Monte Carlo Simulation to Study the Spatial Variation of Ground Motion Associated with Basin Heterogeneities

Peyman Ayoubi*
Domniki Asimaki

Mechanical and Civil Engineering Department, California Institute of Technology, Pasadena, U.S.A

Abstract

Basin effects cause a complex wave interference inside a basin which can be attributed to basin versus bedrock material contrast and edge effect. This will have a significant impact on spatial variation, duration, and intensity of surface ground motion (SGM) during an earthquake. While important, the lack of sufficient information about material properties and stratigraphy of a basin prevents accurate simulation of the phenomena, particularly in high frequency regime. Stochastic analysis and the Monte Carlo technique are suitable approaches to address this issue, where basin material is represented by a correlated random field. In this study, We use a 2D finite element analysis of an idealized-shaped basin subjected to a vertically propagating SV plane wave and investigate the spatial variation of SGM associated with basin effects by assuming a correlated random field to represent basin material. We generate a random medium by adding perturbations to a homogeneous domain with various correlation lengths, coefficient of variations, and autocorrelation functions to evaluate their contribution to SGM. Our results show a difference between the output of homogeneous and stochastic models, where we conclude that the former would not represent basin response, especially in the high-frequency regime correctly. Among the parameters we consider, the coefficient of variation has the most influential impact on surface acceleration. We observe that increasing this parameter decreases the mean value of surface amplification while its standard deviation increases. In addition, correlation length affects the standard deviation of surface acceleration, but it does not significantly impact the mean amplification. As for the autocorrelation function, where we consider von Karman, Gaussian, and exponential, the results show that the trend of surface amplification does not change by choosing a different autocorrelation function. Finally, by comparing the 2D basin versus 1D layered medium, we show that one cannot accurately capture basin response by using a 1D analysis for seismic hazard quantification.

Keywords— Basin effects, Monte Carlo simulation, stochastic analysis, surface ground motion, spatial variation

*Corresponding author: Peyman Ayoubi `peyman.ayoubi@gmail.com`

1 Introduction

Numerical studies of basin effects require a detailed geotechnical and geological information as SGM rely on several parameters, namely bedrock depth [1], the frequency content of earthquake [2], material properties of basin and bedrock [3], among others [4]. However, due to the limited availability of testing and empirical data, simplifying assumptions regarding geometry and material properties is widely adopted. In the bulk of studies, a basin with homogeneous material (or a layered basin with a discrete velocity profile) is assumed (for example [5, 6]), and a single deterministic analysis is performed to analyze ground motion. This neglects the inherent heterogeneity in soil, when even a parametric study cannot fully account for it. This becomes important where an accurate simulation of a basin in high frequency is of interest.

Probabilistic modeling techniques and Monte Carlo (MC) simulations have been used in earth sciences (for example [7,8]) and can address the issue mentioned above by gauging the uncertainty arising from soil material spatial variability, and testing and statistical errors [9]. In seismology, earthquake source modeling [7, 10–12] and ground motion analysis [7, 13], among others, have been the primary usage of probabilistic approaches. On the other hand, engineers perform probabilistic studies to analyze liquefaction, site responses, to name but a few [8, 14–16]. These studies generate a correlated velocity random field to take material heterogeneity into account. For site effects application, the main focus has been on 1D wave propagation, or a rectangular 2D medium [8, 16], and the lack of such analysis for basin configuration is noticeable.

In a basin setting, two types of interaction are of interest. The first one occurs when a wave interacts with the basin as a whole (low-frequency interaction, similar to [4]). The second one happens when a seismic wave interacts with heterogeneities within a basin (high-frequency interaction). The second interaction results in a different basin response depending on the frequency and size of stochastic features [17]. The wave-basin interaction is easier to simulate, as has been done mainly in the literature [4, 6, 18, 19]. Wave-heterogeneities interaction is what most studies miss due to computational challenges, practicality, and insufficient geotechnical information. Neglecting the heterogeneities within a basin cannot be fully justified since it is known that small-scale heterogeneities can impact seismograms drastically through the broadening of body waves [20], distortion of radiation pattern [21], and elongation of surface motion [22].

This study analyzes the spatial variation of SGM associated with basin heterogeneities using a stochastic approach. We consider a 2D elastic basin overlying a bedrock subjected to a vertically propagating SV plane wave of the Ricker type. The basin geometry is assumed to be a generic form (half-cosine shaped similar to what studied by [4]). We use a correlated random field to take the spatial variability of materials into account in the numerical simulations. Parameters such as correlation lengths, autocorrelation function (ACF), and coefficient of variation of a random field are varied to evaluate their impact on SGM. In the following sections, we first explain the correlated random field generation procedure. Next, we present the results and demonstrate how each parameter would affect the spatial variation of SGM. Finally, we show a comparison versus 1D analysis, which is the state of practice.

2 Monte Carlo Simulation

In this section, we first explain the Finite Element Model (FEM). Next, we detail the random field generation procedure. The random field is generated on top of a homogeneous basin, and this homogeneous background is called "background medium/model" throughout this article.

2.1 Finite Element Model

For the numerical simulation, we use SVL [23]. SVL is an open-source, easy-to-use, fast, and extendable C++ finite element code that is designed to optimally perform linear and nonlinear wave-propagation simulations in meso-scale. The software offers various options such as dynamic solvers for time-domain

Basin Properties			Bedrock Properties		
Shear Velocity	Density	Poisson's ratio	Shear Velocity	Density	Poisson's ratio
$[\frac{m}{s}]$	$[\frac{kg}{m^2}]$	$[-]$	$[\frac{m}{s}]$	$[\frac{kg}{m^2}]$	$[-]$
250	2000	0.333	500	2000	0.333

Table 1: Basin and bedrock material properties in background medium.

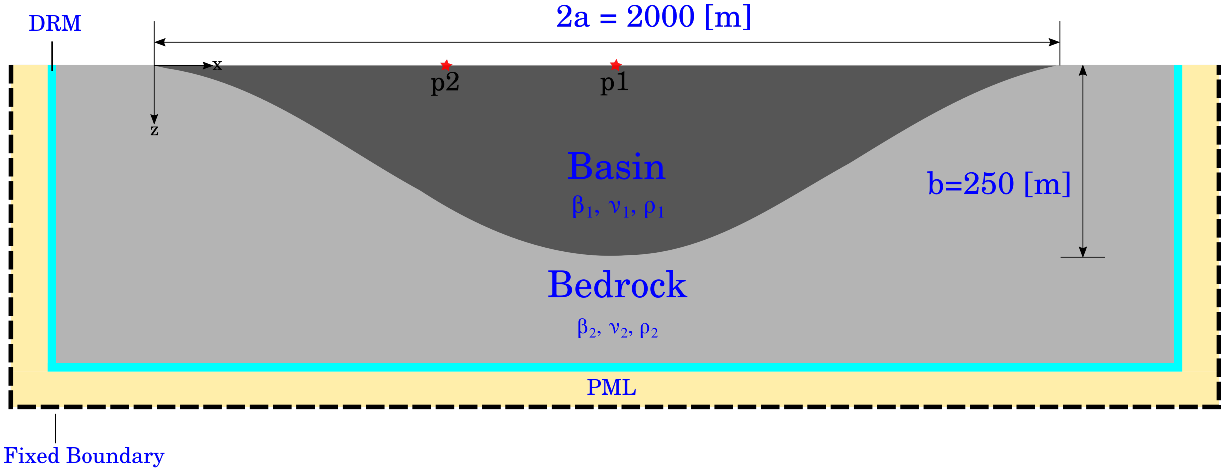

Figure 1: The numerical domain. The dark gray area represents the basin. The light gray shows the bedrock. Cyan and yellow regions demonstrate DRM and PML, respectively. The external boundary is fixed. Two red stars show the location at which we will show results later in this article: in the basin middle, halfway between center and corner.

analyses of inelastic problems, and Perfectly Matched Layer (PML) to efficiently mimic realistic scenarios. We use elastic materials to represent a basin and bedrock. We neglect damping in the basin to prevent exhausting high-frequency energy of the seismic wave and better represent the impact of features on SGM. For the basin configuration, we choose a generic cosine-shaped basin, similar to Figure 1. Table 1 shows the material properties of the background medium, which will be used to generate the correlated random field (details in § 2.2).

The geometry of the basin and dimensionless parameters are derived based on the analysis performed by [4]. We use following dimensionless parameters to define the model: aspect ratio $AR = a/b = 4.0$, material contrast $C = \beta_2/\beta_1 = 2.0$, density ratio $\rho_2/\rho_1 = 1.0$, Poisson's ratio of basin $\nu_1 = 0.333$, and Poisson's ratio of bedrock $\nu_2 = 0.333$. The input motion is shown in Figure 2 which is a Ricker wavelet with the dominant frequency of $5\,Hz$. We expect two separate interactions that might occur inside a basin, namely 1) wave-basin interaction and 2) wave-features interactions in horizontal and vertical directions. The reason for choosing $5\,Hz$ as the dominant frequency of the input motion is the availability of sufficient energy at both low and high frequencies. This means that the Fourier spectral amplitude of input motion has enough energy to interact with the heterogeneities inside a basin.

We use 15 quad elements (each having size of dx) per shortest wavelength in a domain to resolve maximum frequency and discretize the numerical domain. We also satisfy CFL $<$ 0.25 condition by assuming $dt = 0.005$. Perfectly Matched Layers (PML) with a length of $25 \times dx_{bedrock}$ are prescribed at the side and bottom boundaries to absorb scattered wavefield. The outer boundary of PML is fixed. Finally, the input motion is applied using Domain Reduction Method (DRM) at the cyan shaded region of Figure 1.

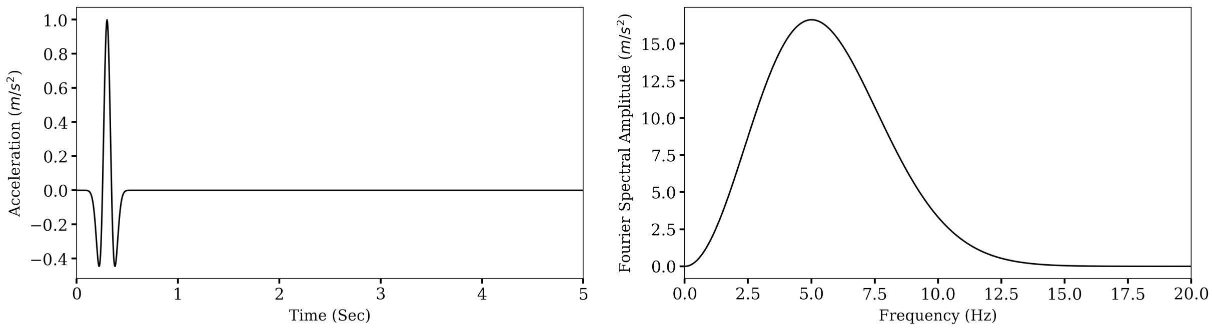

Figure 2: Input acceleration time-series (left) and corresponding Fourier spectral amplitude (right).

