Selection of Random Vibration Theory Procedures for the NGA-East Project and Ground Motion Modeling


Prepared using sagej.cls
Abstract

Traditional ground-motion models (GMMs) are used to compute pseudo-spectral acceleration (PSA) from future earthquakes and are generally developed by regression of PSA using a physics-based functional form. PSA is a relatively simple metric that correlates well with the response of several engineering systems and is a metric commonly used in engineering evaluations; however, characteristics of the PSA calculation make application of scaling factors dependent on the frequency content of the input motion, complicating the development and adaptability of GMMs. By comparison, Fourier amplitude spectra (FAS) represent ground-motion amplitudes that are completely independent from the amplitudes at other frequencies, making them an attractive alternative for GMM development. Random vibration theory (RVT) predicts the peak response of motion in the time domain based on the FAS and a duration, and thus can be used to relate FAS to PSA. Using RVT to compute the expected peak response in the time domain for given FAS therefore presents a significant advantage that is gaining traction in the GMM modeling field. This paper provides recommended RVT procedures relevant to GMM development, which were developed for the NGA-East project. Additionally, an orientation-independent FAS metric—called the effective amplitude spectrum (EAS)—is developed for use in conjunction with RVT to preserve the mean power of the corresponding two horizontal components considered in traditional PSA-based modeling (i.e. RotD50). The EAS uses a standardized smoothing approach to provide a practical representation of the FAS for ground-motion modeling, while minimizing the impact on the four RVT properties (zero\textsuperscript{th} moment, \(m_0\); bandwidth parameter \(\delta\), frequency of zero crossings, \(f_z\); and frequency of extrema, \(f_e\)). Although the recommendations were originally developed for NGA-East, they and the methodology they are based on can be adapted to become portable to other GMM and engineering problems requiring the computation of PSA from FAS.

Keywords
NGA-East, random vibration theory, RVT, peak factor, effective amplitude spectrum, EAS

Prepared using sage.cls [Version: 2017/01/17 v1.20]
Introduction

Traditional ground-motion models (GMMs) are used to compute the expected pseudo-
spectral acceleration (PSA)—as well as other intensity measures (IMs)—for a given
scenario event—defined by an earthquake magnitude, source-to-site distance measure,
etc.. The response spectrum was selected by the engineering community as the intensity
measure (IM) of choice because it is generally a good predictor of the dynamic response
of a wide range of structures. Thus, PSA is also directly related to seismic design code
spectra. Each frequency of the response spectrum corresponds to the peak absolute
value of the time-domain response of a single-degree-of-freedom (SDOF) system with
specified oscillator natural period and damping. The SDOF response depends on the
frequency content and timing (or phasing) characteristics of the ground motion. The
response is largely controlled by amplitudes in the frequency range near to, and lower
than, the natural oscillation frequency of the SDOF system and the bandwidth affecting
the response is therefore frequency-dependent. As a result, scaling factors (e.g., site
amplification, path attenuation, etc.) for the response spectrum are dependent on the
frequency content of the input motion (i.e., spectral shape), even though the SDOF is a
linear system (e.g., Bora et al. 2016; Stafford et al. 2017).

By comparison, Fourier spectra—defined by an amplitude and phase—provide the
complete characterization of the ground motion in the frequency domain. The influence
of source, path, and site contributions to the Fourier amplitude spectrum (FAS) is
a multiplicative process (as in a linear filter), frequency by frequency, so that the
combined effect at a given frequency does not involve the contributions for other
frequencies. As a result, scaling factors—such as site amplification, path attenuation,
site attenuation, etc.—are applicable to any motion regardless of the spectral shape,
making FAS an ideal tool to integrate seismological effects and scaling factors
needed for ground-motion modeling in a mathematically efficient way. The primary
disadvantage of the FAS is that it is not the primary IM for engineering applications
using current design methods. Random vibration theory (RVT) solves this problem
by providing a method to compute the expected peak response in the time domain
(e.g., PGA, PGV, etc.) for a given FAS. By coupling RVT with the suite of SDOF

1 Pacific Gas & Electric, San Francisco, CA 94105
2 University of California, Berkeley, CA 94720
3 United States Geological Survey, Menlo Park, CA
4 University of California, Los Angeles, CA 90095
5 University of Southern California, Los Angeles, CA, 90089
6 TBD
7 TBD
8 National Observatory of Athens
9 The University of Texas, Austin
10 Pacific Engineering & Analysis
11 United States Geological Survey, Golden, CO
12 Geosyntec Consultants, Inc.

Corresponding author:
Albert R. Kottke, 245 Market St MC: N4C, San Francisco, CA 94105
Email: albert.kottke@gmail.com

Prepared using sagej.cls
transfer functions with varying natural frequencies, it becomes possible to calculate the expected acceleration response spectrum that corresponds to the FAS. The GMM development can therefore be completed in FAS space—where many processes are linear—and then converted to PSA using RVT models. For these reasons, the FAS were relied on by the NGA East project (Goulet et al. 2018; Hollenback et al. submitted, this issue).

