# SIMULATION OF NEAR FAULT GROUND MOTIONS USING NEURO-FUZZY NETWORKS AND WAVELET ANALYSIS

Saman EFTEKHAR ARDABILI
*MSc. Student, Department of Civil Engineering, Ahar Branch, Islamic Azad University, Ahar, Iran*
*Eftekhare.Ardabili@Gmail.com*

Amin GHOLIZAD
*Assistant Professor, University of Mohaghegh Ardabili, Ardabil, Iran*
*Gholizad@uma.ac.ir*

**Keywords:** ANFIS, Artificial Records, Near-fault, Pulse-like, Wavelet Packet Transform

## ABSTRACT

The existence of recorded accelerograms to perform dynamic inelastic time history analysis is of the utmost importance especially in near-fault regions where directivity pulses, known as the most important characteristics of these ground motions, impose extreme demands on structures and cause widespread damages. But due to the lack of recorded acceleration time histories, it is common to generate proper artificial ground motions. In this study, in order to generate near-fault pulse-like ground motions, first, it is proposed to extract velocity pulses from an ensemble of near-fault pulse-like ground motions and then simulate nonpulse-type ground motion using Adaptive Neuro-Fuzzy Inference Systems (ANFIS) and wavelet packet transform (WPT). In the next step, the pulse-like ground motion is produced by superimposing directivity pulse on the previously generated nonpulse-type motion in a way that it is compatible with an specified near-field spectrum. Particle swarm optimization (PSO) is employed to optimize both the parameters of pulse model and cluster radius in subtractive clustering and principle component analysis (PCA) is used to reduce the dimension of ANFIS input vectors. Finally, a number of interpretive examples are presented to show how the proposed method works.

## INTRODUCTION

Near-fault ground motions have different characteristics from those of far-fault ground motions. Forward-directivity pulse and permanent displacement so-called "fling step", are the most important ones which should be considered during designing and analyzing the response of structures located near the source. The high-amplitude, long-period velocity pulses are produced by the forward-directivity effects resulting from the pattern of fault dislocation. when fault rapture propagates toward the site with a velocity that is almost equal to shear wave velocity and the direction of fault slip is aligned with the site, this shows itself in the form of velocity pulse in the velocity time history (Somerville et al., 1997). Forward-directivity pulses and fling steps occur in fault normal and fault parallel directions, respectively. The forward-directivity pulses are just considered here for the aims of this study. Imposing extreme demands (such as higher base shears, inter-story drifts and roof displacements) on structures by pulse-like ground motions on the one hand and the lack of recorded near-source acceleration time histories plus the importance of existence of such records in order to perform dynamic inelastic time history analysis on the other, provide researchers with an extra incentive to investigate and present methods in order to generate proper near-fault pulse-like ground motions.

After the 1994 Northridge, California, earthquake, most of the engineers and seismologists were sensible of special effects of pulse-like ground motions on structural damages and started studying characteristics, structural responses of these records and tried to model forward-directivity pulses and simulate pulse-like records. Mavroeidis and Papageorgiou (2003) used Gabor wavelet obtained through multiplying a harmonic oscillator by a bell-shaped function to model pulses and then generated pulse-like ground motions via combining synthetic high-frequency component with the generated long-period pulse. Li and Zhu (2004) presented an equivalent pulse model with pulse period, pulse intensity, number of half-period cycles and contribution ratio as its parameters. They concluded that the pulse period is not the same as the predominant period in the velocity response spectrum, but their ratio tends to remain constant. Tian et al. (2007) used a simple continuous function to simulate pulse-like velocity time history where their equivalent model includes 5 parameters in which two of them refer to pulse period and peak velocity and the rest represent the shape of the pulse. Fan and Dong (2008) generated near-fault pulse-like ground motion by combining real or artificial far-fault nonpulse-type ground motion, filtered using time-frequency filter, with equivalent pulse where the generated motion could reflect the local characteristics of site and the pulse-like characteristics of near-fault ground motion.

