



LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

Conditioned Simulation of Ground Motion Time Series at Uninstrumented Sites using Gaussian Process Regression

A. Tamhidi, N. Kuehn, S. F. Ghahari, A. J.
Rodgers, M. D. Kohler, E. Taciroglu, Y. Bozorgnia

February 22, 2021

Bulletin of the Seismological Society of America

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

1 Conditioned Simulation of Ground Motion Time Series at 2 Uninstrumented Sites using Gaussian Process Regression

3 Aidin Tamhidi^{1,*}, Nicolas Kuehn^{1,†}, S. Farid Ghahari^{1,‡}, Arthur J. Rodgers^{2,§}, Monica D. Kohler^{3,**},
4 Ertugrul Taciroglu^{1,††}, Yousef Bozorgnia^{1,‡‡}

5 ¹*Civil and Environmental Engineering Department, University of California, Los Angeles, CA 90095, USA*

6 ²*Lawrence Livermore National Laboratory, Livermore, CA 94550, USA*

7 ³*Department of Mechanical and Civil Engineering, California Institute of Technology, Pasadena, CA 91125, USA*

9 ABSTRACT

10 Ground motion time series are essential input data in seismic analysis and performance assessment of
11 the built environment. As instruments to record free-field ground motions are generally sparse, methods
12 are needed to estimate motions at locations with no available ground motion recording instrumentation.
13 In this study, given a set of observed motions, ground motion time series at target sites are constructed
14 using a Gaussian Process Regression approach, which treats the real and imaginary parts of the Fourier
15 spectrum as random Gaussian variables. Model training, verification, and applicability studies are carried
16 out using physics-based simulated ground motions of the 1906 M7.9 San Francisco earthquake and M7.0
17 Hayward fault scenario earthquake in northern California. The method's performance is further
18 evaluated using the M7.1 2019 Ridgecrest earthquake ground motions recorded by the Community
19 Seismic Network stations located in southern California. These evaluations indicate that the trained
20 Gaussian Process Regression model is able to adequately estimate the ground motion time series for
21 frequency ranges that are pertinent for most earthquake engineering applications. The trained Gaussian
22 Process Regression model exhibits decent performance in predicting the long-period content of the
23 ground motions as well as directivity pulses.

* Ph.D. Candidate. Corresponding author: Aidin Tamhidi, E-mail: aidintamhidi@ucla.edu

† Research Scientist

‡ Research Scientist

§ Staff Scientist

** Research Professor

†† Professor

‡‡ Professor

24 KEYWORDS: Conditioned Simulation of Ground Motions; Gaussian Process Regression; Spatial
25 Variation of Ground Motions; Regional Seismic Assessment

26 (The authors acknowledge there are no conflicts of interest recorded.)

27 INTRODUCTION

28 While the number of available recorded earthquake ground motions has increased in the last few decades,
29 current sensor networks are still sparse for various site-specific earthquake applications. Thus, an
30 estimation of either ground motion intensity measures (GMIM), e.g. PGA, PGV, spectral response
31 ordinates, or the entire ground motion time series is required to evaluate the damage state or performance
32 level of a specific structure for post-event assessment. As there are only approximately 2000 ground-
33 level stations to record the free-field ground motions in California (Southern California Seismic
34 Network, Northern California Seismic Network, CSMIP), site-specific structural assessments invariably
35 require estimations using interpolation methods.

36 Presently, “ShakeCast,” “ShakeMap,” and the USGS “Did You Feel It?” platforms offer
37 estimates of the shaking level and GMIM after an event using various techniques (Wald et al. 2008;
38 Fraser et al. 2008; Worden et al. 2018; Wald et al. 2012; Lin et al. 2018). Some of these techniques
39 involve estimating the GMIMs at target sites using the surrounding observations (Baker and Chen 2020;
40 Otake et al. 2020; Worden et al. 2018). However, for nonlinear response-history analyses of structural
41 systems, and analysis of the degree and distribution of damage in a structure, the entire ground motion
42 time series is needed. Therefore, the generation of realistic time series is needed at sites where recorded
43 motions are not available (Petrone et al. 2020). The generated motions should be able to capture
44 reasonable variations in the amplitude, phase, and frequency content over an area (Zerva and Zervas
45 2002; Zerva 2009; Chen and Baker 2019) as such spatial variations can have considerable effects,
46 especially on distributed lifeline structures (Adanur et al. 2016; Jayaram and Baker 2009; Zerva et al.
47 2018; Todorovska et al. 2017; Tian et al. 2016).

48 There has been extensive work on “conditioned ground motion simulations” wherein the time
49 series at target sites are constructed using surrounding measurements (Lu et al. 2021; Rodda and Basu
50 2019, 2018; Huang and Wang 2017; Wu et al. 2016; Alimoradi et al. 2015; Zentner 2013; Konakli and
51 Der Kiureghian 2012; Kameda and Morikawa 1992). The majority of conditioned ground motion
52 simulations are based on the use of cross-spectral density (CSD) and auto-spectral density (ASD)
53 functions to determine the covariance between the Fourier series coefficients for neighboring stations
54 (Der Kiureghian 1996; Konakli and Der Kiureghian 2012; Rodda and Basu 2018). The conditioned
55 ground motion simulation results depend on the spatial variability of the motions captured by CSD and
56 ASD. The CSD is determined using coherency functions, and the coefficients of these functions are
57 assigned empirically using data-driven methods (Abrahamson et al. 1991). Moreover, a detailed
58 description of the site properties and wave propagation characteristics is sometimes needed for
59 generating the simulated motions, which can be computationally expensive and thus time-consuming,
60 especially when an ensemble of ultra-dense sites is needed.

61 In this study, the Gaussian Process Regression (GPR) method, also known as Kriging
62 (Rasmussen and Williams 2006), is employed to generate the ground motion time series at target sites
63 where there are no available recording instruments. This method is able to construct the entire ground
64 motion time series decently at the target site using limited input information such as geographical
65 coordinates and the average shear wave velocity in the uppermost 30 m, $V_{s_{30}}$, from each site. Therefore,
66 it is able to estimate the motion time series with lower computational costs in comparison with the
67 aforementioned methods. The GPR method spatially interpolates the real and imaginary parts of the
68 observed frequency content of the neighboring motions using an assumed covariance function to
69 establish the ground motion time series at the target site. The spatial correlation of the ground motions
70 is computed in order to estimate the entire time series at a target site using the observed data set.

71

THEORETICAL BACKGROUND

Suppose the ground motion acceleration time series, $a_s(t)$, at location s , is constructed from N discrete data points, $a_s(t_i)$, $i = 1 \dots N$, at equal time intervals, Δt . The accelerations, $a_s(t_i)$, can then be expressed using their Discrete Fourier Transform (DFT) coefficients A_k (e.g., Oppenheim et al. 1997) as

$$a_s(t_i) = \sum_{k=0}^{N-1} A_k e^{j\omega_k t_i} \quad (1)$$

where

$$A_k = \frac{1}{N} \sum_{i=0}^{N-1} a_s(t_i) [\cos(\omega_k t_i) + j \cdot \sin(\omega_k t_i)] = \mathcal{R}e_k + j \cdot \mathcal{I}m_k. \quad (2)$$

In Eqs. 1 and 2, ω_k is the k^{th} natural frequency (at equal frequency intervals) of the DFT, $j = \sqrt{-1}$ and $\mathcal{R}e_k$ and $\mathcal{I}m_k$ are the real and imaginary parts of the DFT coefficient, A_k , respectively, at the k^{th} frequency.

Here, we assume that $\mathcal{R}e_k$ and $\mathcal{I}m_k$ (at k^{th} frequency, $k = 0, \dots, N-1$) are random Gaussian variables for any location, s , within a region. We also consider that $\mathcal{R}e_k$ at location s , is spatially correlated to $\mathcal{R}e'_k$ at location s' where s and s' are neighbors. A similar assumption for $\mathcal{I}m_k$ is taken. In this study, we aim to implement GPR as a method to estimate the values of $\mathcal{R}e_k$ (and $\mathcal{I}m_k$) at the k^{th} frequency ($k = 0, \dots, N-1$) using the corresponding $\mathcal{R}e'_k$ (and $\mathcal{I}m'_k$) from the surrounding station observations. We then reconstruct the entire acceleration time series at the target location with all estimated $\mathcal{R}e_k$ and $\mathcal{I}m_k$ using Eq. 1.

It is assumed that there is a statistically insignificant correlation between $\mathcal{R}e_k$ (or similarly $\mathcal{I}m_k$) and $\mathcal{R}e_j$ (or similarly $\mathcal{I}m_j$) at the same location, s , for different frequencies k and j , where $k \neq j$, in order to construct the mean estimated ground motion time series. It is worth noting that the mean estimated values

90 for multivariate Gaussian variables (here $\mathcal{R}e$ and $\mathcal{I}m$) are independent of the inter-frequency correlation
 91 between amplitudes at various frequencies; yet, the inter-frequency correlations of the DFT coefficients
 92 need to be accounted for when generating random ground motion realizations (see subsection
 93 “Realizations of Ground Motion”).

94 ***Gaussian Process Regression***

95 GPR is a supervised learning method that has numerous applications in earthquake engineering and
 96 seismology, such as ground-motion time series estimation, post-earthquake damage assessment,
 97 development of performance models of engineering materials, and seismic fragility assessment (Tamhidi
 98 et al. 2019, 2020; Sajedi and Liang 2020; Sheibani and Ou 2020; Landwehr et al. 2016; Sun et al. 2018;
 99 Gentile and Galasso 2020; Ghaderi et al. 2020). A Gaussian Process (GP) (Rasmussen and Williams
 100 2006) is a collection of indexed random variables such that every finite subset is distributed according
 101 to a multivariate normal distribution. In general terms, GP can be understood as a multivariate normal
 102 distribution for infinitely many random variables. More specifically, GP is a distribution over function
 103 $f(\mathbf{x}) \in \mathbb{R}$

$$f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')), \quad (3)$$

104 which reads as “the function value, $f(\mathbf{x})$, at input location \mathbf{x} is drawn from a GP with the mean function,
 105 $m(\mathbf{x})$, and the covariance function $k(\mathbf{x}, \mathbf{x}')$.” As Eq. 3 indicates, a GP is entirely defined by its mean,
 106 $m(\mathbf{x})$, and covariance, $k(\mathbf{x}, \mathbf{x}')$ functions, which are

$$m(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})] \quad (4)$$

$$k(\mathbf{x}, \mathbf{x}') = \mathbb{E}[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))]. \quad (5)$$

107 In Eqs. 4 and 5, \mathbb{E} stands for mathematical expectation. The covariance function $k(\mathbf{x}, \mathbf{x}')$ indicates
 108 the degree of similarity between function values at data points, \mathbf{x} and \mathbf{x}' . In Bayesian non-parametric

109 statistics (Hjort et al. 2010), a GP is often used to specify a prior distribution over possible functions.
 110 Here, we assume that the real and imaginary parts of the DFT coefficients (cf. Eq. 2) are functions of the
 111 location and possibly other geotechnical or seismological parameters such as the local site condition.
 112 Since the functional form is unknown and complicated, we replace it with a GP. In order to estimate the
 113 DFT coefficients, we carry out a GPR over the observed values, \mathbf{f} , which are either the real or imaginary
 114 parts of the DFT coefficients at each frequency.

115 It is worth noting that other regression methods such as Nadaraya-Watson kernel regressions
 116 (Watson 1964; Nadaraya 1964) or Savitzky-Golay filter (Savitzky and Golay 1964) are possible
 117 alternatives for interpolation purposes. Both GPR (see Realizations of Ground Motion subsection) and
 118 Kernel regression methods (Rubin 1981) can estimate the uncertainty of the predicted GP values. The
 119 Kernel regression methods can implement adaptive kernels that change with data (Huang et al. 2014),
 120 and the GPR is able to implement a combination of multiple kernel functions (through multiplication
 121 and summation) to estimate the covariance among observations from a complex function. In this study,
 122 we use GPR as it simultaneously optimizes the covariance function parameters based on observations
 123 and estimates GP values at target locations without imposing a high computational cost.

