

## **Sensitivity of site response analysis on the number of ground motion records and implications for PSHA**

Papaspiliou M.

*Willis Re, 51 Lime Street, London, EC3M 7DQ, UK.*

*(Formerly Department of Civil and Environmental Engineering, Imperial College London, UK)*

Tel: +44(0)2031247637

Email: myrto.papaspiliou@willis.com

Kontoe S.

*Department of Civil and Environmental Engineering, Imperial College London,  
South Kensington Campus, London, SW7 2AZ, UK*

**Abstract** This paper investigates issues related to the number of ground-motion records required for the performance of site response analysis and the inclusion of the site-specific amplification function within probabilistic seismic hazard calculations (PSHA). It explores the minimum number of records required for a robust estimation of the median and standard deviation of the site amplification function, as well as the impact of the selected ground-motion suites on the results of PSHA. Site response analyses are performed using both equivalent linear and nonlinear methodologies. Although the median amplification was observed to be relatively easy to capture, the standard deviation was seen to fluctuate considerably, especially when suites of few records were used. It was observed that in the case of the nonlinear site response analysis 10 records provide relatively stable estimates of the hazard curves for the majority of periods, while in the case of the equivalent linear analysis 20 records or more are required to achieve a similar level of accuracy.

**Keywords:** *Site response, ground-motion selection, PSHA*

## 1 Introduction

A probabilistic seismic hazard analysis (PSHA) is typically performed to estimate the ground motion at rock site conditions and is subsequently modified deterministically to account for site effects, i.e. using only the median value of the site factor and ignoring its uncertainty. However, this procedure underestimates ground motions evaluated probabilistically (Bazzurro and Cornell 2004b; Goulet and Stewart 2009). In order to overcome this shortcoming, site-specific PSHA needs to be performed that considers the nonlinear site effects within the hazard calculations, either using an appropriate ground-motion prediction equation (e.g. Power et al. 2008) or using methods that allow the proper incorporation of site-specific effects (e.g. Bazzurro and Cornell 2004b; Baturay and Stewart 2003). This paper investigates the minimum number of ground motion records required for a robust estimation of the site-specific amplification function and its incorporation within the hazard calculations.

The topics of ground-motion selection and the minimum required number of records have been of concern for many cases, whether it is for the purposes of structural or geotechnical dynamic analysis. Numerous studies have produced guidelines for the performance of structural dynamic analyses and the selection of ground-motion records using different criteria (e.g. Bommer and Acevedo 2004; Goulet et al. 2008; Iervolino et al. 2008; Katsanos et al. 2010; Naeim et al. 2004). However, in the case of geotechnical dynamic analysis, the guidelines available in literature are much more limited (e.g. Jayaram et al. 2011; Kottke and Rathje 2008; Tönük and Ansal 2010).

This paper explores the minimum number of records required for a robust estimation of the site amplification function, as well as the impact of the selected ground-motion suites on the results of PSHA. Site response analyses are performed using both equivalent linear and nonlinear methodologies in order to identify the number of records required for the stable estimation of the site amplification function and the surface hazard curve in each case. The methodology proposed by Bazzurro and Cornell (2004b) involving the transformation of a rock ground-motion prediction equation (GMPE) into a site-specific one is employed.

In engineering practice due to time, budget and resource constraints site-specific site response analysis is usually performed for a limited number of records, while its results are not incorporated into the hazard calculations. It is therefore important to quantify the implications of using suites of a certain number of records on both the site response and the hazard estimates. The incorporation of site effects within the hazard calculations is essential so that the uncertainty associated with the site response can also be taken into consideration. This paper, therefore, explores, not only how many records are required to capture the median amplification function, across a wide range of accelerations and frequencies, but also the number of records required for a robust estimation of its standard deviation. This issue is of particular importance, as site response analyses can be particularly time consuming to perform, especially when advanced methodologies are used that employ complex constitutive models.

## 2 Ground-motion selection and site response analysis

The investigation is performed for the Sylmar County Hospital (SCH) site which is a sandy site in Los Angeles, California consisting of 90m of alluvial deposits. The water table is located 46m below the surface and the average shear wave velocity in the upper 30m is 280m/s (NEHRP class D) (Figs 1a and 1b). The ground-motion records used in this study were selected from the Pacific Earthquake Engineering Research Center's NGA database based on the following criteria:

- Magnitude  $M_w \geq 6.0$
- Source-to-site distance: 0-100km

- Average shear wave velocity in the upper 30m ( $V_{s30}$ ) at recording site  $\geq 760$ m/s
- Geomatrix Class 1: A, B, I (i.e. free-field records)

The values of magnitude  $M_w$  vary between 6.0 and 7.9 and the shortest distance from the recording station to the ruptured area varied from approximately 2 to 95km. The initial selected dataset (53 unscaled records) included only a small number of records with magnitude larger than 7.0 and the PGA of most records did not exceed 0.3g. Therefore the ground-motion records were linearly scaled resulting in a database of 120 acceleration time histories, of both scaled and unscaled records, which capture a wider range of spectral accelerations as shown in Fig. 1c. The use of 120 records is considered sufficient in order to capture the “true” site amplification function and its standard deviation, as far as these can be captured by one-dimensional (1D) site response analysis. This is also in accordance with the number of records used by Bazzurro and Cornell (2004a) for validation purposes of their study (78 records).

Site response analysis is performed using both equivalent linear analysis (SHAKE91) and nonlinear analysis using the modified hyperbolic (MKZ) constitutive model (Matasovic and Vucetic 1993) and the software DMOD-2000 (Geomotions 2007). Fig. 2 shows the shear modulus and damping curves (Darendeli 2001) used for the top layer of the soil column, together with the curves assumed in the nonlinear analysis. The latter have been selected to average the misfit between the shear modulus degradation and damping curves, as the MKZ model cannot produce a perfect fit to both curves, across the entire strain range. The equivalent linear analysis was also based on the shear modulus and damping curves of Darendeli (2001), calculated at the middle point of each layer. More details regarding the site response analysis can be found in Papaspiliou et al. (2012a).

Once the site response analyses have been performed, the site amplification function ( $AF(f)$ ) is obtained using nonlinear regression and following the functional form proposed by Goulet (2008):

$$\ln AF(f) = c_0 + c_1 \ln(S_a^r + c_2) + \varepsilon_{\ln AF(f)} \sigma_{\ln AF(f)} \quad (1)$$

$$\sigma_{\ln S_a^s(f)} = \sqrt{\left(\frac{c_1 \cdot S_a^r(f)}{c_2 + S_a^r(f)} + 1\right)^2 \cdot \sigma_{\ln S_a^r(f)}^2 + \sigma_{\ln AF(f)}^2} \quad (2)$$

where  $AF(f)$  is the frequency dependent amplification function,  $f$  is a generic oscillator frequency,  $S_a^s(f)$  and  $S_a^r(f)$  are the 5% damped spectral acceleration values at the soil surface and at the bedrock respectively,  $\varepsilon_{\ln AF(f)}$  is the standard normal variable,  $\sigma_{\ln AF(f)}$  is the standard error and  $c_0$ ,  $c_1$  and  $c_2$  are constants obtained from the regression analysis.

According to findings by Bazzurro and Cornell (2004a) the estimates of the statistics of the amplification function can be obtained by a limited number of soil dynamic analyses, independent of the method used for the performance of the site response analysis. Bazzurro and Cornell (2004a; 2004b) performed regressions of  $\ln AF(f)$  on  $\ln S_a^r(f)$  based on a large dataset of records. However, they suggest that in practice, given the relatively small values of  $\sigma_{\ln AF(f)}$  for a given value of  $S_a^r(f)$ , only 10 soil response analyses are sufficient to achieve a level of accuracy of  $\pm 10\%$  in estimating the median  $AF(f)$ . However, it should be noted that if a range of hazard levels are of interest, then the records to be used need to have spectral accelerations that span the entire range expected in the hazard analysis. Failure to do so would automatically signify the extrapolation of the amplification function to acceleration levels for which it is not constrained and could lead to severely erroneous amplification and incorrect hazard estimates. Similarly, if more than one spectral period is of interest for the application, then records with a good range of accelerations need to be considered for more than one spectral period as well. For this reason this paper evaluates the sensitivity of the Bazzurro and Cornell (2004b) method to suites of different numbers of records which span over a wide acceleration and period

range. The entire set of 120 records is used to obtain a benchmark site amplification function; subsequently, records are selected to form 11 suites of 10, 15 and 20 records. Records are selected based on their rock spectral acceleration in order to capture a wide range of values, across the period range of interest (0.01s - 3.0s). It is noted that the selection of ground motion records should generally consider the magnitude and distance range that is applicable for the site in question. However it is highlighted that given a spectral acceleration value the prediction of the site amplification function does not significantly improve with the inclusion of additional predictor variables, such as magnitude, particularly at short periods (Bazzurro and Cornell, 2004a). Initially the 10-record suites are selected and they are subsequently enhanced with a further 5 and 10 records, respectively, to improve the range of accelerations captured.

### 3 Impact of the number of records on $AF(f)$

As previously explained, the robust estimation of the median site amplification function and its standard deviation is a critical step of the Bazzurro and Cornell (2004a) methodology. In this section the sensitivity of the median site amplification function on the examined number of ground motion records is investigated considering the results of both equivalent linear and nonlinear site response analysis.

#### 3.1 Sensitivity of the median $AF(f)$

Fig. 3 presents the amplification functions obtained from the 10- 15- and 20-record suites (for  $PGA$ ,  $T=0.2s$  and  $T=1.0s$ ) for the equivalent linear analysis. The individual amplification functions are plotted only over their valid range and therefore not all of them extend through the entire acceleration range covered by the 120-record dataset. At  $PGA$  (i.e.  $T=0.01s$ ), the amplification functions show very little scatter and compare very well with the “benchmark”  $AF(f)$ , even when only 10 records are used. Equivalent linear analysis is known to predict flatter spectral response and to have lower standard deviation in the high-frequency range than nonlinear analysis and empirical methods (e.g. Stewart & Kwok (2008), Lessi-Cheimarios et al 2012). This is due to the use of a constant damping ratio irrespectively of the intensity of each cycle of the ground motion, which leads to the over-damping of low intensity cycles that are often associated with high frequencies. On the other hand, a significant deviation from the median of the 120 records is noted at  $T=0.2s$  and  $T=1.0s$ . In the case of the 10-record suites, the amplification factor for records with  $S_a^r(0.2s)$  larger than about 1g varies by a factor of 4, between 0.5 and 2.0, among the different suites. In this case, at least 20 records are required for the median  $AF(f)$  to be captured within one standard deviation of the “true” median function across the entire rock acceleration range.

Figs. 4a and 4b examine the records that comprise the suites that lie outside the one-standard-deviation interval for  $T= 0.2s$  (i.e. suites No 7 and No 8). The records used in each suite, and particularly in the higher end of  $S_a^r(0.2s)$ , predict significantly different amplification (or de-amplification in the case of Fig. 4a) and considerably affect the entire slope of  $AF(f)$ . Particularly interesting is the difference between the two strongest records in the suites which both have a rock spectral acceleration close to 2.5g. The record of suite No 7 results in an amplification factor of about 0.35, while the record of suite No 8 predicts an amplification factor equal to 1.3. Similar differences are encountered between the two records with  $S_a^r(0.2s)$  close to 1.5g; the No 7 suite record leads to an amplification factor of 0.5, while the other one approaches a value of 2.

