Quantification of Uncertainties in Seismic Ground-Motion Prediction

by

Sreeram Reddy Kotha

Kumulative Dissertation

zur Erlangung des akademischen Grades

"doctor rerum naturalium"

(Dr. rer. nat.)

in der Wissenschaftsdisziplin "Geophysik"

submitted to the

Faculty of Mathematics and Natural Sciences

at the University of Potsdam

Potsdam, 2018

Published online at the
Institutional Repository of the University of Potsdam:
URN urn:nbn:de:kobv:517-opus4-415743
http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-415743

Declaration of Authorship

I, Sreeram Reddy Kotha, hereby declare that this thesis titled, 'Quantification of Uncertainties in Seismic Ground-Motion Prediction' and the work presented in it are my own. I confirm that:

- I have fully acknowledged and referenced the ideas and work of others, whether published or unpublished, in my thesis.

- This dissertation was not submitted for the award of any other degree or diploma in any other institution.

Signed:

_____

Place, Date:

_____

Abstract

The purpose of Probabilistic Seismic Hazard Assessment (PSHA) at a construction site is to provide the engineers with a probabilistic estimate of ground-motion level that could be equaled or exceeded at least once in the structure's design lifetime. A certainty on the predicted ground-motion allows the engineers to confidently optimize structural design and mitigate the risk of extensive damage, or in worst case, a collapse. It is therefore in interest of engineering, insurance, disaster mitigation, and security of society at large, to reduce uncertainties in prediction of design ground-motion levels.

In this study, I am concerned with quantifying and reducing the prediction uncertainty of regression-based Ground-Motion Prediction Equations (GMPEs). Essentially, GMPEs are regressed best-fit formulae relating event, path, and site parameters (predictor variables) to observed ground-motion values at the site (prediction variable). GMPEs are characterized by a parametric median (μ) and a non-parametric variance (σ) of prediction. μ captures the known ground-motion physics i.e., scaling with earthquake rupture properties (event), attenuation with distance from source (region/path), and amplification due to local soil conditions (site); while σ quantifies the natural variability of data that eludes μ. In a broad sense, the GMPE prediction uncertainty is cumulative of 1) uncertainty on estimated regression coefficients (uncertainty on μ, σ_μ), and 2) the inherent natural randomness of data (σ). The extent of μ parametrization, the quantity, and quality of ground-motion data used in a regression, govern the size of its prediction uncertainty: σ_μ and σ.

In the first step, I present the impact of μ parametrization on the size of σ_μ and σ. Over-parametrization appears to increase the σ_μ, because of the large number of regression coefficients (in μ) to be estimated with insufficient data. Under-parametrization mitigates σ_μ, but the reduced explanatory strength of μ is reflected in inflated σ. For an optimally parametrized GMPE, a $\sim$10% reduction in σ is attained by discarding the low-quality data from pan-European events with incorrect parametric values (of predictor variables).

In case of regions with scarce ground-motion recordings, without under-parametrization, the only way to mitigate σ_μ is to substitute long-term earthquake data at a location with short-term samples of data across several locations – the Ergodic Assumption. However, the price of ergodic assumption is an increased σ, due to the region-to-region and site-to-site differences in ground-motion physics. σ of an ergodic GMPE developed from *generic* ergodic dataset is much larger than that of non-ergodic GMPEs developed from region- and site-specific non-ergodic subsets - which were too sparse to produce their *specific* GMPEs. Fortunately, with the dramatic increase in recorded ground-motion data at several sites across Europe and Middle-East, I could quantify the region- and site-specific differences in ground-motion scaling and upgrade the GMPEs with 1) substantially more accurate region- and site-specific μ for sites in Italy and Turkey, and 2) significantly smaller prediction variance σ. The benefit of such enhancements to GMPEs is quite evident in my comparison of PSHA estimates from ergodic versus region- and site-specific GMPEs; where the differences in predicted design ground-motion levels, at several sites in Europe and Middle-Eastern regions, are as large as $\sim$50%.

Resolving the ergodic assumption with mixed-effects regressions is feasible when the quantified region- and site-specific effects are physically meaningful, and the non-ergodic subsets (regions and sites) are defined a priori through expert knowledge. In absence of expert definitions, I demonstrate the potential of machine learning techniques in identifying efficient clusters of site-specific non-ergodic subsets, based on latent similarities in their ground-motion data. Clustered site-specific GMPEs bridge the gap between site-specific and fully ergodic GMPEs, with their partially non-ergodic μ and, σ $\sim$15% smaller than the ergodic variance.

The methodological refinements to GMPE development produced in this study are applicable to new ground-motion datasets, to further enhance certainty of ground-motion prediction and thereby, seismic

hazard assessment. Advanced statistical tools show great potential in improving the predictive capabilities of GMPEs, but the fundamental requirement remains: large quantity of high-quality ground-motion data from several sites for an extended time-period.

Zusammenfassung

Der Zweck der probabilistischen seismischen Gefährdungsbeurteilung (PSHA) auf einer Baustelle besteht darin, den Ingenieuren eine probabilistische Schätzung des Bodenbewegungspegels zu liefern, die mindestens einmal in der Entwurfslebensdauer der Struktur erreicht oder überschritten werden könnte. Eine Gewissheit über die vorhergesagte Bodenbewegung erlaubt es den Ingenieuren, das strukturelle Design sicher zu optimieren und das Risiko von weitreichenden Schäden oder im schlimmsten Fall eines Zusammenbruchs zu minimieren. Es liegt daher im Interesse des Ingenieurwesens, der Versicherung, der Katastrophenvorsorge und der Sicherheit der Gesellschaft insgesamt, die Unsicherheiten bei der Vorhersage der Bodenbewegungs des Entwurfs zu reduzieren.

In dieser Studie, beschäftige ich mich mit der Quantifizierung und Reduzierung der Vorhersageunsicherheit von Regressions-basierten Bodenbewegungsvorhersage-Gleichungen (GMPEs). Im Wesentlichen sind GMPEs am besten angepasste Formeln, die Ereignis-, Pfad- und Standortparameter (Prädiktorvariablen) auf beobachtete Bodenbewegungswerte an der Stelle (Vorhersagevariable) beziehen. GMPEs sind gekennzeichnet durch einen parametrischen Median (μ) und eine nichtparametrische Varianz (σ) der Vorhersage. μ erfasst die bekannte Bodenbewegungs-Physik, d. h. Skalierung mit Erdbebenbrucheigenschaften (Ereignis), Dämpfung mit Abstand von der Quelle (Region/Pfad) und Verstärkung aufgrund lokaler Bodenbedingungen (Standort); während σ die natürliche Variabilität von Daten quantifiziert, die sich dem μ entziehen. In einem weiten Sinne ist die GMPE-Vorhersageunsicherheit kumulativ von 1) Unsicherheit bezüglich der geschätzten Regressionskoeffizienten (Unsicherheit in μ; σ_μ) und 2) der inhärenten natürlichen Zufälligkeit von Daten (σ). Das Ausmaß der μ-Parametrisierung, die Menge und die Qualität der Bodenbewegungsdaten, die in einer Regression verwendet werden, bestimmen die Größe der Vorhersageunsicherheit: σ_μ und σ.

Im ersten Schritt stelle ich den Einfluss der μ-Parametrisierung auf die Größe von σ_μ und σ vor. Überparametrisierung scheint σ_μ zu erhöhen, da die große Anzahl von Regressionskoeffizienten (in μ) mit unzureichenden Daten geschätzt werden muss. Unterparametrisierung mindert σ_μ , aber die reduzierte Erklärungsstärke von μ spiegelt sich in erhöhtem σ wider. Für eine optimal parametrisierte GMPE wird eine $\sim$ 10% ige Verringerung von σ erreicht, indem die Daten niedriger Qualität aus paneuropäischen Ereignissen mit inkorrekten Parameterwerten (von Prädiktorvariablen) verworfen werden.

In Regionen mit wenigen Bodenbewegungsaufzeichnungen, ohne Unterparametrisierung, besteht die einzige Möglichkeit, σ_μ abzuschwächen, darin, langfristige Erdbebendaten an einem Ort durch kurzzeitige Datenproben an mehreren Orten zu ersetzen - die Ergodische Annahme. Der Preis der ergodischen Annahme ist jedoch aufgrund der Unterschiede in der Bodenbewegungsphysik von Region-zu-Region und von Ort-zu-Ort ein erhöhtes σ. σ einer ergodischen GMPE, die aus einem generischen ergodischen Datensatz entwickelt wurde, ist viel größer als die von nicht-ergodischen GMPEs, die aus regions- und standortsspezifischen nicht-ergodischen Teilmengen entwickelt wurden - die zu dünn waren, um ihre spezifischen GMPEs zu erzeugen. Durich den dramatischen Anstieg der erfassten Bodenbewegungsdaten an mehreren Standorten in Europa und im Nahen Osten konnte ich die regions- und standortspezifischen Unterschiede bei der Bodenbewegungsskalierung quantifizieren und die GMPE mit 1) wesentlich genauerer regions verbessern, und standortspezifische μ für Standorte in Italien und der Türkei verbessern, und 2) signifikant kleinere Vorhersagevarianz σ erreichen. Der Vorteil solcher Verbesserungen für GMPEs ist ziemlich offensichtlich in meinem Vergleich von PSHA-Schätzungen von ergodischen gegenüber regions- und standortsspezifischen GMPEs; wo die Unterschiede in den prognostizierten Bodenbewegungsebenen an verschiedenen Standorten in Europa und im Nahen Osten bis zu $\sim$ 50% betragen.

Die Lösung der ergodischen Annahme mit gemischten Regressionen ist machbar, wenn die quantifizierten regions- und standortsspezifischen Effekte physikalisch sinnvoll sind und die nicht-ergodischen Teilmengen (Regionen und Standorte) a priori durch Expertenwissen definiert werden. In Ermangelung von Expertendefinitionen demonstriere ich das Potential von maschinellen Lerntechniken bei der Identifizierung effizienter Cluster von standortsspezifischen nicht-ergodischen Untergruppen, basierend auf latenten Ähnlichkeiten in ihren Bodenbewegungsdaten. Geclusterte standortsspezifischen GMPEs überbrücken die Lücke zwischen standortsspezifischen und vollständig ergodischen GMPEs mit ihrem teilweise nicht-ergodischen μ und $\sim$15% kleiner als die ergodische Varianz.

Die methodischen Verbesserungen der GMPE-Entwicklung, die in dieser Studie heransgearbeitet wurden, sind auf neue Bodenbewegungsdatensätze anwendbar, um die Sicherheit der Bodenbewegungsvorhersage und damit die Bewertung der seismischen Gefährdung weiter zu verbessern. Fortgeschrittene statistische Werkzeuge zeigen ein großes Potenzial bei der Verbesserung der Vorhersagefähigkeiten von GMPEs, aber die grundlegende Anforderung bleibt: eine große Menge an hochwertigen Bodenbewegungsdaten von mehreren Standorten für einen längeren Zeitraum.

Tribute

गुरुर्ब्रह्मा गुरुर्विष्णु र्गुरुर्देवो महेश्वरः

गुरु साक्षात परब्रह्मा तस्मै श्रीगुरवे नमः

Guru Brahma, Guru Vishnu, Guru devo Maheshwara

Guru sakshat, param Brahma, tasmai shri guravay namah

Our creation is that guru (Brahma-the force of creation); the duration of our lives is that guru (Vishnu-the force of preservation); our trials, tribulations, illnesses, calamities and the death of the body is that guru (devo Maheshwara-the force of destruction or transformation). There is a guru nearby (Guru Sakshat) and a guru that is beyond the beyond (param Brahma). I make my offering (tasmai) to the beautiful (shri) remover of my darkness, my ignorance; (Guru) it is to you I bow and lay down my life (namah).

When the student is ready, the teacher appears. I, Sreeram, would like to express my deepest gratitude to my teachers for transforming me. I owe my parents and family at home for culturing a seeker in me, and nurturing me through troubled times. I owe my friends for believing in my capacity. Your love and faith has been my guiding light. I thank you.

I thank my supervisor, Prof. Dr. Fabrice Cotton, for choosing me as his Masters and PhD student, and guiding me through the transformation from an enthusiastic young boy to a cultured scientific mind. Fabrice taught me to identify the key scientific questions, device elegant solutions, and prepare *sexy* figures for a captivating presentation. I owe my academic recognition to Fabrice's leadership, which put me under the mentorship of another brilliant scientist, Dr. Dino Bindi.

Working with Dino has accelerated my research several fold. For every question we had, we formulated a dozen new questions of scientific significance, and with Dino it was never impossible to achieve half of them within a month. I owe my adventurous trait to Dr. Dino Bindi, who has taught me to be a man-of-action, whilst holding a cold beer.

Speaking of scientific brilliance, I would like to thank my co-supervisor, Prof. Dr. Paolo Bazzurro, for teaching me a valuable lesson very early in my scientific career. Paolo once said to me, 'Sreeram, there are many brilliant young scientists who disappeared with a flash – you don't have to be that'. Paolo's sheer intelligence and humility made him my role-model of great mentorship. I thank him for that.

A successful pursuit is a matter of perseverance, discipline, and of course, great colleagues. Working at Helmholtzstr. 6/7 of Potsdam, has been a delight, thanks to the many cheerful friends. A special thanks to Ms. Solveig Strutzke for being a stern, yet extremely lovable facilitator throughout my PhD - in organizing my meetings on to preparation of this dissertation. An applause to Mr. Uwe Lemgo, for cheerfully solving the worst of my computer issues. I wouldn't miss my work a single day and not have a lunch with my dear friends: Nima, Nasim, Djamil, Toño, T.B., Yenshin, Malte, Hoby, Michael, Bojana, Carla, Camilo, Carlos, Luigi, Francesco, Tim, Jeane, Ayleen, Mehdi, Camilla, Sanjay, and many others – who I love very much despite forgetting my *PhD hat*. My best wishes to them. A very special place for Dr. Eleonora Rivalta, who has thinned the boundaries between a great mentor and a dear friend, inspiring me with her romantic view of the massive geophysical phenomena. I appreciate Mr. Christian Bosse, Ms. Steffi Lammers, Dr. Arno Zang, Dr. Graeme Weatherill, Dr. Danijel Schorlemmer, Dr. Dietrich Stromeyer, Dr. Jeoung Sook Yoon, Dr. Sum Mak, Dr. Hoby Razafindrakoto for being such a pleasant and intellectually stimulating company at work.

Last and perhaps the most important, I would like to thank my mother, father and sister for loving and encouraging me, and being so patient despite being so far away from here. And certainly, my epitome of intelligence and affection, Dr. Cristina Voiticovschi-Iosob, who has been my home-away-from-home since the beginning, inspiring and nurturing me to be better than myself, every day, all day. I thank you for making me.

You are my *Gurus*. This dissertation is my tribute to your love. My deepest thanks to you all – *Pranaam*.

Contribution to Publications

Chapter 2:
Dino Bindi, Fabrice Cotton, ***Sreeram Reddy Kotha****, Christian Bosse, Dietrich Stromeyer, and Gottfried Grünthal: Application-driven ground motion prediction equation for seismic hazard assessments in non-cratonic moderate-seismicity areas. Journal of Seismology (2017)*
- I contributed partially to designing the functional form of Ground-Motion Prediction Equation (GMPE), and completely, to the estimation of its asymptotic variances. I also produced the figures comparing the asymptotic variance of the GMPEs against the existing NGA-West2 models.

Chapter 3:
Sreeram Reddy Kotha*, Dino Bindi, Fabrice Cotton: Partially non-ergodic region specific GMPE for Europe and Middle-East. Bulletin of Earthquake Engineering (2016)*
- My co-authors helped in identifying the key needs of next-generation of ground-motion models for use in non-ergodic region-specific probabilistic seismic hazard assessment. I proposed the mixed-effects regression approach to regionalize the GMPEs, designed the study, scripted the regression procedure, developed the model, produced the figures, wrote 90% of the manuscript, and served as the corresponding author of publication. My co-authors contributed to interpretation and presentation of the results, and revision of the manuscript.

Chapter 4:
Sreeram Reddy Kotha*, Dino Bindi, and Fabrice Cotton: From ergodic to region- and site-specific Probabilistic Seismic Hazard Assessment: Method development and application at European and Middle Eastern sites. Earthquake Spectra (2017)*
- In process of developing a region-specific GMPE (above), I identified a few limitations in the existing site-specific probabilistic seismic hazard assessment procedure. Based on which, I proposed refinements to existing approach, designed the study, performed the hazard analysis, produced the figures, wrote 90% of the manuscript, and served as the corresponding author of publication. My co-authors contributed to interpretation and presentation of the results, and revision of the manuscript.

Chapter 5:
Sreeram Reddy Kotha*, Fabrice Cotton, and Dino Bindi: A New Approach to Site Classification: Mixed-effects Ground Motion Prediction Equation with Spectral Clustering of Site Amplification Functions. Soil Dynamics and Earthquake Engineering (2018)*
- My co-authors introduced me to the need for a new site-classification in building design code revision, and educated me on the state-of-practice. I proposed a machine learning approach to site-classification, performed the regression and machine learning procedures, produced the figures, wrote 80% of the manuscript, and served as the corresponding author of publication. My co-authors contributed to interpretation and presentation of the results.

Chapter 6:
Sreeram Reddy Kotha*, Dino Bindi, and Fabrice Cotton: Site-corrected magnitude and region dependent correlations of horizontal peak spectral amplitudes. Earthquake Spectra (2017)*
- During my interaction with engineers, I recognized an unaddressed need for new empirical correlation models of ground-motion amplitudes at different frequencies. Subsequently, I identified scope for improvements in existing models, proposed solutions to meet the needs, designed the study, scripted the algorithms, produced the figures, wrote 90% of the manuscript, and served as the corresponding author of the publication. My co-authors contributed to interpretation and presentation of the results.

List of Figures

FIGURE 1-1: WORKFLOW OF PSHA 1

FIGURE 1-2: IMPACT OF GMM ON HAZARD CURVES AT A SITE 2

FIGURE 1-3: ATTENUATION OF PEAK GROUND ACCELERATION (PGA) WITH DISTANCE (LEFT) AND MAGNITUDE (RIGHT) 3

FIGURE 1-4: REGIONAL DIVERSITY OF STRONG GROUND-MOTION DATA IN NGA-WEST2 DATASET 6

FIGURE 2-1: DISAGGREGATION OVER MAGNITUDE AND DISTANCE SCENARIOS FOR PRELIMINARY HAZARD ASSESSMENT AT A REPRESENTATIVE SITE IN GERMANY 15

FIGURE 2-2: DISTRIBUTION OF MAGNITUDE VERSUS HYPOCENTRAL DISTANCE FOR THE CONSIDERED RECORDINGS 16

FIGURE 2-3: COEFFICIENTS OF THE GMPE CALIBRATED FOR THE JOYNER-BOORE DISTANCE 17

FIGURE 2-4: TOP: PERIOD DEPENDENCE OF THE MAGNITUDE TERMS 18

FIGURE 2-5: VARIANCE, CORRELATION AND GRADIENT OF COEFFICIENTS 19

FIGURE 2-6: UNCERTAINTY IN THE MEAN COMPUTED FOR THE JOYNER-BOORE MODEL 20

FIGURE 2-7: WITHIN-EVENT (Φ), BETWEEN-EVENT (τ), AND TOTAL (Σ) STANDARD DEVIATIONS VERSUS PERIODS FOR THE JOYNER-BOORE MODEL 21

FIGURE 2-8: PREDICTION VERSUS DATA FOR PGA (A) AND PGV (B), FOR THE HYPOCENTRAL DISTANCE MODEL 22

FIGURE 2-9: WITHIN-EVENT RESIDUALS VERSUS DISTANCE AND MAGNITUDE 22

FIGURE 2-10: GEOMETRIES USED TO GENERATE THE SCENARIOS FOR THE SAMMON'S MAP 23

FIGURE 2-11: SAMMON'S MAPS FOR $\boldsymbol{T}$ = 0.02, 0.1, 1, AND 4S 24

FIGURE 2-12: MAGNITUDE SCALING FOR DIFFERENT GMPES A $\boldsymbol{T}$ = 0.02, 0.1, 1.0, AND 4.0S 25

FIGURE 2-13: DISTANCE SCALING FOR DIFFERENT GMPES A $\boldsymbol{T}$ = 0.02, 0.1, 1.0, AND 4.0 S 26

FIGURE 2-14: PERIOD DEPENDENCE OF THE DISTANCE IN THE SAMMON'S PLANE 27

FIGURE 3-1: SCATTER PLOT SHOWING THE DISTRIBUTION OF OBSERVED DATA IN MAGNITUDE-DISTANCE RANGES FOR DIFFERENT EC8 SITE CLASSES FOR EACH REGION ITALY (IT), OTHERS, AND TURKEY (TR) 31

FIGURE 3-2: BETWEEN STATION RESIDUALS (PGA IN LEFT PANEL, SA (2S) IN RIGHT PANEL) PLOTTED AGAINST $\boldsymbol{V_{S30}}$ (M/S) WITH STATIONS SEPARATED INTO REGIONS 33

FIGURE 3-3: $\Delta \boldsymbol{c_3}$ FOR THE THREE REGIONS ACROSS DIFFERENT SPECTRAL TIME PERIODS 34

FIGURE 3-4: RANDOM EFFECTS ON $\boldsymbol{g_1}$ AND $\boldsymbol{g_2}$ ALONG WITH THEIR STANDARD ERRORS 35

FIGURE 3-5: DISTANCE SCALING FOR PGA AND SA (2S) 36

FIGURE 3-6: MAGNITUDE SCALING FOR PGA AND SA (2S) AT SITE WITH $\boldsymbol{V_{S30}}$ = 450M/S 37

FIGURE 3-7: RESPONSE SPECTRA SHOWING THE CUMULATIVE EFFECT OF REGIONAL ADJUSTMENTS TO THE GMPE 38

FIGURE 3-8: COMPARISON OF INDIVIDUAL COMPONENTS OF ALEATORY VARIABILITY IN GMPE 39

FIGURE 3-9: REGIONAL VARIATION OF BETWEEN EVENT RESIDUALS AT SA (1S) 40

FIGURE 4-1: COMPARISON OF GMPES 47

FIGURE 4-2: COMPARISON OF $\boldsymbol{\delta S2S}$ FROM R13 AND NLME 48

FIGURE 4-3: DISTRIBUTION OF $\boldsymbol{\phi_{ss,s}}$ FOR THE STATIONS WITH AT LEAST TWO RECORDS IN RESORCE DATASET 49

FIGURE 4-4: COMPARISON OF $\delta S2S_s$ OBTAINED BY APPLYING THE NLME AND R13 METHODS TO TWO WELL-RECORDED SITES IN RESORCE 50
FIGURE 4-5: COMPARISON OF GMPES USING DATA UNTIL 2009 AND 2012 52
FIGURE 4-6: COMPARISON OF MEAN SITE-TERMS ESTIMATED FOR THE NEW SITE USING R13 AND NLME METHODS 53
FIGURE 4-7: FLOWCHART FOR REGION-SPECIFIC PSHA AND REGION- AND SITE-SPECIFIC PSHA 53
FIGURE 4-8: DISTRIBUTION OF SITES IN RESORCE DATASET 55
FIGURE 4-9: COMPARISON OF ERGODIC HAZARD CURVES WITH PARTIALLY ERGODIC HAZARD CURVES 56
FIGURE 4-10: PERCENT CHANGE IN ESTIMATED GROUND-MOTION VALUES AT EACH RETURN PERIOD, SHIFTING FROM ERGODIC TO REGION-SPECIFIC PSHA 57
FIGURE 4-11: PERCENT CHANGE IN ESTIMATED GROUND-MOTION VALUES AT EACH RETURN PERIOD SHIFTING FROM ERGODIC TO REGION- AND SITE-SPECIFIC PSHA 57
FIGURE 5-1: DATA DISTRIBUTION FOLLOWING THE RECORD SELECTION CRITERIA FOR GMPE REGRESSION AT $\boldsymbol{T = 0.01}$S 62
FIGURE 5-2: DISTANCE SCALING OF GEOMETRIC MEAN OF 5% DAMPED HORIZONTAL PSEUDO SPECTRAL ACCELERATIONS AT $\boldsymbol{T}$ = 0.02S, 0.2S AND 2S. 63
FIGURE 5-3: PARAMETRIC ANALYSIS FOR MAGNITUDE SCALING AT $\boldsymbol{T}$ = 0.02S, 0.2S, 2S 64
FIGURE 5-4: RANDOM-EFFECTS AND RESIDUAL PLOTS FOR GMPE EVALUATION AT $\boldsymbol{T}$ = 0.02S, 0.2S, AND 2S 65
FIGURE 5-5: DISTANCE SCALING OF THE GMPE FIXED-EFFECTS AT $\boldsymbol{T}$ = 0.02S, 0.2S, AND 2S 66
FIGURE 5-6: MAGNITUDE SCALING OF THE GMPE FIXED-EFFECTS AT $\boldsymbol{T}$ = 0.02S, 0.2S, AND 2S 67
FIGURE 5-7: MEDIAN RESPONSE SPECTRA, BETWEEN-EVENT ($\boldsymbol{\tau}$), BETWEEN-SITE ($\boldsymbol{\phi_{S2S}}$), EVENT-AND-SITE CORRECTED STANDARD DEVIATIONS ($\boldsymbol{\phi_0}$) AND THE TOTAL ALEATORY VARIABILITY ($\boldsymbol{\sigma}$) OF THE GMPE IN NATURAL-LOG SCALE. 67
FIGURE 5-8: OPTIMAL NUMBER OF CLUSTERS BASED ON TOTAL WITHIN SUM OF SQUARES (WSS) AND GAP STATISTIC (GS) 68
FIGURE 5-9: 2D VISUALIZATION OF K-MEAN CLUSTERS AND THE K-MEANS AMPLIFICATION FUNCTIONS 70
FIGURE 5-10: SITE AMPLIFICATION FUNCTIONS AND WITHIN-CLUSTER SITE-TO-SITE VARIABILITY 71
FIGURE 5-11: EVALUATION OF SITE RESPONSE PROXIES IN CHARACTERIZING SITE CLUSTERS 73
FIGURE 6-1: REDUCTION OF WITHIN-EVENT STANDARD DEVIATION ($\boldsymbol{\phi}$) TO EVENT-AND-SITE CORRECTED STANDARD DEVIATION ($\boldsymbol{\phi ss}$) BY REMOVING BETWEEN-SITE VARIABILITY ($\boldsymbol{\phi_{S2S}}$) 81
FIGURE 6-2: COMPARISON BETWEEN TOTAL-RESIDUAL CORRELATION MODELS 82
FIGURE 6-3: BETWEEN-EVENT CORRELATIONS FOR $\boldsymbol{M < 5.5}$ EVENTS AND $\boldsymbol{M \geq 5.5}$ 83
FIGURE 6-4: DISTANCE DEPENDENCE OF WITHIN-EVENT RESIDUALS AND EVENT-AND-SITE CORRECTED RESIDUALS FOR TWO SPECTRAL PERIODS (0.02S, 2S) 84
FIGURE 6-5: COMPARISON OF WITHIN-EVENT AND EVENT-AND-SITE CORRECTED RESIDUAL CORRELATIONS AT NEAR SOURCE RECORDS 86
FIGURE 6-6: REGION DEPENDENCE OF EVENT-AND-SITE CORRECTED RESIDUAL CORRELATIONS IN NEAR SOURCE DISTANCE RANGE 86
FIGURE 6-7: ADJUSTMENT FACTOR ($\boldsymbol{10^{\delta S2S}}$) FOR FOUR SITES WITH $\boldsymbol{V_{s30} = 450}$M/S AND 10 RECORDS IN RESORCE DATASET 87
FIGURE 6-8: COMPARISON OF ERGODIC CMS AND PARTIALLY NON-ERGODIC CMS 88

List of Tables

TABLE 3-1: COEFFICIENT TABLE FOR GMPE 42

TABLE 3-2: COEFFICIENT TABLE FOR V_{s30} BASED SITE RESPONSE 42

TABLE 5-1: NUMBER OF STATIONS WITHIN EACH CLUSTER AND WITHIN-CLUSTER SUM OF SQUARES (WCSS) 69

TABLE 5-2: SITE CLASSIFICATION BASED ON ZHAO ET AL. (2006). 72

TABLE 5-3: SITE CLUSTER CHARACTERIZATION BASED ON V_{S10}-V_{s30}-H_{800} RANGES 74

Contents

ABSTRACT I

ZUSAMMENFASSUNG III

TRIBUTE V

CONTRIBUTION TO PUBLICATIONS VI

LIST OF FIGURES VII

LIST OF TABLES IX

1 INTRODUCTION 1

1.1 PURPOSE OF THE STUDY 2

1.1.1 Development of GMPEs 3

1.1.2 Ergodic assumptions in GMPEs 4

1.2 APPROACH OF THE STUDY 5

1.2.1 Ergodic hypotheses in GMPEs 5

1.2.2 Quantification of non-ergodic biases 7

1.2.3 Partially non-ergodic GMPEs 7

1.2.4 Partially non-ergodic Probabilistic Seismic Hazard Assessment 9

1.3 PRESENTATION OF THE STUDY 9

2 APPLICATION-DRIVEN GROUND-MOTION PREDICTION EQUATION FOR SEISMIC HAZARD ASSESSMENTS IN NON-CRATONIC MODERATE-SEISMICITY AREAS 12

2.1 INTRODUCTION 13

2.1.1 Towards an application driven GMPE: Hazard assessment in a moderate seismicity area (Germany) and associated needs 14

2.1.2 NGA-west2 data and GMPE development 15

2.2 RESULTS 17

2.2.1 Fixed effects (median model) 17

2.2.2 Epistemic variability in the median 18

2.2.3 Aleatory variability 20

2.3 DISCUSSIONS 21

2.3.1 Analysis of residuals 21

2.3.2 Comparison with NGA2 and RES14 models 23

2.4 CONCLUSIONS 25

3 PARTIALLY NON-ERGODIC REGION-SPECIFIC GMPE FOR EUROPE AND MIDDLE-EAST 28

3.1 INTRODUCTION 29

3.2 DATASET AND SELECTION CRITERIA 30

3.3 REGRESSION APPROACH 31

3.3.1 Regional variability in apparent anelastic attenuation term 32

3.3.2 Style of faulting terms *32*
3.3.3 Regional variability in site-response as a function of $\boldsymbol{V_{S30}}$ *33*
3.4 RESULTS 34
3.5 DISCUSSION 35
3.5.1 Region dependent distance scaling and V_{S30} *based site response* *35*
3.5.2 Impact of the regionalization on the median predictions *36*
3.5.3 Impact of the regionalization on the model uncertainty *38*
3.5.4 Potential regional differences in magnitude scale *39*
3.6 CONCLUSIONS 40
4 FROM ERGODIC TO REGION- AND SITE-SPECIFIC PROBABILISTIC SEISMIC HAZARD ASSESSMENT: METHOD DEVELOPMENT AND APPLICATION AT EUROPEAN AND MIDDLE EASTERN SITES 43
4.1 INTRODUCTION 44
4.2 DATASET AND ERGODIC GROUND-MOTION PREDICTION EQUATIONS 45
4.3 PARTIALLY NON-ERGODIC SITE-SPECIFIC GMPES 45
4.3.1 Current practice: R13 Approach *45*
4.3.2 Motivation to revise R13 ..... *46*
4.3.3 Nonlinear Mixed Effects Approach *47*
4.4 ESTIMATION OF THE SITE TERM FOR NEW SITES 50
4.5 DEVELOPMENT OF A REGION- AND SITE-SPECIFIC PSHA FRAMEWORK 53
4.5.1 Partially ergodic region- and site-specific GMPE *54*
4.5.2 Source model *54*
4.5.3 Target sites *54*
4.6 COMPARISON OF ERGODIC AND PARTIALLY NON-ERGODIC PSHA 55
4.6.1 Region-specific PSHA versus Ergodic PSHA *56*
4.6.2 Region- and site-specific PSHA versus Ergodic PSHA *57*
4.7 CONCLUSIONS 57
5 A NEW APPROACH TO SITE CLASSIFICATION: MIXED-EFFECTS GROUND-MOTION PREDICTION EQUATION WITH SPECTRAL CLUSTERING OF SITE AMPLIFICATION FUNCTIONS 59
5.1 INTRODUCTION 60
5.2 DATA 61
5.3 GROUND-MOTION PREDICTION EQUATION 62
5.3.1 Parametric regression *63*
5.4 RESULTS 66
5.5 SPECTRAL CLUSTERING ANALYSIS 68
5.6 SITE CLASSIFICATION FROM CLUSTER ANALYSIS 69
5.7 SITE CLUSTERS 69

5.7.1 Site amplification functions: Mean and variability *70*
5.7.2 Site-response proxies *71*
5.8 DISCUSSION AND CONCLUSIONS 74
6 SITE-CORRECTED MAGNITUDE AND REGION DEPENDENT CORRELATIONS OF HORIZONTAL PEAK SPECTRAL AMPLITUDES 77
6.1 INTRODUCTION 78
6.2 A NEW METHOD TO COMPUTE RESIDUAL CORRELATIONS TAKING INTO ACCOUNT REPEATABLE SITE EFFECTS 78
6.3 DATASETS 80
6.3.1 NGA-West2 dataset *80*
6.3.2 RESORCE dataset *80*
6.4 DATASET DEPENDENCE OF TOTAL-RESIDUAL CORRELATIONS 81
6.5 MAGNITUDE DEPENDENCE OF BETWEEN-EVENT CORRELATIONS 82
6.6 WITHIN-EVENT AND EVENT-AND-SITE CORRECTED RESIDUALS 84
6.7 REGIONAL DEPENDENCE OF EVENT-AND-SITE CORRECTED RESIDUALS 85
6.8 APPLICATION: PARTIALLY NON-ERGODIC CONDITIONAL SPECTRA 86
6.9 CONCLUSIONS 87
7 CONCLUSIONS AND OUTLOOK 89
7.1 CONCLUSIONS 89
7.2 OUTLOOK 91
8 REFERENCES 92

1
Introduction

Probabilistic Seismic Hazard Assessment (PSHA) is a procedure to estimate the annual frequency of exceeding (AFE) a ground-motion level at a site (Y_{ref}), provided models describing nearby seismicity, and attenuation of ground-motion with distance from the earthquake rupture. The first component constituting a Seismic Hazard Model, referred to as Seismic Source Models, are usually derived from the seismicity in a region. For example, in the frame of the European Commission project "Seismic Hazard Harmonization in Europe" (SHARE), a homogeneous, European parametric earthquake catalogue is compiled into SHARE European earthquake catalogues in the period 1000-2006 by *Stucchi et al.*, 2013. Using the seismological, geological, tectonic and geodetic information, the parametric catalog is used to construct seismic source models for use in European seismic hazard model, which describe the frequency, size, location, and other rupture parameters. Figure 1-1 is an illustration of PSHA procedure, where the seismic source model is shown in the left-most panel, as a collection of seismic sources capable of producing earthquakes. A sketch of recurrence model of ruptures is shown in the second panel of Figure 1-1, associated with the source seismic source model in the first panel.

