Geophysical Journal International

Geophys. J. Int. (2010) **182,** 1567–1576 doi: 10.1111/j.1365-246X.2010.04704.x

Refined thresholds for non-linear ground motion and temporal changes of site response associated with medium-size earthquakes

Chunquan Wu,[1] Zhigang Peng[1] and Yehuda Ben-Zion[2]

[1] *School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA* 30332, *USA. E-mail: chunquanwu@gatech.edu*
[2] *Department of Earth Sciences, University of Southern California, Los Angeles, CA* 90089-0740, *USA*

Accepted 2010 June 15. Received 2010 June 15; in original form 2010 February 22

SUMMARY

We systematically analyse non-linear effects and temporal changes of site response associated with medium-size earthquakes, using seismic data recorded by the Japanese Strong Motion Network KIK-Net. We apply a sliding-window spectral ratio technique to surface and borehole strong motion records at six sites, and stack results associated with different earthquakes that produce similar peak ground acceleration (PGA). In some cases we observe a weak coseismic drop in the peak frequency when the PGA is as small as ∼20–30 Gal, and near instantaneous recovery after the passage of the direct *S* waves. The percentage of drop in the peak frequency starts to increase with increasing PGA values. We also observe a coseismic drop in the peak spectral ratio for two sites. When the PGA is larger than ∼60 Gal to more than 100 Gal, we observe considerably stronger drops of the peak frequencies followed by logarithmic recovery with time. The observed weak reductions of peak frequencies with near instantaneous recovery likely reflect non-linear response with essentially fixed level of damage, while the larger drops followed by logarithmic recovery reflect the generation (and then recovery) of additional rock damage. The results indicate clearly that non-linear site response may occur during medium-size earthquakes, and that the PGA threshold for *in situ* non-linear behaviour is lower than the previously thought value of ∼100–200 Gal.

Key words: Elasticity and anelasticity; Earthquake ground motions; Site effects; Wave propagation.

GJI Seismology

1 INTRODUCTION

Non-linear site response is associated with deviations from the linear trend predicted by Hooke's elasticity when the amplitude of ground motion exceeds a certain threshold. This phenomenon is well documented in both laboratory experiments (Ostrovsky & Johnson 2001, and references within) and under *in situ* conditions (Beresnev & Wen 1996, and references within). Generally, non-linear response has a strong correlation with the level of ground motion (e.g. Hartzell 1998; Su *et al.* 1998; Johnson *et al.* 2009) along with rock type and other characteristics of the sites (e.g. Beresnev & Wen 1996; Hartzell 1998; Trifunac *et al.* 1999; Tsuda *et al.* 2006). Quantification of soil non-linearity and subsequent consequences for site response are currently key components of estimating seismic hazard, predicting future ground motions (Frankel *et al.* 2000) and designing geotechnical and structural engineering systems on soils (NEHRP 2003).

Comparisons of the site response during strong and weak ground motions provide one of the most typical ways to identify non-linear site behaviour. Several recent studies have also investigated details of the recovery process from non-linear response by computing the temporal evolution of spectral ratios between a target site and a nearby reference site (e.g. Pavlenko & Irikura 2002; Sawazaki *et al.* 2006; Karabulut & Bouchon 2007; Sawazaki *et al.* 2009; Wu *et al.* 2009a, b; Rubinstein 2010). These studies generally identify strong reductions of peak frequencies (i.e. resonant frequencies) measured in the spectral ratios during the strong ground motions of a nearby large earthquake, followed by logarithmic recoveries to the level before the large event. The timescales of the recoveries observed in these studies are quite different. In some cases, the coseismic changes appear to recover on timescales of minutes or less (e.g. Pavlenko & Irikura 2002; Karabulut & Bouchon 2007), while in other cases the recoveries appear to last for years (e.g. Sawazaki *et al.* 2006). Some of the documented differences may stem from different analysis resolutions, but they may also reflect (at least partially) different types of non-linear behaviour.

Lyakhovsky *et al.* (2009) analysed theoretically the response of a solid governed by a non-linear damage rheology and highlighted the existence of two forms of non-linear behaviour. With fixed existing material damage, the non-linear effects are small and the recovery to linear behaviour with reduction of source amplitude is instantaneous. When the source motion leads to increasing material damage, the non-linear effects are considerably larger and the recovery to linear behaviour after the source ceases to operate is logarithmic with time. Both forms of non-linear response of materials to loading are seen clearly in high-resolution laboratory experiments (Pasqualini *et al.* 2007).

Because non-linear effects tend to be clear particularly under strong shaking, most of the previous observational seismological studies on this topic focused on large earthquakes (typically larger than $M_w = 6$). They found that non-linear site response typically manifests beyond an amplitude threshold of 100–200 Gal or dynamic strain of 10^{-5} to 10^{-4} (e.g. Beresnev & Wen 1996). On the other hand, laboratory studies identified non-linear effects for geomaterials under strains as small as 10^{-8} (TenCate *et al.* 2004). Using repeating earthquakes, Rubinstein & Beroza (2004) found reduction of *S*-wave velocities in the near surface (which could be used as a proxy for resonant frequency) after an $M_w = 5.4$ aftershock of the 1989 $M_w = 6.9$ Loma Prieta earthquake. Recently, Rubinstein (2010) used the spectral ratio method to analyse strong ground motion recordings of 13 earthquakes with magnitudes of 3.7–6.5 (including the 2003 $M_w = 6.5$ San Simeon earthquake, 2004 $M_w = 6.0$ Parkfield earthquake and their aftershocks). He found that at the Turkey Flat site, ∼35 Gal of peak ground acceleration (PGA) produces non-linear site effects, suggesting that non-linear effects could be much more common than previously thought.

In a previous study (Wu *et al.* 2009a), we quantified the effects of the input ground motion on the degree of non-linearity by applying the spectral ratio method to the strong motion data recorded by a pair of surface and borehole stations before, during and after the 2004 $M_w = 6.8$ mid-Niigata earthquake sequence. We found that the coseismic peak frequency reduction and the post-seismic recovery time increase with the peak ground velocity (PGV) beyond ∼5 cm s^{-1}, or PGA beyond ∼100 Gal. We were unable to identify clear coseismic drop and post-seismic recovery with PGA less than 100 Gal, mainly because of averaging effects associated with the employed 10-s sliding window. We noted that by reducing the sliding-window size to smaller values, it may be possible to detect more subtle non-linear site response associated with smaller input ground motions.

In this work, we apply the sliding-window spectral ratio method with a small window size to the strong motion data generated by ∼2000 medium-size earthquakes recorded by six pairs of surface and borehole stations (Fig. 1). These include the data analysed previously by Wu *et al.* (2009a), together with a few sites that have shown non-linear responses by other recent studies (Sawazaki *et al.* 2006, 2009; Assimaki *et al.* 2008). Because we have more than 2000 seismic records, we are able to stack the spectral ratios for corresponding PGA values, which result in stable measurements. Our analysis shows that the peak frequency starts to decrease at PGA levels of several tens of Gal, followed in some cases by a near instantaneous recovery. When the PGA values exceed ∼60 Gal to more than 100 Gal, the onset of non-linear response is followed by a gradual recovery. The two different observed forms of non-linear behaviour for different ranges of excitation levels are consistent with the theoretical expectations based on the damage model of Lyakhovsky *et al.* (2009) and the laboratory observations of Pasqualini *et al.* (2007). The low thresholds for the onset of *in situ* non-linear effects, documented in this work and the study of Rubinstein (2010), imply that non-linear models should be considered for smaller earthquakes than previously thought (e.g. Kramer & Paulsen 2004).