To further investigate the above-mentioned differences the response spectra of the two strongest records are shown in Fig. 5, both for rock and soil. Although the records have very similar  $PGA$  and  $S_a^r(0.2s)$  values, there are considerable differences in the longer period range where the record from suite No 7 has a very rich content with large peaks between about  $T=0.7$  and  $T=1.5s$ . The close match of those peaks with the predominant site period (0.95s) means that the dynamic response is significantly

enhanced. The iterative procedure of the equivalent linear analysis converges at very large strains, in excess of 1%, while in the case of suite No 8 strains do not exceed 0.3%. As a result, in the former case considerable stiffness degradation has taken place (convergence at 7% of the maximum stiffness) and damping reaches 20%, whereas in the latter case damping barely exceeds 10%. Such differences between the two records explain the considerable variability in the amplification for 0.2s.

Furthermore it is interesting to note that both records of suite No 7, which led to the severe de-amplification, have very rich content at periods between 0.7 and 1.5s, leading to very large strains and hence de-amplification of the short-period part of the record. On the other hand, the two records of suite No 8, are characterised by a much narrower spectrum, with a predominant period close to 0.2s and much lower accelerations at 1.0s, leading to the significantly larger amplification at the examined range. Clearly, the inclusion of more records in each suite improves significantly the median amplification predictions. The use of a frequency-content parameter could evidently assist in the robust estimation of the standard deviation with a smaller number of records. When records are selected employing as a sole criterion the establishment of a wide rock acceleration range, the use of records with similar frequency content can lead to the underestimation of the standard deviation.

Equivalent linear analysis is often used in engineering practice as, despite its limitations in modelling soil behaviour, it offers a quick and easy way of estimating site response. A comparison to the nonlinear results is performed herein to identify whether the same number of records would be required irrespectively of the method used for the performance of the site response analysis. Fig. 6 presents the amplification functions obtained from the 10- 15- and 20-record suites (for  $PGA$ ,  $T=0.2s$  and  $T=1.0s$ ) for the nonlinear analysis. The median amplification is captured within the one-standard-deviation interval by almost all suites, even those consisting of only 10 records, with the exception of  $T=0.2s$ . Overall the comparison of Figures 3 and 6 demonstrates that the scatter is lower for the nonlinear analysis, allowing the calculation of a robust amplification function with a smaller number of records.

The records of the outlying suite for  $T=0.2s$  of the nonlinear analysis are further examined in Fig. 7. Fig. 7a shows the 10 records comprising that suite and Fig. 7c presents the corresponding 20-record suite, both plotted together with the complete dataset. Clearly, the residuals of the 10-record suite (Fig. 7b) are all negative, while three of the highest intensity records show the largest deviation from the mean leading to the steepening of the amplification function. It should be noted that the rock response spectra of these three records have a predominant spectral period equal to 0.2s (not shown herein for brevity). Although the predominant period is not the most appropriate parameter for the characterisation of the frequency content of a ground motion (Rathje et al. 2004), it does give an indication of a common characteristic among the three outlying records which result in the lowest amplification factors. This suggests that the predominant period could be used to improve the selection of records. As expected, the estimation of  $AF(f)$  improves with increasing number of records. Evidently for the 20-record suite, the  $AF(f)$  plots almost within one standard deviation from the mean. Fig. 7d shows that the larger suite contains records with both positive and negative residuals which push the median amplification closer to the “true” median, making the influence of the three high intensity records less dominant.

### 3.1.1 Sensitivity of the standard deviation, $\sigma_{\ln AF(f)}$

Figs 8 and 9 show the variability in  $\sigma_{\ln AF(f)}$  for the different suites and its comparison to the benchmark value for the equivalent linear and the nonlinear analysis respectively. Clearly, capturing the standard deviation is not as straightforward as capturing the median of  $AF(f)$ . When only a few records are used for the analysis, the ground-motion variability can be significantly under- or overestimated and therefore, a sufficient number of records

need to be selected to ensure that the standard deviation estimation is robust and its true value is captured.

The standard deviation,  $\sigma_{lnAF(0.2s)}$ , of the equivalent linear analysis varies between 0.1 and 0.5 among the suites of 10 records, with suites N<sup>o</sup> 7 and 8 having the lowest  $\sigma_{lnAF(0.2s)}$ . The increase of the record number to 15, limits the standard deviation fluctuation between 0.15 and 0.38, while the use of 20 records brings only modest additional improvement. At  $T=0.01s$ , where the dispersion was already relatively low, small variability is noted in the standard deviations of the different suites, irrespectively of the number of records used. At periods longer than 1.0s the variability in the estimates increases slightly, with the highest differences noted between the estimates of the 10- and 15- record suites.

The nonlinear analysis results show the largest variability among the various periods at  $T=0.2s$ . At this period, when 10 records are used,  $\sigma_{lnAF(f)}$  varies from 0.1 to just over 0.3, while the estimate for the 120 records is equal to 0.185. The use of 15-record suites results in a considerable limitation of the variability in  $\sigma_{lnAF(f)}$  across all periods, while increasing further the number of records to 20 leads to relatively smaller improvements. For  $T=0.2s$ , the standard deviation varies between 0.14 and 0.25 for the 15-records suites, and between 0.16 and 0.22 for the 20-record suites. In the intermediate period range,  $T=0.8s$  and  $1.0s$ , small differences are noted in the standard deviations between the 10-, 15- and 20-record suites. The variability increases again slightly at long periods, while at  $T=3.0s$  a considerable reduction in  $\sigma_{lnAF(f)}$  can only be observed when the suites comprise of 20 records. Fig. 9 clearly shows that the number of records necessary for the robust estimation of the site amplification depends on the amount of scatter in the records. At  $T=0.2s$  and  $3.0s$ , where the standard deviation is higher, a larger number of records is needed for the estimation of the statistics of the site amplification.

The comparison of Figures 8 and 9 indicates that equivalent linear analysis requires a larger number of records than nonlinear analysis in order to estimate the statistics of the site amplification with a similar level of accuracy. Therefore, although the equivalent linear analysis has several advantages in terms of the ease of performance and speed of computation, the need for a larger number of records could limit its attractiveness. The significance of this variation among different suites is further evaluated based on its effece on a probabilistic seismic hazard analysis.

### 3.1.2 Impact of dynamic soil properties on the results

The number of records necessary for the site response analyses depends on the expected standard error for the site. Stiffer sites are expected, in accordance with findings of several studies (e.g., Baturay and Stewart 2003; Goulet 2008), to have amplification functions with smaller variability. As a result, a smaller number of records is expected to be required to capture both the median amplification function and its standard deviation. On the contrary, softer sites can have amplification functions with significantly higher record-to-record variability, meaning that a larger number of records might be required. For the profile examined so far, the nonlinear analysis captured relatively easily the median function with even just 10 records for the majority of cases. This was not the case however for the standard deviation where, individual suites result in larger variation in the estimates of  $\sigma_{lnAF(f)}$ .

In this section the nonlinear analysis is repeated for a medium plasticity clay site with very similar average shear wave velocity in the upper 30m ( $V_{S30}=284m/s$ ) in comparison to the first site, but with distinctly different dynamic soil properties (i.e. stiffness degradation and damping characteristics). The clayey site stratigraphy and the associated shear wave velocity profile are shown in Fig. 10. Although both sites are classified as NEHRP class D, the clayey site is expected to demonstrate more linear response and hence experience lower degradation of stiffness and de-amplification of the ground motion under low intensity shaking when compared to the sandy site. Fig. 11 plots the amplification functions obtained from the 10- 15- and 20-record suites (for  $PGA$ ,  $T=0.2s$  and  $T=1.0s$ ) for the clayey site. The nonlinear analysis captures again the median

amplification within the one-standard-deviation interval by almost all suites and across all spectral accelerations, showing a slight improvement compared to the sandy site. As expected, the clayey site gives higher amplification at the lower intensity range than the sandy site (see Fig.6) which is particularly evident for  $T=1.0\text{sec}$ . Fig.12 presents the variability in  $\sigma_{\ln AF(f)}$  for the different suites of the clayey site and the corresponding benchmark value resulting from the analysis of all 120 records. Overall the clayey site exhibits a similar range of variability between the various suites as the sandy site, with the exception of  $T=0.2\text{sec}$  for which the standard deviation is lower for the 10- and 15-record suites. However, in the case of the clayey site, as the number of records is increased a larger number of suites tend to show a  $\sigma_{\ln AF(f)}$  very close to the benchmark value, particularly for PGA, 0.2s and 3.0s. The above indicates that despite the very similar  $V_{s30}$  of the two profiles, the stiffer response of the clayey site means that there is less variability in the soil response and with the same number of records a slightly better match to the benchmark  $\sigma_{\ln AF(f)}$  is obtained than for the sandy site, in accordance with other research findings.

#### 4 Impact of ground-motion selection on PSHA

The selected ground-motion suites and their respective amplification functions are used to evaluate the impact of ground-motion selection on PSHA and the surface hazard curve. The seismic hazard analysis using the different suites of records is performed using OpenSHA (Field et al. 2003), for a location in California (118.135N, 34.696W). The methodology of Bazzurro and Cornell (2004b), which has been implemented in OpenSHA, allowing the median and standard deviation of the rock Abrahamson and Silva (1997) GMPE to be transformed using site-specific results, has been used. More details regarding the PSHA can be found in Papaspiliou et al. (2012b).

##### 4.1 Equivalent linear analysis

Fig. 13 presents the results of the PSHA for the 10-record suites together with the findings of the hazard analysis when the entire 120-record dataset is used for the derivation of the amplification function. Clearly, at  $T=0.2\text{s}$  the variability in the hazard curves of the 10-record suites is very large, with the spectral acceleration at 2% probability of exceedance in 50 years ranging from 0.9g to almost 3.0g, and dramatically increasing with decreasing probability levels. On the other hand, at PGA level, where the standard deviation of the “true” amplification function was relatively small - in fact smaller than that of the nonlinear analysis- the hazard estimates are more similar and the predicted ground-motions vary only between 0.7 and 0.9g.

In order to explore the sources of these differences among the eleven hazard curves, Fig. 14 presents the transformed ground-motion prediction equations (GMPEs) and the variation of the median  $S_a^s(0.2\text{s})$  and  $\sigma_{\ln S_a^s(0.2\text{s})}$  with distance for a magnitude 7.7 event. This was based on a deaggregation analysis which identified the dominant event as a magnitude 7.7 at a distance of about 14km along the San Andreas fault system (see Papaspiliou et al. (2012b)). The examined site location is situated approximately 15km from the San Andreas fault and, therefore, the hazard at the site is expected to be dominated by the occurrence of events along the San Andreas fault system. Evidently, there is considerable variability in  $\sigma_{\ln S_a^s(0.2\text{s})}$  for the different suites. Apart from the variation in the median predictions, the standard deviation varies from 0.15 to almost 0.6 (ln units) for a distance around 15km. As far as the median spectral accelerations are concerned, suites N° 7 and N° 8, examined earlier, produce the lowest and highest surface accelerations respectively. The median amplification functions at 0.2s of these suites lay the farthest away from the 120-record median, while they were characterised by the lowest standard deviation,  $\sigma_{\ln AF(0.2\text{s})}$ , among all suites.

The effects of the very low standard deviation together with the strong deamplification of suite N° 7 are clearly visible in Fig. 15 and explain the significantly lower hazard

estimates obtained by this suite. On the other hand, the hazard curve for suite N° 8 seems to be driven by the high median amplification resulting in the highest hazard estimate of all curves since it predicts no soil deamplification across the entire spectral acceleration range. In addition, the very mild amplification function slope (a function of the regression parameter  $c_1$  in Eq. 2) affects the modification of the standard deviation term of the GMPE. Consequently, despite the very low  $\sigma_{\ln AF(0.2s)}$ , the standard deviation of the equation is considerably higher than that of suite N° 7 due to the “less negative”  $c_1$  value (Eq. 2), as seen in Fig. 14b. The effects of suite N° 8 are also visible at  $T=1.0s$  where, the median function was seen to shift upwards for spectral accelerations larger than about 0.1g (Fig. 3).