124 One can compute the predictive distribution for function values \mathbf{f}^* at new (target) locations
 125 (without recorded ground motions) by conditioning on the observed data. The joint distribution of
 126 observed data and new simulated data is

$$\begin{bmatrix} \mathbf{f} \\ \mathbf{f}^* \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \boldsymbol{\mu} \\ \boldsymbol{\mu}^* \end{bmatrix}, \begin{bmatrix} \mathbf{K}_{xx} + \sigma_y^2 \mathbf{I} & \mathbf{K}_{xx^*} \\ \mathbf{K}_{x^*x} & \mathbf{K}_{x^*x^*} \end{bmatrix} \right), \quad (6)$$

127 where \mathbf{K}_{xx} denotes the covariance matrix of the DFT coefficients at the observed locations. The entries
 128 of \mathbf{K}_{xx} are calculated from the covariance function via $K_{xx_{ij}} = k(\mathbf{x}_i, \mathbf{x}_j)$, where k denotes the covariance
 129 function between two locations (cf. Eq. 5). Correspondingly, \mathbf{K}_{xx^*} describes the covariance between the
 130 observed DFT coefficients and the estimated ones at the target locations, and $\mathbf{K}_{x^*x^*}$ are the covariances

131 of the DFT coefficients at the target locations. The term σ_y denotes the observation noise; \mathbf{I} is the identity
 132 matrix; and $\boldsymbol{\mu}$ and $\boldsymbol{\mu}_*$ are the prior mean vectors at the observed and target locations, respectively. Here,
 133 the observed ground motions, and subsequently their DFT coefficients, are considered noise-free ($\sigma_y =$
 134 0). The predictive distribution for the function values \mathbf{f}_* at the target locations is then (Rasmussen and
 135 Williams 2006)

$$\mathbf{f}_* | \mathbf{X}_*, \mathbf{X}, \mathbf{f} \sim \mathcal{N}(\boldsymbol{\mu}_*, \boldsymbol{\Sigma}_{**}) \quad (7)$$

where

$$\boldsymbol{\mu}_* = \boldsymbol{\mu} + \mathbf{K}_{x_*x} \mathbf{K}_{xx}^{-1} (\mathbf{f} - \boldsymbol{\mu}) \quad (8)$$

$$\boldsymbol{\Sigma}_{**} = \mathbf{K}_{x_*x_*} - \mathbf{K}_{x_*x} \mathbf{K}_{xx}^{-1} \mathbf{K}_{xx_*} \quad (9)$$

136 and \mathbf{X} denotes the input matrix of the observations, each row of which is one observed location's input
 137 vector including its geographical coordinates and possibly other features. Similarly, \mathbf{X}_* is the input
 138 matrix of all new target locations.

139 The GPR's output and smoothness depend on the computed covariance function which is defined
 140 based on a kernel, $k(r)$, where r is the distance between the input vectors \mathbf{x} and \mathbf{x}' given by Eq. 10.

$$r = \theta \sqrt{\sum_{i=1}^d (x_i - x'_i)^2} \quad (10)$$

141 In Eq. 10, θ is a positive normalizing factor (also known as the inverse of length-scale, l , where $\theta = 1/l$)
 142 and d is the size of the input vector (number of attributes). There are several established covariance
 143 functions such as exponential and Matérn which are given by

$$k_{exp}(r) = \sigma_f^2 \exp(-r), \text{ and} \quad (11)$$

$$k_{Matérn}(r) = \sigma_f^2 \frac{2^{1-\nu}}{\Gamma(\nu)} (\sqrt{2\nu}r)^\nu K_\nu(\sqrt{2\nu}r), \quad (12)$$

144 respectively. In Eq. 12, Γ is the Gamma function $\Gamma(n) = (n - 1)!$; K_ν is modified Bessel function
145 (Abramowitz and Irene 1972); ν is a positive parameter that controls the smoothness of the output
146 function. In Eqs. 11 and 12, σ_f is the variance that governs how uncertain the GPR's estimate is for a
147 given input location.

148 In this study, a single θ value is used to normalize all attributes within an input vector (cf. Eq.
149 10). Such a covariance function is called isotropic. As an alternative, an anisotropic covariance function
150 where each attribute has its own specific length-scale also can be used (Rasmussen and Williams 2006).
151 The θ value specifies the rate of decay for the covariance function. Higher values for θ (smaller length-
152 scale) results in a faster decay of covariance, and subsequently correlation, by increasing the distance.
153 More detailed descriptions of GPR can be found in, e.g., Li and Sudjianto (2005) and Chapters 2 and 4
154 of Rasmussen and Williams (2006).

155 PROPOSED MODELS

156 A proper input vector for the observed and target sites needs to be defined in order to start fitting
157 the GPR. It is possible to consider the homogeneity assumption for regions with fairly uniform site
158 conditions. In this case, all of the GP's stochastic descriptors depend only on the geographical separation
159 distance between the stations (Zerva and Zervas 2002). In this study, we consider two types of input
160 vectors (corresponding to two GPR models) for the stations. These are namely, **Type 1**, the 3D Cartesian
161 components of each station (after converting the geographical coordinates longitude and latitude into 3D
162 Cartesian coordinates), $\mathbf{x} = \{x_1, x_2, x_3\}$, where the homogeneity assumption is valid. In fact, $\{x_1, x_2, x_3\}$
163 are the Cartesian coordinates of the station on the Earth's surface; and **Type 2**, wherein the 3D Cartesian
164 components are stacked up with $\log(V_{s_{30}})$ as the 4th component, $\mathbf{x} = \{x_1, x_2, x_3, \log(V_{s_{30}})\}$, where the
165 homogeneity assumption is invalid.

166 More precisely, the GPR model type 1 is a specific case of the more inclusive GPR model type
 167 2. GPR model type 1 is used here to investigate the applicability of a simpler attribute vector (using only
 168 the 3D Cartesian coordinates), imposing lower computational cost for regions with fairly uniform soil
 169 conditions (i.e., the variation of V_{s30} is negligible).

170 The GPR model type 2 input vector can also be extended to include more attributes of the
 171 locations such as $Z_{1.0}$ (depth to $V_s = 1$ km/s), $Z_{2.5}$ (depth to $V_s = 2.5$ km/s), and R_{JB} (closest distance to
 172 the surface projection of coseismic rupture). In the GPR model type 2, input vector attributes are
 173 normalized, such that the mean and standard deviation of each distribution are zero and one, respectively.
 174 This normalization is required in order to convert all the attributes into a similar range of values.

175 *Model Parameters and Optimization*

176 The parameters of the GPR model are the distance normalizing factor θ , and the GP mean μ and variance
 177 σ_f , which need to be pre-defined to implement the GPR. Denoting the model parameters as $\gamma = (\theta, \mu, \sigma_f)$,
 178 a commonly used method to find the optimum γ is to maximize the log-marginal likelihood of the n
 179 observations given γ , using

$$\log p(\mathbf{f}|\mathbf{X}, \boldsymbol{\gamma}) = -\frac{1}{2}(\mathbf{f} - \boldsymbol{\mu})^T \mathbf{K}_{xx}^{-1}(\mathbf{f} - \boldsymbol{\mu}) - \frac{1}{2} \log |\mathbf{K}_{xx}| - \frac{n}{2} \log 2\pi. \quad (13)$$

180 In Eq. 13, the superscript T indicates transpose operator; and $|\mathbf{K}_{xx}|$ denotes the determinant of the matrix
 181 \mathbf{K}_{xx} . In Eq. 13, $\boldsymbol{\mu}$ and \mathbf{K}_{xx} are functions of θ (Li and Sudjianto 2005). Parameter estimates found by
 182 maximizing Eq. 13 are the Maximum Likelihood Estimates (MLEs). The MLEs have considerable
 183 variance near their optimum solution, as the likelihood function is almost flat close to its extremum,
 184 especially when observations are sparse (Li and Sudjianto 2005). In order to tackle this issue, one can
 185 maximize the penalized log-likelihood (log posterior) rather than log marginal likelihood. Eq. 14 shows
 186 the penalized log-likelihood, $Q(\gamma)$, formulation:

$$Q(\boldsymbol{y}) = -\frac{1}{2}(\boldsymbol{f} - \boldsymbol{\mu})^T \mathbf{K}_{xx}^{-1}(\boldsymbol{f} - \boldsymbol{\mu}) - \frac{1}{2} \log |\mathbf{K}_{xx}| - \frac{n}{2} \log 2\pi - n d p_\lambda(\theta). \quad (14)$$

187 In Eq. 14, $p_\lambda(\theta)$ is a non-negative penalty function for normalizing factor θ . The λ is a non-negative
 188 regularization factor that needs to be tuned using data-driven methods, as elaborated in the next
 189 subsection. There are several choices for the penalty function in Eq. 14, such as the Least Absolute
 190 Shrinkage and Selection Operator (LASSO) (e.g., Tibshirani 1996), and Smoothly Clipped Absolute
 191 Deviation (SCAD) (Fan and Li 2001). In this study, the SCAD penalty function is used and is given by

$$p_\lambda(\theta) = \begin{cases} \lambda\theta & \theta \leq \lambda \\ -\frac{\lambda^2 + \theta^2 - 2a\lambda\theta}{2(a-1)} & \lambda < \theta \leq a\lambda \\ \frac{\lambda^2(a+1)}{2} & a\lambda < \theta \end{cases} \quad (15)$$

192 where a is a constant, which is assumed to be 3.7 based on Fan and Li (2001) who illustrated that a
 193 model's performance is not considerably improved choosing a through a data-driven method. The
 194 penalized log-likelihood, $Q(\boldsymbol{y})$, in Eq. 14, is the log posterior distribution of \boldsymbol{y} , given the observations.
 195 In other words, the Maximum a Posteriori Estimates (MAPEs) of parameters, $\hat{\boldsymbol{y}} = (\hat{\theta}, \hat{\mu}, \hat{\sigma}_f)$, are
 196 employed as an alternative to the commonly used MLEs by maximizing Eq. 14. The GP mean, μ , in Eq.
 197 8, and the variance, σ_f , are updated into $\hat{\mu}$ and $\hat{\sigma}_f$ given $\hat{\theta}$ (e.g., Li and Sudjianto 2005).

198 The GPR is completely defined by its optimized parameters $\hat{\theta}$, $\hat{\mu}$, and $\hat{\sigma}_f$. The regularization factor,
 199 λ , needs to be defined before optimizing these parameters through maximizing $Q(\boldsymbol{y})$. More specifically,
 200 λ governs the derivation of optimized parameters $\hat{\theta}$, $\hat{\mu}$, and $\hat{\sigma}_f$. As a hierarchical view, one can recognize
 201 θ , μ , and σ_f as the parameters of the GPR model, while λ is its hyper-parameter. The process of
 202 optimization of this hyper-parameter is elaborated next.

203

204

205 *Hyper-Parameter Optimization*

206 It is common to use data-driven methods such as Cross-Validation (CV) to find the optimum
207 hyper-parameter values, here the regularization factor, $\hat{\lambda}$. In our case, the “data” to be used in the “data-
208 driven” methodology is a set of “observed” ground motions which is a subset of physics-based simulated
209 ground motions for the 1906 M7.9 San Francisco earthquake. Here, we used broadband ground motions
210 generated using Graves’s hybrid simulation wave propagation code (Aagaard et al. 2008). These ground
211 motion time series are generated at 40,700 locations on a 1.5 km \times 1.5 km uniform grid along three
212 orthogonal directions. Table 1 displays various features of the physics-based simulated ground motions
213 for the 1906 M7.9 San Francisco earthquake. A minimum V_{s30} value of 760 m/s was used for these
214 simulations. Correction factors were applied for site effects at locations with V_{s30} lower than 760 m/s.

215 Two different optimum $\hat{\lambda}$ values must be obtained for the two GPR models introduced in the
216 “Proposed Models” section. For GPR model type 1, 396 locations with the same V_{s30} (560 m/s) within a
217 20 km \times 75 km rectangular region are chosen with the homogeneity assumption. We refer to this region
218 hereafter as the ‘East Bay’ region (Figure 1). For GPR model type 2, two regions where the homogeneity
219 assumption is invalid are chosen. These two regions are hereafter referred to as the ‘Palo Alto’ and
220 ‘South Napa’ regions (Figure 1), with 104 and 111 chosen sites, respectively. The sites within each of
221 the East Bay, Palo Alto, and South Napa regions are randomly split into a training set (80% of the total
222 number of sites) which makes up the “observed” ground motions, while the remaining 20% are
223 considered the test set (target sites) (Figure 2).