2 DATA AND ANALYSIS PROCEDURE

2.1 Seismic data

The analysis employs strong motion data recorded by six stations (MYGH04, NIGH06, NIGH12, SMNH01, IWTH04 and IWTH05) in the Japanese Digital Strong-Motion Seismograph Network KiK-Net operated by the National Research Institute for Earth Science and Disaster Prevention (Aoi *et al.* 2000). The network consists of 659 stations with an uphole/downhole pair of strong-motion seismometers. Each KiK-Net unit consists of three-component accelerometers and a data logger having a 24-bit analogue-to-digital converter with a sampling frequency of 200 Hz.

The six stations used in this study are chosen mainly because previous studies have identified clear non-linear effects at these stations during large events (Sawazaki *et al.* 2006; Assimaki *et al.* 2008; Sawazaki et al. 2009;Wu *et al.* 2009a). We also find that the observed temporal changes in peak frequencies and peak spectral ratios (maximum of the spectral ratios) at these stations are much clearer than those at other stations, allowing us to better detect potential temporal changes associated with relatively small ground motions. Additional details on the network and site conditions can be found in Table S1, Fig. S1 (see the Supporting Information section for both), or at the KiK-Net website (http://www.kik.bosai.go.jp/kik/index_en.shtml).

In the subsequent analysis we utilize a total of 2204 events that occurred between 1999 January and 2008 May and recorded by the six surface and borehole strong motion sensors (see also Table S2). The magnitudes of most events range from 3 to 5, and the hypocentral depths range from 5 to 70 km. The recorded PGAs of most events range from 0 to 100 Gal, generally less than the 100–200 Gal threshold for non-linear site effects suggested by previous studies (e.g. Chin & Aki 1991; Beresnev & Wen 1996).

2.2 Analysis procedure

The analysis procedure follows overall those of Sawazaki *et al.* (2006, 2009) and Wu *et al.* (2009a). The main analysis steps are as follows: we use 6-s time windows that are moved forward by 2 s for all waveforms recorded by the surface and borehole stations. The zero time corresponds to the hand-picked *S*-wave arrival, and we use the time corresponding to starting point of each window to denote the time of that window. In such case, the window corresponding to the *S*-arrival window (i.e. zero time) fully captures the strongest motion during the *S*-wave. We have tested various window lengths and sliding values. Our testing results indicate, as noted by Wu *et al.* (2009a), that increasing window length would miss the temporal changes over small timescales. On the other hand, decreasing window length would reduce the stability of the results. The employed 6-s window appears to be a well-balanced value for the examined data between temporal resolution and stability. All possible seismic phases, including pre-event noise, *P*, *S* and coda waves are analysed together. Next, we remove the mean value of the traces and apply a 5 per cent Hanning taper to both ends. We compute the Fourier power spectra of the two horizontal components and take the square root of the sum to get the amplitude of the vector sum of the two horizontal spectra. The obtained spectra are smoothed by applying the mean smoothing algorithm from the subroutine 'smooth' in the Seismic Analysis Code (Goldstein *et al.* 2003), with half width of 0.5 Hz. We tested different smoothing window sizes, and found that smoothing window with half width of 0.5 Hz could remove most of the noisy spikes while not changing the overall shape of spectra significantly. The spectral ratios are obtained by dividing the combined horizontal spectra of surface stations by the spectra of the borehole stations.

Fig. 2 shows an example of the original acceleration records at NIGH06 generated by an *M* 5.3 event on 2004 October 25 and the 6-s windows used to compute the spectral ratios for the direct *S* and

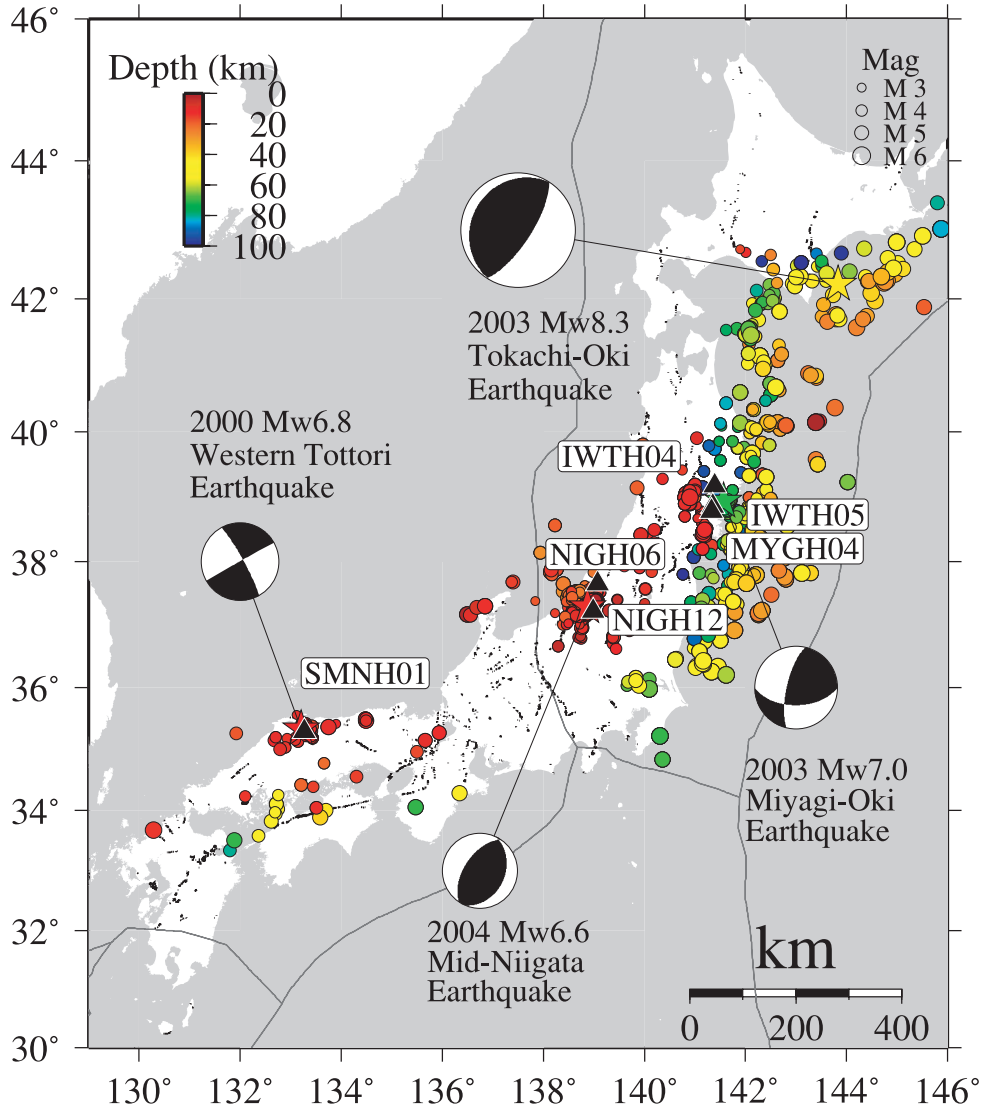

Figure 1. Map of the study region in Japan. Epicentres of the 2000 M_w 6.8 Western Tottori earthquake, 2003 M_w 8.3 Tokachi-Oki earthquake, 2003 M_w 7.0 Miyagi-Oki earthquake and 2004 M_w 6.6 mid-Niigata earthquake are shown in stars together with their moment tensor solution. Other events are shown in circles. The size of circle indicates the magnitude of each event and colour shows the depth with red being shallow and blue being deep. Locations of the 6 KiK-Net stations used in this study are shown in black triangles. The black lines show the active faults in this region and the grey lines denote the subduction plate boundaries.

coda waves. By comparing the spectral ratios, we identify a clear shift of resonant frequency around 4 Hz to lower values and a drop in peak amplitude for the direct *S* waves (Fig. 2c). Both of these features are hallmarks of non-linear site effects (e.g. Beresnev & Wen 1996).