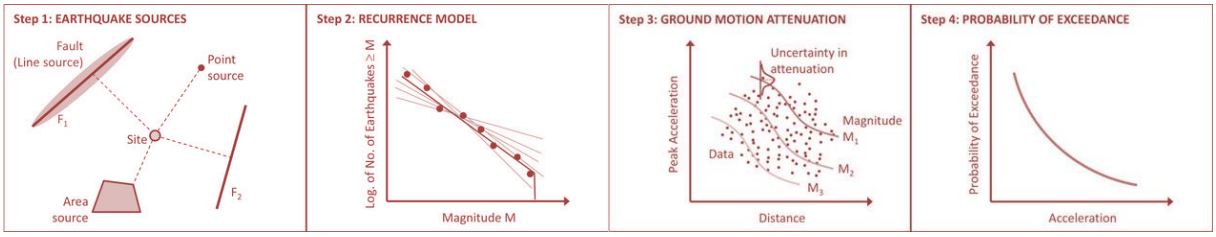


Figure 1-1: Workflow of PSHA

The second component of a seismic hazard model are the Ground-motion Models (GMM), describing the scaling of ground-motion intensity with rupture parameters such as magnitude (M), source-to-site distance (R), and local soil conditions (S). For example, *Woessner et al.* (2015) selected 14 GMMs for the European hazard model, to characterize the expected ground-motions for all magnitude, depth and distance ranges, and for all geological conditions in Europe. The center-right panel of Figure 1-1 shows the scatter-plot of recorded ground-motion data (points) from which the GMMs (lines) are derived. Typical GMMs characterize the expected ground-motions not as a unique value, but a Gaussian distribution conditioned on the source and site parameters (M, R and S) – i.e. a median ground-motion (μ) and variability (σ) for each combination (M, R, S). Given a seismic source model and GMM, the hazard integral (Eq. 1) cumulates the annual frequency of exceedance, AFE of Y_{ref} at a site, from all prospective ruptures defined in the seismic source model. $1 - P(y < Y_{ref}|M,R,S)$ component of eq. (1) estimates the probability of exceeding a reference ground-motion level as described the GMM $\sim \mathcal{N}(\mu, \sigma^2)$. A more elaborate formulation of eq. (1) contains $1 - P(y < Y_{ref}|M,R,S,\sigma)$ as an integrand, with an additional $d\sigma$ differential term. AFE calculated for a suite of Y_{ref} at a site represent the hazard estimate as an exceedance probability curve, commonly referred to as the Hazard Curve. The right-most panel of Figure 1-1 shows the hazard curve for the site in the left panel of Figure 1-1.

$$AFE(y_S \geq Y_{ref}) = \nu \iint_{M,R} P(M)P(R)\left(1 - P(y < Y_{ref}|M,R,S)\right) dMdR \qquad (1)$$

Empirically developed parametric seismic source models and GMMs contain uncertainties in choice of parametrization and assigned parametric values. Eq. (1) integrates the model-specific uncertainties into total uncertainty in hazard estimates. For example, Figure 1-2 illustrates the impact of GMM uncertainties on the hazard curves: AFE of a suite of ground-motion levels at a hypothetical site. An incorrect GMM

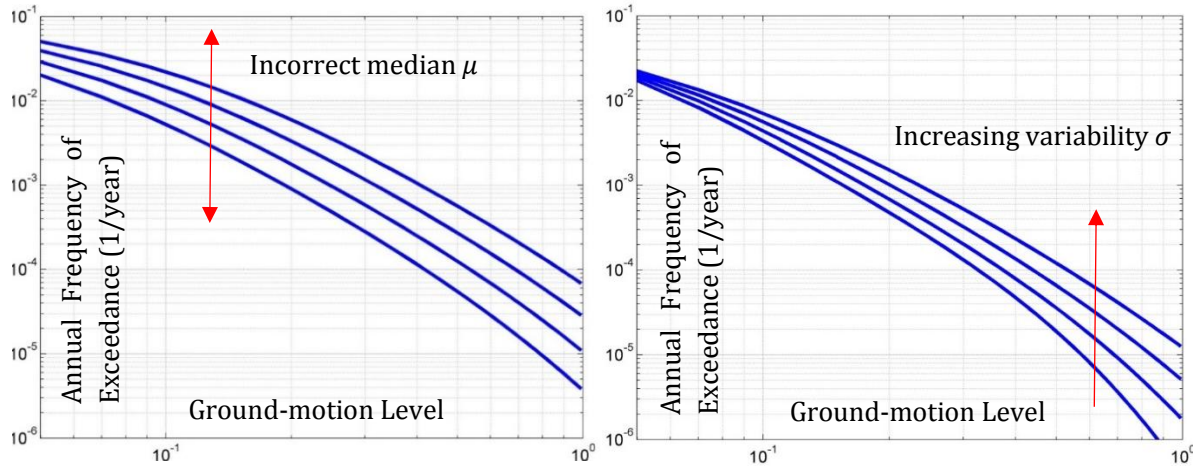


Figure 1-2: Impact of GMM on hazard curves at a site. (left panel) An incorrect median μ prediction effects the hazard estimates at all return periods, (right panel) A larger variability σ increases the probability of large ground-motions.

median μ leads to an over/under-estimation of hazard at a site, while a large variability σ (with μ unchanged) results in higher frequency of exceeding of large ground-motion levels (e.g. *Bommer and Abrahamson*, 2006). In combination, a biased μ and large σ dramatically over-estimate hazard at a site, which leads to over-conservative design of structures at the site. Therefore, it is of key engineering interest to curtail GMM uncertainties.

1.1 Purpose of the study

In this study, I investigate the uncertainties related to GMMs, through a review of the fundamental assumptions made in their development, and the resulting impact on PSHA. A prevalent approach to developing parametric GMMs is to derive predictor functions explaining the observed attenuation of ground-motion from past earthquakes. Such regression based GMMs are called the Ground-Motion Prediction Equations (GMPEs). GMPEs relate the observed ground-motion (e.g. Peak Ground Acceleration, PGA) to the earthquake rupture size (e.g. Moment Magnitude, M_w), source-to-site distance (Joyner-Boore distance, R_{JB}) and site local geological conditions (e.g. time averaged shear-wave velocity, V_S). GMPEs characterize the observed ground-motion (Y) in terms of a median prediction (μ) and residuals ($\varepsilon = \ln(Y) - \ln(\mu)$). Since the observed ground-motions are assumed to be lognormally distributed, the regression residuals ε follow a normal distribution with zero-mean and variance σ^2 – therefore, $\ln(Y) = \mathcal{N}(\ln(\mu), \sigma^2)$.

Figure 1-3 shows a typical GMPE prediction plotted against the observed ground-motion intensity (here, PGA) scaling with distance (R_{JB}) and magnitude (M_w). The markers show the 'scatter' of observed PGA values (Y). The solid black lines represent the median (μ) of the GMPE, while the density distribution silhouettes illustrate normal variability (σ). The grey ribbon illustrates the GMPE regression uncertainty (σ_μ) due to insufficient data. Natural variability of observations, efficiency of parametrization, and regression method regulate the size of GMPE standard deviation σ and median uncertainty σ_μ. Since μ, σ and σ_μ are critical parameters in estimating the AFE of Y_{ref} at a site, the focus of this study is on improving the accuracy and precision of a GMPE.

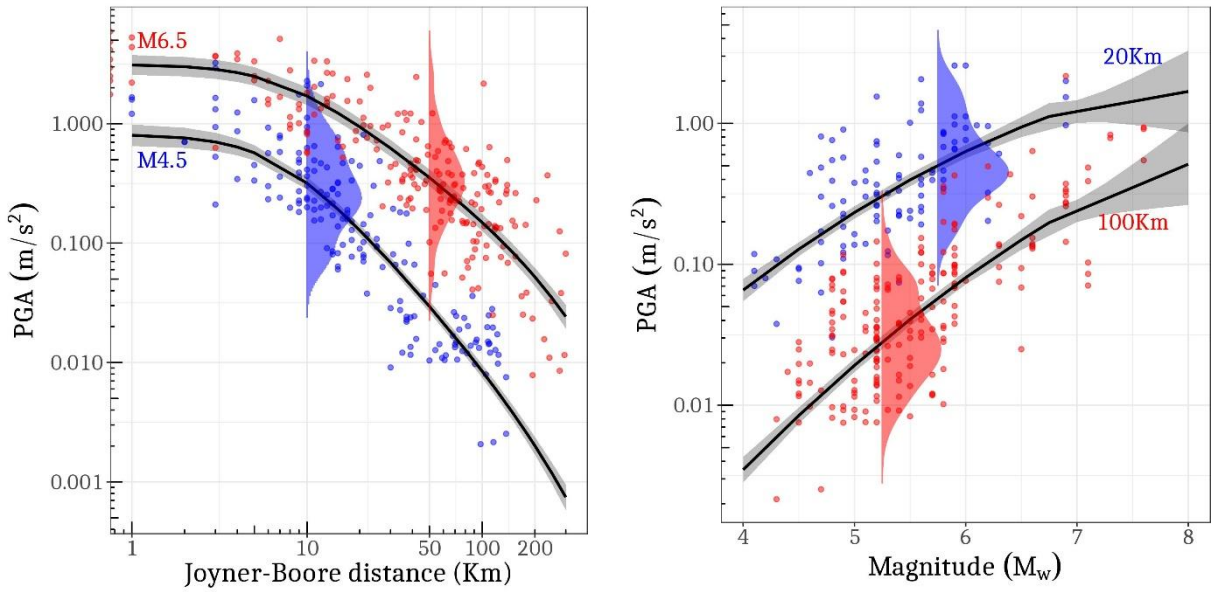


Figure 1-3: Attenuation of Peak Ground Acceleration (PGA) with distance (left) and magnitude (right). Black line represents the GMPE median (μ) and the kernel distribution illustrates lognormal variability (σ) of observed PGA. The grey ribbon visualizes the uncertainty on median (σ_μ).

1.1.1 Development of GMPEs

GMPEs are essentially regressed formulations of observed ground-motion scaling. The procedure in deriving one involves a series of steps, each with relevant assumptions and associated impact on μ, σ and σ_μ:

Step 1. Compilation of ground-motion datasets

Development of a GMPE is subject to availability of good quality ground-motion data in a region. For this purpose, seismically hazardous regions, where strong earthquakes in the past are known to have caused extensive damage to infrastructure and loss of life, are densely instrumented with accelerometric and seismological stations. Seismologists use the data from such networks to approximate the location, size, geometry, and other rupture parameters, constituting the 'event metadata'. While Geotechnical experts parametrize the geological conditions at each station, and formulate the 'site metadata'. Each event-site combination in the dataset has an associated seismogram, which contains the information on ground-motion attenuation required in GMPE development.

Processed seismograms from the ground-motion networks are compiled into a ground-motion dataset with event and site metadata, and the per-record ground-motion values. In Engineering Seismology, Spectral Amplitude (SA) is the preferred scale to measure ground-motion intensity. SA is the maximum response amplitude (acceleration/velocity/displacement) of a 5% (or any other damping level) damped Single Degree of Freedom (SDOF) oscillator with a fundamental resonance period T, when excited with a ground-motion recording. For $T = 0$, the peak acceleration response of an 'infinitely stiff' SDOF oscillator would be Peak Ground Acceleration (PGA), and for other periods (e.g. $T = 0.1$s, 1s, 2s etc.) the acceleration responses are the Spectral Accelerations (SA). A vector of SAs, i.e. responses of a suite of 5% damped SDOF oscillators with different fundamental periods, is called the Response Spectrum. A ground-motion dataset contains event and site metadata, and the 5% damped response spectrum for each strong motion recording.

The preliminary contributors to GMPE uncertainty are the inconsistencies in event and site metadata (e.g. Fig. 4 in *Ktenidou et al.*, 2017). For instance, erroneous event location translates into incorrect event-site distance measurement, biasing the distance-scaling component of GMPE μ. While erroneous moment-magnitude (M_w) effects the median scaling of ground-motion with rupture size. Similarly, with site

metadata, site parameters may contain errors from measurement (e.g. array size and inversion techniques) and inference techniques (e.g. spatial extrapolation of subsurface soil conditions). In essence, uncertainties in predictor variables (M, R and S), if not accounted during GMPE development, bias the median prediction μ and contribute to variability σ.

Step 2. Choice of parametrization

Given the event and site metadata, GMPE developers perform a non-parametric analysis to understand the scaling of SAs with different event, path, and site parameters. It is then a choice of the developer to select a few efficient parameters to model the observed SA scaling. However, the availability of suitable parameters is limited by the quality/absence of metadata, in which case the GMPE developer is forced to under-parametrize the attenuation phenomenon. If a known physical process cannot be modelled due to lack of parametrization, the resulting GMPE is limited in its explanatory power, which translates into the σ (e.g. *Bindi et al.*, 2017).

Step 3. Regression method

Regression refers to a family of predictive modelling techniques. In GMPE development, the prevalent method is Linear/Nonlinear Regression, where a continuous dependent/prediction variable (SA or other measures) is related to a set of continuous or discrete independent/predictor variables (M, R, S, etc.) through a best-fit line. Based on the non-parametric analysis, the GMPE developer frames a set of predictor functions to serve as the median μ of GMPE. Regression is then performed to obtain the best-fit, which is essentially a weighted combination of predictor variables constituting μ, that minimizes the residual variability σ. The weights, technically the regression coefficients, are conditional estimates reflecting the relative contribution of each predictor variable in explaining the observed SA scaling. Indeed, the choice of predictor variables regulates the size of σ, but there is an additional uncertainty on regression coefficients.

Uncertainty in μ is quantified as σ_μ, which is distinct from the GMPE total standard deviation σ. For GMPEs developed from densely instrumented high seismicity regions, where large varieties of earthquake scenarios were recorded, σ_μ is relatively small compared to σ. In case of GMPEs for low seismicity regions, where recorded data is insufficient to constrain the regression coefficients, σ_μ is comparable in size to σ. To overcome the data insufficiency and to reduce σ_μ, a typical practice in GMPE development is to make the so-called 'Ergodic Assumption' in Compilation of ground-motion data – the core topic of this study. However, the price of ergodic assumption is increased variability in data, and a consequently large residual variability σ.

Contributions to the total standard deviation σ of GMPE appear in every step of its development. The natural randomness inherent to compiled ground-motion data that cannot be explained with available event, path, and site parameterization is considered as Aleatory Variability. The uncertainty in event and site parametrization, choice of parameters (predictor variables) in GMPE, and regression coefficients associated with the chosen parameters, are consolidated as Epistemic Uncertainty. Efforts made in the Compilation of ground-motion data stage to achieve precise and exhaustive parametrization of event and site metadata reduce both the aleatory and epistemic uncertainty. However, the ergodic assumption made to reduce σ_μ comes at the cost of increased aleatory variability σ. The purpose of this study is to decompose and reduce σ, which means to relax the ergodic assumption.

1.1.2 Ergodic assumptions in GMPEs

The aim of PSHA is to predict hazard as the annual frequency of exceedance of a design ground-motion level at a site in a region. Prediction requires the probable location and size of a future earthquake, and in addition, the probable ground-motion intensity (at the site) given such an event occurs. It is customary in PSHA to extrapolate the location and size of future events based on the observed seismicity in region. Such spatiotemporal extrapolation of seismicity (through seismic source models) requires that GMPEs predict

the ground-motions for events that were rarely/never observed in the region, solely based on recorded ground-motion from a few past earthquakes in the region. GMPEs derived from the very limited ground-motion data in a moderate-low seismicity region become inapplicable for extraordinary scenarios, because of the very large σ_μ, i.e. the regression coefficients are poorly constrained for parametric values outside the range of recorded parameters in a region.

To overcome this limitation, a necessary ergodic assumption is made in the Compilation of ground-motion data step of GMPE development. The purpose of the ergodic assumption is to assemble ground-motion data with a wide range of event, path and site metadata, in order to keep the σ_μ reasonably small for every credible combination of parameters. Essentially, the ergodic assumption allows substituting the (ideal) temporal sampling of ground-motion recordings in a region (or a site) with spatial sampling from several regions with similar tectonics. For example, in deriving a GMPE for Active Shallow Crustal (ASC) regions, ground-motion data from multiple ASC regions is compiled into a unique global ASC dataset (e.g. *Ancheta et al.*, 2014). Any differences in strong motion attenuation across the various ASC regions are assumed negligible (e.g. *Douglas*, 2004). When theoretically plausible regional differences in the compiled data are not accounted in the empirical model (GMPE), all region-specific physical effects are rendered as generic natural variability, implying an inaccurate μ and large σ for all constituent regions.

Several studies discussed the many levels of ergodic assumption in GMPE development (e.g. *Anderson and Brune*, 1999; *Al Atik et al.*, 2010; *Lin et al.*, 2011). Some of the dominant contributors to the ergodic σ were identified, and methods were proposed for their quantification in non-ergodic GMPEs. However, to identify and quantify the systematic differences requires repeated measurements of the physical effects. With sufficient data, appropriate parametrization and regression technique, the ergodic assumption can be relaxed and the σ reduced.

1.2 Approach of the study

Purpose of this study is to improve the accuracy and precision of GMPE by relaxing the ergodic assumption(s) in its development: measured as reduction of ergodic bias in μ, and removal of ergodic variability from σ. Therefore, in the study, I use global strong motion datasets compiled of data from many regions, investigate various ergodic hypotheses, quantify non-ergodic biases, develop partially non-ergodic GMPEs and measure the impact of reduced GMPE ergodicity on hazard estimates.

1.2.1 Ergodic hypotheses in GMPEs

The three steps in GMPE development effect the GMPE performance: (1) incorrect estimates of event and site metadata during Compilation of ground-motion data, (2) inefficient Choice of parametrization, and (3) ergodic assumption in GMPE Regression method. Recent and ongoing research in source and site characterization significantly improved the quality of event and site metadata, thereby curtailing the first contribution to σ. Extensive parametrization allowed GMPE developers to test combinations of predictor variables that best capture the attenuation of SAs, thus subduing the second source of variability. The third, requiring an ergodic assumption to constrain the regression coefficients, remains the key focus of research. In fact, there are multiple ergodic hypotheses made in compiling global strong motion datasets, of which a few are extensively studied in the literature given their dominance on σ:

1. Regional ergodicity is a holistic ergodic hypothesis, where the event, path, and site physics, are considered independent of any region-specific peculiarities. NGA-West2 dataset by *Ancheta et al.* (2014) was compiled specifically for development of NGA-West2 GMPEs for use in PSHA in Western United States (WUS). However, the dataset contains Active Shallow Crustal (ASC) events from several regions across the globe. Figure 1-4 illustrates the distribution of event magnitude (M_w) and Joyner-Boore distance from the recording site (R_{JB}) in the NGA-West2 dataset. It is interesting to note that, since California lacks data from large $M_w \geq 7.5$, data from ASC events occurring in Taiwan, China and

Alaska were used to populate the sparse *M*-*R* ranges of the dataset. The resulting GMPEs become regionally ergodic, unless region-specific differences in ground-motion scaling are identified and accounted. Theoretically, regional differences in ground-motion scaling can be expected from the macro-spatial dissimilarities in crustal thickness, velocity, rigidity, and other structural properties. However, in order to quantify statistically significant regional differences, repeated observations sampling the crust within each region are required, along with parametrization of relevant crustal characteristics.

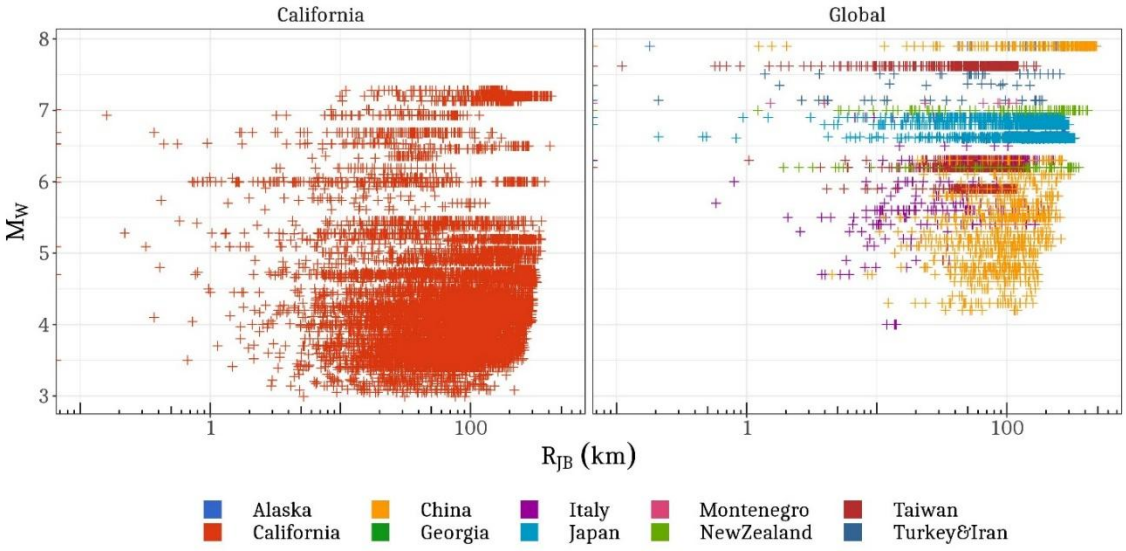


Figure 1-4: Regional diversity of strong ground-motion data in NGA-West2 dataset

2. Event ergodicity is typically a within-region ergodic hypothesis, where earthquake ruptures identified with similar parametric values are assumed to produce similar ground-motions. Events with the same magnitude (usually M_w) and focal mechanism (SoF) are treated as producing identical ground-motions. In terms of event location, ruptures with identical size and geometry, but occurring on different fault systems are assumed to produce identical ground-motions. Both the assumptions are debatable given the spatial variability of stress-fields and near-fault geology. Since the available rupture parametrization is insufficient, customary practice in GMPE development is to quantify (1) event-to-event variability for well recorded ruptures (e.g. *Brillinger and Preisler*, 1984; *Joyner and Boore*, 1993) and (2) location-to-location variability for well sampled fault systems (e.g. *Lin et al.*, 2011; *Villani and Abrahamson*, 2015).
3. Path ergodicity is another within-region ergodic hypothesis in which the attenuation of ground-motion is assumed identical across every conceivable wave propagation path between events and sites. The basis for path ergodicity is an idealization of Earth's crust as a homogeneous and isotropic half-space, and the point-source approximation of finite ruptures. However, to identify path-specific peculiarities in a GMPE requires path parametrization (e.g. *Spudich*, 2013) and repeated observations along each path. A few studies attempted parametrizing the path based on geographical indices and crustal properties, while others introduced event-to-site azimuth (along with distance) as an additional parameter in GMPEs to quantity path-to-path variability (e.g. *Lin et al.*, 2011; *Villani and Abrahamson*, 2015).
4. Site ergodicity is the third within-region ergodic hypothesis, which allows ground-motion data from several sites in a region to be compiled for GMPE regressions (*Anderson and Brune*, 1999). Local site conditions, such as soil stratum, topography, geology, etc., are known to dramatically amplify/attenuate (up to 10 times) specific ranges of ground-motion frequencies (e.g. *Boore*, 2004; *Thompson and Wald*, 2016). Such site dependent modifications are usually referred to as 'site response'. Despite this, site ergodicity in GMPEs assumes that all sites characterized by similar parametric values (site metadata) show identical response to seismic excitation (e.g. *Boore et al.*, 2011; *Thompson et al.*, 2012;

Puglia et al., 2015). A shortcoming of this hypothesis is concerning the parametrization of site response. Several site response studies proposed various parameters as optimal predictor variables to capture the complex site behavior (e.g. *Allen and Wald*, 2009; *Luzi et al.*, 2011; *Héloïse et al.*, 2012; *Derras et al.*, 2017). Given the definition of PSHA is to predict ground-motion levels at a specific site, site ergodicity is a rather strong assumption – thus a key focus of this study.

Region level ergodic assumption is the overarching hypothesis encompassing event, path, and site level ergodic assumptions. Therefore, resolving the latter three partially breaks the region level hypothesis. Although our goal is to achieve a fully non-ergodic GMPE capable of predicting accurate event-path-site specific μ with minimal σ, empirical and theoretical limitations allow only a partially non-ergodic GMPE.

1.2.2 Quantification of non-ergodic biases

An ergodic dataset is composed of several non-ergodic subsets. A non-ergodic bias is the mean deviation of non-ergodic observations from the mean of ergodic observations. Robust estimates of non-ergodic biases can be used to adjust the (generic) ergodic μ to obtain a (specific) non-ergodic μ_0 for any well-populated subset (i.e. region/event/path/site-specific subset of data). Removing non-ergodic biases from the residuals 'narrows' their scatter, and thereby reduces the residual standard deviation σ.

In order to estimate non-ergodic biases, repeated measurements sampling the non-ergodic physical characteristics are necessary. For instance, to resolve site ergodicity requires each site to have recorded multiple nearby earthquakes, i.e., sufficient site-specific data to quantify site-specific non-ergodic biases. A non-zero bias also implies that the site parameters used in developing a functional form for μ are inefficient in explaining the site response. Therefore, non-ergodic biases can be analyzed to derive additional site parameters, which can be then introduced into μ as new predictor variables. However, over-parametrization while improving the explanatory strength of a GMPE often reduces its predictive capability (e.g. *Shmueli*, 2010; *Bindi*, 2017). Additionally, if accurate site-specific parametric values are not available, e.g. for new sites not in the regressed dataset, then the applicability of GMPE is limited only to appropriately parametrized sites (e.g. *Thompson et al.*, 2012). Similar is the case with event and path ergodicity – only that sampling and parametrization of path effects is considerably more difficult. While all ergodic assumptions are made in the data compilation stage, partial non-ergodicity in GMPEs can be achieved at regression stage by properly accounting non-ergodic biases.

1.2.3 Partially non-ergodic GMPEs

A generic GMPE median is composed of predictor functions for magnitude, distance, site dependent scaling of ground-motion, as in $\ln(\mu) = F_M + F_R + F_S$ (e.g. *Brillinger and Preisler*, 1984). μ is an ergodic prediction because the GMPE is derived over a compendium ground-motion dataset assuming all events, paths, and sites with similar parametric values as identical. In this perspective, a fully non-ergodic prediction would require multiple recordings for every set of parametric values from every possible, past and future, combinations of event, path, and site parameters – which is impractical, because the choice of parametrization is itself an uncertainty. Under-parameterization decreases explanatory strength and increases ergodicity, while over-parametrization may explain better the observations but at the price of decreased predictive capability (e.g. *Bindi*, 2017). It is unlikely that any number of predictor variables can fully explain the natural variability. Therefore, parametrization to remove ergodicity is perhaps beyond the capacity of current scientific knowledge.

A more practical alternative is to quantify and use the non-ergodic biases as adjustments to ergodic predictions, to yield non-ergodic predictions. For this purpose, a widely used regression technique in GMPE development is the linear/nonlinear mixed-effects regression (LME/NLME, e.g. *Brillinger and Preisler*, 1984; *Abrahamson and Youngs*, 1992; *Stafford*, 2014). As a regression method, NLME produces a GMPE median μ and a residual standard deviation σ. An additional benefit crucial to our goal is that, NLME can estimate the non-ergodic biases for any user defined non-ergodic subsets. In the NLME terminology, the

non-ergodic biases are called Random-Effects (RE), in the sense that no physically meaningful combination of predictor variables can explain/predict characteristics specific to the non-ergodic subsets. The physics of ground-motion scaling that can be explained by a combination of predictor functions (μ), is handled by the Fixed-Effects (FE) – thus the name mixed-effects regressions. Essentially, FE alone gives the generic ergodic μ, and the combination FE + RE gives a specific non-ergodic median μ_0, where RE is a group-specific non-ergodic bias/adjustment.

The key NLME benefit is that the GMPE's FE and RE together constitute the μ_0, while the residuals, filtered for non-ergodic biases as RE, quantify the non-ergodic variability σ_0. Note that fully non-ergodic GMPEs are hard to achieve due to empirical and theoretical limitations, so μ_0 and σ_0 are at best partially non-ergodic. Also, μ_0 is a generic symbol for non-ergodic median, whose subscript changes depending on the level of non-ergodicity. For example, a site-specific non-ergodic prediction would be μ_s, s being the site index.

In this study, I mentioned four ergodic hypotheses in GMPE development: region, event, path, and site level ergodicity. Considering region level ergodicity as the holistic hypothesis, REs can be estimated at the event, path and site levels. In addition, NLME allows defining groups of non-ergodic subsets to investigate if groups of regions/events/paths/sites with similar parameters, collectively, show a significant bias. A few relevant grouping levels are as follows:

Event-to-event: Ruptures with similar parameters (e.g. magnitude, focal mechanism) are often observed to produce very different ground-motions (e.g. *Joyner and Boore*, 1981; *Brillinger and Preisler*, 1984); possibly, from dissimilar physical characteristics that could not be parametrized and predicted (e.g. rupture velocity, slip distribution). However, for a well-recorded event, the event-specific ground-motion data can be viewed as the event-specific non-ergodic subset. The mean deviation of event-specific SAs from the ergodic GMPE prediction for given event-specific parameters, is the event-specific non-ergodic bias, commonly referred to as event-term δB_e. For a dataset featuring several well-recorded events, δB_e is estimated for each. The resulting random distribution of δB_e represents the event-to-event variability.

NLME estimates δB_e (of all events in the dataset) as a normally distributed random variable with zero-mean and an event-to-event variance τ^2. Thus $\delta B_e \sim N(0, \tau^2)$ describes the between-event variability of the dataset with respect to the parametric GMPE. The ergodic μ can be adjusted with an event-specific δB_e, to yield a partially non-ergodic event-specific prediction $\ln(\mu_e) = \ln(\mu) + \delta B_e$, in which case the GMPE variability would be a partially non-ergodic σ_0 devoid of event-to-event variability. In cases where δB_e for a prospective event cannot be predicted, it is necessary to use the ergodic μ, with τ reintroduced into GMPE variability, as in $\sigma = \sqrt{\sigma_0^2 + \tau^2}$.

Event location-to-location: NLME allows grouping multiple non-ergodic subsets to estimate a group-specific bias. Given a dataset with several well-recorded events, and/or several events occurring within a seismogenic zone, e.g. a fault system, seismic source depth, or other seismic localization, a GMPE developer may be interested in studying a location-to-location variability of observed ground-motions (e.g. *Lin et al.*, 2011; *Villani and Abrahamson*, 2015). As with event-to-event variability, a location-to-location variability can be estimated as $\delta L2L \sim N(0, \tau_L^2)$. The location-specific $\delta L2L$ is a repeatable effect with better predictability than the δB_e, provided an a priori regionalization of events is available. Otherwise, spatial distribution of δB_e can be queried for evidence of event regionalization.

Path-to-path: The wave propagation path from the event to site is a complex geometry with multiple internal reflections, refractions and scattering. Prior to attempting path parametrization, defining path is itself a complex issue (e.g. *Scherbaum et al.*, 2004). In GMPEs however, a 3D wave path is simplified into a straight-line connecting event location (e.g. epicenter) and the site location. For small magnitude events, a point-source approximation of rupture allows at least two definitions of path: epicenter to site, hypo-

center to site. For larger finite ruptures, the alternatives are several, including a few accounting the hypocentral depth. Irrespective of definition, the only path parameter used in GMPEs is the event-to-site distance (in km).

Path ergodicity assumes the rate of attenuation (per unit distance) as constant across the region, which is a rather strong assumption given the spatial variability of crustal properties. Path-to-path variability ($\delta P2P$) is therefore a large contributor to σ. To estimate $\delta P2P$, a few studies grouped paths connecting seismogenic zones and network of sites, while others attempted azimuth-dependent grouping of paths connecting a single event to network of sites (e.g. *Dawood and Rodriguez-Marek*, 2013; *Villani and Abrahamson*, 2015; *Landwehr et al.*, 2016). Despite, there are a very few datasets that facilitate investigating $\delta P2P$.

Site-to-site: As with δB_e, site-terms can be estimated for well-recorded sites with sufficient strong motion data from multiple nearby earthquakes. Assuming the local site conditions do not change over time or with successive earthquakes, site-specific deviation from ergodic μ is the site-term $\delta S2S_s$ (e.g. *Rodriguez-Marek et al.*, 2013). Since $\delta S2S_s$ are NLME estimated random-effects, $\delta S2S_s \sim N(0, \phi_{S2S}^2)$ where ϕ_{S2S} is the between-site variability of the ergodic dataset. Compared to the δB_e and $\delta P2P$ terms, which are difficult to parametrize and predict, the site-specific $\delta S2S_s$ are highly repeatable terms. Given sufficient time, any instrumented site in a seismically active region can accumulate enough ground-motion data to constitute its site-specific non-ergodic subset. The ease of estimation, time-independent repeatability, and relative ease of parametrization, makes site non-ergodicity a widely practiced concept in PSHA.

Region-to-region: Regional ergodicity is partially broken when any of the above non-ergodic biases are filtered from the overall variability. On the other hand, if any of the above cannot be quantified (due to lack of sufficient data); macro-spatial regional differences can be resolved by grouping events, paths and sites into encompassing regions. However approximate, region-specific RE quantify regional differences in geology, crustal properties, lithography, and other macro-spatial features that are likely to have a cumulative effect on the attenuation process.

Along with the above, multiple non-ergodic subsets can be grouped together to estimate group-specific random-effects. For example, it is a common practice to group several soft soil sites to estimate the average amplification factors with respect to a rock site group. Similarly, events can be grouped based on their hypocentral depth, and paths based on their traversal or not, across large crustal asperities. Such predefined groups allow the GMPE developers to investigate the extent of (average) group-specific deviations from ergodic predictions.

1.2.4 Partially non-ergodic Probabilistic Seismic Hazard Assessment

Probabilistic Seismic Hazard Assessment for a site in a region can be performed given a seismic source model and a GMPE. If the GMPE used is applicable to the region's tectonic regime and is ergodic, the resulting PSHA is ergodic as well: a non-ergodic GMPE results in a non-ergodic PSHA. Non-ergodic GMPEs predict group-specific median ground-motions with a smaller variability, which means the physics of ground-motion scaling is better explained. However, to measure the practical benefit of shifting from ergodic to non-ergodic GMPEs is preferable in terms of change in hazard level at a site. In this study, I make use of peer-reviewed seismic source models and open-source PSHA software to illustrate the benefit of shifting from ergodic to non-ergodic GMPEs.

1.3 Presentation of the study

In describing the purpose of and approach in the study, I identified the key ergodic assumptions in GMPEs, and introduced our approach to estimate the non-ergodic biases and the impact of GMPE non-ergodicity on PSHA. The outcomes of this study are presented in chapters, wherein the concepts are further elaborated.

Chapter 2: In this chapter, we worked with the high quality NGA-West2 dataset (*Ancheta et al.*, 2014) with qualified metadata – thus, subduing the initial contribution to σ from incorrect metadata. NGA-West2 is an ergodic dataset compiled of strong motion data from several high seismicity Active Shallow Crustal regions across the globe. The event, path, and site metadata are extensively parameterized, allowing the GMPE developers to model the highly sophisticated NGA-West2 GMPEs. Using these models as benchmark, we: (1) develop a new under-parametrized GMPE for application in low seismicity regions, (2) evaluate the impact of under-parametrization on σ, and (3) demonstrate the dependence of σ_μ on data sufficiency and parametrization.

Chapter 3: The pan-European dataset, RESORCE (*Akkar et al.*, 2014), was compiled of strong motion data from large $M7$ events in Iran and Turkey together with smaller $M4.5$-$M5.5$ events in Italy. Regionally ergodic GMPEs derived from this dataset, were thus capable of predicting ground-motions for events $4.5 \leq M_w \leq 7$ at distances $0\text{km} \leq R_{JB} \leq 300\text{km}$. If only the Italian data were used to derive a GMPE, σ_μ would be too large for $M7$ scenarios – and vice versa for smaller $M5$ events, if only Turkish data were used. Assuring that σ_μ is relatively small for our choice of parametrization; we look at the other two contributors to σ: incorrect metadata and regional ergodicity. In this chapter, using NLME (*Bates et al.*, 2014) we develop a new region-specific GMPE to demonstrate the reduction in σ from, (1) accounting regional differences in attenuation between Italy and Turkey, and (2) discarding data with incorrect event metadata.

Chapter 4: The partially non-ergodic region-specific GMPE developed in previous chapter predicts region-specific ground-motions for Italy and Turkey, and is accompanied by a σ smaller than its ergodic counterpart. In this chapter, we go a step further and develop a region- and site-specific GMPE, i.e. a non-ergodic GMPE capable of predicting site-specific ground-motions. Since region- and site-specific non-ergodic biases are removed from the residual variability, the σ is reduced to a significantly smaller σ_0. In process, we devise a framework to perform region- and site-specific PSHA, which we then use to illustrate the practical benefit of shifting from ergodic to region- and site-specific PSHA.

Until this point, we focused on isolating the non-ergodic biases/trends from the σ and introducing them into μ, thereby improving the accuracy and precision of GMPE in predicting partially non-ergodic region- and site-specific response spectra. In the later chapters, we focus more on practical application of non-ergodic GMPEs.

Chapter 5: In engineering applications, especially in earthquake resistant building design, a practical necessity is to predict site-specific response spectra for future events. An accurate prediction is possible with site-specific GMPEs involving $\delta S2S_s$. Ideally, if a site has enough strong motion data in its non-ergodic subset, one could estimate the $\delta S2S_s$, use it to adjust the non-ergodic GMPE median μ_0, and predict site-specific ground-motions as in $\ln(\mu_s) = \ln(\mu_0) + \delta S2S_s$. However, most sites are either unsuitable for instrumentation or there is not enough time to gather site-specific strong motion data, and estimate $\delta S2S_s$. Therefore, for sites with no strong motion data, we need to predict a $\delta S2S_s$.

In this chapter, we tackle an engineering approach to predicting $\delta S2S_s$ for new sites: site classification. Site classification relies on the ergodic assumption that sites with similar parametric values also show similar response to seismic excitation. In this context, the response is measured in terms of $\delta S2S_s$. Since a new site has no recorded strong motion data, site classification is a practical approach to predicting an approximate $\delta S2S_s$ based solely on its parametric values. However, there are a few challenges in this approach regarding: (1) optimal number of site classes, (2) efficient site parameters that define site classes, (3) between- and within-class site response variability compared to single-site and ergodic variability. In response, we propose a novel approach to site classification, where we estimate $\delta S2S_s$ for several sites in Japan (using KiK-net dataset), classify sites through spectral clustering technique, and test various site parameters for efficiency in distinguishing the classes. This new approach is possible only with the region- and site-specific GMPEs developed in previous chapters.

Chapter 6: In the previous chapter, we established a method to predict $\delta S2S_s$ from site metadata. The approximate $\delta S2S_s$ can be used to estimate site-specific response spectrum at sites with no strong motion data, which is significantly more accurate and precise than an ergodic RS response spectrum. However, more advanced Earthquake Engineering applications require not just the response spectra but the Conditional Spectra (*Baker*, 2011).

