Development and Verification of Regression and Application Software Tools for Non-ergodic Ground-Motion Models

G. Lavrentiadis[1,2†], E. Seylabi[3], N.M. Kuehn[1], X. Meng[4], A.R. Kottke[5], Y. Bozorgnia[1], C.A. Goulet[4]

[1]University of California, Los Angeles; [2]California Institute of Technology; [3]University of Nevada, Reno; [4]University of Southern California; [5]Pacific Gas & Electric Company

†glavrent@caltech.edu

Abstract

Non-ergodic ground-motion models (NGMMs) are a promising development in probabilistic seismic hazard analysis (PSHA) as they have the potential of reducing the ground-motion aleatory variability. This reduction in aleatory variability is accompanied by epistemic uncertainty in regions with sparse recordings or a systematic shift in the median ground motion in regions with dense recordings. The use of NGMMs can have a large impact on the seismic hazard, especially at long return periods which are important for critical infrastructure.

Gaussian Process Regression (GPR) – with spatially varying coefficients for modeling the source and site systematic effects and cell-specific anelastic attenuation for modeling the systematic path effects – is a flexible and robust modeling technique for developing NGMMs. As part of this work, open-source computer tools and instructions have been developed to show the steps toward developing and applying NGMMs in the GPR framework. The developed tools are written in Python and R using the statistical packages STAN and INLA. The software tools were tested against synthetic data sets with known non-ergodic effects. Different implementations of the developed software were evaluated for scalability, universality, precision, and model complexity.

Introduction

Research Objectives:

- Develop open-source software tools for the derivation and application of NGMMs
- Assist in the adoption of NGMM, in general, and these tools, in particular, through technical reports, workshops, and instructional videos
- Perform verification exercises with synthetic datasets to evaluate the accuracy, scalability, and model complexity of the developed tools

Ergodic versus Non-ergodic:

Ergodic GMMs:

- Developed with global data from similar tectonic environments
- Stable median GM, large aleatory variability
- GM predictions independent of source and site location

Non-ergodic GMMs:

- Separate the repeatable source, path, and site effects from the aleatory variability for the GMM
- Smaller aleatory variability but larger epistemic uncertainty in the absence of site-specific data
- GM predictions are source and site location dependent
- Can result in large changes in the seismic hazard

There is no change in the amount of information between ergodic and non-ergodic NGMMs:

Epistemic Terms / Aleatory terms

Residuals: $\delta W_{es} + \delta B_e \approx \overbrace{\delta L2L_e + \delta P2P_{es} + \delta S2S_s}^{\text{Epistemic Terms}} + \overbrace{\delta W^0_{es} + \delta B^0_e}^{\text{Aleatory terms}}$

St. Dev.: $\phi^2 + \tau^2 \approx \tau^2_{L2L} + \phi^2_{P2P} + \phi^2_{S2S} + \phi^2_0 + \tau^2_0$