Baker (2007) used self-similarity revealing capability of the wavelet analysis to extract velocity pulses from velocity time histories and then developed a quantitative criterion for classifying a ground motion as "pulse-like". Mukhopadhyay and Gupta (2013) used smoothening technique to extract directivity pulses from accelerograms directly and then represented "pulse index" based on the value of maximum fractional signal energy contribution by any half-cycle of the velocity time history for identifying pulse-like records. They also proposed using Mexican Hat function as the equivalent pulse models.

Ghodrati et al. (2012) used residual nonpulse-type records after pulse extraction by Baker (2007) method for training PSO-based neural networks and combined the produced artificial nonpulse-type ground motion with mathematical pulse model proposed by Mavroeidis and Papageorgiou (2003) to obtain pulse-like ground motions. However, comparing Figure 1a and b reveals that extracting pulses using Baker's method via subtracting wavelets repetitively makes the residual ground motion lose more information than just pulse itself, only because of the wavelet shape. Multiple pulses are also treated the same as single ones in this method. Moreover, the mathematical model of Mavroeidis and Papageorgiou (2003) doesn't matches the extracted velocity pulses precisely. Therefore, in this study, it is proposed to use Mukhopadhyay and Gupta's (2013) pulse extraction technique and their proposed pulse model to generate artificial near-fault pulse-like ground motion.

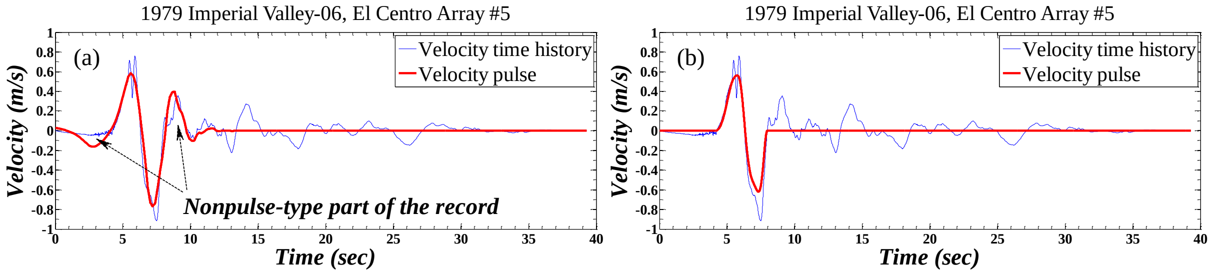


Figure 1. Extracted velocity pulse: (a) Baker's method (2007), (b) Mukhopadhyay and Gupta's method (2013).

In this study, a novel algorithm is presented in order to generate artificial pulse-like ground motion which its response spectrum is compatible with a near-field target spectrum. Adaptive Neuro-Fuzzy Inference System (ANFIS), Wavelet Packet Transform (WPT), Discrete Wavelet Transform (DWT), Particle Swarm Optimization (PSO) and Principal Component Analysis (PCA) are used to achieve the desired goal. Smoothening method of pulse extraction is used here to extract directivity pulses. After pulse extraction, the residual ground motions are used to train ANFIS networks. In fact, ANFIS has been used to generate the high-frequency components of the nonpulse-type ground motions and the equivalent pulse model has been adopted to replicate the intermediate- to long-period directivity pulses of the near-field ground motions.

## PULSE EXTRACTION

After studying different methods of pulse extraction, it is proposed to use the one proposed by