2.2 Random Field

In a deterministic scenario, due to a lack of sufficient information about soil mechanical properties, one needs to either assume fixed material properties for a basin or adopt a layered velocity profile. However, the Earth's crust is far from being homogeneous or isotropic, the reason we are using a random field to describe the material inside a basin. This does not mean that the basin is a fully randomized medium while correlations exist. The correlated random field we are using requires 5 parameters, namely 1) mean (μ), 2) Coefficient of Variation (COV), 3) Autocorrelation functions (ACF), and correlation lengths in 4) horizontal (θ_x) and 5) vertical (θ_z) directions. In practice, for a geotechnical application, skewed distributions are used for shallow layers and a symmetric distribution for deep [24]. In this study, given the depth of the basin, we utilize a symmetric distribution. In addition, due to the arbitrary geometry of the basin (as opposed to a rectangular domain), we use Karhunen-Loève expansion to generate a correlated random field [25, 26]. Similar to [26], algorithm 1 shows the procedure to generate a correlated random field.

Algorithm 1 Generation of a correlated random field.

1: The **initialization stage:** Define random field parameters:

$$\mu, COV, ACF, \theta_x, \theta_z.$$

2: **Correlation matrix generation:** Generate the correlation matrix (M) based on the FE mesh nodes and above parameters. If there are n nodes in a basin, $M \in R^{n \times n}$.

3: **Calculating eigenvalues and eigenvectors of M:** Calculate eigenpairs (λ_i, ϕ_i) of correlation matrix M, where λ_i is i-th eigenvalue and ϕ_i is the i-th eigenvector.

4: **Estimating correlated random field:** The correlated random field can be calculated as:

$$V_s(x) = \mu + \sigma \Sigma_{i=1}^{m} \sqrt{\lambda_i} Z_i \phi_i \quad (1)$$

where Z_i is an independent sample drawn from the standard normal distribution. m is estimated using the Cholesky decomposition.

The following subsections detail different components of correlated random field generation of Algorithm 1.

2.2.1 von Karman Autocorrelation Function

The correlation structure of a random field can be characterized using an ACF in the spatial domain or its power spectral density (PSD) in the Fourier domain [10]. In this study, we use von Karman [27] ACF

No.	Properties			
	μ (m/s)	COV	θ_x (m)	θ_z (m)
1	250	0.2	50	20
2			100	20
3			100	40
4			200	20
5			200	40
6		0.3	100	20
7			200	20
8		0.4	50	20
9			50	40
10			100	20
11			100	40
12			200	20
13			200	40

Table 2: Statistical parameters for different models of this study.

as shown in Eq. 2 [10]. von Karman ACF is an appropriate choice for solid Earth material [28], and we use it as the main ACF in future analysis.

$$C(r) = \frac{G_H(r)}{G_H(0)} \tag{2}$$

where $G_H(r) = r^H K_H(r)$. Here, H is the Hurst exponent (we assume $H = 0.1$), K_H is the modified Bessel function of the second kind (order H). r is called the characteristic length and is given as a function of horizontal and vertical correlation lengths $r = \sqrt{\frac{x^2}{\theta_x^2} + \frac{z^2}{\theta_z^2}}$. Having set the ACF, the remaining four parameters will be discussed in the following subsection. We will also discuss the impact of ACF on SGM in § 3.4 by considering two other ACFs, namely Gaussian and exponential.

2.2.2 Statistical Parameters

Table 2 shows values for the parameters mentioned before. To generate the random medium, we use Gaussian distribution similar to [7]. We define the distribution using mean μ and coefficient of variation ($COV = \sigma/\mu$, σ is the standard deviation). The combination of values of Table 2 are used to study the effect of horizontal and vertical correlation lengths, and COV. $\mu = 250\,m/s$ (taken from background medium), $COV = 0.2$, 0.3, 0.4, $\theta_x = 50\,m$, 100, $200\,m$, and $\theta_z = 20\,m$, $40\,m$ are used. Based on the frequency content of input motion, the values of θ_x and θ_z provide an impact on the resultant surface response at different frequency ranges, as will be discussed later. These values result in 13 separate numerical models as shown in Table 2. The parameters are chosen to cover a reasonable range of values based on the geometry of the basin and introduce a sufficient variation in material properties in a basin. In the table, the first 5 models has $COV = 0.2$. Models 6 and 7 are added to better represent the impact of COV when moving from $COV = 0.2$ to $COV = 0.4$. The last 6 models have $COV = 0.4$.

2.2.3 Statistical Significance of Monte Carlo Simulation

An essential parameter in doing Monte Carlo simulation is the number of realizations to guarantee statistical significance. We measure statistical significance by repeating the simulations until the convergence rate (CR, Eq. 3) is less than 5% for a specific intensity measure (IM) of interest. We examined PGA

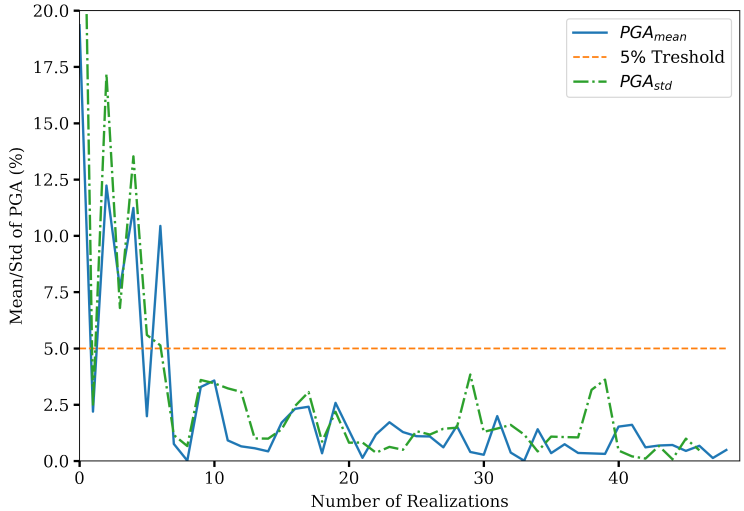

Figure 3: Convergence of the Monte Carlo simulation for mean and standard variation of PGA at the middle of #3 basin.

similar to [8]. PGA is chosen for model #3 as an example given that this model has parameters that would result in the worst-case scenario.

$$CR = \frac{|IM_i - IM_{i+1}|}{IM_i} \tag{3}$$

where i is the realization number. Figure 3 shows the variation of the mean and standard deviation of PGA in percent in the middle of the basin of model #3 from Table 2. To assure the statistical significance of all models in our analysis, we use 50 realizations. Note that basin configuration is helping with the statistical significance. This is the reason why a higher realization number is necessary for non-basin geometries. For example, in [8], 100 realizations are used to satisfy the 5% criterion.

3 Results

In this section, we present the results of the numerical simulations. We elaborate on them by examining both time domain and frequency domain responses. We define the time domain amplification factor as PGA at each point of basin surface divided by the far-field PGA (it's 2 due to the doubling effect). In addition, seismogram synthetics of surface motion is also shown to represent the spatio-temporal variation of SGM better. On the other hand, the frequency domain amplification factor is defined as the amplitude of the Fourier transform of a time-series divided by the corresponding Fourier transform amplitude of rock outcrop. We also assess the impact of stochastic patches on Arias Intensity and the fundamental frequency of a basin. Comparisons are made by considering a point at the center of the basin ($p1$), see Figure 1). In the first subsection, we study one realization from each model #1, #2, #3 together with the background model. Note that *benchmark* is a stochastic case while the *background* is a homogeneous model. We use the benchmark model to observe how a stochastic model differs from the background model. Next, we examine different statistical parameters that shape a correlated random field (Table 2) by studying the ensemble of realizations both in time and frequency domains.

3.1 Benchmark

As the benchmark example, we use one realization from three models and compare the results with the background model. The models are $\theta_x = 100\,m$, $\theta_z = 20\,m$, $COV = 0.2$ (model #2), $\theta_x = 50\,m$, $\theta_z = 20\,m$, $COV = 0.2$ (model #1), and $\theta_x = 100\,m$, $\theta_z = 20\,m$, $COV = 0.4$ (model #3). These examples are used to portray the effect of a random medium on SGM and how it is different from the background model. We specifically focus on COV and θ_x here, and we leave θ_z and ACF for a later section. Each column of Figure 4 shows responses corresponding to the top-row velocity profiles (Figure 4-a). As can be seen, the left column shows the background model, the second from left is model #2, the second from the right is model #1, and the right column shows model #3. Note that these realizations are examples of possible correlated random fields, and the results do not speak for the ensemble of realizations. The background model is a homogeneous basin, and the value of shear wave velocity equals the value in Table 1. By comparing one realization versus the other two, one can see the effect of θ_x and COV on the generated random field. Increasing θ_x results in a stretched stochastic patches in a basin (comparison of model #1 and model #2) while increasing COV results in a more pronounced shear wave velocity variation in a basin (comparison of model #1 and model #3). These changes will directly impact the wave interference in a basin. For the comparison, we will consider point $P1$ of Figure 1.

Figure 4-b demonstrates horizontal and vertical time-series at $p1$. It turns out that the amplitude of horizontal and vertical acceleration is lower than the background model due to the scattering effect of heterogeneity patches. Given the horizontally polarized input motion, the accentuated vertical component is of importance for all the realizations. This will introduce a torsional motion in the basin which could impact long infrastructure components, such as pipelines. In addition, an elongation in time-series duration is observable, especially for model #3, which is in addition to the fact that even a homogeneous basin would increase the duration of oscillation on the surface compared to the rock outcrop. Therefore, adding the heterogeneities will exacerbate the situation. This phenomenon happens for both horizontal and vertical components.

Figure 4-c portrays the surface amplification in time-domain. It shows the division of PGA at each point on basin surface with respect to PGA of rock outcrop in x direction (which is twice the input PGA). This figure better quantifies the spatial variation of surface amplification. While the middle points show both lower (model #3) or higher (model #2) amplification with respect to the background model in the horizontal direction, the SGM is exacerbated due to the existence of a randomized medium as we move toward basin edges. In addition, a significant amount of amplification occurs in the vertical direction in the basin center while the input motion is purely horizontal. These observations are a testament to the insufficiency of a homogeneous representation to fully take variation and magnitude of SGM into account, where the asymmetry and large variation of amplification with distance better resemble what is observed in Nature even for a simplified symmetric basin geometries adopted in this study. Note that the computational cost is the main drawback of such an analysis in practice. In this study, since we are using 50 realizations, each model of Table 2 has 50 times more computational cost than the background model on average. This can be justified in high-importance projects where basin response may cause an irreversible impact on basin surface.