3 RESULTS

After processing all the data, we obtain 13 7980 spectral ratios, including 2204 spectral ratio traces for the direct *S* windows. First we sort all the traces for each station by their PGA and stack them in different PGA ranges (0–500 Gal with an increment of 1 Gal). The PGA for each trace is measured from the amplitude of the vector sum of the two horizontal-component ground acceleration records within the 6-s window. Then we identify the peak spectral ratios and peak frequencies for the stacked traces. Fig. 3 shows all the spectral ratio traces against the corresponding PGAs for the six stations. A general pattern is that the peak frequency drops to lower values when the PGA is between 10 and 100 Gal, although the actual threshold varies for different stations and it is hard to quantify by visual inspection. In addition, the spectral ratio traces in such plots have different seismic phases and contain both the peak motion in each seismogram (which may produce non-linearity) as well as the following weaker motion (at which time the site response may be associated with a recovery process).

To better quantify the degree and process of non-linear response in the spectral ratios for the medium-size events, we select data with PGA less than 100 Gal and focus on the temporal evolutions of the spectral ratios before, during and after the direct *S* waves (which generally have the largest amplitude of all the seismic phases for medium-size events). We group the events in different PGA ranges (0–100 Gal with an increment of 10 Gal) and stack the spectral ratio traces from 12 s before to 60 s after the *S*-wave arrival time within each group. As before, we tested stacking the trace using different PGA ranges, and the obtained results are similar. Next, we identify

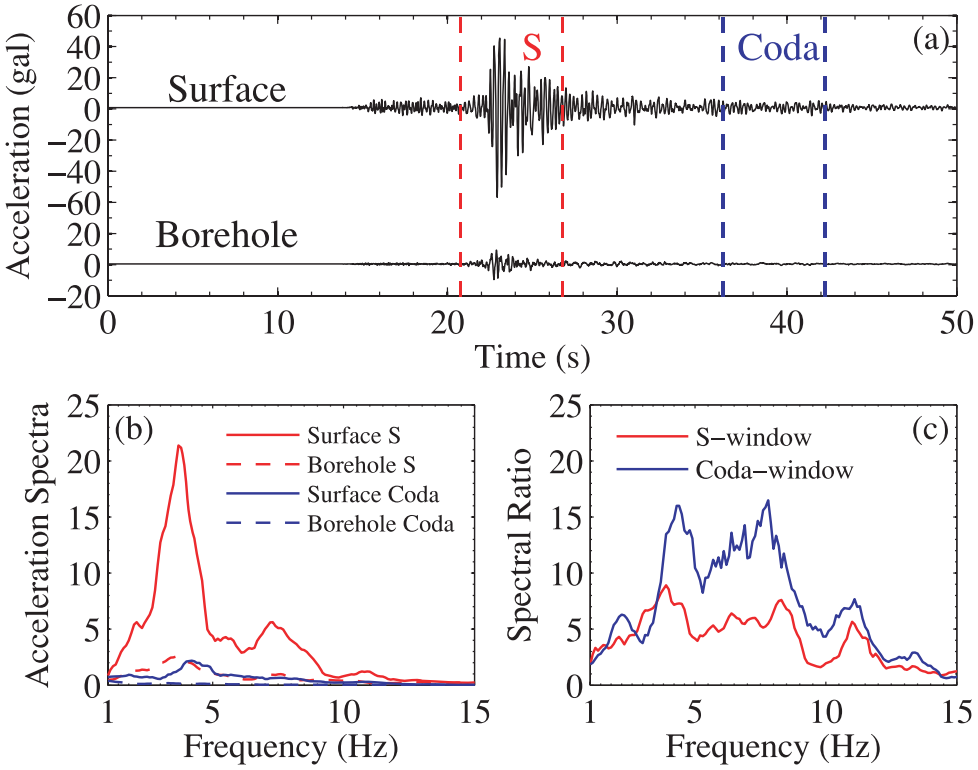

Figure 2. (a) East-component ground accelerations recorded at the station NIGH06 generated by an *M* 5.3 earthquake on 2004 October 25. Surface recording is shown at the top and borehole recording is shown at the bottom. The red and blue dashed lines indicate the direct *S* and coda window that are used to compute the acceleration spectra in (b) and spectral ratios in (c).

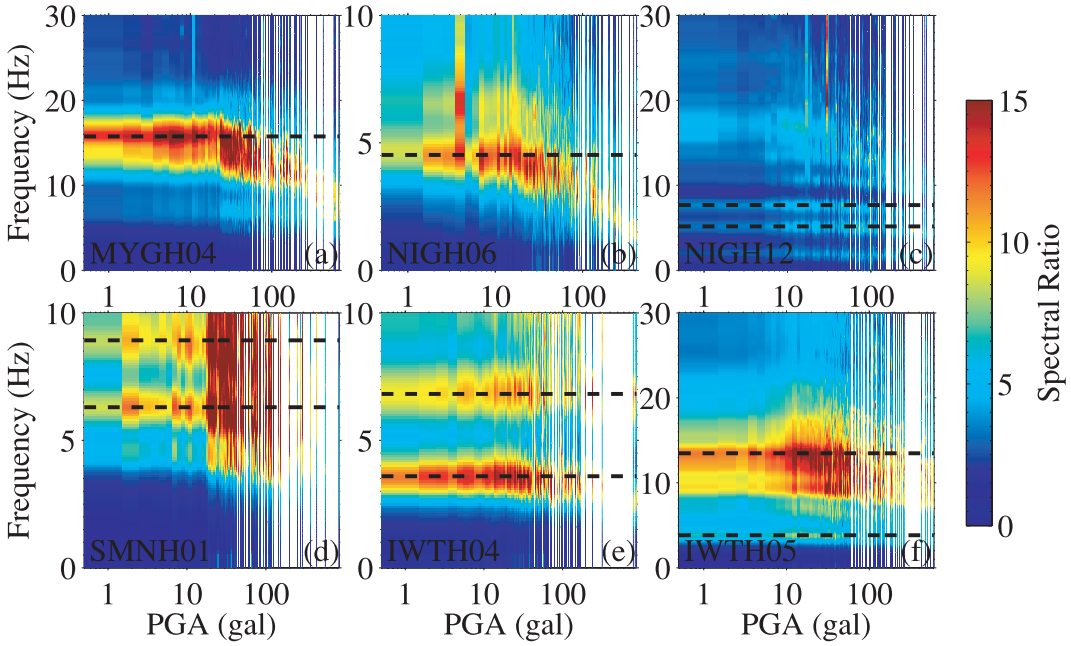

Figure 3. Stacked spectral ratios from all the 6-s windows plotted against PGA for the six stations (a–f). All the spectral ratio traces for each station are stacked based on the PGA of each sliding window from 0 to 1000 Gal with a step of 1 Gal. The *x*-axis shows the PGA value in log scale, and the *y*-axis shows the frequency. The spectral ratio value are colour-coded with red being high and blue being low. Gaps represent no data. The horizontal dashed lines indicate the resonance frequencies without non-linearity measured from 0–10 Gal.