Mukhopadhyay and Gupta (2013). In this method, pulses are categorized into three groups: pulses Type 1 with a large half-cycle in the middle and two small adjacent half-cycles, pulses Type 2 with two comparable half-cycles, and multiple pulses. The extraction process consists of 3 main steps: (i) determination of pulse-time window, (ii) smoothening acceleration time history in order to exclude the incoherent high-frequency part of the accelerogram and identify long-period directivity pulse through the equation $y_i = 1/4\ x_{i-1} + 1/2\ x_i + 1/4\ x_{i+1}$, and third applying adjustments which include changing both the first and the last sharp-varying part of the pulse before and after the first and the last peak\trough to slow- or linear-varying one, and correcting the baseline, because the velocity and displacement pulses don't reach zero at the last instance of the pulse. Here, for the baseline correction, polynomial fits of zero and first order are performed to the first and third part of the entire displacement pulse signal before and after pulse-time window, respectively. Then, baseline is corrected using spline fit for the second part of the displacement pulse inside the window. After extracting the first pulses, the same procedure is conducted again on the residual record to have the second pulse extracted if possible. In the case of multiple pulses, the first and second or rarely third pulses of Type 1 or 2 are extracted from the record. The extracted velocity pulses of multiple-type is shown in Figure 2.

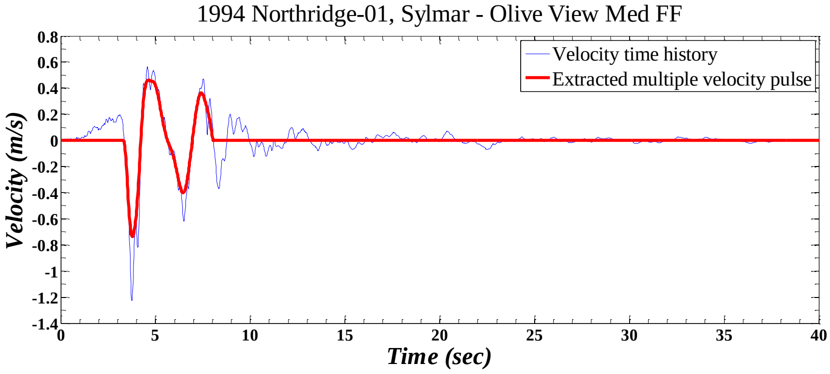


Figure 2. Extracted multiple velocity pulse.

## WAVELET ANALYSIS

The low frequency component forms the most important segment of many signals, so decomposing a signal into its frequency components is counted as the most important application of signal analysis tools. The discrete wavelet transform (DWT) and wavelet packet transform (WPT) are of those kinds of tools that provide such possibility where signal is decomposed into two low- and high-frequency components called approximation and detail signals, respectively. Wavelet decomposition tree for both methods, including three levels of decomposition, is shown in Figure 3 where each approximation coefficient is split into its approximation and detail coefficients in the next level in DWT. However, details as well as approximations are decomposed into their approximation and detail coefficients at each level in WPT. Both methods are reversible, that is, it is possible to reconstruct the original signal from its coefficients via their inverse transforms IDWT and IWPT. Each of the detail coefficients cover certain frequency range in DWT.

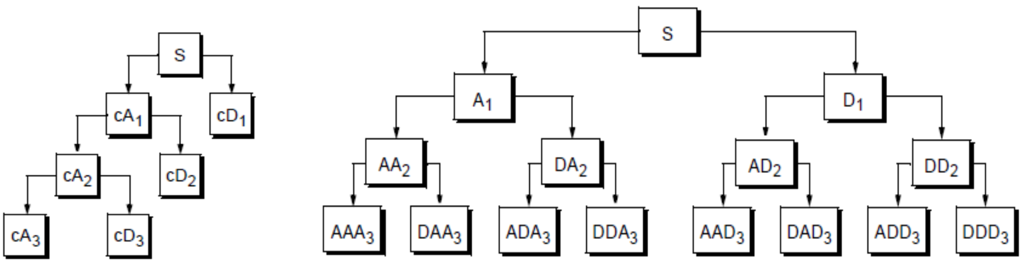


Figure 3. Left: DW decomposition tree, Right: WP decomposition tree (MATLAB, 2002).