Ergodic Components / Non-ergodic Components

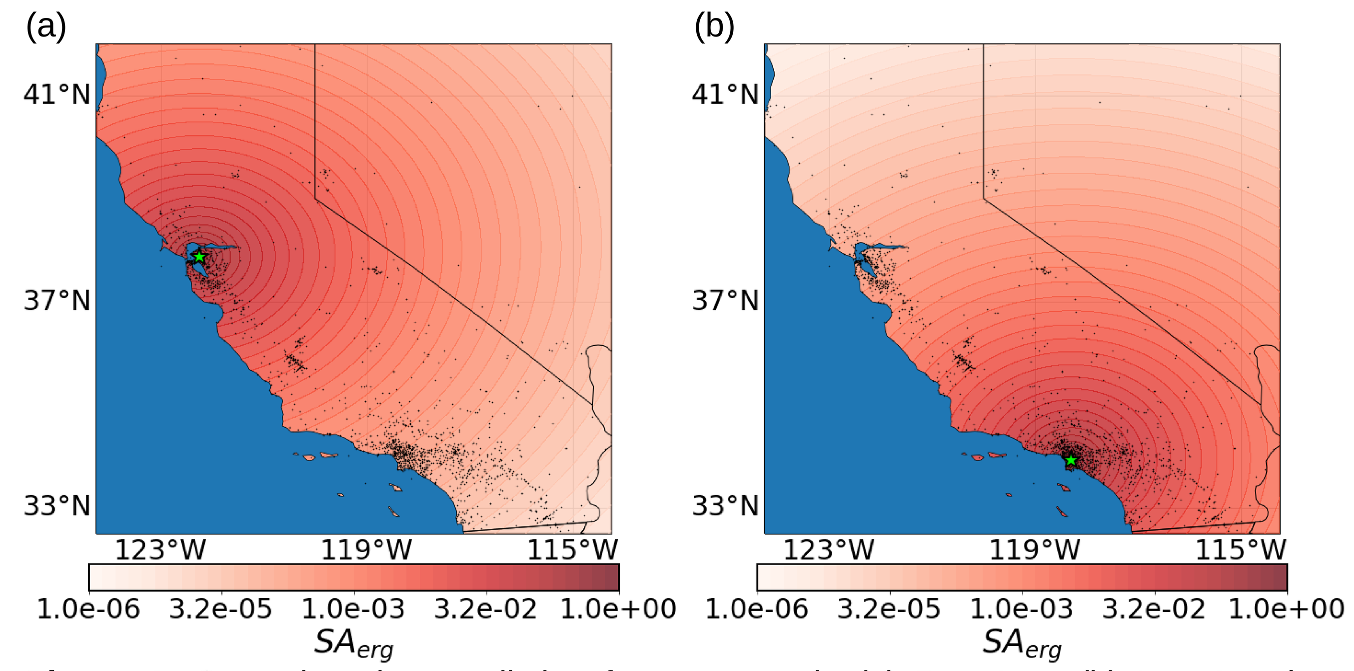

Figure 1: Ground-motion prediction for an event in (a) Bay Area (b) Los Angeles, using an ergodic GMM