Conditional spectrum describe the distribution of SAs at $T \neq T_0$, on the condition that SA at $T = T_0$ exceeds a certain ground-motion level, where $T_0$is the fundamental period of resonance of a structure at the site. To determine the conditional spectra, empirical correlation of peak spectral amplitudes for a period range are needed. The basis for correlation of SAs (across a range of T values) is that, within each ground-motion record the ground-motion amplitudes at different frequencies are correlated. Since GMPEs ignore such within-record correlation of SAs, empirical correlation of SAs are determined from the random-effects and residuals. For instance, empirical correlation of SAs in source (event) spectrum are derived from δB_e, and those related to path are derived from residuals corrected for source and site effects. Essentially, a record with larger than predicted median SA at $T = T_1$ will also show a systematically larger SA at $T = T_2$ depending on the correlation coefficient ρ_{T_1,T_2}.

Traditionally, correlation of δB_e are assumed independent of any rupture characteristics, which is an ergodic hypothesis stating that all ruptures produce identical distribution of SAs. Similar is the case with path terms $\delta P2P$, where relative attenuation of different SAs is assumed to be identical across any path or region. In this chapter, we challenge these assumptions to develop non-ergodic correlation models.

2

Application-driven Ground-Motion Prediction Equation for Seismic Hazard Assessments in non-Cratonic Moderate-Seismicity Areas

D. Bindi[1], F. Cotton[1, 2], **S. R. Kotha**[1, 2], C. Bosse[1], D. Stromeyer[1] and G. Grünthal[1]

[1]German Research Centre for Geosciences GFZ, Potsdam, Germany

[2] University of Potsdam, Potsdam, Germany

Abstract

We present a Ground-Motion Prediction Equation (GMPE) for probabilistic seismic hazard assessments (PSHA) in low-to-moderate seismicity areas, such as Germany. Starting from the NGA-west 2 flat-file (*Ancheta et al.*, 2014), we develop a model tailored to the hazard application in terms of data selection and implemented functional form. In light of such hazard application, the GMPE is derived for hypocentral distance (along with the Joyner-Boore one), selecting recordings at sites with $v_{S30} \geq 360$m/s, distances within 300 km, magnitudes in the range 3 to 8 (being 7.4 the maximum magnitude for the PSHA in the target area). Moreover, the complexity of the considered functional form is reflecting the availability of information in the target area. The median predictions are compared with those from the NGA-west 2 models and with one recent European model, using the Sammon's map constructed for different scenarios. Despite the simplification in the functional form, the assessed epistemic uncertainty in the GMPE median is of the order of those affecting the NGA-west2 models for the magnitude range of interest of the hazard application. On the other hand, the simplification of the functional form led to an increment of the apparent aleatory variability. In conclusion, the GMPE developed in this study is tailored to the needs for applications in low-to-moderate seismic areas and for short return periods (e.g., 475 years); its application in studies where the hazard is involving magnitudes above 7.4 and for long return periods is not advised.

Keywords: Ground-Motion Prediction Equations; moderate seismicity region; NGA-West2

2.1 Introduction

Since 2003, the Pacific Earthquake Engineering Research Center (PEER) is conducting a large research program to develop the next generation of ground-motion prediction equations (GMPEs) for shallow crustal earthquakes in active tectonic regions. The second phase of this project (called NGA-West2) concluded in 2014 and provided important results, including a strong motion database of recorded ground-motions (*Ancheta et al.*, 2014) and a set of peer-reviewed GMPEs (*Abrahamson et al.*, 2014).

Several recent hazard projects have shown that the models developed by NGA-west projects may be of interest not only for active regions but also for non-cratonic and lower seismicity regions. NGA-west ground-motion models have been selected as part of ground-motion logic tree to compute recent probabilistic seismic hazard assessments (PSHA) in Europe (*Delavaud et al.*, 2012), Switzerland (*Edwards et al.*, 2016) and Germany (*Grünthal et al.*, 2017). These models are also widely used in regions where active faults have not been identified for most seismic sources , which means that rock hazard computations have to be conducted for distributed seismicity (area sources or zoneless approach) without taking into account the refinements introduced by recently developed GMPEs (directivity and hanging wall effects, non-linear site effects, basin effects).

These projects show that existing NGA-west2 GMPEs, despite their high quality, are then not fully fitting the needs of regional hazard computations in moderate seismicity areas. The main problems encountered in the application of the NGA-west2 models in low-to-moderate seismic areas are the following:

1) Modern GMPEs use definitions of the source-to-site distance that reflect the dimensions of the fault rupture for larger earthquakes better than point-source measures relative to the epicenter or hypocenter. This is a positive development since it more realistically reflects the fact that energy is released from the crust around the entire fault rupture during a large earthquake. However, seismic source configurations defined for PSHA in low-to-moderate seismicity areas almost invariably include areas of distributed point-source seismicity. Point-source simulations can be enhanced to include simulations of virtual extended ruptures. These adaptations are computationally demanding and not easily implemented. As suggested by *Bommer and Akkar* (2012), there is a need to compute pairs of equations, one using an extended-source distance metric, the other a point-source measure. To our knowledge, such pairs of equations have not been performed by the NGA-west 2 project.
2) A problem often encountered in the application of the NGA-west 2 GMPEs based on complex functional forms is related to the availability of suitable metadata in the target region. In low to moderate seismicity regions the source and site characterizations are generally not as detailed as in the data set used to derive the GMPE (host region). In such cases, the GMPEs are applied in simplified forms, where one or several variables (e.g., basin depth, hanging wall foot wall effects) are constrained to default values. This operation should be accompanied by either a proper handling of the epistemic uncertainty introduced when fixing some variables, or by propagating the uncertainty to the aleatory component. Both choices imply some additional work and expert decisions.
3) The hazard computed at a given location depends on both the seismic source model and on the ground-motion model. Hazard computations in low-to-moderate seismicity areas are particularly dependent on the GMPE magnitude scaling around magnitude 5.5-6. Some NGA-west 2 models have chosen functional forms with a magnitude hinge around $M = 5.5$. Such a choice has a low impact on hazard computations in high seismicity regions but a larger one in moderate seismicity regions. This application-driven practical issue motivates the development of functional forms adapted to moderate seismicity areas.

To overcome these problems, we derived a new GMPE using the high quality PEER flat file but tuning the complexity of the model to the information available in moderate seismicity regions. This development is also motivated by the needs of hazard computations in low-to-moderate seismicity areas: pairs of equations (one using an extended-source distance metric, the other a point- source measure), focus on stiff soil and rock site conditions, specific magnitude-scaling analysis in the magnitude range 5.5-6.

The use of simpler GMPEs (point source distance metric, lower number of input parameters) has however two main drawbacks, which will be analyzed in the following:

1) The aleatory variability of GMPEs (σ) has a strong impact on the results of PSHA at long return periods. The σ values are indeed an estimate of the apparent aleatory variability since they are evaluated with respect to the chosen model. The use of simpler models implies larger σ and the impact of such increase has to be carefully evaluated.
2) GMPEs are used as part of a logic tree or selected as a backbone equation. In both cases, it is important to evaluate the GMPEs epistemic uncertainty of the median (particularly in the magnitude range of interest) and the proximity of the model with other published models, e.g., using Sammon's maps (*Scherbaum et al.*, 2010).

This article is organized as follow. First, we discuss the motivations that led us to derive a GMPE tailored to our specific hazard application (*Grünthal et al.*, 2017). Then, after describing the functional form and the data considered for the GMPE development, we discuss the epistemic uncertainty in the median and the aleatory variability. Finally, the comparisons of the median predictions with NGA-West2 GMPEs are presented in terms of Sammon's map and Trellis charts.

2.1.1 Towards an application driven GMPE: Hazard assessment in a moderate seismicity area (Germany) and associated needs

This study is part of the German Hazard map project accomplished on behalf of the Deutsches Institut für Bautechnik (DIBt; The Centre of Competence in Civil Engineering). The new version of the national PSHA should predict uniform hazard spectra (UHS) for any site within Germany, hazard maps for spectral accelerations, peak ground accelerations, and disaggregation for the hazard levels of 10%, 5% and 2% exceedance probability within 50 years (*Grünthal et al.*, 2017). All hazard calculations had to be performed for $v_{S30} = 800$ m/s, where v_{S30} is the time-averaged shear-wave velocity of the top 30 m. Induced events are excluded from the study. Epistemic uncertainties have been explored both for the seismicity and ground-motion model part.

The tectonic context of Germany (*Grünthal et al.*, 2017) is complex with active structural elements mainly along the chain of the Rhine Graben up to rather stable parts towards the north and northeast. Because of this complexity, GMPEs logic tree used in past seismic hazard studies for this part of the West European Platform (e.g. *Delavaud et al.*, 2012) included equations for active crustal regions (ACR).

The use of ACR models calibrated to the NGA-west database was also motivated by recent stochastic models and GMPE testing performed in Western Europe. In France, *Beauval et al.* (2012) tested several GMPEs: the NGA-west 1 *Abrahamson and Silva* (2008) model was ranked as one of the best models. *Drouet and Cotton* (2015) developed and tested a new stochastic model based on data recorded in the French Alps and their resulting model is consistent with GMPEs derived for active crustal regions (e.g. *Boore and Atkinson*, 2008). In Switzerland, *Edwards and Fäh* (2013) and *Cauzzi et al.* (2015a) proposed stochastic ground-motion models of the Swiss Foreland and the Swiss Alpine region. They also showed that Swiss stochastic ground-motion models are broadly consistent with the NGA-west-1 *Chiou and Youngs* (2008) model.

The GMPEs logic tree implemented for updating the seismic hazard in Germany is composed by three main branches, each of them including one or more models derived from different data sets. In particular, while the first branch includes models derived for Europe using the RESORCE data set (*Akkar et al.*, 2014), the other two branches are relevant to GMPEs calibrated considering global ACR data sets (*Grünthal et al.*, 2017). The model of *Cauzzi et al.* (2014b), calibrated over a data set mostly populated by Japanese earthquakes, is considered for one of the two branches while the other was reserved to a GMPE derived from the NGA2-west flat file (*Ancheta et al.*, 2014). The seismicity model used to derive the German hazard map is based mainly on area sources. Disaggregation analyses have been performed over magnitude and distance scenarios for preliminary hazard assessments for several representatives sites (Figure 2-1). The

hazard (for the return period $RP = 475$ years) is controlled mainly by earthquakes of moderate magnitude ($M < 5.5$) at distance below 25 km. The disaggregation results suggest a significant impact of the functional form chosen to define the magnitude scaling between $M = 5$ and $M = 6$ (Figure 2-1). For example, the kink in the magnitude scaling of the BSSA14 model (*Boore et al.*, 2014) around the hinge magnitude $M_w = 5.5$ increases the relative contribute to hazard of scenarios for magnitude between 5 and 6 and distances around 20 km. In the following, we use the PEER-NGA2 flat file to derive a new GMPE whose functional form is selected for the specific hazard study of interest.

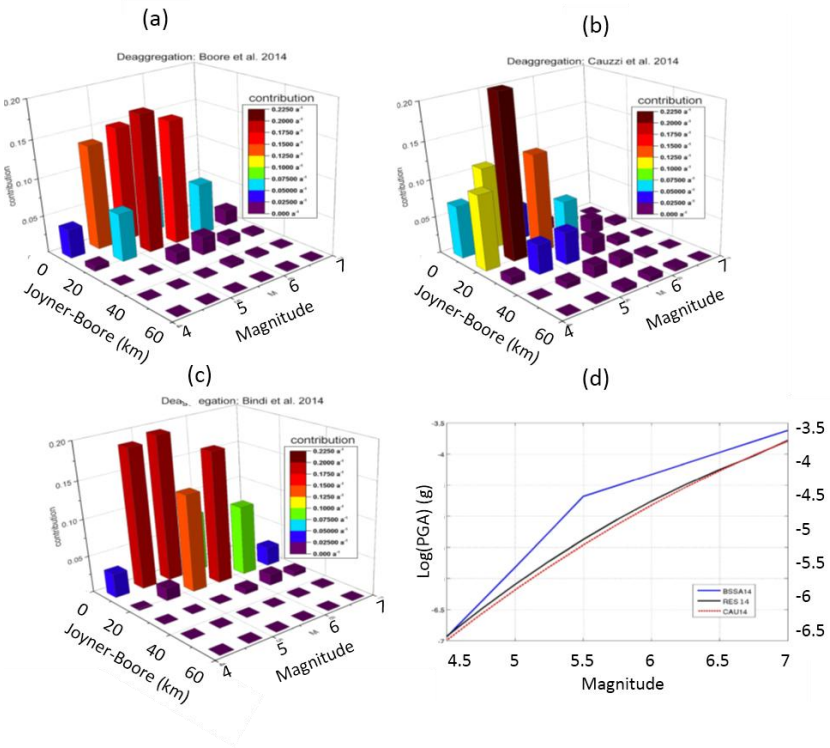


Figure 2-1: Disaggregation over magnitude and distance scenarios for preliminary hazard assessment at a representative site in Germany. (a) BSSA14 (Boore et al., 2014) model; (b) Cauzzi et al. (2015b) model (CAU14); (c) Bindi et al. (2014) model (RES14). In panel (d), the PGA magnitude scaling for the three GMPEs is compared considering a vertical strike slip earthquakes at 30 km ($V_{S30} = 800$m/s).

2.1.2 NGA-west2 data and GMPE development

To develop a global GMPE for logic tree implemented in the hazard assessment for Germany, we consider the PEER-NGA2 flat file (*Ancheta et al.*, 2014). In particular, starting from the *Campbell and Bozorgnia* (2014) data selection, further selection criteria related to the specific application are applied to the data set. We selected only recordings for distances less than 300 km and stations with v_{S30} computed or estimated from shear wave measurements (i.e., v_{S30} code either 0 or 1; *Ancheta et al.*, 2014). Moreover, since the hazard application is performed for rock condition, only $v_{S30} \geq 360$ m/s are selected in order to limit possible bias in the median due to not properly modelled site effects for soft sites (i.e., neglecting nonlinear effects). Figure 2-2 shows the scatterplot for the distribution of magnitude with hypocentral distance. The selected data set is composed by 4692 recordings from 242 earthquakes and 1025 stations. The [16th, 50th, 84th] percentiles of the magnitude, Joyner-Boore and hypocentral distributions are [3.7, 4.3, 6.7], [43.8, 111.9, 209.5] km, and [49.7, 117.2, 217.8] km, respectively. In particular, for magnitudes above 4.5, hypocentral distances below 10 km are almost not sampled. Regarding the selected stations, the percentiles for v_{S30} are [393, 511, 786] m/s, while the number of recordings for Eurocode8 class A and B are 689 and 4013, respectively (Figure 2-2).

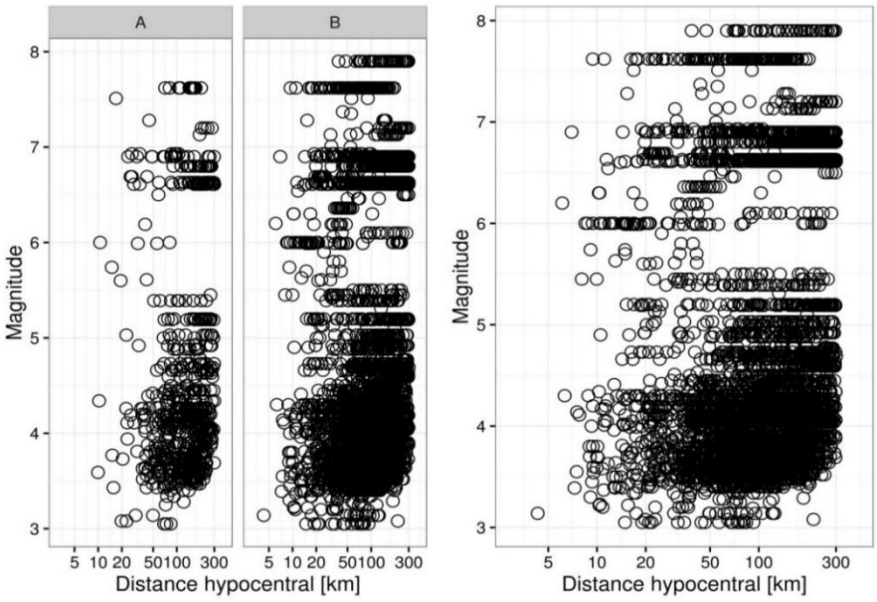


Figure 2-2: Distribution of magnitude versus hypocentral distance for the considered recordings. In the left panels, the distributions for classes A and B of Eurocode 8 are shown separately.

Regarding the functional form, the involved explanatory variables should reflect the metadata availability in the target region and the requirements of the specific hazard assessment. For example, since the hazard is computed for rock site conditions ($v_{S30} = 800$ m/s) and in a low-to-moderate seismicity area, only the linear site effect is implemented, without correction for basin effects. Also, since the hazard assessment is based on source area models, the hypocentral distance is preferred, although the model is also derived for the Joyner-Boore distance, following (*Bommer and Akkar*, 2012). Moreover, extended source effects such as hanging-wall/foot-wall terms are not modelled. For the aforementioned reasons, the following functional form is considered:

$$\ln Y = e_1 + F_D(R, M) + F_M(M) + F_S \tag{1}$$

Where the distance F_D and magnitude F_M functions are given by:

$$F_D(R_{JB}, M) = [c_1 + c_2(M - M_{ref})] ln\left(\sqrt{{R_{JB}}^2 + h^2}/R_{ref}\right) + c_3\left(\sqrt{{R_{JB}}^2 + h^2} - R_{ref}\right) \tag{2}$$

$$F_D(R_{hypo}, M) = [c_1 + c_2(M - M_{ref})] ln(R_{hypo}/R_{ref}) - c_3(R_{hypo} - R_{ref}) \tag{3}$$

$$F_M(M) = \begin{cases} b_1(M - M_{ref}) + b_2(M - M_{ref})^2, M < M_h \\ b_3(M - M_h) + b_1(M_h - M_{ref}) + b_2(M_h - M_{ref})^2, otherwise \end{cases} \tag{4}$$

$$F_S = sA \ \ln(V_{S30}/800) \tag{5}$$

In Eq. (2) and (3), the Joyner-Boore (R_{JB}) and hypocentral (R_{hypo}) distances are considered, respectively. The reference distance R_{ref} have be set equal to 1 km, the reference magnitude M_{ref} to 4.5 (i.e., close to the 50th percentile of the cumulative number of recordings versus magnitude). The hinge magnitude M_h introduced to handle the saturation in the magnitude scaling, is set equal to 6.5, that is, slightly above the values suggested by data (6-6.2), to move the kink in the magnitude scaling at a magnitude larger than those controlling the hazard (Figure 2-1). After preliminary tests, the style of faulting term is not considered because not justified in term of bias-variance trade-off, using the AIC parameter (*Akaike*, 1973). The regression is performed using a mixed effect approach (*Abrahamson and Youngs*, 1992; *Bates et al.*, 2014), accounting for the between-event residuals as random effect on the offset depending on the earthquake grouping level. The models are calibrated for 5%-damped pseudo-acceleration response spectra, considering 90 periods ranging between 0.01 and 4s.

2.2 Results

In the following paragraphs, we discuss the results in terms of: fixed effects, epistemic uncertainty in the median and aleatory variability.

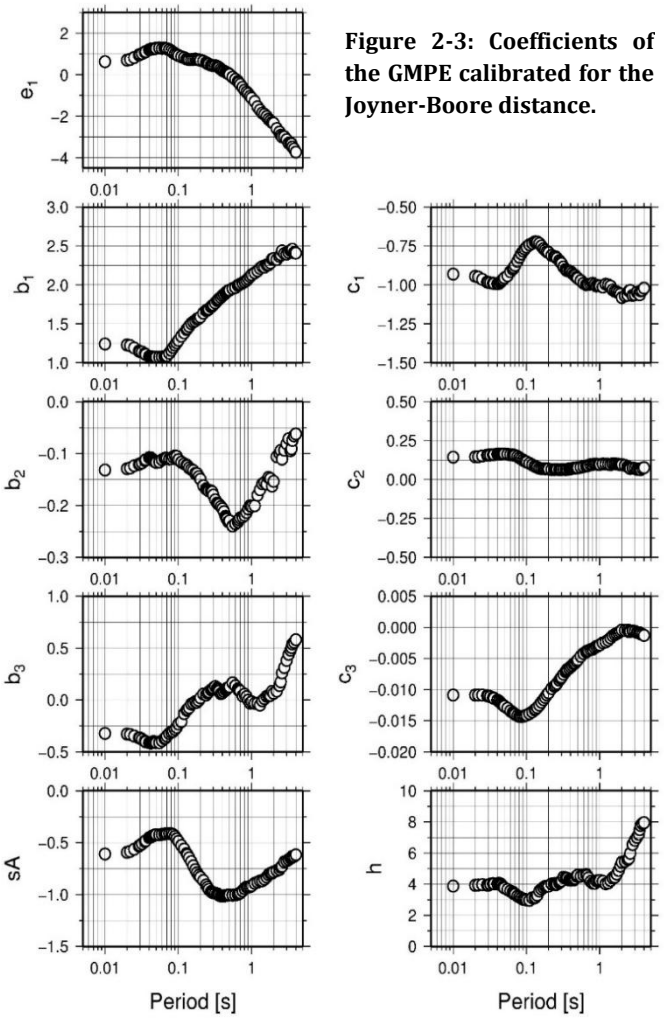


Figure 2-3: Coefficients of the GMPE calibrated for the Joyner-Boore distance.

2.2.1 Fixed effects (median model)

The obtained coefficients are listed in Tables 1 and 2 included in the Electronic Supplements. Figure 2-3 shows the variability with period of the coefficients obtained for the GMPE implementing the Joyner-Boore distance. The trends of the parameters with periods determine the scaling of the model with respect to the explanatory variables. For example, the decrease (in absolute value) of coefficient c_3 with period reflects the decrease of the attenuation proportional to distance, sometimes referred to as anelastic attenuation term, although this interpretation is strictly valid only in the Fourier domain, e.g. *Bora et al.* (2016), which almost vanishes above 2s. The effect of this term is largest at 0.1s and, for periods smaller than 0.03s, it is constant. The coefficient c_2 in Eq. (2) and (3) controls the magnitude dependence of the attenuation with the logarithm of distance (sometimes referred to as geometrical spreading, in analogy with a model for Fourier). As shown in Figure 2-3, c_2 is positive and almost constant. Its effect on $\ln Y$ depends on the sign of $(M\text{-}M_{ref})$: for $M > M_{ref}$, the c_2 term reduces the attenuation with distance while, for $M < M_{ref}$, it increases the attenuation with distance. Then, the distance attenuation is more significant for small magnitudes. This effect is more evident for short periods (below 0.05s) and almost vanishes between 0.1 and 0.6s. Another example is the site coefficient sA (Eq. 5). For velocity lower than 800m/s, the term $\ln(V_{S30}/800)$ is negative. Then, the trend of sA in Figure 2-3 implies that the site amplification effects are larger between 0.03 and 0.1s for velocities larger than 800m/s and between 0.2 and 1s for velocities smaller than 800m/s. Since the different terms in Eq. (2) through (5) could be affected by mutual trade-offs, Figure 2-4 shows the overall dependencies of the predictions on magnitude and distance, i.e., by grouping all terms depending on these explanatory variables. In the top panel, the period dependence of the term F_M given by Eq. 4 is shown for two different magnitudes (i.e., 4 and 6). Since also the c_2 term includes the magnitude, the plot is repeated for two different distances (continuous line for $R = 30$km; dashed lines for $R = 100$km). The dependence on the source recalls the shape of the response spectra, with a sharper increase with decreasing period for smaller magnitude. For short periods, the curve flattens and it is almost independent on the frequency of the oscillator. The dependence on distance is weak, that is the role of the c_2 term with respect the terms with b_1, b_2 and b_3, and only appreciable for short periods (see the discussion above about c_2). The period dependence of F_D (Eq. 3) is shown in the bottom panel of Figure 2-4, for four scenarios defined by $M6$ and $M4$ at 30 and 100km. By comparing the curves for $R = 30$ and 100km, we see that the overall effect of c_1 is to scale the

prediction. For a given distance, the impact of the magnitude (through the term with c_2) becomes significant for periods below 0.15s, where large magnitude are less attenuated than the small ones, as previously discussed for the period dependence of c_2.

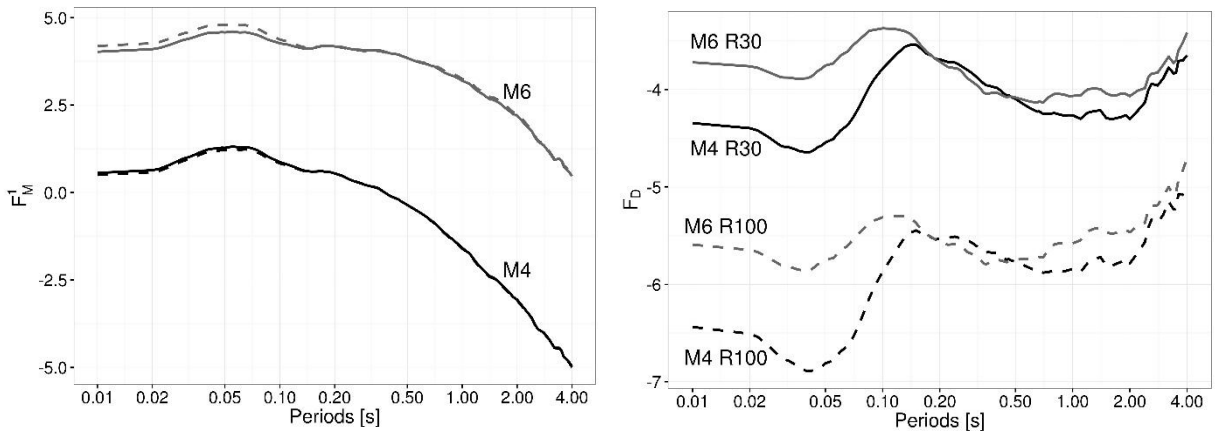


Figure 2-4: Top: Period dependence of the magnitude terms. controlled by parameters b_1, b_2, b_3 and c_2 (F_M^1 is given by the F_M term in equation 4 plus the c_2 term of equation 3), for magnitude 6 (gray) and 4 (black), and two distances ($R = 30$km, continuous line; $R = 100$km, dashed lines). Bottom: Period dependence of the distance terms controlled by c_1, c_2 and c_3 (F_D term, see equation 3), for magnitude 6 (gray) and 4 (black) and two distances ($R = 30$km, continuous lines; $R = 100$km, dashed lines).

2.2.2 Epistemic variability in the median

The variance-covariance matrix of the model quantifies the uncertainties of coefficients (diagonal elements) and their mutual trade-offs (off-diagonal elements). Following *Al Atik and Youngs* (2014), the variance-covariance matrix and the matrix of the gradients of the model with respect to the coefficients can be used to assess the epistemic uncertainty in the median model (see *Al Atik and Youngs* (2014) for a detailed discussion of the methodology):

$$var[\overline{lnY}]_{x_0} = J_0^T \left[varCov_{x_i} \right] J_0 \qquad (6)$$

Where, the Jacobian matrix J_0 is evaluated in the predictive location x_0 and the variance-Covariance matricx $varCov$ is evaluated in the data points x_i used to develop the model. Figure 2-5 shows the ingredients to assess the variance of $\ln(SA)$ at 0.1s. For graphical reasons, the variances of the model coefficients and the correlation matrix are shown instead of $varCov$. The standard deviation σ_μ of the median (i.e., the square root of the left hand side term in Eq. (6) quantifies the epistemic uncertainty in the median due to the combined effects of limited data availability and implemented functional form.

The largest variances are those of e_1 and b_3 (Figure 2-5, panel a), while the largest trade-off occurs between c_1 and c_3, between e_1 and c_1, between b_1 and c_2, and between e_1 and c_3 (Figure 2-5, panel b), reflecting of the trade-off between the source and attenuation terms. Following Eq. (6), the uncertainty on the median is controlled by the product of these terms with the values of the gradient of the model with respect to the coefficients, evaluated for the predictive scenarios. Figure 2-5, panel (c), shows the gradients for different magnitude and distance combinations, and for $v_{S30} = 600$m/s. It is worth noting that, in the case of hypocentral model, the model is linear with respect to the coefficients and therefore the gradients are period independent. The dependence of σ_μ on period eventually arises from the variance-covariance matrix. On the contrary, for the Joyner-Boore model, the derivative of F_D in Eq. (2) with respect to the coefficients c_1, c_2, c_3, and h depends on the model coefficients, making the gradients period dependent.

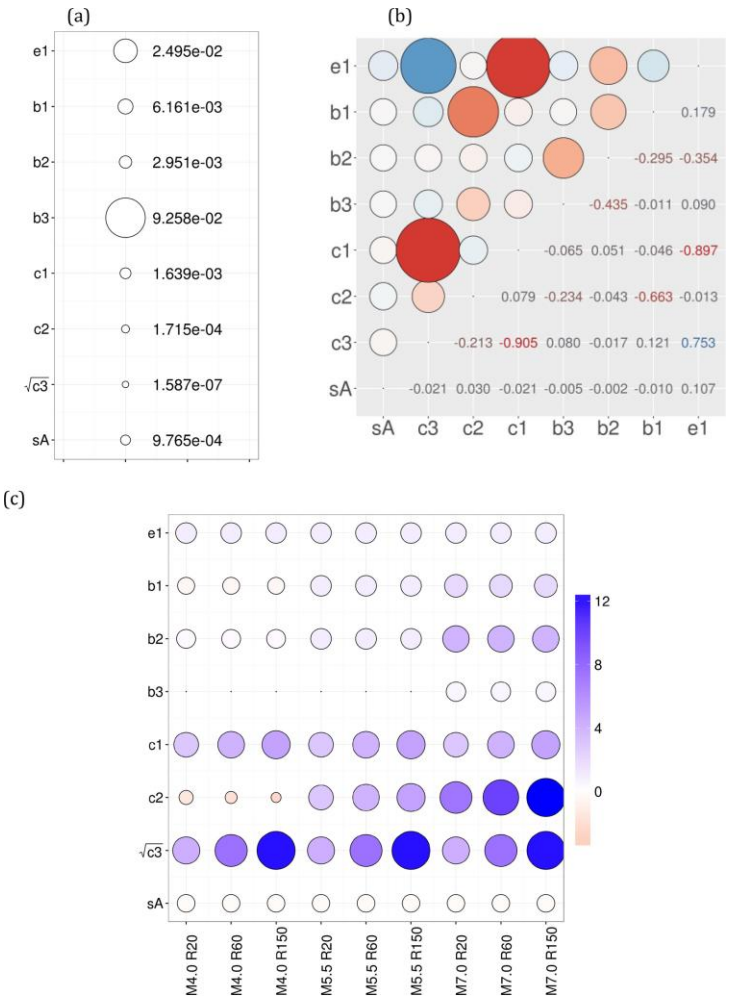


Figure 2-5: Variance, correlation and gradient of coefficients. (a) Variance of the coefficients for $T = 0.1$s, considering the model for hypocentral distance. (b) Correlation of the coefficients at 0.1s, evaluated at the data points used to develop the model. (c) each column includes the gradient of the model with respect to one coefficient (from e_1 to sA, see equations 2 through 5), evaluated at a given magnitude and distance (for example, the column M 4.0 R20 is evaluate for magnitude 4 at 20km, M7.0 R60 for magnitude 7 at 60, and so forth). The v_{S30} velocity is fixed to 600 m/s. Please note that the square root of variance for c_3 (panel a) and the square root of the derivative with respect to c_3 (panel c) are considered.

Figure 2-6 shows σ_μ for different magnitude and distance scenarios. Panels (a) and (b) show that σ_μ for magnitude 6 and $v_{S30} = 800$m/s (i.e., a typical scenario of interest for our application) is weakly dependent on distance and is of the order of σ_μ modelled for NGA2 GMPEs (*Al Atik and Youngs*, 2014). For distances shorter than 10 km, σ_μ slightly increases with decreasing distance, reflecting the fact that short distances are weakly constrained by data. In the magnitude range from 4 to 6 (and for a distance of 25 km and $V_{S30} = 800$m/s), σ_μ is weakly dependent on magnitude, while it increases outside this range, in particular above magnitude 7. The bump around magnitude 6.5 is a consequence of introducing the hinge magnitude for handling the saturation with magnitude. The overall dependence on period is weak, as shown in panel (c). For magnitude 6, σ_μ is between the models for normal and reverse faulting derived for NGA2 while, as already shown in panel (b), larger values are obtained for magnitude 8. The increases of σ_μ for periods longer than 2s is stronger for small magnitudes, which less constrain the ground-motion at low frequencies. Regarding the dependence of σ_μ on v_{S30} (here not shown), it is negligible. Finally, the overall contribution of σ_μ to the mean response spectra uncertainty are shown in panel (d), for different magnitudes, at a distance of 25 km and for $v_{S30} = 800$m/s.

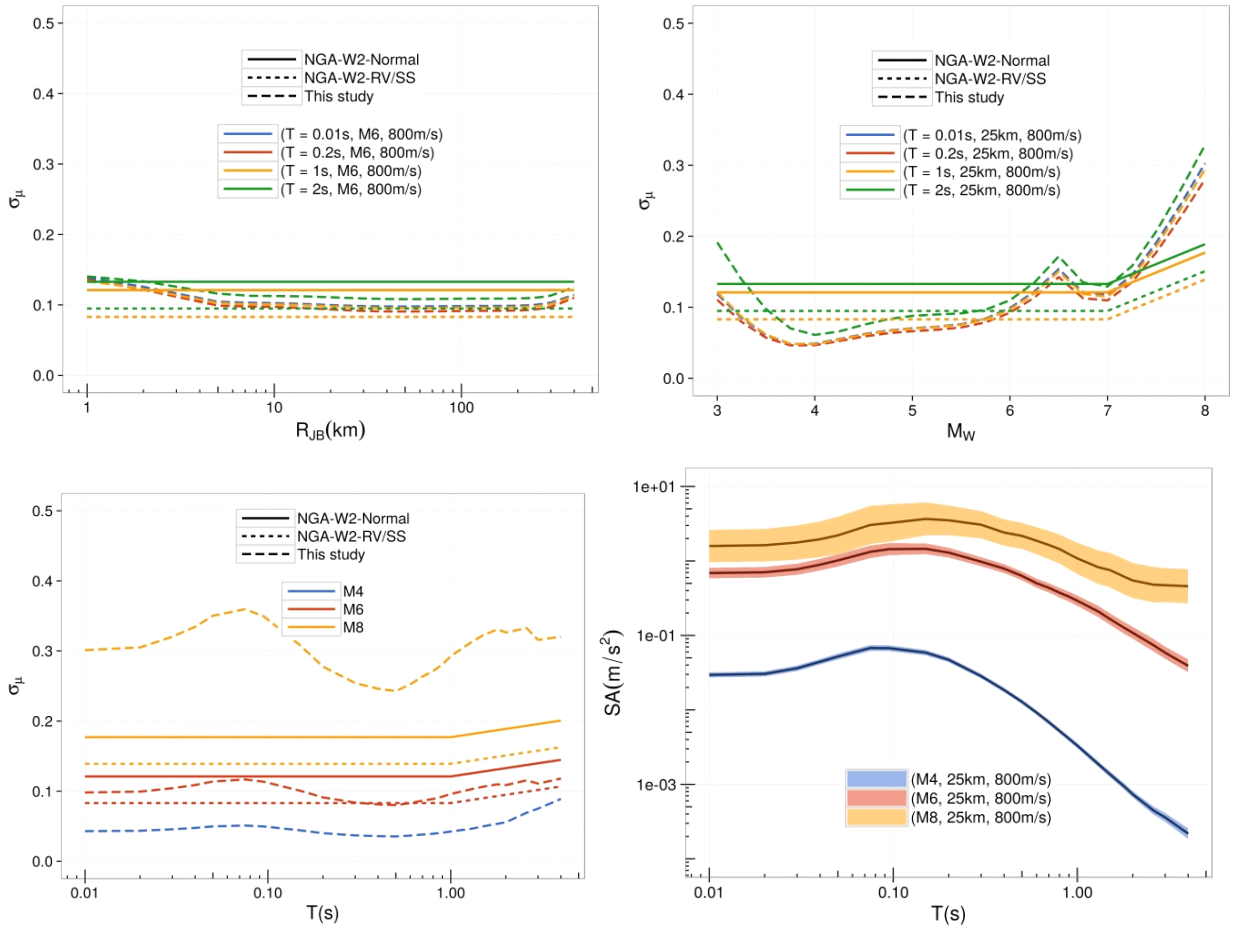


Figure 2-6: Uncertainty in the mean computed for the Joyner-Boore model. (a) Dependence of σ_μ on distance, for different periods (colors); (b) Dependence of σ_μ on magnitude, for different periods; (c) Dependence of σ_μ on periods for different magnitude (colors) at a distance of 25 km; (d) overall effect of σ_μ on the mean response spectra, for three different magnitudes at a distance of 25 km (V_{s30}=800 m/s). In panels (a), (b), and (c) the continuous and dotted lines represent the models proposed by Al-Atik and Youngs (2014) for NGA2 for normal and reverse or strike faults, respectively.