## NEURO-FUZZY NETWORKS

The solution of fuzzy logic (FL) or fuzzy inference systems for decision-making problems is based on membership functions and fuzzy if-then rules. There are two types of fuzzy inference systems: (i) Mamdani, (ii) Sugeno. These systems vary somewhat in the way outputs are determined (MATLAB, 2014). In this study Sugeno or Takagi-Sugeno (TSK) model is used. "If *Input1* = $x$ and *Input2* = $y$, then *Output* is $z = ax + by + c$" is a typical rule in a Sugeno fuzzy model in which $x$ and $y$ are linguistic variables or fuzzy sets or membership functions and $z$ is the output membership function in Sugeno systems which can be either linear or constant. In this paper, Gaussian membership function has been used as the input membership function and is defined as follows:

$$\mu_{A_i}(x) = \exp\left\{-\left(\frac{x - c_i}{\sigma_i}\right)^2\right\} \tag{1}$$

Where $c_i$ is cluster center and $\sigma_i$ is standard deviation. Linear membership function for output variable is used in this study. Subtractive clustering is also employed to determine both fuzzy if-then rules and membership functions parameters. Combination of fuzzy logic and neurocomputing, leading to neuro-fuzzy systems, has highest visibility in soft computing and plays important role in the induction of rules from observations. An effective method developed by Dr. Roger Jang for this purpose is called Adaptive Neuro-Fuzzy Inference System (ANFIS). It's easy to create a fuzzy system to match any set of input-output data via ANFIS. A hybrid method consisting of backpropagation algorithm and least squares estimation is used here to tune parameters of the input and the output membership functions, respectively. The subtractive clustering algorithm provides fuzzy inference system (FIS) with minimum number of rules and determines parameters of membership functions of FIS. Here, PSO is used to optimize the neighborhood radius ($r_a$) in subtractive clustering. A range between 0.15 and 0.3 is considered for $r_a$'s variations in ANFIS networks.

## PROPOSED METHOD

Response spectra of near-fault ground motions have an amplification in pulse period region and it's not caught by the Boore and Atkinson (2008) median prediction model, however, their attenuation model can predict the residual ground motion spectra decently (Baker, 2008). It can be shown that the response spectra of the original pulse-like records are in good agreement with the near-fault prediction model proposed by Rupakhety et al. (2011) and the so called narrow-band amplification region is well described by this model. The pseudo-acceleration response spectra of pulse-like record, residual, Boore and Atkinson (2008) model and Rupakhety et al. (2011) model for the record with corresponding pulse of multiple-type, are shown in Figure 4.

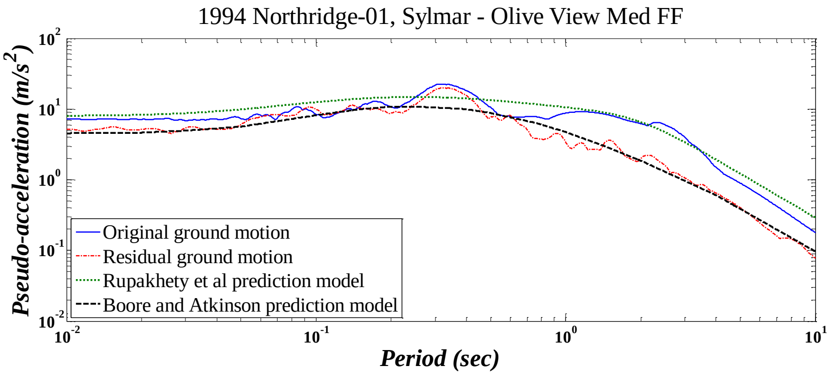


Figure 4. Response spectra of multiple velocity pulse.