224 A five-fold CV procedure is implemented over the training set (observed ground motions) within
225 each region in order to select the best regularization factor, $\hat{\lambda}$, for the corresponding GPR model. The
226 accuracy criterion for this selection is the normalized root mean square error (NRMSE) between the
227 exact (physics-based simulated) and the estimated (conditioned simulated) ground motions’ 5%-damped
228 pseudo-spectral acceleration (PSA) at the target site. The NRMSE is computed as

$$NRMSE = \sqrt{\frac{1}{\tau} \sum_{i=1}^{\tau} \frac{(PSA_i - \widehat{PSA}_i)^2}{\widehat{PSA}_i^2}} \quad (16)$$

229 where, τ equals 85 which is the number of periods included in the PSA ranging from 0.1 s to 20 s, and
 230 PSA_i and \widehat{PSA}_i are the predicted and exact ground motions' PSA values at the i^{th} period, respectively. A
 231 lower NRMSE value indicates a greater similarity between the estimated and exact response spectra.
 232 One can use Fourier Amplitude Spectrum (FAS) NRMSE; however, the response spectrum is smoother
 233 than the FAS, which makes the NRMSE criterion better suited to measure the degree of similarity
 234 between the estimated and exact ground motions. The NRMSE value computed across all frequencies
 235 for the FAS is highly sensitive to the rapid variations of amplitude from one frequency to another,
 236 whereas such changes are far smaller for the PSA. Moreover, the PSA spectrum is representative of the
 237 GMIM, which is commonly used for engineering applications.

238 The following steps are taken to select $\hat{\lambda}$. First, we randomly split the training (observed) data set
 239 into five separate folds. For each λ_{test} to be evaluated, we carry out the following procedure:

240 1. For each fold $i = 1, \dots, 5$:

241 1.1 Find the optimum parameters, $\hat{\theta}$, $\hat{\mu}$, and $\hat{\sigma}_f$ for the observed motions within all folds
 242 except the i^{th} fold using λ_{test} and maximizing $Q(\boldsymbol{\gamma})$ in Eq. 14. These parameters need
 243 to be found for each frequency and for both real and imaginary parts of the DFT
 244 coefficients.

245 1.2 Estimate the ground motion time series at each site within the i^{th} fold using the posterior
 246 mean (Eq. 8) for the DFT coefficients, using $\hat{\theta}$, $\hat{\mu}$, and $\hat{\sigma}_f$ determined in step 1.1.

247 1.3 Compute the NRMSE between the estimated (step 1.2) and exact ground motion
 248 response spectra (Eq. 16) at each site within the i^{th} fold, and store their averages as
 249 $Error_i$.

250 2. Take the average of $Error_i$ ($i = 1, \dots, 5$) i.e., $Error_{avg}$ and record it as being associated with
 251 λ_{test} .

252 Eventually, we choose the λ_{test} with the lowest $Error_{avg}$ computed in step 2 as the optimized
 253 regularization factor, $\hat{\lambda}$.

254 **Covariance Function Selection**

255 We investigate the performance of the model using three different covariance functions:
 256 exponential (cf. Eq. 11), and Matérn with $\nu = 1.5$ and $\nu = 2.5$ (shown in Eqs. 17 and 18), in order to
 257 find the optimized covariance function for the GPR model. The values $\nu = 1.5$ and $\nu = 2.5$ are widely
 258 used for Matérn covariance functions in GPR applications (Rasmussen and Williams 2006). Exponential
 259 covariance functions have been used in the estimation of spatially distributed GMIMs (Jayaram and
 260 Baker 2009). Matérn covariance functions have also been used to model the spatial correlation of ground
 261 motions for developing non-ergodic ground motion models (Kuehn and Abrahamson 2020.)

$$k_{\nu=1.5}(r) = \sigma_f^2(1 + \sqrt{3}r)\exp(-\sqrt{3}r) \quad (17)$$

$$k_{\nu=2.5}(r) = \sigma_f^2\left(1 + \sqrt{5}r + \frac{5}{3}r^2\right)\exp(-\sqrt{5}r) \quad (18)$$

262 We use the South Napa training set (Figure 2.c) to conduct the five-fold CV procedure for GPR
 263 models type 2 constructed with each of the three aforementioned covariance functions. First, the
 264 optimized regularization factor, $\hat{\lambda}$, is obtained for each of the GPR models; then, the average NRMSE
 265 ($Error_{avg}$ described in “Hyper-Parameter Optimization” subsection) for the corresponding obtained $\hat{\lambda}$
 266 is determined for each GPR model. Table 2 illustrates derived $\hat{\lambda}$ values for each covariance function as

267 well as the average NRMSE obtained for the corresponding model over the South Napa training set in
268 both the Fault-Normal (FN) and Fault-Parallel (FP) directions. As is illustrated in Table 2, the Matérn
269 with $\nu = 1.5$ covariance function outperformed the other two covariance functions. Thus, in this study,
270 we use the Matérn covariance function with $\nu = 1.5$ (cf. Eq. 17) to establish the GPR models type 1 and
271 2.

272 Table 3 displays the $\hat{\lambda}$ values for the FN and FP directions for each model using the Matérn ($\nu =$
273 1.5) covariance function. The CV procedure yielded the same $\hat{\lambda}$ values within the South Napa and Palo
274 Alto study regions for the GPR model type 2. It is worth noting that the optimized regularization factor,
275 $\hat{\lambda}$, is dependent on the density of the observations (recall that this is the number of observed sites divided
276 by the area of the network). We observe that a smaller number of available observations (lower density
277 of observed sites) leads to higher required regularization factor values, consistent with Li and Sudjianto
278 (2005). The observation densities for the East Bay and Palo Alto (or similarly South Napa) regions are
279 0.26 sites/km² and 0.29 sites/km², respectively.

280 MODEL EVALUATION

281 The performance of the proposed GPR models is evaluated next. To do so, we compare the estimated
282 ground motions generated by the GPR model with the exact ground motions at the same sites. The
283 following procedure is used to estimate the ground motion time series at any target site:

- 284 1. Given the observed ground motions (training set), the model parameters, $\hat{\theta}$, $\hat{\mu}$, and $\hat{\sigma}_f$ are obtained
285 at each frequency for the real and imaginary parts of the DFT coefficients using the
286 corresponding $\hat{\lambda}$ given in Table 3.
- 287 2. The posterior means (Eq. 8) at the desired sites for the DFT coefficients are calculated for each
288 frequency using the values of $\hat{\theta}$, $\hat{\mu}$, and $\hat{\sigma}_f$ from step 1.
- 289 3. The entire ground motion time series is constructed using Eq. 1.

290 The tuned $\hat{\lambda}$ based on the 1906 M7.9 San Francisco earthquake subset (as elaborated in the
291 previous section) is used to validate the GPR models' estimation of ground motions in simulated data
292 sets of the 1906 M7.9 San Francisco (Aagaard et al. 2008) and the M7.0 Hayward fault scenario
293 earthquake (Rodgers et al. 2019), as well as the M7.1 2019 Ridgecrest Earthquake recorded by the
294 Community Seismic Network (Clayton et al. 2020).

295 *The 1906 M7.9 San Francisco simulated motions*

296 The training set for each study region (East Bay, Palo Alto, and South Napa) for the 1906 M7.9 San
297 Francisco earthquake simulated motions are shown in Figure 2. The corresponding GPR model is
298 implemented for each region to estimate the ground motion time series at each test site within the test
299 set (colorful circular points in Figure 3). Figure 3 illustrates the distribution of NRMSE between the
300 estimated and exact motions' linear response spectra 5%-damped RotD50 (Boore 2010) values. In Figure
301 3, there are three chosen test sites for each region. The prediction results for the RotD50 spectrum,
302 velocity time series, and FAS are shown for selected sites in the East Bay, Palo Alto, and South Napa
303 regions in Figures 4, 5, and 6, respectively. Table 4 summarizes the 1906 San Francisco test set's
304 NRMSE for FN, FP, and RotD50 linear response spectra.

305 Regarding the RotD50 spectrum NRMSE, Figure 3 demonstrates that the GPR model is able to
306 estimate the ground motion time series at most of the target sites reasonably well. The estimation can be
307 less accurate for sites at the boundaries of the network (as shown in Figure 3), where there is a less
308 uniform distribution of observations. By comparing the results of Figure 3b and Figure 3c, it is apparent
309 that the estimation accuracy for the sites far away from the causative fault might be higher than for those
310 close to the fault (also shown in Table 4). This could be due to the use of an isotropic covariance function,
311 which allocates a uniform correlation to the surrounding locations based on the separation distance. The
312 GPR model predictions can be improved by employing an anisotropic covariance structure, which uses
313 different normalizing factors for each attribute to compute the separation distance (Rasmussen and

314 Williams 2006) for regions closer to the fault. Figure 3b indicates that the test sites with less accurate
315 ground motion estimation (higher RotD50 NRMSE) within the Palo Alto region are mainly restricted to
316 the edge stations; yet the trained GPR model is able to predict ground motions for sites close to the fault
317 appropriately. As evidenced in Figure 4, the GPR model type 1 is capable of estimating the entire ground
318 motion time series decently for the structural period ranges pertinent to most earthquake engineering
319 applications.

320 Figures 5a and 5b show that the GPR model is able to estimate the long-period pulses along the
321 FN direction due to the directivity effect for the sites far away from the epicenter but close to the fault
322 (sites 1 and 2 in Figure 3b).

323 Figures 4-6 show that the RotD50 response spectrum and FAS errors are lower for longer periods,
324 whereas the difference between the estimated and the exact spectra increases for the shorter periods. This
325 might be due to two reasons: first, the short-period motions of the 1906 San Francisco earthquake are
326 constructed stochastically (Aagaard et al. 2008), which results in lower correlations for the short-period
327 content of the neighboring motions. Thus, the GPR estimation for the short-period motion could be less
328 accurate than it is for longer-period motion. Second, the motions are less well-correlated to each other
329 at higher frequencies and longer geographical separation distances because of the smaller wavelengths
330 associated with those frequencies. Therefore, short-period waves of the ground motions may be less
331 accurately synthesized, especially when the neighboring stations are not sufficiently close to each other.
332 This phenomenon is observed in existing ‘lagged coherency’ models where the lagged coherency
333 between two stations, as a representative of the correlation between the frequency content, drops with
334 increasing frequency and separation distance (e.g., Liao and Zerva 2006; Rodda and Basu 2018;
335 Abrahamson et al. 1991).

336 Figure 7 displays the DFT coefficients’ real part $\hat{\theta}$ values for the GPR model type 2 implemented
337 within the Palo Alto and South Napa study regions along the FN and FP directions. As indicated, we

338 incorporated the effects of variation in soil conditions in these regions. Figure 7 demonstrates the $\hat{\theta}$
339 growth as a function of increasing frequency. A similar observation exists for the imaginary part $\hat{\theta}$
340 values. It is recognizable from Eq. 10 and 17 that covariance (and subsequently correlation) among the
341 observed values decreases with increasing $\hat{\theta}$ (equivalently decrease of length-scale). In other words,
342 there is a lower correlation between the higher-frequency content of the ground motions, which is
343 consistent with the established lagged coherency models.

344 ***M7.0 Hayward Fault scenario earthquake simulated motions***

345 We evaluate the performance of the trained GPR on another simulated earthquake data set, which was
346 not used during the hyper-parameter optimization procedure. To do this, the M7.0 Hayward fault
347 scenario earthquake simulated ground motions (Rodgers et al. 2019) are employed. In the present study,
348 motions for the 3D model ('3D TOPO') are used to evaluate the accuracy of the trained GPR's estimation.
349 The 3D TOPO Earth model has a $V_{s_{min}} = 500$ m/s; therefore, the simulation results for the sites with $V_s >$
350 500 m/s are more reliable. The 3D subsurface material properties and the topography (3D TOPO)
351 simulations are obtained based on the United States Geological Survey (USGS) model (USGS 2018).
352 There is a total of 2301 locations within a 120 km \times 80 km rectangular domain on a uniform 2 km \times 2
353 km grid for which velocity time series are generated along the FN, FP, and vertical directions.

354 In this study, 326 locations with fairly uniform site conditions (520 m/s $> V_{s_{30}} > 500$ m/s) are
355 chosen within the East Bay study region (Figure 8). As all sites are located on a fairly uniform site
356 condition, the GPR model type 1 is implemented to estimate the test sites' ground motion time series.

357 About 80% of the 326 locations are randomly selected as the training set, while the remaining
358 20% are set aside as the test set. The observation density of the training set is about 0.25 stations/km²,
359 which makes the $\hat{\lambda}$ obtained for the GPR model type 1 (with approximately the same observation density
360 in Table 3) usable to estimate the ground motion time series at target sites. The distribution of the training

361 set and test sites is shown in Figure 8. In addition, the distribution of the NRMSE between the estimated
362 and “exact” (physics-based simulated) motions’ RotD50 spectra are shown in Figure 8. Two test sites
363 are shown in Figure 8, for which the prediction results are shown in Figure 9.

