NON-STATIONARY FREQUENCY CONTENT CHARACTERIZATION OF SEISMIC ACCELEROGRAMS VIA THE SLOPE "ALPHA" OF THE WAVELET-BASED MEAN INSTANTANEOUS PERIOD

A. Margnelli[1], A. Giaralis[2]

[1] *PhD Candidate, Dept. of Civil Engineering, City University London; Associate Director, AKT-II, alessandro.margnelli@akt-uk.com*
[2] *Senior Lecturer, Dept. of Civil Engineering, City University London, agathoklis@city.ac.uk*

Abstract

Typical recorded acceleration traces of seismic ground motions (GMs) exhibit a time-varying frequency composition, however, all GM properties currently considered in earthquake engineering to quantify the structural damage potential of GMs do not explicitly characterize the time-evolving trends of GM frequency content. Indeed, the intensity measures (IMs) and record selection criteria adopted in the context of performance-based earthquake engineering (PBEE) do not account for the non-stationary frequency content of GMs in a straightforward manner. In this regard, this paper considers a novel scalar quantity termed alpha, "a", defined by the average slope (angle) of the wavelet-based mean instantaneous period (MIP) to characterize the temporal evolution of the mean frequency content of recorded GMs. Specifically, the MIP is the time-varying first-order average along the frequency or, equivalently, along the period axis of the wavelet-based GM spectrogram (squared magnitude of the GM wavelet transform) treated as a GM energy distribution on the time-frequency/period plane. Consequently, a captures the rate by which the mean frequency content of GMs changes in time. Linear regression analyses are undertaken involving a dataset of 684 GMs from 30 seismic events of Magnitude 6.5<M<8 and distance to rupture plane 20km<R_{rup}<120km to quantify statistical/empirical correlation trends between a and well-established GM properties, namely the peak ground acceleration (PGA), peak ground velocity (PGV), and the mean Fourier-based frequency (T_m), oftentimes used as intensity measures (IMs) and record selection criteria in PBEE. Further, regression analyses are undertaken to probe into the relationships between a and important seismological and local site characteristics, namely M, R_{rup}, and the shear wave velocity V_s30. It is found that no significant correlation exist between a and M, Rrup, or PGA. However, it was established that a is well-correlated with the average frequency content of GMs as captured by T_m and by PGV: the lower the average frequency content, the larger the value of a tends to be, that is, the faster the time evolution (transition) of the average frequency content is from higher to lower frequencies. Further, the reported numerical data indicate that the level of the above correlation depends on the intensity of GMs in terms of PGA (conditional on PGA): a is larger for fixed T_m as PGA increases and a increases faster as T_m increases for larger PGA values. Moreover, GMs recorded on softer soils are more likely to have larger a conditional on PGA, a phenomenon that is attributed to the fact that soft soils under strong seismic shaking exhibit strong non-linear behavior that enriches the low frequency content of free field recorded GMs. Lastly, the influence of a in predicting the peak inelastic structural response is assessed within the PBEE framework through a standard sufficiency statistical test on PGA and PGV, treated as non-structure specific IMs in conducting incremental dynamic analysis (IDA) for a hysteretic oscillator with strength and stiffness degradation representing a benchmark 12-storey reinforce concrete frame exposed to the above set of 684 GMs. The considered statistical tests demonstrate that the non-stationary average frequency content of GMs as captured by a influences the peak inelastic structural response at collapse. Overall, the herein furnished results establish the validity and usefulness of a in characterizing the evolutionary frequency content of GMs and suggests that a should be considered as a record selection criterion in undertaking IDA using PGA and PGV as IMs.

Keywords: wavelet analysis, non-stationary frequency content, mean instantaneous period, sufficiency of intensity measure

1. Introduction

Typical field recorded acceleration time-histories of earthquake induced strong ground motions (GMs) exhibit time-evolving amplitude and frequency content due to the dispersion of different types of propagating seismic waves arriving at the recording station at different time instants. Several researchers considered various joint time-frequency signal analysis tools, including the wavelet transform, to capture the time-evolving characteristics of the frequency content of GMs [1-4]. Further, such tools have been used for GM classification associated with near-fault signatures such as forward-directivity pulses [5] and for stochastic simulation of GMs using non-stationary stochastic models [4,6]. More importantly, it has been argued that the time-evolving frequency content of GMs influences the hysteretic response of seismically excited yielding structures. This issue has been studied from the structural dynamics and from the damage detection perspectives by means of analytical stochastic dynamics techniques (e.g. [7]), signal analysis tools used in a deterministic manner [1-3], and simulation-based studies considering stochastic models to represent the non-stationary features of the strong ground motion [8].

Still, the evolving frequency content of GMs is not taken into account by any of the commonly-used scalar parameters to characterize the structural damage potential of GMs (see e.g. [9]). For instance, recent research work focusing on the effects of frequency content of GMs to the structural response of yielding structures [10,11] consider the Fourier transform-based mean period T_m in [12] for GM characterization and classification which provides only for the averaged in time spectral composition of the GMs; it does not carry any information on its time-varying trends. Accordingly, none of the non-structural intensity measures (IMs), such as the peak ground acceleration (PGA), and the peak ground velocity (PGV), or of the GM record selection criteria, extensively used to facilitate seismic risk quantification within the performance-based earthquake engineering (PBEE) framework, consider explicitly the evolution of GM frequency content. Further, the structure-dependent IMs used in PBEE, such as the spectral acceleration at the fundamental natural period, $Sa(T_1)$, (e.g., [13-15]) account for the non-stationary GM frequency content only implicitly as it reflects to elastic or inelastic structural response. Nevertheless, this does not allow for examining the influence of the salient non-stationary frequency trends of different GMs to the response of yielding structures in a systematic manner, either from a structural dynamics viewpoint (e.g., identification of "moving resonance" phenomena [7] or "period elongation" phenomena [20]) or from a purely statistical viewpoint (e.g., peak inelastic response variability as obtained from incremental dynamic analysis- IDA [13]), and, therefore, to gauge its contribution to the seismic risk in the context of PBEE.

In addressing some of the above open issues, the authors recently demonstrated that the time-varying mean instantaneous period (MIP), computed from wavelet-based joint time-period representations of ensembles of recorded GMs, correlates well with the MIP of acceleration response time-histories of inelastic single-degree-of-freedom (SDOF) oscillators near collapse [16]. Note that the MIP is a time-history on the time-period plane tracking the variation of the mean signal frequency content in time by averaging along the frequency axis the signal wavelet coefficients at each time instant. In this respect, MIP is closely related to the mean instantaneous frequency (MIF) whose usefulness for treating GMs and inelastic structural response signals was established in [2,3]. In particular, it was found in [16] that the MIPs of the response signals tend to converge to the GM MIP in a point-wise manner as stronger inelastic behavior is exhibited in conducting IDA for a particular ensemble of far-field GMs. Furthermore, it was observed that the *slope* of the ensemble average *response* MIP for the near-collapse limit state lies close to the ensemble average *excitation* (GM) MIP.