In this study, first, all directivity pulses of three types are extracted from an ensemble of near-fault ground motions and considering that residual ground motions are compatible with Boore and Atkinson (2008) prediction model, they are used to train ANFIS networks in order to simulate spectrum compatible nonpulse-type ground motion. Eventually, pulse-type ground motion is obtained by superimposing previously generated high-frequency nonpulse-type component with long-period directivity pulse model. The directivity pulse model based on Mexican Hat function is employed here as the long-period component of near-fault ground motions due to its resemblance with the extracted pulse of Type 1 (Mukhopadhyay and Gupta, 2013):

$$v_{MH}(t) = A\left(1 - \frac{t^2}{\sigma^2}\right)e^{-\frac{t^2}{2\sigma^2}} \tag{2}$$

where $A$ is amplitude of the function, and $\sigma$ has a relationship with dominant period of pulse via the following equation:

$$\sigma = 0.2220 T_{v,MH} \tag{3}$$

PSO is applied to optimize the parameters of pulse models so that the ultimate synthetic record is in good agreement with the Rupakhety et al. (2011) near-field attenuation model in the narrow-band amplification region specifically.

## INTERPRATIVE EXAMPLE

To evaluate the performance of the proposed method, 23 records among 91 near-fault records studied by Baker (2007) are chosen according to the site soil conditions and also their significant duration. All the records were rotated into the fault-normal orientation prior to any other pre-processing. All the records have $175 \leq V_{s30} \leq 375$ meter per second, that is, they are recorded in a site of type 3. Pulses of all accelerograms are extracted and significant duration of accelerograms after pulse extraction is smaller than 20 seconds. All accelerograms are discretized at 0.01 second. The peak ground acceleration (PGA) of all residual accelerograms are scaled to 1g. To make all residuals have equal durations of 30 seconds, first, significiant duration of all is selected using Terifunac and Brady (1975) method, pieces of original record are then added to this extracted duration to the extent that they are set to have 30 seconds total duration and shifted in a way that PGA of all occur at the same time (here in 8 seconds) due to efficient and convinient training of the ANFIS networks. A series of zeros are added to the records for which their total durations are shorter than the specified duration. The time interval between the five percent and the ninety-five percent of the acceleration cumulative energy, the integral of the square of acceleration, is defined as significant duration here. The pseudo-velocity response spectra of all accelerograms are calculated by solving the single degree of freedom equation for earthquaeke ground motion using linear interpolation method at 1000 equally spaced points of periods between 0.01-10 sec, in logarithmic scale.

Calculating pseudo-velocity response spectrum at 1000 discrete frequencies as the input of the ANFIS networks, we are dealing with a thousand-dimensional problem so that PCA, a data compression tool, is used to reduce the input space dimension. To this end, just 22 eigenvectors corresponding to 22 largest eigenvalues are chosen providing a reasonably close approximation, therefore, for the 23 records used to train the ANFIS networks, the compressed space equals:

$$[Y]_{22\times23} = [Q]_{1000\times22} * [X]_{1000\times23} \tag{4}$$

in which $[X]$ includes spectral values in real space for 23 records and thousand frequency points (dimensions), $[Y]$ is the matrix of spectral values in feature/compressed space including 22-dimension, and $[Q]$ is the eigenvectors matrix. Matrix $Y$ is used as the input vectors of the ANFIS.

Then, wavelet packet transform is applied to decompose the residual accelerograms into wavelet packet coefficients. The output layer of a single ANFIS network consists of just one node, so let take the *k*th wavelet packet coefficient in the *i*th level of decomposition and *j*th packet as the output of each ANFIS network:

$$c_j^{\,i}(k)=\int_{-\infty}^{+\infty} a_g(t)\psi_{j,k}^{\,i}(t)dt \tag{5}$$

where $a_g(t)$ is earthquake ground acceleration and $\psi^i_{j,k}(t)$ is the mother wavelet. In this study, Daubechies mother wavelet of order 10 (db10) is used. The accelerograms are transformed into their first level of wavelet packet decomposition coefficients. There are 2 packets (just an approximation and a detail coefficients) at the first level and each packet includes 1509 points, Therefore, 3018 ANFIS networks are trained using PCA coefficients of the response spectra and single points of wavelet packet coefficients as the input and output of the networks, respectively.