364 Table 5 summarizes the NRMSE of the linear response spectrum in both the FN and FP directions
365 and for RotD50. Table 5 shows that the M7.0 Hayward fault scenario earthquake test set, which was not
366 used during the hyper-parameter optimization, resulted in a higher average NRMSE for the RotD50
367 spectrum in comparison with the 1906 M7.9 San Francisco simulated motion data set. In addition, the
368 Hayward fault simulations consider the topography of the region, while the 1906 San Francisco
369 simulations were carried out for a horizontal surface. Moreover, the M7.0 Hayward fault simulations
370 used wave propagation for all frequencies, while the 1906 San Francisco simulations implemented
371 stochastic noise for frequencies above 1 Hz. Therefore, the Hayward fault data set includes more
372 complexity, which might lead to a higher estimation error than for the 1906 San Francisco data set.

373 Figure 8 demonstrates that the RotD50 spectra of the estimated ground motions have acceptable
374 NRMSE for the majority of the test locations for the M7.0 Hayward fault simulated motions. Moreover,
375 Figure 8 illustrates the applicability of the trained GPR model in the prediction of ground motion time
376 series for the sites close to the fault (less than 4 km away) and close to the epicenter (less than 9 km
377 away). For a few locations mainly located at the boundary edges of the simulation network, the estimated
378 ground motion time series are less accurate. Figure 9a demonstrates that the GPR model decently
379 predicted the long-period pulses due to the directivity effect for site 1. Figure 9 illustrates that the ground
380 motion time series estimation is more accurate for the long-period motions, while the shorter-period
381 shear waves might be less accurately predicted. This feature of our results is consistent with those
382 generated using the 1906 San Francisco earthquake.

383 ***2019 M7.1 Ridgecrest Earthquake***

384 It is useful to examine the trained GPR model's prediction with an actual recorded earthquake dataset.
385 To do so, we used motions of the M7.1 2019 Ridgecrest earthquake that were recorded by the
386 Community Seismic Network (CSN) within the northern Los Angeles basin (Clayton et al. 2020; Kohler
387 et al. 2020; Filippitzi et al. 2021) as another test set for the GPR model type 2. The M7.1 Ridgecrest
388 earthquake ground motions were sparsely recorded in the epicentral region as there are only about three
389 seismic stations within a $20 \text{ km} \times 20 \text{ km}$ area surrounding the epicenter (USGS ShakeMap for 2019
390 M7.1 Ridgecrest earthquake). Therefore, an area with an adequate number of recording stations is chosen
391 to evaluate the trained GPR model's performance. In this study, we chose 151 CSN ground-level stations
392 that had recorded the 2019 M7.1 Ridgecrest earthquake. The site condition, V_{s30} , of the recording stations
393 is estimated using a proxy-based model as described in Ahdi et al. (2020).

394 Of the total 151 recording stations, 149 are considered as the training set, while the remaining
395 two stations are set aside as the test stations. The observation density for 149 observed sites distributed
396 over a 492 km^2 region is about $0.30 \text{ stations/km}^2$, which makes the $\hat{\lambda}$ obtained for the GPR model type
397 2 (for $0.29 \text{ stations/km}^2$ density) applicable. Figure 10 shows the distribution of the observed stations
398 (training set) and their site conditions, V_{s30} , as well as the test stations for which the ground motion time
399 series are estimated. The trained GPR model type 2 is implemented for the conditioned simulation
400 procedure. As indicated, this GPR model is capable of incorporating variations in local soil conditions.
401 Figure 11 displays the prediction results for the two test stations shown in Figure 10. The predicted time
402 series are reliable only within the mutually usable frequency bandwidth (Ancheta et al. 2014) among all
403 observed motions, which is the reliable frequency range after the noise removal of the recorded motions.
404 Figure 11 displays the estimated and the exact (recorded) ground motions' RotD50 and FAS within the
405 overlapping usable frequency bandwidth of the observed motions.

406 Figure 11 shows that the results for ground motion time series estimation for the 2019 M7.1
407 Ridgecrest earthquake sequence are auspicious. The peak ground velocity (PGV) and the long-period

408 pulses of the recorded motions (Filippitzi et al. 2021) are captured fairly accurately at both test stations.
409 This implies that the GPR model type 2 can generate ground motions with acceptable accuracy. We plan
410 to investigate further the applicability of the GPR model type 2 using a broader set of recorded
411 earthquake data sets.

412

413 ***Realizations of Ground Motion***

414 The trained GPR model provides the posterior mean vector and posterior covariance matrix for the DFT
415 coefficients at each frequency for the target sites based on Eqs. 8 and 9. In this study, there is only one
416 target site to estimate the DFT coefficients at each prediction step. Thus, Eqs. 8 and 9 provide the DFT
417 coefficients' real and imaginary parts' posterior mean and posterior standard deviation at each frequency.

418 In addition, we estimate the correlation between $\mathcal{R}e_k$ and $\mathcal{I}m_k$ at k^{th} frequency, $k = 0, \dots, N-1$, at the
419 target site using the observed ground motion DFT coefficients at the same frequency. Therefore, we can
420 generate pairs of $(\mathcal{R}e_k, \mathcal{I}m_k)$ 2×1 random sample vectors having the 2×1 mean vector and 2×2
421 covariance matrix of the real and imaginary parts at each frequency. These generated samples can then
422 be converted to samples of amplitude, $|A_k|$ (cf. Eq. 2). We estimate the logarithmic mean and standard
423 deviation of amplitudes at each frequency using the $|A_k|$ samples. Eventually, we implement the inter-
424 frequency correlation model established by Bayless and Abrahamson (2019) to develop the covariance
425 matrix of $\log(|A_k|)$ for all frequencies, $k = 0, \dots, N-1$. We generate 150 multivariate Gaussian random
426 samples of FAS using the established $N \times 1$ mean vector and $N \times N$ covariance matrix. These FAS samples
427 are then combined with the phase spectrum constructed with mean estimated real and imaginary parts at
428 the target site to generate 150 random ground motion realizations.

429 Figure 12a depicts 150 random generated ground motion realizations' 5%-damped response
430 spectra as well as the logarithmic mean of those samples' response spectra along the East-West direction.
431 In addition, Figure 12a demonstrates that the estimated ground motion time series using mean DFT

432 coefficients has a similar response spectrum (solid black line in Figure 12a) to the logarithmic mean of
433 the realizations' response spectra. Figure 12b shows 68% confidence interval (mean \pm standard
434 deviation) of the ground motions' realizations on a logarithmic scale. It is observable in Figure 12b that
435 the estimated ground motion's response spectrum has higher uncertainty at shorter periods, while this
436 uncertainty decreases at longer periods. As evidenced in Figure 12b, the recorded motion's response
437 spectrum is located within the 68% confidence interval at most periods within the useable frequency
438 bandwidth.

439

440 CONCLUSION AND DISCUSSION

441 A novel approach to estimate the entire ground motion time series at a target location using the observed
442 surrounding motions was developed. The generated motions at target (uninstrumented) sites can be used
443 for site-specific nonlinear structural analysis as well as quantification of spatial damage distribution.

444 In this study, GPR was employed to estimate time series at the target sites through the real and imaginary
445 parts of the DFT coefficients. To do so, the GPR model's hyper-parameter, λ , was tuned using ground
446 motions of the physics-based simulated 1906 San Francisco earthquake. Two GPR models were
447 developed: one applicable to the homogeneous regions (relatively uniform site condition) and the other
448 usable in regions with considerable local site condition variation. The optimized $\hat{\lambda}$ for these models are
449 applicable for the regions with an approximately similar observation density. Both models demonstrated
450 acceptable performance for estimation of the ground motion, as well as the response spectra for the 1906
451 M7.9 San Francisco and M7.0 Hayward fault simulated ground motions. In addition, our investigation
452 demonstrated the applicability of the trained GPR model for estimation of the 2019 M7.1 Ridgecrest
453 earthquake recorded ground motions.

454 The trained GPR models estimated the long-period pulses decently. The estimation of the
455 motions for locations at the edges of the network, where there is a non-uniform distribution of
456 observations or regions with fewer observations, may not be as accurate as those at other locations. In
457 addition, the length-scale parameter of the covariance functions demonstrated that there is a higher
458 correlation for the long-period content of the ground motions compared to the short-period content within
459 a region. Therefore, the conditioned simulated ground motions are generally more reliable in the long-
460 period range than those at short periods. In addition, we incorporated the posterior mean and standard
461 deviation of the DFT coefficients as well as the inter-frequency correlations among neighboring
462 frequencies to generate random realizations of ground motions at the target site. The ground motion
463 realizations depicted that the uncertainty of the estimated ground motions is higher for the short periods.

464 The GPR models can be expanded by considering other site attributes such as $Z_{1.0}$, $Z_{2.5}$, and R_{JB}
465 as well as combining the covariance functions. In addition, using an anisotropic covariance function,
466 especially for regions closer to the fault, may improve the estimation.

467 DATA AND RESOURCES

468 The 1906 M7.9 San Francisco earthquake simulated ground motions were provided by Robert W. Graves
469 (Aagaard et al. 2008.) The RotD50 and orthogonal directions linear response spectra of the ground
470 motions were constructed using the R package for computation of earthquake ground motion response
471 spectra (Wang et al. 2017) which is accessible through <https://peer.berkeley.edu/peer-reports> (last
472 accessed February 2021.) The M7.0 Hayward fault scenario earthquake simulated motions (Rodgers et
473 al. 2019) were provided by Arthur J. Rodgers. The M7.1 2019 Ridgecrest earthquake data recorded by
474 the Community Seismic Network were obtained from <http://csn.caltech.edu/data/> (last accessed
475 February 2021). The processed recorded motions for the M7.1 2019 Ridgecrest earthquake can be
476 retrieved from <https://www.risksciences.ucla.edu/nhr3/gmdata> (last accessed February 2021). The

477 average shear wave velocity values, V_{s30} , at each CSN station were provided by Pengfei Wang using the
478 proxy-based model (Ahdi et al. 2020).

479 ACKNOWLEDGMENTS

480 This study was partially supported by the University of California, Los Angeles (UCLA) Graduate
481 Fellowship to the first author, which is gratefully acknowledged. Partial supports of the National Science
482 Foundation (Award number 2025310), California Department of Transportation and Pacific Gas &
483 Electric Company are also fully appreciated. Arthur Rodgers' work was performed under the auspices
484 of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-
485 AC52-07NA27344. Any opinions, findings, conclusions, or recommendations expressed in this
486 publication are those of the authors and do not necessarily reflect those of the supporting agencies. The
487 authors would like to also thank Dr. Robert Graves for providing the simulated ground-motion dataset
488 of the 1906 event and Prof. Tadahiro Kishida for his efforts in organizing and processing it. Drs. Sean
489 Ahdi and Pengfei Wang have kindly assisted in the estimation of V_{s30} values at the recording stations of
490 the 2019 Ridgecrest earthquake. The authors also benefitted from constructive discussions with Dr.
491 Silvia Mazzoni. The Comments from two BSSA anonymous reviewers are greatly appreciated.

492 REFERENCES

- 493 Aagaard, Brad T., Thomas M. Brocher, David Dolenc, Douglas Dreger, Robert W. Graves, Stephen Harmsen, Stephen
494 Hartzell et al. "Ground-Motion Modeling of the 1906 San Francisco Earthquake, Part II: Ground-Motion Estimates
495 for the 1906 Earthquake and Scenario Events Ground-Motion Modeling of the 1906 San Francisco Earthquake, Part
496 II." *Bulletin of the Seismological Society of America* 98, no. 2 (2008): 1012-1046.
- 497 Abrahamson, N. A., J. F. Schneider, and J. C. Stepp. "Empirical spatial coherency functions for application to soil-structure
498 interaction analyses." *Earthquake spectra* 7, no. 1 (1991): 1-27.
- 499 Abramowitz, Milton, and Irene A. Stegun. "Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical
500 Tables." National Bureau of Standards Applied Mathematics Series 55. Tenth Printing. (1972).
- 501 Adanur, Suleyman, Ahmet C. Altunisik, Kurtulus Soyuluk, A. Aydin Dumanoglu, and Alemdar Bayraktar. "Contribution of
502 local site-effect on the seismic response of suspension bridges to spatially varying ground motions." *Earthquakes and
503 Structures* 10, no. 5 (2016): 1233-1251.
- 504 Ahdi, Sean Kamran, Silvia Mazzoni, Tadahiro Kishida, Pengfei Wang, Chukwuebuka C. Nweke, Nicolas M. Kuehn, Victor
505 Contreras, Badie Rowshandel, Jonathan P. Stewart, and Yousef Bozorgnia. "Engineering characteristics of ground
506 motions recorded in the 2019 Ridgecrest earthquake sequence." *Bulletin of the Seismological Society of
507 America* 110, no. 4 (2020): 1474-1494.