The use of 20 records in Fig.15 offers a significant improvement in the variability of the hazard estimates, in contrast to the 15 records whose use showed limited advantages (not shown herein for brevity). A reduction in the spread of the curves is also noted for the longer spectral periods, but the variability remains high at  $T=0.2s$ .

## 4.2 Nonlinear analysis

Figs. 16 and 17 present the results of the PSHA for the 10- and 20-record suites respectively based on nonlinear site response analysis. Looking at the 2% probability of exceedance in 50 years level (2475yr return period),  $PGA$  varies between 0.62 and 0.9g for the 10-record suites and between 0.6 and 0.7 for the 20-record suites. The nonlinear analysis captured for most suites the median  $AF(f)$  within one standard deviation, irrespective of the number of records, and thus the inclusion of more records mostly offered an improvement in the estimation of the standard deviation.

The hazard curves show that the reduction in variability achieved by the addition of more records in the dataset has a relatively small effect on  $PGA$ . However, at  $T=0.2s$ , where the largest variations in the median  $AF(f)$  and  $\sigma_{\ln AF(f)}$  were observed, the use of more records has a clear effect. At this period and for the same APE, it is seen that when only 10 records are used, the surface spectral accelerations range from 1.1 to 2.0g. The inclusion to the suites of five further records limited the range of  $S_a^s(f)$  values between 1.3 and 2.0g (the 15 record-suites are not shown for brevity), while the 20-record suites result in values between 1.4 and 2.0g. Although the reduction in the variability of  $\sigma_{\ln AF(f)}$  was mostly achieved by increasing the number of records from 10 to 15, a large number of the 20-record suites produce hazard curves identical to that of the 120-record dataset (Fig. 11). It is thus clear that an improvement is achieved for the majority of suites.

Considering the two site response analysis methods, the superiority of the nonlinear analysis over the equivalent linear one is clearly demonstrated. For example, at  $T=0.2s$  the variability in the hazard curves of the equivalent linear 20-record suites is almost the same as the 10-record suites of the nonlinear analysis. This is also true for the hazard curves at  $T=1.0$  and  $T=3.0s$ . For  $T=0.2s$  it is seen (20-record suite, Fig. 15) that at 2% probability of exceedance in 50 years, the equivalent linear ground-motion predictions using 20 records vary from 1.1 to 2.2g, with the corresponding estimates for the nonlinear 10-record suites (Fig.16) varying from just 1.4 to 2.0g. It is evident from the above analysis and earlier findings that the equivalent linear analysis requires more records for the estimation of the hazard than the nonlinear analysis for the considered site.

## 5 Conclusions

This paper investigates the sensitivity of the amplification function and its standard deviation when different suites, of different records and numbers, are used, as well as the impact of the selected ground-motion suites on the results of PSHA. The sensitivity to the above issues was examined for both equivalent linear and nonlinear analysis considering a NEHRP class D sandy site. The “benchmark” median amplification and its standard deviation was obtained by the analyses of 120 acceleration time histories, of both scaled and unscaled records, which captured a wide range of spectral accelerations. The use of



120 records was assumed sufficient to capture the “true” site amplification function and its standard deviation. Although the median amplification was relatively easy to capture by the different suites of records, the standard deviation was seen to fluctuate considerably, especially when suites of few records were used. It should also be noted that the equivalent linear analysis suffered from larger scatter than the nonlinear analysis, particularly at  $T=0.2s$ . As a result, a larger number of records were required for the robust estimation of the median amplification function and its standard deviation when equivalent linear site response analysis was performed.

The nonlinear analysis was repeated for another Class D site, but which was characterised predominantly by clay deposits resulting in distinctly different dynamic soil degradation properties. A slight improvement was noted both in capturing the median amplification and in the variability of the standard deviation.

Furthermore, the results of the hazard analyses were significantly affected by the use of different suites of ground-motion records for the performance of the site response analysis. The predominance of sigma over the PSHA results and the high variation in  $\sigma_{lnAF(f)}$  for the equivalent linear analysis resulted in considerable variation of the hazard estimates, irrespective of the number of records used. The nonlinear site response analysis achieves much more stable results, with a smaller number of records and thus prompts for its use for more robust estimates using fewer records. Specifically, it was seen that 10 records provide relatively stable estimates of the hazard curves for the majority of periods, while in the case of the equivalent linear analysis 20 records or more are required to achieve a similar level of accuracy. Use of 20 records in the case of nonlinear analysis can have some advantages and could be used when higher levels of accuracy are thought to be needed. It should be noted though that these observations refer to the examined site and that more sites with different depths and shear wave velocity profiles would need to be examined to draw generic conclusions.

Finally, the examination of suites of both nonlinear and equivalent linear analyses showed a correlation between the amplification factors and the frequency content of the records. It is therefore expected that consideration of a parameter that describes the frequency content of the records should assist in the selection of records with significantly different characteristics. Although the predominant spectral period mentioned herein is not considered to be the most useful parameter, the existence of various other scalar parameters describing the frequency content of ground motions, such as the mean period  $T_m$  and the average spectral period  $T_{avg}$  (Rathje et al. 2004), could prove particularly useful in selecting records covering a wide rock acceleration and frequency content range.

## ACKNOWLEDGEMENTS

The authors would like to gratefully acknowledge the valuable contribution of Prof. Julian Bommer to this research. The authors are also grateful for the anonymous reviews of this paper which helped to improve the clarity and completeness of the manuscripts considerably. This work has been funded by Engineering and Physical Sciences Research Council (EPSRC) and Bodossaki Foundation.

## References

- Abrahamson NA, Silva W J (1997) Empirical response spectra attenuation relations for shallow crustal earthquakes. *Seismol Res Lett* 68:94- 127
- Bazzurro P, Cornell CA (2004a) Ground-motion amplification in nonlinear soil site with uncertain properties. *Bull. Seismol Soc Am* 94(6):2090-2109
- Bazzurro P, Cornell CA (2004b) Nonlinear soil-site effects in probabilistic seismic-hazard analysis. *Bull. Seismol Soc Am* 94(6):2110-2123
- Baturay MB, Stewart JP (2003) Uncertainty and bias in ground-motion estimates from ground response analyses. *Bull Seismol Soc Am* 93(5):2025-2042

- Bommer JJ, Acevedo AB (2004) The use of real earthquake accelerograms as input to dynamic analysis. *J Earthq Eng* 8(1):43-91.
- Darendeli MB (2001) Development of a new family of normalised modulus reduction and material damping curves. Dissertation, University of Texas
- Field EH, Jordan TH, Cornell CA (2003) OpenSHA: a developing community-modeling environment for seismic hazard analysis. *Seismol Res Lett* 74(4):406-419.
- Geomotions (2007) D-MOD2000 - a computer program package for seismic response analysis of horizontally layered soil deposits, earthfill dams, and solid waste landfills. Geomotions LLC, Washington, User's Manual.
- Goulet AC (2008) Improving the characterization of seismic hazard for performance-based earthquake engineering design. Dissertation, University of California
- Goulet AC, Watson-Lamprey J, Baker JW, Haselton CB, Luco N (2008) Assessment of ground motion selection and modification (GMSM) methods for non-linear dynamic analyses of structures. *Geotechnical Earthquake Engineering and Soil Dynamics IV*, Sacramento, California
- Goulet CA, Stewart JP (2009) Pitfalls of deterministic application of nonlinear site factors in probabilistic assessment of ground motions. *Earthq Spectra* 25(3):541-555
- Iervolino IG, Maddaloni G, Cosenza E (2008) Eurocode 8 compliant real record sets for seismic analysis of structures. *J Earthq Eng* 12:54-90.
- Jayaram N, Lin T, Baker JW (2011) A computationally efficient ground-motion selection algorithm for matching a target response spectrum mean and variance. *Earthq Spectra* 27(3):797-815
- Katsanos EI, Sextos AG, Manolis GD (2010) Selection of earthquake ground motion records: a state-of-the-art review from a structural engineering perspective. *Soil Dyn Earthq Eng* 30:157-169.
- Kottke A, Rathje EM (2008) A semi-automated procedure for selecting and scaling recorded earthquake motions for dynamic analysis. *Earthq Spectra* 24(4):911-932.
- Lessi-Cheimarios A, Kontoe S, Stafford (2012) Comparison of the Total Uncertainty Associated With Alternative Approaches to Site Response Analysis. 15<sup>th</sup> World Conference on Earthquake Engineering. Lisbon, Portugal.
- Matasovic N, Vucetic M (1993) Cyclic characterization of liquefiable sands. *J Geotech Eng* 119(11):1805-1822
- Naeim F, Alimoradi A, Pezeshk S (2004) Selection and scaling of ground motion time histories for structural design using genetic algorithms. *Earthq Spectra* 20(2):413-426
- Papaspiliou M, Kontoe S, Bommer JJ (2012a) On the incorporation of site response into PSHA; part I: issues on the performance of site response analysis. *Soil Dyn Earthq Eng*, 42:302-315.
- Papaspiliou M, Kontoe S, Bommer JJ (2012b) On the incorporation of site response into PSHA; part II: impact on the surface hazard curve. *Soil Dyn Earthq Eng*, 42: 316-330.
- Power M, Chiou B, Abrahamson NA, Bozorgnia Y, Shantz T, Roblee C (2008) An overview of the NGA project. *Earthq Spectra* 24(1):3-21.
- Rathje, E. M., Faraj, F., Russell, S., Bray, J. (2004) Empirical relationships for frequency content parameters of earthquake ground motions. *Earthq Spectra* 20(1):119-144.
- Stewart, J.P. and Kwok, A.O. (2008). "Nonlinear seismic ground response analysis: code usage protocols and verification against vertical array data," in *Geotechnical Engineering and Soil Dynamics IV*, May 18-22, 2008, Sacramento, CA, ASCE Geotechnical Special Publication No.181, D. Zeng, M.T. Manzari, and D.R. Hiltunen (eds.), 24 pages (electronic file)
- Tönük, G, Ansal, A (2010) Selection and scaling of ground motion records for site response analysis. 14<sup>th</sup> European Conference on Earthquake Engineering, Ohrid, FYROM.

## Figure Captions

**Fig. 1** Site stratigraphy, shear-wave velocity profile (after Kottke, 2006) and acceleration response spectra of the 120 rock records used in the ground response analyses

**Fig. 2** Darendeli (2001) shear modulus degradation and damping ratio curves and fit of the MKZ model for the top layer of the investigated site

**Fig. 3** Variability of the median amplification function for suites of 10, 15 and 20 records for  $T=0.01, 0.2$  and  $1.0s$  (equivalent linear analysis). Solid black line represents the median  $AF(f)$  for the 120-record suite and the dashed black lines the  $\pm$  one-standard deviation curves

**Fig. 4** Amplification obtained by each record in the dataset for  $T=0.2s$  highlighting (a) the records used in suite No 7 and (b) the records used in suite No 8

**Fig. 5** Rock (solid lines) and surface (dashed lines) acceleration response spectra for records of suites N° 7 and N° 8

**Fig. 6** Variability of the median amplification function following the nonlinear site response analysis for suites of 10, 15 and 20 records and  $T=0.01, 0.2$  and  $1.0s$ . Solid black line represents the median  $AF(f)$  for the 120-record suite and the dashed black lines the  $\pm$  one-standard deviation curves

**Fig. 7** (a) Amplification obtained by each record in the dataset for  $T=0.2s$  highlighting the 10 records used in suite N° 11, (b) residuals of records for  $T=0.2s$  highlighting those used in suite N° 11, (c) amplification obtained by each record in the dataset for  $T=0.2s$  highlighting the 20 records used in suite N° 11 and (d) residuals of records for  $T=0.2s$  showing in grey those of the 20-record suite

**Fig. 8** Variability of the standard deviation,  $\sigma_{\ln AF(f)}$ , for the different suites of 10, 15 and 20 records following the equivalent linear analysis.