The RVT relies on extreme value statistics, which describe the distribution of peak values of the underlying time-varying seismogram. Because of complexities of phase angles inherent to seismograms, there is no exact solution between the FAS and time-domain peak value. The RVT procedure includes numerous steps, and there are a number of possible solutions, each with their own advantages and disadvantages for each step. One of the NGA-East project objectives was to develop at least one set of GMMs using this FAS-RVT-PSA approach pioneered by Campbell (1981) extended by Bora et al. (2014, 2015). NGA-East required that the RVT procedure provide accurate results over the widest frequency range possible and was the motivation for the creation of the RVT Working Group (co-authors of the current paper), tasked to develop appropriate recommendations. The recommendations for some RVT steps are based on the judgement of this working group, while other recommendations are more formally based on the similarity between directly computed PSA and RVT-based PSA values. The full RVT Working Group report is documented in Kottke et al. (2018) and summarized in this paper.

The paper is structured as follows: (1) overview of the RVT approach; (2) selection of peak factors to use for the GMM development application; (3) alternative duration models that are consistent with the RVT process; (4) recommendations; and (5) down-sampled, orientation-independent FAS, referred to as the effective amplitude spectrum (EAS) for use in GMM development. The recommendations provided in this paper were used to develop a suite of GMMs for the NGA-East project (Hollenback et al. 2015, submitted, this issue). In addition, the down-sampled, smoothed EAS were computed and included as data products for the NGA-East (Goulet et al. 2014, submitted, this issue), NGA-West2 (Ancheta et al. 2013), and NGA-Sub (Bozorgnia and Stewart 2020) databases.

**Introduction to random vibration theory for ground motion characterization**

Random-vibration theory (RVT) is a method to statistically represent time series that allows for the calculation of the expected (i.e., mean) time-domain peak value, given an underlying FAS source spectrum and duration, with one calculation; bypassing more computationally intensive time-domain simulations that would otherwise be required. The details of the RVT calculation may be found in Wang and Rathje (2018b); Boore (2003). RVT has been used by seismologists and engineers for decades for a wide range of applications in seismology and engineering (e.g., Hanks 1979; Hanks and McGuire 1981; Boore 1983; Campbell 2003; Rathje and Ozbey 2006; Kottke and Rathje 2013). In practice, RVT can be separated into: (1) a theoretical framework that relates the frequency content and duration of the motion to the distribution of the peak...
time-domain response, and (2) the empirical corrections that are used to improve the accuracy of RVT where assumptions begin to limit the methodology.

In RVT, the ground motion is characterized by the power spectral density (PSD). For a time varying signal \( x(t) \), the PSD can be computed by:

\[
\text{PSD}(f) = \frac{|X(f)|^2}{D} \tag{1}
\]

where \( f \) is frequency and \( X(f) \) is the FAS and \( D \) is the ground-motion duration. Seismologists typically separate the frequency content (i.e., \( X(f) \)) and the duration, instead of combining them into the PSD. If an RVT motion is defined as the average of a set of time series, then the FAS should correspond to the mean power, which is computed by the mean value of the squared Fourier amplitudes (Boore 2003).

RVT defines a peak factor (PF) that relates the root-mean-squared (\( \text{rms} \)) response (\( x_{\text{rms}} \)) to the absolute peak response (\( x_{\text{max}} \)). An assumption that a ground motion is a stationary stochastic process with no change in the probability distribution over the duration interval (i.e., stationarity) is required for the development of the PF formulation. Although earthquake ground motions often violate this assumption, the RVT method has shown to provide good agreement with time series when coupled with corrections for non-stationarity (e.g., Boore and Joyner 1984; Liu and Pezeshk 1999; Boore and Thompson 2012, 2015).

Consider a time-varying signal \( x(t) \) with its associated FAS \( X(f) \): the \( \text{rms} \) value of the signal \( x_{\text{rms}} \) is a measure of the average value over a given time interval, \( D_{\text{rms}} \), and is computed from the integral of the time series over that time interval by:

\[
x_{\text{rms}} = \sqrt{\frac{1}{D_{\text{rms}}} \int_0^{D_{\text{rms}}} |x(t)|^2 dt} \tag{2}
\]

Parseval’s theorem relates the integral of a time series to the integral of its Fourier transform such that Equation 2 can be rewritten in terms of the FAS of the signal:

\[
x_{\text{rms}} = \sqrt{\frac{2}{D_{\text{rms}}} \int_0^{\infty} |X(f)|^2 df} = \sqrt{\frac{m_0}{D_{\text{rms}}}} \tag{3}
\]

where \( m_0 \) is defined as the zero\(^{th} \) moment of the FAS. The \( n^{th} \) moment of the FAS is defined as:

\[
m_n = 2 \int_0^{\infty} (2\pi f)^n |X(f)|^2 df \tag{4}
\]

and is used as a measure of the frequency content of the ground motion. Using these spectral moments, characteristics such as frequency of zero crossings (\( f_z \)) and extrema (\( f_e \)) (i.e., frequency of peaks in the signal) can be calculated. By combining \( f_z \) and \( f_e \) with the duration of ground-motion duration (\( D_{\text{gm}} \)), the number of zero crossings (\( N_z \)) and extrema (\( N_e \)) can be calculated. For further details see Wang and Rathje (2018a).

**Peak factor formulations**

There are different models for the peak factors, called peak factors (\( \text{PF}= \frac{x_{\text{max}}}{x_{\text{rms}}} \)) that relate the maximum statistics parameters mentioned above (\( N_z, N_e, f_z, f_e, \) etc.)
to the expected maximum value. In this section, we summarize PF models commonly used in ground motion studies. These models were explored because of the formulation assumptions, the validation efforts, and applications of the method. Additional models are presented in Kottke et al. (2018). Other PF formulations exist that were not considered as part of this effort and that may benefit from a similar evaluation (e.g., Winterstein and Cornell 1985).