Stopping criteria in training networks is to reduce model error for check data considering the training duration of the network. In other words, cluster radius is chosen using PSO so that model mean square error meets its lowest value as evaluating the check data. This will result in selecting the most appropriate cluster radius for forming an FIS and avoiding overfitting of the network. The check records are the ones for which the least error is occurred before overfitting.

Generally, generation of near-field pulse-like ground motion consists of two parts; the first part is to generate high-frequency nonpulse-type Boore and Atkinson compatible artificial record, and the second is to superimpose long-period directivity pulse. MATLAB software is used for coding each section. After training the networks, providing PCA coefficients of the response spectra as input of the networks, one can obtain wavelet packet coefficients of the artificial record. Then, by applying inverse wavelet packet transform on the coefficients, the artificial record is obtained. To improve the results of the compatibility, the synthetic record is decomposed again using descrete wavelet transform and the detail coefficients are modified iteratively for the *j*th level (Mukherjee & Gupta, 2002a and b), that is:

$$cD_j^{Mod}=cD_j\times\frac{\int_{T_{1j}}^{T_{2j}} PSV(T)_{Tar}\,dt}{\int_{T_{1j}}^{T_{2j}} PSV(T)_{Calc}\,dt} \tag{6}$$

$$T_{1j}=2^{j}\Delta t\,,\; T_{2j}=2^{j+1}\Delta t \tag{7}$$

where $T_{1j}$ and $T_{2j}$ are the period range of detail coefficient in the *j*th level of DWT and $\Delta t$ is the time step of $a_g(t)$. $PSV(T)_{Tar}$ is the target pseudo-velocity response spectrum and $PSV(T)_{Calc}$ is the calculated pseudo-velocity response spectrum of artificial record. The iterative process continues so far obtaining a satisfying compatibility between the target and the synthetic record spectra, then ultimate modified spectrum compatible artificial accelerogram is obtained by applying IDWT. Eventually, final near-fault pulse-like ground motion is produced by superimposing pulse on the previously generated accelerogram in a way that there is a good compatibility between Rupakhety (2011) near-field attenuation spectrum and final generated near-fault pulse-like record.

Accordingly, the efficiency of the trained networks using accelerograms belonging to the train and check data set is validated. Figure 5 and 6 show the test of the networks for a record from the train data set and one from check data set, respectively. A complete compatibility between the spectra and accelerograms of the generated records and the original one can be seen in Figure 5 and a sensible compatibility is concluded for records from check data as shown in Figure 6.

The generated pulse-like record with $M_w=7.3$, $r=6$ km, $V_{s30}=320$ m/s and *fault type=ss*, and associated acceleration and velocity pulses with optimized parameters and its response spectra before and after pulse addition are shown in figure 7.

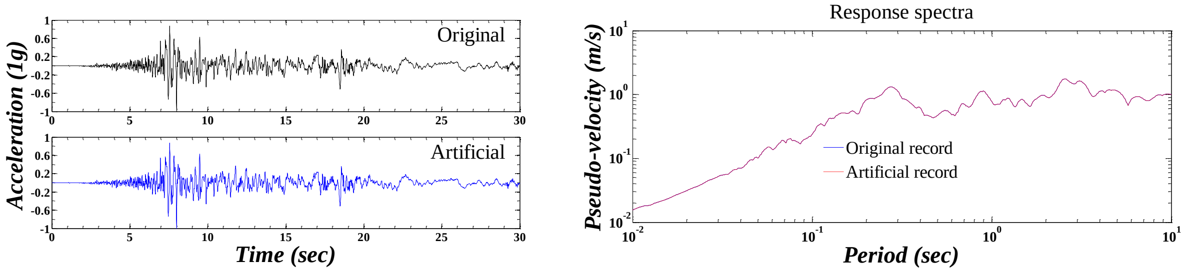


Figure 5. Comparison of original and generated records belong to train set (1979 Imperial Valley-06, Brawley Airport).