2.2.3 Aleatory variability

Figure 2-7 shows the aleatory variability in terms of period dependent τ (between-event), ϕ (within-event) and σ (total aleatory variability). To provide a term of comparison, Figure 2-7 also reports the models for BSSA14 (*Boore et al.*, 2014) and RES14 (*Bindi et al.*, 2014). Since BSSA14 is heteroscedastic, its standard deviations are evaluated for magnitudes 4 and 7 (representing the range of main interest for the hazard application in Germany), at a distance of 40km, and for $v_{S30} = 800$m/s.

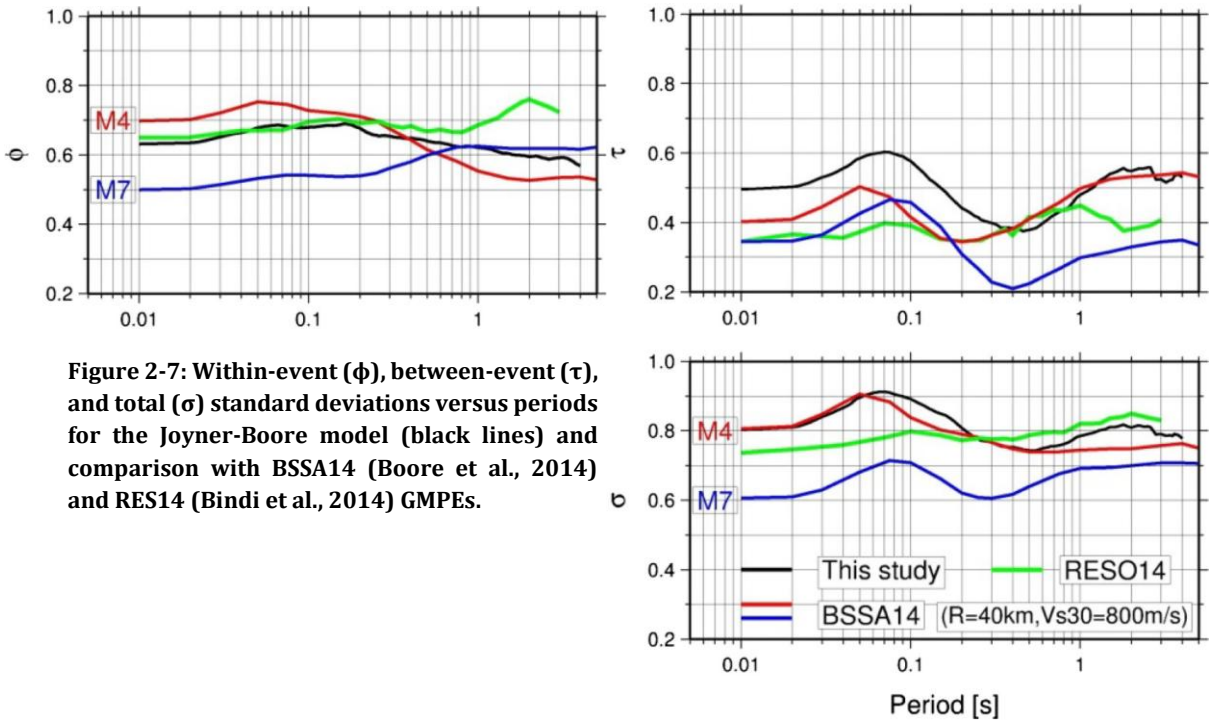


Figure 2-7: Within-event (ϕ), between-event (τ), and total (σ) standard deviations versus periods for the Joyner-Boore model (black lines) and comparison with BSSA14 (Boore et al., 2014) and RES14 (Bindi et al., 2014) GMPEs.

Considering the simplification applied to the functional form, σ values larger than those of BSSA14 were expected. Indeed, for short periods, σ is very close to the values of BSSA14 for $M4$ while, for longer periods, it approaches the values of RES14 (which was calibrated for Europe). The main differences among the models are observed for τ. For periods shorter than 0.3s, τ is larger than the values of the other models while for longer periods it overlaps to BSSA14 evaluated for magnitude 4. The values for RES14 are smaller, probably as a consequence of a more regional composition, since it is manly composed by earthquakes occurred in Italy and Turkey. Regional differences are also present in the European data but they are mainly affecting the distance scaling and site effects (e.g., *Kotha et al.*, 2016). The largest contribution to σ is coming from ϕ. Below 0.2s, ϕ from this study is similar to RES14 while, for longer periods, it is smaller than RES14 and close to the BSSA14 one evaluated for magnitude 7. In conclusion, the aleatory variability of the derived model is close to the variability of BSSA14 for low magnitude and, since we do not allowed σ to be heteroscedastic, larger than the BSSA14 one for large magnitude. The simplifications applied to the functional form mainly affect the source variability for periods below 0.4s and the record-to-record variability for the longer periods.

2.3 Discussions

In the following, the derived model is discussed in terms of residuals analysis and by comparing the predictions with those from previous models.

2.3.1 Analysis of residuals

The explanatory power of the models is evaluated through the analysis of the residual distributions. Figure 2-8 shows the prediction versus distance for PGA and PGV, considering magnitudes 4 and 6.7, and $v_{S30} = 800$m/s. The predictions are compared to observations selected in ± 0.2 range with respect the magnitude used for the prediction, and considering all available stations (circles for stations with $v_{S30} \geq 800$m/s, triangles for $v_{S30} < 800$m/s).

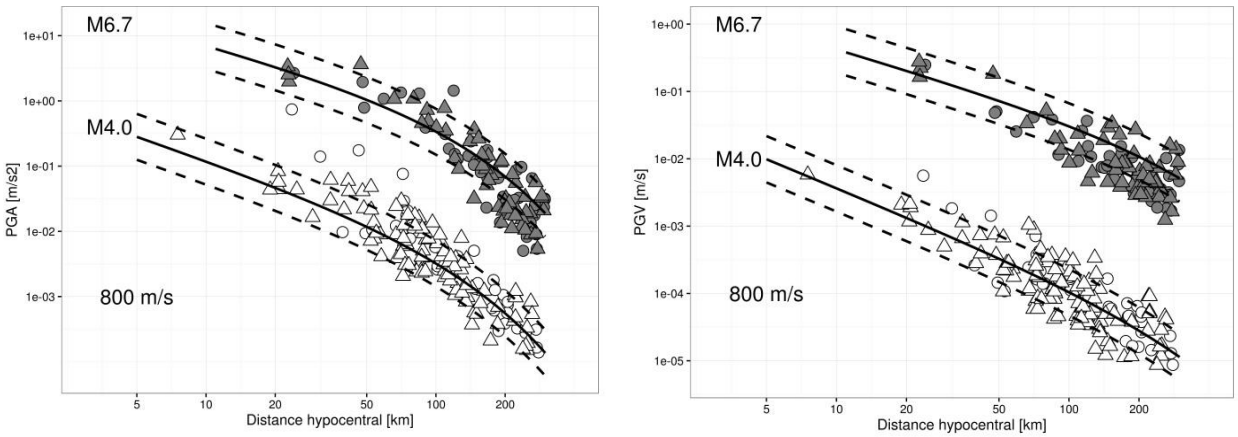


Figure 2-8: Prediction versus data for PGA (a) and PGV (b), for the hypocentral distance model. The median (continuous lines) ± 1σ (dashed lines) are computed for magnitude 4 and 6.7, and vs30=800 m/s. Symbols are observations for M=4.0±0.25 (white) and 6.7±0.25 (gray), considering stations with vs30≥800 m/s (circles) and <800 m/s (triangles).

The model captures well the trend and the variability in the data, with perhaps the tendency of overestimating the ground-motion for large magnitude at large distances. To quantify the overall agreement between data and predictions, the within and between residuals are shown in Figure 2-9 against distance and magnitude, respectively. Generally, the within event residuals (left panels) for both selected periods (i.e., 0.1 and 1.0s) do not show any significant trend with distance, except for a slight tendency to underestimate the spectral acceleration at distance smaller than 20 km for $T = 0.1$s. Regarding the between event residuals (right panels), they are shown accordingly to their focal mechanism. The data set is dominated by strike-slip (SS) events (181 earthquakes), shown as black circles, while the number of events with normal (NF) and reverse (RF) mechanisms are 16 and 45, respectively. While the between event distribution is unbiased when considered as a whole, the average residuals computed separately for the three style of faulting classes [SS, NF, RF] are [0.007, -0.116, 0.013] at 0.1s, and [0.005, -0.096, 0.009] at 1s. Therefore, the model tends to slightly overestimate the spectral acceleration for normal events, although the large standard deviations (of the order of 0.4) make all these values not significantly different from zero.

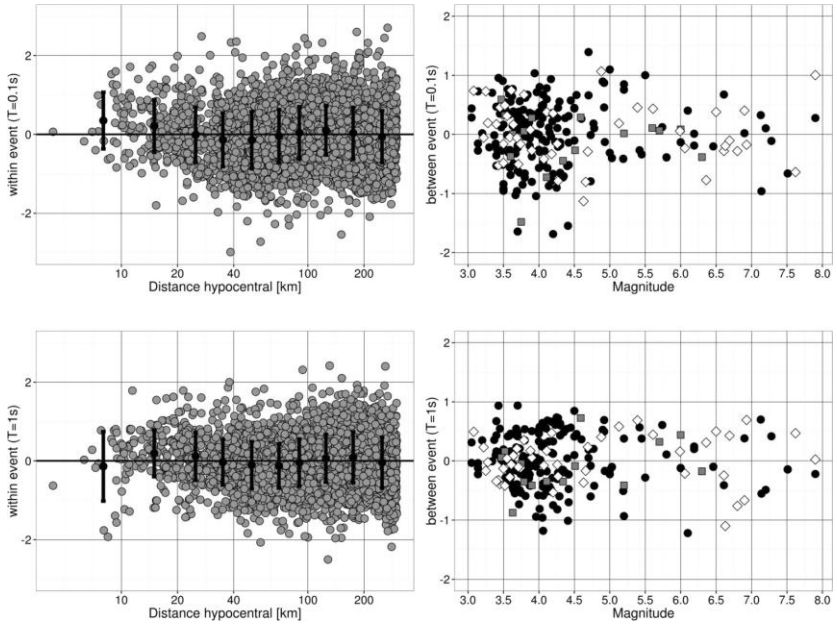
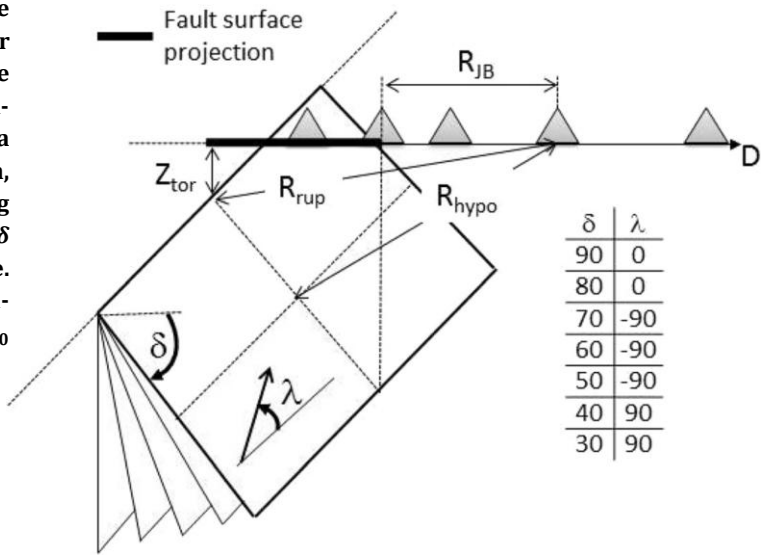


Figure 2-9: Within-event residuals versus distance and magnitude. Top. Within event residuals versus distance (left) and between event residuals versus magnitude (right), for spectral acceleration at 0.1s. The symbols in the between event plot indicate different focal mechanisms (circle: strike slip; square: normal; diamonds: reverse). In the bottom panels, the same distributions are shown but for 1s.

2.3.2 Comparison with NGA2 and RES14 models

The median predictions of the model calibrated in this study are compared to four different NGA2 models and to one European model. The models considered are: BSSA14 (*Boore et al.*, 2014), CY14 (*Chiou and Youngs*, 2014), ASK14 (*Abrahamson et al.*, 2014), CB14 (*Campbell and Bozorgnia*, 2014), RES14 (*Bindi et al.*, 2014). The *Idriss* (2014) model is not used here because its application is suggested above magnitude 5 which does not fit with the needs of hazard computations in moderate seismicity areas. Both the hypocentral and Joyner-Boore versions of the model derived in the present study are discussed here. The comparison is performed in terms of Sammon's map (*Scherbaum et al.*, 2010) and Trellis plot. Since the implemented GMPEs use different distance definitions, the comparison is performed for a set of a-priori defined source scenarios (Figure 2-10).

Figure 2-10: Geometries used to generate the scenarios for the Sammon's map. For each source, the distances required by the considered GMPEs are computed for six stations located at different distances along a line perpendicular to the strike direction, being all stations located on the hanging wall. The combinations between the dip δ and rake λ angles are given in the figure. For each combination, four different magnitudes are generated (i.e., 4, 5, 6, and 7). v_{S30} is fixed to 800m/s.



In particular, 7 different fault configurations are adopted varying the dip δ and the rake λ angles (Figure 2-10 shows the list of the 7 considered combinations). For each fault model, 4 different magnitudes are selected to generate the fault extension (i.e., 4, 5, 6, and 7). Regarding the station locations, they are arranged along a line orthogonal to the strike, located over the hanging wall at distances equal to Δ=[0.01, 0.1, 0.2, 0.5, 1, 2] degrees. For all stations, $v_{S30} = 800$m/s is used. For each models, those parameters of NGA2 GMPEs like Z_{tor}, $Z_{2.5}$, Z_{10}, etc., are set equal to default values suggested by the GMPE's authors and no regional attributes are considered. For each source-station combination, the distances required by the GMPEs are computed (i.e., rupture distance R_{rup}, Joyner-Boore distance R_{JB}, and hypocentral distance R_{hypo}, being the latter computed locating the hypocenter in the middle of the fault). In total, 168 source-station combinations are generated, which are used to compile the multi-dimensional vectors for the Sammon's maps. In order to provide a reference in the Sammon's map, the mixture of the 4 considered NGA2 GMPEs is computed with equal weights, and indicated with MIX in. Moreover, artificial scaling with distance and magnitude are applied to MIX, to add further reference points in the maps. In particular, M+ and M++ in Figure 2-11 refer to MIX with added the term $0.25(M\text{-}M_{ref})$ and $0.5(M\text{-}M_{ref})$, respectively. Similar definitions apply to M- and M--, but the artificial scaling is in this case subtracted. Regarding the distance scaling, R+ and R- correspond to adding $0.25\ln(R_{rup})$ or subtracting $0.25\ln(R_{rup})$ to MIX, respectively. To make it easier the comparison between different maps, we have applied translation and rotation to the Sammon's maps in order to locate MIX always in the origin of the coordinate system and R+ along the positive x-axis (Figure 2-11). If necessary, a reflection with respect to the x-axis is finally applied.

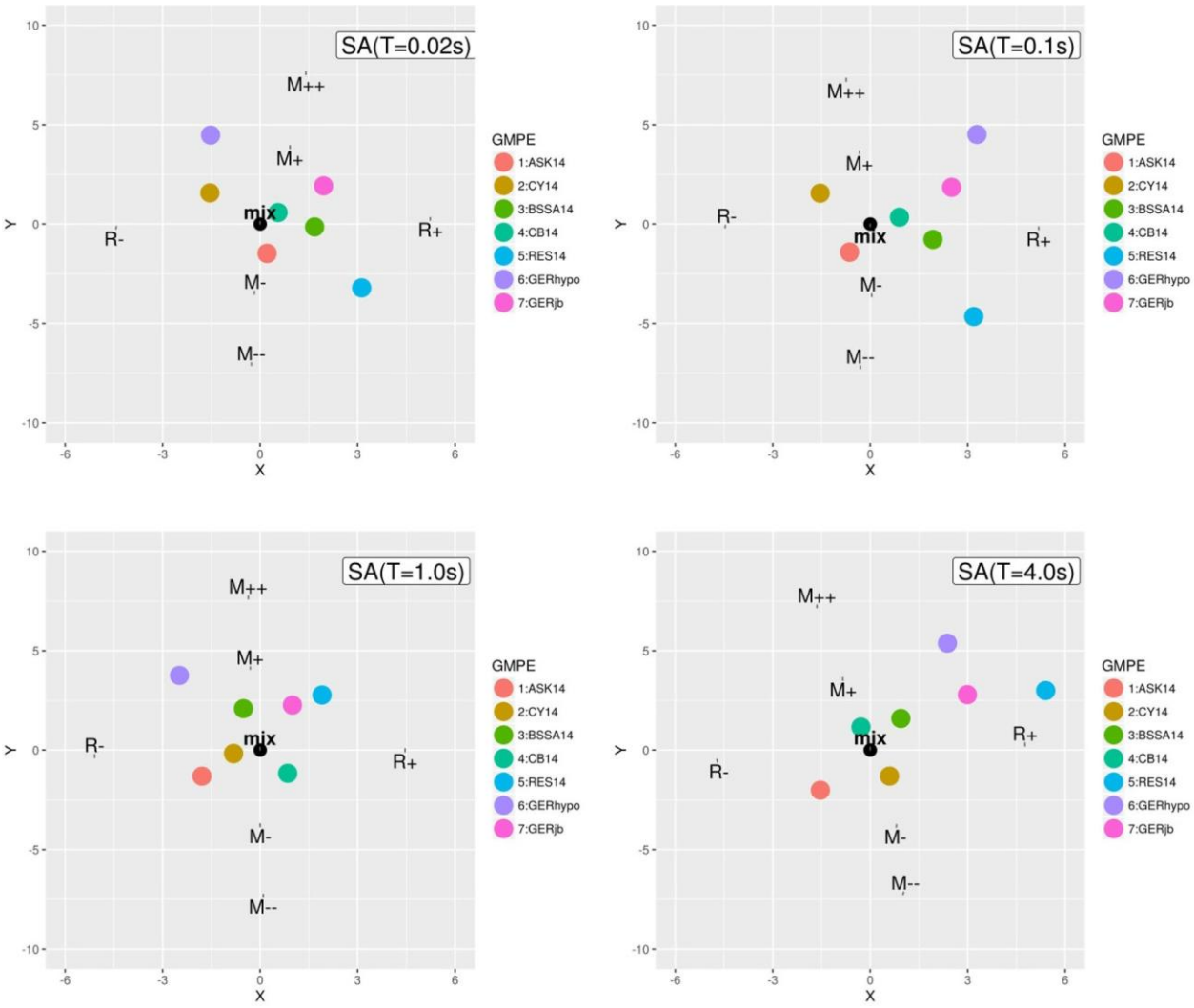


Figure 2-11: Sammon's maps for T = 0.02, 0.1, 1, and 4s, considering the source scenarios described in Fig. 10. The considered GMPEs are represented by colored circles while MIX is the reference model computed as mixture of the four considered NAGA2 GMPEs with equal weights. M+, M++, M−, M—, R+, and R− are GMPEs where artificial either magnitude or distance scaling are added to MIX (see text for details).

The inter-point distances in Figure 2-11 are an estimate of the GMPEs proximity in predicting similar ground-motion levels (*Scherbaum et al.*, 2010), assessed from the feature vectors constructed considering the scenarios in Figure 2-10. Considering the distances among the NGA-west2 GMPEs shown in Figure 2-11, the model derived in this study for the Joyner-Boore distance (GERjb) and for periods up to 1s, is close to the NGA-west2 group. On the contrary, GERjb for longer periods (i.e., 4s in Figure 2-11) and the hypocentral model (GERhypo) for all periods predict significant different ground-motion. In particular, GERhypo and GERjb show a stronger magnitude scaling with respect to MIX, being the differences more evident for GERhypo. This is confirmed by the magnitude scaling shown in Figure 2-12, in particular for short distances and $T = 0.1$s. Regarding the scaling with distance, the Sammon's maps show a stronger decay of Rhypo than MIX (i.e., Rhypo is along the R- direction) at $T = 0.02$ and 0.1s while the attenuation is weaker (R+ direction) at 0.1 and 4s. The Trellis plots in Figure 2-13 confirm these overall trends, although with differences depending on specific scenarios (e.g., at $T = 0.1$ s for M 6.5, Rhypo shows a stronger attenuation than MIX). The Sammon's maps also provide information for the other models. For example, RES14 shows a weaker magnitude scaling than MIX for 0.02 and 0.1s (see also Figure 2-11) and a general weaker attenuation with distance than MIX (see also Figure 2-13). It is worth noting that 4s is beyond the range of applicability of RES14 suggested by the authors.

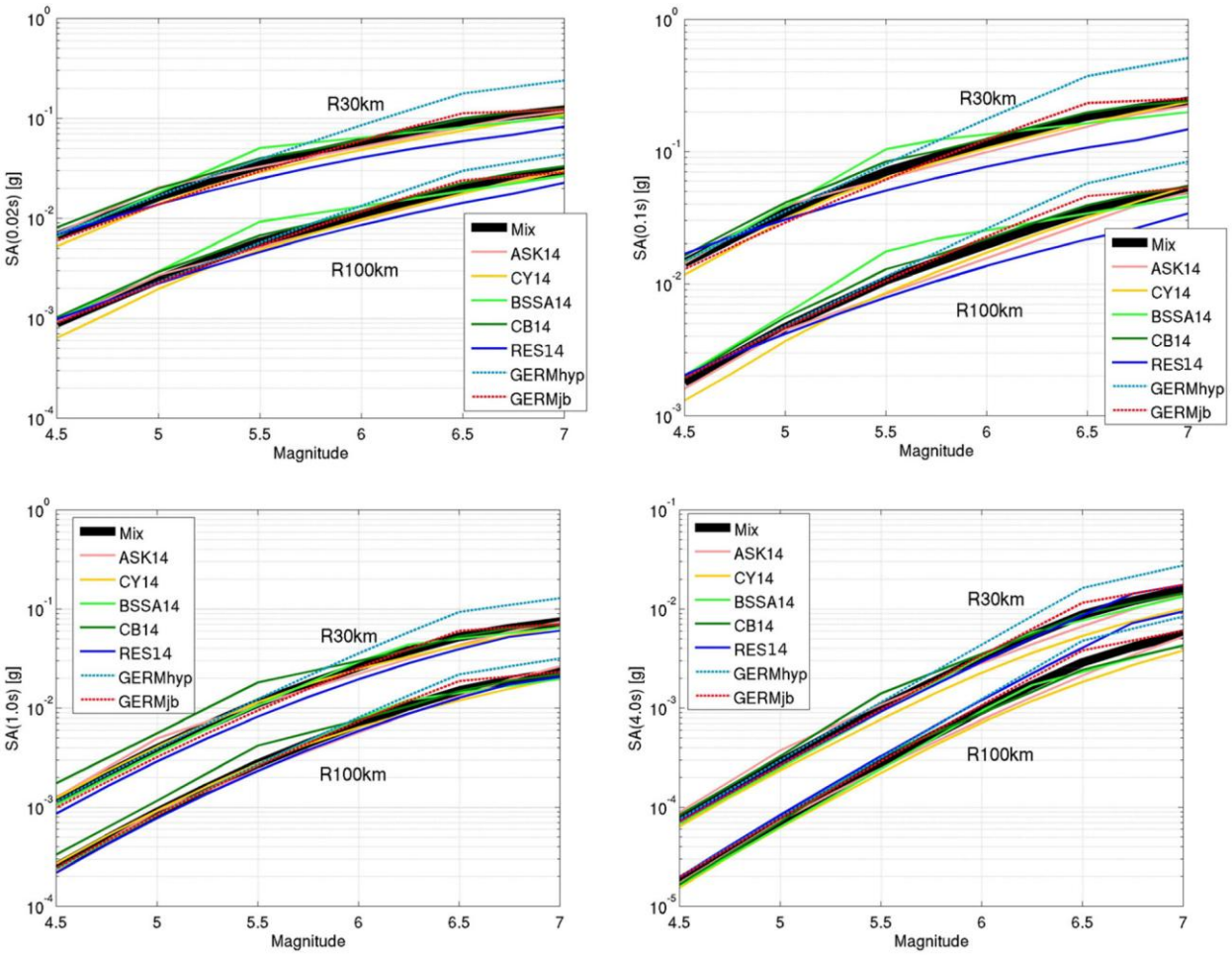


Figure 2-12: Magnitude scaling for different GMPEs at T = 0.02, 0.1, 1.0, and 4.0s, considering two different distances (i.e., 30 and 100km), V_{s30} = 800m/s, and a vertical strike slip earthquake. The model MIX is the mixture of the four considered NGA2 models considered with equal weights (see text for details).

Figure 2-14 summarizes the results of the Sammon's map analysis, showing the distance between each considered model and the reference one (i.e., the mixture of the NGA2 models), as function of periods. The predictions from the model derived in this study for the Joyner-Boore distance are close to the NGA-west2 ones, in particular for short periods. For long periods, the Joyner-Boore model is close to CY14 and BSSA14 while between 0.3 and 2s, the predictions are closer to the pan European model (RES14). Figure 2-14 confirms that the hypocentral model derived, where a point source is considered for computing the distances, shows larger differences in the prediction of the ground shaking for the considered scenarios.

2.4 Conclusions

Motivated by its application in the update of the seismic hazard assessment for Germany, we developed a ground-motion prediction equation (GMPE) in this study tailored to such specific needs. Starting from the high quality NGA-west2 flat file, we constructed our model taking into consideration the requirements from the specific hazard application in a low-to-moderate seismicity area, being the following the main ones: a model implementing a point source measure of distance (i.e. hypocentral) along with an extended source metric (Joyner-Boore in our case); develop a model for a reference rock condition of v_{S30} = 800m/s, avoiding possible bias due to low velocities (see also the discussions in *Idriss*, 2014); a GMPE with a smooth magnitude scaling around magnitude 5.5, which control the hazard at short return period in the target area; a complexity of the model suitable for its application in a low-to-moderate seismic area, that is, a functional form not requiring a-priori assumptions of variables not known in the target area, that would imply additional assumptions for refining the aleatory variability model.

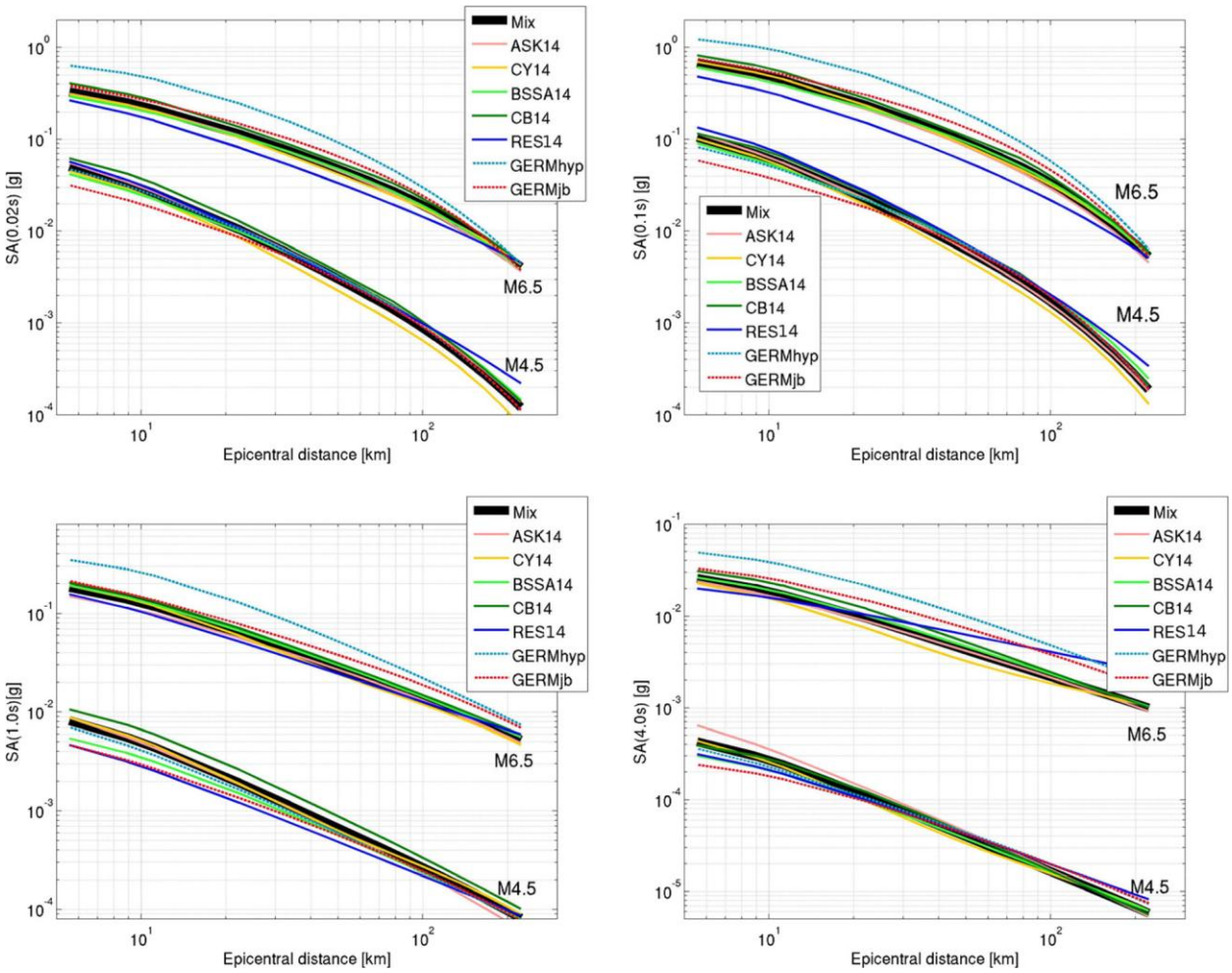


Figure 2-13: Distance scaling for different GMPEs at T = 0.02, 0.1, 1.0, and 4.0s, considering two different magnitudes (i.e., 4.5 and 6.5), V_{s30} = 800m/s, and a vertical strike slip earthquake. The model MIX is the mixture of the four considered NGA2 models considered with equal weights (see text for details).

The simplification in the functional form with respect to the NGA-west2 GMPEs had the effect of increasing σ. Indeed, the obtained values of σ are close to the NGA-west2 values for small magnitudes and periods shorter than about 0.6s whereas, for longer periods, σ increases to the values observed for the *Bindi et al.* (2014) Pan European model (RES14). This increased variability level for large magnitudes hamper the applicability of the model derived in this study for those applications where long mean return periods are of concern, such as site-specific hazard assessments. On the other hand, the followed approach can be of interest for many other applications, such as shake maps or earthquake early warning, as well as for the development of GMPEs for new intensity measure (e.g. *Koufoudi et al.*, 2015). The comparison between the median predictions with those from the NGA-west2 and RES14 models in terms of Sammon's map shows that the predictions from our GMPE derived for the Joyner-Boore distance are closer to the NGA-west2 ones than the RES14 model whereas, for the GMPE implementing the hypocentral distance, larger differences are observed in the magnitude and distance scaling. Furthermore, the analysis of the covariance matrix shows that the epistemic uncertainty in the median of the model calibrated for the hypocentral distance, controlled by both the functional form and the data availability, is of the order of those affecting the NGA-west 2 models for magnitudes smaller than 7.5.

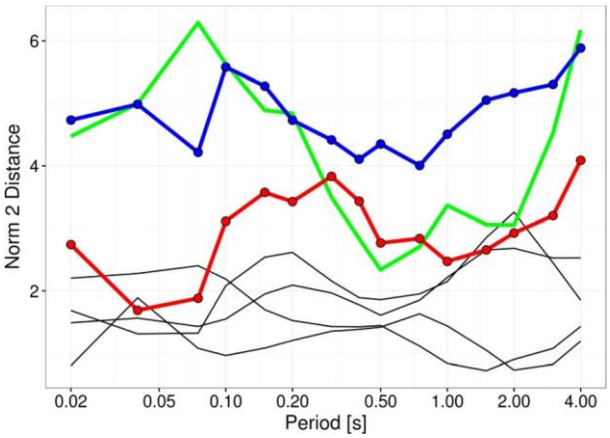


Figure 2-14: Period dependence of the distance in the Sammon's plane, between each considered model in Fig. 11 and the reference one (MIX). The results for the NGA2 models are shown in black, RES14 in green, while the models derived in study for the Joyner-Boore and hypocentral distances are shown in red and blue, respectively.

In conclusion, the suggested ranges of applicability for the GMPE derived in this study for the Joyner-Boore distance are between magnitude 3 and 8, distances shorter than 300 km and v_{S30} larger than 360m/s (i.e., class A and B of Eurocode 8); for the hypocentral distance GMPE, the suggested ranges are from 10 to 300 km and magnitudes between 3 and 7.

Acknowledgments

The need for the study summarized in this paper was born in the frame of the research project partly financed by the DIBt (https://www.dibt.de/) for the reappraisal of the seismic hazard of Germany for the National Annex to the updated Eurocode 8. We thank two anonymous Reviewers and the Editorial board who helped us to improve our work. Scripts downloaded from the Baker Research Group in Stanford (http://stanford.edu/~bakerjw/ GMPEs.html, last accessed May 2016) were used to compute the predictions from the NGA2 models; we thanks the authors for sharing their programs. The R software (*R Development Core Team*, 2008; http://www.r-project.org, last accessed May 2016) has been used in this study to perform the regressions. In particular, the package lme4 (*Bates et al.*, 2015; https://cran.r-project.org/web/packages/lme4/news.html, last accessed May 2016) has been used for mixed-effect regressions. The NGA2-West flat file is available at the PEER web page (http://peer.berkeley.edu/ngawest2/final-products/, last accessed May 2016).

3
Partially Non-Ergodic Region-Specific GMPE for Europe and Middle-East

S. R. Kotha[1, 2], D. Bindi[1], F. Cotton[1, 2]

[1]German Research Centre for Geosciences GFZ, Potsdam, Germany

[2] University of Potsdam, Potsdam, Germany

Abstract

The ergodic assumption considers the time sampling of ground shaking generated in a given region by successive earthquakes as equivalent to a spatial sampling of observed ground-motion across different regions. In such cases the estimated aleatory variability in source, propagation, and site seismic processes in ground-motion prediction equations (GMPEs) is usually larger than with a non-ergodic approach. With the recently published datasets such as RESORCE for Europe and Middle-East regions, and exploiting algorithms like the Non-Linear Mixed Effects Regression it became possible to introduce statistically well-constrained regional adjustments to a GMPE, thus 'partially' mitigating the impact of the assumption on regional ergodicity. In this study, we quantify the regional differences in the apparent attenuation of high frequency ground-motion with distance and in linear site amplification with V_{S30}, between Italy, Turkey, and rest of the Europe-Middle-East region. With respect to a GMPE without regional adjustments, we obtain up to 10% reduction in the aleatory variability σ, primarily contributed by a 20% reduction in the between-station variability. The reduced aleatory variability is translated into an epistemic uncertainty, i.e. a standard error on the regional adjustments which can be accounted for in the hazard assessment through logic-tree branches properly weighted. Furthermore, the between-event variability is reduced by up to 30% by disregarding in regression the events with empirically estimated moment magnitude. Therefore, we conclude that a further refinement of the aleatory variability could be achieved by choosing a combination of proxies for the site response, and through the homogenization of the magnitude scales across regions.

Keywords: Ground-Motion Prediction Equations, Europe and the Middle-East, RESORCE, Regional variations, Non-ergodicity, Nonlinear Mixed Effects Regression

3.1 Introduction

Reliability of the ground-motion predicted by empirical models mostly depends on the characteristics of the underlying calibration dataset. In the framework of seismic hazard assessment, the motivation behind compilation of a large strong motion dataset which includes recordings from different regions is twofold: first, to improve the magnitude-distance data distribution, and sampling different source characteristics and site conditions; second, to allow the calibration of models complex enough to describe the main physical processes contributing to the variability of the ground-motion. The current practice for computing the seismic hazard is based on an ergodic assumption, where the aleatory variability, i.e. the standard deviation sigma (σ) of ground-motion prediction equation (GMPE), includes the regional differences in ground-motion. If on one hand the ergodic assumption allows to replace the time sampling of ground shaking generated in a given region by successive earthquakes with a spatial sampling of ground shaking observed across different regions, on the other hand it increases the aleatory variability associated with source, propagation, and site seismic processes. Allowing regional differences in the GMPE 'partially' removes this ergodicity by translating the aleatory variability into epistemic uncertainty which, in statistical sense, is the modelling uncertainty in region-specific adjustments.