**Fig. 9** Variability of the standard deviation,  $\sigma_{\ln AF(f)}$ , for the different suites of 10, 15 and 20 records following the nonlinear analysis (regression analysis did not converge for the 15-record suites N° 9 and N° 10 at  $3.0s$ ).

**Fig. 10** Site stratigraphy and shear-wave velocity profile adopted for the clayey site.

**Fig. 11** Variability of the median amplification function for suites of 10, 15 and 20 records for  $T=0.01, 0.2$  and  $1.0s$  for the nonlinear analysis of the clayey site. Solid black line represents the median  $AF(f)$  for the 120-record suite and the dashed black lines the  $\pm$  one-standard deviation curves

**Fig. 12** Variability of the standard deviation,  $\sigma_{\ln AF(f)}$ , for the different suites of 10, 15 and 20 records following the nonlinear analysis of the clayey site.

**Fig. 13** Hazard curves for various spectral periods showing the effect of the use of different 10-record suites for the estimation of the site amplification using equivalent linear site response analysis and its incorporation in PSHA. Black curve is obtained from the use of the entire 120-record dataset

**Fig. 14** (a) Variation of median  $S_a^s(0.2s)$  for the different 10-record suites of equivalent linear analysis (coloured lines) for a magnitude 7.7 event; (b) variation of the standard deviation  $\sigma_{\ln S_a^s(0.2s)}$  for the same magnitude event and suites of records. Black line represents the estimate for the entire 120-record dataset

**Fig. 15** Hazard curves for various spectral periods showing the effect of the use of different 20-record suites for the estimation of the site amplification using equivalent linear site response analysis and its incorporation in PSHA. Black curve is obtained from the use of the entire 120-record dataset

**Fig. 16** Hazard curves for various spectral periods showing the effect of the use of different 10-record suites for the estimation of the site amplification using nonlinear site response analysis and its incorporation in PSHA. Black curve is obtained from the use of the entire 120-record dataset

**Fig. 17** Hazard curves for various spectral periods showing the effect of the use of different 20-record suites for the estimation of the site amplification using nonlinear site response analysis and its incorporation in PSHA. Black curve is obtained from the use of the entire 120-record dataset

Fig1

[Click here to download colour figure: Fig1.eps](#)

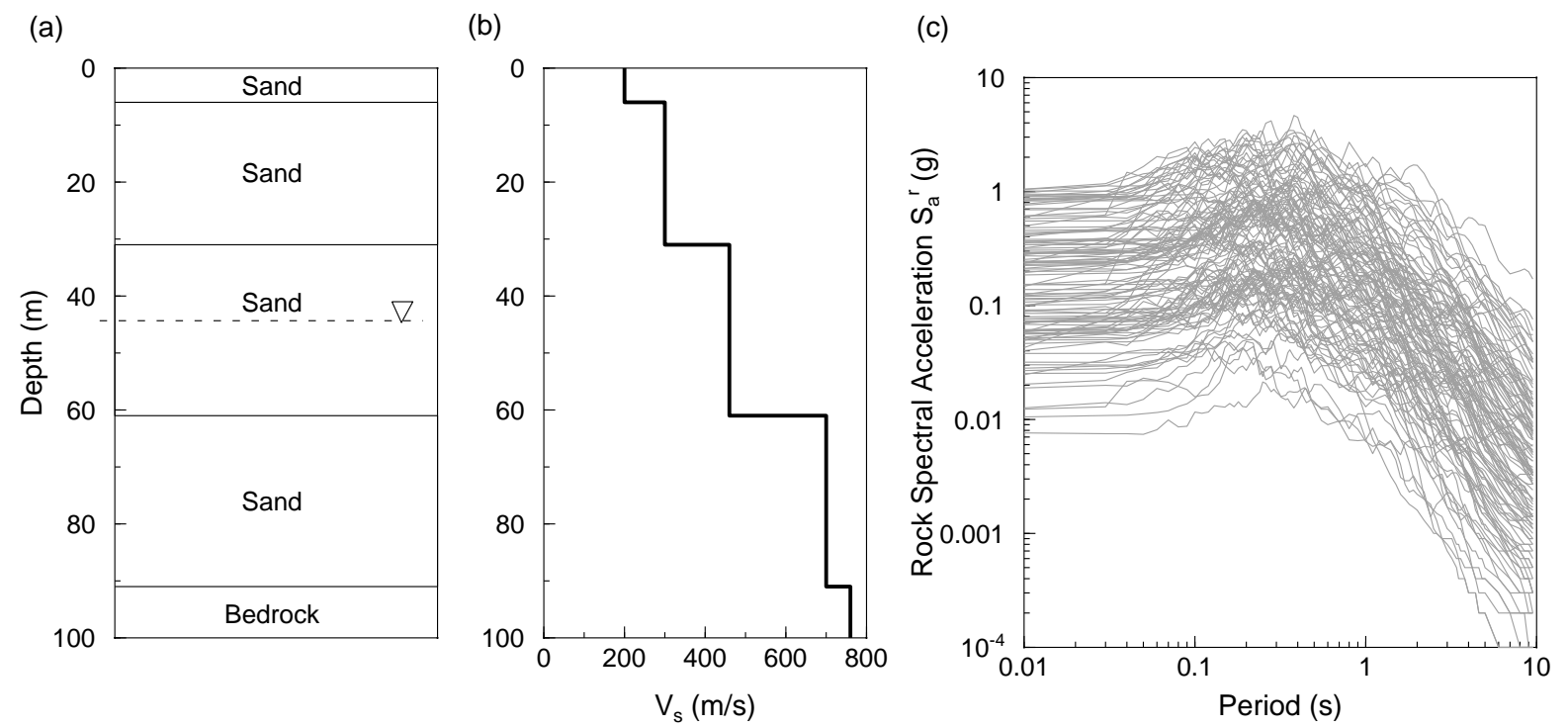
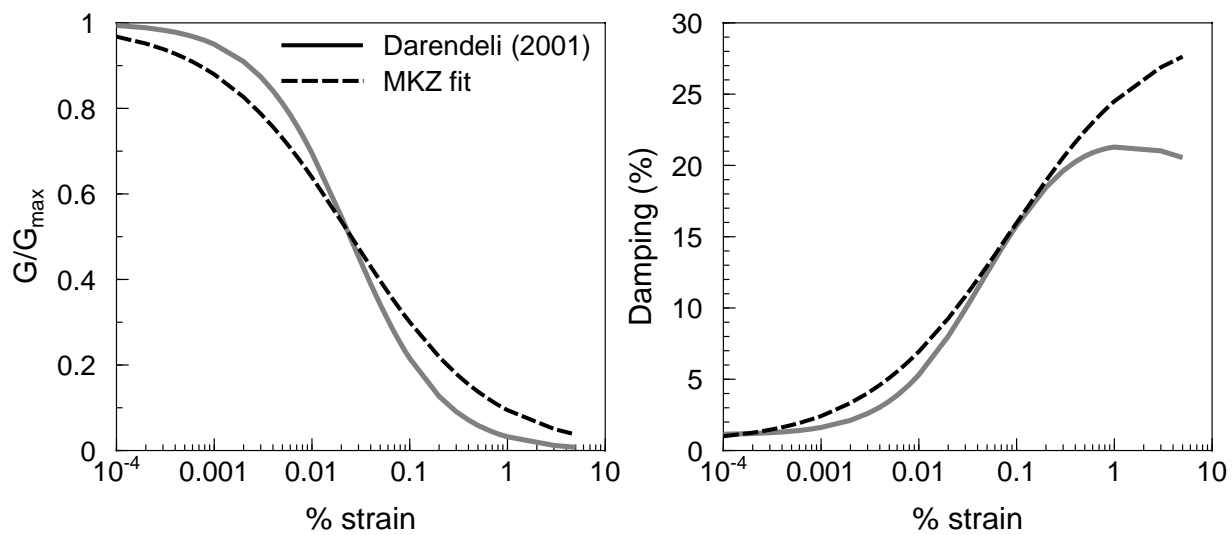