Lastly, Figure 4-d and Figure 4-e demonstrate the surface seismogram synthetic in horizontal and vertical directions, respectively. A relatively noticeable difference between randomized and background models is the "activation" of more local interactions and accentuated reverberation. As can be seen, more intense oscillations happen on the basin surface in the stochastic case, which is due to the fine-scale stochastic features. This could have a significant impact on structures and infrastructure objects that would not be affected by the low-frequency component of the seismic wave.

Figure 5 demonstrates the difference between each realization by focusing on their relative behavior with respect to the rock outcrop in the frequency domain. This figure portrays the Fourier spectral amplification. The Fourier spectral amplification is defined as the amplitude of the division of the Fourier transform of a time-series in a specific location ($p1$ in this case) with respect to the corresponding Fourier transform

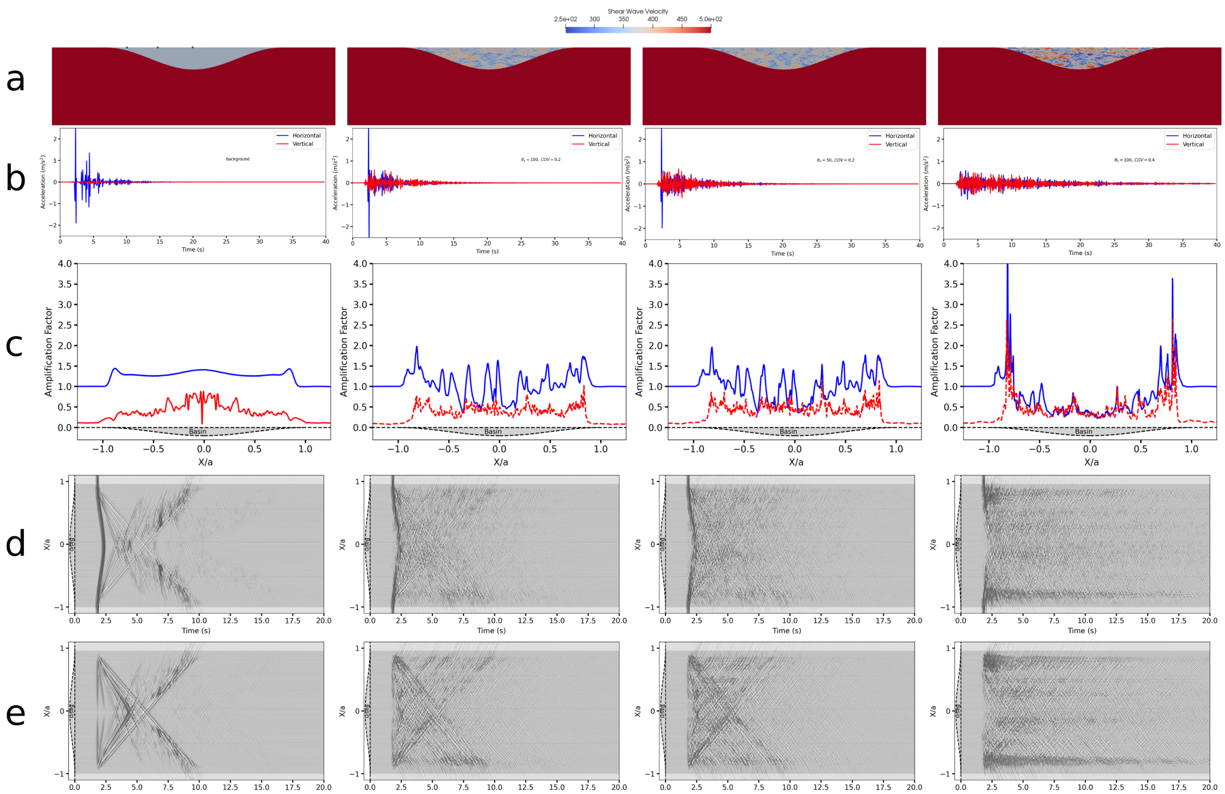

Figure 4: a) Background model and one realization of model #2 ($\theta_x = 100\,m$, $\theta_z = 20\,m$, $COV = 0.2$), model #1 ($\theta_x = 50\,m$, $\theta_z = 20\,m$, $COV = 0.2$), model #3 ($\theta_x = 100\,m$, $\theta_z = 20\,m$, $COV = 0.4$), b) acceleration time-series at point $p1$, c) surface amplification in horizontal and vertical directions, d) seismogram synthetic for surface acceleration in horizontal direction, e) seismogram synthetic for surface acceleration in vertical direction.

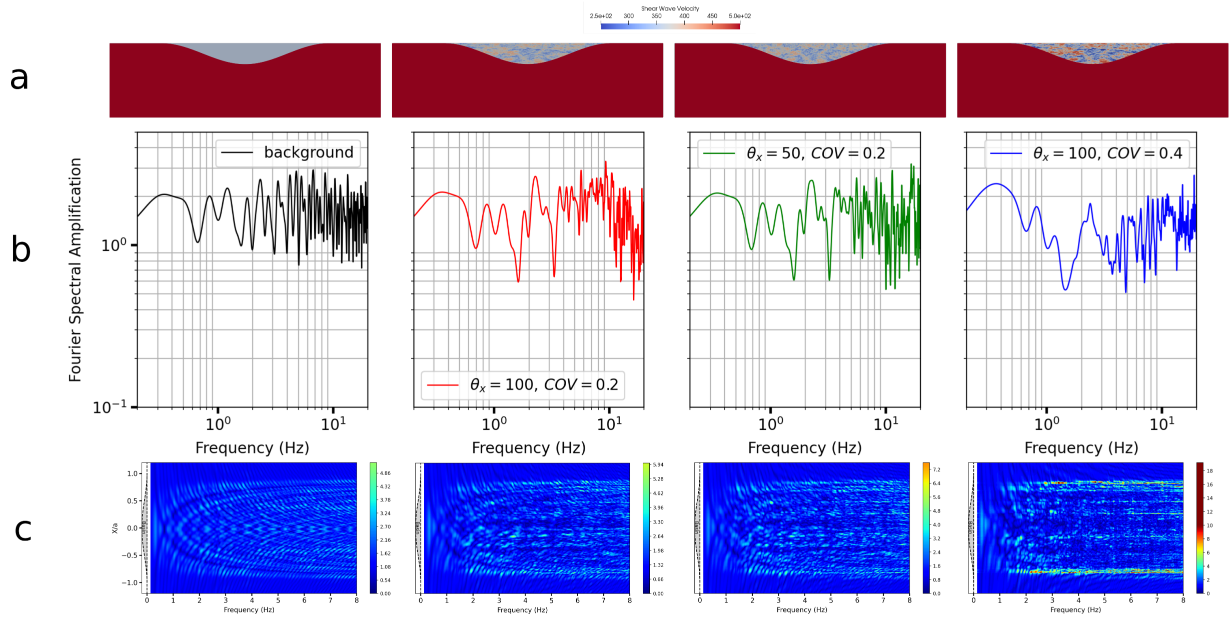

Figure 5: Similar to Figure 4. b) Fourier spectral amplification with respect to rock outcrop. c) equivalent of seismogram synthetics of Figure 4-d except that it shows Fourier amplification at different frequencies. The colorbar is capped at 10.

on the rock outcrop. In this figure, each column shows the results of the corresponding realization at the top, including the background model, similar to Figure 4. In Figure 5-b, the general trend of amplification is similar between background model and realizations, which demonstrates the impact of basin geometry on the surface response. As was seen in Figure 4, it also depends on the location at which the results are being plotted, and the response varies as one moves along the basin. Stochastic patches also impact a basin's fundamental frequency, which will be explained in a later section. Finally, Figure 5-c shows a response similar to Figure 4-d except here Fourier spectral amplification is portrayed. The symmetry in the background model no longer exists in stochastic cases. A more considerable amplification at basin edge exists for stochastic cases, which does not exist in the background model. This is a confirmation of the previous observations in Figure 4-c, where a significant amplification happened at basin corners.

In the following subsections, we examine the impact of statistical parameters on SGM. COV is studied in § 3.2, Correlation lengths (θ_x and θ_z) in § 3.3, and ACF in § 3.4. The following figures are being discussed in later sections: 1) mean and standard deviation of Fourier spectral amplification with respect to the rock outcrop, 2) similar to Fourier spectral amplification, response spectra amplification and its standard deviation will be discussed, 3) significance duration ratio is computed with respect to the background medium. Significance duration is defined as the duration between 5% and 95% of final Arias Intensity, and 4) fundamental frequency ratio is computed by dividing the fundamental frequency of the stochastic model with respect to the background medium.

3.2 Effect of Coefficient of Variation

COV defines the standard deviation of shear wave velocity in the random field as shown in Algorithm 1. By increasing COV, the random field would take a wider range about the mean. This introduces larger shear wave velocity contrasts within a basin, which results in more scattering. For example, given the mean shear wave velocity of $250\,{}^{m}/_{s}$ (see Table 1), $COV = 0.2$ results in standard deviation of $50\,{}^{m}/_{s}$ while $COV = 0.4$ introduces a standard deviation of $100\,{}^{m}/_{s}$. Such a difference would significantly impact the shear wave velocity in a basin and its response under seismic loading.

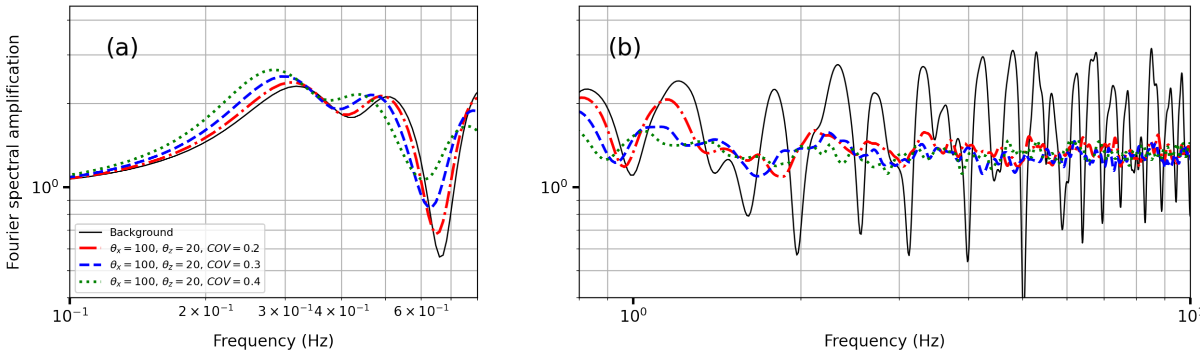

Figure 6: Mean Fourier spectral amplification at point $p1$ for three models with $\theta_x = 100\,[m], \theta_z = 20\,[m]$ and different COVs. Background model is shown in black. Fourier spectral amplification is defined as the ratio of Fourier transform of acceleration time-series at point $p1$ with respect to the rock outcrop. a) frequency range of $0.1\,[Hz]$ to $0.8\,[Hz]$, b) frequency range of $0.8\,[Hz]$ to $10.0\,[Hz]$.