The collection of data from different regions with similar tectonic features (e.g. shallow crustal active regions, stable continental regions, etc.) was performed in the past under the assumption that the trans-regional and between-country variability of the ground-motion was either negligible or otherwise difficult to model due to the limitation in the sampling properties of the compiled datasets e.g. *Douglas* (2004a, 2004b). As an example, the NGA-West models (*Abrahamson and Silva*, 2008a) were derived from a dataset including recordings from multiple regions (mainly California, Taiwan, Japan) without modelling the regional effects. Later studies on the applicability of the NGA models to Europe (e.g. *Stafford et al.*, 2008) highlighted the general agreement between predicted median values and the observations. The main difference was a faster distance attenuation observed in European data with respect to California; in agreement with previous findings (*Douglas*, 2004a). Moreover a detailed comparison between the NGA models and strong motion data recorded in Italy (*Scasserra et al.*, 2009) confirmed that it was possible to improve the predictive performance of NGA models for Europe by applying regional corrections to the attenuation with distance terms and to the overall scaling parameters (offset and pseudo-depth).

Extension of the NGA database into NGA-West2 (*Ancheta et al.*, 2014) with introduction of several small magnitude events mainly from California, and moderate to large size earthquakes from other regions of the world, promoted the interest in evaluating regional effects in the ground-motions. As a consequence, the most recent GMPEs developed from NGA-West2 include correction terms accounting for regional effects. Many authors (e.g. *Boore et al.*, 2014; *Chiou and Youngs*, 2014) introduced regional differences in the anelastic attenuation coefficient and the site term related to depth of basin. Regional differences in the Vs30 scaling were also considered (e.g. *Abrahamson et al.*, 2014), while information available in the dataset is not enough to constrain correction factors for other parameters.

RESORCE strong motion database (*Akkar et al.*, 2014a) was compiled with recordings from different European and Middle-East countries, and was used to derive several GMPEs (*Douglas et al.*, 2014). While these models do not account for regional differences in ground-motion scaling, recent studies highlighted the presence of regional effects either between selected countries (e.g. between Turkey and Iran by *Kale et al.*, 2015), or among different tectonic regions in Europe (*Gianniotis et al.*, 2014). Ignoring the regional differences in ground-motion scaling may result in an inflated residual standard deviation, and correction for regional bias in the median ground-motion can be a first step towards 'partially non-ergodic' region-specific PSHA. With such a goal in mind, this study focuses on identification of systematic regional differences in ground-motion scaling in Europe. Following the previous efforts of developing GMPE using RESORCE dataset (*Douglas et al.*, 2014 and reference therein), we derive a new GMPE based on a relatively simple functional form which will still be able to capture the main features of ground-motion-scaling (*Bindi et al.*, 2014). However, unlike in previous studies, a non-linear mixed effect regression (NLMER by

Bates et al., 2014) approach is applied where the regional differences are estimated as random effects applied to different model parameters. The advantages of using NLMER in place of the traditional random effect algorithm by *Abrahamson and Youngs* (1992) are discussed by *Stafford* (2014). For example, group specific adjustments can be estimated for any of the regression coefficients in a statistically correct way making NLMER much more extendable than traditional approaches. We identify the statistically significant random effects and the regional adjustments for relevant parameters are provided as final result.

3.2 Dataset and selection criteria

The most recent Pan-European GMPEs (*Douglas et al.*, 2014) are based on the RESORCE strong motion dataset (http://www.resorce-portal.eu/). RESORCE extends the previous pan-European strong motion dataset (*Ambraseys et al.*, 2004) with recently compiled Greek, Italian, Swiss and Turkish accelerometric archives (*Akkar et al.*, 2014a). In this study, starting from the 2013 release of RESORCE, we performed a preliminary data selection to exclude the poor quality or unprocessed records, or those records lacking the three components of ground-motion; then, we applied the following criteria to select the input data for regression:

- Given the recent interest in considering small magnitude earthquakes for assessing the hazard in several regions of Europe (http://projet-sigma.com/ScientificObjectives.html), records from events with moment magnitudes larger than or equal to 4 are considered.
- Only focal depths shallower than 35km, and distances (Joyner-Boore, R_{JB}, or epicentral R_{epi}) shorter than 300km are selected. The epicentral distance, R_{epi} is used to approximate R_{JB} when the latter is unspecified, but only when $M \leq 5$ and $R_{epi} \geq 10$ km. For larger magnitudes and smaller epicentral distances, records without R_{JB} are disregarded.
- For each oscillator period T, only those recording filtered with high pass corner frequency (f_{hp}) smaller than or equal to $1/(1.25\,T)$, i.e. $f_{hp} \leq 0.8\,f_{oscillator}$ (*Abrahamson and Silva*, 1997). For example, for $T = 1$s ($f_{oscillator} = 1$Hz), we considered only recordings with f_{hp} smaller than or equal to 0.8Hz; for $T = 4$s ($f_{oscillator} = 0.25$Hz), we chose $f_{hp} \leq 0.2$ Hz. Single recorded earthquakes are not selected.
- We consider only recordings from sites with known or inferred V_{S30}.

In RESORCE, the moment magnitude is provided either as directly computed (e.g. from the moment tensor solutions), or converted from other magnitude scales (e.g. local magnitude or surface wave magnitude) using country-based empirical regressions (see *Akkar et al.*, 2014a, for details). Earthquakes with M_w derived through empirical regressions are not considered in this study. Considering the unbalanced composition of the dataset, we categorize the contributing regions into three groups: Italy, Turkey, and Others, where the latter collects data from all the countries contributing to RESORCE with less than 200 selected records. Although a regionalization based on the tectonic settings (e.g. *Delavaud et al.*, 2012) could be more appropriate to explore regional differences in ground-motion, we opt for a country-based categorization that reflects the structure followed for data compilation. The filtered dataset is composed of 1251 recordings, with 659 recordings from Turkey (TR), 378 from Italy (IT), 214 in Others group; primarily contributed to by Greece, Montenegro, Iran, and France.

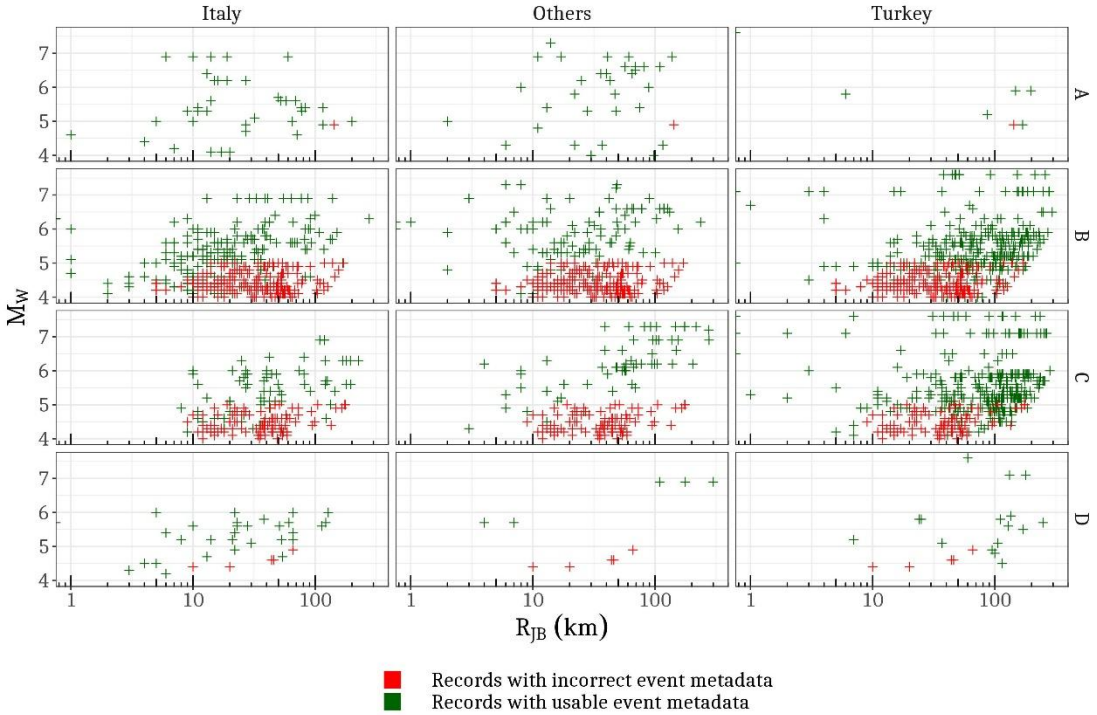


Figure 3-1: Scatter plot showing the distribution of observed data in Magnitude-Distance ranges for different EC8 site classes for each region Italy (IT), Others, and Turkey (TR). The red markers correspond to events without a computed moment magnitude but only an empirically estimated/converted moment magnitude, and consequently excluded from the regression. Green markers show the final distribution of records that are used for PGA regression.

In terms of magnitude range, distance range, and site characterization, the dataset is unbalanced among the regions. *Douglas* (2007) showed that the predicted median ground-motions are not well-constrained away from the centroid of data, especially for sparse datasets. Figure 3-1 shows the magnitude – distance distribution of recordings in our dataset, categorized according to different regions and soil classes. For example, there are very few recordings from Turkey in site class A (rock with $V_{S30} > 800$m/s), which means that when a GMPE is derived from the compendium dataset without regional distinction, the estimated site response for class A could be controlled by contributions from Italy and Other regions, even though the class A rock response in Turkey could be significantly different. Similarly, for distances larger than 100km and empirical site response of class B (stiff soil with 800m/s $> V_{S30} >=$ 360m/s) and Class C (soft soil with 360m/s $> V_{S30} >=$ 180m/s) the predictions could be controlled by strong motion recordings from Turkey. Moreover, preliminary non-parametric analysis (here not shown) suggest that the average slope of distance scaling is different among the regions, hinting for possible regional differences in the distance scaling of high-frequency ground-motions, which we could quantify as a regional variation during the GMPE regression. Based on these evidences, in the following we seek for ground-motion regional variations related to the scaling with distance and to the site response.

3.3 Regression approach

Different models were derived from RESORCE dataset performing either a parametric regression (e.g. *Akkar et al.*, 2014b; *Bindi et al.*, 2014) or following non-parametric approaches (e.g., *Derras et al.*, 2012; *Hermkes et al.*, 2014). The parametric regression approaches were applied using the random effects methodology of *Abrahamson and Youngs* (1992), where the residuals are split into between-event (δB_e), and within-event (δW_{es}) residuals. The GMPE functional forms used were relatively simple with respect to those implemented within the NGA-West2 project (e.g. *Abrahamson et al.*, 2014), reflecting the detail of information available in the RESORCE metadata. With the aim of investigating the presence of regional effects in ground-motion variability, we also follow a parametric regression approach but using a non-

linear mixed effect approach (NLMER, e.g. *Bates et al.*, 2014). Following *Bindi et al.* (2014), we consider the following functional form:

$$\ln(GM) = e_1 + F_D(R, M) + F_M(M) + \delta B_e + \delta B_s + \varepsilon \quad (1)$$

$$F_D(R, M) = \left[c_1 + c_2\left(M - M_{ref}\right)\right] \ln\left(\frac{\sqrt{R^2 + h^2}}{R_{ref}}\right) + \left(c_3 + \Delta c_{3,r}\right)\left(\sqrt{R^2 + h^2} - R_{ref}\right) \quad (2)$$

$$F_M(M) = \begin{cases} b_1(M - M_h) + b_2(M - M_h)^2 & \text{for } M < M_h, \text{where } M_h = 6.75 \\ b_3(M - M_h) & \text{for } M \geq M_h \end{cases} \quad (3)$$

In Eq. (1), e_1 is the global off-set parameter; F_D, and F_M are the distance and magnitude scaling components as defined in Eq. (2) and (3), respectively; δB_e and δB_s are random effects on e_1 describing the between-event and between-station variability, respectively (*Stafford*, 2014; *Al Atik et al.*, 2010); ε is the residual distribution accounting for the aleatory variability. In the following, the standard deviation of the between-event and residual distributions are indicated with the symbols τ and ϕ_0, respectively. The hinge magnitude M_h is fixed at 6.75 and the parameter b_3, which controls the saturation with magnitude, is not constrained to be positive (i.e. the over-saturation at magnitudes greater than 6.75 is allowed). As in *Bindi et al.* (2014), the reference moment magnitude M_{ref} and reference Joyner-Boore distance R_{ref} are set at M5.5 and 1km, respectively.

The major contributor to 'Others' group in terms of recordings is Greece (137), followed by Montenegro (35), and Iran (20). We performed several preliminary regressions considering different number of geographical categories, including attempts of isolating the Greek recordings from Others. In order to get reliable regional adjustments for the anelastic attenuation, a minimum number of recording per category (i.e., per country) was needed. Since the adjustment factor for Greece, once isolated from Others, was not significantly different from zero at 95% confidence interval, we kept the Greece recordings inside the Others category.

It is worth noting that we only introduced a regional adjustment factor for the apparent anelastic attenuation coefficient (i.e .c_3 in Eq. 1), but the magnitude scaling component (F_M in Eq. 1) is constrained by the data from all regions. When asked for a random-effect on a regression parameter (e.g. regional adjustment to c_3 in Eq. 1) for each level in the group (levels being Italy, Turkey, and Others), the NLME algorithm estimates scalar additive adjustments which follow a standard-normal distribution. Therefore, the GMPE regression-coefficient c_3 without any regional-adjustments (i.e. without adding $\Delta c_{3,\mathrm{r}}$ to c_3), is a generic anelastic attenuation coefficient without a regional bias.

3.3.1 Regional variability in apparent anelastic attenuation term

In Eq. (2), we introduce a country-based random effect $\Delta c_{3,r}$ on parameter c_3, where r represents the three selected regions, i.e., r = IT, Others, TR. Coefficients c_1, c_2 and c_3 in the scaling with distance F_D, correspond to the geometrical spreading, magnitude-dependent geometrical spreading, and apparent anelastic attenuation, respectively, although these names should be strictly used only for a model based on Fourier spectral amplitudes. Coefficient c_3 is constrained to being less than or equal to 0 for all spectral periods to disallow oversaturation at longer distances as in *Bindi et al.* (2014). Preliminary trials showed that for long periods ($>$ 1s), c_3 is taking a positive value and has a large negative correlation with c_1, and a positive correlation with e_1. Since a Student's t-test confirms that it is anyway losing is significance, c_3 and the associated regional variations are fixed at zero for periods longer than 1s.

3.3.2 Style of faulting terms

Dependence of the median ground-motion on style of faulting (SoF) is generally accounted through a period-dependent SoF specific adjustment to the median. Trial regressions including SoF adjustment factors

on the offset showed that the estimates were not well constrained and had large standard errors. In RESORCE the distribution of recorded focal mechanisms among different regions is strongly unbalanced since in Italy most of the events are normal and very few strike-slip events, unlike in Turkey. Moreover, reverse faulting events are very few in the dataset. Considering that the odd distribution of SoF among the regions could result in a trade-off with the regional random effects on the offset, and also based on a preliminary non-parametric analysis of the dataset that showed no clearly distinguishable differences among the distance scaling of ground-motion between different SoF, we chose to drop the SoF term from the functional form.

3.3.3 Regional variability in site-response as a function of V_{S30}

In the model described by Eq. (1), site effects are captured by the between-station terms, which account for the systematic station-specific deviation in offset with respect to the generic prediction for the population. Figure 3-2 shows that δB_s scales with V_{S30} indicating that V_{S30} is a first order proxy for describing site response. Large scatter around the best fit model suggests that a combination with other proxies is needed to better capture the complexity of site response (e.g. *Cadet et al.*, 2008; *Luzi et al.*, 2011). Besides the clear region-dependent scaling with V_{S30}, Figure 3-2 suggests that the distributions of velocities for three regions are compatible with the assumption of a region-dependent reference velocity (i.e., the value of V_{S30} corresponding to zero-crossing of δB_s). Hence, we perform a further mixed-effect regression considering the following model:

$$\delta B_s = \left(g_1 + \Delta g_{1,r}\right) + \left(g_2 + \Delta g_{2,r}\right) \ln(V_{S30}) + \delta S2S_s \qquad (4)$$

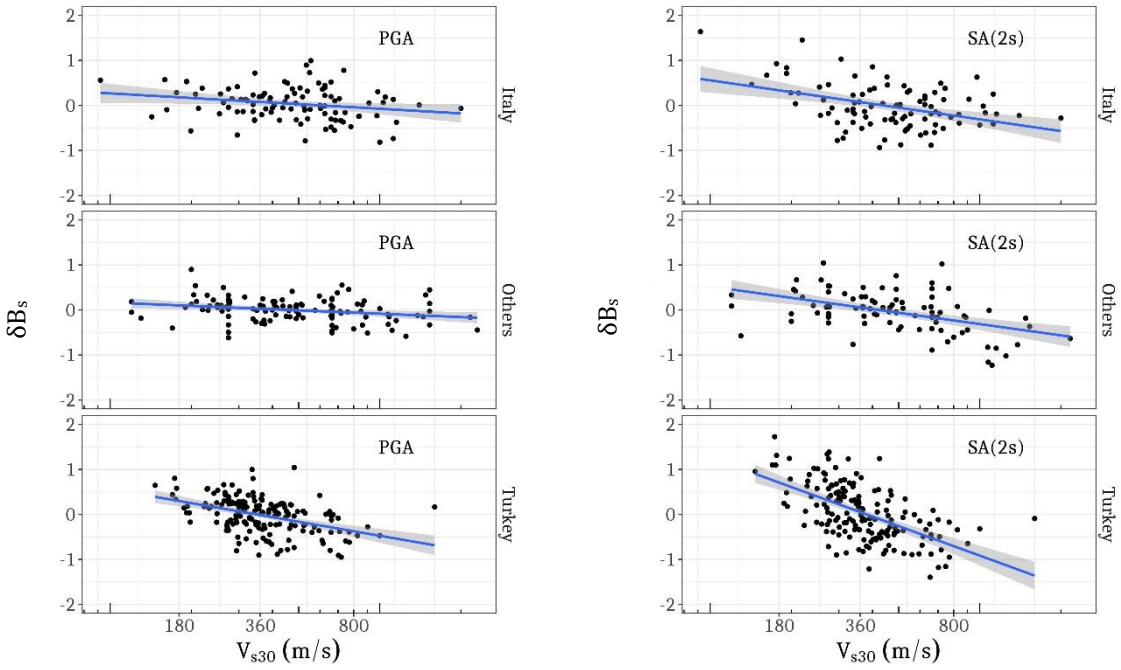


Figure 3-2: Between station residuals (PGA in left panel, SA (2s) in right panel) plotted against V_{S30} (m/s) with stations separated into regions. The blue line is a regression fit of residuals as a function $ln(V_{S30})$. The grey ribbon shows the standard error on regression fit. Difference in slope of the regression fit shows regional difference in linear site-amplification (g_2), difference in x-intercept shows the regional difference in reference V_{S30} (V_{ref}).

Regional effects on site term are captured by the random effects $\Delta g_{1,r}$ on offset g_1, and $\Delta g_{2,r}$ on the slope with V_{S30}. In Eq. (4), $\delta S2S_s$ represents the systematic deviation of recordings for individual station with respect to the model accounting also for the scaling with V_{S30}. The standard deviation of $\delta S2S_s$ is the between-station variability (ϕ_{S2S}) in the GMPE. It is worth noting that non-linear site amplification effects are not considered in the present study. Moreover, since the attenuation of high frequency ground-motion can be a result of both anelastic attenuation and site effects, it is worth checking for a possible correlation (or a trade-off) between the parameters c_3 and g_2, as well as between the estimated regional variations

$\Delta c_{3,r}$ and $\Delta g_{2,r}$. The results (here not shown) do not highlight any significant correlation among these parameters.

3.4 Results

The presence of ground-motion regional variations in RESORCE dataset are modelled by allowing the site response component and the decay of ground-motion with distance to be region specific. The fixed and random effects parameters relevant to regression (1) and (4) are listed in Table 3-1 and Table 3-2. At each period, the mixed effect regression provides both the global c_3 value and the estimated deviation $\Delta c_{3,r}$ for each region (r), computed as random effect on c_3 in a region group. Figure 3-3 shows the random effects at different periods along with the associated 95th percent confidence interval, the standard error (grey ribbon). Regional variations in c_3 are shown only until spectral period of 1s beyond which, along with c_3, they are constrained to zero. The apparent anelastic attenuation is higher for Italy than for Turkey or Others regions; a trend similar to that observed by *Boore et al.* (2014) in their within-event residuals which showed a faster distance-decay in Italy (and Japan) compared to Turkey (and China). The physical interpretation of the differences between the attenuation in Italy and Turkey is beyond the aim of our paper. A comparison of results available in literature for those physical properties that can influence the anelastic attenuation (e.g. velocity and attenuation topographic maps; heat flow distribution; etc.) is not straightforward because of the different implemented methodologies, the different investigated spatial scales, and the different data analyzed In any case the standard errors on $\Delta c_{3,r}$ are small enough to indicate that the regional corrections at short periods are statistically significant. These standard errors represent the modelling uncertainty of regional adjustments to anelastic attenuation component and can be handled through ground-motion logic trees. The $\Delta c_{3,r}$ random effects for different periods are listed in Table 3-1 along with their standard errors.

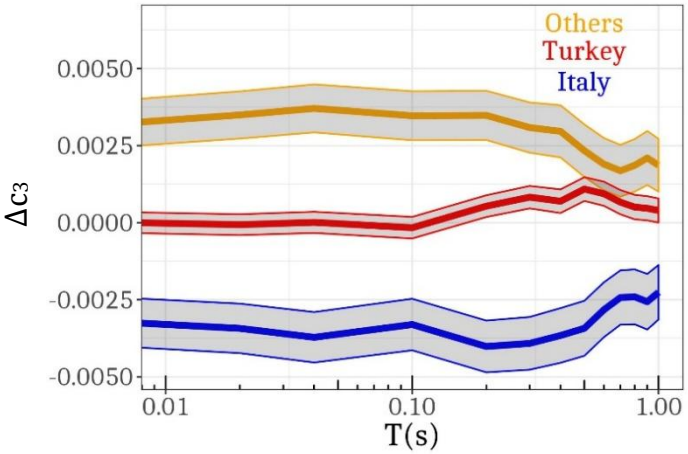


Figure 3-3: Δc_3 for the three regions across different spectral time periods. Beyond spectral period of 1s, c_3 in the regression is constrained to 0 with no regional variations. Grey-ribbon shows the 95% confidence interval about the median.

Regarding the site response term, by allowing the offset g_1 to vary among regions, we can account for regional differences in the reference V_{S30} while with g_2 we quantify the regional differences in scaling with V_{S30}. Figure 3-4 shows the random effects $\Delta g_{1,r}$ and $\Delta g_{2,r}$ for different periods, along with the estimation errors. A larger value of g_1 (and a smaller g_2) indicates a smaller reference V_{S30} for that region according to Eq. (4). Figure 3-1showed that the largest fraction of recordings from Turkey come from EC8 soil class B and C stations compared to Italy and Others groups where the stations are more evenly distributed across soil classes. This means the 'centroid' V_{S30} (modal V_{S30} value) of the data is lower for Turkey as indicated by the higher positive Δg_1 value for Turkey in Figure 3-4. Also seen in Figure 3-2 is the stronger scaling with V_{S30} for Turkey indicated by a larger negative value for g_2 in Figure 3-4. It is worth noting that by allowing regional variations in these two components of GMPE we move a fraction of the aleatory variability into epistemic uncertainty, quantified through the standard error on $\Delta c_{3,r}$, $\Delta g_{1,r}$

and $\Delta g_{2,r}$. These standard errors can be reduced by collecting more ground-motion data from the regions.

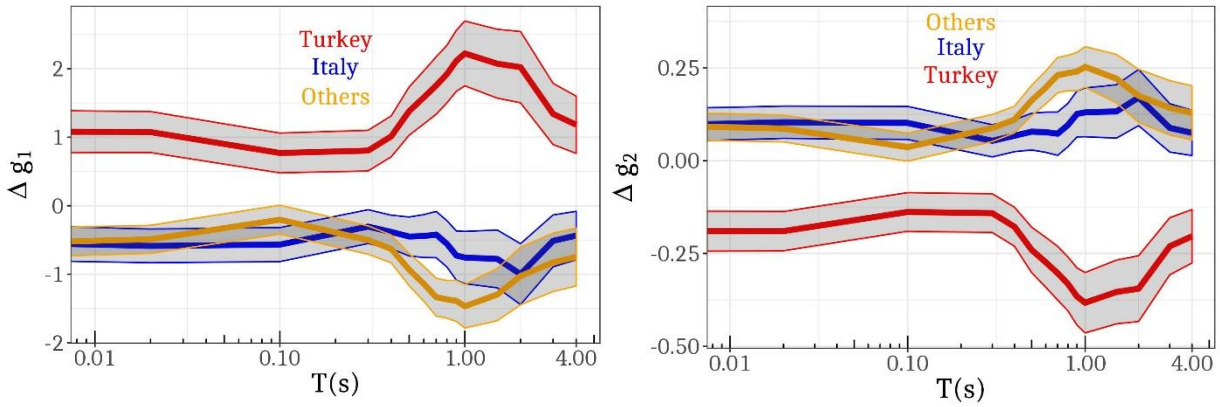


Figure 3-4: Random effects on g_1 and g_2 along with their standard errors. Δg_1 and Δg_2 are estimated as a correlated-random effects. Grey-ribbon shows the 95% confidence interval about the median.

3.5 Discussion

In the previous sections, we derived a GMPE from the European-Middle-East dataset (RESORCE), including regional (i.e. country-based) adjustments. Following recent studies, we introduced corrections for the ground-motion decay and for the scaling with Vs30 in terms of random effects.

3.5.1 Region dependent distance scaling and V_{S30} based site response

The regionalization of distance attenuation has been described by a region-dependent apparent anelastic attenuation model (Figure 3-3). As also observed by *Chiou* (2012), the geometric spreading (term dependent on logarithm of distance) and of the anelastic (term dependent on distance) contribution to the attenuation show a high degree of correlation. Studies dealing with the parametrization of Fourier amplitude decay with distance in terms of geometrical spreading and anelastic attenuation shows that the trade-off between these two terms cannot be resolved using only the spectral amplitude information (e.g. *Oth et al.*, 2011; *McNamara et al.*, 2014). Although the model for Fourier spectral amplitude is not strictly applicable to response spectra (*Bora et al.*, 2014), a similar situation arises with the GMPE, where the period-dependent terms controlling the linear decay with distance (i.e. c_1 and the magnitude correction c_2) are in trade-off with c_3, controlling the decay with the logarithm of distance. Since different wave types (body waves and surface waves) and phases (direct waves and reflect waves as SmS) contribute to the attenuation with distance over different distance and period ranges, the geometrical terms could be affected by regional bias related, for example, to differences in focal depths and crustal thickness (*Cotton et al.*, 2006; *Douglas*, 2007). Therefore, we tested a model including a correlated regional variation on the parameters controlling the distance scaling (c_1, c_2 and c_3), or considering combination of them (e.g. c_1 and c_3). Statistical tests using ANOVA (*R Core Team*, 2013; *Chambers and Duval*, 2008) do not show appreciable improvements in prediction power of GMPE (e.g., comparing the Akaike Information Criterion values, performing significance tests, or analyzing the residual distributions). The estimated regional variations in anelastic attenuation ($\Delta c_{3,r}$) are similar to the ones in the simpler model discussed in previous section, and the random effects on c_2 ($\Delta c_{2,r}$) either have 0 values at high frequencies or large standard errors (encompassing 0) at low frequencies, which makes it not a well constrained regression parameter. We finally preferred not to include regional variations in c_1 and c_2 in our model.

By considering region specific reference V_{S30}, we observed remarkable differences in the site term scaling with V_{S30}. In particular, Figure 3-2 shows that the slope of the between station random effects with V_{S30} is larger in Turkey than in the other two regions, both at short and long periods. Regional effects in the site term were already recognized in the NGA-West2 models. For example, *Abrahamson et al.* (2014) included regional corrections in the V_{S30} scaling for Taiwan, Japan and China with respect to California. As discussed in *Boore et al.* (2014), the observed regional variability of the site effects can be a consequence

of using a simplified proxy (i.e. V_{S30}) to capture the site amplification which in fact depends on many other factors, such as the soil depth. Previous studies showed that regional differences in the depths of typical soil profiles lead amplification functions with peaks occurring at different period also for site sharing similar V_{S30} (e.g. *Atkinson and Casey*, 2003; *Ghofrani et al.*, 2013), as observed when comparing sites in Japan with those in California. Previous work (e.g. *Boore et al.*, 2011) showed that the correlation of V_{S30} with the shear-wave velocity at different depths (either shallower or deeper than 30m) is regional dependent. In particular, *Boore et al.* (2011) suggested that the differences in the correlation observed for Japan with respect to California, or Europe, could be ascribed to differences in the selection of the strong motion sites, since Japanese stations are mostly installed on stiff or rock material. Similar considerations could be applied also to discuss the differences observed for Italy and Turkey. Anyway, without any detailed analysis of the velocity profiles for the analyzed stations, any conclusion would be speculative and we left this investigation for future studies.

Finally, in the NGA-West2 models the soil depth effect is considered through $\Delta Z_{1.0}$ (depth of basin to rock with V_{S30} of 1000m/s), and a regionalization for this term is also considered. The site information included in RESORCE does not allow including soil depth in the model for site effects.

3.5.2 Impact of the regionalization on the median predictions

The impact of regional adjustments on distance scaling (Figure 3-5) and magnitude scaling (Figure 3-6) obtained with and without allowing the regional corrections in the regressions are compared. Included in these figures is an 'Initial' model which is a GMPE without any regional variations with functional form as in Eq. (5). Note that in Eq. (1) the regional variability in site-response is left to be examined using Eq. (4), while for the 'Initial' GMPE without regional variability a generic site response term $g * \ln(V_{S30})$ is included in the median (Eq. 5).

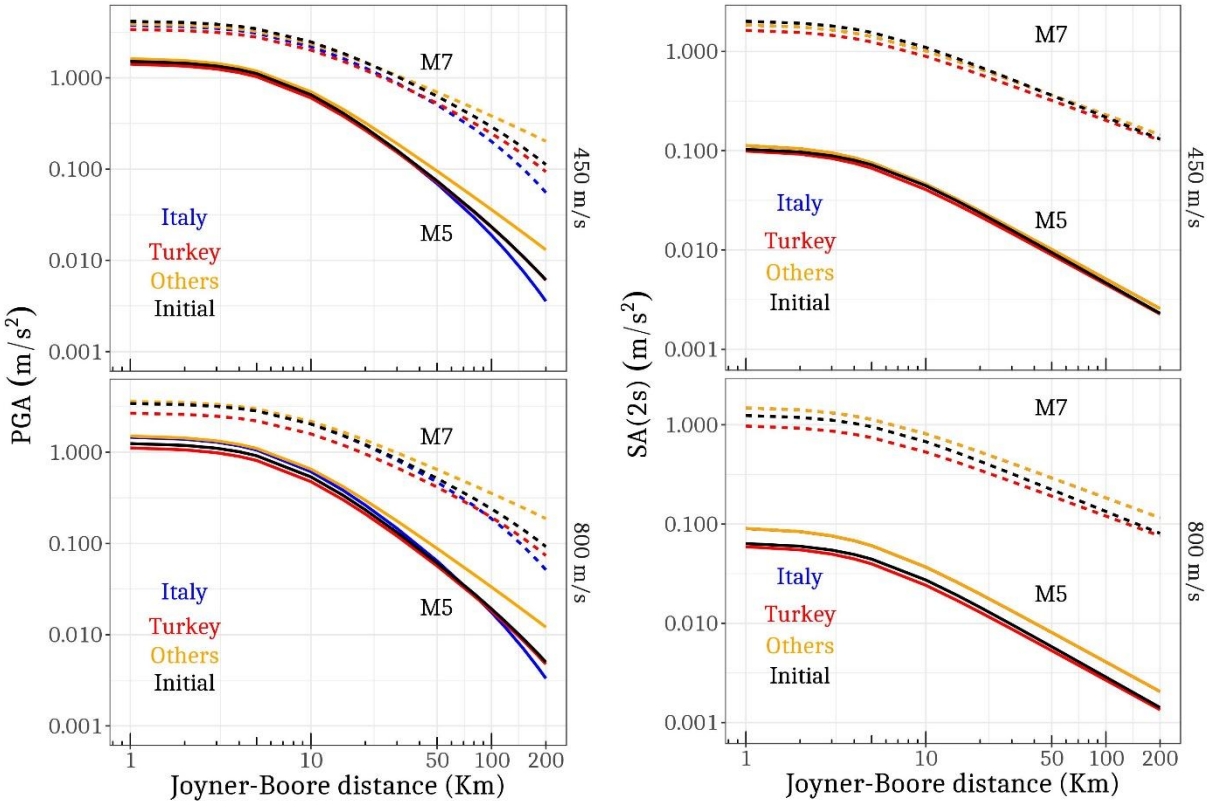


Figure 3-5: Distance scaling for PGA and SA (2s), at site with V_{S30} = 450m/s (above panels), and V_{S30} = 800m/s, for $M5$ and $M7$. Comparison of distance scaling with GMPE accounting regional variations in anelastic attenuation (slope of the curves) and V_{S30} scaling (offset of the curves), against the 'Initial' GMPE obtained from regression without accounting regional variations.

In the left panel of Figure 3-5, regional differences in high frequency ground-motion are observed as difference among the offset of the curves, and in slope of the curve at distances greater than 50km, which are a combination of V_{S30} scaling and anelastic attenuation effects. In the right panel however, which is for lower frequency ground-motion, the differences are solely due to variations in V_{S30} scaling, more pronounced for rock sites (800m/s). Similarly, in magnitude scaling (Figure 3-6) the differences in offset of the curves are a combination of regional variations in anelastic attenuation and V_{S30} scaling.

$$\ln(GM) = e_1 + F_D(R, M) + F_M(M) + g * \ln(V_{S30}) + \delta B_e + \delta B_s + \varepsilon \quad (5)$$

Cumulative effect of all regional adjustments across spectral periods is shown in the response spectra (Figure 3-7). For a site with V_{S30} of 450m/s at distance 10km, regional variations in anelastic attenuation and site response are negligible at all spectral periods. On the other end is a site with V_{S30} of 800m/s located 100km from the seismic source; in this case both anelastic attenuation and site response terms are significantly different across the regions. At the same site, for spectral periods larger than 1s the regional differences are solely contributed to by differences in site response. The two intermediate scenarios, V_{S30} of 450m/s at distance 100km, and V_{S30} of 800m/s at distance 10km show effect of regional differences in anelastic attenuation, and site response scaling with V_{S30} respectively. For example, at a rock site (800m/s) located 25km from a rupture of magnitude $M6.5$ the predicted ground-motion at spectral frequency of 3Hz is $1.51m/s^2$ in Italy, $1.47m/s^2$ in Turkey, and $1.96m/s^2$ in Others region. The differences in predicted ground-motion are significant across regions after correcting the GMPE median for regional bias.

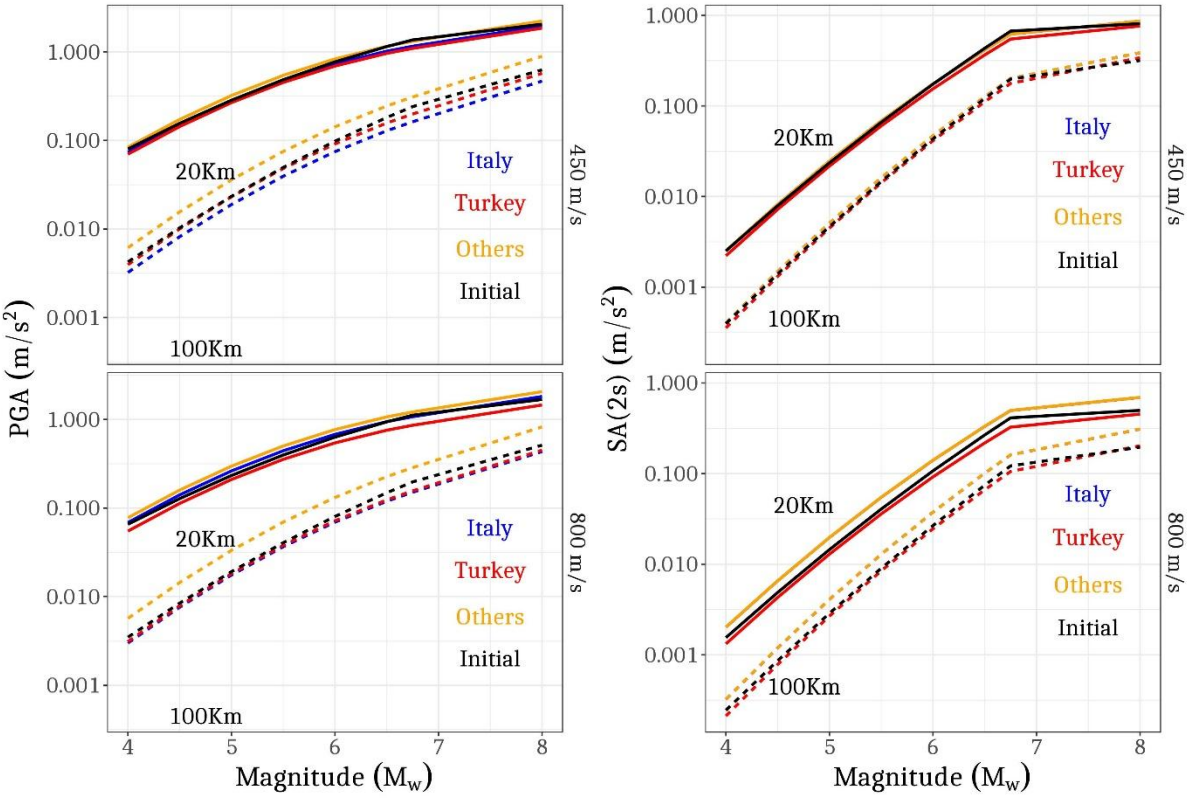


Figure 3-6: Magnitude scaling for PGA and SA (2s) at site with V_{S30} = 450m/s (above panels), and V_{S30} = 800m/s, for Joyner-Boore distances 20km, and 100km. Comparison of magnitude scaling with GMPE accounting regional variations in anelastic attenuation and V_{S30} scaling (offset of the curves), against the 'Initial' GMPE obtained from regression without accounting regional variations.