- 508 Alimoradi, Arzhang, and James L. Beck. "Machine-learning methods for earthquake ground motion analysis and
509 simulation." *Journal of Engineering Mechanics* 141, no. 4 (2015): 04014147.
- 510 Ancheta, Timothy D., Robert B. Darragh, Jonathan P. Stewart, Emel Seyhan, Walter J. Silva, Brian S-J. Chiou, Katie E.
511 Wooddell et al. "NGA-West2 database." *Earthquake Spectra* 30, no. 3 (2014): 989-1005.
- 512 Baker, Jack W., and Yilin Chen. "Ground motion spatial correlation fitting methods and estimation uncertainty." *Earthquake
513 Engineering & Structural Dynamics* (2020).
- 514 Bayless, Jeff, and Norman A. Abrahamson. "An Empirical Model for the Interfrequency Correlation of Epsilon for Fourier
515 Amplitude SpectraAn Empirical Model for the Interfrequency Correlation of Epsilon for Fourier Amplitude
516 Spectra." *Bulletin of the Seismological Society of America* 109, no. 3 (2019): 1058-1070.
- 517 Boore, David M. "Orientation-independent, nongeometric-mean measures of seismic intensity from two horizontal
518 components of motion." *Bulletin of the Seismological Society of America* 100, no. 4 (2010): 1830-1835.
- 519 CESMD (2019). Center for engineering strong motion data, <https://strongmotioncenter.org> (last accessed February 2021)
- 520 Chen, Yilin, and Jack W. Baker. "Spatial Correlations in CyberShake Physics-Based Ground-Motion Simulations." *Bulletin
521 of the Seismological Society of America* 109, no. 6 (2019): 2447-2458.
- 522 Clayton, Robert W., Monica Kohler, Richard Guy, Julian Bunn, Thomas Heaton, and Mani Chandy. "CSN-LAUSD network:
523 A dense accelerometer network in Los Angeles Schools." *Seismological Research Letters* 91, no. 2A (2020): 622-
524 630.
- 525 Fan, Jianqing, and Runze Li. "Variable selection via nonconcave penalized likelihood and its oracle properties." *Journal of
526 the American statistical Association* 96, no. 456 (2001): 1348-1360.
- 527 Fraser, William A., David J. Wald, and Kuo-Wan Lin. "Using ShakeMap and ShakeCast to Prioritize Post-Earthquake Dam
528 Inspections." In *Geotechnical Earthquake Engineering and Soil Dynamics IV*, pp. 1-10. (2008).
- 529 Filippitzi, F., M. D. Kohler, T. H. Heaton, R. W. Graves, R. W. Clayton, R. G. Guy, J. J. Bunn, K. M. Chandy, "Ground
530 motions in urban Los Angeles from the 2019 Ridgecrest earthquake sequence", *Earthquake Spectra*, early version
531 available online, doi:10.1177/87552930211003916, (2021).
- 532 Gentile, Roberto, and Carmine Galasso. "Gaussian process regression for seismic fragility assessment of building
533 portfolios." *Structural Safety* 87 (2020): 101980.
- 534 Ghaderi, Aref, Vahid Morovati, and Roozbeh Dargazany. "A Bayesian Surrogate Constitutive Model to Estimate Failure
535 Probability of Rubber-Like Materials." arXiv preprint arXiv:2010.13241 (2020).
- 536 Hjort, Nils Lid, Chris Holmes, Peter Müller, and Stephen G. Walker, eds. "Bayesian nonparametrics". Vol. 28. Cambridge
537 University Press, (2010).
- 538 Huang, Chen, Yicong Liang, Xiaoqing Ding, and Chi Fang. "Generalized joint kernel regression and adaptive dictionary
539 learning for single-image super-resolution." *Signal processing* 103 (2014): 142-154.
- 540 Huang, Duruo, and Gang Wang. "Energy-compatible and spectrum-compatible (ECSC) ground motion simulation using
541 wavelet packets." *Earthquake Engineering & Structural Dynamics* 46, no. 11 (2017): 1855-1873.
- 542 Jayaram, Nirmal, and Jack W. Baker. "Correlation model for spatially distributed ground-motion intensities." *Earthquake
543 Engineering & Structural Dynamics* 38, no. 15 (2009): 1687-1708.
- 544 Kameda, Hiroyuki, and Hitoshi Morikawa. "An interpolating stochastic process for simulation of conditional random
545 fields." *Probabilistic Engineering Mechanics* 7, no. 4 (1992): 243-254.
- 546 Kiureghian, A. Der. "A coherency model for spatially varying ground motions." *Earthquake engineering & structural
547 dynamics* 25, no. 1 (1996): 99-111.
- 548 Konakli, Katerina, and Armen Der Kiureghian. "Simulation of spatially varying ground motions including incoherence,
549 wave-passage and differential site-response effects." *Earthquake Engineering & Structural Dynamics* 41, no. 3
550 (2012): 495-513.
- 551 Kohler, Monica D., Filippos Filippitzi, Thomas Heaton, Robert W. Clayton, Richard Guy, Julian Bunn, and K. Mani Chandy.
552 "2019 Ridgecrest Earthquake Reveals Areas of Los Angeles That Amplify Shaking of High-Rises." *Seismological
553 Research Letters*, 91, no. 6 (2020): 3370-3380.
- 554 Kuehn, Nicolas M., and Norman A. Abrahamson. "Spatial correlations of ground motion for non-ergodic seismic hazard

- 555 analysis." *Earthquake Engineering & Structural Dynamics* 49, no. 1 (2020): 4-23.
- 556 Landwehr, Niels, Nicolas M. Kuehn, Tobias Scheffer, and Norman Abrahamson. "A nonergodic ground-motion model for
557 California with spatially varying coefficients." *Bulletin of the Seismological Society of America* 106, no. 6 (2016):
558 2574-2583.
- 559 Li, Runze, and Agus Sudjianto. "Analysis of computer experiments using penalized likelihood in Gaussian Kriging
560 models." *Technometrics* 47, no. 2 (2005): 111-120.
- 561 Liao, S., and A. Zerva. "Physically compliant, conditionally simulated spatially variable seismic ground motions for
562 performance-based design." *Earthquake engineering & structural dynamics* 35, no. 7 (2006): 891-919.
- 563 Lin, K., D. J. Wald, C. A. Kircher, D. Slosky, K. Jaiswal, and N. Luco. "USGS SHAKECAST SYSTEM
564 ADVANCEMENTS." (2018).
- 565 Lu, Xinzhen, Qingle Cheng, Yuan Tian, and Yuli Huang. "Regional Ground-Motion Simulation Using Recorded Ground
566 Motions." *Bulletin of the Seismological Society of America* (2021).
- 567 Nadaraya, Elizbar A. "On estimating regression." *Theory of Probability & Its Applications* 9, no. 1 (1964): 141-142.
- 568 Oppenheim, Alan V., Alan S. Willsky, and S. Hamid Nawab. "Signals and systems Prentice Hall." Inc., Upper Saddle River,
569 New Jersey 7458 (1997).
- 570 Otake, Ryota, Jun Kurima, Hiroyuki Goto, and Sumio Sawada. "Deep Learning Model for Spatial Interpolation of Real-Time
571 Seismic Intensity." *Seismological Society of America* 91, no. 6 (2020): 3433-3443.
- 572 Petrone, Floriana, Norman Abrahamson, David McCallen, and Mamun Miah. "Validation of (not-historical) large-event near-
573 fault ground-motion simulations for use in civil engineering applications." *Earthquake Engineering & Structural
574 Dynamics* (2020).
- 575 Rasmussen, Carl Edward, and Christopher KI Williams. "Gaussian Processes for Machine Learning", the MIT
576 Press. Cambridge, Mass (2006).
- 577 Rodda, Gopala Krishna, and Dhiman Basu. "Spatial variation and conditional simulation of seismic ground motion." *Bulletin
578 of Earthquake Engineering* 16, no. 10 (2018): 4399-4426.
- 579 Rodda, Gopala Krishna, and Dhiman Basu. "On Conditional Simulation of Spatially Varying Rotational Ground
580 Motion." *Journal of Earthquake Engineering* (2019): 1-36.
- 581 Rodgers, Arthur J., N. Anders Petersson, Arben Pitarka, David B. McCallen, Bjorn Sjogreen, and Norman Abrahamson.
582 "Broadband (0–5 Hz) fully deterministic 3D ground-motion simulations of a magnitude 7.0 Hayward fault
583 earthquake: Comparison with empirical ground-motion models and 3D path and site effects from source normalized
584 intensities." *Seismological Research Letters* 90, no. 3 (2019): 1268-1284.
- 585 Rubin, Donald B. "The bayesian bootstrap." *The annals of statistics* (1981): 130-134.
- 586 Sajedi, Seyed Omid, and Xiao Liang. "A data-driven framework for near real-time and robust damage diagnosis of building
587 structures." *Structural Control and Health Monitoring* 27, no. 3 (2020): e2488.
- 588 Savitzky, Abraham, and Marcel JE Golay. "Smoothing and differentiation of data by simplified least squares procedures.
589 " *Analytical chemistry* 36, no. 8 (1964): 1627-1639.
- 590 Sheibani, Mohamadreza, and Ge Ou. "The development of Gaussian process regression for effective regional post-earthquake
591 building damage inference." *Computer-Aided Civil and Infrastructure Engineering* (2020).
- 592 Southern California Earthquake Data Center. [Online]. Available at:
593 <https://service.scedc.caltech.edu/SCSNStationMap/station.html> (last access February 2021)
- 594 Sun, Han, Henry Burton, Yu Zhang, and John Wallace. "Interbuilding interpolation of peak seismic response using spatially
595 correlated demand parameters." *Earthquake Engineering & Structural Dynamics* 47, no. 5 (2018): 1148-1168.
- 596 Tamhidi, A., Kuehn, N. M., Kohler, M. D., Ghahari, F., Taciroglu, E., & Bozorgnia, Y., "Ground-Motion Time-Series
597 Interpolation within the Community Seismic Network using Gaussian Process Regression: Application to the 2019
598 Ridgecrest Earthquake". Poster Presentation at 2020 SCEC Annual Meeting, (2020).
- 599 Tamhidi, A., Kuehn, N., Bozorgnia, Y., Taciroglu, E., & Kishida, T. "Prediction of Ground-Motion Time-Series at an
600 arbitrary location using Gaussian Process Interpolation: Application to the Ridgecrest Earthquake". Poster
601 Presentation at 2019 SCEC Annual Meeting, (2019).

- 602 Tian, Li, Xia Gai, Bing Qu, Hongnan Li, and Peng Zhang. “Influence of spatial variation of ground motions on dynamic
603 responses of supporting towers of overhead electricity transmission systems: An experimental study.” *Engineering*
604 *Structures* 128 (2016): 67-81.
- 605 Tibshirani, Robert. “Regression shrinkage and selection via the lasso.” *Journal of the Royal Statistical Society: Series B*
606 (Methodological) 58, no. 1 (1996): 267-288.
- 607 Todorovska, Maria I., Haiping Ding, and Mihailo D. Trifunac. “Coherency of Synthetic Earthquake Ground Motion for the
608 Design of Long Structures: Effect of Site Conditions.” In *International Collaboration in Lifeline Earthquake*
609 *Engineering* 2016, pp. 427-434. Reston, VA: American Society of Civil Engineers, (2017).
- 610 U.S. Geologic Survey (2018). 3-D geologic and seismic velocity models of the San Francisco Bay region, available at
611 [https://www.usgs.gov/natural-hazards/earthquake-hazards/science/3-d-geologic-and-seismic-velocity-models-san-](https://www.usgs.gov/natural-hazards/earthquake-hazards/science/3-d-geologic-and-seismic-velocity-models-san-francisco)
612 [francisco](https://www.usgs.gov/natural-hazards/earthquake-hazards/science/3-d-geologic-and-seismic-velocity-models-san-francisco) (last accessed February 2021).
- 613 U.S. Geological Survey (2019). M7.1-2019 Ridgecrest Earthquake Sequence, available at
614 <https://earthquake.usgs.gov/earthquakes/eventpage/ci38457511/executive> (last accessed May 2021.)
- 615 Wald, David Jay, Vincent Quitoriano, Charles Bruce Worden, Margaret Hopper, and James W. Dewey. “USGS “Did You
616 Feel It?” internet-based macroseismic intensity maps.” *Annals of geophysics* 54, no. 6 (2012).
- 617 Wald, David, Kuo-Wan Lin, Keith Porter, and Loren Turner. “ShakeCast: Automating and improving the use of ShakeMap
618 for post-earthquake decision-making and response.” *Earthquake Spectra* 24, no. 2 (2008): 533-553.
- 619 Wang, Pengfei, Jonathan P. Stewart, Yousef Bozorgnia, David M. Boore, and Tadahiro Kishida. *R Package for Computation*
620 *of Earthquake Ground Motion Response Spectra*. No. 2017/09. Report, (2017).
- 621 Watson, Geoffrey S. “Smooth regression analysis.” *Sankhyā: The Indian Journal of Statistics, Series A* (1964): 359-372.
- 622 Worden, C. Bruce, Eric M. Thompson, Jack W. Baker, Brendon A. Bradley, Nicolas Luco, and David J. Wald. “Spatial and
623 spectral interpolation of ground-motion intensity measure observations.” *Bulletin of the Seismological Society of*
624 *America* 108, no. 2 (2018): 866-875.
- 625 Wu, Yongxin, Yufeng Gao, Ning Zhang, and Dayong Li. “Simulation of spatially varying ground motions in V-shaped
626 symmetric canyons.” *Journal of Earthquake Engineering* 20, no. 6 (2016): 992-1010.
- 627 Zerva, Aspasia. “Spatial variation of seismic ground motions: modeling and engineering applications.”, Crc Press, (2009).
- 628 Zerva, Aspasia, Mohammad Reza Falamarz-Sheikhabadi, and Masoud Khazaei Poul. “Issues with the use of spatially variable
629 seismic ground motions in engineering applications.” In *European Conference on Earthquake Engineering*
630 *Thessaloniki, Greece*, pp. 225-252. Springer, Cham, (2018).
- 631 Zerva, Aspasia and Zervas Vassilios. “Spatial variation of seismic ground motions: an overview.” *Applied Mechanics*
632 *Reviews* 55, no. 3 (2002): 271-297.
- 633 Zentner, Irmela. “Simulation of non-stationary conditional ground motion fields in the time domain.” *Georisk: Assessment*
634 *and Management of Risk for Engineered Systems and Geohazards* 7, no. 1 (2013): 37-48.