Fig2

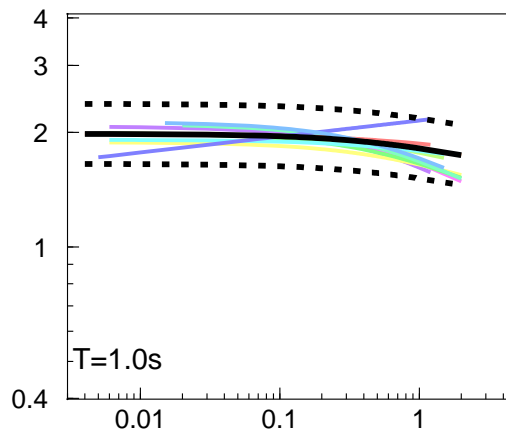
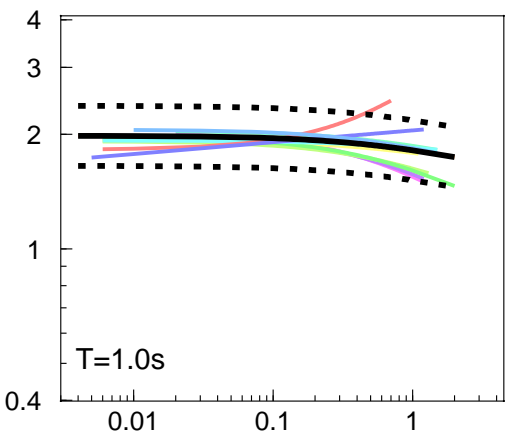
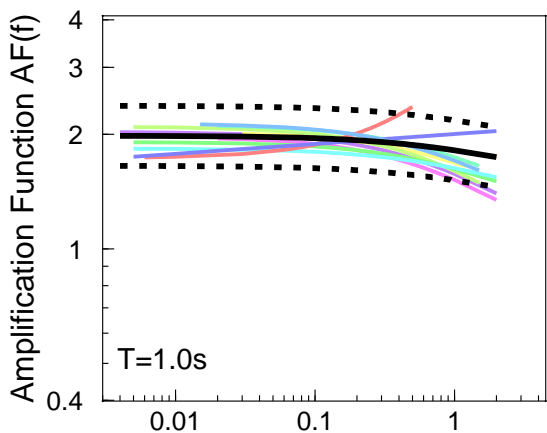
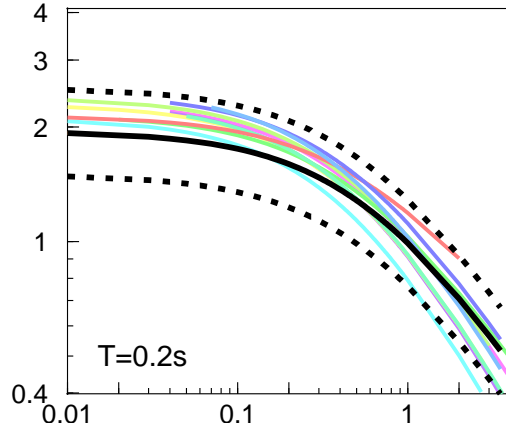
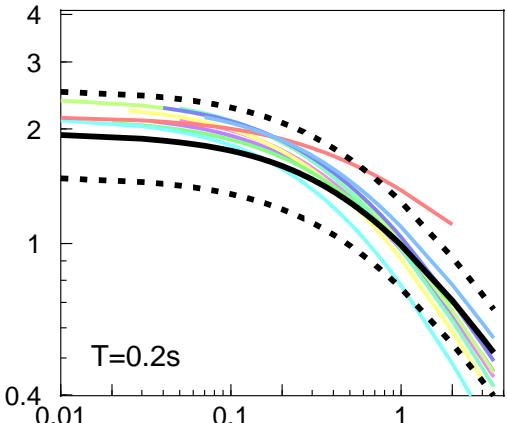
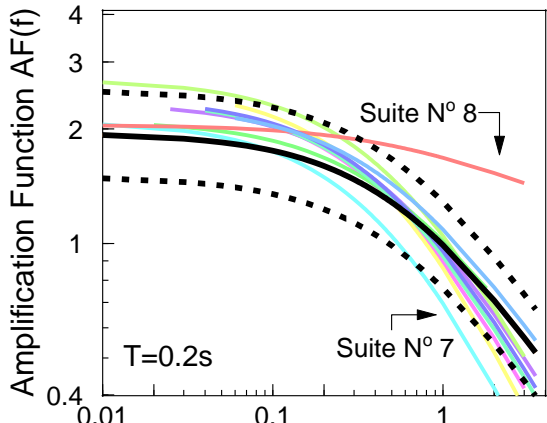
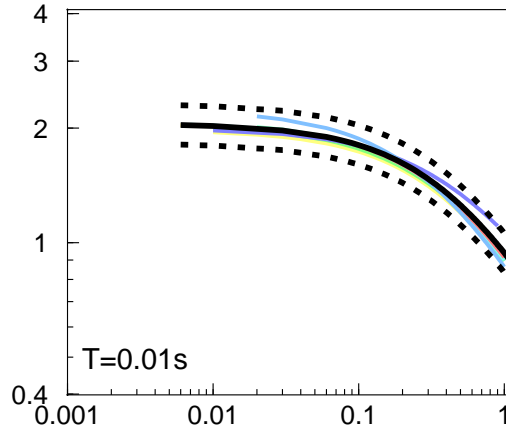
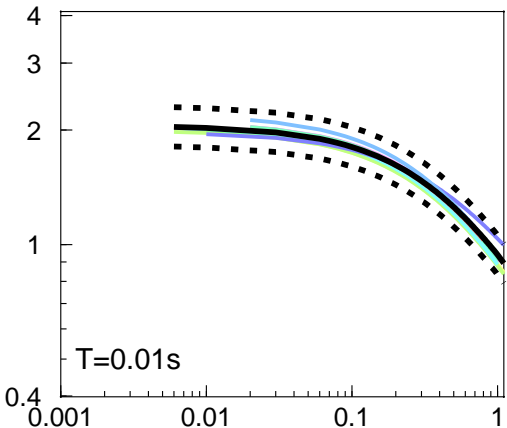
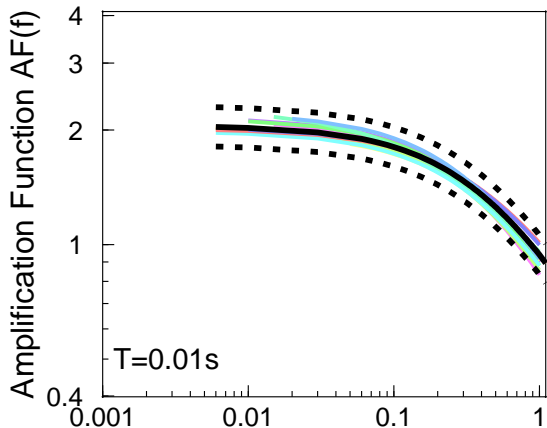
[Click here to download colour figure: Fig2.eps](#)



10-record suites

15-record suites

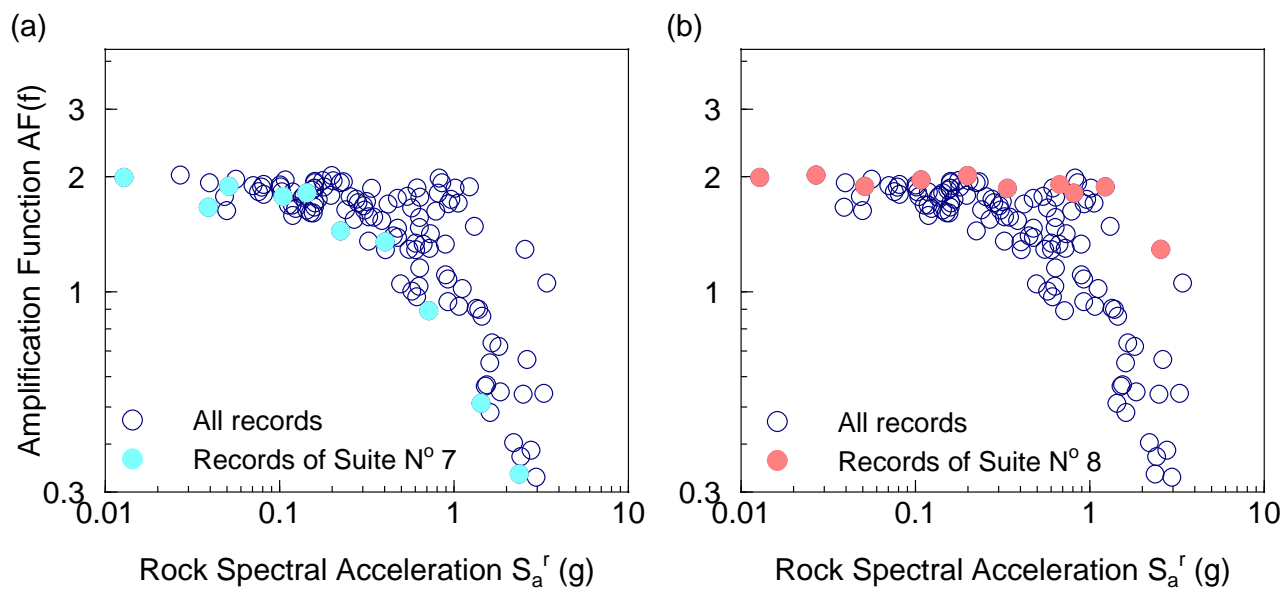
20-record suites



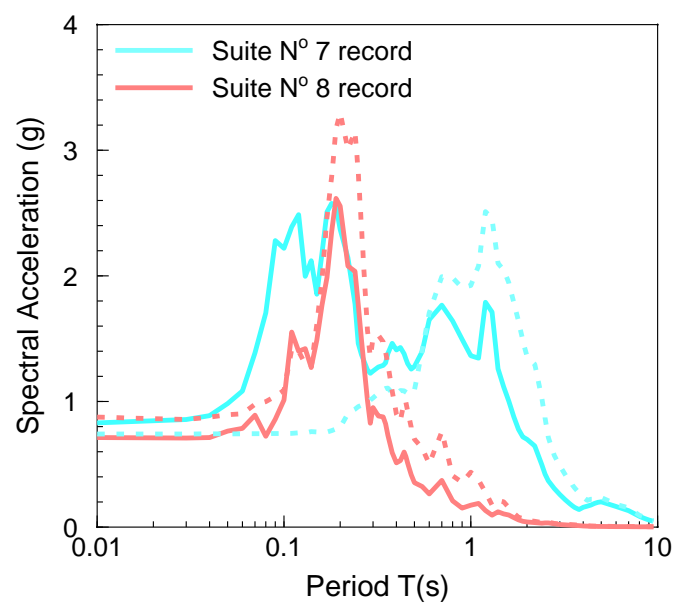
Rock Spectral Acceleration  $S_a^r$  (g)

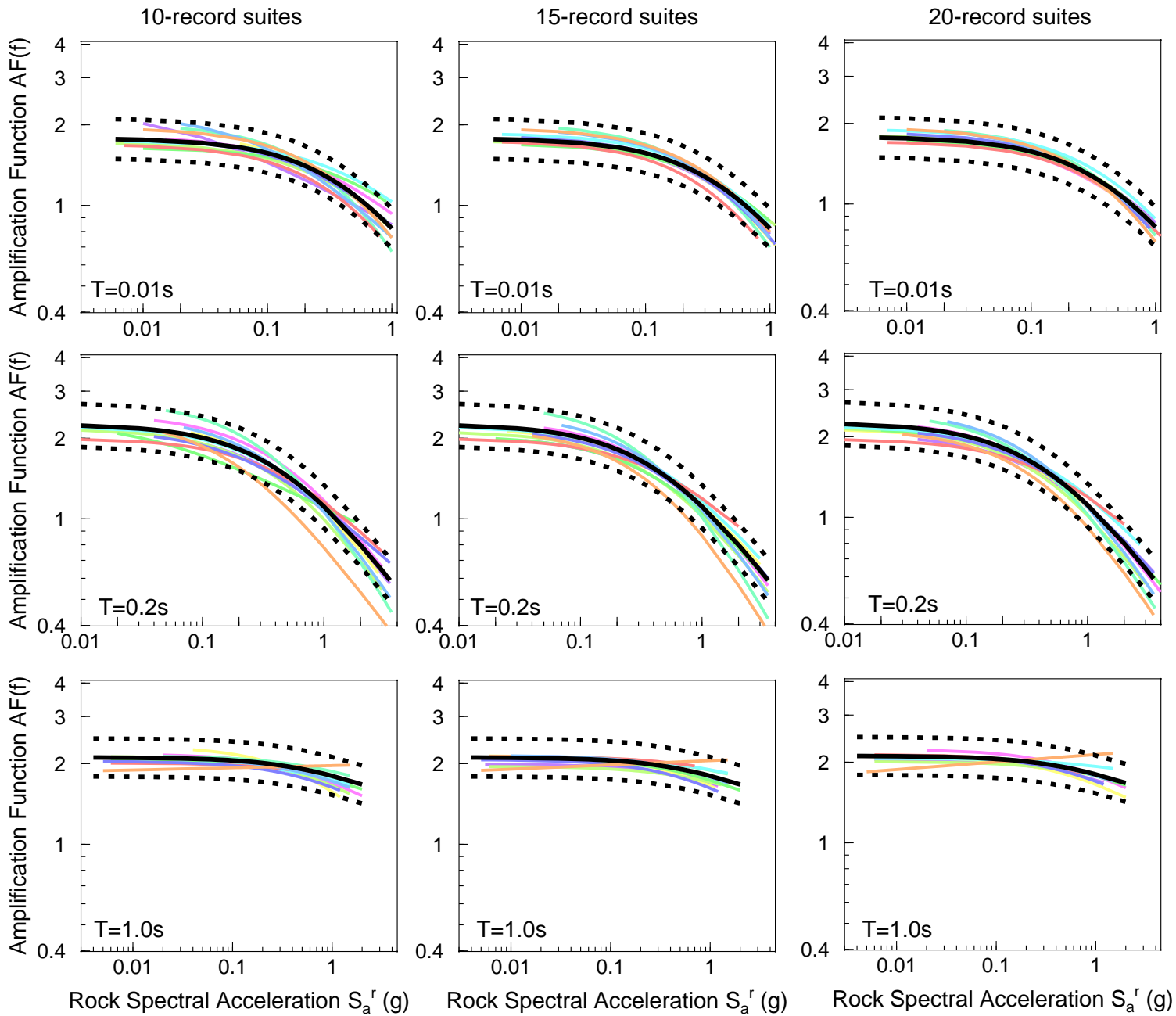
Rock Spectral Acceleration  $S_a^r$  (g)

Rock Spectral Acceleration  $S_a^r$  (g)