The above findings motivated the herein work focusing on the average slope of the MIP of recorded GMs, hereafter termed angle "alpha" *a*, and aiming to quantify its relationship with well-established seismological parameters and widely used non-structure dependent IMs in a statistical sense. This is achieved by means of standard linear regression analyses pertaining to a large databank of far-field GMs. The remainder of this paper is structured as follows: Section 2 provides for a brief mathematical background and definitions of MIP and *a* along with their properties; Section 3 furnishes statistical data and comments on the relation of *a* to seismological parameters (M, R_{rup}), soil conditions (shear wave velocity V_s30), PGA, and PGV; Section 4 considers testing the most widely used non-structure IMs PGA and PGV for sufficiency with respect to *a* in undertaking IDA for a hysteretic bilinear oscillator with strength and stiffness degradation calibrated ; and Section 5 summarizes conclusions and points to future research directions.

2. Mathematical Background and Definitions

2.1 Fourier transform-based mean period T_m

Consider the continuous-time acceleration GM trace x(t) of finite duration T_o and its discrete-time version $x[n]=$ $x(n\cdot\Delta t)$ with $n=0,1,2,\ldots,N-1$ sampled with a time step Δt, such that $T_o= N\cdot\Delta t$. The discrete Fourier transform (DFT), defined as

$$\hat{X}[k] = \frac{1}{N}\sum_{n=0}^{N-1} x[n]e^{-i\omega_k n} \tag{1}$$

decomposes/projects $x[n]$ onto an orthogonal basis of harmonically related discrete-time sinusoidal functions with frequencies $\omega_k= 2\pi k/N$ (in rad/s). In Eq.(1) i denotes the imaginary unit. The magnitude of the complex-valued Fourier coefficients $\left|\hat{X}[k]\right|$ can be viewed as a measure of similarity between the signal $x(t)$ and a non-decaying in time sinusoidal function of frequency ω_k (i.e., single frequency component). Therefore, the magnitude of the DFT achieves a representation of the average over time frequency composition of $x(t)$ on the frequency ω axis or, equivalently, on the period $T=2\pi/\omega$ axis, with the sharpest possible resolution. Further, the DFT is energy preserving (i.e., $\left(x[n]\right)^2 \propto \left|\hat{X}[k]\right|^2$) and, therefore, the square magnitude of the Fourier coefficients can be interpreted as a distribution of the signal energy in the frequency domain. This interpretation allows for defining the Fourier-based mean period within the frequency band [0.25 20] Hz of interest for GMs [26]

$$T_m = \frac{\sum_{k=K_1}^{K_2}\left|\hat{X}[k]\right|^2 \dfrac{2\pi}{\omega_k}}{\sum_{k=K_1}^{K_2}\left|\hat{X}[k]\right|^2} \tag{2}$$

where K_1 and K_2 are the closest integers to $0.25N$ and $20N$, respectively. The mean period T_m has been shown to represent better the (mean) frequency content of GMs (see e.g., [26]) compared to alternative response spectrum-based metrics and it commonly used in the literature to quantify the impact of the mean frequency content of GMs to the response of yielding structures (e.g. [10,11]). However, T_m does not capture the time-varying frequency content of GMs which may significantly influence the inelastic structural response (e.g., [1,7,8]). This is because the magnitude of the DFT does not carry any temporal information (i.e., it is not possible to know the location in time of each identified frequency component). Motivated by the above studies and recognizing that wavelet-based signal representations have been successfully used to represent the time-varying frequency content of both GMs and inelastic structural response signals (e.g., [3,16]), the wavelet-based time-varying mean instantaneous period (MIP) is defined in the next sub-section to characterize the evolution of the mean frequency content of GMs.

2.2 The continuous wavelet transform and the mean instantaneous period (MIP)

The continuous wavelet transform (CWT) of a discrete-time signal $x[n]$ is defined as (e.g., [17])

$$W(s,n) = \sum_{n'=0}^{N-1} x[n']\psi^*\left(\frac{(n'-n)\Delta t}{s}\right) \tag{3}$$

The above transformation can be viewed as a decomposition/projection of $x[n]$ onto a collection of oscillatory localized in time functions ("wavelets"). These functions are generated by scaling, via the positive scale parameter s, and by translating in time a single finite energy function $\psi(t)$ ("mother wavelet"), which may be complex-valued. In the last equation the superscript (*) denotes complex conjugation. Importantly, the scaling operation and the oscillatory form of the wavelets are the salient features that allow for interpreting the square magnitude (spectrum) of the wavelet coefficients $|W(s,n)|^2$ as an estimator of the signal energy distribution on the time-frequency or, equivalently, on the time-period plane. This is because the scale parameter s can be associated with an effective frequency $\omega_{eff}(s)=\omega_p/(\Delta t\cdot s)$ (or equivalently with an effective period $T_{eff}(s)=2\pi/\omega_{eff}(s)$), where ω_p is the predominant frequency of the unscaled (i.e., $s=1$) mother wavelet. To this end, a wavelet-based time-varying mean instantaneous period (MIP) is herein considered defined as (see also [2,3])

$$\text{MIP}[n]=\text{MIP}(n\Delta t) = \frac{\sum_{s=S_1}^{S_2}\left|W(s,n)\right|^2 T_{eff}(s)}{\sum_{s=S_1}^{S_2}\left|W(s,n)\right|^2} \quad for \quad floor\left(\frac{t_{05}}{\Delta t}\right) \le n \le ceil\left(\frac{t_{95}}{\Delta t}\right) \tag{4}$$

band-limited within the same frequency range as the T_m in Eq. (2). That is, in the last equation, S_1 and S_2 are integers specifying scales with effective frequencies of 0.25Hz and 20Hz, respectively. Note that the above MIP is a time-dependent function corresponding to the time frame in which the central 90% of the total signal energy lies. That is, in Eq. (5) t_{05} and t_{95} are the time instants at which 5% and 95% of the total signal energy, respectively, has been released, while the operators *floor(·)* and *ceil(·)* round their arguments to the closest small and large integer value, respectively. This time frame corresponds to the effective duration of GMs [9]. Therefore, the MIP provides an estimate of the temporal evolution of the mean period T_m within the effective GM duration. To this effect, it can be viewed as a generalization of T_m, in the sense that by temporal averaging of MIP, the following *estimation* of the T_m in Eq. (2) can be reached (see also [16])