Figures 6 and 7 show the Fourier spectral amplification and standard deviation. Figure 6 is divided into two subfigures, one covering the frequency range of $0.1\,Hz$ to $0.8\,Hz$ and the other $0.8\,Hz$ to $10\,Hz$. In these figures, we focus on three models in addition to the background. For most of the analysis of this study, model #2 is the benchmark, and for each statistical parameter, the impact is investigated by changing values with respect to this model. By holding θ_x and θ_z fixed, and changing COV, in Figure 6-a, one can see that the fundamental frequency of basin moves toward lower frequency as we increase COV. This can be attributed to the higher variation of shear wave velocity in the basin, resulting in an accentuated lateral propagation of the seismic wave. The changes in the fundamental frequency will be better quantified in Figure 9. Lower COV value results in a medium that is more similar to the background model, which is the reason that responses of models with $COV = 0.2$ better resembles the background medium than $COV = 0.3$ and $COV = 0.4$. In Figure 6-b, the background medium shows more oscillatory amplification in comparison to the mean of stochastic models. However, the advantage of a stochastic simulation is to provide a range in which the amplification value could happen, i.e., considering the standard deviation is an essential part of the analysis. Given that the medium is generated using Gaussian distribution, $\pm 2\sigma$ (σ is the standard deviation. We limit velocity values to be within $\pm 2\sigma$ of μ.) of mean is probable to happen for each model. This could result in a higher amplification than the background model.

Figure 7 shows the standard deviation of Fourier spectral amplification. Regardless of the frequency range, models with higher COV have a higher standard deviation in frequencies less than $2\,[Hz]$. This is due to the enhanced shear wave velocity variation inside the basin. By moving toward higher frequencies, the impact of the range of shear wave velocity values becomes less pronounced since all the heterogeneities would interact with the seismic wave inside the basin, which will produce a a similar standard deviation for all models.

Figure 8-a is similar to Figure 6 except that it shows the mean response spectra amplification and its standard deviation. Response spectra amplification shows a cleaner representation of low-period (high-frequency). As can be seen, by adding stochasticity to the basin, the amplification decreases and higher values of COV results in de-amplification. Figure 8-b shows the standard deviation of response spectra amplification, and similar to Figure 7, it shows that a higher COV would result in a higher standard deviation.

To further quantify the impact of COV on significance duration and fundamental frequency of a basin, Figure 9 demonstrates the two outputs as a ratio with respect to the background medium. The circles show the mean in this figure, and bars demonstrate 1 standard deviation from the mean. Figure 9-a shows the significant duration ratio. Significance duration is defined as the time between 5% to 95% (T_{5-95})

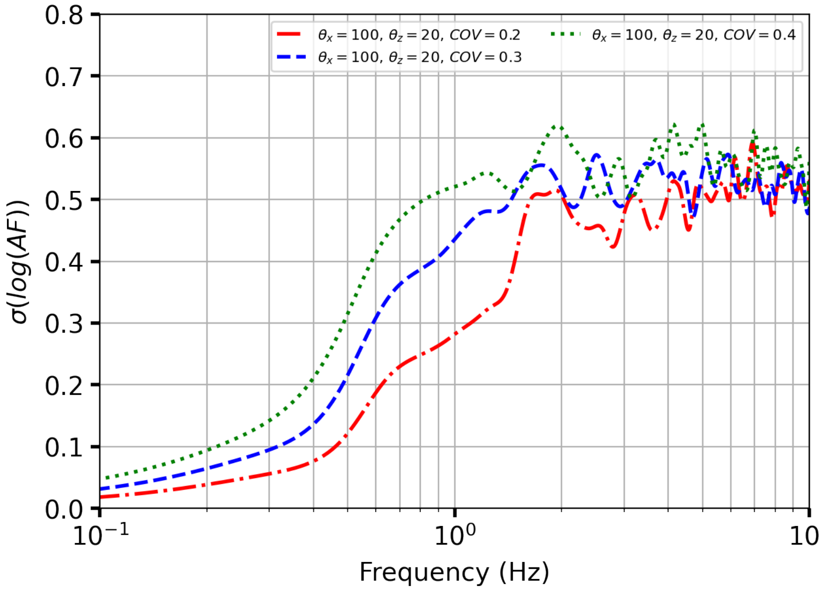

Figure 7: Standard deviation of natural logarithm of Fourier spectral amplification of models in Figure 6.

of Arias Intensity final value. As can be seen, increasing COV results in a higher mean and standard deviation, which can be explained by enhanced scattering of seismic waves for larger COV. $COV = 0.2$ shows almost $1.5 - 2$ increase and $COV = 0.4$ has values as high as $5 - 6$ times the background model. $COV = 0.4$ shows a strong reverberating effect due to heterogeneities which significantly elongate the shaking duration. Figure 9-b, on the other hand, demonstrates the fundamental frequency (f_0) ratio. Increasing COV results in a lower f_0, as was mentioned before. The trend of the mean ratio seems to be almost linear as a function of COV and does not change with correlation lengths. Note that although the mean f_0 ratio decreases as COV goes up, the standard deviation of the ratio increases. This figure shows that the fundamental frequency can go as low as 0.9 times the background medium.

3.3 Effect of Correlation Length

The next component of a correlated random field that we will examine is the correlation lengths in horizontal (θ_x) and vertical (θ_z) directions. As was shown previously, we consider different values for each direction, namely $\theta_x = 50,\ 100,\ 200\,m$ and $\theta_z = 20,\ 40\,m$. The values are chosen based on two criteria: 1) the size of patches is large enough to be "seen" by the seismic wave, and 2) we have enough patches in the basin in each direction. The second criterion is to ensure enough interaction between the seismic waves and heterogeneities. θ_x and θ_z show the distance from a point above which two points are not correlated. Changing the size of patches affects the interaction of seismic waves (patches behave as points of scattering) and cause reflection and refraction of waves within a basin. This might eventually affect the observed SGM. Increasing the size of patches could push the wave-patch interaction toward a larger wavelength (or lower frequency).

Assuming an average of $250\,{}^{m}/_{s}$ for shear wave velocity and a maximum frequency of $12\,Hz$, the minimum wavelength is $\sim 21\,m$. By decreasing the frequency, the wavelength increases, and different interaction scales are expected depending on the size of heterogeneities. For the horizontal component, the interference between seismic waves and stochastic patches happens due to mode conversion at the basin corner or the decomposition of the incoming wave when entering the basin, dependent on basin-edge effect and shear wave velocity contrast. In general, increasing patch size shows the impact on surface acceleration at a lower frequency, especially in terms of standard deviation, as will be shown later.

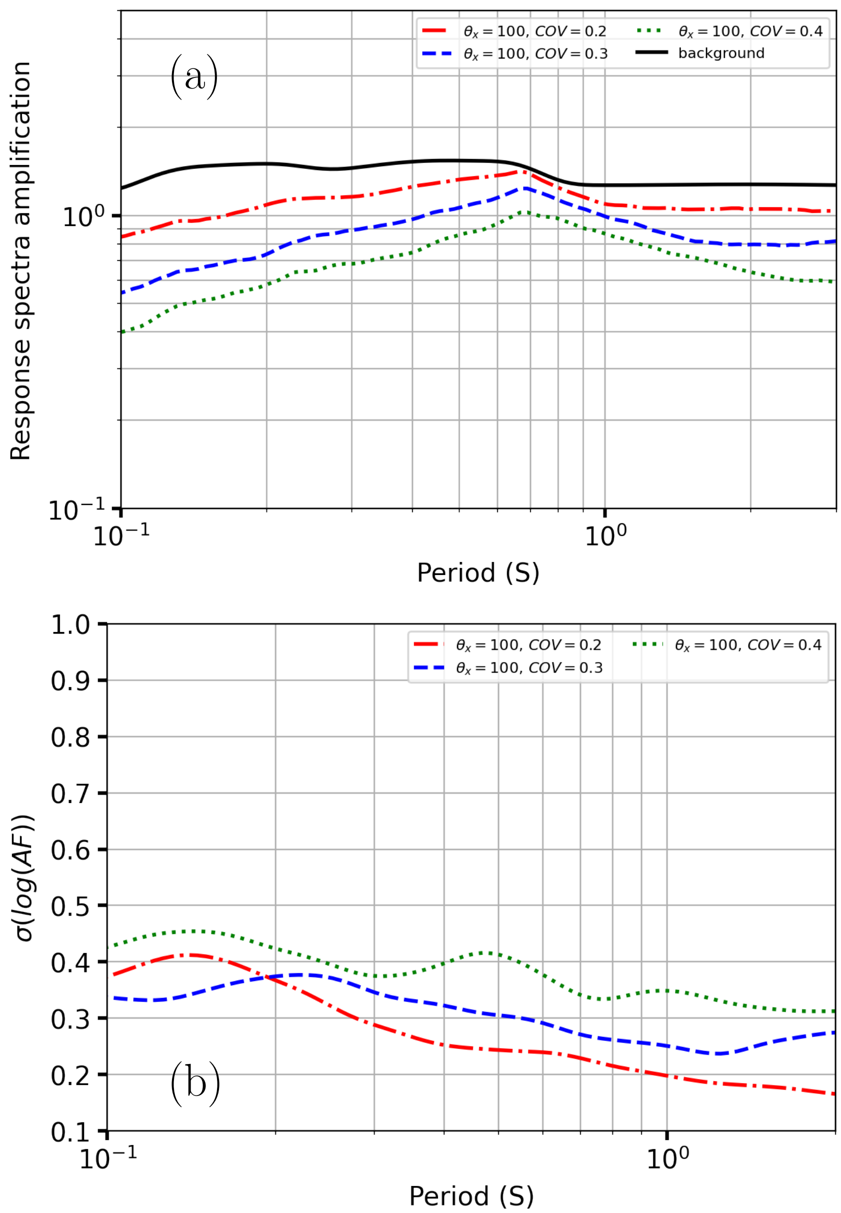

Figure 8: a) Mean and b) standard deviation of response spectra amplification for point $p1$ for three models with $\theta_x = 100\,[m], \theta_z = 20\,[m]$ and different COVs.

