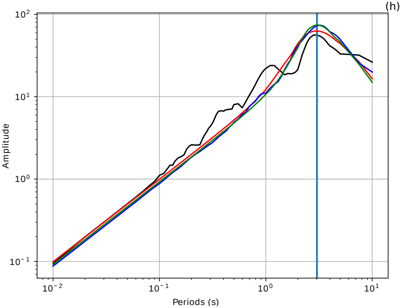


Figure: Spectral Response of waveform (black), Ricker wavelet (red), Chang et al. (2016) (green) and Shahi & Baker (2014) (blue).

Results

	Pulse at PGV	**Pulse at other place**
Ricker Wavelet	290	26
3rd order Morlet Wavelet	4	0
4th order Morlet Wavelet	0	0

	Pulse at PGV	**Pulse at other place**
Strike Slip Fault	111	10
Normal Fault	62	7
Reverse Fault	121	9

Results

Waveform Name	**T_p Shahi & Baker 2014**	**PI**	**T_p Chang et al. 2016**	**E_p**	**T_p**	**Waveform Energy**	**Wavelet Power Spectrum Energy**
TCU078	0	-0.71	1	0.50	3.60	74.15	63.79
Vineyard Canyon 1E	0	-1.63	1	0.50	1.27	47.90	37.02
Brawley Airport	0	-2.4	0	0.29	6.05	57.91	47.53
D08C	3.89	1.90	0	0.30	0	29.74	23.41
AQK	2.04	0.69	1.7	0.38	0	34.76	24.33
Pacoima Dam	0.78	7.69	0.7	0.38	0	39.78	19.19
KJMA	1.09	5.82	1	0.35	0	38.34	19.24
Port Island	2.7	5.94	2.1	0.39	0	32.35	18.21

Conclusion

- Ricker wavelet analysis gives a higher resolution in the time domain, which is more suitable for determining the exact timing of the pulse.
- A Ricker wavelet is better than Morlet wavelets for mimicking the pulse part of the earthquake signal based on residual analysis.
- Our method is reproducing the spectral periods of the pulses, which makes the method convincing.
- Most of the velocity pulses occurred at PGV. However, it is worth mentioning that pulses may occur also in other intervals of the signal.
- This study has correlated with previous studies while expanding the information about the pulse shaped signal such as determining the pulse that occurred other than PGV region.

Identification of Near Fault Pulse Shaped Signals With Machine Learning Algorithms

Deniz Ertuncay[1], Andrea De Lorenzo[2], Giovanni Costa[1], Eric Medvet[2]

University of Trieste Department of Mathematics and Geosciences SeisRaM Working Group
Universtiy of Trieste, Department of Engineering and Architecture, Machine Learning Lab

29/05/2019

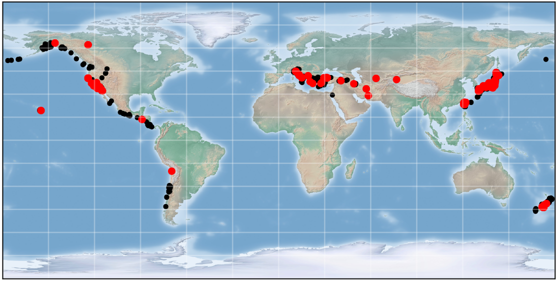
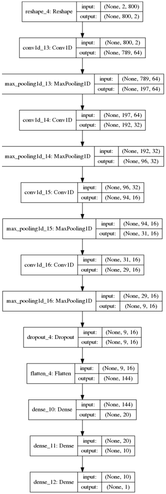


Figure: Spatial distribution of the waveforms.

1. 17581 earthquake waveform from 388 of crustal earthquakes
2. Arrival of the P wave picked manually
3. Signals downsampled to 20Hz
4. 456 impulsive signals found by Shahi and Baker (2014)
5. 407 impulsive signals found by Chang et al. (2016)
6. 458 impulsive signals found by our method
7. 442 impulsive signals detected manually

Data Augmentation

Since the ratio between impulsive (positive) and non impulsive (negative) signals are too low we incremented the number of impulsive signals by generating artificial ones.
For each positive signal we generate as many signals as the ratio between positives and negatives.
Artificial impulsive signals are created adding a zero mean Gaussian noise with 0.1 standard deviation.



Activation function: ReLu (final layers: sigmoid)
Loss function: binary cross-entropy and the learning rate is set accordingly with Adam optimization.
Weights: Glorot normal initializer.
Training Set: %75, **Test Set**: %25
The training divived into two: training and validation.
Training: to effectively train the network, **Validation**: to stop the learning if the loss function start to grow up.
Accordingly with the cross-validation procedure, these steps have been repeated 10 times, varying the portions used as training and testing.

Results

	FPR	FNR
Manual Picking	0.023	0.249
Shahi and Baker (2014)	0.000	0.007
Chang et al. (2016)	0.500	0.008
Our Method	0.000	0.009

Conclusion

- In order to train the model, manually picking is necessary and picking the pulse shaped signals visually may cause disagreement with previous works. It is due to fact that identify the signal as pulse shape is subjective in some cases. In the future we'll use the results of previous studies to create a new model.
- Our method detects non impulsive signals very well but it needs improvement on impulsive signals.

What's Next?

- Adding synthetic waveforms instead of adding Gaussian noises to existing waveforms.
- k^2 method will be used in order to create the synthetic waveforms (planar fault)
- The idea behind is to have full azimuthal coverage around the fault line for several distance intervals.
- ML algorithm will see, in theory, new impulsive signals that we do not have on our databases due to the azimuthal gaps.