$$T_m^{MIP} = \frac{1}{N}\sum_{n=0}^{N-1}\text{MIP}[n]. \tag{5}$$

An important concern in the CWT is the choice of the wavelet analysis function in Eq. (3) which is highly application-dependent. For the purposes of this work, Morlet wavelets, which are widely used to derive "smooth" and well-localized in time wavelet spectra $|W(s,n)|^2$ and were historically the first waveforms to be used in conjunction with the CWT [18] are considered. Specifically, the Morlet wavelet in Eq. (3) is defined as

$$\psi\left(\frac{(n'-n)\Delta t}{s}\right) = \sqrt{\frac{\Delta t}{s}}\pi^{-1/4}\exp\left(i\omega_p\frac{(n'-n)\Delta t}{s}\right)\exp\left(-\left(\frac{(n'-n)\Delta t}{4s}\right)^2\right) \tag{6}$$

where the term in front of the first exponential function is included to ensure that Morlet wavelets at all scales have unit energy. For illustration, contour plots of the Morlet wavelet spectrum on the time-period plane are provided in Figs.1(c) and 1(d) (warmer colors indicate higher amplitude values) for two different GMs plotted in Figs. 1(a) and 1(b), respectively, obtained from the PEER NGA-West2 database (http://ngawest2.berkeley.edu/). These wavelet spectra have been generated by considering a sufficient number of discrete scales corresponding to a frequency range of interest in earthquake engineering to achieve a meaningful signal representation as detailed in [17]. On the same plot, the so-called "cone of influence" [17] is indicated by a red broken line outside of which the wavelet spectrum is severely affected by end-effects and, therefore, is not reliable. Lastly, the MIP computed from Eq. (4) is superposed on Figs1(c) and 1(d) (thick broken black line), defined within the white rectangular window on the time-period plane.

2.3 The average slope "alpha" α of the MIP

The above discussed MIP is useful to visualize qualitatively the evolutionary trend of the mean frequency content of GMs, however, being a time-history on the time-period plane, it bears limited practical merit to serve as an index quantifying the evolving frequency content of GMs; a scalar quantity is desirable for the task. For this purpose, it is herein proposed to consider the average slope of the MIP "alpha" a as a scalar that captures the average evolutionary trend in time of the mean frequency content of GMs. Rather than providing an impractical mathematical definition, it is deemed preferable to define α graphically as shown in Figs. 1(e) and 1(f). In particular, in the latter figures the MIP of the GMs in Figs. 1(a) and 1(b) are plotted (blue continuous curves) on the time-period plane as extracted from the corresponding Morlet wavelet spectra. Next, a linear function is fit to the MIP using standard least squares minimization (broken black line), and a is defined as the slope of the linear function fitted to MIP measured in degrees. Therefore, a can be effectively viewed as the average slope of the MIP in time within the effective duration of the GM. The angle a is assigned a positive value when the mean GM frequency decreases in time (on the average) or, equivalently, when the GM period increases in time, as shown in Fig.1(e), while a is assigned a negative value when the mean GM frequency increases in time or, equivalently,

when the GM period decreases in time, as shown in Fig.1(f). Following seismological considerations, it is expected that typical horizontal far-field GMs with no near-fault signatures would normally have a positive *a* since the seismic waves that arrive first at a recording station are dominated by higher frequencies while lower frequencies kick in at later times (e.g., [4,6,9]). Indeed, this is true for the vast majority of the 684 far-field GM components examined in the next section (Table 1). Nevertheless, there are also far-field GMs with *a<0*, as the recorded acceleration treated in the right panels of Fig.1. Careful examination of the Morlet wavelet spectra and the associated MIPs reveal that GMs with negative *a* values are typically characterized by the late appearance of significant bursts of energy carried at slightly higher frequencies compared to the mean frequency content at the beginning of the GM (see also Figs. 1(c) and 1(d) vis-à-vis). Overall, visual record-by-record qualitative inspection of wavelet spectra and MIPs of all the 684 GMs considered in this study suggests that the value of *α* (in degrees) reflects well the actual average-in-time evolutionary trend of the mean frequency content as captured by Morlet wavelet transformation. The following section includes further discussion on the rationale and validity of *a* to serve as an index of the frequency non-stationarity of GMs in view of pertinent statistical analyses gauging its correlation with seismological parameters and established GM properties.
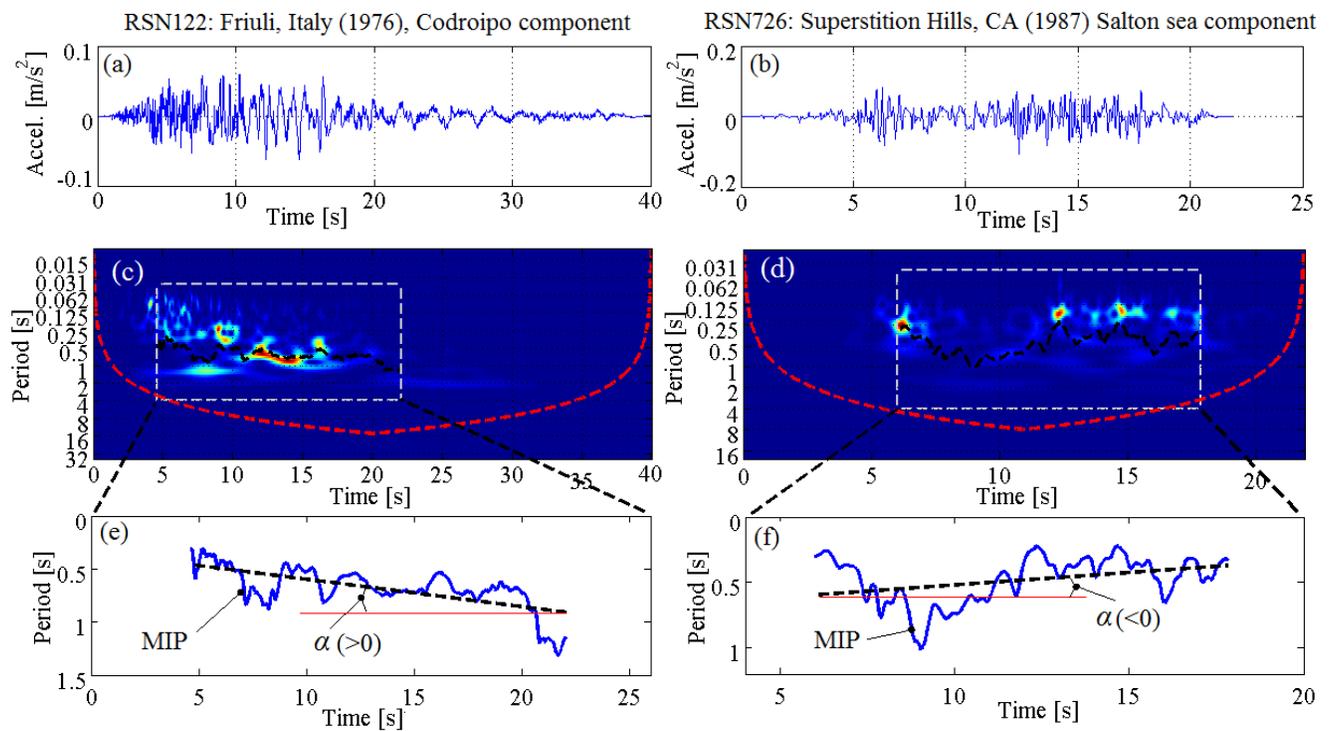
Fig. 1 – (a), (b): Acceleration time-histories, (c), (d): Morlet wavelet contour plots with mean instantaneous period (MIP) superposed, and (e), (f): slope *α* of MIPs, for two different accelerograms.