635

636

637

638

639

640

641

642

AUTHOR AFFILIATIONS

643

644 Aidin Tamhidi, Ph.D. Candidate, University of California, Los Angeles, Department of Civil and
645 Environmental Engineering, 90095, Los Angeles, CA; aidintmahidi@ucla.edu

646

647 Nicolas Kuehn, Research Scientist, University of California, Los Angeles, J. Garrick Institute for the
648 Risk Science, 94804, Richmond, CA; kuehn@ucla.edu

649

650 S. Farid Ghahari, Research Scientist, University of California, Los Angeles, Department of Civil and
651 Environmental Engineering, 90095, Los Angeles, CA; ghahari@seas.ucla.edu

652

653 Arthur J. Rodgers, Staff Scientist, Lawrence Livermore National Laboratory, 94550, Livermore, CA;
654 arodgers@llnl.gov

655

656 Monica D. Kohler, Research Professor, California Institute of Technology, Department of Mechanical
657 and Civil Engineering, 91125, Pasadena, CA; kohler@caltech.edu

658

659 Ertugrul Taciroglu, Professor, University of California, Los Angeles, Department of Civil and
660 Environmental Engineering, 90095, Los Angeles, CA; etacir@ucla.edu

661

662 Yousef Bozorgnia, Professor, University of California, Los Angeles, Department of Civil and
663 Environmental Engineering, 90095, Los Angeles, CA; yousef.bozorgnia@ucla.edu

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

TABLES

683

Table 1. 1906 M7.9 San Francisco physics-based simulated wave propagation parameters from Aagaard et al. (2008)

684

Domain			Resolution		Features			
Length (km)	Width (km)	Maximum depth	Bandwidth	Minimum V_s	Topography	Water	Material Properties	Attenuation
555	162	45	$T > 1.0$ sec	760 m/s	Bulldozed	Sediment filled	USGS 05.1.0	Graves(Aagaard et al. 2008)

685

686

Table 2. optimized regularization factor, $\hat{\lambda}$, and obtained average NRMSE over the South Napa training set

Covariance Kernel	$\hat{\lambda}$ (FN)	$\hat{\lambda}$ (FP)	$Error_{avg}$ (FN)	$Error_{avg}$ (FP)
Exponential	1.3	1.3	0.36	0.36
Matérn ($\nu = 1.5$)	0.7	0.7	0.28	0.28
Matérn ($\nu = 2.5$)	0.7	0.7	0.30	0.31

687

688

Table 3. optimized regularization factor, $\hat{\lambda}$, for the models type 1 and type 2

GPR model type 1		GPR model type 2	
Fault-Normal	Fault-Parallel	Fault-Normal	Fault-Parallel
1.2	1.2	0.7	0.7

689

690

Table 4. 1906 M7.9 San Francisco test set's NRMSE (normalized root mean square error) for model type 1 and type 2

691

Model Type	Study Region	FN		FP		RotD50	
		Average	Standard Deviation	Average	Standard Deviation	Average	Standard Deviation
Type 1	East Bay	0.23	0.08	0.23	0.08	0.19	0.07
Type 2	Palo Alto	0.34	0.29	0.38	0.38	0.31	0.36
Type 2	South Napa	0.23	0.06	0.26	0.1	0.19	0.05

692

693

Table 5. M7.0 Hayward fault scenario earthquake simulated motions test set's NRMSE (normalized root mean square error) for model type 1 and type 2

694

Model Type	Study Region	FN		FP		RotD50	
		Average	Standard Deviation	Average	Standard Deviation	Average	Standard Deviation
Type 1	M7.0 Hayward Fault	0.28	0.07	0.31	0.11	0.25	0.07

695

696

697

698

FIGURE CAPTIONS

699

700

701 **Figure 1.** a) Aagaard et al. (2008) 1906 M7.9 San Francisco earthquake simulated ground motions
702 domain and b) the study regions corresponding to the type 1 (East Bay) and type 2 Gaussian Process
703 regression models (Palo Alto and South Napa)

704

705 **Figure 2.** Distribution of the training and test sets for the a) East Bay, b) Palo Alto, and c) South Napa
706 study regions within the 1906 San Francisco simulated motions domain

707

708 **Figure 3.** The distribution of the test set's NRMSE for the 5%-damped RotD50 spectrum for the GPR
709 model a) type 1 in East Bay, b) type 2 in Palo Alto, and c) type 2 in South Napa study regions

710

711 **Figure 4.** The RotD50, velocity time series, and FAS of the predicted as well as the exact motions
712 along Fault-Normal direction for the chosen test sites: a) No. 1, b) No. 2, and c) No. 3 within East Bay
713 study region

714

715 **Figure 5.** The RotD50, velocity time series, and FAS of the predicted as well as the exact motions
716 along Fault-Normal direction for the chosen test sites a) No. 1, b) No. 2, and c) No. 3 within the Palo
717 Alto study region

718

719 **Figure 6.** The RotD50, velocity time series, and FAS of the predicted as well as the exact motions
720 along Fault-Normal direction for the chosen test sites a) No. 1, b) No. 2, and c) No. 3 within the South
721 Napa study region

722

723 **Figure 7.** $\hat{\theta}$ for real part (Re) covariance functions along a) Fault-Normal and b) Fault-Parallel
724 directions within Palo Alto and South Napa study regions

725

726 **Figure 8.** Distribution of the training and test set as well as the test set's NRMSE for the 5%-damped
727 RotD50 spectrum for the M7.0 Hayward fault scenario earthquake simulated motions study region

728

729 **Figure 9.** The RotD50, velocity time series, and FAS of the predicted as well as the exact motions
730 along Fault-Normal direction for the chosen test sites a) No. 1 and b) No. 2 within the M7.0 Hayward
731 fault scenario earthquake simulated motions study region

731

732 **Figure 10.** Distribution of the training set and test stations for the CSN network

733

734 **Figure 11.** The RotD50, velocity time series, and FAS of the predicted as well as the exact motions
735 along East-West direction for the chosen test stations a) No. 1 and b) No. 2 within the CSN network
736 that recorded the 2019 M7.1 Ridgecrest earthquake

737 **Figure 12.** The 5%-damped PSA along East-West direction at test station No. 1 within the CSN
 738 network that recorded the 2019 M7.1 Ridgecrest earthquake for a) 150 random ground motion
 739 realizations and b) 68% Confidence Interval

740
 741

FIGURES

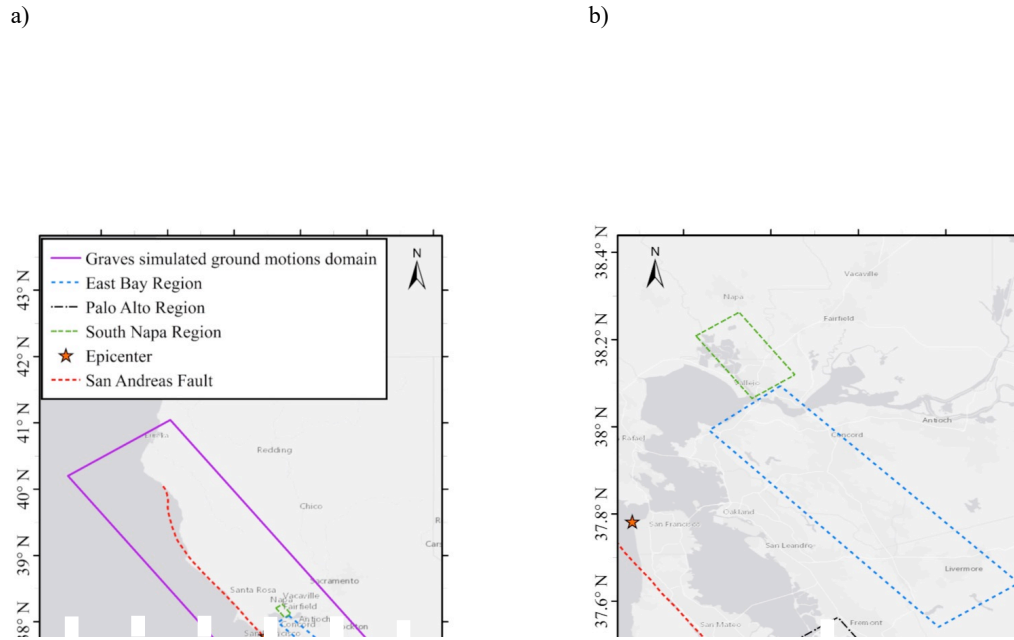


Figure 1. a) Aagaard et al. (2008a) 1906 M7.9 San Francisco earthquake simulated ground motions domain and b) the study regions corresponding to the type 1 (East Bay) and type 2 Gaussian Process regression models (Palo Alto and South Napa)

742
 743

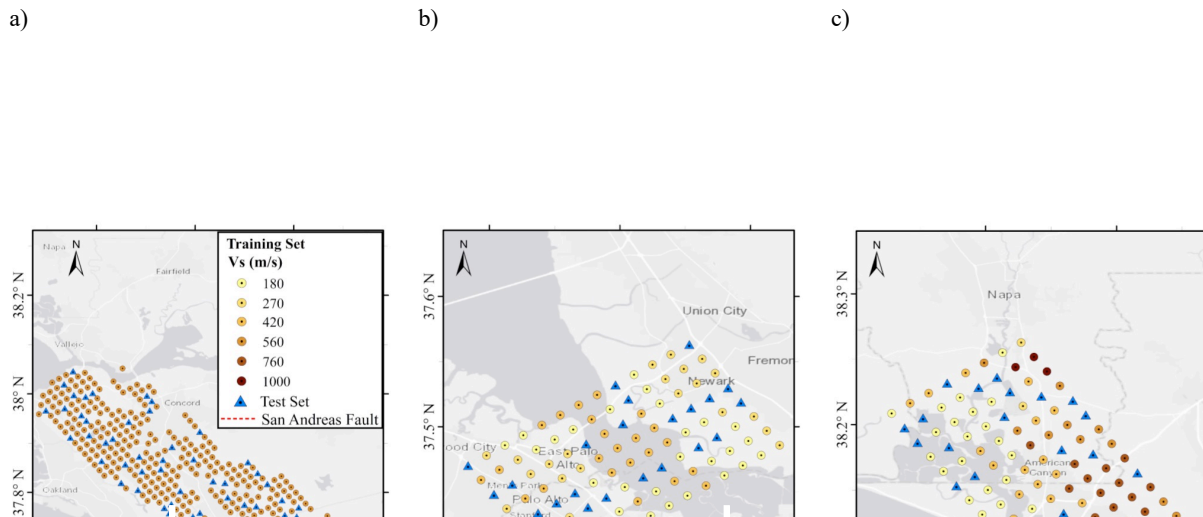


Figure 2. Distribution of the training and test sets for the a) East Bay, b) Palo Alto, and c) South Napa study regions within the 1906 San Francisco simulated motions domain