the peak spectral ratio and peak frequency for the stacked trace in each PGA range. Fig. 4 illustrates the procedure for data recorded at site NIGH06. The results show again a sudden drop of peak spectral ratio and a shift of the spectral peak to lower frequencies during the *S*-wave arrival, and these changes become clear when the PGA exceeds certain threshold. Similar analyses using data at the other sites indicate that the threshold varies from ∼20 to ∼80 Gal among the different stations, and the degree of

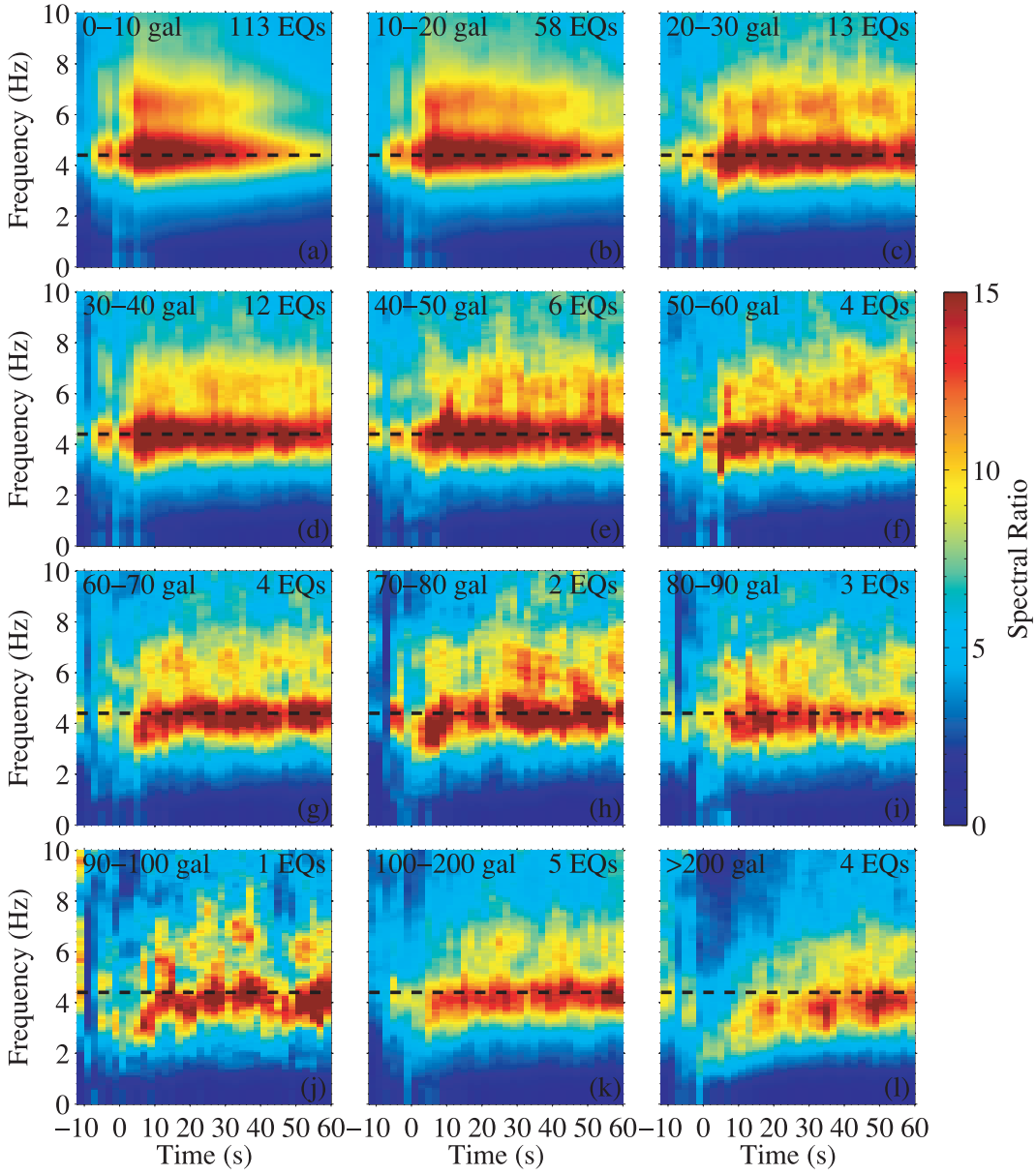

Figure 4. Stacked spectral ratios for station NIGH06 plotted against the traveltime relative to the *S*-wave arrival. Spectral ratio traces from 12 s before to 60 s after the *S*-wave arrival are stacked based on the PGA of that event from 0 to 100 Gal with a step of 10 Gal (a–j). We also include the results between 100–200 Gal and >200 Gal for comparison (k–l). The spectral ratio value are colour-coded with red being high and blue being low. The PGA range and number of the stacked events are marked at the left- and right-hand side on the top of each panel. The horizontal dashed lines indicate the reference resonance frequency of 4.4 Hz.

non-linearity (i.e. percentage drop of peak frequency) also varies for different stations.

Using the obtained results, we examine the relationship between the observed temporal changes and the input ground motion, by relating the percentage of drop in peak frequency and peak spectral ratio with the PGA range for the stacked traces. We use the measured peak frequency and peak spectral ratio for the stacked trace in the PGA range 0–10 Gal as the reference values. To test the validity of the reference value, we have compared the stacked spectral ratio trace in the PGA range of 0–10 Gal with those traces from low-amplitude coda-windows, and found that they are nearly identical (Fig. 5a). We also checked individual spectral ratio traces in the PGA range of 0–10 Gal, and found no significant variation in peak frequency (Fig. 5b).

We calculate the percentage reduction for both peak frequency and peak spectral ratio associated with the *S* waves, by comparing the 6-s windows immediately after the *S*-wave arrivals with the reference values. The peak frequency drop (PFD) and peak spectral ratio drop (PSRD) are defined by

$$\mathrm{PFD} = (f_r - f_s)/f_r, \tag{1}$$

$$\mathrm{PSRD} = (\mathrm{PSR}_r - \mathrm{PSR}_s)/\mathrm{PSR}_r, \tag{2}$$

where f_r and PSR_r are the reference peak frequency and peak spectral ratio, respectively, and f_s and PSR_s are the peak frequency and peak spectral ratio measured at the 6-s window immediately after the *S*-wave arrival, respectively. Because the degree of non-linearity is site specific, we use 10 per cent of the PFD or PSRD measured