**Cartwright and Longuet-Higgins**

Cartwright and Longuet-Higgins (1956), abbreviated as CLH56, studied ocean waves and developed functions to predict peak wave height based on characteristics of the wave train. CLH56 considers a stationary process over a time interval and computes the expected peak value. Boore (1983, 2003) modified the PF equation by changing the variables to remove an integrable singularity. The resulting PF equation is:

\[
\text{PF} = \frac{x_{\text{max}}}{x_{\text{rms}}} = \sqrt{2} \int_0^\infty \left\{ 1 - \left[ 1 - \varepsilon \exp(-z^2) \right]^{N_e} \right\} dz
\]  

where \(z\) is the variable of integration and \(\varepsilon\) is defined as ratio of number of zero crossings \((N_z)\) to the number of extrema \((N_e)\):

\[
\varepsilon = \frac{N_z}{N_e} = \sqrt{\frac{(m_2)^2}{m_0 \cdot m_4}}
\]

where \(N_z\) and \(N_e\) are defined in Equations (6) and (8).

Equation 5 assumes each peak is statistically independent (i.e., a Poisson process) and the distribution of peaks is one-sided. For narrow-band motions \((\varepsilon = 1)\), the peaks follow a Rayleigh distribution. For broadband motions \((\varepsilon = 0)\), the peaks follow a Gaussian distribution. The one-side distribution is inconsistent with the response spectrum, which considers the two-sided maximum.

The CLH56 PF has been used extensively by the seismological community (e.g., Boore 2003; Campbell 2003). Since it’s inception, there have been several modifications of the CLH56 PF. Davenport (1964), abbreviated as D64, simplified Equation 5 to an asymptotic form, which results in faster calculations at the cost of decreased accuracy. Given current computing power, this simplification is not recommended, but is included in comparisons for completeness.

**Vanmarcke (1976)**

Vanmarcke (1976), abbreviated herein as V76, extended the PF formulation to include the potential for clumping in time of zero crossings. This formulation no longer assumes statistically independent peaks over time (i.e., Poisson process). The cumulative distribution function (CDF) of the peak values from Equation (29) of V76 is as follows:

\[
F_x(x) = \left[ 1 - \exp \left( -\frac{x^2}{2} \right) \right] \cdot \exp \left\{ -\frac{N_z \left[ 1 - \exp \left( -\sqrt{\frac{\pi}{x} \cdot \delta_x \cdot x} \right) \right]}{1 - \exp \left( -x^2/2 \right)} \right\}
\]  

Prepared using sagej.cls
where $\delta_e$ is an empirical factor defined as:

$$\delta_e = \delta^{1+b} \tag{8}$$

$\delta$ is an alternative measure of bandwidth defined by:

$$\delta = \sqrt{1 - \frac{(m_1)^2}{m_0 \cdot m_2}} \tag{9}$$

and the value of $b$ was estimated to be 0.20 based on fitting to numerical simulations of a simple oscillator to a Gaussian white-noise input. Equation 7 provides the CDF of the PF and does not provide for direct calculation of the expected PF. Instead, a direct calculation can be obtained from the equation:

$$E[x] = \int_{0}^{\infty} x \cdot f_x(x) \, dx \tag{10}$$

where

$$f_x(x) = \frac{\partial F_x(x)}{\partial x} \tag{11}$$

The derivative can be calculated analytically. However, because $F_x$ is continuous and only non-negative, the expected value may be computed directly from $F_x$ by:

$$E[x] = \int_{0}^{\infty} [1 - F_x(x)] \, dx \tag{12}$$

The PF is then computed by substituting Equation 7 into Equation 12.

Der Kiureghian (1980), abbreviated as DK80, assessed the Davenport (1964) PF formulation and identified that the Poisson assumption of crossings tends to overestimate the mean and underestimate the variance of the peak, and developed a modified $N_z$ based on empirical observation and use of the Vanmarcke (1976) bandwidth measure $\delta$.

The community has been slowly moving from CLH56 to the V76 PF. Since 2015, SMSIM uses the V76 peak factor formulation (Boore and Thompson 2015). Kottke and Rathje (2013) compared time series and RVT seismic site response analyses, and showed significant differences between the results, which in part was due to the assumptions in the CLH56 peak factor. Work by Wang and Rathje (2016) has shown that using the V76 PF formulation reduces these differences.

Peak factor comparisons

The PF formulations presented in the previous section can be separated into two categories: (1) those based on CLH56; and (2) those based on V76. The CLH56 formulation factor assumes statistical independence between the local peaks of a random process, which is a significant approximation for a narrow-band process (Der Kiureghian 1980). Recall that the D64 and DK80 factors are approximations of the CLH56 and V76 factors, respectively.

To illustrate the differences between the PF formulations, the PFs were computed for two earthquake scenarios. The scenarios are defined with FAS computed using a
point-source seismological model using parameters for Eastern North America from Atkinson and Boore (1998) and Boore and Joyner (1997), and documented in Campbell (2003), for a magnitude $M_4$ at 5 km and a $M_7$ at 100 km. The FAS of the ground motion and three 5%-damped SDOF oscillators is shown in Figure 1 (a) and (b). The response of the SDOF oscillator distorts the shape of FAS depending on the frequency of the oscillator and the characteristics of the ground motion.