744

745

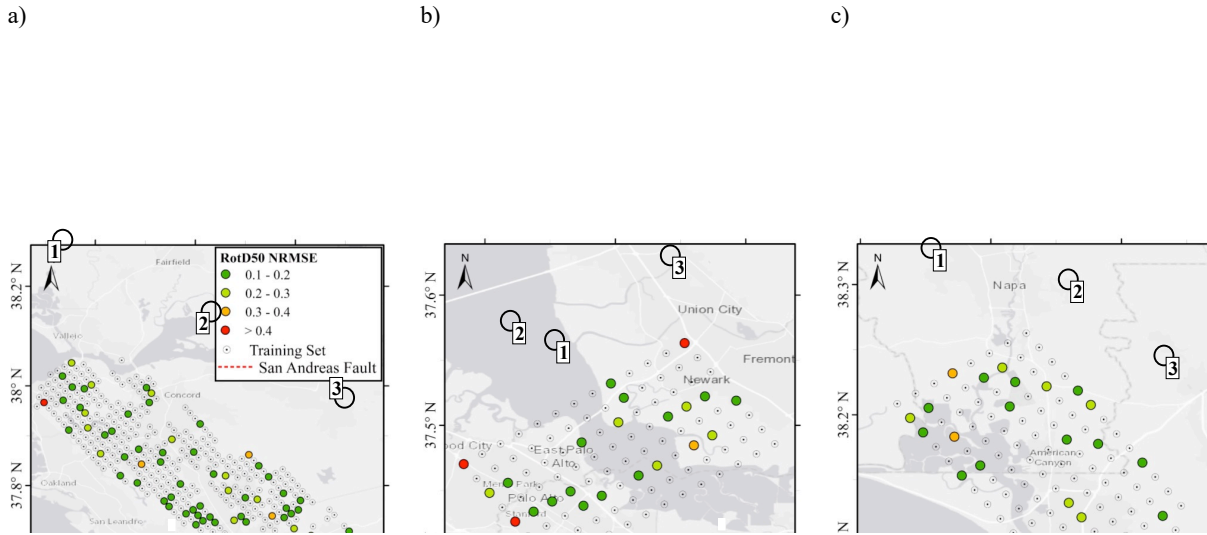


Figure 3. The distribution of the test set's NRMSE for the RotD50 spectrum (5% damping) for the GPR model a) type 1 in East Bay, b) type 2 in Palo Alto, and c) type 2 in South Napa study regions

746

747

748

749

750

751

752

753

754

755

756

757

758

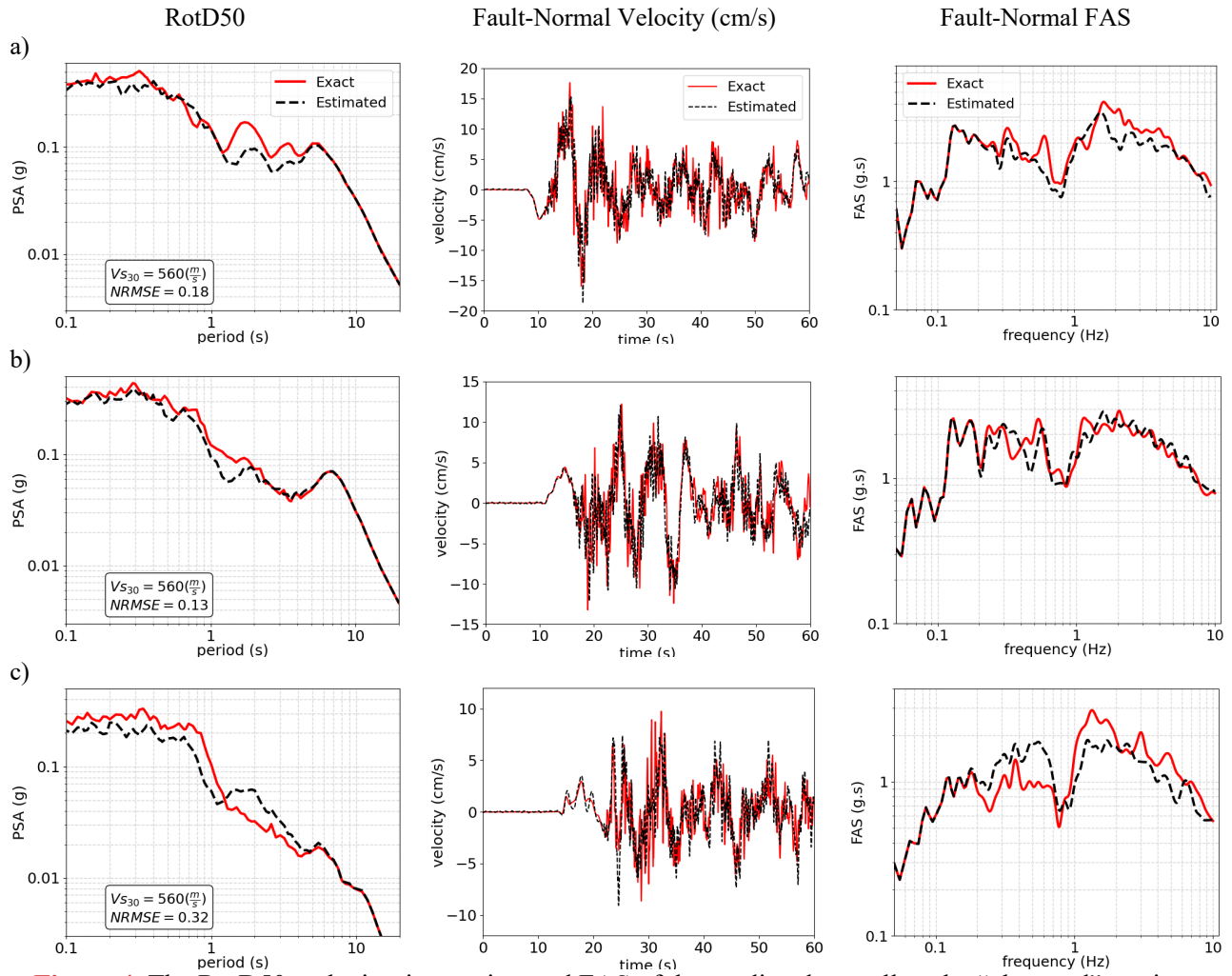


Figure 4. The RotD50, velocity time series, and FAS of the predicted as well as the “observed” motions along Fault-Normal direction for the chosen test locations: a) No. 1, b) No. 2, and c) No. 3 within East Bay study region

759

760

761

762

763

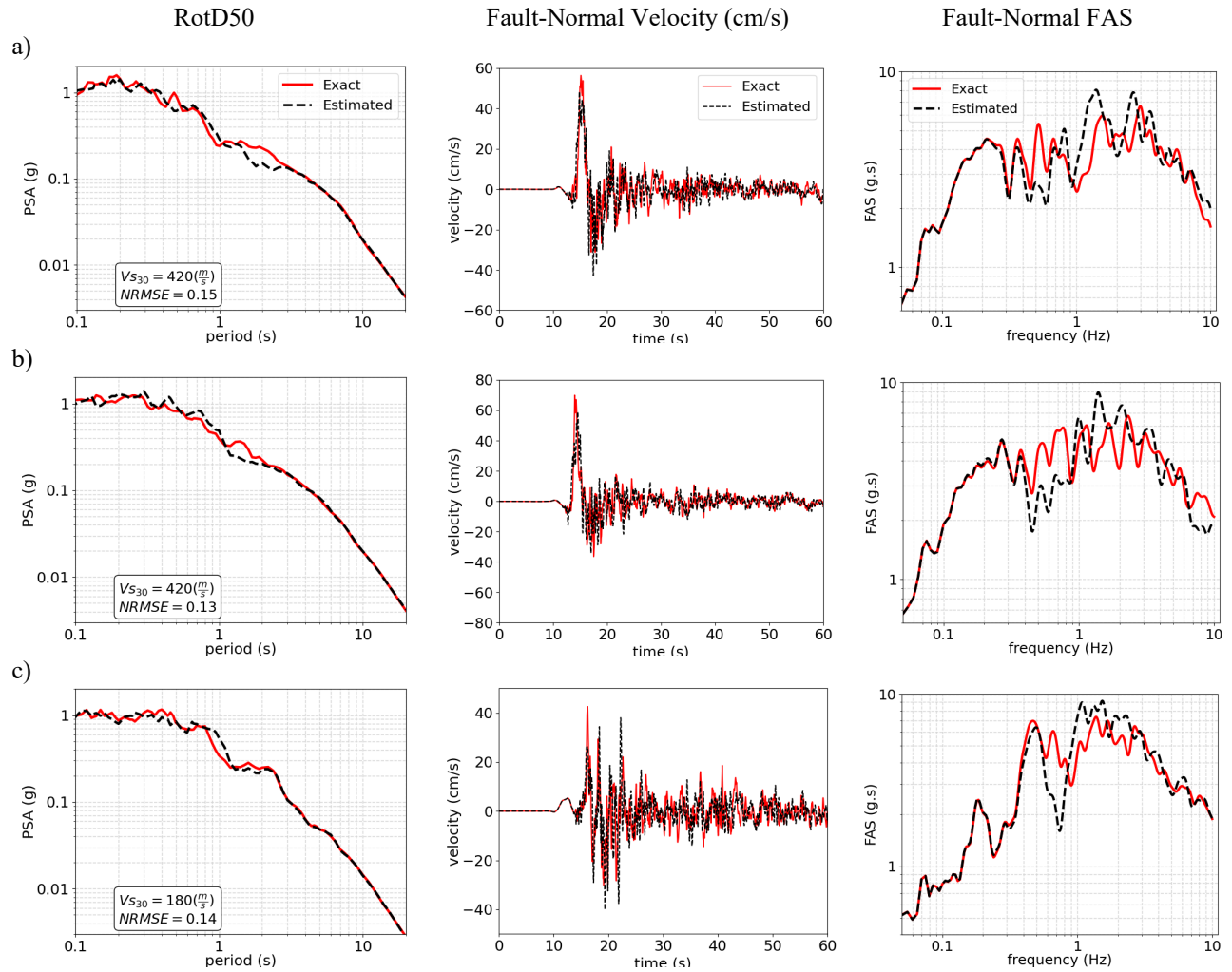


Figure 5. The RotD50, velocity time series, and FAS of the predicted as well as the “observed” motions along Fault-Normal direction for the chosen test stations a) No. 1, b) No. 2, and c) N. 3 within the Palo Alto study region

767
768
769
770
771
772
773
774
775

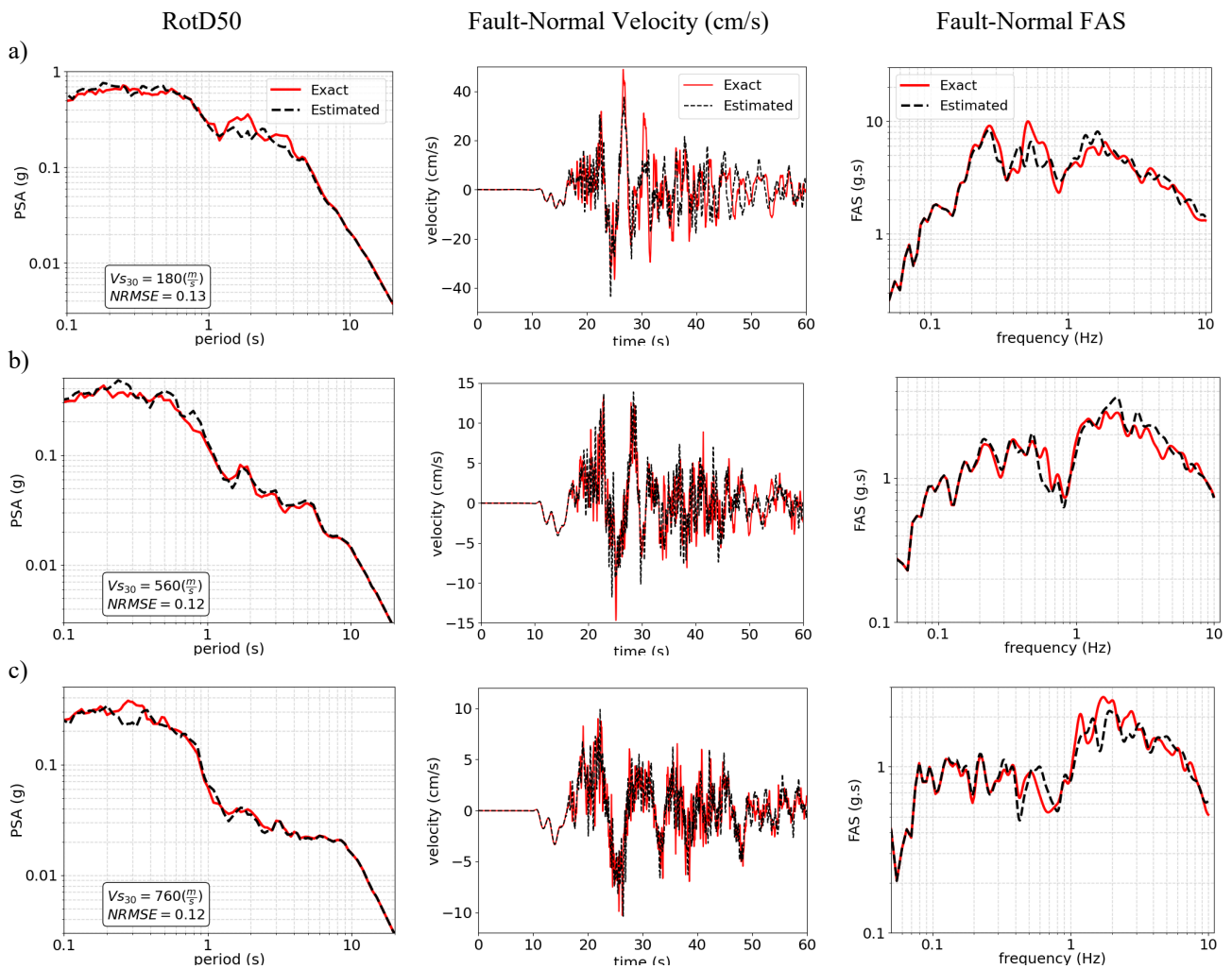


Figure 6. The RotD50, velocity time series, and FAS of the predicted as well as the “observed” motions along Fault-Normal direction for the chosen test stations a) No. 1, b) No. 2, and c) No. 3 within the South Napa study region