3. Statistical relationships of *α* with GM properties and seismological parameters

The angle *a* defined in the previous section is computed for 684 GMs from 30 different seismic events with magnitude range 6.5<M<8 and distance to rupture plane range 20km<R_{rup}<120km retrieved from the PEER NGA-West2 Ground Motion Database (http://ngawest2.berkeley.edu/). Unscaled and unfiltered GMs along the "as-recorded" direction were considered and no filter was applied on the faulting type, while "pulse-like" GMs were excluded [5]. Table 1 lists the seismic events, number of GMs considered per event, magnitude, and faulting mechanism. In computing the *a* for each GM, the MIP of the Morlet wavelet spectrogram is computed as detailed in the above section, upon base-line adjustment of the GM signal by acausal high-pass filtering with a fourth-order Butterworth filter with 0.13Hz cut-off frequency (see also [19] and references therein).

Table 1 – Catalogue of GMs used in this study

Event	date	Records number	Magnitude	Mechanism
Northwest California	02/09/1941	2	6.6	strike slip
Borrego Mtn, El Centro Array	04/09/1968	2	6.5	strike slip
San Fernando	02/09/1971	32	6.61	Reverse
Friuli - Italy	05/06/1976	4	6.5	Reverse
Imperial Valley-06	10/15/1979	9	6.53	strike slip
Irpinia - Italy	11/23/1980	8	6.9	Normal
Ierissos - Greece	08/06/1983	1	6.7	strike slip
Taiwan SMART1(25)	09/21/1983	9	6.5	Reverse
Borah Peak ID-01	10/28/1983	8	6.88	Normal
Superstition Hills-02	11/24/1987	4	6.54	strike slip
Spitak Armenia	12/07/1988	1	6.77	Reverse Oblique
Loma Prieta	10/18/1989	62	6.93	Reverse Oblique
Cape Mendocino	04/25/1992	2	7.01	Reverse
Northridge-01	01/17/1994	120	6.69	Reverse
Kobe Japan	01/16/1995	12	6.9	strike slip
Nenana Mountain Alaska	10/23/2002	2	6.7	strike slip
Kern County	07/21/1952	3	7.36	Reverse
Tabas Iran	09/16/1978	2	7.35	Reverse
Trinidad	11/08/1980	3	7.2	strike slip
Taiwan SMART1	11/14/1986	15	7.3	Reverse
Landers	06/28/1992	16	7.28	strike slip
Gulf of Aqaba	11/22/1995	1	7.2	strike slip
Duzce Turkey	11/12/1999	5	7.14	strike slip
Caldiran Turkey	11/24/1976	1	7.21	strike slip
Manjil Iran	06/20/1990	4	7.37	strike slip
Hector Mine	10/16/1999	38	7.13	strike slip
Kocaeli Turkey	08/17/1999	13	7.51	strike slip
Chi-Chi Taiwan	09/20/1999	302	7.51	Reverse Oblique
Sitka Alaska	07/30/1972	1	7.68	strike slip
St Elias Alaska	02/28/1979	2	7.54	Reverse

Next, standard linear regression analysis is undertaken for the GM dataset of Table between *a* and three different GM properties, namely PGA, PGV, and T_m. For each GM, the values of PGA and PGV reported in the PEER NGA-West2 database are used, while the estimate of T_m in Eq. (5) is adopted as this is consistent with the wavelet-family-dependent MIP used in the analysis [16]. Further, a similar regression analysis is also undertaken between *a* and the shear wave velocity V_s30 value (a representative measure of the local soil site conditions) as defined and reported in the same database. The aim of these analyses is to quantify the regression slope coefficients

between a and PGA, PGV, T_m, and $V_{s,30}$ which, upon qualitative interpretation, can serve as evidence that a is a valid index to capture the non-stationary frequency trends of GMs, rather than a signal analysis artefact.