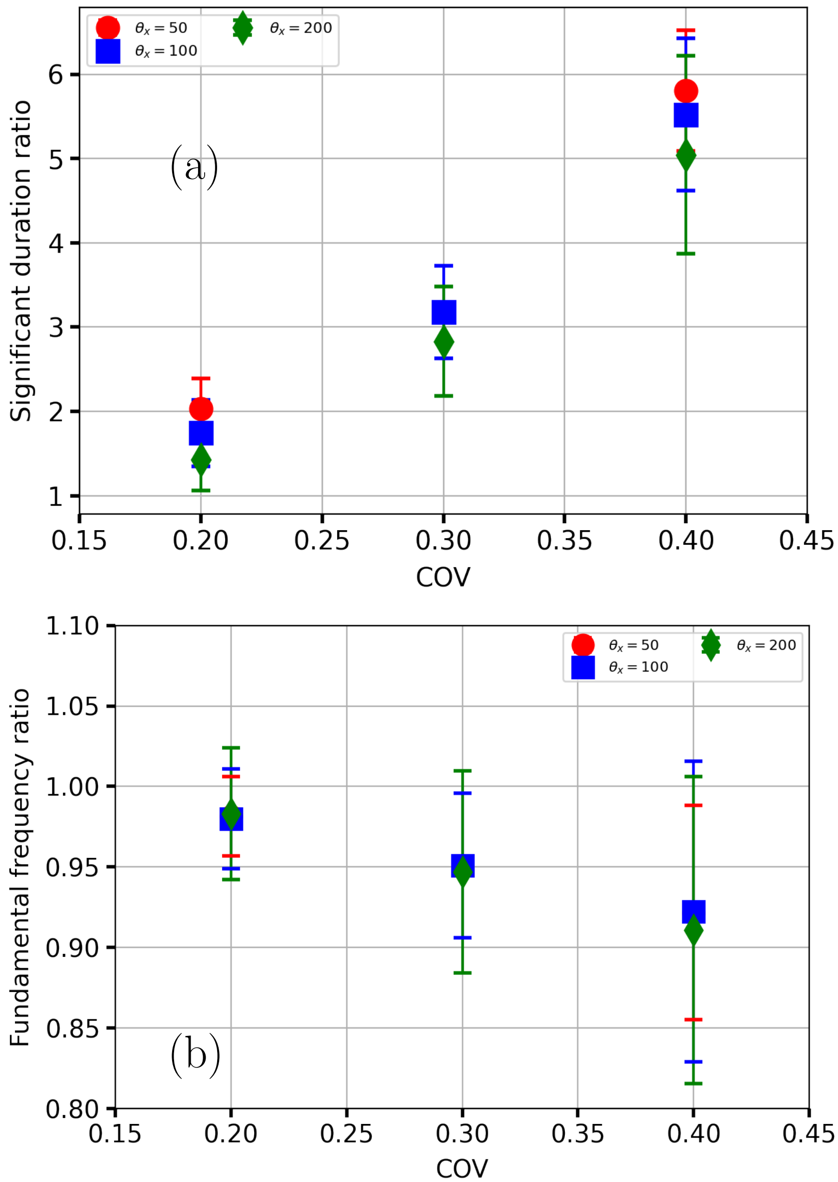

Figure 9: Time and frequency domain response for ensemble of realizations for 8 models of Table 2 with $\theta_z = 20\,[m]$. Results are shown for $p1$. Circle symbols shows $\theta_x = 50\,[m]$, square shows $\theta_x = 100\,[m]$, and diamond shows $\theta_x = 200\,[m]$. a) significance duration ratio, b) fundamental frequency ratio.

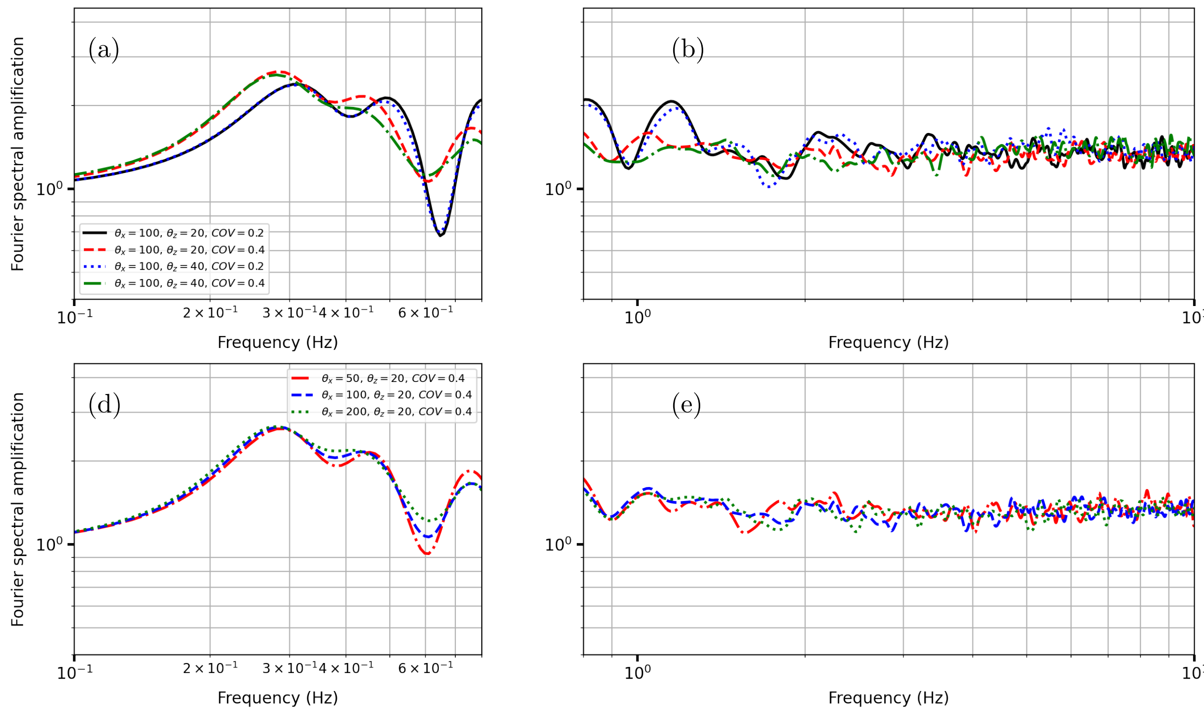

Figure 10: a,b) mean Fourier spectral amplification to study the impact of θ_z. Four models with fixed $\theta_x = 100\,[m]$, and $\theta_z = 20, 40\,[m]$, and $COV = 0.2, 0.4$. c, d) Fourier spectral amplification to study the impact of θ_x. Three models with fixes $\theta_x = 20\,[m]$ and $COV = 0.4$, and three values of $\theta_x = 50, 100, 200\,[m]$.

In deciding the importance of correlation lengths, focusing on a frequency range is necessary. The seed for random number generation is fixed to guarantee the reproducibility of a correlated random field for all the models shown here. This means that for a fixed COV, for a certain realization, models with different θ_x will have similar material representation except for a stretch or shrink of patches based on the magnitude of θ_x. Similar behavior is expected for vertical correlation. Therefore, a patch with $\theta_x = 200\,m$ would interact with lower frequency components than a patch with $\theta_x = 100\,m$.

Figure 10 shows two sets of comparison. Figures 10-a and b show the mean amplification of four models. These figures are supposed to show the impact of θ_z on surface amplification. We fix $\theta_x = 100\,[m]$ for this figures to assess how θ_z would change the mean amplification. We do not fix COV to investigate whether the impact of θ_z changes with COV. The role of θ_z can be explained by considering the overall configuration of the model. As for the first arrival of the SV plane wave, the horizontal component of motion shown here is impacted by θ_z since the wave is horizontally polarized but vertically propagated. However, the influence is not noticeable for lower COV. It becomes more pronounced for larger COV as a separation between curves for $\theta_z = 20\,[m], COV = 0.4$ and $\theta_z = 40\,[m], COV = 0.4$ happens. As for the θ_x, Figures 10-c and d portray a similar pattern as Figures 10-a and b. Note that in this figure, we only use $COV = 0.4$ to investigate the worst case scenario. The influence of θ_x on surface amplification is through affecting surface generated waves and decomposed body waves inside the basin due to the interference between seismic wave and material heterogeneities. These figures show that the level of impact of θ_x is not as much as to produce a meaningful difference between different models in terms of mean amplification.

To further study correlations lengths, Figure 11 shows the standard deviation of amplification at point $p1$ for vertical and horizontal correlation lengths. The influence of correlation lengths can be better seen in standard deviation plots. For θ_z, Figure 11-a shows two groups of graphs, which are separated by COV.

As was previously observed, COV is the dominant parameter in the correlated random field. For θ_x, in Figure 11-b, model with $\theta_x = 200\,m$ has a higher standard deviation than the model with $\theta_x = 100\,m$ (at a fixed COV) at lower frequencies. However, the difference narrows as we move toward higher frequencies. Note that in the comparisons shown here, since the values are chosen realistically, and are relatively close to each other, the distinction may not be as significant as one would expect. The difference would have been easier to distinguish if we had for instance $\theta_x = 200\,m$ and $\theta_x = 20\,m$. Similar observations about correlation lengths were also observed in other studies such as [8] for a rectangular domain.

As the final illustration to confirm our previous observation about the impact of correlation length on SGM, Figure 12 shows significance duration ratio and fundamental frequency ratio as a function of θ_x. Three difference COVs are shown with different symbols. The horizontal axis shows θ_x, and the vertical axis indicates an output of interest. As can be seen, for both significance duration ratio and fundamental frequency ratio, changing θ_x does not significantly impact the outcome, which confirms our earlier findings.

3.4 Effect of Autocorrelation Function

Up to now, we have used Von-Karman ACF to generate a correlated random field. There are other common autocorrelation functions that have been used in the literature, namely Gaussian (Eq. 4) and exponential (Eq. 5).

$$C(r) = e^{-r^2} \tag{4}$$

$$C(r) = e^{-r} \tag{5}$$

In this section, we examine the difference in basin surface response given different ACFs. For this comparison, we use model #2 from Table 2 and generate realizations based on different ACFs. Figure 13 shows three example realizations. Gaussian has the smoothest variation of shear wave velocity over stochastic patches as was shown repetitively before, for example [7]. Since the seed number is fixed, the general shape of heterogeneity patches is similar. The differences arise as to how fast/slow the correlation decays over distance.

Figure 14 shows the comparison results. As can be seen, the trend of amplification doesn't change for different ACFs. In Figure 14-a, the mean Fourier spectral amplification curves follow a similar path for all frequency ranges with a minimal difference between the models. For large frequencies, exponential ACF has a higher amplification over a wide frequency range. Figure 15 shows the standard deviation of the natural logarithm of amplifications. Gaussian has a higher standard deviation value in comparison to others in low frequency. This is due to the smooth drop of Gaussian ACF (Eq. 4), which constructs a larger chunk of heterogeneity. More interaction with seismic waves could result in a more significant reflection/refraction in the basin, increasing fluctuating behavior and standard deviation on the surface. As for higher frequency, there is no clear ACF that dominates the other two. We conclude that in the scenario studied in this article, ACF is not an influential parameter on surface response.