3.5.3 Impact of the regionalization on the model uncertainty

Introducing regional differences reduced the aleatory variability at the cost of an increased epistemic uncertainty in GMPE. The increase in modelling uncertainty is captured by standard errors on regional adjustments, while the reduction of variability is captured by decrease in standard deviation of GMPE given by Eq. (6)

$$\sigma = \sqrt{\tau^2 + \phi_{s2s}^2 + \phi_0^2} \qquad (6)$$

Figure 3-8 shows the comparison of standard deviations between the model with and without regional variations 'Initial'. There is a 5-10% reduction in the total standard deviation (σ) by introducing regional variations, primarily from the reduction of between-station variability (ϕ_{s2s}) by 13-20%. Reduction in residual (ϕ_0) standard deviation is small (<2%). There is no noticeable change in between-event standard deviation (τ).

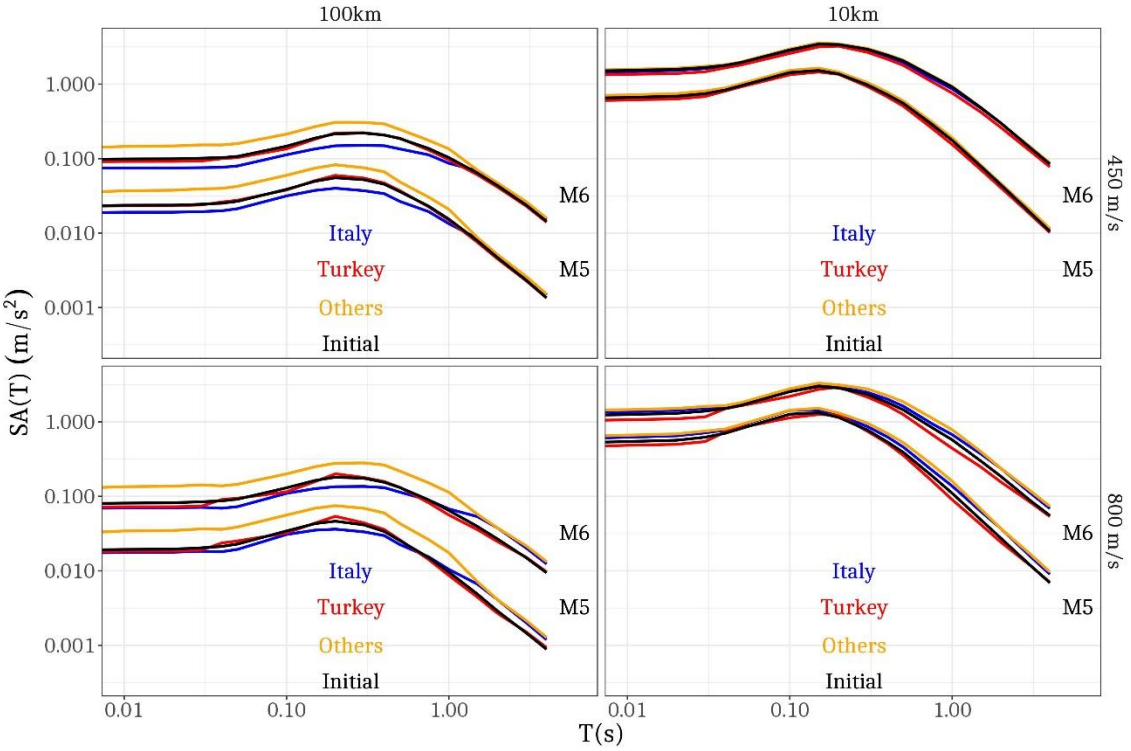


Figure 3-7: Response spectra showing the cumulative effect of regional adjustments to the GMPE. Most significant differences are observed for rock sites (V_{S30} = 800m/s) at distances larger than 50km.

Improvement in median prediction of the GMPE by correcting regional bias with regional adjustments is quantified in terms of Akaike Information Criterion (AIC), which is a measure of the relative quality of a statistical model for a given set of data penalized by the number of model parameters. Introducing the regional variations in this case increases the number of regression parameters by 3, yet a smaller AIC value of the model with regional variations justifies its increased complexity.

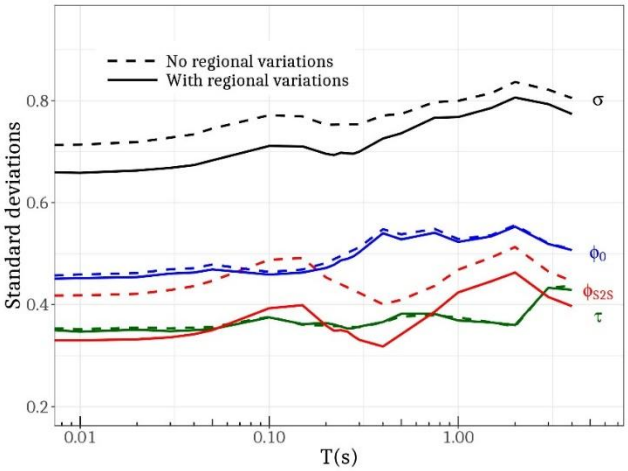


Figure 3-8: Comparison of individual components of aleatory variability in GMPE (standard deviations) between the model with regional variations (solid line) and without regional variations 'Initial' (dashed lines).

3.5.4 Potential regional differences in magnitude scale

The between-event residual δB_e can be used to evaluate the impact of considering earthquakes with converted moment magnitude from other magnitude scales (local magnitude, surface-wave magnitude, and body-wave magnitude). In Figure 1, the recordings relevant to these earthquakes are shown in red and mainly correspond to magnitude smaller than 5 in Turkey. Figure 3-9 is a box-plot of δB_e at SA(1s) for each country in the regressed dataset. The scatter in δB_e from considering events with both computed (from moment tensor solutions) and empirically estimated M_w is larger than that when considering only those with computed M_w (refer to *Akkar et al.*, 2014a, for details on empirical estimation of M_w). In Figure 3-9, this reduction in scatter can be seen as a shift of the country-wise median towards 0, from left to right panel. Within-country scatter shown as the height of the box-plot has also reduced, especially in case of Turkey. Filtering out events with empirically estimated M_w reduced the between-event standard deviation of the GMPE (τ) by an average of 10% (and a maximum of 30%) across the periods, without losing constrain on other regression parameters (i.e., increase in standard error of estimate of coefficients). We note that this filter primarily removes small magnitude events from Turkey (less than $M5$), which could also be the reason for decrease in τ. A further study could be focused on examining the regional differences in moment tensor solutions based computed M_w which, once homogenized, may allow analyzing other regional differences in source physical parameters.

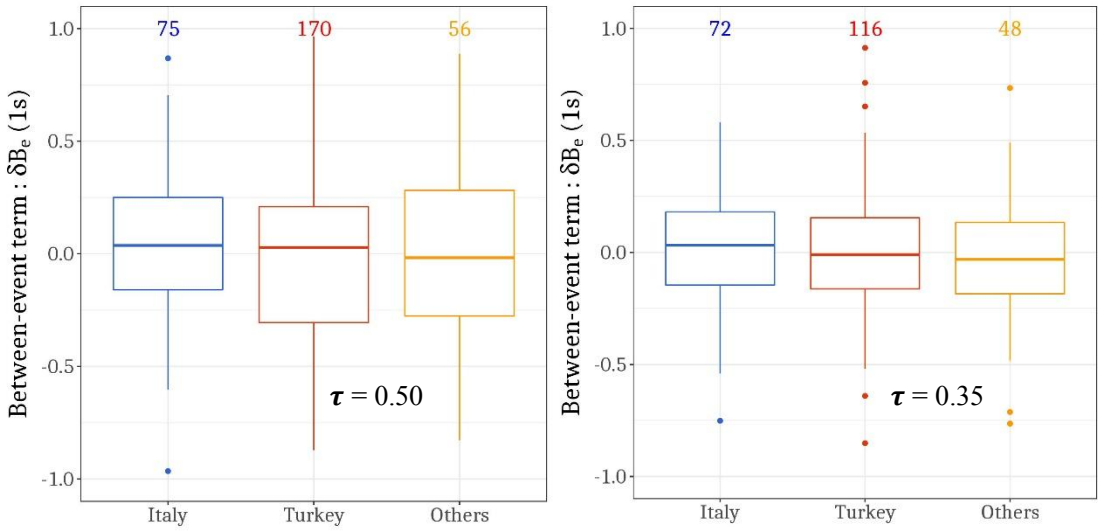


Figure 3-9: Regional variation of between event residuals at SA (1s). Box-plot the median (50th) and the quartiles (5th, 25th, 75th and 95th). The left panel shows residuals from all the events whose M_w is either computed (as calculated from moment tensor solutions) or empirically estimated (for details refer to Akkar et al., 2014a). Right panel shows the residuals from only the events whose M_w is computed, and not empirically estimated/converted. The decrease in height of the box plots reflects a decrease in between event variability within and across regions.

3.6 Conclusions

RESORCE database and the Non-linear mixed effects regression tools allowed analyzing and quantifying regional variations in ground-motion data for Europe – Middle-East regions. The GMPE is developed specifically for active crustal earthquakes in Europe – Middle-East regions, and we do not recommend using it elsewhere without a prior compatibility check. The dataset is strongly unbalanced across the contributing regions in terms of magnitude, distance and recording station site classification. If separate GMPEs were to be developed for each of the regions, then the applicability of each GMPE would be strongly limited in magnitude, distance, and site V_{S30} range. By allowing regional variability only on specific terms (anelastic attenuation and site response), and estimating all the regression coefficients (magnitude scaling, geometric spreading) using the entire dataset we overcome this limitation. In its current form, the GMPE is recommended to be used for following scenarios:

- Active crustal earthquakes magnitude range from 4 to 7.6: since the magnitude distribution is symmetric around the median magnitude of M5.5, neither the small nor the large events are likely to bias the prediction.
- Sites with V_{S30} from 180 to 1000m/s: Even though the range of V_{S30} used in regression is 90m/s to 2000m/s, the bulk of data is within 200-600m/s. We suggest using the GMPE in a range narrower than its underlying dataset, and especially not to extrapolate beyond the suggested V_{S30} limits.
- Joyner-Boore (R_{JB}) distances up to 200km: The GMPE is calibrated with data up to 300km with the bulk of data from within 150km.
- Partially non-ergodic region specific seismic hazard assessment by adjusting the GMPE median (Table 3-1) and linear site-amplification model (Table 3-2) with the provided regional adjustments. The reported standard errors are estimated as square-root of conditional variances estimated by the Markov Chain Monte Carlo bootstrap method available in LME4.0 package in *R* (*Bates et al.*, 2014). These values can also be used as epistemic uncertainty on the regional adjustments. Since the underlying distribution is not known, the epistemic uncertainty can be assumed to be normally distributed and modeled using a three-point distribution that maintains the mean and the standard deviation of the original distribution. Under such an assumption the upper and lower limits on regional adjustments

can be set as ±1.6 times the standard error, with logic tree weights 0.2, 0.6 and 0.2 for upper, middle and lower branches respectively.

At the moment statistically significant regional variations in apparent anelastic attenuation, and V_{S30} scaled linear site response could be captured and accounted in the new GMPE; thereby correcting median for regional bias and deflating the total variability by 5-10% depending on the spectral period. Regional differences in distance scaling found in this study are in agreement with recently published studies. The largest reduction in GMPE standard deviation comes from allowing regional variations in the site response component. This variability could be further reduced by using a combination of site-response proxies, instead of V_{S30} alone. Another large reduction in standard deviation comes from using only the events with moment tensor solutions based moment magnitude in regression, at the cost of losing many small magnitude events. It is desirable to plug such data losses by homogenizing the magnitude scale across regions. In summary, a decrease in aleatory variability of ground-motion prediction as demonstrated in this study is accompanied by a new epistemic uncertainty on estimated regional adjustments, which in turn may only be reduced by improving the underlying datasets.

Acknowledgments

We are very thankful to John Douglas, and the anonymous reviewer for their insightful remarks and suggestions to improve the manuscript. We also wish to thank Dietrich Stromeyer, Amir Hakimhashemi, Olga-Joan Ktenidou, Sanjay Singh Bora, and colleagues at section 2.6 of GFZ German Research Centre for Geosciences for their invaluable contributions in understanding the mathematics involved, interpretation of our results, and the thorough internal review of the manuscript.

Table 1. Coefficient table for GMPE

t	e_1	b_1	b_2	b_3	c_1	c_2	c_3	h	τ	$\emptyset_0$	σ	$\Delta c_{3,IT}$	$\Delta c_{3,Others}$	$\Delta c_{3,TR}$	SE($\Delta c_{3,IT}$)	SE($\Delta c_{3,Others}$)	SE($\Delta c_{3,TR}$)
pgv	0.773	0.483	-0.101	-0.021	-1.198	0.229	-0.00008	5.845	0.349	0.496	0.683	-0.00189	0.00142	0.00050	0.00077	0.00074	0.00035
pga	2.982	-0.363	-0.195	-0.406	-1.231	0.272	-0.00395	6.390	0.350	0.451	0.657	-0.00326	0.00326	0.00000	0.00079	0.00076	0.00034
0.01	3.002	-0.366	-0.193	-0.412	-1.236	0.272	-0.00385	6.425	0.347	0.452	0.657	-0.00334	0.00341	-0.00007	0.00080	0.00076	0.00034
0.02	3.064	-0.368	-0.192	-0.425	-1.251	0.273	-0.00375	6.336	0.351	0.454	0.661	-0.00343	0.00349	-0.00006	0.00080	0.00076	0.00034
0.03	3.128	-0.378	-0.183	-0.440	-1.267	0.278	-0.00371	6.108	0.348	0.461	0.672	-0.00356	0.00364	-0.00008	0.00081	0.00077	0.00034
0.04	3.223	-0.414	-0.168	-0.487	-1.299	0.291	-0.00377	6.096	0.350	0.463	0.681	-0.00372	0.00371	0.00001	0.00082	0.00078	0.00034
0.05	3.304	-0.478	-0.165	-0.497	-1.321	0.301	-0.00388	6.086	0.352	0.469	0.704	-0.00374	0.00378	-0.00005	0.00083	0.00079	0.00035
0.10	3.757	-0.666	-0.232	-0.341	-1.342	0.295	-0.00522	7.658	0.375	0.459	0.702	-0.00330	0.00347	-0.00016	0.00084	0.00079	0.00035
0.15	3.877	-0.404	-0.226	-0.214	-1.212	0.243	-0.00693	7.468	0.362	0.463	0.693	-0.00371	0.00338	0.00033	0.00084	0.00080	0.00035
0.20	3.578	-0.217	-0.231	-0.122	-1.048	0.207	-0.00792	6.030	0.364	0.472	0.700	-0.00402	0.00348	0.00054	0.00084	0.00080	0.00035
0.30	3.482	0.107	-0.226	-0.042	-0.966	0.159	-0.00701	5.123	0.357	0.503	0.725	-0.00391	0.00308	0.00083	0.00085	0.00081	0.00036
0.40	3.340	0.243	-0.233	0.010	-0.947	0.142	-0.00539	4.750	0.366	0.540	0.737	-0.00366	0.00296	0.00070	0.00089	0.00085	0.00038
0.50	3.220	0.392	-0.191	-0.236	-0.946	0.163	-0.00497	4.580	0.382	0.528	0.767	-0.00343	0.00234	0.00109	0.00089	0.00085	0.00038
0.75	2.998	0.667	-0.169	-0.178	-0.972	0.144	-0.00197	4.685	0.382	0.541	0.769	-0.00229	0.00175	0.00052	0.00088	0.00084	0.00039
1.00	2.880	0.837	-0.176	-0.114	-0.990	0.128	-0.00094	5.392	0.369	0.523	0.786	-0.00226	0.00186	0.00039	0.00088	0.00086	0.00039
1.50	2.312	1.127	-0.127	-0.094	-0.948	0.139	0.00000	4.553	0.365	0.534	0.806	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
2.00	1.684	1.079	-0.159	-0.222	-0.911	0.162	0.00000	4.309	0.360	0.553	0.793	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
3.00	1.057	1.474	-0.039	0.052	-0.855	0.160	0.00000	4.365	0.433	0.519	0.774	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
4.00	0.755	1.775	0.035	0.302	-0.852	0.143	0.00000	4.990	0.429	0.507	0.683	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

Units of estimated PGV is (m/s), PGA and SA (t) are estimated in (g)

Table 2. Coefficient table for V_{S30} based site response

t	g_1	g_2	$\emptyset_{S2S}$	$\Delta g_{1,IT}$	$\Delta g_{1,Others}$	$\Delta g_{1,TR}$	$\Delta g_{2,IT}$	$\Delta g_{2,Others}$	$\Delta g_{2,TR}$	SE($\Delta g_{1,IT}$)	SE($\Delta g_{1,Others}$)	SE($\Delta g_{1,TR}$)	SE($\Delta g_{2,IT}$)	SE($\Delta g_{2,Others}$)	SE($\Delta g_{2,TR}$)
pgv	2.188	-0.364	0.314	-0.296	-1.082	1.378	0.051	0.186	-0.236	0.286	0.230	0.349	0.049	0.039	0.060
pga	1.407	-0.234	0.330	-0.360	-0.678	1.038	0.063	0.119	-0.182	0.258	0.212	0.314	0.045	0.037	0.055
0.01	1.399	-0.233	0.330	-0.351	-0.663	1.013	0.062	0.116	-0.178	0.256	0.210	0.310	0.045	0.037	0.054
0.02	1.382	-0.230	0.332	-0.379	-0.655	1.034	0.067	0.115	-0.182	0.253	0.208	0.307	0.045	0.037	0.054
0.03	1.312	-0.218	0.336	-0.376	-0.652	1.028	0.066	0.115	-0.182	0.252	0.207	0.304	0.044	0.037	0.054
0.04	1.244	-0.207	0.342	-0.409	-0.606	1.014	0.072	0.107	-0.179	0.255	0.210	0.307	0.045	0.037	0.054
0.05	1.163	-0.194	0.350	-0.439	-0.562	1.001	0.078	0.099	-0.177	0.259	0.213	0.311	0.046	0.038	0.055
0.10	0.962	-0.160	0.393	-0.344	-0.390	0.734	0.061	0.070	-0.131	0.261	0.220	0.304	0.047	0.039	0.054
0.15	1.066	-0.177	0.399	-0.072	-0.341	0.413	0.013	0.062	-0.075	0.212	0.184	0.233	0.039	0.034	0.043
0.20	1.207	-0.200	0.359	-0.094	-0.403	0.497	0.017	0.072	-0.089	0.224	0.192	0.256	0.040	0.034	0.046
0.30	1.462	-0.243	0.331	-0.109	-0.660	0.769	0.019	0.115	-0.134	0.264	0.217	0.316	0.046	0.038	0.055
0.40	1.779	-0.296	0.318	-0.206	-0.777	0.983	0.036	0.135	-0.171	0.261	0.213	0.318	0.045	0.037	0.055
0.50	2.236	-0.373	0.342	-0.294	-1.113	1.406	0.050	0.191	-0.242	0.308	0.247	0.375	0.053	0.043	0.065
0.75	2.931	-0.488	0.386	-0.329	-1.501	1.831	0.057	0.258	-0.315	0.351	0.281	0.430	0.060	0.048	0.074
1.00	3.348	-0.558	0.424	-0.603	-1.581	2.184	0.103	0.271	-0.374	0.392	0.321	0.479	0.067	0.055	0.082
1.50	3.395	-0.566	0.446	-0.720	-1.309	2.028	0.123	0.223	-0.346	0.429	0.384	0.505	0.073	0.065	0.086
2.00	3.337	-0.556	0.463	-0.952	-1.009	1.962	0.162	0.172	-0.334	0.449	0.419	0.520	0.076	0.071	0.089
3.00	2.964	-0.493	0.415	-0.394	-0.831	1.226	0.069	0.145	-0.214	0.354	0.390	0.417	0.062	0.068	0.073
4.00	2.707	-0.451	0.397	-0.341	-0.791	1.021	0.060	0.138	-0.178	0.332	0.386	0.387	0.058	0.067	0.068

Table 3-1: Coefficient table for GMPE

Table 3-2: Coefficient table for V_{S30} based site response

4

From Ergodic to Region- and Site-Specific Probabilistic Seismic Hazard Assessment: Method Development and Application at European and Middle Eastern Sites

S. R. Kotha[1, 2], D. Bindi[1], F. Cotton[1, 2]

[1]German Research Centre for Geosciences GFZ, Potsdam, Germany

[2] University of Potsdam, Potsdam, Germany

Abstract

The increasing numbers of recordings at individual sites allows quantification of empirical linear site-response adjustment factors ($\delta S2S_s$) from the Ground-Motion Prediction Equation (GMPE) residuals. The $\delta S2S_s$ are then used to linearly scale the ergodic GMPE predictions to obtain site-specific ground-motion predictions in a partially non-ergodic Probabilistic Seismic Hazard Assessment (PSHA). To address key statistical and conceptual issues in the current practice, we introduce a novel empirical region- and site-specific PSHA methodology wherein, (1) site-to-site variability (ϕ_{S2S}) is first estimated as a random-variance in a mixed-effects GMPE regression, (2) $\delta S2S_s$ at new sites with strong motion are estimated using the a-priori ϕ_{S2S}, and (3) the GMPE site-specific single-site aleatory variability $\sigma_{ss,s}$ is replaced with a generic site-corrected aleatory variability σ_0. Comparison of region- and site-specific hazard curves from our method against the traditional ergodic estimates at 225 sites in Europe and Middle-East shows an approximate 50% difference in predicted ground-motions over a range of hazard levels - a strong motivation to increase seismological monitoring of critical facilities and enrich regional ground-motion datasets.

Keywords: Site-specific Ground-Motion Prediction Equations, Europe and the Middle-East, Nonlinear Mixed Effects Regression

4.1 Introduction

Current site-specific seismic hazard analyses are generally performed using probabilistic methods. When dealing with a specific site, the typical methodology involves using a Ground-Motion Prediction Equation (GMPE) to estimate the rock ground-motion at depth and its associated variability. Ground-motions are then propagated to the ground surface by site response analysis. Recent studies have shown difficulties: (1) in adjusting GMPEs calibrated on soft surface rock motions to hard-rock conditions at depth, (2) in simulating complex 2D/3D site response using geotechnical information which may be only 1D, and (3) in correctly evaluating the site-specific variability (e.g. *Baturay and Stewart*, 2003; *Goulet and Stewart*, 2009; Thompson et al. 2009, *Afshari and Stewart*, 2015).

Along with the development of strong motion networks in the last twenty years, multiple recordings from several events recorded at individual sites are becoming readily available. The rapid increase of the number of records at a single site is a strong motivation for the development of new strategies to compute region- and site-specific PSHA (e.g. *Bradley*, 2013; *Faccioli et al.*, 2015; *Villani and Abrahamson*, 2015; *Douglas and Aochi*, 2016). Several GMPE residual analysis studies, performed at sites where many earthquakes were recorded, identified repeatable site-specific effects and a better estimate of «single-site» ground-motion variabilities (e.g. *Atkinson*, 2006; *Lin et al.*, 2011; *Rodriguez-Marek et al.*, 2011; *Afshari and Stewart*, 2015). The within-event residuals can be regarded as the sum of a site factor ($\delta S2S_s$), and event-and-site corrected residual ($\delta WS_{es,s}$) (e.g. *Al Atik et al.*, 2010; *Rodriguez-Marek et al.*, 2013). The site factor represents the systematic deviation of the observed linear amplification at this site from the median amplification predicted by the model using simple site classification such as the time-averaged shear-wave velocity in the uppermost 30 meters at the site, V_{S30}.

The resulting approach is commonly referred to as partially non-ergodic. This approach is attractive because it allows one to compute PSHA taking into account site-specific and data-driven (empirical) linear amplification factors without the cost of adjusting rock-GMPEs at depth or computing complex site responses. However, the development of partially non-ergodic PSHA is not straightforward, and several key questions need to be addressed:

1) *Rodriguez-Marek et al.* (2013) and *Faccioli et al.* (2015) identified as key requirements for application in PSHA that the median value of $\delta S2S_s$ must be properly estimated, along with the epistemic uncertainty in site term $\delta S2S_s$ and the single-site sigma $\sigma_{ss,s}$.
2) *Stafford* (2014) suggested that the mixed-effects algorithm of *Abrahamson and Youngs* (1992) used to estimate between-event random-effects and within-event residuals of a GMPE, may not support the subsequent step of manipulating within-event residuals to obtain site term $\delta S2S_s$ - without consistently updating the GMPE median. Here we assess the benefit of using a more sophisticated Non-linear Mixed Effects Regression algorithm (NLME, *Bates et al.*, 2014) to simultaneously estimate between-site ($\delta S2S_s$) and between-event random-effects, and the impact of such residual decomposition on the GMPE median.
3) Only sites with a certain minimum number of records are eligible for site-specific PSHA. The definition of this threshold and the impact of the number of recordings on the reliability of $\delta S2S_s$ are still open questions in the GMPE community. Consequently, partially non-ergodic site-specific PSHA has been performed for only a few well-recorded sites, while comparisons with standard ergodic approaches for multiple-sites over a wide spatial scale has not yet been attempted.

In this work, we refine the partially non-ergodic PSHA framework to better meet these requirements. We first compare two different approaches for estimating the *linear only* site-response adjustment $\delta S2S_s$ and its uncertainty for sites featuring in the GMPE strong motion dataset. We then extend this comparison to the estimation of $\delta S2S_s$ for new sites whose strong motion data were not included in the GMPE regression. In the process, we discuss the need for a minimum number of records and the uncertainties to be accounted for a reliable site-specific hazard assessment. Finally, we apply our partially non-ergodic region- and site-specific PSHA approach to 225 sites in pan-Europe, taking advantage of the recent dissemination

of RESORCE pan-European strong-motion dataset (*Akkar et al.*, 2014). The hazard curves obtained for the 225 sites using a classical ergodic PSHA are compared with region- and site-specific hazard curves computed using the partially non-ergodic approach proposed in this study.

4.2 Dataset and Ergodic Ground-Motion Prediction Equations

In this study, we use the Europe and Middle Eastern strong motion dataset RESORCE. The RESORCE project dataset includes 5882 recordings from 1814 earthquakes occurring in Europe and the Middle East with magnitude range from 3.0 to 7.8. Turkey and Italy are the two main contributors of crustal earthquakes in this dataset; the rest of the Europe and Middle East regions (labeled 'Others' hereon) contribute smaller numbers. For example, France, Georgia, Greece, Iran, and Montenegro contribute fewer than 50 records individually to the dataset. In total, there are 225 sites in the dataset with known V_{S30} and at least two recordings.

The most recent pan-European GMPEs (*Douglas et al.*, 2014) are based on the RESORCE strong motion dataset (http://www.resorce-portal.eu/). The GMPE developed by *Bindi et al.* (2014), whose functional form for median peak spectral acceleration is shown in Eq. (1), is one such model calibrated using records with $M_w \geq 4.0$, focal depth $\leq$ 35km, $R_{JB} \leq$ 300km from the RESORCE dataset. Details regarding record selection and random-effects regression method (*Abrahamson and Youngs*, 1992) are available in *Bindi et al.* (2014). We consider *Bindi et al.* (2014) as an ergodic GMPE, not including any region or site-specific adjustments. The ergodic aleatory variability, sigma (σ), of the GMPE is estimated as in Eq. (2) where τ is the standard deviation of between-event residuals (δB_e), and ϕ is the standard deviation of within-event residuals (δW_{es}).

$$\log_{10}(\mu) = e_1 + F_D(R, M) + F_M(M) + F_{SoF} + g * log_{10}\left(\frac{V_{S30}}{800}\right) \quad (1)$$

$$\sigma = \sqrt{\tau^2 + \phi^2} \quad (2)$$

4.3 Partially Non-Ergodic Site-Specific GMPEs

4.3.1 Current practice: R13 APPROACH

Previous studies (e.g. *Rodriguez-Marek et al.*, 2014; *Faccioli et al.*, 2015) extensively discussed the site-specific PSHA procedure. Essentially, the site-specific PSHA begins with a site-specific GMPE, where the median and standard deviation of an ergodic GMPE are modified to a site-specific median and standard deviation. Since this approach was formalized by *Rodriguez-Marek et al.* (2013), we refer to it hereafter as R13. R13 requires within-event residual processing at a site (δW_{es} in Eq. 3) to isolate the 'mean' site-specific residual ($\delta S2S_s$ in Eq. 3) and event-and-site corrected residuals (δWS_{es} in Eq. 3). For well-recorded sites, the $\delta S2S_s$ in Eq. (3) can be considered as a site-specific adjustment to the GMPE median. The standard deviation of event-and-site corrected residuals at that site is the site-specific aleatory variability ($\phi_{ss,s}$). Median of the ergodic GMPE in Eq. (1) can be adjusted to obtain a site-specific GMPE median as shown in Eq. (4), where the ergodic GMPE is modified with the site-specific adjustment ($\delta S2S_s$). The site-specific sigma ($\sigma_{ss,s}$), estimated through Eq. (5), replaces the ergodic sigma (σ) shown in Eq. (2).

$$\delta W_{es} = \delta WS_{es} + \delta S2S_s \quad (3)$$

$$l\log_{10}(\mu_s) = l\log_{10}(\mu) + \delta S2S_s \quad (4)$$

$$\sigma_{ss,s} = \sqrt{\tau^2 + \phi_{ss,s}^2} \quad (5)$$

In R13, since the site-specific median and sigma are estimated with a small sample of data, usually 10 recordings or more, standard errors on both the $\delta S2S_s$ and $\sigma_{ss,s}$ need to be accounted (*Abrahamson and Hollenback*, 2012; *Faccioli et al.*, 2015). The key points discussed in the context of R13 are:

1. **Estimation of site term ($\delta S2S_s$)**: Site terms are estimated as the mean of site-specific within-event residuals ($\delta W_{es,s}$ in Eq. 6), where n_s is the number of recordings at the site.
2. **Epistemic uncertainty on $\delta S2S_s$** is estimated using Eq. (7), where between-site variability (ϕ_{S2S}) is the standard deviation of the $\delta S2S$ random variable.
3. **The site-specific variability ($\phi_{ss,s}$)** used in place of the ergodic within-event ϕ is the standard deviation of site-specific event-and-site corrected residuals (Eq. 8). Since $\phi_{ss,s}$ is estimated using a limited site-specific sample of recordings, it is provided with an epistemic uncertainty (Eq. 9), in which SD($\phi_{ss,s}$) is the standard deviation of the distribution of $\phi_{ss,s}$.

$$\delta S2S_s = \frac{\sum \delta W_{es}}{n_s} \tag{6}$$

$$SE(\delta S2S_s) = \frac{\phi_{S2S}}{\sqrt{n_s}} \tag{7}$$

$$\phi_{ss,s} = SD(\delta WS_{es}) \tag{8}$$

$$\phi_{ss,s,epistemic} = \frac{SD(\phi_{ss,s})}{\sqrt{n_s}} \tag{9}$$

The epistemic uncertainty on $\phi_{ss,s}$ translates into epistemic uncertainty on the site-specific sigma ($\sigma_{ss,s}$). *Faccioli et al.* (2015) accommodated this uncertainty by considering a lower and upper percentiles on $\sigma_{ss,s}$ (Eq. 10), where τ is the standard deviation of between-event residuals (δB_e), identical to that in Eq. (5).

$$\sigma^u_{ss,s} = \sqrt{\left(\phi_{ss,s} + \phi_{ss,s,epistemic}\right)^2 + \tau^2}$$

$$\sigma_{ss,s} = \sqrt{\phi^2_{ss,s} + \tau^2} \tag{10}$$

$$\sigma^l_{ss,s} = \sqrt{\left(\phi_{ss,s} - \phi_{ss,s,epistemic}\right)^2 + \tau^2}$$

4.3.2 Motivation to revise R13

Although R13 is a well-established intuitive approach to derive a site-specific GMPE from an ergodic GMPE, *Stafford* (2014) pointed out a few statistical shortcomings in the estimation of the site-specific adjustments ($\delta S2S_s$) following a traditional *Abrahamson and Youngs* (1992) random-effects regression. In the context of GMPE regression analysis, the random effect approach was first introduced by *Brillinger and Preisler* (1985) and *Abrahamson and Youngs* (1992) to handle imbalanced strong-motion datasets. In particular, mixed-effect regressions were proposed to avoid bias in the predicted median when a dataset featured well-recorded *atypical* earthquakes. By introducing a random effect on the earthquake grouping level, atypical earthquakes were identified by their large between-event residuals. For site-specific GMPEs, and ergodic GMPEs in general, it is also important to consider the dataset from an *atypical* sites viewpoint. Since there is no site-to-site distinction or site-level grouping in the *Abrahamson and Youngs* (1992) random-effects algorithm, the site-response component in the median (scaling with V_{S30}) and estimation of site-specific residuals may be biased by the presence of atypical sites. Such bias in the median site-response is not uncommon in GMPEs. GMPEs tend to be centered on the densest part of the dataset, and usually the number of rock sites ($V_{S30} \geq$ 760m/s) and relevant records are fewer compared to those from softer soils (V_{S30}~450m/s). Due to such imbalances in the dataset, in terms of number of records at sampled site types, GMPE predictions are less accurate for less recorded yet typical site conditions. Thus, large differences among GMPEs tend to occur in rock ground-motion predictions. We are motivated to investigate such tendencies with the (strongly imbalanced) RESORCE strong motion dataset, and to revise R13 with the aim of deriving a site-specific GMPE.

4.3.3 Nonlinear Mixed Effects Approach

According to *Stafford* (2014), it is recommended to estimate the site-to-site variability as a between-site random-variance *during* the regression, allowing an 'interaction' between coefficient 'g' of the V_{S30} site-response scaling (within $\log_{10}(\mu)$ of Eq. 1) and $\delta S2S_s$ in Eq. (4), similar to the interaction between magnitude-scaling component and between-event random-effects (δB_e) of a GMPE. We call this approach NLME approach in this context, abbreviating Non-Linear Mixed Effects. To demonstrate the differences between NLME and R13 approaches, we derive two GMPEs with identical median functional form and record selection as in *Bindi et al.* (2014) (Eq. 1 and 2), but with different partitioning of residuals: (1) in R13, between-event (δB_e) and within-event residuals (δW_{es}) and, (2) in NLME, between-event, between-site ($\delta S2S_s$), and event-and-site corrected residuals (δW_{es}). From here on, the first GMPE is referred to as R13-GMPE to designate it as the first step in the R13 method. The R13-GMPE within-event residuals are later partitioned into site-terms and event-and-site corrected residuals to develop site-specific GMPEs. The second GMPE, where the site-terms are estimated as random-effects during the regression, is the first step of the NLME approach for partially non-ergodic PSHA introduced in this study. In subsequent plots, it is labeled as NLME-GMPE.

4.3.3.1 Comparison of the median response spectra

Median response spectra from the two GMPEs for typical hazard scenarios are shown in Figure 4-1a. In the scenario (M 4.5, 10km, 760m/s) of Figure 4-1a, R13-GMPE (blue) predicts 20% higher SA(0.1s) ground-motion compared to NLME-GMPE (red). Figure 4-1a shows larger differences in median spectra for V_{S30} = 760m/s (rock site) compared to V_{S30} = 360m/s. Similar but smaller differences are observed at sites with V_{S30} = 180m/s. Such features are typical of imbalanced datasets where the GMPE is centered on the densest part of the dataset which in the case of RESORCE, corresponds to M5.0-M6.0 and 180m/s ≤ V_{S30} ≤ 760m/s (soil type B and C according to Eurocode8 classification). Note that in Figure 4-1a, although the differences observed are smaller than those usually observed across different GMPEs (e.g. NGA-West2 models), they are solely from the degree of residual partitioning, preserving the GMPE record selection and functional form.