Table 2 – Regression analysis results between α and four different GM properties

Y	R²	Standard error	Coefficient	Coefficient Value	95% confidence interval		p-value	t-value
PGA	0.0060	1.14	C_0 (intercept)	1.12	0.9827	1.264	0.0000	15.66
			C_1 (slope)	1.21	0.045	2.379	0.0417	2.04
PGV	0.076	1.10	C_0 (intercept)	0.78	0.641	0.929	0.0000	10.66
			C_1 (slope)	0.035	0.026	0.044	0.0000	7.53
T_m^{MIP}	0.12	1.076	C_0 (intercept)	0.42	0.23	0.61	0.0000	4.40
			C_1 (slope)	1.13	0.90	1.37	0.0000	9.46
$V_{S,30}$	0.03	1.127	C_0 (intercept)	1.69	1.48	1.91	0.0000	15.80
			C_1 (slope)	-0.001	-0.0016	-0.0006	0.0000	-4.65
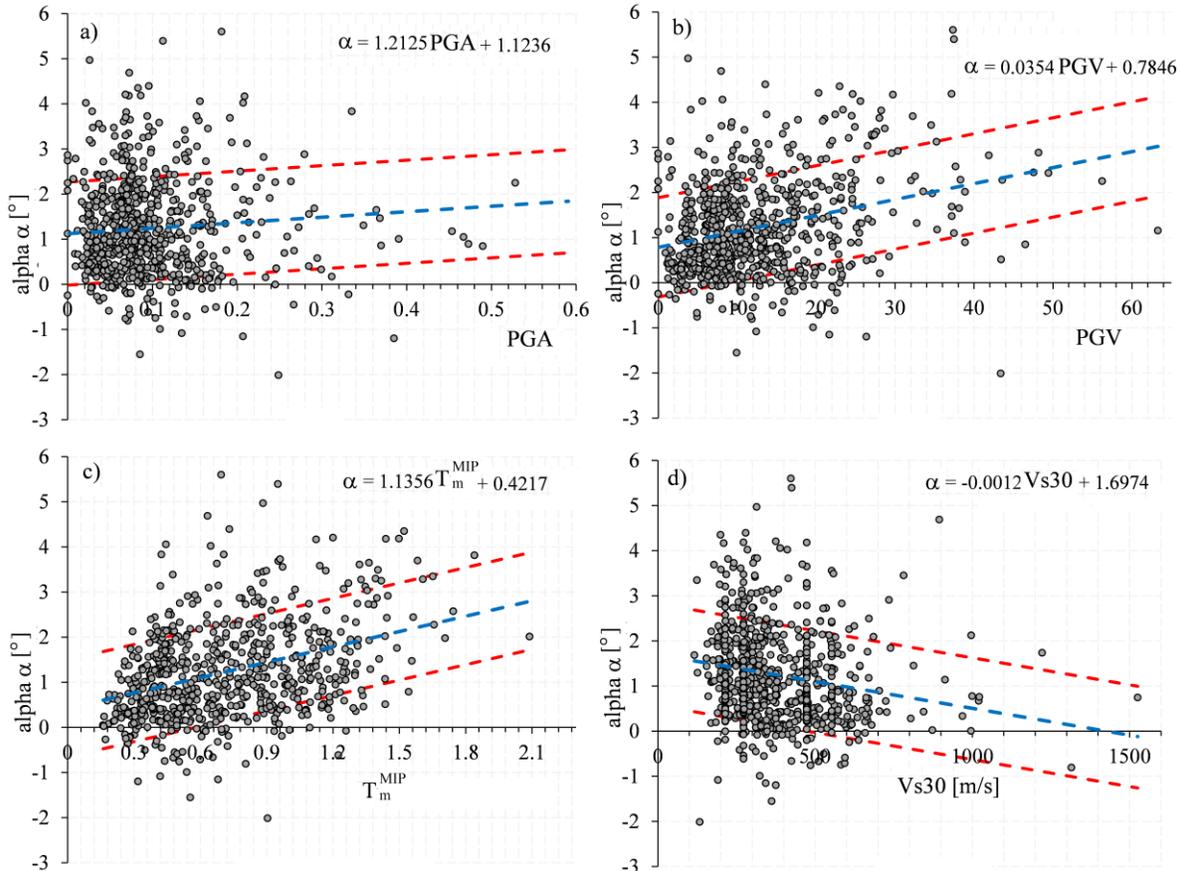
Fig. 2 – Linear regression analyses of α with (a): PGA, (b): PGV, (c): T_m^{MIP}, (d): Vs30.

Figure 2 plots clouds of the 684 data points along with the mean and the mean plus/minus one standard deviation lines obtained from linear regression analysis of the form

$$a = C_0 + C_1 \cdot Y \tag{7}$$

where Y is the GM property under consideration. Table 2 collects the coefficient of determination R^2, the standard errors and the values and confidence intervals of the determined regression coefficients C_o and C_1 as well as their p-values and t-values testing the null hypothesis (i.e., $C_o=0$ or $C_1=0$) for all regression analyses undertaken. It is deduced from Fig.2(a) that, on the average, a is not significantly affected by the PGA as the mean fitted linear function to the (α, PGA) pairs of values is almost flat. Indeed, the fact that the p-value and the t-value of the slope C_1 coefficient of the regression analysis are close to 0.05 and 2.0, respectively (Table 2), confirm that the either way small value of the slope is not statistically significant. However, the average value of a increases appreciably as PGV increases in Fig.2(b) and slope coefficient of the regression analysis is statistically significant. These trends can be intuitively justified by considering that GMs with larger PGV values tend to be richer in low frequencies (this is because the velocity trace of GMs is derived by integration of the acceleration trace which is a low-pass filtering operation [20] suppressing the higher frequency components and accentuating the lower frequencies); *and* by noticing that the lower frequency components in a typical GM usually appear at later times compared to the high frequencies due to the early arrival of the p-waves (see e.g., Fig. 1(c) for a typical example). Therefore, it is natural to expect that the higher the PGV value in a typical GMs is, the more significantly the mean frequency content shifts in time from the higher to the lower frequencies. And if this shift is to be accommodated within roughly the same effective duration (note that in Table 1 most GMs considered are associated with a relatively narrow magnitude range, 7<M<7.5, and hence with roughly the same effective duration as the latter correlates well with the magnitude [9]), then the angle a (i.e., rate of change of the mean frequency content from higher to lower frequencies for $\alpha>0$) attains higher values.

The validity of the above reasoning is further reinforced by examining the average a-T_m^{MIP} trend in Fig.2(c). Specifically, it is seen that α increases with increasing mean frequency (averaged over all times). Hence, it is confirmed that the rate of change in time of the mean frequency content is, on the average, higher for GMs with rich mean low frequency content. The latter observation has also been reported in [6] in which the rate of change of the average frequency content was used as one of the parameters defining a non-stationary GM stochastic model but was extracted from a databank of recorded GMs based on the average zero-crossing rate of GMs: a very different time-domain approach from the wavelet-based one herein adopted. Turning the attention to the a-V_s30 trend in Fig.2(d), it is observed that a decreases as local soil conditions becomes "stiffer" [9], that is, as the value of V_s30 increases. This trend can be readily justified by taking as a fact that a is higher for GMs with richer low frequency content and by considering that soft soils shifts the frequency content of the GMs towards lower frequencies [9]. Alternatively, by reversing the above line of arguments, Fig.2(d) can be used as further evidence that the temporal rate of change of the mean frequency content is higher for GMs that are richer in low frequencies.