For both motions, the PF is computed using the CLH56, D64, V76, and DK80 formulations. For these calculations, $N_z$ was limited to be greater than 1.33, which is a threshold recommended by Veneziano and Van Dyck (1985, section 3, vol. 3, eq. B-35, p. B-80). The computed PFs are shown in Figure 1 (c) and (d). The differences between each of the formulations vary with respect to input motion and oscillator frequency. For the $M_7$ event, the considered PFs are clustered into two groups: CLH56 and D64, demonstrate relatively consistent behavior, as do V76 and DK80. These results are not surprising, as the two groups of PFs are based on differing assumptions. Additionally, the differences between the approximate solutions (D64 and DK80) and the full solutions (CLH56 and V76) are shown in this figure. For the $M_4$ event, the significant differences between CLH56 and D64 at frequencies less than $\sim 6$ Hz are due to the limited range over which the asymptotic approximation is accurate. Note that the limitation of $N_z \geq 1.33$ also has the effect of causing the DK80 PF to be a constant ($\sim 1.51$) for frequencies less than $\sim 10$ Hz because the approximation is a function of $N_z$.

**Ground-motion duration for RVT calculations**

In the context of ground motions, “duration” is the term used to refer to different definitions of time-length intervals. When discussing ground-motion duration, it is important to understand the specific definition of the metric. In some cases, duration can refer to the actual physical duration of ground shaking at a site based on threshold or rules; in other cases, as is the case for some RVT applications, it can be a fitted parameter with units of time, but it does not relate to the traditional concept of shaking duration.

**Significant durations**

For many ground-motion applications, the significant duration, bracketed duration, and uniform duration definitions are often used to characterize the duration of ground-motion shaking. Herein, only the significant duration, as introduced by Trifunac and Brady (1975), will be discussed. It is based on the normalized integral of the squared acceleration (i.e., the Husid (1969) curve) and is closely associated with the $rms$ acceleration.

The Arias intensity ($I_a(t)$) is the cumulative squared acceleration and at time $t$ is computed by:

$$I_a(t) = \frac{\pi}{2g} \int_0^t (\alpha(t))^2 \, dt$$  \hfill (13)
where $x(t)$ is the acceleration at time $t$ in units of gravity (g) (Arias 1970). The Husid curve at time $t$ ($h(t)$) is defined as:

$$h(t) = \frac{I_a(t)}{I_a(T)}$$

(14)

where $T$ is the total duration of the record. The Husid curve ranges from 0.0 at the start of a record to 1.0 at the end of a record. The significant duration $D$ (i.e., Arias or Husid duration) is defined as the difference in the times at which $h(t)$ equals two specific percentages of its final value. For example, $D_{5-95}$ refers to the time difference of the 5%-95% occurrence of the squared acceleration.

A fundamental assumption underlying RVT is that the signal is stationary, which means no change in the probability distribution over the duration internal. Based on that, Vanmarcke and Lai (1980) and Ou and Herrmann (1990) argued that the duration window should be chosen such that a stationary process can approximate the signal. Their criterion for determining that a time window is near-stationary is that its Husid plot increase approximately linearly throughout that time window. They analyzed records from the Eastern Canada Telemetered Network with epicentral distances ranging from 135–994 km and concluded that the $D_{5-75}$ window is approximately stationary. They validated the use of this window with RVT by comparing the measured peaks to the peaks predicted by RVT, noting that the spectrum should be estimated only from the signal within the $D_{5-75}$ time window when durations are used in the RVT calculations.

The stochastic method of ground-motion simulation, as developed by Boore (1983), is a frequency-based technique that uses $D_{5-95}$ as the total duration of ground shaking and the calculation of $x_{rms}$, and uses a duration metric adjusted as a function of the oscillator period to compute the PF. Thus, $D_{5-95}$ may still be preferred as the metric for ground-motion duration for use within the Boore (1983) framework, for a number of reasons: (1) the duration of ground-motion excitation is conceptually distinct from the duration of a responding oscillator (as in Boore and Thompson 2012, 2015); (2) it provides consistency and continuity with the Boore formulation (Boore 1996, 2003); and (3) it removes the ambiguity with IM-compatible durations (defined below) because the duration is quantified from the time series. Note that Boore and Thompson (2014) proposed using an effective $D_{5-95}$ duration that is computed as $2 \cdot D_{20-80}$ to avoid variability in the estimated duration arising from early arriving P waves or long surface wave durations.

Since publication of our recommendations for the NGA-East project, Kolli and Bora (2021) has suggested that significant duration metric of $D_{5-75}$ provides least variation in residual spread when compared with other duration metrics—including RVT optimized durations. Future studies should consider the arguments presented in Kolli and Bora (2021) when selecting a duration metric.

**IM-compatible ground motions**

Bora et al. (2014) proposed a duration that is rooted in the RVT framework. It is different from other duration measures in that: (1) it is not based on an acceleration time series; and (2) it is not a “duration” in a physical sense, but rather an empirically
estimated coefficient suitable for calculating response spectra using RVT. The duration proposed in Bora et al. (2014) is defined as the duration that minimizes the misfit between the response spectrum computed in time domain and the response spectrum calculated from the observed FAS with RVT. Thus, in an application that aims to predict response spectra, this duration measure is ideally suited for the RVT framework. Essentially, the duration is treated as a parameter that is obtained by fitting the RVT response spectrum to the analytical response spectrum. Therefore, one should refrain from associating any physical meaning with this duration: it is to be used as a parameter within RVT to predict response spectra together with the FAS GMM. The advantage of treating the duration as a tuning parameter is that it absorbs any misspecification of the model. Forward calculations need to be made with the same assumptions/formulations as the ones used in the inversion of the duration. This technique was also used by Hollenback et al. (2015, submitted, this issue) in which IM-compatible durations were regionalized based on the localization of the ground motion records.