4 Comparison between 1D and 2D analysis

In this section, we assess how the analysis so far would compare with the state-of-practice. In practice, $1D$ wave propagation is the common approach for seismic hazard quantification. We intend to examine whether a $1D$ wave analysis is able to account for the response of a $2D$ heterogeneous basin. We choose three stochastic models, namely $\theta_x = 100\,[m]$, $\theta_z = 20\,[m], COV = 0.2$, $\theta_x = 100\,[m]$, $\theta_z = 40\,[m], COV =$

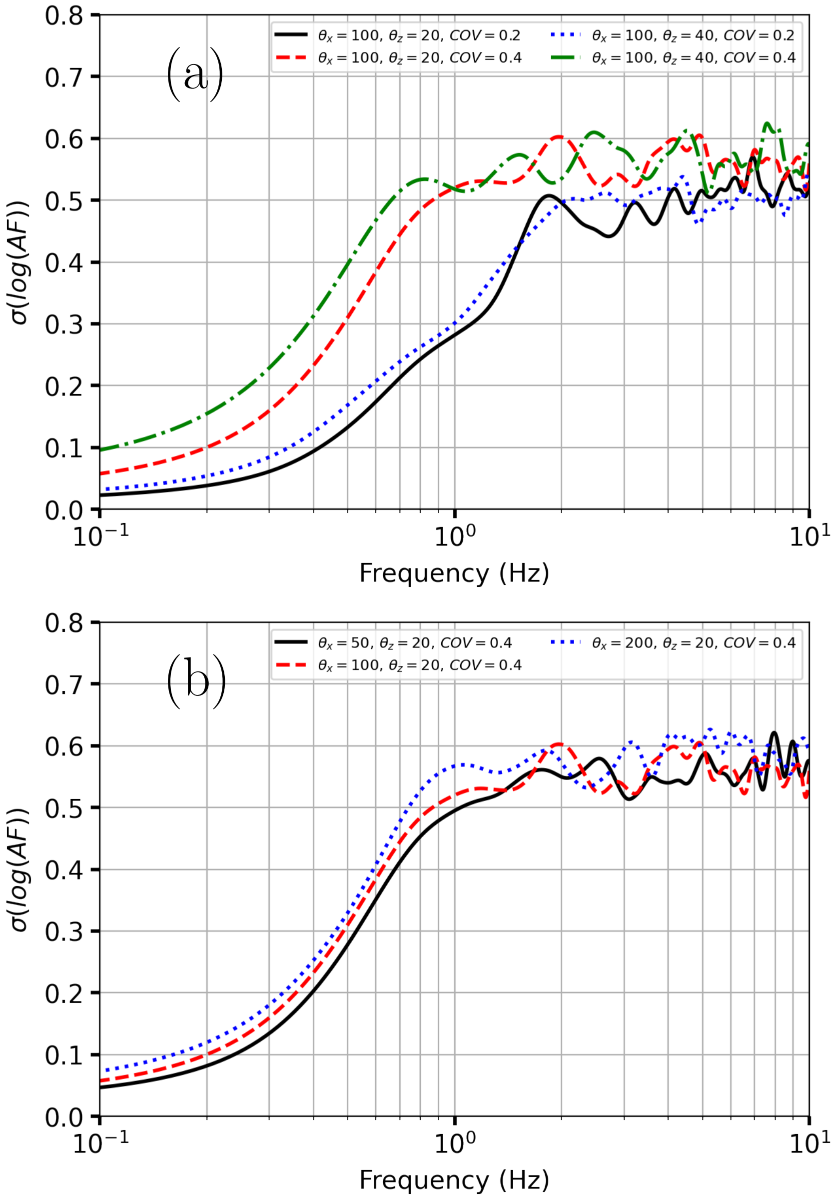

Figure 11: Standard deviation of natural logarithm of amplification for a) models with $\theta_x = 100$, $\theta_z = 20, COV = 0.2$, $\theta_x = 100\,\theta_z = 20, COV = 0.4$, $\theta_x = 100$, $\theta_z = 40, COV = 0.2$, and $\theta_x = 100$, $\theta_z = 40, COV = 0.4$, and b) $\theta_x = 50$, $\theta_z = 20, COV = 0.4$, $\theta_x = 100$, $\theta_z = 20, COV = 0.4$, and $\theta_x = 200$, $\theta_z = 20, COV = 0.4$.

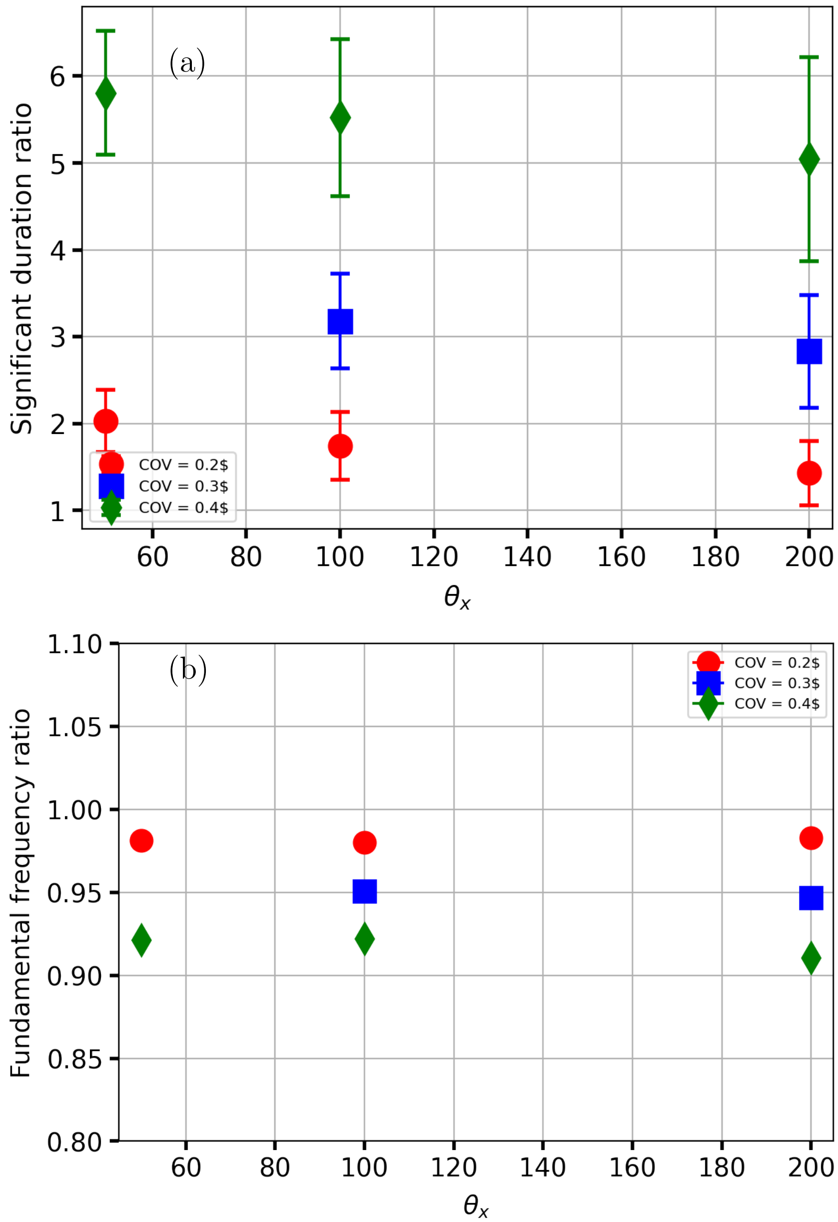

Figure 12: Time and frequency domain response for ensemble of realizations of 8 models of Table 2 with $\theta_x = 20\,[m]$ with respect to the background medium. The circle symbol shows $COV = 0.2$, the square shown symbol shows $COV = 0.3$, and diamond symbol shows $COV = 0.4$. Each point shows the mean of a model. a) significance duration ratio, b) fundamental frequency ratio. The difference with Figure 9 is that the horizontal axis is θ_x.

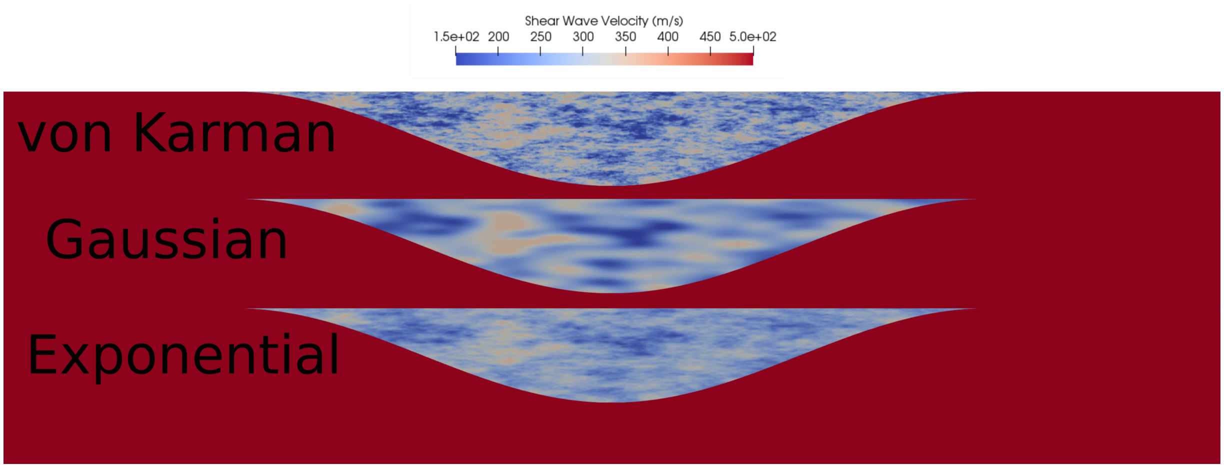

Figure 13: Example realizations with different ACF.