Overall, the statistical data furnished in Table 2 and Fig.2 suggest that despite the large scattering/variability of the a with all 4 considered scalars (note the small R^2 values of the regression analyse), it is seen that a is mostly related to the mean frequency content: GMs with lower frequency content tend to have larger a values. To gain an insight on the potential dependency of the angle a on the amplitude of the GM acceleration trace as captured by the PGA, further linear regression analyses are undertaken between a-T_m^{MIP} and a-V_s30 upon dividing the GM dataset of Table 1 into 3 different bins according to their PGA: (i) PGA≤0.05g; relatively low intensity GMs, (ii) 0.05g<PGA<0.15g; medium intensity GMs, and (iii) PGA≥0.15g; relatively high intensity GMs. The mean regression lines between a-T_m^{MIP} and a-$V_{s,30}$ are plotted in Figs. 3(a) and 3(b), respectively, for all 3 bins on top of data points clouds color-mapped according to the considered PGA-based classification. It is observed that as PGA increases, the average value of a becomes more sensitive to the values of both the T_m^{MIP} and V_s30. Indeed, the rate by which the statistical average value of a increases as the low frequency content of GMs becomes richer depends significantly on the PGA: the slope of the red regression line corresponding to high intensity GMs is significantly steeper from the other two regression lines (pertinent t statistic tests [21] were considered verifying that the slope differences in Fig.3 are statistically significant). This result indicates that although PGA is not well-correlated with a directly (at least not as much as the PGV in Figs. 2(a) and 2(b)), it does influence the expected value of a significantly for GMs rich in low frequency content. This trend can be attributed to the fact that typical GMs with relatively high PGA *and* PGV values are characterized by an early in time significant in amplitude high frequency

content, which drives the MIP towards high frequencies (short periods) at the beginning of the GM, *and* by rich low frequency content kicking in later in time, which shifts the MIP towards long periods at a fast rate (i.e., with a large *a* value). The above observation is also verified by visual inspection of the Morlet spectra contour plots and MIPs of GMs with PGA>0.15g. Furthermore, Fig. 2(b) suggests that the expected (average) value of *a* is more sensitive to the soil conditions for high intensity GMs. Indeed, for the PGA≤0.05g bin, the average regression line of *a* with respect to the soil stiffness is flat: the expected value of *a* is not sensitive to soil conditions. However, the average *a* values increase as softer soils *and* higher PGA values are considered. This trend can be readily attributed to the fact that soft soils exhibit stronger non-linear behaviour under intense (high amplitude) seismic shaking compared to stiff soils [9], which reflects on the frequency content of GMs becoming richer in low frequencies.

Lastly, the regression lines in Fig.4 suggest that the magnitude and the distance to rupture R_{rup} (at least within the considered ranges of 6.5<M<8 and 20km< R_{rup} <120km, respectively) do not influence *a* in a direct manner.
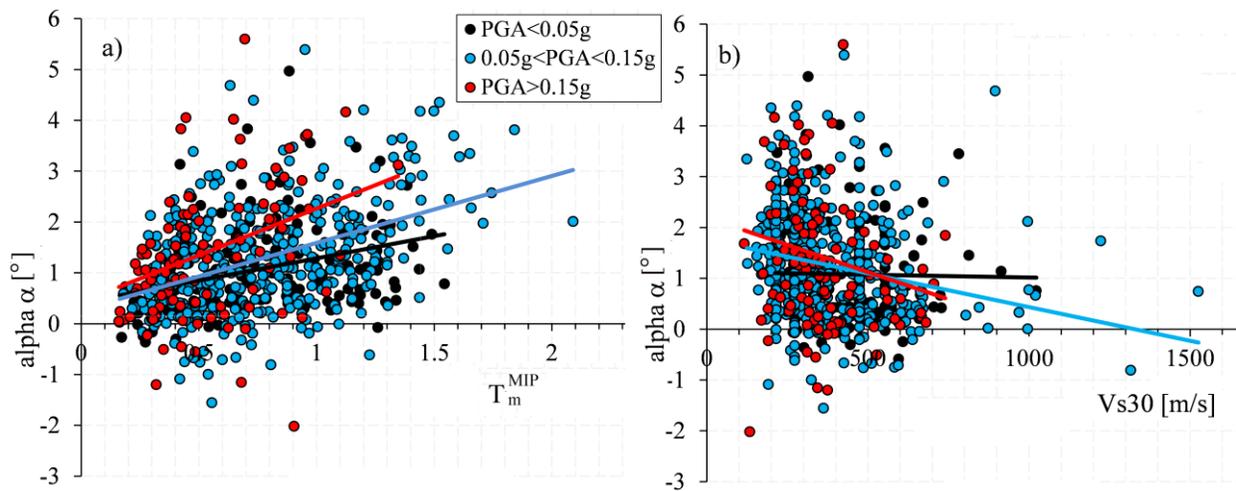
Fig. 3 – Regression analysis of *α* with (a): T_m^{MIP} and (b): Vs30, for the GM dataset of Table 1 classified in 3 different PGA bins.
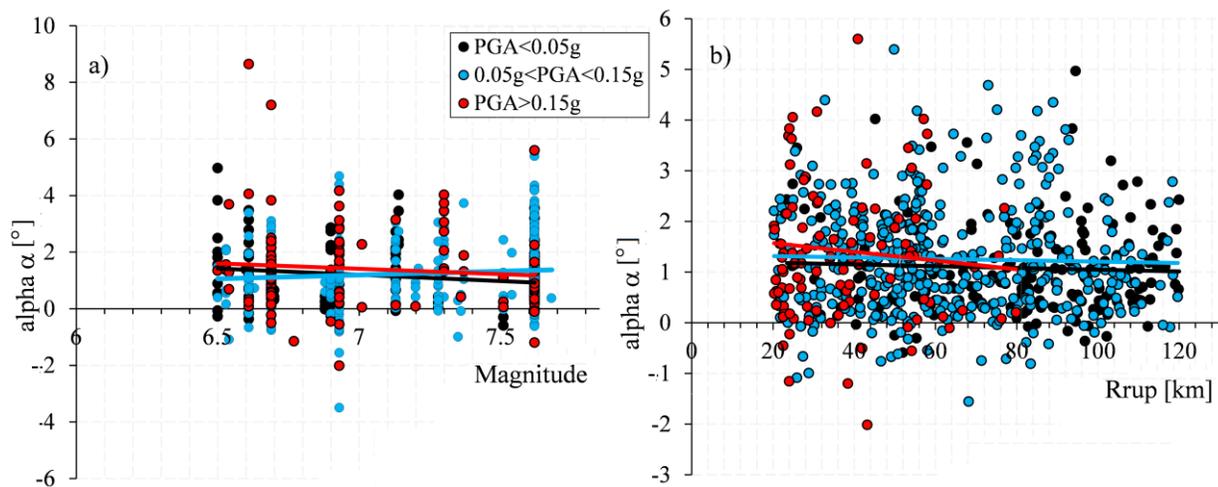
Fig. 4 – Regression analysis of *a* with (a): magnitude M and (b): distance to rupture R_{rup}, for the GM dataset of Table 1 classified in 3 different PGA bins.