**Correction for non-stationarity**

Non-stationarity (i.e., time-varying statistical properties) exists in both \( x_{\text{rms}} \) and the PF. Ou and Herrmann (1990) demonstrated that for peak values such as PGA and PGV, the careful selection of a time window can overcome much of the non-stationarity problem. But for resonant systems, such as SDOF oscillators or layered soils, the problem of non-stationarity is an issue. As described in Van Houtte et al. (2018), while there are theoretical corrections for non-stationarity in both \( x_{\text{rms}} \) and the PF, empirical corrections have focused on the non-stationarity of PF. There are two methods for adjusting the PFs to correct for non-stationarity: (1) modifying the duration of the motion, or (2) directly scaling the PF.

Various methods exist for correcting differences between RVT-predicted spectral accelerations and time-domain simulations. Vanmarcke (1976) considered the time-dependent response of a SDOF oscillator and developed a non-stationary factor by integrating the transient squared-amplification function over all frequencies. This non-stationarity factor corrects for the fact that an oscillator may not reach steady-state conditions over the duration of the ground motion (Toro and McGuire 1987). Other researchers (e.g., Boore and Joyner 1984; Liu and Pezeshk 1999; Boore and Thompson 2012, 2015) have developed correction factors based on comparison with time-domain simulations and adjustment of the \( x_{\text{rms}} \) duration. Non-stationarity may be an issue for downstream applications. For examples, a FAS GMM could be modified by site-specific site transfer functions to capture the site effects. Doing this may introduce non-stationarity issues. While not considered in this study, Wang and Rathje (2018b,a) developed correction factors for non-stationarity for application of RVT to site response and calculation of surface response spectra. These factors were developed through adjustment of the duration based on characteristics of the site transfer function.

The use of empirical corrections provides a mechanism to improve the RVT prediction of a response spectrum from a source FAS and duration, and thus are recommended. However, these corrections do not address the underlying non-stationarity that is not fully captured by the PF, which warrants future work.
**Recommendations for PSA computations**

**Peak factor formulations**

Our recommendations for the PF are based on the fundamental assumptions of the different PFs, and are not based on quantitative comparisons of RVT-based response spectra to directly computed response spectra. Instead, we first recommend that the asymptotic approximation (DK80 and D64) not be used since the computational efficiency gained by these approximations is no longer warranted and the approximations can introduce bias. Thus, the Cartwright and Longuet-Higgins (1956), and Vanmarcke (1976) factors are candidate models. Further, we recommend the use of the V76 PF over the CLH56 because V76 does not assume independence of the peaks, and is thus a more robust solution. Furthermore, it has been shown to provide better estimates of site response by Wang and Rathje (2016), which might be part of ground motion development. These PFs should be used in conjunction with the corrections for non-stationarity proposed by Boore and Thompson (2012, 2015). The V76 PF should not include the V76 correction for non-stationarity as this was not included in Boore and Thompson (2015).

**Duration**

For purposes of calculating oscillator response, the corrections for the oscillator rms durations developed by Boore and Thompson (2012, 2015) for response spectra calculations are recommended. To be consistent with these corrections, the effective $D_{5-95}$ computed by $2 \cdot D_{20-80}$ is the preferred duration metric, as it was used in the development of the corrections. Rather than using a physical duration (e.g., $D_{5-95}$), another reasonable option is to compute new rms durations optimized to minimize PSA bias could be computed from a region-specific database; that approach was applied by Hollenback et al. (2015, submitted, this issue) in the development of their GMMs. $D_{5-75}$ is a good alternative (Vanmarcke and Lai 1980; Ou and Herrmann 1990; Kolli and Bora 2021), however the corrections for non-stationarity may need to be adjusted if it is selected for use.

**Effective amplitude spectrum**

Prior to developing ground-motion models based on the FAS, a standard processing framework must be defined. For NGA-East, the ground-motion metric selected for ground-motion prediction is RotD50 (Boore 2010) for PGA, PGV and 5%-damped PSA. RotD50 consists in an “average” representation of the ground motion that is independent of the orientation of the sensors. To ensure compatibility with this metric, we developed the effective amplitude spectrum (EAS), a single representation of the horizontal acceleration FAS that is not dependent on the orientation of the two orthogonal components of recorded ground motion and is smoothed by a standardized process that maintains RVT statistical properties. The unsmoothed EAS is given by:

$$EAS(f) = \sqrt{\frac{1}{2} [FAS_{H1}(f)^2 + FAS_{H2}(f)^2]}$$  \hspace{1cm} (15)
where FAS\textsubscript{H1} and FAS\textsubscript{H2} are the FAS of the first and second as-recorded horizontal components of a three-component acceleration time series (Figure 2a). As a matter of convention, the EAS are provide for acceleration. The calculation of the EAS is based on the average power of the two time series, which is computed by the mean value of the squared Fourier amplitudes (Boore 2003), as mentioned above.