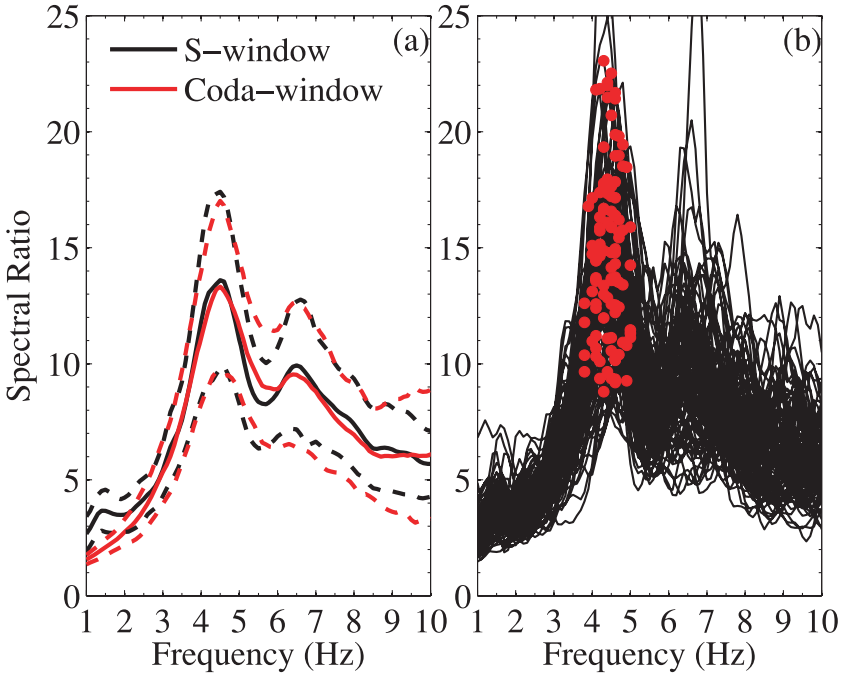

Figure 5. (a) Comparison of the stacked spectral ratios at station NIGH06 from the direct-*S* windows (black) and the *S*-coda windows (red) for all the events in the PGA range of 0–10 Gal. The solid and dashed lines show the average spectral ratios and the standard deviations. (b) The spectral ratio traces from the direct-*S* windows used to generate the black solid line in (a). The red dots mark the lower resonance peaks of the spectral ratios.

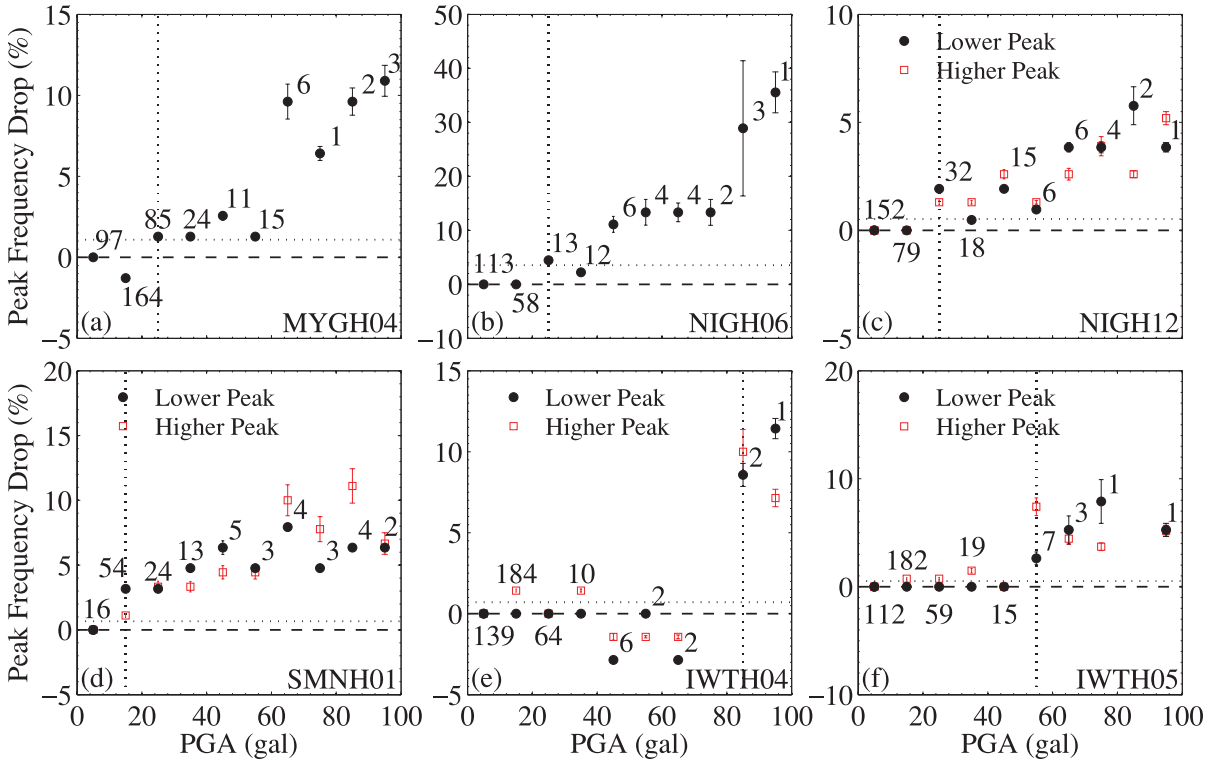

Figure 6. Percentage drop of the peak frequency plotted against the PGA for the six stations (a–f). The black solid circles and red open squares mark the values measured from the 6-s windows immediately after the direct *S*-wave arrivals for the lower and higher resonance peaks, respectively. The vertical solid bar centred at each data point shows the standard deviation, with the number of stacked events by its side. The horizontal dashed line indicates the zero level. The horizontal dotted line marks 10 per cent of the value measured at 90–100 Gal. The vertical dotted line marks the inferred PGA threshold of non-linearity. The station name is marked at the bottom-right corner of each panel.

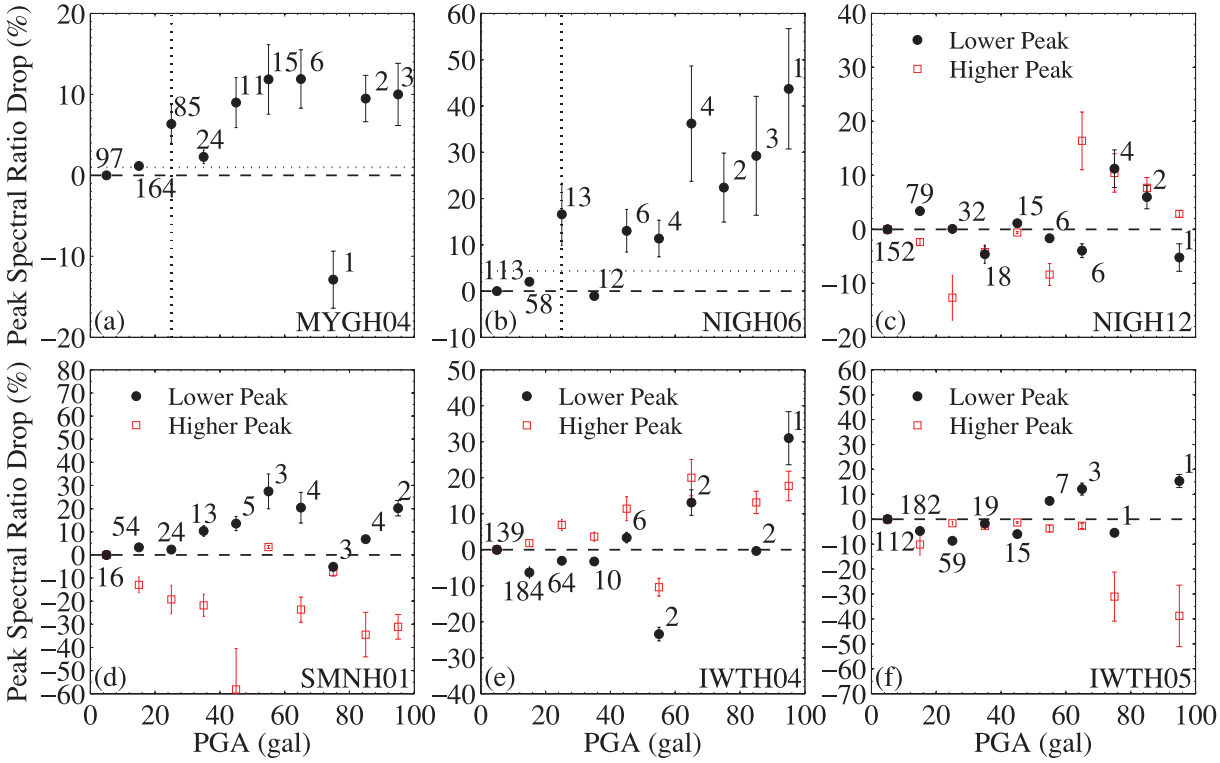

Figure 7. Similar plot as Fig. 6 for percentage drop of the peak spectral ratio.