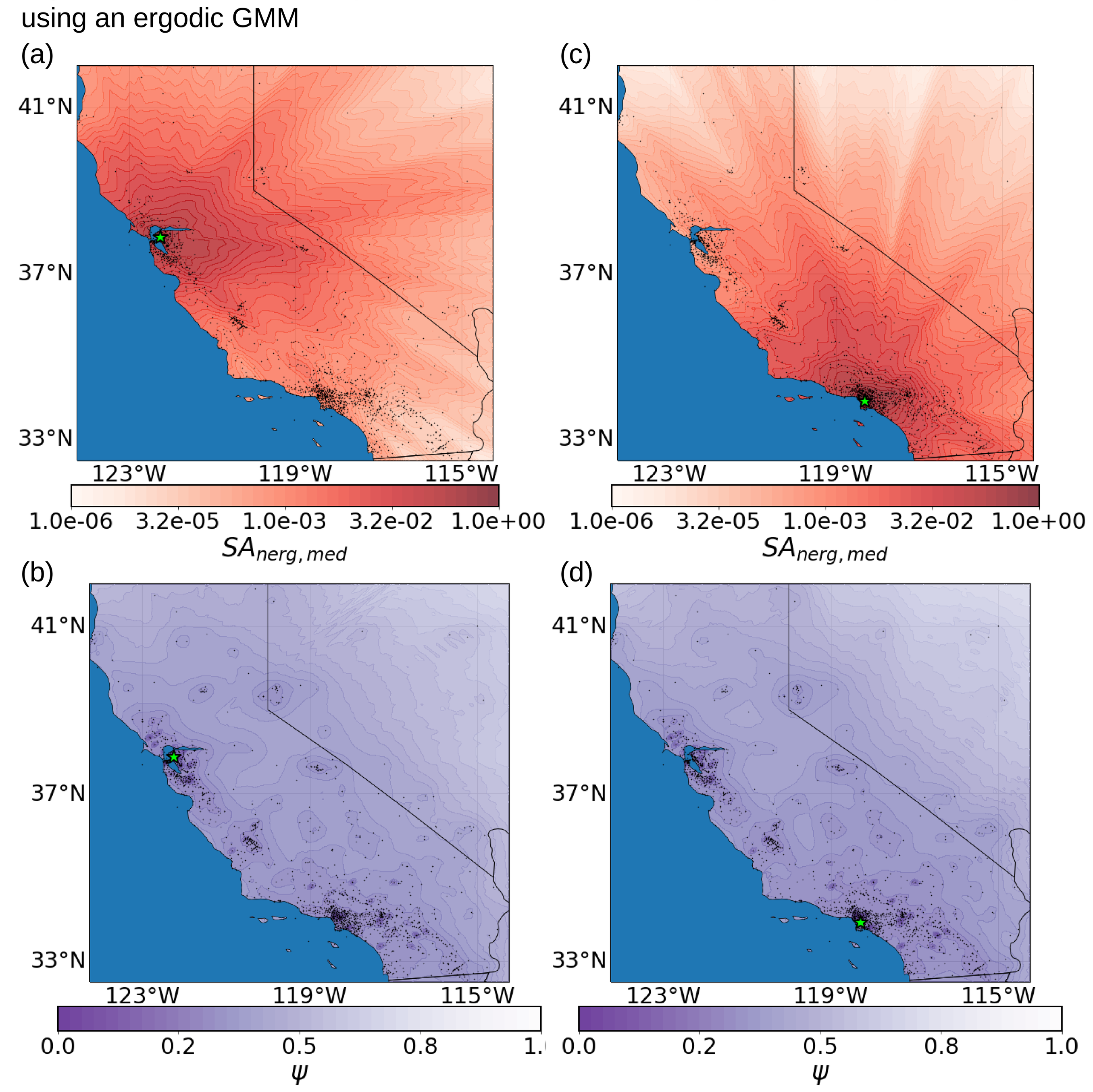

Figure 2: Ground-motion prediction for an event in Bay Area (a) median estimate (b) epistemic uncertainty, and an event in Los Angeles (c) median estimate and (d) epistemic uncertainty using a non-ergodic GMM

Model Framework

In the varying coefficient framework, the non-ergodic coefficients vary as a function of source, site coordinates, etc. The ergodic base model is developed using a large global dataset; it contains the primary scaling relationships (e.g., magnitude scaling), as well as empirical and analytical constraints (e.g., large magnitude saturation and non-linear site amplification). The non-ergodic coefficients are expressed as Gaussian processes, their correlation structure is a function of the source or site coordinates, and they are estimated based on the regional dataset.

$$f_{nerg}(M, R_{rup}, \ldots, \vec{t}_E, \vec{t}_S) = \underbrace{f_{erg}(M, R_{rup}, \ldots)}_{\text{Ergodic Base Model (e.g., large magnitude scaling)}} + f_E(\underbrace{\vec{t}_E}_{\text{Earthquake coordinates}}) + f_S(V_{S30}, \underbrace{\vec{t}_S}_{\text{Site coordinates}}) + f_P(R_{rup}, \vec{t}_E, \vec{t}_S)$$

Example modeling for systematic site effects: $f_S(V_{S30}, \vec{t}_S) = \delta c_{1a,S}(\vec{t}_S) + \delta c_{1b,S}(\vec{t}_S)$

$$\delta c_{1a,S} \sim N(0, \kappa_{1a,S}(\vec{t}_S, \vec{t}_S'))$$

The cell-specific anelastic attenuation approach is used to capture the systematic path effects.

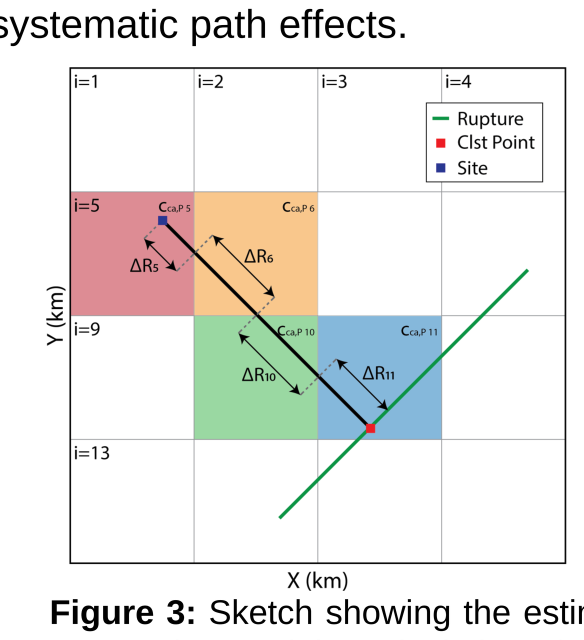

Figure 3: Sketch showing the estimation of the total anelastic attenuation, and the calculation of the cell path segments