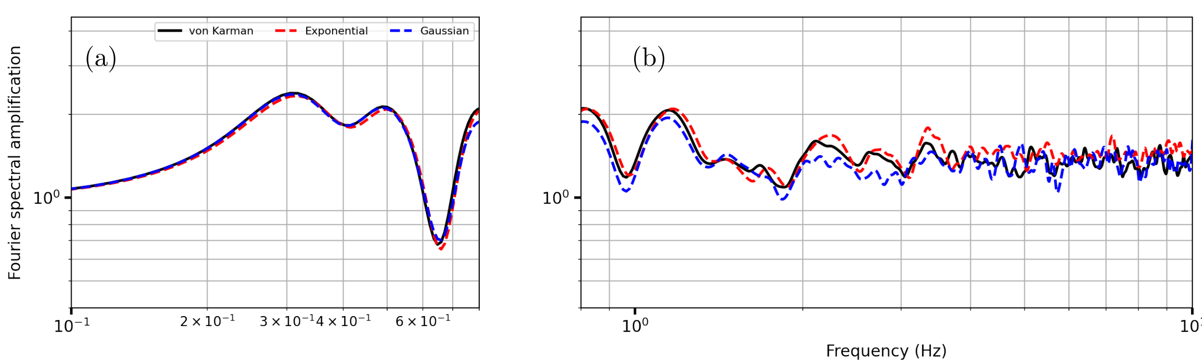

Figure 14: geometric mean of Fourier spectral amplification at point $p1$ for different frequency range with different ACFs.

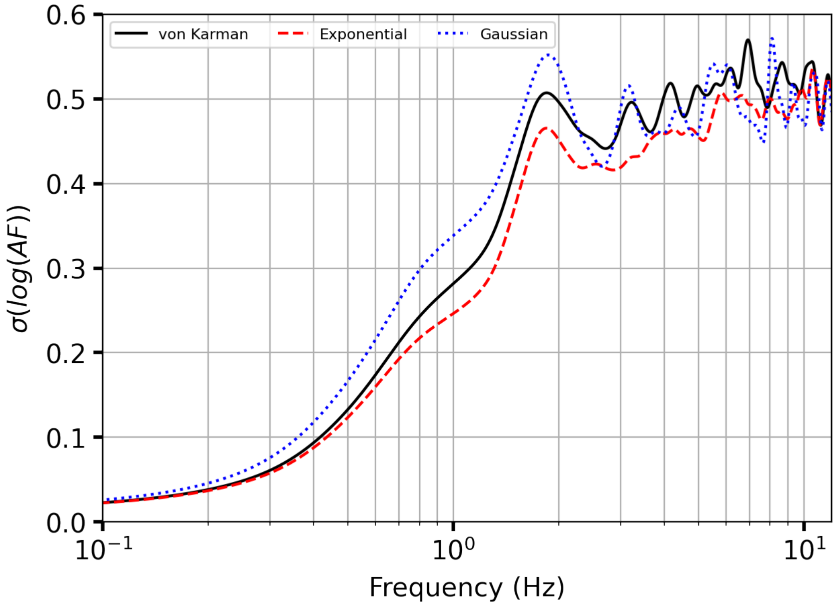

Figure 15: standard deviation of natural logarithm of Fourier spectral amplification at point $p1$ with different ACFs.

0.2, $\theta_x = 100\,[m]$, $\theta_z = 20\,[m], COV = 0.4$, and extract $1D$ shear wave velocity profile underneath point $p1$ for each realization. We simulate each column subjected to the same input as $2D$ models and compare the responses in terms of 1) mean Fourier spectral amplification and standard deviation and 2) mean response spectra amplification and standard deviation. Note that $p1$ has the closest condition to a $1D$ soil column since it is the farthest from the corners. As one moves toward a basin's edge, the corner effects would introduce a significant $2D$ wave interference to the surface ground motion.

Figure 16 shows the Fourier spectral amplification mean and standard deviation. The amplification is defined as the Fourier transform of basin response divided by surface response of $1D$ column. As can be seen in Figure 16-a, we observe a different surface response in comparison to the $1D$ analysis. In the $1D$ analysis, a layered soil medium is assumed while the $2D$ effects are neglected. Therefore, we expect significantly different resultant wavefields on basin surfaces both in low and high frequencies. The response in the low-frequency regime comes mostly from the $2D$ effects of the whole basin. On the other hand, heterogeneities would play an important role in the difference in high-frequency range. In this figure, the $COV = 0.2$ shows both amplification and de-amplification in low-frequencies ($\leq 0.6\,[Hz]$). This observation is due to the fundamental behavior of $1D$ columns and basins. Similar behavior is observed for $COV = 0.4$. Note the shift toward lower frequency for higher COV as was mentioned before. As the frequency increases, the impact of heterogeneities accentuates. For $COV = 0.2$, the resultant amplification although not zero, is not significant since the $2D$ basins resemble background medium for this COV. We observe that both models show a consistent amplification over all frequencies above $1\,[Hz]$. However, for the model with $COV = 0.4$, the difference is significant compared to the $1D$ analysis. Figure 16-b portrays the difference from the lens of standard deviation. While the model with higher COV shows a clear divergence from $1D$, the other two models still demonstrate a noticeable standard deviation. This means that $1D$ analysis is not appropriate for estimating a heterogeneous basin response even for lower COVs.

Response spectra mean amplification and the standard deviation (Figure 17) is another way to investigate the capability of a $1D$ analysis to capture the $2D$ response of a basin with heterogeneities. The amplification is defined similarly to the Fourier spectral amplification of Figure 16. Figure 17 shows how the mean

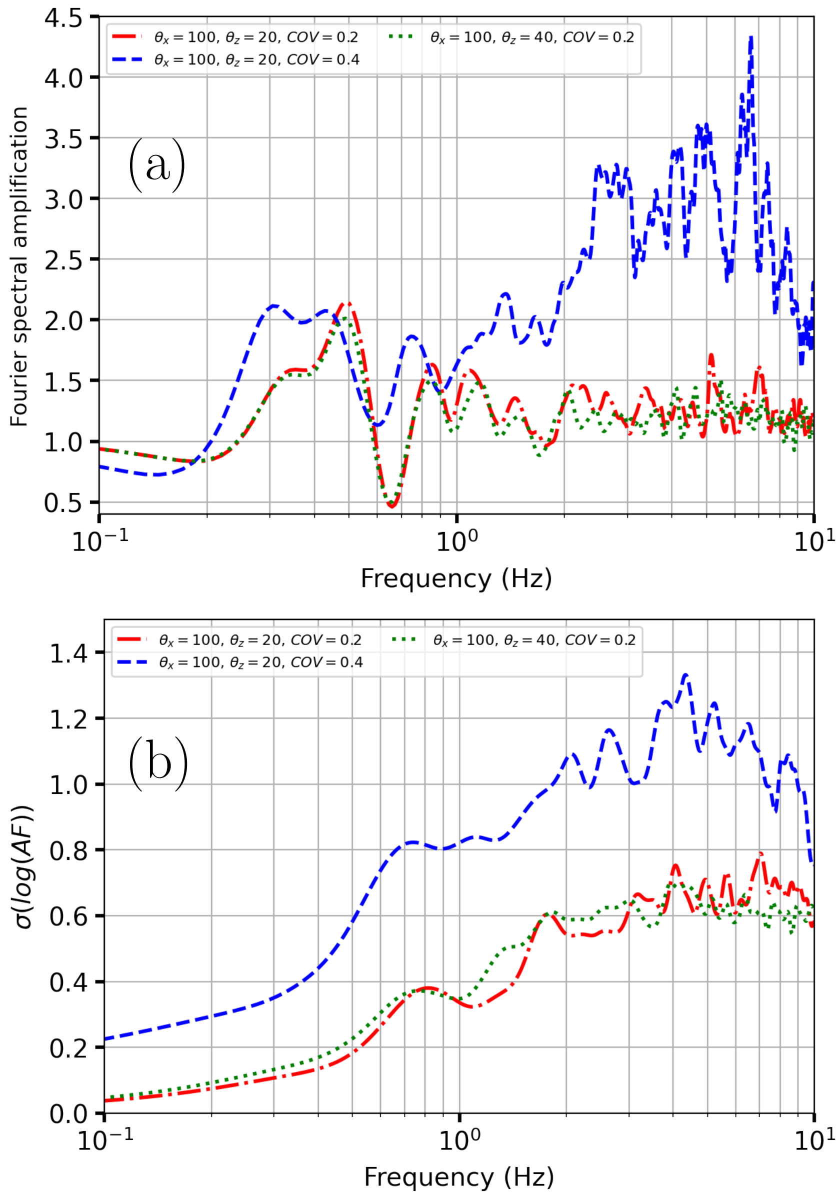

Figure 16: a) Mean and b) standard deviation of Fourier spectral amplification. Fourier spectral amplification is defined as the ratio between Fourier transform of time-series at point $p1$ of $2D$ model versus $1D$ analysis of a column underneath $p1$.

response spectra and standard deviation vary for different periods. As for response spectra amplification, the $COV = 0.4$ shows the maximum amplification in low period (equivalent to high frequency) and it decreases as we move to $COV = 0.2$. This is in agreement with Figure 16, and it can be seen that $1D$ analysis is not able to properly account for $2D$ effects inside a basin.

5 Summary and conclusions

In this article, we investigated the effect of material heterogeneity on basin surface acceleration during an earthquake. By means of Monte Carlo simulation, we generate various realizations of an elastic basin velocity field and examined effects of coefficient of variation (COV, § 3.2), correlation lengths (θ_x and θ_z, § 3.3), and autocorrelation function (ACF, § 3.4) on the spatial variation of surface ground motion. Since the focus of this study is on the material heterogeneity effects, we do not consider different basin geometries and an idealized cosine shape with $AR = 4$ and $\beta_2/\beta_1 = 2$ are assumed for numerical analysis (Figure 1, derived from [4]). The model is subjected to a vertically propagating SV plane wave of Ricker type with a dominant frequency of $5\,Hz$. We assume $\theta_x = 50$, 100, $200\,m$ and vertical $\theta_z = 20$, $40\,m$, $COV = 0.2$, 0.3, 0.4, and three ACFs, namely Von Karman, Gaussian, and exponential. For each parameter, we study its effect on surface ground motion in both time and frequency domains using 1) Fourier spectral amplification, 2) response spectra amplification, 3) significance duration ratio, and 4) fundamental frequency ratio. Following conclusions have been drawn through analyzing the results.