4. Sufficiency of the non-structure-specific IMs PGA and PGV with respect to *a*

Having established the validity and usefulness of *a* to represent the non-stationary trend of the mean frequency content and its relationship with key GM properties, a further study is herein undertaken to examine whether and to what extend *a* carries information not included in the PGA and PGV in gauging/predicting the structural damage potential of GMs. This study is motivated by the fact that PGA and PGV are the most commonly adopted non-structure-specific intensity measures (IMs) used to predict certain engineering demand parameters (EDPs) characterizing the seismic response of yielding structures within the PBEE framework [13-15, 22-24]. Although it is well-recognized that PGA is not an efficient IM in predicting EDP values, several studies demonstrated that PGV can be an efficient IM, especially for the case of relatively flexible structures [22-24]. For the purposes of this work, a standard sufficiency IM test is adopted (see e.g. [15, 23]) to assess a potential capability of *a*, beyond that of PGA and PGV, to predict the peak (collapse) inelastic drift $\theta_{\max}$: arguably, the most widely used EDP in earthquake engineering applications (see also [24]). In particular, the considered statistical test involves first undertaking regression analysis between the IM *Y* (e.g., PGA or PGV) and an EDP (e.g., $\theta_{\max}$). Under the common assumption that the EDP-IM relationship follows a power law, the linear regression model is written as

$$\ln(\theta_{\max}) = \ln(C_o^*) + C_1^* \ln(Y),\tag{8}$$

where $\ln(C_o^*)$ and C_1^* are the intercept and the slope regression coefficients of the ln(EDP)-ln(IM) relationship, respectively. Next, a second regression analysis is undertaken between the residuals $\varepsilon|Y$ of the previous regression analysis and *a* as in

$$\varepsilon|Y = C_o' + C_1'a,\tag{9}$$

and the attention is focused on the value of the slope C_1' and its statistical significance as captured by its p-value. This is because the regression in Eq.(9) is a one-predictor (i.e., the *a*) case, therefore a hypothesis test on $C_1' = 0$ is a test on $R^2 = 0$ of the last regression which seeks to determine whether *a* does not correlate with the residuals $\varepsilon|Y$ and, hence, whether *Y* already encompasses the information carried by *a* in predicting the EDP values.
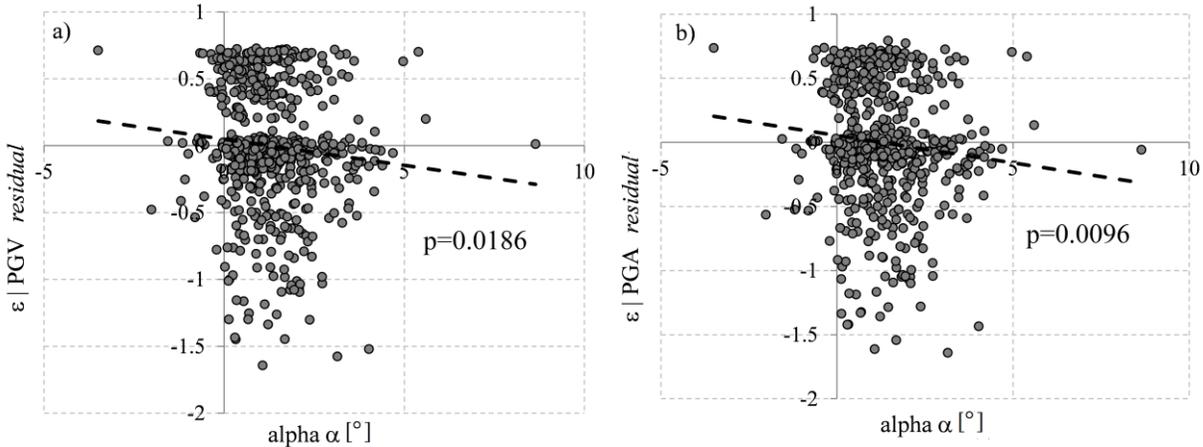
Fig. 5 – (a): Regression analysis of $\varepsilon|PGV$ residuals with *a* (*p-value=0.0186*), and (b): Regression analysis of $\varepsilon|PGA$ residuals with *a* (*p-value=0.0096*).

Following the above methodology, clouds of $\varepsilon|PGV$ and $\varepsilon|PGA$ residuals with *a* are plotted in Fig.5 (count: 684) along with the fitted linear regression lines. The underlying $\theta_{\max}$ (EDP) values used in the regression analysis in Eq. (8) were obtained by performing IDA for the 684 GMs considered in the previous section to a single-degree-of-freedom bilinear hysteretic oscillator with strength and stiffness degradation following the model in [25] as implemented in the OpenSees finite element platform. The bilinear backbone curve of the considered oscillator has been calibrated, using the N2 pushover method, against a regular benchmark 12-storey r/c frame designed according to the current European seismic code (Eurocode 8) for high ductility class [11]. The pre-yield natural frequency of the oscillator is 0.966s and a viscous damping of 5% has been assumed. More details on the adopted

oscillator can be found in [11] and in the references therein. It is seen in the plots of Fig. 5 that there is a non-negligible linear trend (slope C_1') in the fitted regression lines and, more importantly, these trends are statistically significant since the p-values (also reported in Fig. 5) are significantly lower than 0.05 [21]. These results demonstrate that a has a statistically significant effect to the peak inelastic structural response and this effect is more prominent when PGA is used as an IM as opposed to the PGV.