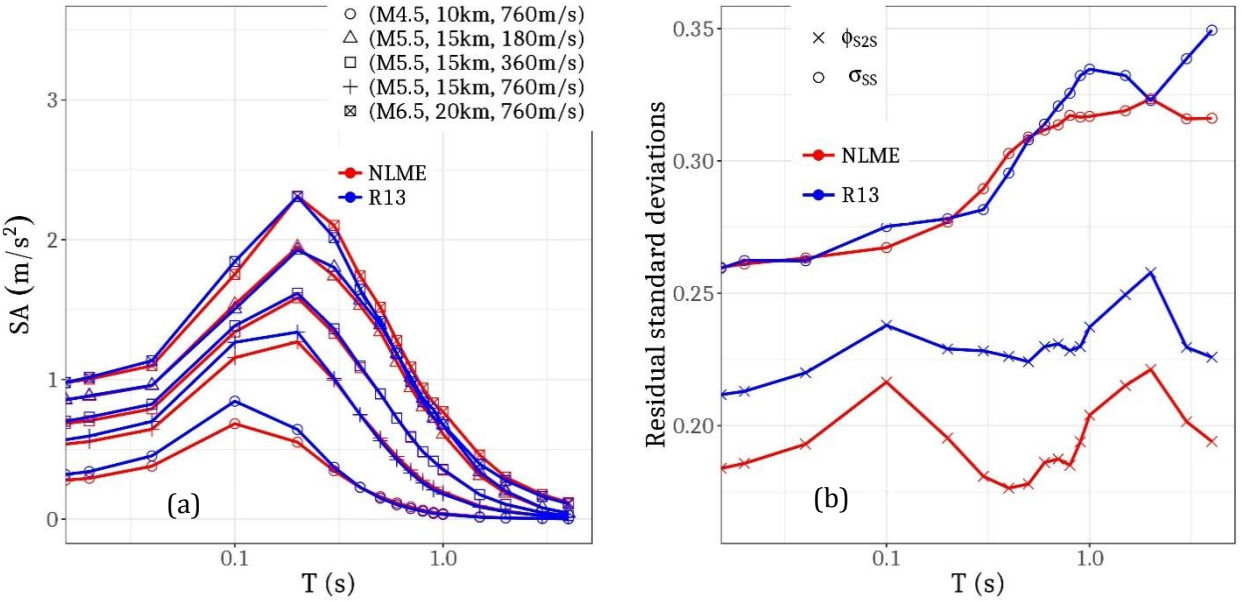


Figure 4-1: Comparison of GMPEs: R13 corresponds to the GMPE where the site-terms are estimated as the mean of site-specific within-event residuals, and NLME refers to the GMPE whose site-terms are estimated as site-specific random effects during regression. (a) Comparison of predicted medians for typical (M, R, V_{S30}) scenarios (b) Comparison of residual standard deviation values.

4.3.3.2 *Comparison of residual standard deviations*

Figure 4-1b shows that the site-corrected total standard deviations σ_{ss} from the two approaches are very similar. By construction, the site-terms ($\delta S2S_s$), estimated as random-effects in NLME, follow a normal distribution (*Bates et al.*, 2014), and those from R13 site-specific GMPE follow a normal distribution by virtue of the Central Limit Theorem (CLT). We performed the Shapiro-Wilk normality test (*Royston,* 1992) to confirm that $\delta S2S$ from R13 method follows a normal distribution as well. ϕ_{S2S} is the standard deviation of the random variable $\delta S2S$, and quantifies the site-to-site variability of the dataset. Figure 4-1b shows that despite using the same set of sites (and records) in both regressions, ϕ_{S2S} from NLME is on an average 20% smaller than its R13 counterpart.

4.3.3.3 *Comparison of site-terms*

a) Distribution of $\delta S2S_s$: Figure 4-2a shows the distribution of $\delta S2S_s$ obtained from the two methods for the sites in RESORCE dataset. The R13-estimated δS2S (distribution of $\delta S2S_s$ at PGA and SA(1s)) for sites with fewer than five records shows a wider spread than that from the NLME approach (left column of Figure 4-2a). In the right column of Figure 4-4a, corresponding to sites with more than five records, NLME yields a marginally wider distribution compared to R13. From these diagrams, we infer that R13 is likely to concentrate the well-recorded sites (in this case, those having more than five records) close to the center of the distribution, rendering a large number of sparsely-sampled sites as uncharacteristic (outlier) sites. This explains the R13 overestimate of site-to-site variability ϕ_{S2S} observed in Figure 4-1b.

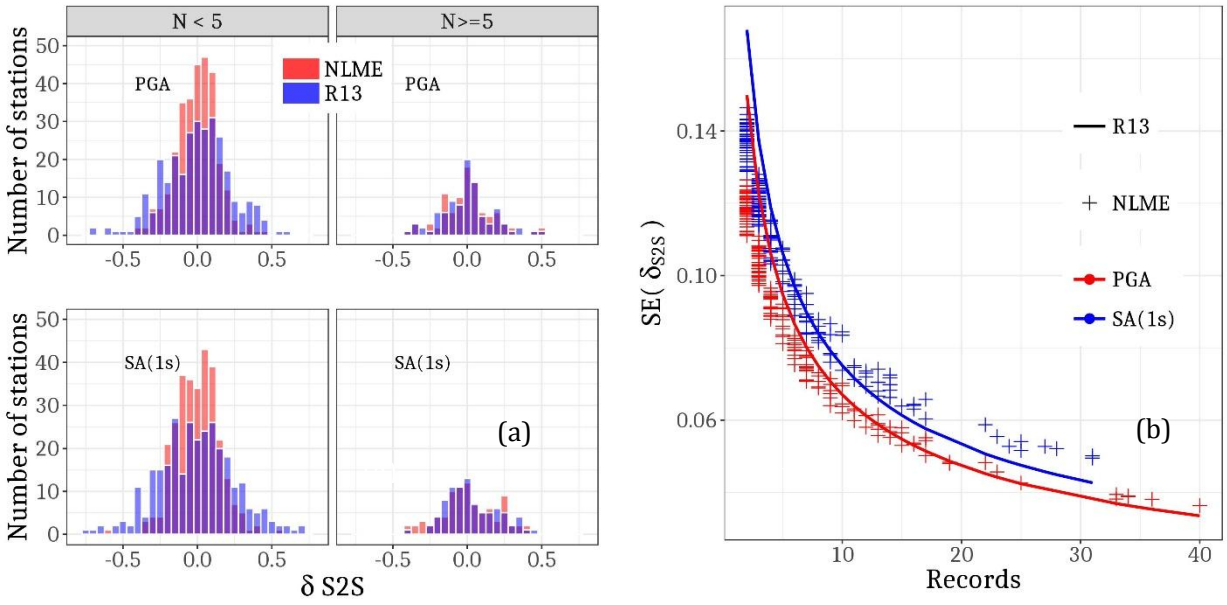


Figure 4-2: Comparison of $\delta S2S$ from R13 and NLME. (a) Random variable $\delta S2S$ from NLME and R13 for sites with fewer than five records (left column), and more than five records (right column) in RESORCE; (b) SE($\delta S2S_s$) from NLME and R13 against number of recordings at the site.

b) Standard error on $\delta S2S_s$ estimated from R13 is conceptually different from that in NLME. According to Eq. (7), the standard error in R13 is statistically 'an estimate of how far the sample mean is likely to be from the population mean', where the sample mean $\delta S2S_s$ belongs to the ($\delta S2S$) normal distribution with zero-mean and standard deviation ϕ_{S2S}. The $\delta S2S_s$ is estimated using a sample of within-event residuals (Eq. 6); its standard error (Eq. 7) decreases with increasing number of records (n_s) at the site. Therefore, any two sites in the dataset with the same number of records is attributed the same standard error on $\delta S2S_s$. The solid lines (color coded for spectral periods) corresponding to R13 in Figure 4-2b show that SE($\delta S2S_s$) decreases as the number of records at a site increases. SE($\delta S2S_s$) provided by the NLME approach are the conditional variance estimates derived from a Markov-Chain-Monte-Carlo procedure. The NLME labelled markers in Figure 4-2b are site-specific SE($\delta S2S_s$), that

is, two sites with the same number of records may have different estimation errors. However, for practical purposes and as shown by Figure 4-2b, the differences in SE($\delta S2S_s$) from the two methods can be considered not significant. We therefore adopt Eq. (7) from R13 in our NLME approach.

c) **Mean $\delta S2S_s$** are compared in Figure 4-4 along with their 70% confidence interval (±1 SE($\delta S2S_s$)) for two well-recorded sites in the RESORCE dataset. We compared the $\delta S2S_s$ estimates for 38 sites with more than 10 records in the RESORCE dataset, mostly from Italy and Turkey. Our comparison showed that in most cases, the (period dependent) differences in mean $\delta S2S_s$ are smaller than their SE($\delta S2S_s$). In a few cases however, for example the site 5401 (V_{S30} = 412m/s) in Figure 4-4a, the differences in $\delta S2S_s$ at longer spectral periods (T = 1s) exceed the SE($\delta S2S_s$). In conjunction with the negligible differences in median ground-motion at T = 1s for sites with V_{S30} close to 400m/s (Figure 4-1a), such non-negligible differences in $\delta S2S_s$ can lead to significant differences in site-specific ground-motions predicted by the two methods.

The above comparisons motivate us to estimate $\delta S2S$ as random effects rather than as means of residuals (Eq. 6). The benefit is twofold: (1) to avoid bias in the GMPE median for less sampled site-classes (e.g. rock sites and very soft sites) and in the site-terms $\delta S2S_s$, and (2) to accurately estimate ϕ_{S2S}, which is necessary in Eq. (7) to estimate SE($\delta S2S_s$).

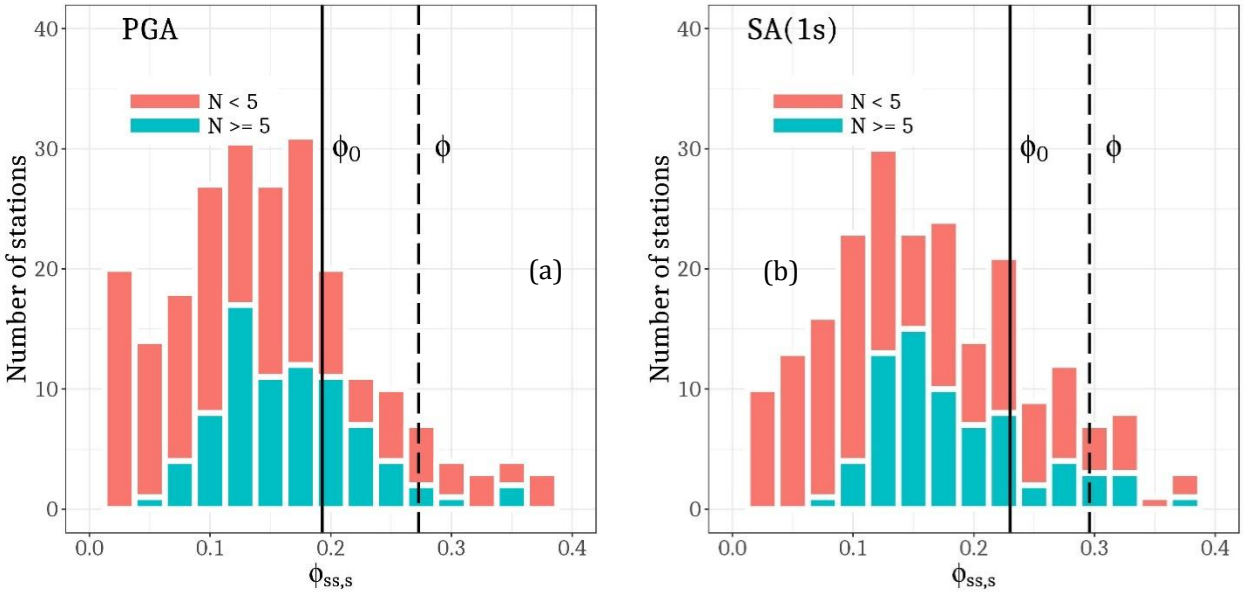


Figure 4-3: Distribution of $\phi_{ss,s}$ for the stations with at least two records in RESORCE dataset at PGA (left) and SA(1s) (right). Stations with at least five recordings are shown in red, the others in blue. The event-and-site corrected variability ϕ_0 and the ergodic within-event variability ϕ are also indicated.

4.3.3.4 Comparison of site-specific variability

In the R13 approach, the standard deviation of the site-specific event-and-site corrected residuals (δWS_{es} at a site) is the so-called site-specific single-site aleatory variability ($\phi_{ss,s}$). *Faccioli et al.* (2015) considered an epistemic uncertainty of $\phi_{ss,s}$ governed by the number of records at the site (Eq. 9). However, as part of the method refinement, in this study we withdrew from the concept of site-specific single-site aleatory variability for the following reasons:

1) A site may have several recordings sampling spatially similar source-to-site paths. In such cases, the resulting $\phi_{ss,s}$ could be an underestimate of the 'true' site-specific aleatory variability in case of propagation medium with significant lateral and vertical heterogeneities.
2) The record-to-record aleatory variability captured by $\phi_{ss,s}$ at any given site is representative of a limited magnitude and distance ranges sampled by the available recordings of events occurring in the vicinity of the site. Since such limited ranges could not fully explore those considered for a PSHA

at the site, the site-specific single-site aleatory variability $\phi_{ss,s}$ may not be a valid representation of the variability required in the PSHA.

Along with the above reasons, use of $\phi_{ss,s}$ limits the application of site-specific PSHA only to well-recorded sites, where the epistemic uncertainty on $\phi_{ss,s}$ is small (Eq. 10 used in *Faccioli et al.* 2015). The aim of method refinement in this study is to perform site-specific PSHA at as many sites as possible in Europe and Middle East, even those with very few records. To avoid issues of misestimating ground-motion variabilities in a site-specific PSHA, we suggest dropping the concept of site-specific single-site aleatory variability ($\phi_{ss,s}$). We recommend using a generic event-and-site corrected variability, ϕ_0, common to all regions and sites. The equivalent of ϕ_0 in the R13 method is ϕ_{ss}, estimated as the standard deviation of the δWS_{es} distribution (*Rodriguez-Marek et al.,* 2013). Figure 4-3 shows the distribution of $\phi_{ss,s}$ from sites with at least two records in RESORCE dataset. In comparison, ϕ_0 is larger than $\phi_{ss,s}$ of majority of sites in the dataset. Conceptually, ϕ_0 is the record-to-record aleatory variability associated with propagation and other sources of variability across all source – site paths in the dataset. Hereon, we refer to the generic site-corrected aleatory variability estimated with τ and ϕ_0 (instead of ϕ in Eq. 2) as σ_0, and use it with a site-specific GMPE median in Eq. (4).

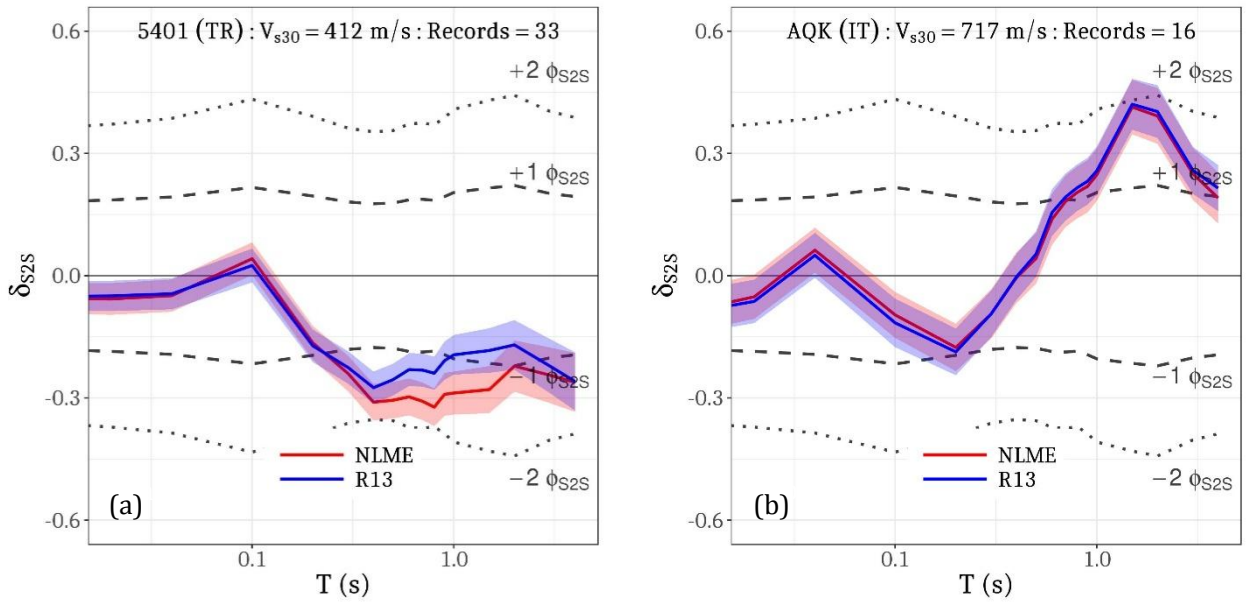


Figure 4-4: Comparison of $\delta S2S_s$ obtained by applying the NLME and R13 methods to two well-recorded sites in RESORCE. (a) Site 5401 in Turkey, (b) Site AQK in Italy. (mean and 70% confidence interval)

4.4 Estimation of the site term for new sites

In the previous sections, we compared the R13 and NLME approaches to derive site-specific GMPEs, provided the sites and records were already present in the regressed strong motion dataset. The R13 method, earlier presented in *Abrahamson and Hollenback* (2012), can also be applied for new sites and events not featured in the GMPE strong motion dataset. As a first step, the between-event random effects (δB_e) for the new event e can be estimated using Eq. (10) of *Abrahamson and Youngs* (1992), shown as Eq. (11) in this study:

$$\delta B_e = \frac{\tau^2 \sum_{s=1}^{n_e} \left(\log\left(GM_{e,s}\right) - \log\left(\mu_{e,s}\right)\right)}{n_e \tau^2 + \phi^2} \qquad (11)$$

In Eq. (11), n_e is the number of available recordings for event e; $GM_{e,s}$ is the observed ground-motion for event e recorded at station s; $\log\left(\mu_{e,s}\right)$ is the corresponding predicted median value; τ and ϕ are the between-event and within-event standard deviations from the GMPE in use. We caution that the Eq. (11) is

applicable only when total aleatory variability σ of the GMPE is homoscedastic. In case of a heteroscedastic ϕ model (as in NGA-West2 GMPEs), Eq. (8-11) of *Stafford* (2012) are more appropriate for estimating the between-event residual δB_e. In this study however, we consider Eq. (11) to be sufficient. Once δB_e for all events are removed from the $\log(G_{e,s}) - \log(\mu_{e,s})$, we obtain the δW_{es} for the new site s, then Eq. (6) and (7) from the R13 method can then be applied to compute $\delta S2S_s$ and SE($\delta S2S_s$). In this study, we propose an alternative to Eq. (6). Under the assumption that the site-specific adjustment ($\delta S2S_s$) for any new site will be from population of sites (the $\delta S2S$ random variable) used to estimate ϕ_{S2S} (as with $\delta S2S_s$ and τ), we suggest using Eq. (12) to estimate $\delta S2S_s$ of the new site.

$$\delta S2S_s = \frac{\Phi_{S2S}^2 \sum_{e=1}^{n_s} \delta W_{es}}{n_s \phi_{S2S}^2 + {\phi_0}^2 + \tau^2} \tag{12}$$

Where, n_s is the number of records at the new site from events not included in the GMPE regression. Note that Eq. (12) is an equivalent of Eq. (11) but to estimate site-terms as between-site random-effects. For the NLME approach introduced in this study, we note that SE($\delta S2S_s$) can still be estimated using Eq. (7) as in R13. It is however important to use an a-priori estimate of ϕ_{S2S} in both Eq. (7) and (12). At the time of this study, few GMPEs (e.g. *Kotha et al.*, 2016; *Luzi et al.*, 2014) provide the dataset-dependent ϕ_{S2S}, thereby limiting the usage of Eq. (12). In cases where ϕ_{S2S} is available for a dataset:

i. For sites with no strong motion records, $\delta S2S_s$ can be set to zero (center of the normal random variable $\delta S2S$) with an SE($\delta S2S_s$) equal to ϕ_{S2S} (*Abrahamson and Hollenback,* 2012). This essentially means that ergodicity cannot be resolved for sites with no records. Instead, ground response analysis based methods, e.g. *Rodriguez-Marek et al.* (2014), can be used to perform a site-specific PSHA at sites with no strong motion recordings. However, using an analytical (not empirical as is in this study) estimate of $\delta S2S_s$ demands a more thorough analysis of SE($\delta S2S_s$), which depends not only on the number of records but also on the choice of 1D vs 2D models, the uncertainty in soil shear-wave velocity profile at the site, and quality of other input parameters.
ii. For sites with a single strong motion record, a theoretical estimate of $\delta S2S_s$ can be obtained using a single within-event residual in Eq. (12). Note that in such cases even though the SE($\delta S2S_s$) (equal to ϕ_{S2S} when $n_s = 1$ in Eq. 7) is large, the median site-specific ground-motion (Eq. 4) may be significantly different from the ergodic prediction depending on how positively or negatively large is the unique δW_{es} (and $\delta S2S_s$).
iii. With increasing numbers of records, Eq. (6) asymptotically approaches Eq. (12) (*Abrahamson and Youngs*, 1992), while Eq. (7) governs the epistemic uncertainty. In a similar context, *Rathje et al.* (2010) demonstrated that a stable median surface response spectrum could be attained with as few as 5 input ground-motions through an equivalent-linear site response analysis. Indeed, it is desirable to have a large number of recordings to well-constrain a $\delta S2S_s$, but in our proposed methodology, we find it reasonable to trade the unavailability of records for a large SE($\delta S2S_s$). For instance, the uncertainty SE($\delta S2S_s$) = 70% of ϕ_{S2S} for sites with two recordings, 45% of ϕ_{S2S} for sites with 5 recordings, and 32% of ϕ_{S2S} for sites with 10 recordings.

To compare the estimates of $\delta S2S_s$ for new sites with new strong motion recordings using Eq. (12) and Eq. (6), we performed the following experiment with the RESORCE dataset:

i. All strong motion data recorded after December 2008 are removed from the RESORCE dataset to obtain a temporally truncated dataset RESORCE_2009. As a result, all strong motion data from L'Aquila sequence of 2009 (Italy) and later are removed from RESORCE, leaving the sites in that region with very few or no recordings.
ii. A GMPE is regressed on the truncated dataset using the same record selection criteria, and residual partition into between-event, between-site and event-and-site corrected residuals. The new GMPE derived from RESORCE_2009 is thus labelled '2009' as opposed to the GMPE using all strong motion data, '2012'. The GMPE '2012' is the same as the NLME-GMPE derived in previous sections.

Figure 4-5a shows that the predicted median response spectra from the two GMPEs ('2009' and '2012') are identical for all practical purposes despite removing 167 recordings, 34 sites, and 18 events from the dataset. Figure 4-5b shows that the between-site variability ϕ_{S2S} is largely unchanged from removal of 34 sites in Italy, while σ_0 is identical among the two GMPEs. Interestingly these differences are smaller than those noticed in Figure 4-1a and Figure 4-1b. The fraction of data removed from the full dataset seems to have no effect on the median prediction and the residual standard deviations of the GMPEs.

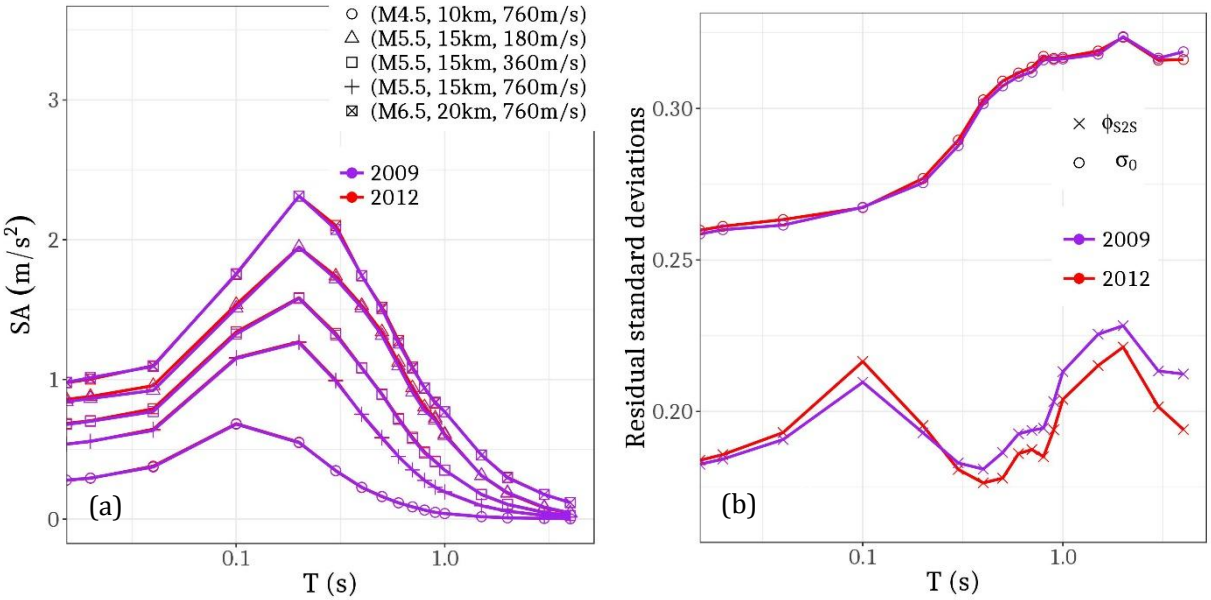


Figure 4-5: Comparison of GMPEs using data until 2009 and 2012: 2009 corresponds to the GMPE considering only recorded strong motion data up to 2009, 2012 corresponds to the GMPE regressed from the full RESORCE dataset including the L'Aquila sequence. (a) Comparison of predicted medians for typical (M, R, V_{S30}) scenarios at a site. (b) Comparison of residual standard deviations.

Using the '2009' GMPE and Eq. 11, we first estimate the event-terms for the 18 events that occurred between 2009 and 2012. In the next step, to estimate $\delta S2S_s(T)$ one can use: (1) Eq. (6) of the R13 method, (2) Eq. (12) introduced in this study or, (3) re-derive the GMPE introducing the new data into the dataset. The first and second approaches do not require a new GMPE regression. The third approach, accessible only to the GMPE developers, in this case yields the '2012' GMPE, whose $\delta S2S_s(T)$ are our benchmark estimates. Figure 4-6 compares the $\delta S2S_s(T)$ estimates from the Eq. (6) and Eq. (12) against the benchmark values at two new stations: AQV, located in central Italy close to L'Aquila and MI03, a temporary station installed in the same area (Onna village) after the 2009 L'Aquila earthquake. RESORCE_2009 contains one recording for AQV and none for MI03, which increased to 13 and six recordings respectively by the end of 2012. The comparisons in Figure 4-6 exemplify the results obtained for all the station considered in the numerical test. In particular, while $\delta S2S_s$ from Eq. (12) always fall within the mean $\pm$1SE($\delta S2S_s$) of the benchmark values, the results from Eq. (6) show trends similar to the benchmark but sometimes with large differences in the amplitude. In case of AQV, at $T = 0.1$s Eq. (6) (R13) estimated $\delta S2S_s = 0.40$, while Eq. (12) (this study) estimated $\delta S2S_s = 0.32$, which implies a site-specific amplification from the R13 approach approximately 27% larger than from NLME.

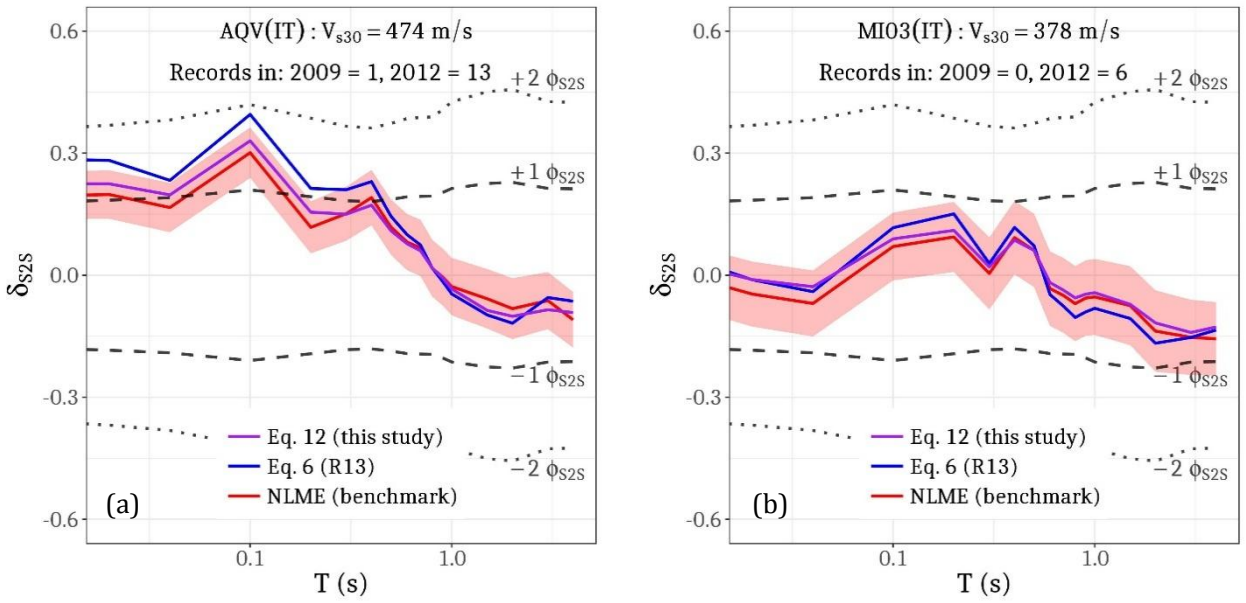


Figure 4-6: Comparison of mean site-terms estimated for the new site using R13 and NLME methods: Eq.(6) of R13 andEq.(12).of this study against the benchmark estimate as random-effect from a GMPE derived with all data until 2012.

4.5 Development of a Region- and Site-specific PSHA framework

Following the discussions in the previous sections, we propose a framework for region- and site-specific probabilistic hazard assessment (Figure 4-7). The framework is constructed over a GMPE including regional attributes in the model for the median and with variance partitioned into between-event (τ), between-site (ϕ_{S2S}), and event-and-site corrected (ϕ_0) components. In the region-specific workflow (black and green blocks in Figure 4-7) the hazard calculation is performed in a single step considering the region-specific median and all the three variance components ($\sigma = \sqrt[2]{\tau^2 + \phi_{S2S}^2 + \phi_0^2}$). In the region- and site-specific workflow (black and red blocks in Figure 7) the hazard is assessed in two steps. In the first step, preliminary hazard curves are estimated using the region-specific median and only two variance components ($\sigma_0 = \sqrt[2]{\tau^2 + \phi_0^2}$). In a second post-processing step, the reference ground-motions levels of the preliminary hazard curves are multiplied with the site-specific $e^{\delta S2S_s}$ to obtain region- and site-specific hazard curves.

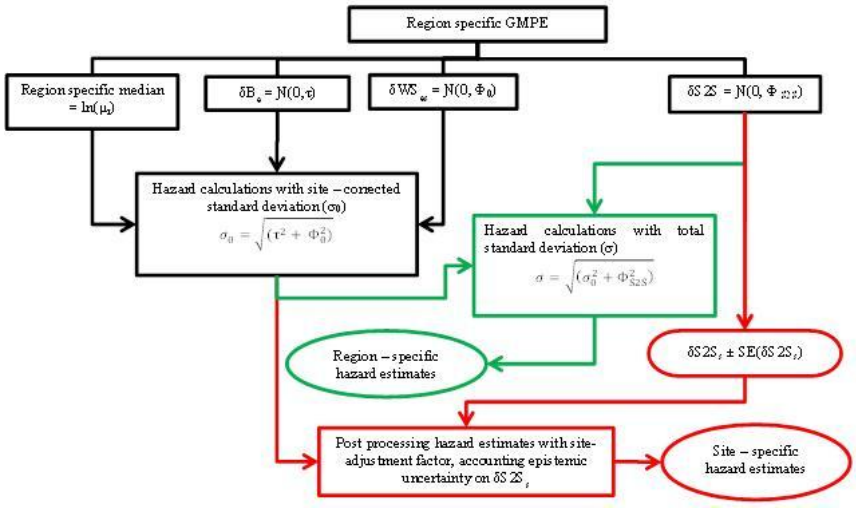


Figure 4-7: Flowchart for region-specific PSHA and region- and site-specific PSHA: Workflow for region-specific (black and green) and region- and site-specific PSHA (black and red).

It is important to note that the workflow of Figure 4-7 is only applicable when δS2Ss and SE(δS2Ss) are magnitude and distance independent (*Iervolino*, 2016). Within the scope of this study, using the pan-European RESORCE dataset, wherein a large fraction of data is from small-moderate magnitude events ($M5$-$M6$) recorded at epicentral distances larger than 10km, it was not possible to observe and empirically model a nonlinear site-response model. The obtained $\delta S2S_s$ are then considered as resulting from linear site amplifications independent from magnitude and distance. SE($\delta S2Ss$) however could depend on the recorded seismicity. For instance, Figure 4-2b shows that sites with equal number of records exhibit unequal NLME estimated SE($\delta S2S_s$), which implies SE($\delta S2S_s$) is not solely dependent on the number of records but the sample of records used to estimate $\delta S2Ss$. To manage this complexity in our workflow, we use Eq. (7) to estimate an approximate magnitude and distance independent SE($\delta S2Ss$). For the pan-European region, we however consider the workflow in Figure 4-7 as a relatively simple and computationally efficient procedure to perform region- and site-specific PSHA at a large number of sites in a region. With a significantly larger dataset (e.g. NGA-West2 dataset) with several well-sampled sites recording both small and large events at near source distances, the magnitude and distance dependence of both site-specific linear and non-linear site-response adjustment factors is worth further investigation.

4.5.1 Partially ergodic region- and site-specific GMPE

Bindi et al. (2014) is the ergodic GMPE considered in this study. Using the NLME algorithm (*Bates et al.*, 2014), *Kotha et al.* (2016) regionalized the *Bindi et al.* (2014) GMPE for Italy and Turkey. In particular, regional adjustments were introduced into the distance decay and linear site-response components of the GMPE. Together these regional adjustments account for possible regional differences in crustal attenuation properties and average V_{S30} profiles. Consequent of regional adjustments to the V_{S30} scaling term is that the resulting NLME-estimated $\delta S2S_s$ in this study are corrected for regional differences. Note that while *Bindi et al.* (2014) uses the common-log, *Kotha et al.* (2016) uses the natural-log of SAs.

4.5.2 Source model

The source model used in this study is adopted from the SHARE seismic hazard model (*Woessner et al.*, 2015). The SHARE source model consists of three alternative source models in the source model logic tree: (1) an area source model (2) a fault and background source model and (3) a gridded seismicity source model. The area source model is a collection of polygonal seismic sources with distributed seismicity in Europe and Middle East, and is given the highest weight (50%) in the SHARE source model logic tree. For simplicity we only consider the SHARE area source model in this study.

4.5.3 Target sites

We used a subset of the 384 sites with V_{S30} provided in the RESORCE dataset. 225 of these sites have at least two recordings (green symbols in Figure 4-8), while 80 have more than 5 records (red symbols in Figure 4-8), and 38 have 10 or more records. Prior works recommended using at least 5 recordings to constrain the linear response at site (e.g. *Rathje et al.* 2010), but for the purpose of this study, which is to demonstrate the shift from ergodic to partially non-ergodic hazard assessment at a large number of sites, we perform a region- and site-specific PSHA for all the 225 sites with at least two recordings - using only sites with more than 5 recordings did not significantly affect the conclusions of this study.

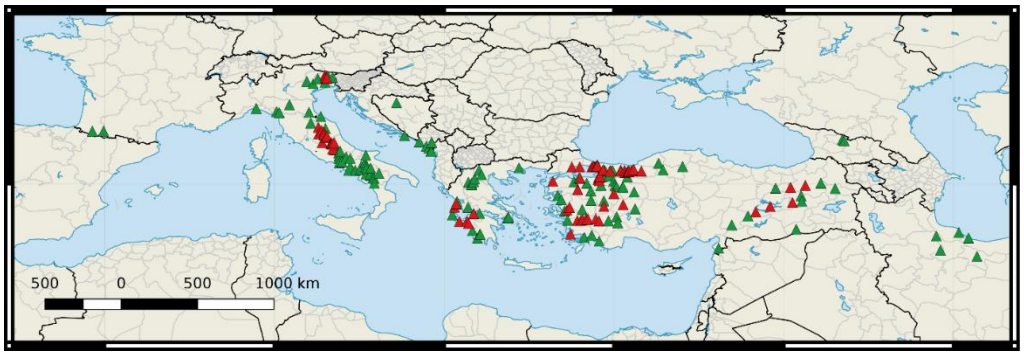


Figure 4-8: Distribution of sites in RESORCE dataset, with at least two recordings (green) and more than five recordings (red).