in the PGA range of 90–100 Gal for each station as the threshold to judge whether non-linearity occurs. Then we repeat the process for other resonance peaks if available. The measured PGA threshold of non-linearity varies (Fig. 6) from ∼20 to ∼80 Gal among all the stations. Fig. 6 also shows that the percentage of drop in peak frequency starts to increase with PGA higher than the threshold for each station. As previous studies (Wu *et al.* 2009a, b) suggested, the uncertainty for peak spectral ratio is much larger than that for peak frequency (see also Fig. 5b). We are able to identify systematic drop of peak spectral ratio only for two stations MYGH04 and NIGH06 (Fig. 7). For the other four stations there is no clear correlation between the percentage of drop in peak spectral ratio and PGA, probably due to the large uncertainties. Because of this, we focus in the discussion on the temporal changes in peak frequency.

4 DISCUSSION

To relate the observed temporal changes in surface-to-borehole spectral ratios with non-linear site response, the following conditions need to be satisfied: (1) the source spectrum should be identical for the station pair; (2) the propagation path should also be the same for the station pair; (3) the reference site should have no significant site response and (4) the reference spectra should be stable over time. Because the surface and borehole stations are very close to each other (on the order of 100 m) as compared to the distances between the earthquake sources and stations (typically more than 5 km), the seismic sources and the propagation path are almost identical for the colocated surface and borehole stations. Hence their source and path terms are effectively cancelled by taking the spectral ratio and the first and second conditions are satisfied. We also examined the spectra of borehole stations and found no significant peaks or troughs, and no clear variation over time, so the third and fourth conditions are also satisfied to a very good approximation.

Non-linear site effects (i.e. reduction of shear modulus and increased damping) induced by strong motions of medium-size earthquakes are likely to be the main origin of the observed variations in the peak frequencies and peak spectral ratios (e.g. Pavlenko & Irikura 2002; Sawazaki *et al.* 2006; Karabulut & Bouchon 2007; Sawazaki et al. 2009; Wu *et al.* 2009a, b; Rubinstein 2010). Assuming a soft soil layer over a half-space bedrock with large impedance contrast (Dobry *et al.* 2000), the fundamental resonant frequency f of the soil layer can be computed by

$$f = \frac{V_s}{4H}, \tag{3}$$

where V_s and H are the S-wave velocity and thickness of the surface soil layer, respectively. From the site profiles (see the KiK-Net website), we found significant velocity contrast between the near-surface layers and the underlying bedrocks for the six stations, so the aforementioned assumption is justified. Using eq. (3), together with the velocity and layer thickness values from the site profiles, we compute fundamental resonance frequencies of 13.8, 4, 7.7, 6.6, 2.8 and 3.2 Hz for the top layers where strongest velocity contrast occurs at stations MYGH04, NIGH06, NIGH12, SMNH01, IWTH04 and IWTH05, respectively. The computed resonance frequencies agree reasonably well with the observed values of 15.7, 4.4, 7.7, 6.3, 3.5 and 3.8 Hz from the spectral ratios without significant non-linear effects (i.e. in the PGA range of 0–10 Gal). Hence we infer that the observed non-linear effects are mainly constrained within the top few tens of metres for the six sites.

As mentioned, the results of Fig. 6 indicate that the peak frequency starts to decrease when the PGA is on the order of several tens of Gal, above which there is a clear correlation between the percentage of drop in peak frequency and PGA level. These

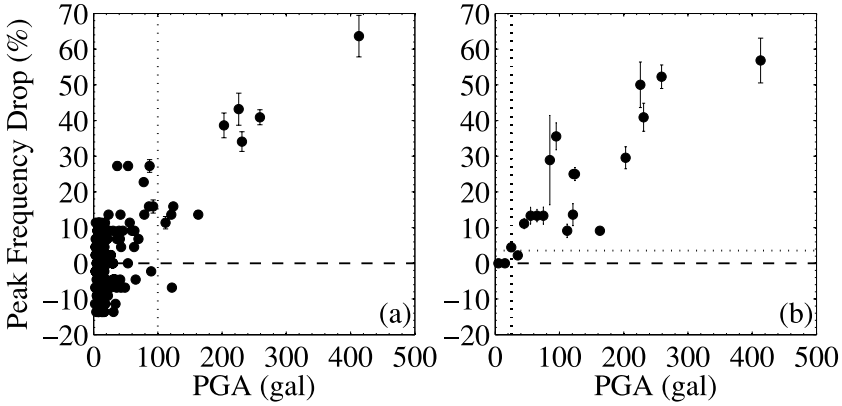

Figure 8. (a) Percentage drop of the peak frequency measured from 10-s windows plotted against the PGA at NIGH06 from Wu *et al.* (2009a). The solid circles mark the values measured from the 10-s windows immediately after the direct *S*-wave arrivals. The vertical solid bar centred at each data point shows the standard deviation. The horizontal dashed line indicates the pre-mainshock value. The vertical grey dotted line marks the PGA value of 100 Gal. (b) Similar plot as (a) for 6-s windows. Values below 100 Gal are measured from the stacked spectral ratio traces (Fig. 6b), and values above 100 Gal are measured from individual even. The horizontal dotted line marks 10 per cent of the value measured at 90–100 Gal. The vertical dotted line marks the inferred PGA threshold of non-linearity.

observations, together with the results of previous studies (e.g. Hartzell 1998; Su *et al.* 1998; Johnson *et al.* 2009; Wu *et al.* 2009a), support the view that non-linear effects strongly correlate with the level of ground motion. However, the ∼20–30 Gal PGA threshold revealed by our analysis is lower than the value of 100–200 Gal found by most previous studies (e.g. Beresnev & Wen 1996). This is perhaps not too surprising, because as far as we know none of the previous studies except the recent work by Rubinstein (2010) investigated non-linear site effects with PGA less than 100 Gal (typically associated with medium-size events). The non-linear effects with PGA less than 100 Gal are generally much weaker than the traditional non-linear site effects associated with stronger shaking (Fig. 8b). Since most previous studies mainly focused on individual events (rather than stacked data), the subtle non-linear effects are quite likely to be buried by low signal-to-noise ratio (SNR). From the comparison of our results (Fig. 8b) with the results of Wu *et al.* (2009a) without stacking (Fig. 8a), we could see a clearer image of the PFD in the PGA range of 0–100 Gal after stacking. Hence, the stacking technique utilized in this study helps to reveal the otherwise buried small non-linear effects in a statistical way. However, stacking requires a significant amount of strong motion recordings, which was not available until the last decade. In addition, reducing the sliding-window size from 10 to 6 s helps to identify subtle coseismic changes associated only with the direct *S* waves, which also causes some differences in the results for PGA > 100 Gal (Fig. 8).