5. Concluding Remarks

A novel wavelet-based scalar quantity termed alpha, a, was proposed to characterize the temporal evolution of the mean frequency content of recorded GMs. It is defined as the average slope (angle) of the time-varying mean instantaneous period (MIP) extracted from the wavelet coefficients of GMs bounded in time within the GM effective duration and bandlimited within the [0.25 25]Hz frequency range. Morlet wavelets were considered in the wavelet transformation of GMs as they yield relatively smooth MIPs in time. Pertinent linear regression analyses involving 684 GMs with no near-fault directivity effects was undertaken to quantify the relationship of a with GM properties PGA, PGV, and mean frequency content T_m, with seismological parameters, M, R_{rup}, and with the shear wave velocity V_s30. No significant correlation was found between a and M, R_{rup}, or PGA. However, it was established that a is well-correlated with the average frequency content of GMs as captured by T_m and by PGV: the lower the average frequency content, the larger a tends to be, that is, the faster the time evolution (transition) of the average frequency content is from higher to lower frequencies. Further, the reported numerical data indicate that the level of the above correlation depends on the intensity of GMs in terms of PGA (conditional on PGA): a is larger for fixed T_m as PGA increases and a increases faster as T_m increases for larger PGA values. Moreover, GMs recorded on softer soils are more likely to have larger a conditional on PGA, a phenomenon that is attributed to the fact that soft soils under strong seismic shaking exhibit strong non-linear behavior that enriches the low frequency content of free field recorded GMs. Lastly, sufficiency statistical tests on a with the residuals of regression analyses between peak inelastic drifts θ_{max} of a bilinear hysteretic SDOF structure estimated through IDA for the previous 684 GMs with PGA and with PGV used as IMs were also conducted. The considered structure includes strength and stiffness degradation effects and is used as proxy of a 12-storey r/c frame. These statistical tests demonstrate that the non-stationary average frequency content of GMs as captured by a influences the peak inelastic structural response at collapse, as captured by θ_{max}. Overall, the herein furnished results establish the validity and usefulness of a in characterizing the evolutionary frequency content of GMs and suggests that a should be considered as a record selection criterion in undertaking IDA using PGA and PGV as IMs. Further research, currently undertaken by the authors, is warranted to confirm the applicability of a for pulse-like GMs and to test whether structure-specific IMs (e.g., $Sa(T_1)$) are sufficient with respect to a especially for pulse-like GMs and for long period structures.

6. Acknowledgements

7. References

[1] Wang J, Fan L, Qian S, Zhou J (2002): Simulations of non-stationary frequency content and its importance to seismic assessment of structures. *Earthquake Engineering and Structural Dynamics*, **31**, 993–1005.

[2] Spanos PD, Giaralis A, Politis NP (2007): Time- frequency representation of earthquake accelerograms and inelastic structural response records using the adaptive chirplet decomposition and empirical mode decomposition. *Soil Dynamics Earthquake Engineering*, **27**, 675- 689.

[3] Spanos PD, Giaralis A, Politis NP, Roessett JM (2007): Numerical treatment of seismic accelerograms and of inelastic seismic structural responses using harmonic wavelets. *Computer-Aided Civil and Infrastructure Engineering*, **22**, 254-264.

[4] Jian F, Yanping Z, Xiaohong L (2014): Establishment of a time-varying modified Kanai-Tajimi non stationary stochastic model of seismic records based on the S-transform. *Journal of Earthquake Engineering*, **18**, 876-890.

[5] Baker JW (2007): Quantitative classification of near-fault ground motions using wavelet analysis. *Bull Seism Soc Am*, **97**,1486–1501.

[6] Rezaeian S, Der Kiureghian A (2010): Stochastic Modeling and Simulation of Ground Motions for Performance-Based Earthquake Engineering. *Technical Report PEER 2010/02*, Pacific Earthquake Engineering Research, Berkeley, USA.

[7] Beck JL, Papadimitriou C (1993): Moving resonance in nonlinear response to fully nonstationary stochastic ground motion. *Probabilistic Engineering Mechanics*, **8**, 157–167.

[8] Vetter C, Taflanidis AA (2014): Comparison of alternative stochastic ground motion models for seismic risk characterization, *Soil Dynamics and Earthquake Engineering*, **58**, 48-65.

[9] Kramer SL (1996): Geotechnical Earthquake Engineering, Prentice-Hall, New Jersey.

[10] Kumar M, Castro JM, Stafford PJ, Elghazouli AY (2011): Influence of the mean period of ground motion on the elastic dynamic response of single and multi-degree of freedom systems. *Earthquake Engineering and Structural Dynamics*, **40**, 237-256.

[11] Katsanos EI, Sextos AG, Elnashai AS (2014): Prediction of inelastic response periods of buildings based on intensity measures and analytical model parameters, *Engineering Structures*, **71**, 161-177.

[12] Rathje EM, Abrahamson NA, Bray JD (1998): Simplified frequency content estimates of earthquake ground motions, *ASCE Journal of Geotechnical and Geoenvironmental Engineering*, **124**, 150-159.

[13] Vamvatsikos D, Cornell CA (2004): Applied incremental dynamic analysis. *Earthquake Spectra*, **20**, 523-553.

[14] Baker JW, Cornell CA (2005): A vector-valued ground motion intensity measure consisting of spectral acceleration and epsilon, *Earthquake Engineering and Structural Dynamics*, **34**, 1193-1217.

[15] Tothong P, Luco N (2007): Probabilistic seismic demand analysis using advanced ground motion intensity measures, *Earthquake Engineering and Structural Dynamics*, **36**, 1837-1860.

[16] Margnelli A, Giaralis A (2015): Wavelet-Based Mean Instantaneous Period (MIP) Of Ground Motions And Time-Varying Period Elongation Of Earthquake Excited Hysteretic Structures. *Earthquake Risk and Engineering towards a Resilient World SECED 2015,* Cambridge, UK.

[17] Torrence C, Compo GP (1997): A Practical Guide to Wavelet Analysis. *Bull Am Met Soc*, **79**, 61-78.

[18] Goupillaud P, Grossman A, Morlet J (1984): Cycle-octave and related transforms in seismic signal analysis. *Geoexploration,* **23**, 85-102.

[19] Giaralis A, Spanos PD (2009): Wavelet-based response spectrum compatible synthesis of accelerograms-Eurocode application (EC8). *Soil Dynamics Earthquake Engineering,* **29**, 219-235.

[20] Worden K (1990): Data processing and experiment design for the restoring force surface method, part I: integration and differentiation of measured time data. *Mechanical Systems and Signal Processing*, **4**, 295-319.

[21] Howell DC (2007): Statistical methods for psychology, 6th Edition, Thompson Wadsworth, Belmont.

[22] Bommer JJ, Alarcon JE (2006): The Prediction and Use Of Peak Ground Velocity. *Journal of Earthquake Engineering*, **10**, 1-1.

[23] Mollaioli F, Lucchini A, Cheng Y, Monti G (2013): Intensity measures for the seismic response prediction of base-isolated buildings. *Bulletin of Earthquake Engineering*, **11**, 1841-1866.

[24] Jalayer F, Beck JL, Zareian F (2012): Analyzing the sufficiency of alternative scalar and vector intensity measures of ground shaking based on information theory. Journal of Engineering Mechanics, ASCE, 138, 307-316.

[25] Ibarra LF, Medina R, Krawinkler H (2005): Hysteretic models that incorporate strength and stiffness degradation, *Earthquake Engineering and Structural Dynamics*, **34**, 1489- 1511.