Anelastic Attenuation:

$$f_{atten,P} = \underbrace{\Delta\vec{R}}_{\text{Cell-path segments}} \cdot \underbrace{\vec{c}_{ca,P}}_{\text{Cell attenuation coefficients}}$$

Fully non-ergodic GMM:

$$f_{nerg}(M, R_{rup}, \ldots, \vec{t}_E, \vec{t}_S) = f_{erg}(M, R_{rup}, \ldots) + f_E(\vec{t}_E) + f_S(\vec{t}_S) + \Delta\vec{R} \cdot \vec{c}_{ca,P} - \underbrace{c_{a,erg} R_{rup}}_{\text{Avoid double counting}}$$

Developed Tools

A series of Jupyter notebooks were created containing the software tools for (https://github.com/NHR3-UCLA/ngmm_tools):

- Input data preparation
- Selection of prior distributions
- NGMM regression using STAN and INLA
- GM prediction with NGMMs

A docker image is provided to help with the cross-platform compatibility of the provided tools (https://hub.docker.com/r/nhr3webhub/ngmm-tools)

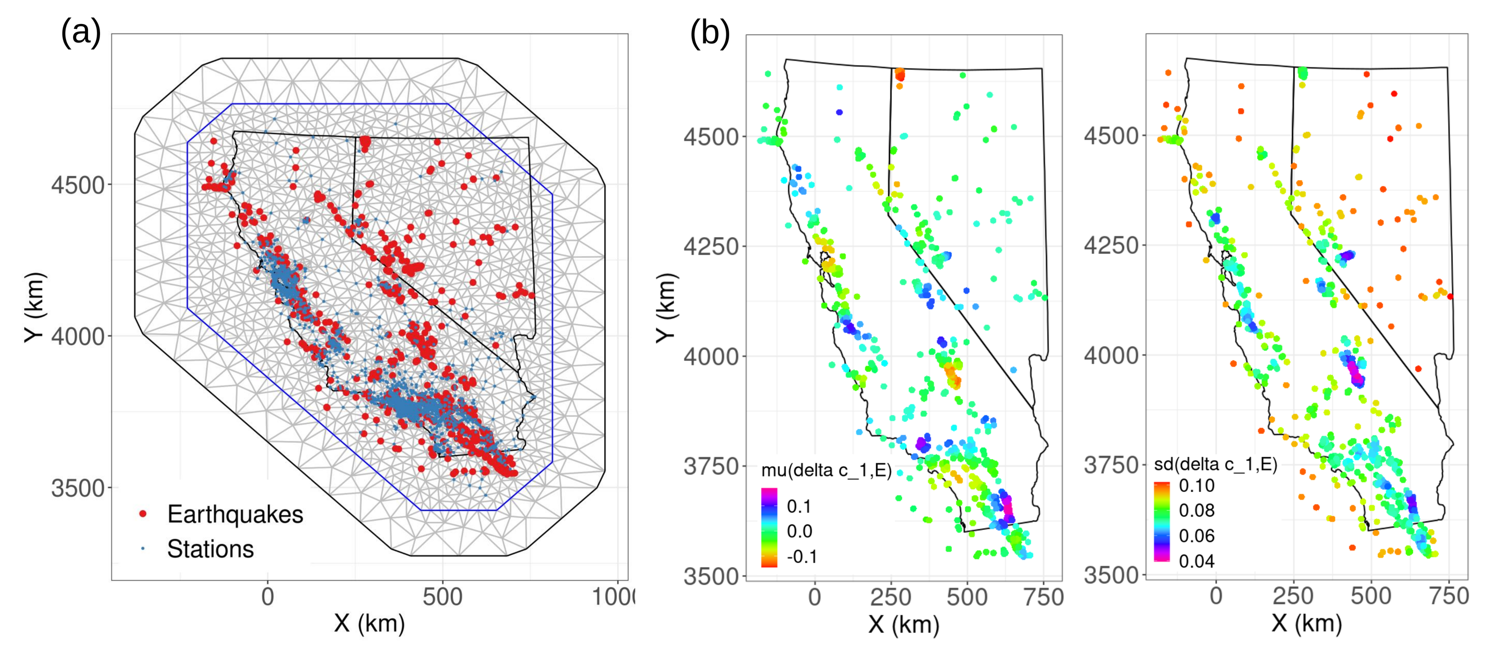

Figure 4: Example output figures of Jupyter notebook for INLA regression. (a) Mesh of spatially varying coefficients (b) Mean estimate and epistemic uncertainty of spatially varying event term using the NGAWest 3 CA synthetic dataset.