Unlike PSA, which can be computed at any specific oscillator frequency, the corresponding frequencies of Fourier amplitude spectrum are a function of both the time step (sampling frequency) and length of the time series. As part of the NGA-East uniform record processing routine, time series were all zero-padded out to longer durations so as to make the frequency step, \( \Delta f \), at which the FAS was computed as uniform as possible. The choice of \( \Delta f \) was a compromise between record length and reasonably similar \( \Delta f \), and it resulted in the two \( \Delta f \) defined in Table 1 based on the samples per second in the original time series (see also Goulet et al. 2014; Kishida et al. 2016; Goulet et al. submitted, this issue).

**Table 1. Frequency sampling of the NGA-East dataset.**

<table>
<thead>
<tr>
<th>Data</th>
<th>Type 1</th>
<th>Type 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samples per second</td>
<td>10, 20, or 40</td>
<td>50, 100, or 200</td>
</tr>
<tr>
<td>Time step, ( \Delta f ) (sec)</td>
<td>0.10, 0.05, or 0.025</td>
<td>0.02, 0.01, 0.005</td>
</tr>
<tr>
<td>Duration (sec)</td>
<td>3276.8</td>
<td>2621.44</td>
</tr>
<tr>
<td>Length</td>
<td>( 2^{15}, 2^{16}, \text{or } 2^{17} )</td>
<td>( 2^{17}, 2^{18}, \text{or } 2^{19} )</td>
</tr>
<tr>
<td>Frequency step (Hz)</td>
<td>0.00030518</td>
<td>0.0003815</td>
</tr>
</tbody>
</table>

The resulting FAS for these processed records contain very large numbers of frequency points. The resulting FAS were not practical for the FAS-based GMM development, for which the regression is performed frequency-by-frequency. With so many points varying within even fairly narrow frequency ranges, it was not possible to extract meaningful amplitudes without considering some kind of smoothing or averaging. This led to the development of a smoothing technique that would make the FAS values represent a certain frequency band, retain the RVT-relevant properties of the FAS but that would reduce the number of points considered in the computations. Hence the smoothing technique and level were carefully selected so as not to introduce bias relative to the original dataset.

The selection of the smoothing method and parameters is described below for the Vanmarcke PFs that were selected by Hollenback et al. (2015, submitted, this issue). The methodology and the principles behind it are portable to other PF formulations, and could be adapted to other applications. In the context of RVT performed using the Vanmarcke PFs (Equation 7), there are four FAS quantities controlling the process in addition to duration: the zero\textsuperscript{th} spectral moment \( (m_0) \), a measure of the ground motion bandwidth \( (\delta) \), the frequency of zero crossings \( (f_z) \), and the frequency of extrema \( (f_e) \). The smoothing criterion for the current application was as follows: that the smoothed, down-sampled EAS lead to similar quantities as the complete EAS for the four properties relevant to RVT. Down sampling is the process of interpolating the FAS at specific frequencies using an average computed by the weights of the smoothing.
operator. The purpose of down sampling is to characterize the EAS with fewer points, while maintaining the relevant information.

Several smoothing windows of various widths were considered and tested, including Hamming, trapezoidal and triangular windows. Such windows applied to linear sampling of frequencies led to bias in spectral shape, especially at low frequencies, and did not allow a close fit to the four quantities listed above. To prevent these issues, the RVT working group selected the Konno and Ohmachi (1998), abbreviated KO98, smoothing window, which is defined based on the log sampling of frequencies. The KO98 window weights are defined by:

$$W(f) = \left\{ \frac{\sin [b \log_{10}(f/f_c)]}{b \log_{10}(f/f_c)} \right\}^4$$ (16)

where $W$ is the weight defined at frequency $f$ for a window centered at frequency $f_c$ and defined by window parameter $b$. Window parameter $b$ can be defined in terms of the bandwidth in log units of the smoothing window as:

$$b = \frac{2\pi}{b_w}$$ (17)

where $b_w$ is the bandwidth of the smoothing window in log units. The KO98 smoothing window was selected because it resulted in little-to-no bias on the amplitudes of the smoothed EAS when compared to the unsmoothed EAS. Other smoothing windows considered (the Hamming, trapezoidal, and triangular windows) did not satisfy that criterion. Figure 3 shows the KO smoothing window for $f_c = 5 \text{ Hz}$ and four different values of $b_w$.

The four RVT properties selected for calibration are all functions of several different $n^{th}$ order moments of the oscillator frequency response function, which depends on the square of the FAS of the ground motion. Thus, the smoothing and down-sampling is performed on the square of the EAS. All EAS in the NGA-East database were smoothed and down-sampled for six different combinations of smoothing window bandwidth, $b_w$, and number of frequency points for down-sampling (Table 2). To determine the combination that would have the least impact on subsequent GMM development, the four properties selected for RVT calibration (namely, $m_0$, $\delta$, $f_z$, and, $f_e$) were calculated for the original EAS and the smoothed and down-sampled EAS.