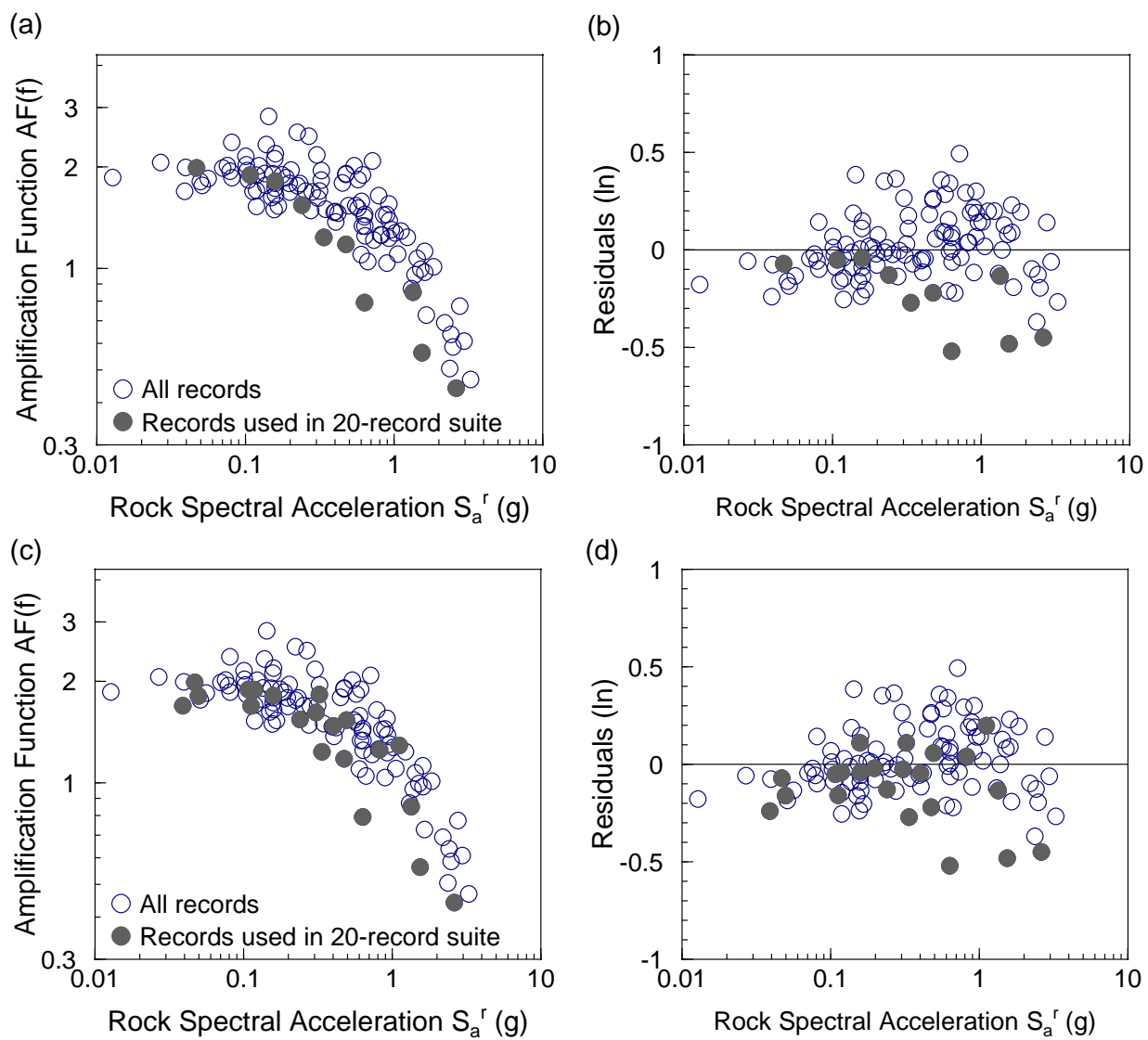


Fig8

[Click here to download colour figure: Fig8.eps](#)

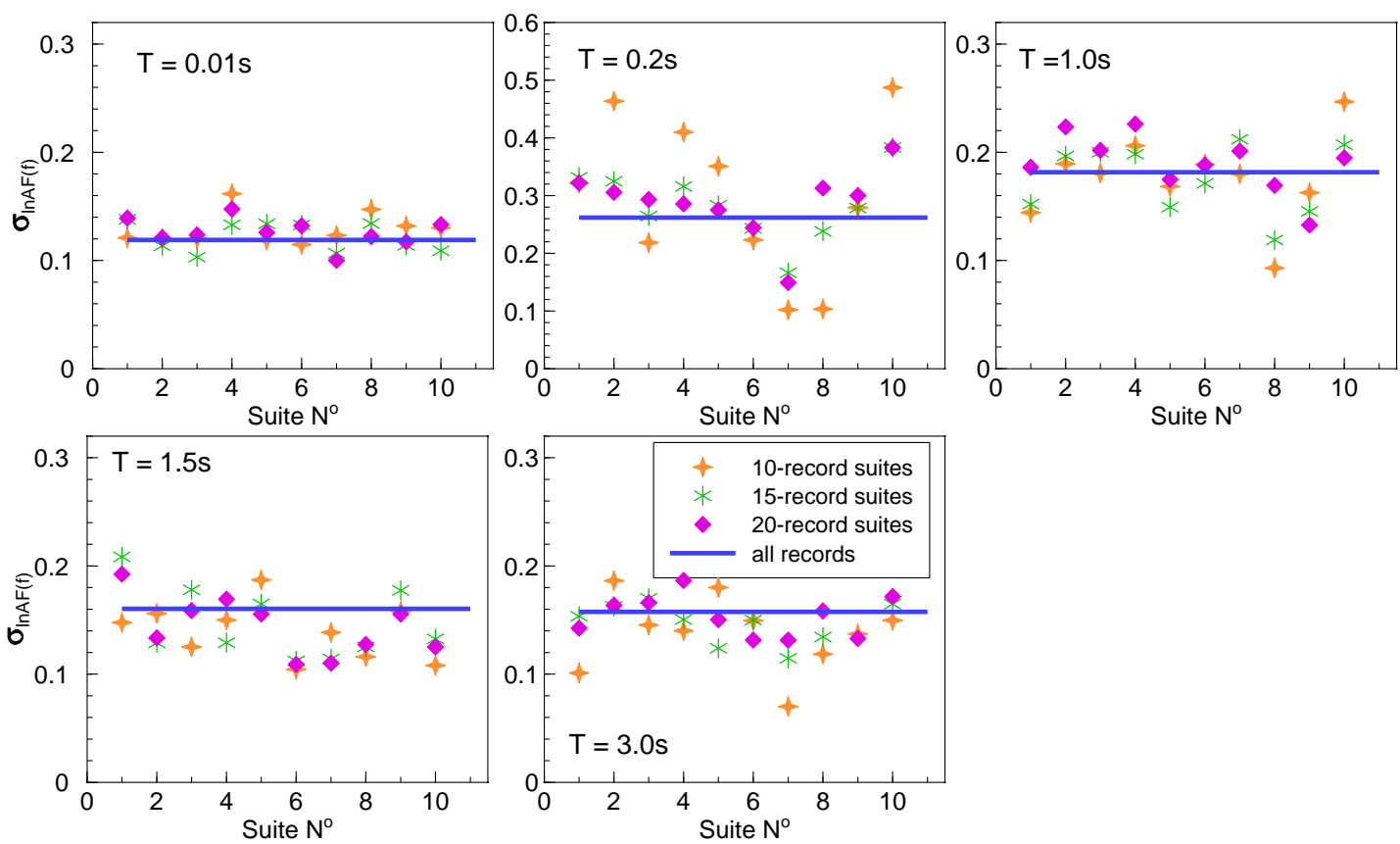


Fig9

[Click here to download colour figure: Fig9.eps](#)

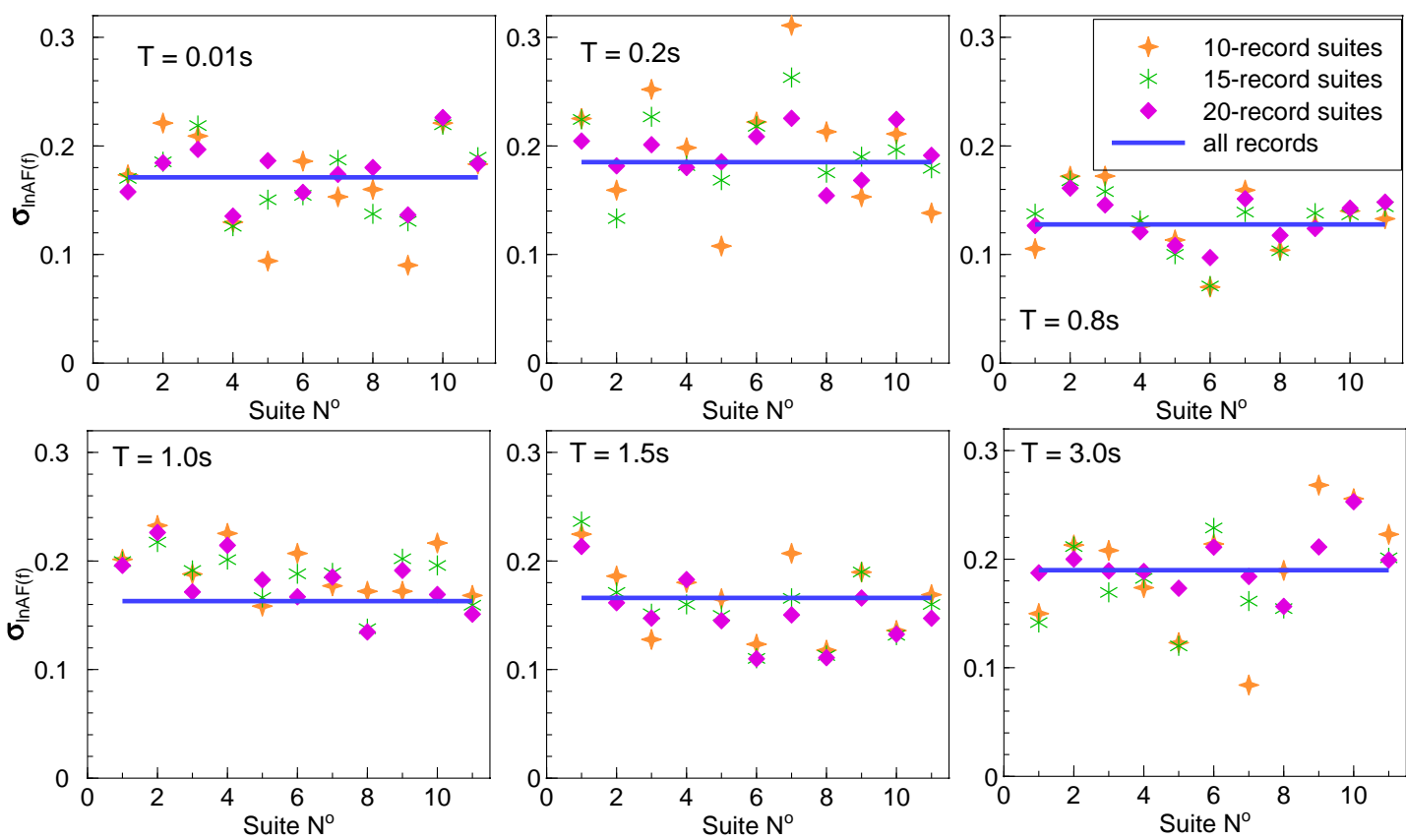


Fig10

[Click here to download colour figure: Fig10.eps](#)