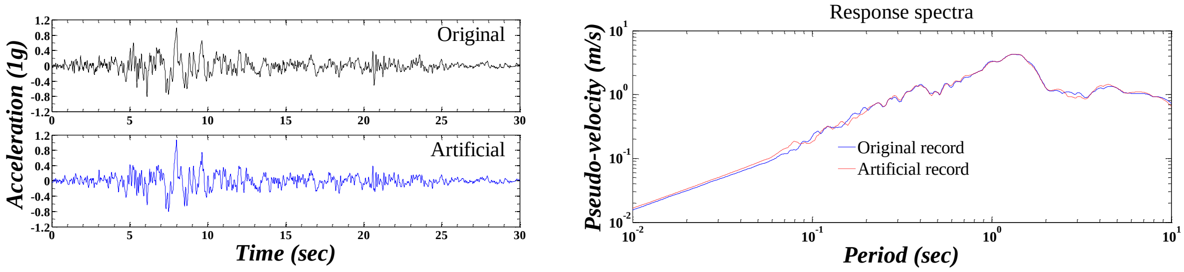


Figure 6. Comparison of original and generated records belong to check data set (1992 Landers, Yermo Fire Station).

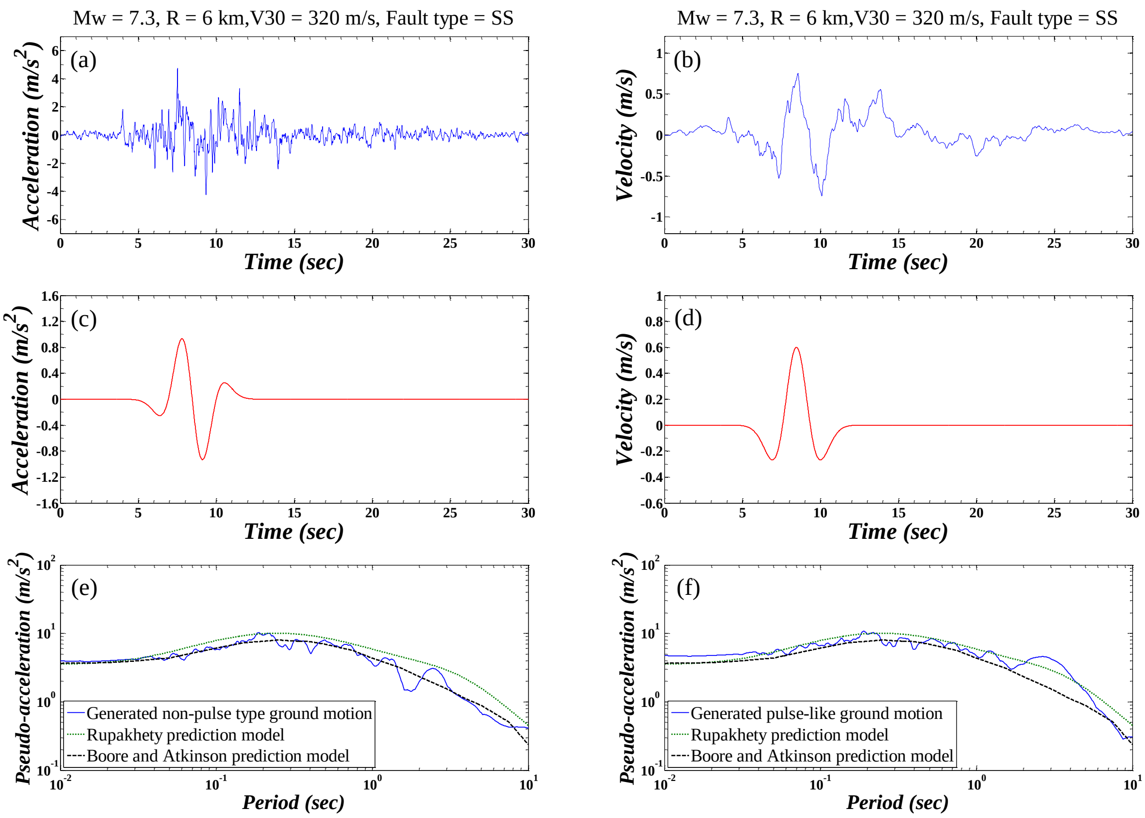


Figure 7. (a) Generated pulse-like acceleration time history, (b) Generated pulse-like velocity time history, (c) Acceleration pulse, (d) Velocity pulse, (e) Response spectra before pulse addition, (e) Response spectra after pulse addition.