The comparison was done at three different oscillator periods covering the full range of periods to be defined by the GMMs: 0.01 s, 0.2 s, and 10 s. The value of each property from the smoothed and down-sampled EAS was compared to that of the original EAS, for all the records and for each oscillator period. We quantified the number of records in the database whose smoothed and down-sampled RVT properties fell within a given percentage of the original RVT properties. The percentage of records which had smoothed and down-sampled RVT properties within ±1% range of original RVT properties are listed in Table 2. Based on this criterion, the combination of $b_w = 1/30$ ($b = 188.5$) and 100 frequency points per decade—a decade is $1 \log_{10}$ cycle, for example 1 to 10 Hz—was identified as having the least impact on the four RVT calibration properties. Hence, this combination was selected for smoothing and down-sampling the NGA-East database (Goulet et al. 2014, submitted, this issue) and for the GMM development by (Hollenback et al. 2015, submitted, this issue). Close inspection
of the records falling outside of the 1% range led to the elimination of those records on the basis of peculiar spectral shape around a limited frequency band.

**Table 2.** Percentage of records in NGA-East database that have RVT properties of the smoothed and down-sampled EAS within \( \pm 1\% \) of the RVT properties of the original EAS. Values above 90% are in bold.

<table>
<thead>
<tr>
<th>Oscillator period 0.01 sec (100 Hz)</th>
<th>( PPD )</th>
<th>( b_w )(^\dagger)</th>
<th>( m_0 )</th>
<th>( \delta )</th>
<th>( f_z )</th>
<th>( f_e )</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 1/15</td>
<td>46</td>
<td>49</td>
<td>86</td>
<td>88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 1/30</td>
<td>18</td>
<td>20</td>
<td>47</td>
<td>57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 1/30</td>
<td>35</td>
<td>39</td>
<td>80</td>
<td>85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 1/50</td>
<td>19</td>
<td>21</td>
<td>54</td>
<td>63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100 1/30</td>
<td>99</td>
<td>98</td>
<td>100</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100 1/100</td>
<td>24</td>
<td>26</td>
<td>64</td>
<td>70</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Oscillator period 0.2 sec (5 Hz) | 30 1/15 | 21    | 15    | 87    | 92    |        |        |
|---------------------------------|--------|--------|--------|--------|--------|        |        |
| 30 1/30                         | 12    | 10    | 66    | 78    |        |        |
| 50 1/30                         | 25    | 24    | 89    | 94    |        |        |
| 50 1/50                         | 14    | 13    | 73    | 83    |        |        |
| 100 1/30                        | 99    | 98    | 100   | 100   |        |        |
| 100 1/100                       | 17    | 15    | 79    | 88    |        |        |

| Oscillator period 10 sec (0.1 Hz) | 30 1/15 | 22    | 37    | 60    | 76    |        |        |
|---------------------------------|--------|--------|--------|--------|--------|        |        |
| 30 1/30                         | 11    | 14    | 32    | 37    |        |        |
| 50 1/30                         | 30    | 34    | 63    | 70    |        |        |
| 50 1/50                         | 19    | 19    | 40    | 43    |        |        |
| 100 1/30                        | 94    | 99    | 99    | 100   |        |        |
| 100 1/100                       | 37    | 28    | 58    | 56    |        |        |

\(^\dagger\) Number of frequency points per decade

\(^\dagger\) Width of smoothing window \( b_w \) (fraction of decade)

The recommended KO98 smoothing window (\( b = 188.5 \)) is compared with the normalized 5%-damped SDOF transfer function and a KO98 smoothing window with \( b \) of 20 in Figure 4. In analysis of horizontal-to-vertical spectral ratio processing, KO98 with a \( b \) of 20 is commonly used. The recommended value for calculation of the EAS is much narrower. This is because the goal of this smoothing window was to preserve the RVT characteristics of the motion, rather than match the SDOF transfer function or provide a smooth, easily interpreted signal (Figure 2b). By smoothing over a very narrow band of frequencies, the operator uses highly correlated Fourier amplitudes, which may explain the better reproduction of the RVT parameters. During tuning of the smoothing parameters, it was observed that the overlap is required for the downsampled FAS to maintain the full FAS properties across all metrics.
A possible future improvement to the definition of EAS would be to make use of the multitaper spectra (Thomson 1982). The multitaper spectra tends to be smoother than the Fourier spectra, and also provides estimates of the uncertainty of the spectrum.

Demonstration of proposed process
As a final demonstration, a set of records from the NGA-East (Goulet et al. 2014, submitted, this issue) ground motion database are selected for comparison with the procedure. Records are selected with the following characteristics:

- Closest distance between 0 and 300 km.
- Magnitude greater than 3.5.
- Quality Flag equal to zero; this flag limits residuals to be within $\pm 4 \cdot \sigma_{1n}$ for PGA, PGV, and PSA(T=0.05 s) relative to Atkinson and Boore (2011) as defined by Goulet et al. (2014, submitted, this issue)

These criteria were chosen to be consistent with the data used to develop (Goulet et al. 2018; Hollenback et al. submitted, this issue) GMM. These criteria reduce the number of records from 9,382 to 1,025. For each of the motions, the V76 PF was used to compute the spectral acceleration from the EAS spectrum of the motion. The duration used in the RVT analysis was adjusted to minimize the misfit by minimizing the sum-of-squared error above 1 Hz and limited to the highest usable frequency. The misfit (difference in natural log-space between observed and that predicted by RVT) is shown in Figure 5. An individual time series (TS) provides a realization of the maximum response, while RVT provides the expected peak time-domain value for a moment with those characteristics. Thus, it is not expected the TS and RVT will agree perfectly. Never-the-less, there is good agreement over the optimized range (above 1 Hz) with a mean bias of about zero (0.001) and mean standard deviation of 0.16 in natural-log units. The standard deviations at frequencies less than 0.5 Hz are due to non-stationarity of the time series at these low frequencies and other assumptions of RVT not holding (i.e. Gaussian distribution), as well as the duration not being optimized for these low frequencies.