- COV defines the range of shear wave velocity in a medium with respect to a mean, and increasing COV results in a more considerable material contrast within a basin. This parameter affects the interference pattern of seismic waves and surface acceleration significantly. Our analysis shows that COV is the most influential parameter (among the three we examined here) on surface ground acceleration. By increasing COV, we observed elongation in significance duration and a decrease in the fundamental frequency of the basin. In addition, a significant increase in the standard deviation of surface amplification is observed.
- Correlation length (θ_x and θ_z) is another parameter of interest, which changes the size of heterogeneity patches in a medium that could affect wave-patch interaction. We observed that correlation lengths are not as important a factor as COV on surface response. Given that the input motion is a vertically propagating horizontally polarized wave, the size of θ_z affects the vertically propagating wave while contributing toward the horizontal component of surface motion. θ_x works differently by affecting horizontally propagating waves and contributing to the vertical component of motion as SV waves propagate in the basin. Different mechanisms could happen for Rayleigh and P-waves in the basin due to mode conversion and edge-induced surface waves. In sum, the impact of correlation lengths on mean response is not significant for scenarios studied here, but it changes the standard deviation of surface amplification in different frequency ranges depending on the magnitude.
- The autocorrelation function (ACF) is the last parameter we studied. While von Karman ACF is an appropriate choice for studying solid Earth, we have also examined two other common ACFs: Gaussian and exponential. We observed that although there are slight differences between the mean and standard deviation of resultant surface response for different ACFs, the discrepancies are insignificant.
- Finally, we compared the results of heterogeneous basins versus the $1D$ site analysis, which is the standard approach in practice. We extracted all $1D$ columns underneath point $p1$ from all realizations of 3 models. We performed a series of $1D$ analyses to examine how close the response of a $1D$ column would be to $2D$ basins. Using amplification mean and standard deviation, we showed that $1D$ could not fully account for the $2D$ phenomena in a basin with heterogeneities.The fact that $1D$ might not be a good approximation was previously known for homogeneous basins, but here, it was shown that the fact still holds for realistic (stochastic) basins.

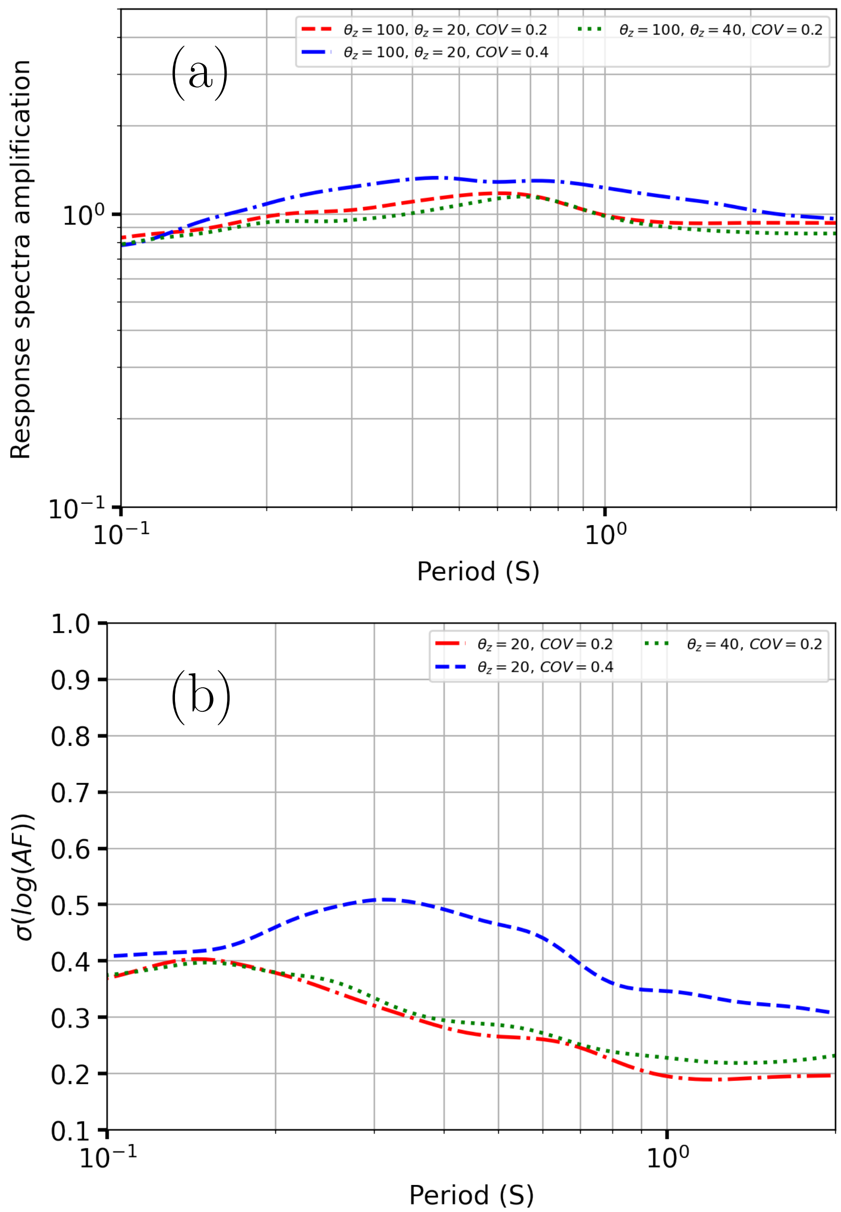

Figure 17: a) Mean and b) standard deviation of response spectra amplification. Response spectra amplification is defined as the ratio between response spectra of time-series at point $p1$ of $2D$ model versus $1D$ analysis of the column underneath $p1$.

Acknowledgments

References

[1] M Dravinski. Influence of interface depth upon strong ground motion. *Bulletin of the Seismological Society of America*, 72(2):597–614, 1982.

[2] H.L Wong and M.D Trifunac. Surface motion of a semi-elliptical alluvial valley for incident plane SH waves. *Bulletin of the Seismological Society of America*, 64(5):1389–1408, 1974.

[3] P.Y Bard and M Bouchon. The seismic response of sediment-filled valleys-Part II. The case of incident P and SV waves. *Bulletin of Seismological Society of America*, 70(5):1921–1941, 1980.

[4] Peyman Ayoubi, Kami Mohammadi, and Domniki Asimaki. A systematic analysis of basin effects on surface ground motion. *Soil Dynamics and Earthquake Engineering*, 141, 2021.

[5] P Moczo, J Kristek, P.Y Bard, S Stripajova, F Hollender, Z Chovanova, M Kristekova, and D Sicilia. Key structural parameters affecting earthquake ground motion in 2D and 3D sedimentary structures. *Bulletin of Earthquake Engineering*, 16:2421–2450, 2015.

[6] J P Narayan and S Kumar. Effects of soil layering on the characteristics of basin-edge induced surface waves. *Acta Geophysica*, 57(2):294–310, 2009.

[7] A Frankel and R.W Clayton. A finite-difference simulation of wave propagation in two-dimensional random media. *Bulletin of Seismological Society of America*, 74(6):2167–2186, 1984.

[8] E El Habar, C Cornou, D Jongmans, D.Y Abdelmassih, and F Lopez-Caballero. Influence of 2D heterogeneous elastic soil properties on surface ground motion spatial variability. *Soil Dynamics and Earthquake Engineering*, 123(1):75–90, 2019.

[9] R Popescu. Effects of Soil Spatial Variability on Liquefaction Resistance: Experimental and Theoretical Investigations. *In Proc., 4th Int. Symposium on Deformational Characteristics of Geomaterials*, 1:73–94, 2008.

[10] P.M Mai and G.C Beroza. A spatial random field model to characterize complexity in earthquake slip. *Journal of Geophysical Research: Solid Earth*, 107(B11):ESE–10, 2002.

[11] O Zielke, M Galis, and P.M. Mai. Fault roughness and strength heterogeneity control earthquake size and stress drop. *Geophysical Research Letters*, 44(2):777–783, 2017.

[12] N Nakata and G.C. Beroza. Stochastic characterization of mesoscale seismic velocity heterogeneity in Long Beach, California. *Geophysical Journal International.*, 203(1), 2015.

[13] A Frankel, E Wirth, N Marafi, J Vidale, and W Stephenson. Broadband Synthetic Seismograms for Magnitude 9 Earthquakes on the Cascadia Megathrust Based on 3D Simulations and Stochastic Synthetics, Part 1: Methodology and Overall Results. *Bulletin of the Seismological Society of America*, 108:2347–2369, 2018.

[14] D Boore. Simulation of ground motion using the stochastic method. *Pure and applied geophysics*, 160:635–676, 2003.

[15] D Assimaki, A Pecker, R Popescu, and J Prevost. Effects of spatial variability of soil properties on surface ground motion. *Journal of Earthquake Engineering*, 7(1):1–44, 2003.

[16] M Rota, C.G. Lai, and C.L. Strobbia. Stochastic 1D site response analysis at a site in central Italy. *Soil Dynamics and Earthquake Engineering*, 31:626–639, 2011.

[17] S Takemura, T Furumura, and T Maeda. Scattering of high-frequency seismic waves caused by irregular surface topography and small-scale velocity inhomogeneity. *Geophysical Journal International*, 201:459–474, 2015.

[18] H Kawase and K Aki. A study on the response of a soft basin for incident S, P, and Rayleigh waves with special reference to the long duration observed in Mexico City. *Bulletin of the Seismological Society of America*, 79(5):1361–1382, 1989.

[19] F Gelagoti, R Kourkoulis, I Anastasopoulos, and G Gazetas. Nonlinear dimensional analysis of trapezoidal valleys subjected to vertically propagating SV waves. *Bulletin of the Seismological Society of America*, 102(3):999–1017, 2012.

[20] T Saito, H Sato, M Ohtake, and K Obara. Unified explanation of envelope broadening and maximum-amplitude decay of high-frequency seismograms based on the envelope simulation using the Markov approximation: forearc side of the volcanic front in northeastern Honshu, Japan. *Journal of Geophysical Research: Solid Earth*, 110, 2005.

[21] K Sawazaki, H Sato, and T Nishimura. Envelope synthesis of shortperiod seismograms in 3-D random media for a point shear dislocation source based on forward scattering approximation: application to small strike-slip earthquakes in southwestern Japan. *Journal of Geophysical Research: Solid Earth*, 116, 2011.

[22] K Aki and B Chouet. Origin of coda waves: source, attenuation, and scattering effects. *Journal of geophysical research*, 80(23):3322–3342, 1975.

[23] D Kusanovic, E Seylabi, A Kottke, and D Asimaki. Seismo-vlab: A parallel object-oriented platform for reliable nonlinear seismic wave propagation and soil-structure interaction simulation. *To be submitted to Computer Methods in Applied Mechanics and Engineering*, 2020.

[24] R Popescu, J. H Prevost, and G Deodatis. Spatial variability of soil properties: two case studies. *Geotechnical special publication*, 75:568–579, 1998.

[25] H Harbrecht, M Peters, and M. Siebenmorgen. Efficient approximation of random fields for numerical applications. *Numerical Linear Algebra with Applications*, 22(4):596–617, 2015.

[26] S Pezzuto, A Quaglino, and M. Potse. n sampling spatially-correlated random fields for complex geometries. *International Conference on Functional Imaging and Modeling of the Heart*, pages 103–111, 2019.

[27] A. Ishimaru. Wave propagation and scattering in random media. *New York: Academic press.*, 2, 1978.

[28] W.H Savran and K.B Olsen. Model for small-scale crustal heterogeneity in Los Angeles basin based on inversion of sonic log data. *Geophysical Journal International*, 205:856–863, 2016.