Verification Exercise

- A synthetic dataset representative of the anticipated NGA West3 CA was created for future-proofing of the developed software
- Using the synthetic datasets, the developed tools were evaluated and improved regarding their accuracy, scalability, universality, and model complexity

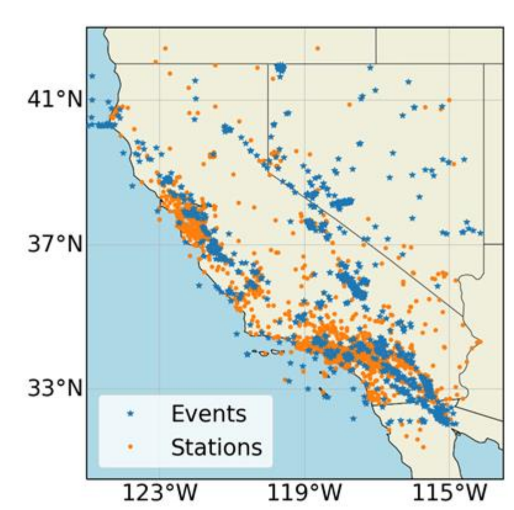

Figure 5: Event and Station distribution of NGAWest3 CA synthetic dataset.

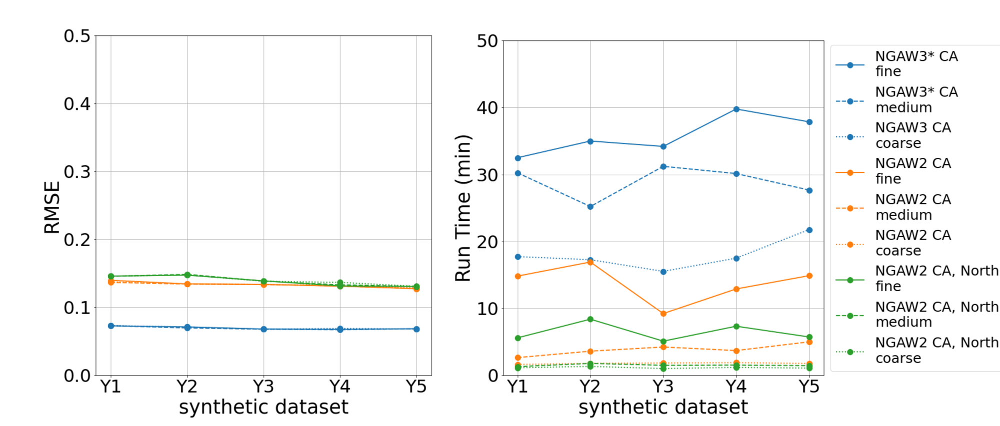

Figure 6: Goodness-of-fit and run-time comparison of INLA for different dataset sizes and mesh coarseness.

Additional Information

- Lavrentiadis G., Kuehn N., Bozorgnia Y., Seylabi E., Meng X., Goulet C., Kottke A. "Non-ergodic Methodology and Modeling Tools (Report GIRS-2022-04)." Natural Hazards Risk and Resiliency Research Center, The B. John Garrick Institute for the Risk Sciences, University of California, Los Angeles (2022)
- Lavrentiadis G., Abrahamson N.A., Nicolas K.M., Bozorgnia Y., Goulet C.A., Babič A., … Walling M. "Overview and Introduction to Development of Non-Ergodic Earthquake Ground-Motion Models." Bulletin of Earthquake Engineering (2022)
- Project webpage: https://www.risksciences.ucla.edu/nhr3/ngmm

Acknowledgements

Support for the development of this report and non-ergodic methodology and modeling tools was provided by the California Department of Transportation (Caltrans) and Pacific Gas & Electric Company (PG&E). The support of these organizations is gratefully appreciated. The opinions, findings, conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the study sponsors, the Natural Hazards Risk and Resiliency Research Center (NHR3), or the Regents of the University of California.

We also like to thank Dr. Norman A. Abrahamson for the insightful discussions during the course of this work. Numerous individuals have been working on the development and applications of non-ergodic ground motions. Their efforts and cooperation are greatly appreciated..