## CONCLUSIONS

A new method has been developed based on wavelet analysis, neuro-fuzzy networks, PSO and PCA to generate near-fault pulse-like ground motions. First, directivity pulses were extracted. It was noticed that the Boore & Atkinson prediction model resembles the spectra of the residual records, therefore, first nonpulse-type ground motions are simulated using learning abilities of ANFIS networks and multi-resolution wavelet packet transform to expand the relationship between PCA coefficients of the response spectra and each points of wavelet packet coefficients. An illustrative example was shown in which good results of spectrum compatibility for the generated nonpulse-type records was obtained. At the end, directivity pulse models were used to generate final near-fault pulse-like ground motion which was compatible with Rupakhety near-fault model. Except for the records and their response spectra, nothing else is needed in this method to produce near-fault records.

## REFERENCES


Baker JW (2008) Identification of near-fault velocity pulses and prediction of resulting response spectra, *Geotechnical Earthquake Engineering and Soil Dynamics IV*, Sacramento, CA, p 10

Baker JW (2007) Quantitative classification of near-fault ground motions using wavelet analysis, *Bulletin of the Seismological Society of America*, 97(5): 1486–1501

Boore D and Atkinson G (2008) Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s, *Earthquake Spectra*, 24: 99-138

Fan J and Dong P (2008) Simulation of artificial near-fault ground motions based on the S-transform, *Congress on Image and Signal Processing*, 518-522

Ghodrati A G, Abdolahi Rad A, Aghajari S and Khanmohamadi Hazaveh N (2012) Generation of near-field artificial ground motions compatible with median-predicted spectra using PSO-based neural network and wavelet analysis, *Computer-Aided Civil and Infrastructure Engineering*, 27: 711–730

Li X and Zhu X (2004) Study on equivalent velocity pulse of near-fault ground motions, *ACTA Seismologica Sinica*, 17(6): 697–706

MATLAB (2002) Reference Guide, *the Math Works Inc*

MATLAB (2014) Reference Guide, *the Math Works Inc*

Mavroeidis GP and Papageorgiou AS (2003) A mathematical representation of near-fault ground motions, *Bulletin of the Seismological Society of America*, 93(3): 1099–1131

Mukherjee S and Gupta K (2002) Wavelet-based characterization of design ground motions, *Earthquake Engineering & Structural Dynamics*, 31: 1173-1190

Mukherjee S and Gupta K (2002) Wavelet-based generation of spectrum-compatible time-histories, *Earthquake Engineering & Structural Dynamics*, 22: 799-804

Mukhopadhyay S and Gupta VK (2013) Directivity pulses in near-fault ground motions—I: identification, extraction and modelling, *Soil Dynamics and Earthquake Engineering*, 50: 1-15

Rupakhety R, Sigurdsson SU, Papageorgiou AS and Sigbjörnsson R (2011) Quantification of ground-motion parameters and response spectra in the near-fault region, *Bull Earthquake Eng*, 9: 893-930

Somerville PG, Smith NF, Graves RW and Abrahamson NA (1997) Modification of empirical strong ground motion attenuation relations to include the amplitude and duration effects of rupture directivity, *Seismological Research Letters*, 68(1): 199–222

Tian Y, Yang Q and Lu M (2007) Simulation method of near-fault pulse-type ground motion, *ACTA Seismologica Sinica*, 20(1): 80–87

Trifunac MD and Brady AG (1975) A study of the duration of strong earthquake ground motion, *Bull, Seism, Soc, Am*, 65: 581-626