The PFDs found in this study for PGA < 100 Gal are typically within the duration of the *S* wave (e.g. Fig. 4). In some cases (e.g. 60–70, 70–80 and 90–100 > 200 Gal), we can identify the logarithmic type of recovery found by various previous studies (e.g. Sawazaki *et al.* 2006; Wu *et al.* 2009a). However, such recovery is not clear in other cases with lower PGA values. One possible explanation is that the observed weak non-linear effects are induced in damaged rocks by increasing wave amplitude without causing additional material damage. This is consistent with the analytical and numerical results of Lyakhovsky *et al.* (2009) for a continuum non-linear damage rheology model. The results from that study show a few percentage shift in the resonance frequency with increasing ground motion at a constant damage level, and tens of percentage shift in the resonance frequency with increasing damage level. These two forms of non-linear effects are observed clearly in the laboratory experiments of Pasqualini *et al.* (2007).

Based on the results of Fig. 6, the percentage of drop in peak frequency is generally very small (<5 per cent) at the onset of non-linearity, except for station IWTH04. At this station, the point between 60 and 70 Gal is missing, which may blur the true picture. When the PGA value is approaching 100 Gal, the percentage drop at station NIGH06 could be as large as 35 per cent. Therefore, when the PGA level is relatively low the ground motion is probably not strong enough to cause additional damage to the medium. In this case weak non-linear effects with increasing ground motion for material with frozen damage level or very minor additional damage is likely to be the major mechanism for our observations. On the other hand, when the PGA level is relatively high (i.e. approaching or large than 100 Gal), additional damage is generated and we start to observe much larger changes in the peak frequency (Fig. 8), followed by logarithmic recovery (Fig. 4). We have tried to quantify the recovery timescale for the coseismic PFD (see Wu *et al.* 2009a for additional analysis details). However, because the coseismic drop in peak frequency is small (Fig. 6, especially for lower PGA ranges), and the logarithmic recovery on short timescales is very fast (Fig. 4), we were not able to identify a sharp boundary between the near instantaneous recovery and the rapid logarithmic recovery. Based on the stacked spectral ratios for the six sites (e.g. Fig. 4), the higher PGA threshold for production of significant additional damage and considerably stronger non-linear effects (i.e. tens of percentage drop in peak frequency followed by a clear logarithmic type of recovery) ranges from ∼60 Gal to more than 100 Gal.

We note that higher pre-existing damage makes the medium more susceptible to further damage, since it is generally easier to nucleate and grow cracks in material with higher crack density. Rubinstein & Beroza (2004) and Rubinstein (2010) used the correlation between pre-exiting damage and susceptibility to additional damage to explain non-linear effects associated with increasing damage induced by M 4–5.5 aftershocks of the 1989 M_w 6.9 Loma Prieta earthquake and 2004 M_w 6.0 Parkfield earthquake. Here we attempt to resolve non-linear response that is not associated with further increase of the damage. We believe this is the situation in the cases with near instantaneous recovery to linear behaviour in the times windows after the direct *S* waves for the following two reasons. First, we examined the damage recovery time of both large and small earthquakes utilized in this study and by Wu *et al.* (2009a) and found that the maximum recovery time to 90 per cent of the

reference value is less than 200 s. Since most aftershocks occurred well beyond this time interval, most of the damage caused by the preceding large earthquakes has been recovered at the occurrence time of each earthquake we utilized, and is continuously decreasing further. Some low-amplitude damage remains for longer durations, as documented in studies with repeating earthquakes (e.g. Rubinstein & Beroza 2004; Peng & Ben-Zion 2006), and contributes to the overall long-term damage evident by the lower seismic velocities in the near surface material and fault zone structures (e.g. Li *et al.* 1990; Peng *et al.* 2003; Liu *et al.* 2005; Lewis & Ben-Zion 2010). Such remaining and decreasing transient damage will produce some apparent non-zero recovery time of the non-linear behaviour. However, this damage was produced by the previous events rather than by the motion that generates the subtle non-linear effects we observe. Second, we also identified weak non-linear effects for medium-size earthquakes before the $M > 6$ events in each site, and we found no clear change in the non-linear response for medium-size earthquakes before and after large events. Hence, we conclude that some of the non-linear effects we observe for the medium-size earthquakes reflect non-linear response of damaged rocks without additional increments to the levels of damage.

According to Fig. 6, the PGA onset and degree of non-linearity appear to vary significantly at different sites, so the PGA threshold probably also depends on the site conditions. We examined the site conditions in terms of the site classification by the average *S*-wave velocity (V_{S30}) in the upper 30 m of the site (NEHRP 2003), soil types at the top layers of each site and the *S*-wave velocity contrast, but none of these properties shows any clear correlation with the PGA onset (see Table S1). It is likely that the existing rock damage at each site controls the onset of non-linearity (e.g. Lyakhovsky *et al.* 2009). A useful quantification of the near-surface damage layer is the crack density, which could be estimated from analyses of seismic scattering and shear wave splitting (e.g. Revenaugh 2000; Liu *et al.* 2005; Chao & Peng 2009). However, such analyses are beyond the scope of this paper. In any case, the results obtained in this study, together with the recent work of Rubinstein (2010), suggest that non-linear site response can be induced by medium-size (*M* 4–5) earthquakes and can occur more commonly than previously thought. The observed weak form of non-linearity with near instantaneous recovery is likely associated with near constant material damage. As noted before, laboratory studies have identified non-linear effects for geomaterials under strains as small as 10^{-8} (TenCate *et al.* 2004), which is much smaller than the dynamic strains (10^{-6}–10^{-5}) induced by the earthquakes utilized in this study. Observing non-linear effects associated with such weak motion under *in situ* conditions will require the development of additional analysis techniques and high resolution data. Although the significance of weak non-linearity is case specific, quantification of non-linearity at different degrees will help to further advance our understanding of site response to shaking caused by earthquakes.

ACKNOWLEDGMENTS

We thank the National Research Institute for Earth Science and Disaster Prevention (NIED) for providing us with the strong motion records of KiK-Net. We thank Reiji Kobayashi for providing the knet2sac program to convert original k-net data format into SAC format, and Wei Li for helping with the data download and insightful discussions. The study was funded by the National Science Foundation (grant EAR-0908310) and the Southern California Earthquake Center (based on NSF Cooperative Agreement EAR-0106924 and USGS Cooperative Agreement 02HQAG0008). The paper benefited from comments by two anonymous referees and Editor Massimo Cocco.

REFERENCES

Aoi, S., Obara, K., Hori, S., Kasahara, K. & Okada, Y., 2000. New Japanese uphole/downhole strong-motion observation network: KiK-net, *Seism. Res. Lett,* **72,** 239.

Assimaki, D., Li, W., Steidl, J. & Tsuda, K., 2008. Site amplification and attenuation via downhole array seismogram inversion: a comparative study of the 2003 Miyagi-Oki aftershock sequence, *Bull. seism. Soc. Am.,* **98,** 301–330.

Beresnev, I. & Wen, K., 1996. Nonlinear soil response - A reality?, *Bull. seism. Soc. Am.,* **86,** 1964–1978.

Chao, K. & Peng, Z., 2009. Temporal changes of seismic velocity and anisotropy in the shallow crust induced by the 1999 October 22 M6. 4 Chia-Yi, Taiwan earthquake, *Geophys. J. Int.,* **179,** 1800–1816.