Summary
Motivated by the need for NGA-East to develop FAS-based GMMs, the NGA-East RVT Working Group was created and tasked to provide recommendations on the most appropriate RVT-related procedures for that purpose. This was a fairly specific task. However, RVT is used in several applications such as site response and structural analysis, which were not considered by this effort.

For the purpose of median GMM development for PSA and based on their relevance to the engineering and seismological communities, two PF formulations are candidates: (1) Cartwright and Longuet-Higgins (1956), and (2) Vanmarcke (1976). The complete form of the V76 formulation is recommended because asymptotic formulations, such as DK80, are not accurate over the range of bandwidths considered by NGA-East (0.01-10 sec.). The V76 PF formulation is preferred over CLH56 because of the underlying assumptions of statistically independent peaks in CLH56.
For the duration, $D_{5-95}$, computed by $2D_{20-80}$, has been successfully used and validated by Boore (2003); Boore and Thompson (2014) and should be used with calibrations from Boore and Thompson (2012, 2015). Alternatively, we recommend that an IM-compatible ground-motion duration model be developed from the database. This duration is not a physical characteristic of the time series, instead it is a model parameter that minimized the misfit between PSA computed from the time series and PSA computed from FAS with RVT.

We also developed an orientation-independent average horizontal acceleration FAS metric called the effective amplitude spectrum (EAS) to use in conjunction with RVT to represent the mean power of the corresponding two horizontal components considered in traditional PSA-based modeling (i.e. RotD50). The benefit of this EAS definition is that it provides a standard method for computing the FAS that allows different data sets and models to be compared. In defining the EAS, we combined the concept of uniform frequency-domain sampling with tuned parameters for the Konno and Ohmachi (1998) smoothing technique to make the EAS a more practical representation of the FAS for ground-motion modeling. Provided that the frequency step is small enough, we recommend using a Konno and Ohmachi (1998) smoothing window with a width ($b_w$) of 1/30 and 100 frequency points per decade. This smoothing window was identified as having the least impact on the four RVT calibration properties; therefore it was selected for smoothing and down-sampling of the NGA-East database (Goulet et al. 2014, submitted, this issue). Although the details and final recommendations were developed for the NGA-East database and modeling, most of the considerations and analysis techniques are portable to other applications.

Data and Resources

The RVT PF formulations discussion in this article are available in the Python library pysra (Kottke 2020b). Implementation of the effective amplitude spectrum including a fast calculation of the Konno and Ohmachi (1998) smoothing window using Numba is available in the Python library pykooh (Kottke 2020a). The times series used in the comparisons are available at PEER ground motion website: https://ngawest2.berkeley.edu/site (Goulet et al. 2014, submitted, this issue).

Acknowledgments

This study was sponsored by the Pacific Earthquake Engineering Research Center (PEER) as part of the NGA-East research project, and was funded by the U.S. Nuclear Regulatory Commission (NRC), the U.S. Department of Energy (DOE), the Electric Power Research Institute (EPRI), and Pacific Gas & Electric (PG&E), with participation of the U.S. Geological Survey (USGS). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the sponsoring agencies, PEER, or the Regents of the University of California. We thank Armen Der Kiureghian for his thoughtful feedback and Nicolas Kuehn for his relentless assistance. The comments of two anonymous reviewers, the Editor, and the USGS reviewers that helped us to significantly improve the presentation of our work.
References


Kotke et al.

1687–1707.


Figure 1. (a & b) Fourier amplitude spectra (FAS) for two example earthquakes (black line). The FAS are modified by SDOF transfer functions with a damping of 5% and oscillator natural frequencies \( f_n \) of 0.1, 3, and 100 Hz. (c & d) Peak factors for the two example earthquakes computed by Cartwright and Longuet-Higgins (1956) [CLH56], Davenport (1964) [D64], Vanmarcke (1976) [V76], and Der Kiureghian (1980) [DK80] PF formulations for a range of oscillator natural frequencies at 5% damping.
Figure 2. Acceleration Fourier amplitude spectrum from Cap-Rouge event (M$_{\text{w}}$ 5.2 on 11/5/1997) recorded at station US.GOGA. (a) The H$_1$ (0°) and H$_2$ (90°) components are plotted along with the unsmoothed effective amplitude spectrum (EAS, Equation 15). (b) The EAS smoothed by the recommended KO98 window with $b = 188.5$ and provided at 100 point per decade.
Figure 3. Comparison of the Konno-Ohmachi smoothing window with four different smoothing bandwidths ($b_w$): 1/100, 1/50, 1/30, and 1/15.
Figure 4. Comparison of the KO98 weight with $b$ of 188.5 compared with $b$ of 20, commonly used in processing of spectral ratios, and a normalized 5%-damped SDOF transfer function.
Figure 5. Residuals of the misfit between the RVT with the V76 PF and optimized duration and the time series spectral accelerations. (a) Residuals at each frequency visualized with boxplots. The box spans from the lower (Q1) to upper (Q3) quartiles of the data with a line at the median. The whisker extends to the range of the data, but excludes data considered to be fliers. The whisker is limited to a maximum length of $1.5 \cdot (Q3 - Q1)$. (b) Standard deviation of the residuals.