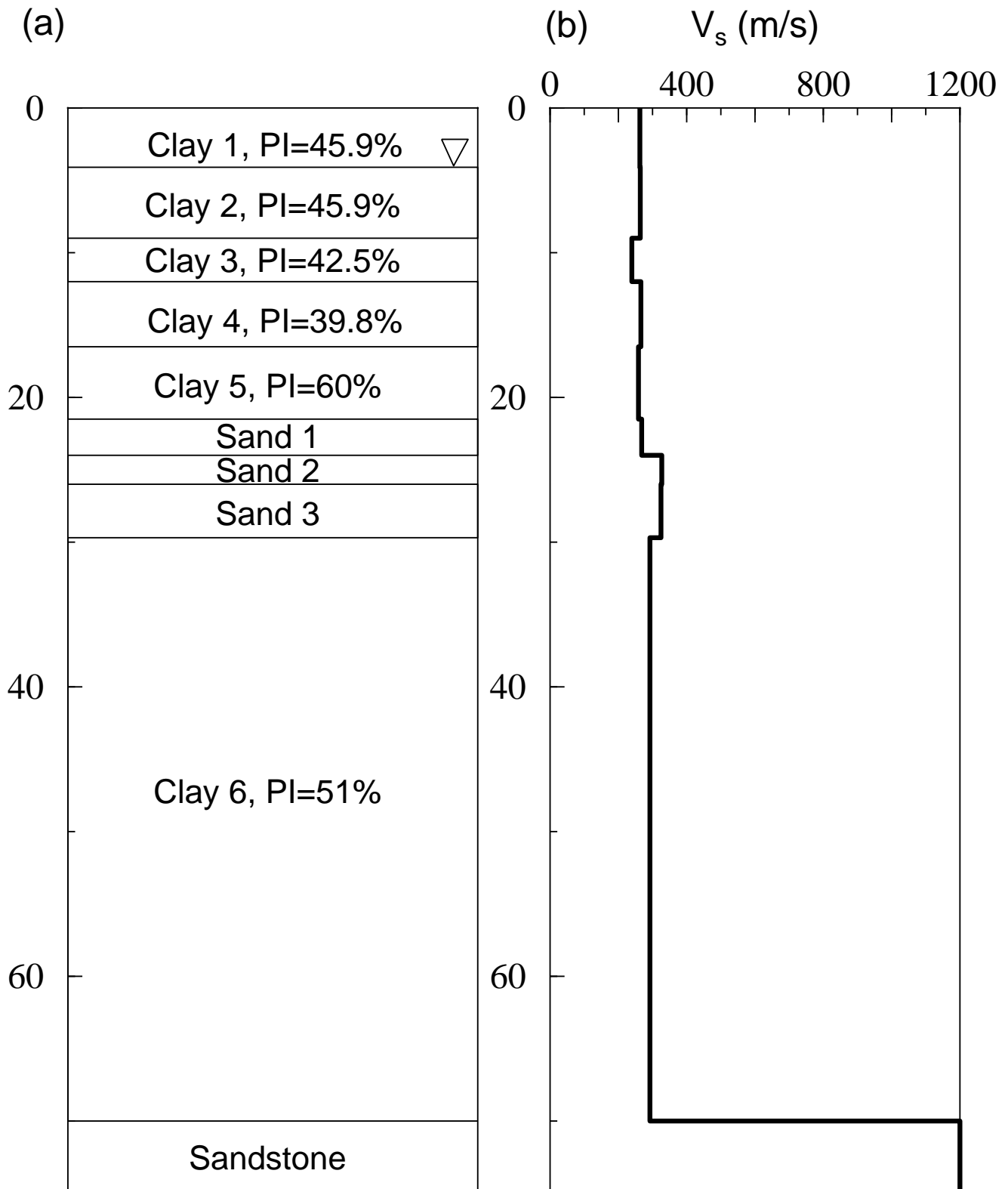


Fig11

[Click here to download colour figure: Fig11.eps](#)

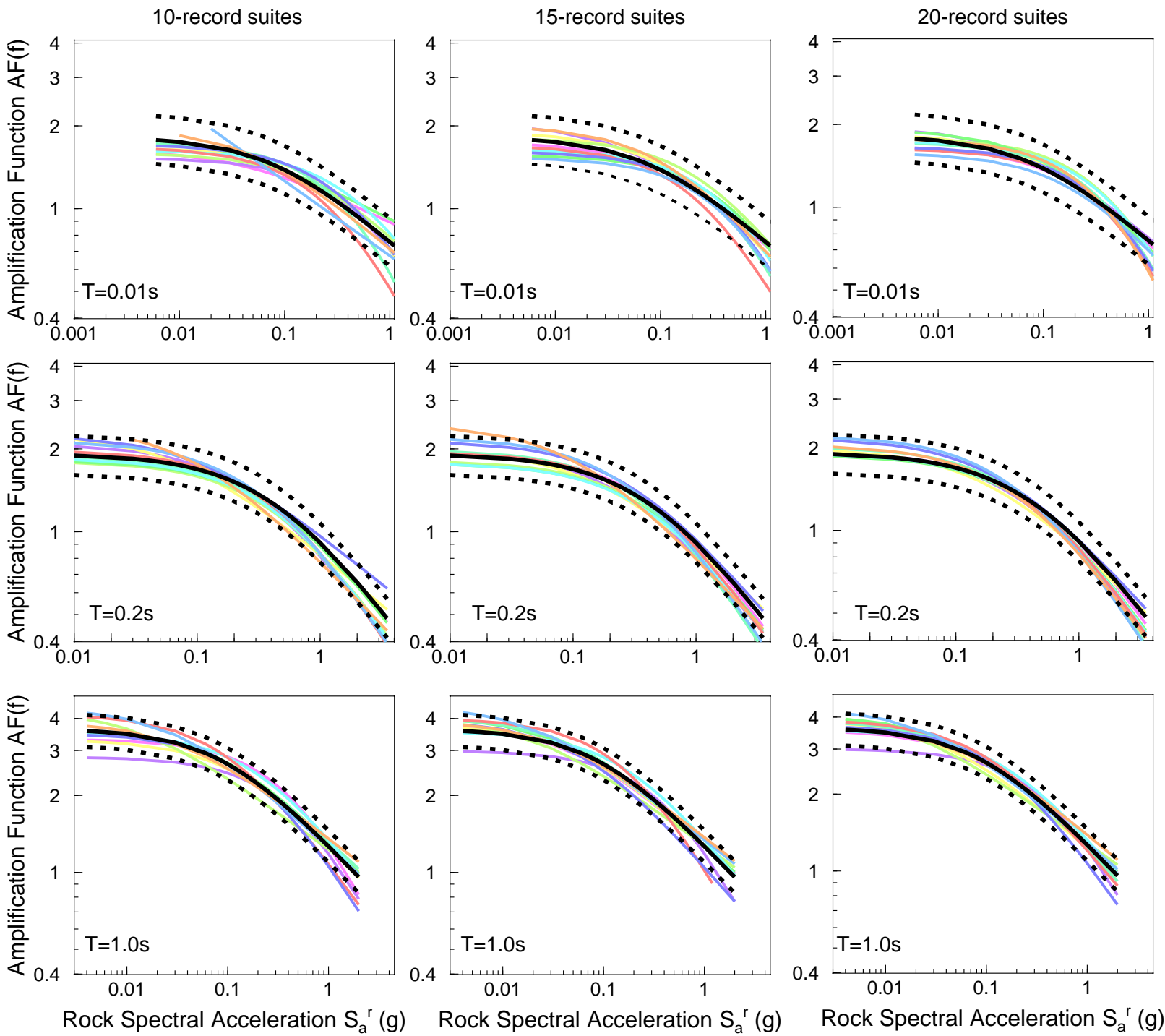


Fig12

[Click here to download colour figure: Fig12.eps](#)

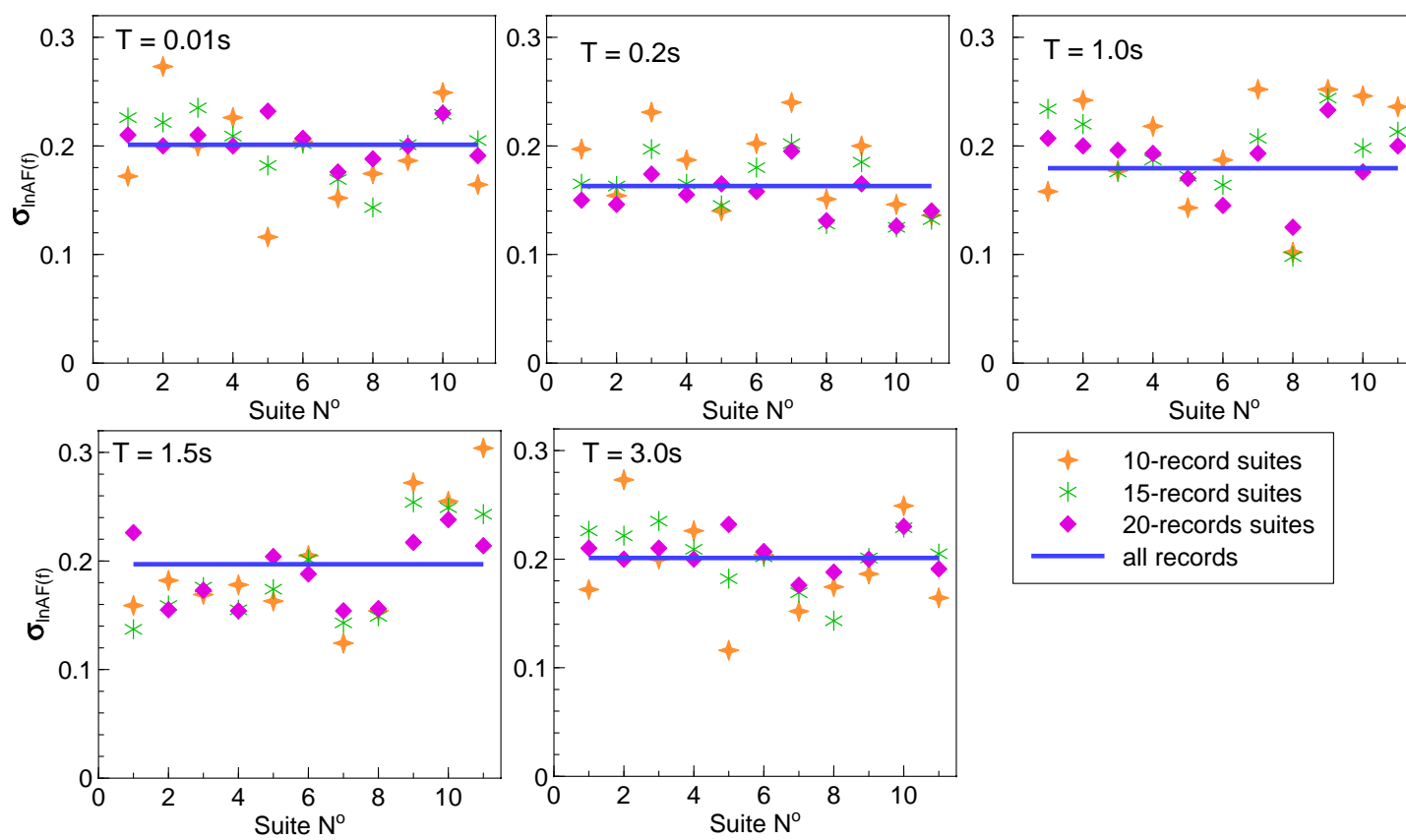
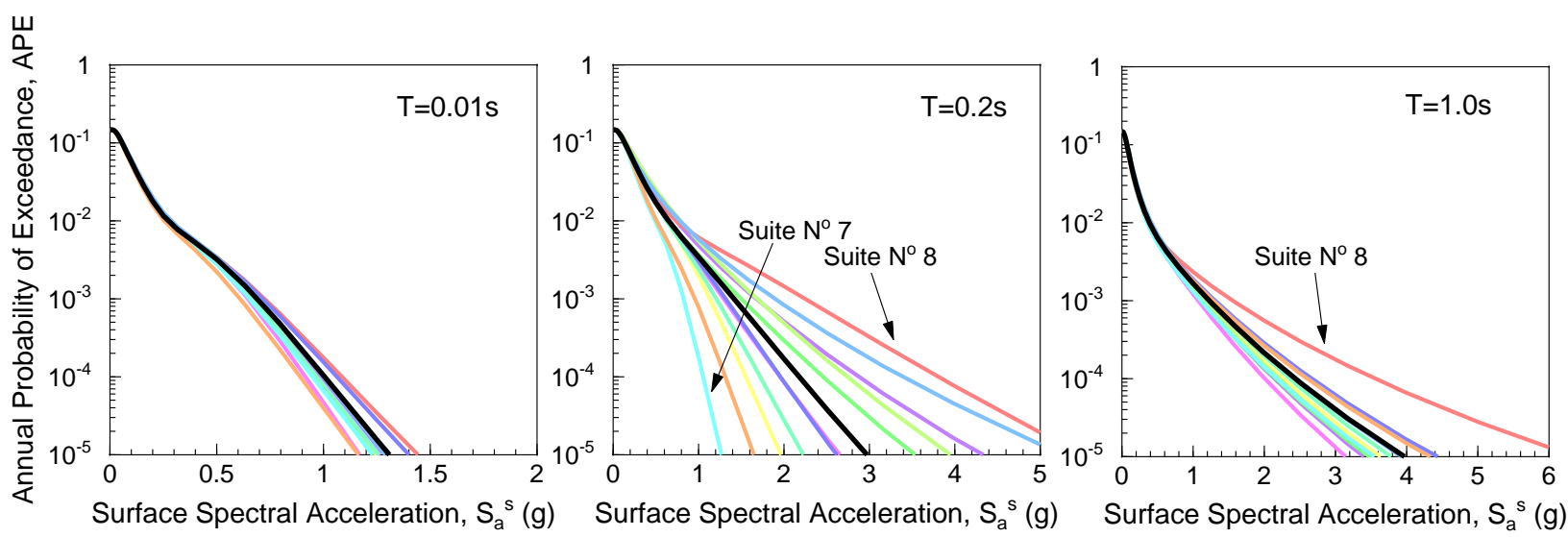