776

a)

b)

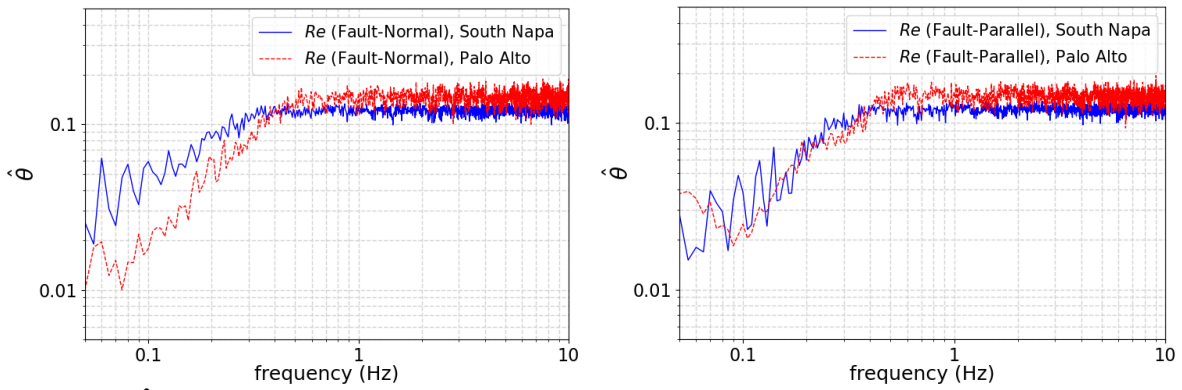


Figure 7. $\hat{\theta}$ for real part (*Re*) covariance functions along a) Fault-Normal and b) Fault-Parallel directions within Palo Alto and South Napa study regions

777

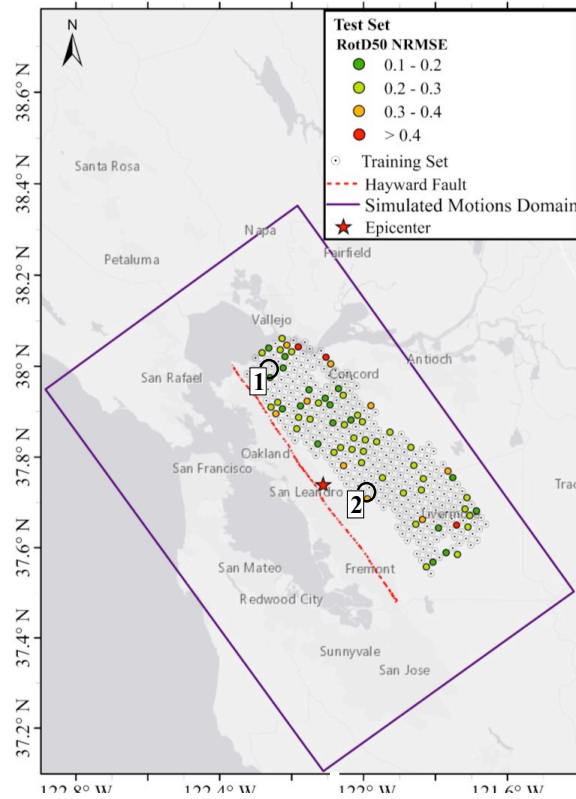


Figure 8. Distribution of the a) training and test set and b) test set's NRMSE for the RotD50 spectrum (5% damping) for the M7.0 Hayward fault simulated motions study region

778

779

780

781

782

783

784

785

786

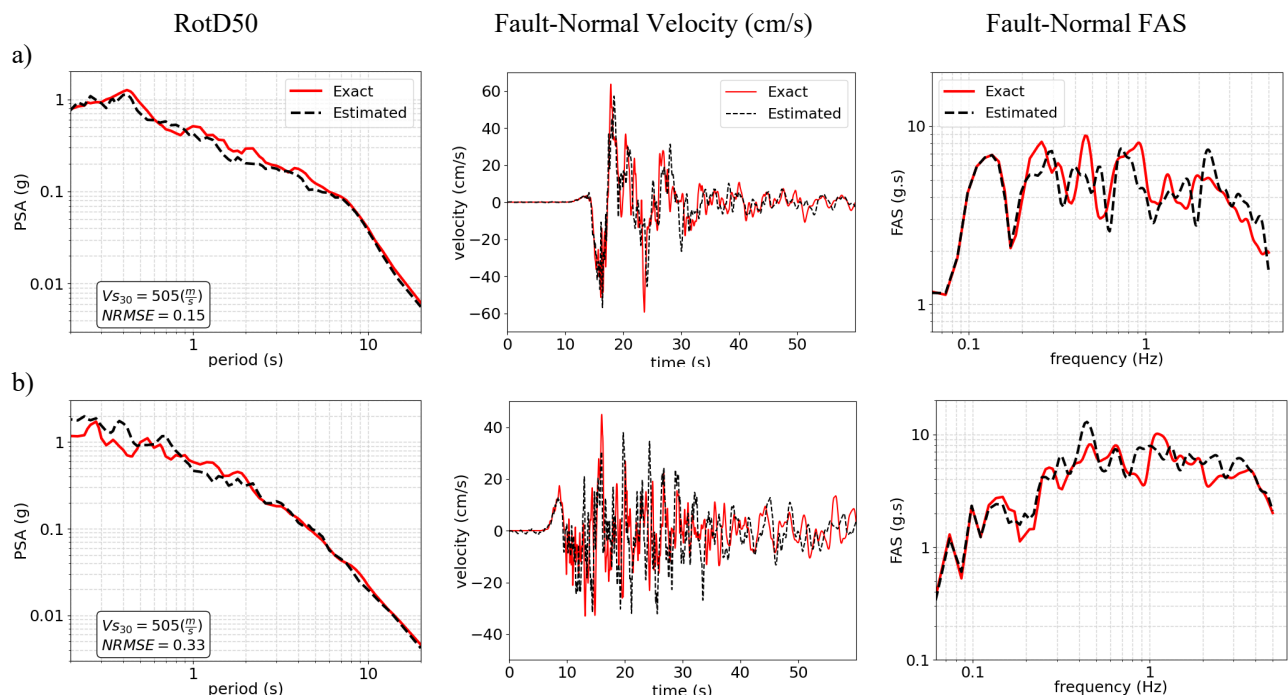


Figure 9. The RotD50, velocity time series, and FAS of the predicted as well as the “observed” motions along Fault-Normal direction for the chosen test locations a) No. 1 and b) No. 2 within the scenario M7.0 Hayward fault simulated motions study region

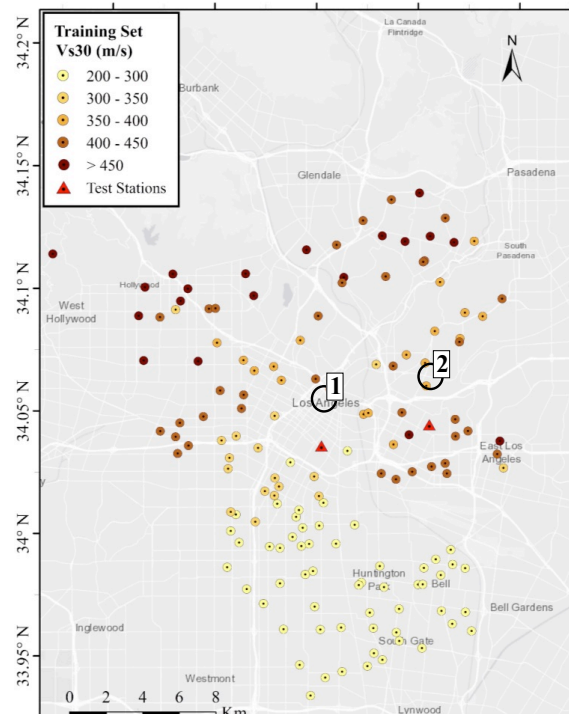
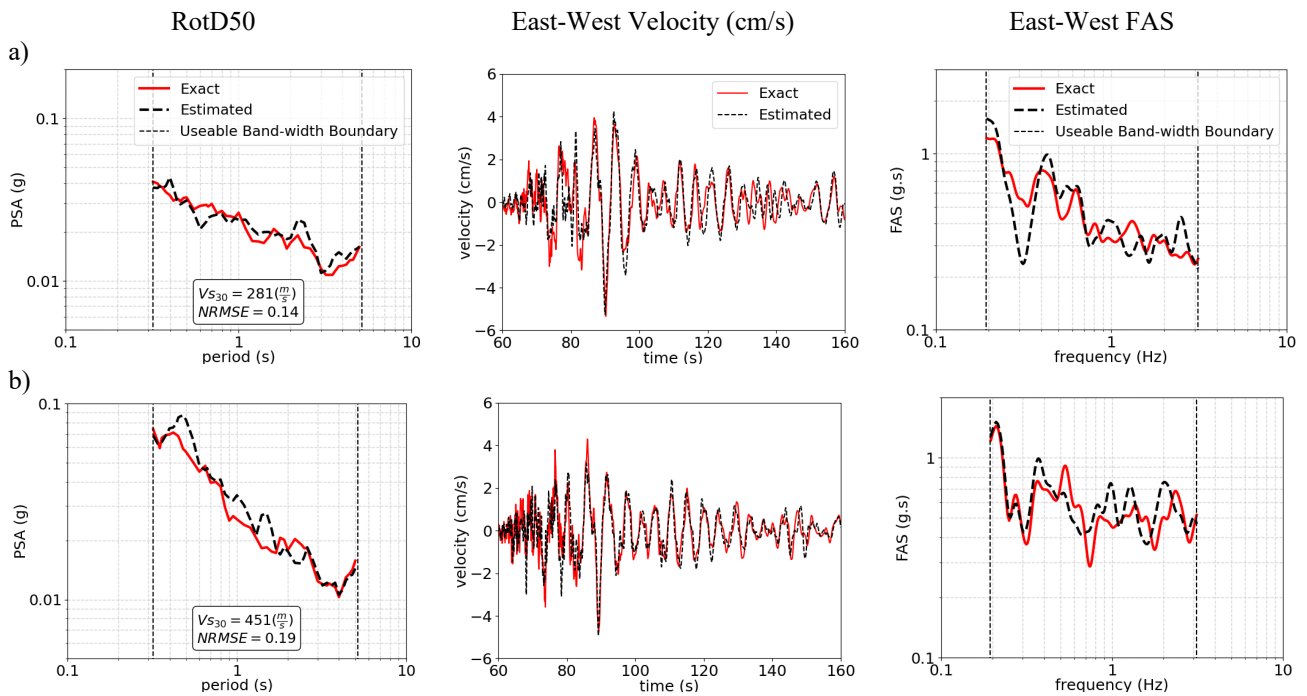


Figure 10. Distribution of the training set and test locations for the CSN network



| **Figure 11.** The RotD50, velocity time series, and FAS of the predicted as well as the observed motions along East-West direction for the chosen test locations a) No. 1 and b) No. 2 within the CSN network recorded the 2019 M7.1 Ridgecrest earthquake

790

791