Chin, B. & Aki, K., 1991. Simultaneous study of the source, path, and site effects on strong ground motion during the 1989 Loma Prieta earthquake: a preliminary result on pervasive nonlinear site effects, *Bull. seism. Soc. Am.,* **81,** 1859–1884.

Dobry, R., *et al*. 2000. New site coefficients and site classification system used in recent building seismic code provisions, *Earthq. Spectra,* **16,** 41–67.

Frankel, A., *et al*. 2000. USGS national seismic hazard Maps, *Earthq. Spectra,* **16,** 1–19.

Goldstein, P., Dodge, D., Firpo, M. & Minner, L., 2003. SAC2000: signal processing and analysis tools for seismologists and engineers. in *The IASPEI International Handbook of Earthquake and Engineering Seismology, Part B, Chap 85.5,* eds Lee, W.H.K., Kanamori, H., Jennings, P. C. & Kisslinger, C., Academic Press, London.

Hartzell, S., 1998. Variability in nonlinear sediment response during the 1994 Northridge, California, earthquake, *Bull. seism. Soc. Am.,* **88,** 1426–1437.

Johnson, P., Bodin, P., Gomberg, J., Pearce, F., Lawrence, Z. & Menq, F., 2009. Inducing in situ, nonlinear soil response applying an active source, *J. geophys. Res.,* **114,** B05304, doi:10.1029/2008JB005832.

Karabulut, H. & Bouchon, M., 2007. Spatial variability and non-linearity of strong ground motion near a fault, *Geophys. J. Int.,* **170,** 262–274.

Kramer, S. & Paulsen, S., 2004. Practical use of geotechnical site response models, **10**, in *Proceedings of International Workshop on Uncertainties in Nonlinear Soil Properties and their Impact on Modeling Dynamic Soil Response,* University of California, Berkeley.

Lewis, M.A. & Ben-Zion, Y., 2010. Diversity of fault zone damage and trapping structures in the Parkfield section of the San Andreas Fault from comprehensive analysis of near fault seismograms, *Geophys. J. Int.,* submitted.

Li, Y., Leary, P., Aki, K. & Malin, P., 1990. Seismic trapped modes in the Oroville and San Andreas fault zones, *Science,* **249,** 763–766.

Liu, Y., Teng, T. & Ben-Zion, Y., 2005. Near-surface seismic anisotropy, attenuation and dispersion in the aftershock region of the 1999 Chi-Chi earthquake, *Geophys. J. Int.,* **160,** 695–706.

Lyakhovsky, V., Hamiel, Y., Ampuero, J. & Ben-Zion, Y., 2009. Non-linear damage rheology and wave resonance in rocks, *Geophys. J. Int.,* **178,** 910–920.

NEHRP, 2003, *NEHRP Recommended Provisions for Seismic Regulations for New Buildings and Other Structures (FEMA 450),* National Earthquake Hazards Reduction Program (NEHRP), Building Seismic Safety Council, Washington, DC.

Ostrovsky, L. & Johnson, P., 2001. Dynamic nonlinear elasticity in geomaterials, *Rivista del Nuovo Cimento,* **24,** 1–46.

Pasqualini, D., Heitmann, K., TenCate, J., Habib, S., Higdon, D. & Johnson, P., 2007. Nonequilibrium and nonlinear dynamics in Berea and Fontainebleau sandstones: low-strain regime, *J. geophys. Res,* **112,** B01204, doi:01210.01029/02006JB004264.

Pavlenko, O. & Irikura, K., 2002. Nonlinearity in the response of soils in the 1995 Kobe earthquake in vertical components of records, *Soil. Dyn. Earthq. Eng.,* **22,** 967–975.

Peng, Z. & Ben-Zion, Y., 2006. Temporal changes of shallow seismic velocity around the Karadere-Düzce branch of the north Anatolian Fault and strong ground motion, *Pure appl. Geophys.,* **163,** 567–600.

Peng, Z., Ben-Zion, Y., Michael, A. & Zhu, L., 2003. Quantitative analysis of seismic fault zone waves in the rupture zone of the 1992 Landers, California, earthquake: evidence for a shallow trapping structure, *Geophys. J. Int.,* **155,** 1021–1041.

Revenaugh, J., 2000. The relation of crustal scattering to seismicity in Southern California, *J. geophys. Res.,* **105,** 25403–25422.

Rubinstein, J., 2010. Nonlinear strong ground motion in medium magnitude earthquakes near Parkfield, CA, *Bull. seism. Soc. Am.,* submitted.

Rubinstein, J. & Beroza, G., 2004. Nonlinear strong ground motion in the ML 5.4 Chittenden earthquake: Evidence that preexisting damage increases susceptibility to further damage, *Geophys. Res. Lett,* **31,** L23614, doi:23610.21029/22004GL021357.

Sawazaki, K., Sato, H., Nakahara, H. & Nishimura, T., 2006. Temporal change in site response caused by earthquake strong motion as revealed from coda spectral ratio measurement, *Geophys. Res. Lett.,* **33,** L21303, doi:21310.21029/22006GL027938.

Sawazaki, K., Sato, H., Nakahara, H. & Nishimura, T., 2009. Time-lapse changes of seismic velocity in the shallow ground caused by strong ground motion shock of the 2000 Western-Tottori earthquake, Japan, as revealed from coda deconvolution analysis, *Bull. seism. Soc. Am.,* **99,** 352–366.

Su, F., Anderson, J. & Zeng, Y., 1998. Study of weak and strong ground motion including nonlinearity from the Northridge, California, earthquake sequence, *Bull. seism. Soc. Am.,* **88,** 1411–1425.

TenCate, J., Pasqualini, D., Habib, S., Heitmann, K., Higdon, D. & Johnson, P., 2004. Nonlinear and nonequilibrium dynamics in geomaterials, *Phys. Rev. Lett.,* **93,** doi:10.1103/PhysRevLett.1193.065501.

Trifunac, M., Hao, T. & Todorovska, M., 1999. On the reoccurrence of site specific response, *Soil. Dyn. Earthq. Eng.,* **18,** 569–592.

Tsuda, K., Archuleta, R. & Jamison, S., 2006. Confirmation of nonlinear site response: case study from 2003 and 2005 Miyagi-Oki earthquakes, *Bull. seism. Soc. Am.,* **96,** 926–942.

Wu, C., Peng, Z. & Assimaki, D., 2009a. Temporal changes in site response associated with strong ground motion of 2004 Mw6. 6 mid-Niigata earthquake sequences in Japan, *Bull. seism. Soc. Am.,* **99,** 3487–3495.

Wu, C., Peng, Z. & Ben-Zion, Y., 2009b. Non-linearity and temporal changes of fault zone site response associated with strong ground motion, *Geophys. J. Int.,* **176,** 265–278.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Tables with site and event information.

Table S1. Station locations, and the site conditions including V_{S30} (m s^{-1}), site classification, max V_S constrast, soil types and nonlinearity threshold (Gal) for the six sites.

Table S2. List of the information of all the events utilized in this study. Fields 1 to 12 are: event ID, year, day of year, hour, minute, second, magnitude ($M_{\rm jma}$), latitude, longitude, depth (km), PGA (Gal) and PGV (cm s^{-1}), respectively.

Figure of V_p and V_s profiles

Figure S1. V_p (red lines) and V_s (blue lines) profiles for the six sites.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.