Fig13

[Click here to download colour figure: Fig13.eps](#)



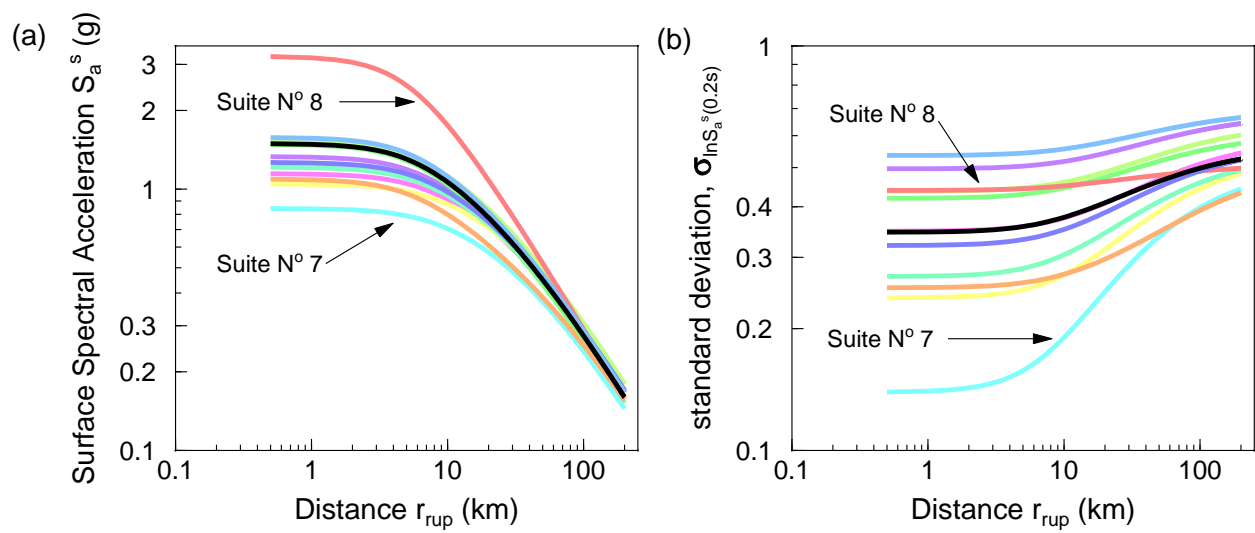


Fig15

[Click here to download colour figure: Fig15.eps](#)

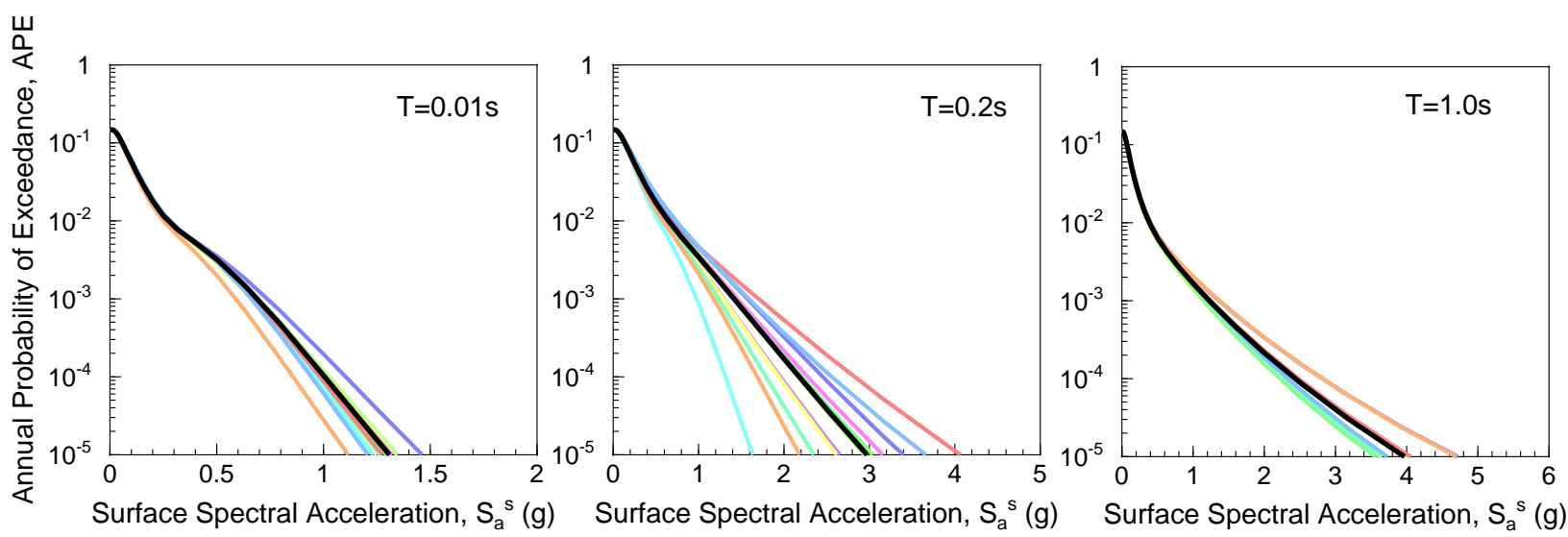


Fig16

[Click here to download colour figure: Fig16.eps](#)

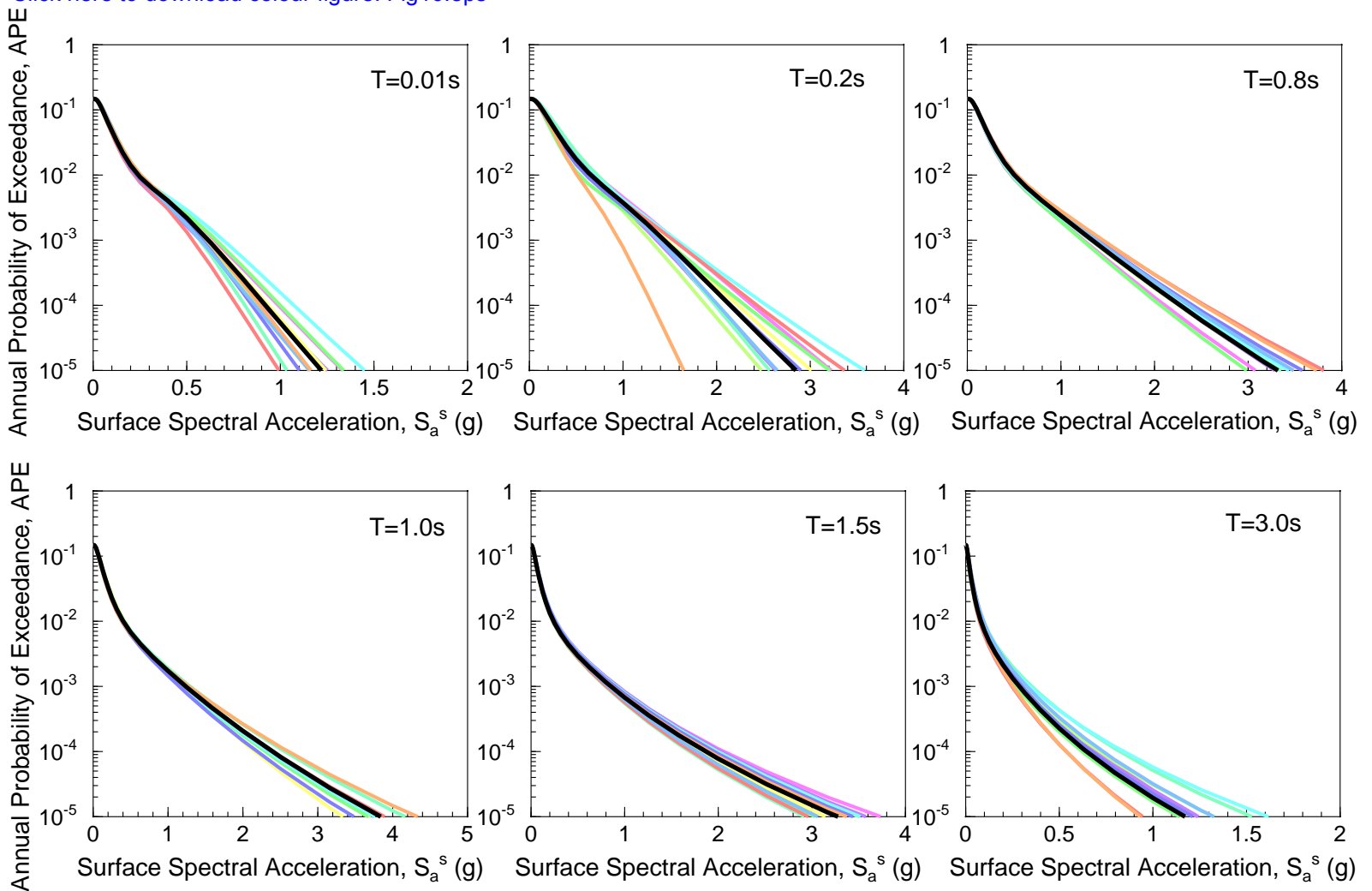


Fig17

[Click here to download colour figure: Fig17.eps](#)