The seismic source, ground-motion and site models are input into the integration-based Classical PSHA calculator in OpenQuake (*Crowley et al.*, 2015). We obtained ergodic, region-specific, and region- and site-specific Peak Ground Acceleration (PGA) hazard estimates for the 225 sites from the RESORCE dataset with V_{S30} available.

4.6 Comparison of ergodic and partially non-ergodic PSHA

We visualized hazard estimates from the three methods using hazard curves for PGA (g) against return period (years). Figure 4-9 exemplifies the hazard curves at four well-recorded sites from the three PSHA approaches. The epistemic uncertainty on the site-specific hazard curves (red lines) is also considered, where SE($\delta S2S_s$) translates into the upper and lower percentiles of the site-specific hazard estimates (red dashed lines). Although it is important in real applications to consider also the epistemic uncertainty affecting the region-specific hazard curves (green lines), for the sake of graphical clarity the upper and lower percentiles of the region-specific hazard estimates are not shown in this figure. Figure 4-9 shows that, moving from ergodic to region-specific or to site-specific approaches, the hazard can either increase or decrease depending on the trade-off between aleatory variability reduction and site adjustment factor. To get a more complete view of the change in the hazard curves when the ergodic assumption is relaxed, we show the comparison for the 225 selected sites in Europe and Middle East.

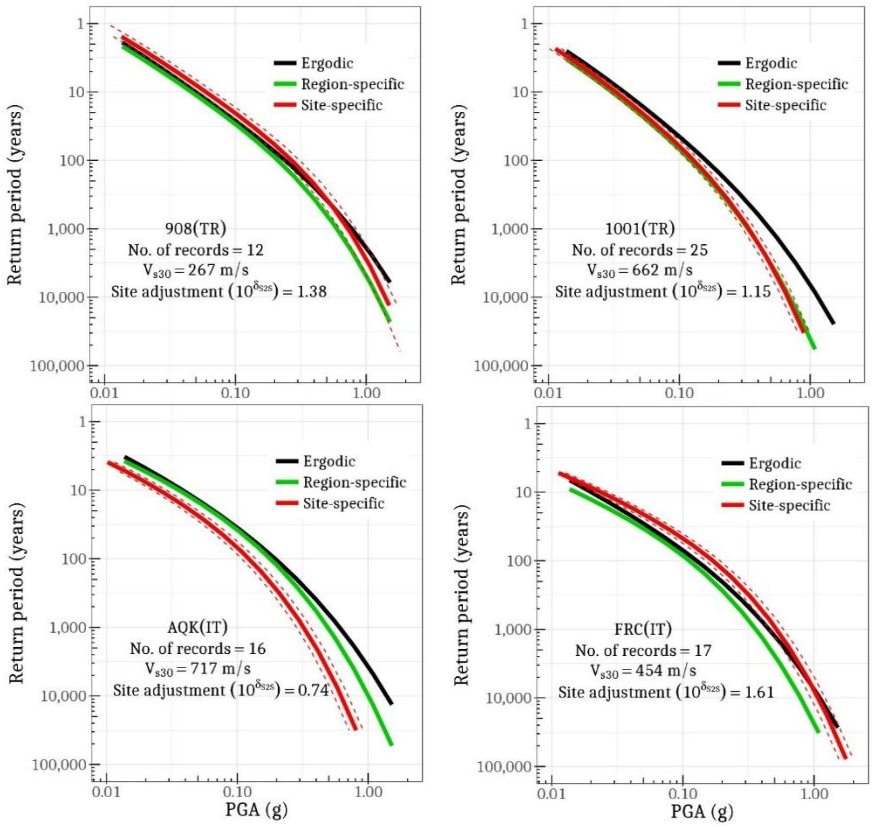


Figure 4-9: Comparison of ergodic hazard curves with partially ergodic hazard curves. In each panel, site name (e.g. 908 in the top left panel), country (TR: Turkey, IT: Italy), number of records, value of V_{S30} and site specific adjustment factor (AF = $e^{\delta_{S2S}}$) are provided.

4.6.1 Region-specific PSHA versus Ergodic PSHA

To analyze the change in hazard estimates due to shift from ergodic (*Bindi et al.*, 2014) to region-specific GMPEs (*Kotha et al.*, 2016); we employ a visually intuitive plot shown in Figure 4-10. In Figure 4-9 it is evident that at the same return period, the predicted ground-motions from the ergodic and regional PSHA are different. Figure 4-10 highlights these differences in terms of percent change in predicted ground-motion values (GMV) at different return periods for all 225 sites. The change in predicted GMV is calculated at each return period using Eq. (13).

$$\Delta GMV(\%) = \frac{100 * (GMV_{region} - GMV_{ergodic})}{GMV_{ergodic}} \quad (13)$$

From Figure 4-10, GMPE regionalization appears to have a significant impact on the hazard estimates at a large number of sites. For example, at the site AQK (IT) the 1000 years return period ergodic PGA is $0.63g$, while the region-specific prediction is $0.51g$, which is a 20% reduction due to GMPE regionalization. We also note that the changes are region-dependent. For example, for the sites in Turkey using a Turkey-specific GMPE median (and a smaller σ), the hazard estimates at all return periods for all sites are reduced. It is remarkable that predicted ground-motion has decreased by more than 25% at a 1000 year return period.

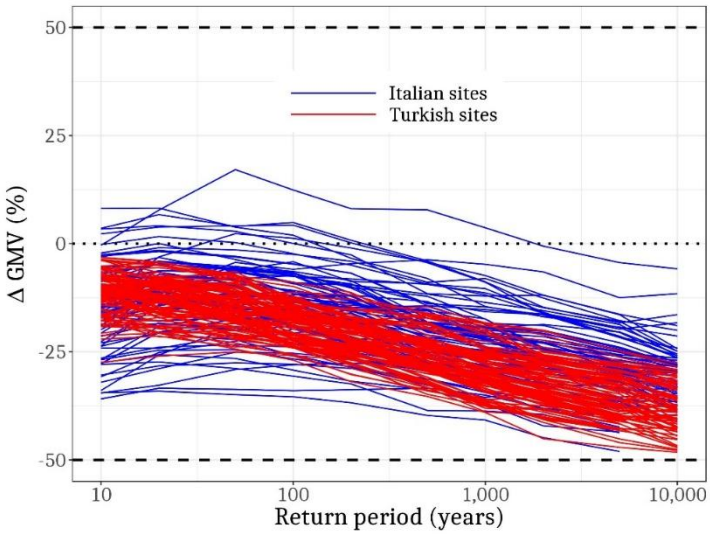


Figure 4-10: Percent change in estimated ground-motion values at each return period, shifting from ergodic to region-specific PSHA. The results are color-coded according to the region.

4.6.2 Region- and site-specific PSHA versus Ergodic PSHA

We performed region- and site-specific PSHA for the 225 sites in RESORCE dataset with at least two records. Similar to Figure 4-10, in Figure 4-11 we provided the difference in predicted GMVs as result of shifting from ergodic to region- and site-specific PSHA for the 225 sites. Also indicated in the figure are the four well-recorded sites from Figure 4-9 along with their adjustment factors (AF = $e^{\delta S2S}$). While region-specific PSHA predicts around 25% change in predicted GMVs (depending on the region, site and return period), the region- and site-specific PSHA predicts even larger changes, in some cases exceeding 50%. For example, at the site AQK (IT) the 1000 year return period ergodic PGA is 0.63g (Figure 4-9) and region-specific prediction is 0.51g (20% decrease), while the region- and site-specific prediction is 46% lower at 0.34g.

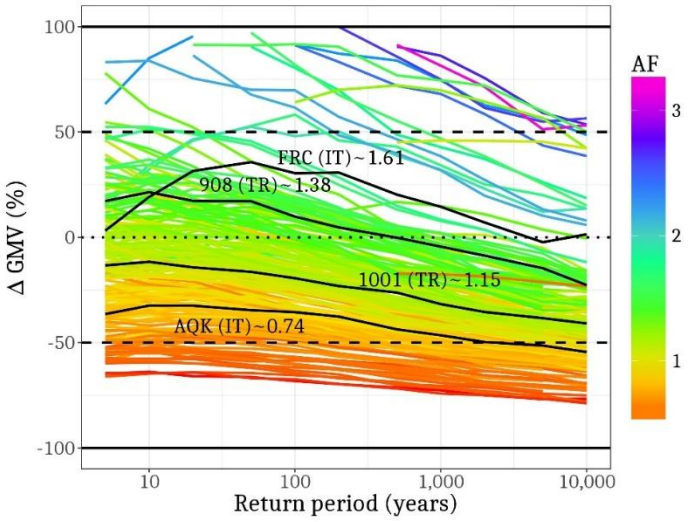


Figure 4-11: Percent change in estimated ground-motion values at each return period shifting from ergodic to region- and site-specific PSHA. The results for 225 sites in Europe-Middle East are color coded according to the site response adjustment factor (AF = $e^{\delta S2S}$). The black curves exemplify the results for the four selected sites (FRC and AQK in Italy; 908 and 1001 in Turkey) shown in Figure 4-9.

4.7 Conclusions

In the framework of Probabilistic Seismic Hazard Assessment, the seismic source model, the ground-motion model and the site model exhibit inherent natural randomness and modelling uncertainties. For Ground-Motion Predictive Equations (GMPEs), the aleatory variability is taken into account by sigma (σ) which, under the ergodic assumption, is inflated with source, path, and site variabilities across several regions. With the availability of large strong motion datasets, it is possible to gradually relax the ergodic assumption: increasing strong motion data from various regions allows regionalization of GMPEs, while increased amounts of data from individual sites are used to develop site-specific GMPEs.

Residual analyses show that linear site-response can be empirically quantified as site-specific linear response adjustment factors ($\delta S2S_s$) with an epistemic uncertainty. $\delta S2S_s$ is then introduced into an ergodic GMPE whilst replacing the ergodic variability (σ) with a single-site site-specific aleatory variability ($\sigma_{ss,s}$). In this study, we introduce a novel region- and site-specific PSHA framework with necessary statistical refinements to estimation of $\delta S2S_s$ and its epistemic uncertainty. We suggest estimating $\delta S2S_s$ for new sites as a random effect rather than as means of residuals, which requires a priori estimation of the site-to-site variability ϕ_{S2S}. The epistemic uncertainty of $\delta S2S_s$ decreases with increasing number of records at the site. Both $\delta S2S_s$ and its uncertainty estimates require ϕ_{S2S}, we therefore suggest it to be estimated as a (dataset-dependent) between-site random-variance in the mixed-effects regression of a GMPE.

A large number of strong motion records at a site are required to constrain and use the site-specific single-site aleatory variability $\phi_{ss,s}$ in a site-specific PSHA. Moreover, it is customary in PSHA applications to assume a wider range of possible magnitudes, distances, source – site azimuths, than those recorded at a site. In such cases $\phi_{ss,s}$ may be an over- or underestimation depending on the collected sample of strong motion recordings at a site. In this study, we aim to apply empirical site-specific PSHA at as many sites as possible, even with a very few strong motion recordings. For this purpose, we chose to drop the concept of $\phi_{ss,s}$, and replace it with a generic event-and-site corrected variability ϕ_0, common for all sites. The refinements to estimation of $\delta S2S_s$, its uncertainty, and the introduction of ϕ_0, allows us to apply empirical site-specific PSHA at any site with a strong motion recording, provided all the uncertainties are carried.

We evaluated changes in hazard estimates at 225 sites in Europe-Middle East using region-specific and region- and site-specific GMPEs instead of an ergodic GMPE. The differences in hazard estimates are computed as percent change in predicted PGA ground-motion for a site and return period due to shift from ergodic to region-specific to region- and site-specific PSHA. Based on our observations, we expect around 25% variation with region-specific GMPEs, and variations as large as 50% with region- and site-specific GMPEs. The large changes observed in predicted ground-motion values are a strong motivation to develop region- and site-specific GMPEs which in turn require increased seismological monitoring of critical facilities and urban areas, thus driving the shift towards fully non-ergodic PSHA.

Acknowledgments

We thank Dr. Julian Bommer, Dr. Brendon Bradley, an anonymous reviewer, and the editorial board for their valuable insights and encouragement to revise the manuscript. We extend our appreciation to Dr. Anne Strader and Ms. Solveig Strutzke for their kind assistance in improving the manuscript presentation. The methods presented in this paper have been developed within the framework of the EPOS-IP European GMPE service (EU-H2020 project n° 676564).

5
A New Approach to Site Classification: Mixed-Effects Ground-Motion Prediction Equation with Spectral Clustering of Site Amplification Functions

S. R. Kotha[1, 2], F. Cotton[1, 2], D. Bindi[1]

[1]German Research Centre for Geosciences GFZ, Potsdam, Germany

[2] University of Potsdam, Potsdam, Germany

Abstract

With increasing amount of strong motion data, Ground-Motion Prediction Equation (GMPE) developers are able to quantify empirical site amplification functions ($\Delta S2S_s$) from GMPE residuals, for use in site-specific Probabilistic Seismic Hazard Assessment. In this study, we first derive a GMPE for 5% damped Pseudo Spectral Acceleration (g) of Active Shallow Crustal earthquakes in Japan with $3.4 \leq M_w \leq 7.3$ and $0 \leq R_{JB} < 600$ km. Using k-mean spectral clustering technique, we then classify our estimated $\Delta S2S_s (T = 0.01s - 2s)$ of 588 well-characterized sites, into 8 site clusters with distinct mean site amplification functions, and within-cluster site-to-site variability ~50% smaller than the overall dataset variability (ϕ_{S2S}). Following an evaluation of existing schemes, we propose a revised data-driven site classification characterized by kernel density distributions of V_{S30}, V_{S10}, H_{800}, and predominant period (T_G) of the site clusters.

Keywords: Mixed-effects regression, Ground-Motion Prediction Equation, Site classification, Spectral Clustering Analysis, Empirical Site Amplification Functions

5.1 Introduction

Current seismic code provisions take into account the significant role of local site conditions on earthquake shaking. Their influence is described through appropriate elastic design spectra based on different site categories. The main parameter proposed for soil categorization is the V_{S30}, i.e. the time-based average value of shear wave velocity (V_S) in the upper 30 m of the soil profile. This parameter has been introduced by *Borcherdt and Glassmoyer* (1992) and *Borcherdt* (1994) as a means to classification of sites for building codes. For example, Eurocode 8 (*Code P*, 2005) and *Rey et al.,* (2002) recommends a site classification based on V_{S30}, and two families of spectral shapes depending on the seismic activity level of area (Type I for active areas, and Type II for moderately active areas).

A number of authors (*Castellaro et al.,* 2008; *Kokusho and Sato,* 2008; *Lee and Trifunac,* 2010, *Héloïse et al.,* 2012) have drawn attention to the limitations of V_{S30} parameter, which is only a proxy and cannot describe alone the physics of site amplification across a broad period (or frequency) range. A number of other proxies (or combinations of proxies) were proposed, coupling information on the shallow impedance and the overall sedimentary thickness. There are several recent studies aimed at developing new and more refined site classification schemes taking into account these additional information (e.g., *Cadet et al.,* 2008; *Gallipoli and Mucciarelli,* 2009; *Luzi et al.,* 2011). For example, *Pitilakis et al.* (2013) introduced a more refined classification using H_{800} (depth to seismic bedrock with $V_S = 800$m/s), $V_{S,av}$ (average shear-wave velocity of the soil column) and fundamental period (f_0). In total *Pitilakis et al.* (2013) suggested 12 site classes for the two European seismicity classes (Type I and Type II). Defining new classifications schemes is however highly challenging because of a few technical issues:

- Only a minimum sufficient number of classes is desirable. The optimal choice of the number of classes is however difficult to define. Ideally the site-to-site variability within each site class should be small compared to a less resolved site classification which, to our knowledge, was not quantitatively analyzed. Moreover, enough recorded strong motion data within each class is seldom available to define statistically well-constrained amplifications factors.
- Only few studies (e.g., *Derras et al.*, 2016) tested the relative efficiency of the various site conditions proxies (e.g., H_{800}, f_0, and V_{S30}) to predict soil amplifications. There is often little consensus on the way to choose and combine the site proxies.
- Site class definitions should avoid unphysical discontinuities in amplification coefficients at the boundaries of adjacent classes. However, such discontinuities are to be expected when using discrete site classes, as opposed to continuous functions of site-response proxies.

In order to resolve some of these issues we explore a new approach to derive a new site classification and site amplification functions. Our aim is to develop a data-driven classification scheme with minimal a priori conditions. For this purpose we adopt the following steps:

1. We take advantage of a high quality dataset featuring several well-characterized sites recordings multiple earthquakes in a region. In this study, we use the KiK-net dataset built by *Dawood et al.* (2016), consisting of 1164 shallow crustal events recorded at 644 sites with several geotechnical site parameters available – e.g. V_{S30} and H_{800} values have been directly derived from down-hole measurements of V_S profile. Further description of the data set is provided in the section titled Data.
2. The empirical site amplification factors are products of a Ground-Motion Prediction Equation (GMPE) mixed-effects analysis. Essentially, we develop a site-specific GMPE from the selected strong motion dataset following the steps described in *Rodriguez-Marek et al.* (2013) and *Kotha et al.* (2017). Details on the GMPE development and mixed-effects analysis is provided in section Ground-Motion Prediction Equation
3. The site amplification factors obtained in the second step are subject to spectral clustering analysis to identify sites with similar response. An optimal number of classes is chosen to minimize both: the site-to-site variability within each site cluster/class and the similarity of their mean amplification

functions. In section Spectral Clustering Analysis, we provide a description of the technique and its application.

4. In the final step, we check the compatibility of various site-response proxies with site clusters obtained in the third step. Site-response proxies ($H_{800}, V_{S30}, V_{S10}$) are not used a priori to define the classes, but a posteriori to characterize the statistical clustering of site-response. In section Site Classification, we introduce the revised site classification scheme, mean site amplifications associated with each class, and site-to-site variability of amplification within each site class.

5.2 Data

In this study, we use the Kiban-Kyoshin network (*Okada et al.*, 2004) database compiled by *Dawood et al.* (2016) for ground-motion studies. A step-by-step automated protocol used to systematically process about 157,000 KiK-net strong ground-motion recordings obtained between October 1997 and December 2011 is elucidated in *Dawood et al.* (2016) and related appendices. A flatfile with all the metadata and the pseudo spectral acceleration (PSA) of the processed records is uploaded to NEEShub (https://nees.org/resources/7849). In addition to the waveform processing by *Dawood et al.* (2016), we make a more GMPE specific record selection for our regression:

1. *Dawood et al.* (2016) remarked that the hypocentral location and M_w obtained from the F-net catalog are more reliable than the values reported in the KiK-net data files. They matched the KiK-net records to F-net earthquakes and classified the match into five categories (A through E) depending on the error margins on location and M_{JMA}. Category A represents the strictest criteria, Category D contains earthquakes that were manually matched, and Category E contains earthquakes for which no match was found. In our study, we choose only the Category A events, which constitute about 89% of the records
2. While most of the GM records in the dataset correspond to subduction earthquakes, we choose only the Active Shallow Crustal (ACRsh) events classified using the *Garcia et al.* (2012) algorithm. However, to filter out any subduction intra-slab and deep continental events, we chose only the ACRsh events whose F-net reported hypocentral depth is ≤35km (as in the HANSR1 criteria of *Garcia et al.* (2012))
3. Most of the KiK-net sites provide 3-component recordings at both surface and borehole sites. In our study, we use only choose the surface recordings at sites with measured V_{S30} available
4. Each record is associated with a high-pass corner frequency (f_c) which limits the maximum usable period $T_{max} \leq \frac{1}{f_c}$ of the record in a GMPE regression. Since the dataset is compiled from an automatic recording processing procedure described in *Dawood et al.* (2016), we applied a more conservative limit of $T_{max} \leq \frac{0.5}{f_c}$. First, we choose only those event and site combinations for which all the 6-component GMs (at surface and borehole) show a Signal-to-Noise ratio (SNR) ≥ 3 in the bandwidth f_c – 30Hz. Then, for regression at each spectral period (T) we select only those records whose $T_{max} \geq T$
5. Finally, we choose only the earthquakes with at least three usable records after all the selection criteria above are cleared. In doing so, the number of usable records for the GMPE regression at $T = 0.01$s falls from 157,000 to 15,896. The number of usable records further decreases to 6462 at $T = 2$s. The data distribution for GMPE regression at $T = 0.01$s is shown in Figure 1. In all there are 850 events with $3.4 \leq M_w \leq 7.3$, 641 sites with $106 \leq V_{S30} \leq 2100$m/s, and 15,896 records with $0 \leq R_{JB} \leq 543$km

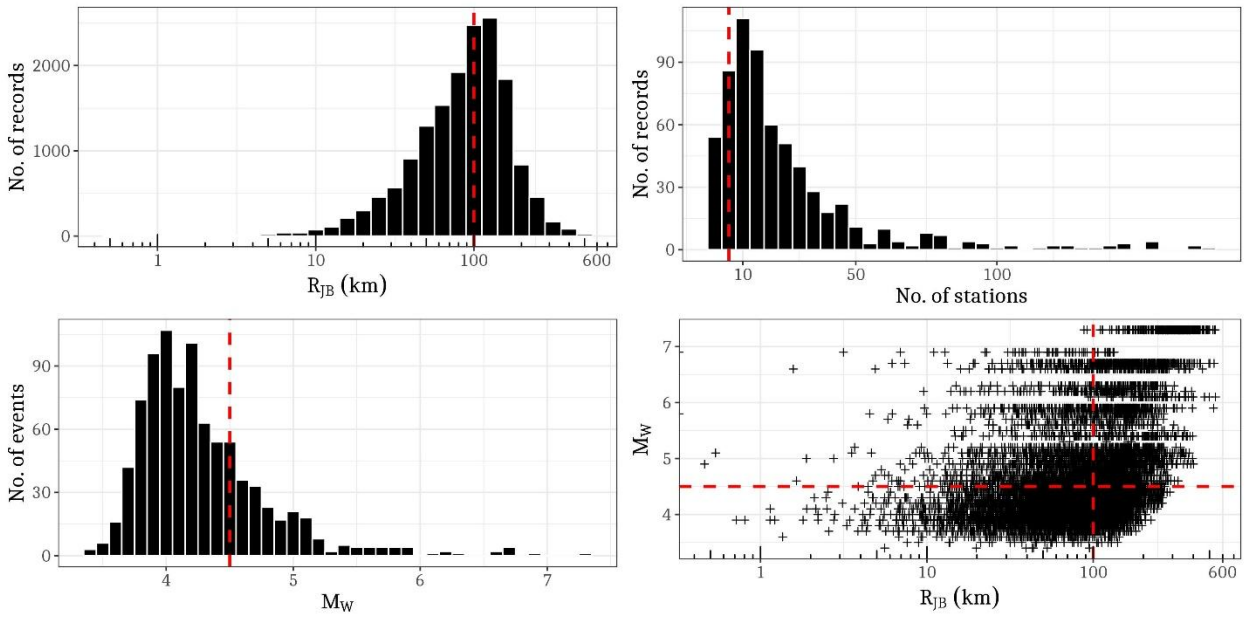


Figure 5-1: Data distribution following the record selection criteria for GMPE regression at $T = 0.01$s: (top-left panel) Distance distribution of usable records, (top-right panel) number of records per station, (bottom-left panel) magnitude distribution of usable records, (bottom-right panel) magnitude-distance scatter plot of usable records.

5.3 Ground-Motion Prediction Equation

Using a mixed-effects regression approach (as in *Abrahamson and Youngs*, 1992; *Kotha et al.*, 2016), we derive a GMPE for the geometric-mean of (5% damped) horizontal Pseudo Spectral Acceleration (PSA) at 33 values of T between 0.01s and 2s.

$$ln(PSA) = f_R(M_w, R_{JB}) + f_M(M_w) + \delta B_e + \delta S2S_s + \delta WS_{e,s} \quad (1)$$

In Eq. (1), the parametric functions $f_R(M_w, R_{JB})$and $f_M(M_w)$ capture the scaling of PSAs with distance and magnitude, respectively, and they are referred to as fixed-effects. δB_e is the between-event random-effect quantifying the systematic deviation of observed ground-motions associated to an event e with respect to the GMPE fixed-effects prediction. $\delta S2S_s$ is the site-specific random-effect for a site s, which can be used to scale the GMPE prediction to a site-specific prediction (e.g., *Rodriguez-Marek et al.*, 2013; *Kotha et al.*, 2017a). $\delta WS_{e,s}$ is the regression residual capturing record-to-record variability (combination of e and s), and can be investigated for other repeatable effects (e.g., *Boore et al.*, 2014; *Kotha et al.*, 2017b; *Baltay et al.*, 2017). The period dependent random-effects and the residuals follow orthogonal normal distributions as $\delta B_e = \mathcal{N}(0, \tau)$, $\delta S2S_s = \mathcal{N}(0, \phi_{S2S})$ and $\delta WS_{e,s} = \mathcal{N}(0, \phi_0)$, where τ is event-to-event or between-event variability, ϕ_{S2S} captures the site-to-site or between-site variability, and ϕ_0 is the event-and-site corrected or residual aleatory variability. Note that the ϕ_0 in this study is the same as the single-station standard deviation ϕ_{ss} of *Rodriguez-Marek et al.* (2013). The total aleatory variability of the dataset with respect to a GMPE is $\sigma = \sqrt{\tau^2 + \phi_{S2S}^2 + \phi_0^2}$.

It is worth noting that Eq. (1) does not include any site-response component in its fixed-effects, unlike the usual practice of including a parametric function of V_{S30}. The site-specific random effects $\delta S2S_s$ absorb all the site-specific response, and serve as the empirical site-specific amplification functions in our proposed site classification scheme. Given that the large fraction of data used in constraining $\delta S2S_s$ is from events with $M_w < 5$ and $R_{JB} > 25$km (Figure 1), these empirical amplification functions capture only the average linear soil response. Therefore, our site classification is solely based on classification linear soil response. To constrain the non-linear soil response in $\delta S2S_s$ would require more data from larger and closer events.

5.3.1 Parametric regression

We develop a GMPE following a multi-step approach where we first calibrate the magnitude-dependent distance scaling function $f_R(M_w, R_{JB})$ and then use distance-corrected observations to calibrate the magnitude scaling function $f_M(M_w)$. At each step, we perform a mixed-effects regression using LMER algorithm of *Bates et al.* (2014) implemented in *R Development Core Team* (2010), estimating both the δB_e and $\delta S2S_s$ random-effects. In doing so, we ensure that the regression coefficients are unbiased by a well-recorded events or sites (e.g. *Boore et al.*, 2014; *Kotha et al.*, 2017a).

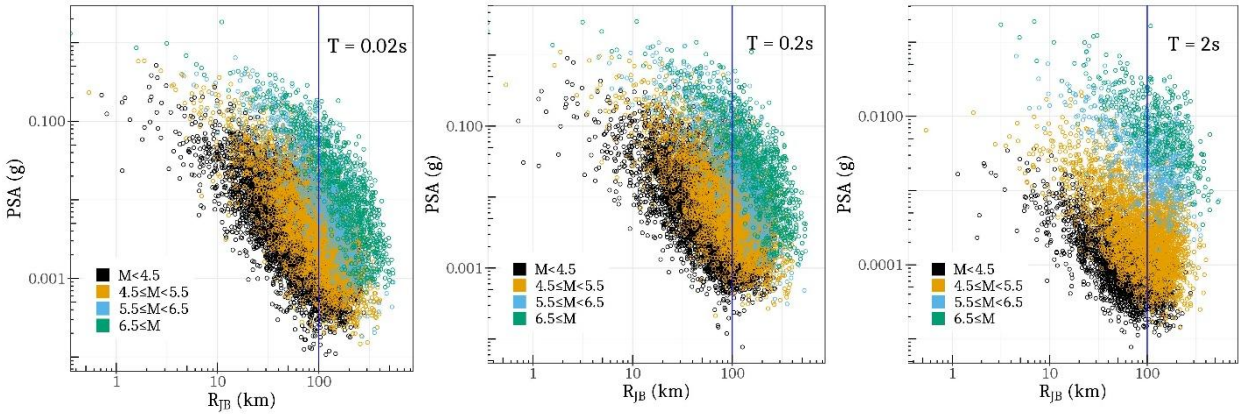


Figure 5-2: Distance scaling of Geometric Mean of 5% damped horizontal Pseudo Spectral Accelerations at T = 0.02s, 0.2s and 2s.

5.3.1.1 Distance scaling: $f_R(M_w, R_{JB})$

The first step in our multi-step regression procedure is to derive the distance scaling component $f_R(M_w, R_{JB})$ of Eq. (1). Figure 5-2 shows the observed *PSA*s at $T = 0.02$s, 0.2s and 2s against the Joyner-Boore distance metric (R_{JB}). The scatter plot is color coded according to magnitude ranges of the observations to identify magnitude dependence of distance scaling.

The distance scaling of *PSA* shows a magnitude-dependent near-source saturation effect, extending to about 5km for $M_w \leq 4.5$ and up to 20km for $M_w \geq 6.5$ events. This effect is generally modeled by introducing the so-called effective-depth or h-parameter in the GMPE distance scaling fixed-effect component. For an equivalent point-source simulation of finite ruptures, we adopted the effective-depth formula (Eq. 2) derived by *Yenier and Atkinson* (2015) from a global dataset of well-recorded earthquakes (including California, Italy, Japan, New Zealand, Taiwan, and Turkey).

$$\ln h = 2.303 \max[(-0.05 + 0.15M_w), (-1.72 + 0.43M_w)] \quad (2)$$

Figure 5-2 also suggests that short (0.02s) and intermediate (0.2s) period *PSA*s attenuate more rapidly beyond 100 km. To capture the apparent anelastic attenuation of high frequency *PSA*s, we introduce a hinge-distance in the definition of $f_R(M_w, R_{JB})$ and model it as in Eq. (3). Note that the only magnitude dependence in distance scaling is from h (Eq. 2). The coefficients c_1, c_2, and c_3 estimated in this step are held constant for the later steps of GMPE regression.

$$f_R(M_w, R_{JB}) = \begin{cases} c_1 \ln\sqrt{R_{JB}^2 + h^2} & R_{JB} < 100km \\ c_1 \ln\sqrt{100^2 + h^2} + c_2 \ln\left(\frac{R_{JB}}{100}\right) + c_3(R_{JB} - 100) & R_{JB} \geq 100km \end{cases} \quad (3)$$

5.3.1.2 *Magnitude scaling:* $f_M(M_w)$

The recorded *PSA*s corrected for $f_R(M_w, R_{JB})$ yield the expected *PSA* at reference distance $R_{ref} = 1$km. Per-event averages of distance scaled *PSA*, hereafter PSA_{Rref}, are the near-source ground-motions corrected for attenuation effects. The GMPE magnitude scaling component $f_M(M_w)$ is then derived through a weighted mixed-effects regression of the PSA_{Rref}.

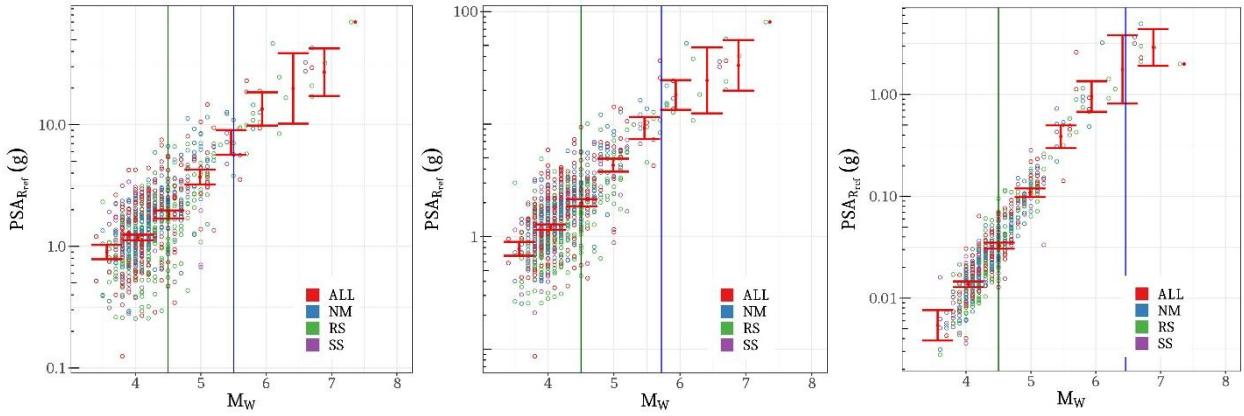


Figure 5-3: Parametric analysis for magnitude scaling at T = 0.02s, 0.2s, 2s. PSA_{Rref} (g) are the per-event averages of recorded PSA_s, corrected for distance scaling (Eq. 3). Colors indicate the focal mechanism of events: ALL – unknown mechanism, NM – Normal faulting, RS – Reverse-slip fault, SS – Strike-slip fault.

Figure 5-3 shows PSA_{Rref} against M_w, color coded according to the focal mechanism of the events. Earlier, *Zhao et al.* (2016) reported that ACRsh events with Normal (NM) focal mechanism produce higher amplitudes compared to Strike-slip (SS) and Reverse events. However, our non-parametric analysis showed no clear distinction of observed PSA_{Rref} with rupture focal mechanism. In addition, our parametric mixed-effects regression showed statistically insignificant variability of magnitude scaling with focal mechanism. Therefore, in this GMPE, we chose to not include any focal mechanism term.

Several recent GMPEs developed for applicability over a wide magnitude range (e.g. $3 < M_w < 8$), adopted piece-wise linear (or a high degree polynomial) expressions in their magnitude scaling components. For example, *Campbell and Bozorgnia* (2014) allowed three period-independent break-points in their magnitude scaling relation at $M_w = 4.5$, 5.5 and 6.5, (*Boore et al.*, 2014) used a single period-dependent magnitude break-point ranging from $M_w = 5.5$ for $T \leq 0.1$s up to $M_w = 6.2$ for $T \geq 0.4$s, while *Zhao et al.* (2016) used a single period-independent break-point at $M_w = 7.1$. Piece-wise magnitude scaling expressions allow variability of magnitude scaling gradient in different M_w-ranges. Especially for imbalanced datasets with several small events and fewer large events, such artificial break-points in magnitude scaling ensure that *PSA*s scale differently for small and large magnitude ranges. For instance, in Figure 5-3, we notice that PSA_{Rref} ($T = 0.02$s) scales more gradually (less positive gradient) for $M_w \leq 4.5$ compared to $M_w > 4.5$. Similarly, PSA_{Rref} ($T = 2$s) scale rapidly for $M_w < 6.5$ compared to $M_w \geq 6.5$ range, where the scaling gradient is close to zero.

Based on our non-parametric analysis, we formulated our $f_M(M_w)$ to have two break-points: 1) at reference magnitude $M_{ref} = 4.5$ to separate the numerous small events ($M_w < 4.5$) from fewer intermediate-large events ($M_w \geq 4.5$), and 2) a period-dependent hinge-magnitude (M_h), to allow over-saturation of PSA_{Rref} for large events. The period-dependence of M_h is inspired by reasoning of *Boore et al.* (2014), where visual inspection of NGA-West2 data suggested M_h to increase with period from $M_h = 5.5$ at $T \leq 0.1$s to $M_h = 6.2$ at $T \geq 0.4$s. However, unlike *Boore et al.* (2014), we allowed M_h to monotonically increase beyond $T = 0.4$s to reach $M_h = 6.46$ at $T = 2$s. Stochastic simulations (e.g. *Schmedes and Archuleta*, 2008) and empirical evidence suggest that ground-motions saturate (or even over-saturate) at close distances from large magnitude events. To account such physical effects, we formulated $f_M(M_w)$ for our GMPE as in Eq. (4).

$$f_M(M_w) = a + \begin{cases} b_1(M_w - M_{ref}) & M_w < M_{ref} \\ b_2(M_w - M_{ref}) & M_{ref} \le M_w < M_h \\ b_2(M_h - M_{ref}) + b_3(M_w - M_h) & M_h \le M_w \end{cases} \quad (4)$$

5.3.1.3 *Random-effects:* δB_e *and* $\delta S2S_s$

During the regression of $f_R(M_w, R_{JB})$ and $f_M(M_w)$ we allow the algorithm to estimate conditional values of δB_e and $\delta S2S_s$. These intermediate estimates are however not the final random-effects of the GMPE, but only to prevent a well-recorded event or site from biasing the fixed-effects coefficient estimates (as discussed in *Kotha et al.*, 2017, and *Stafford*, 2014). The final estimate of random-effects are obtained after correcting the observed *PSA*s for both magnitude and distance scaling effects (as in *Boore et al.*, 2014). For a record from eth event at sth site the residual $\varepsilon_{e,s} = \ln(PSA_{e,s}) - f_R(M_e, R_{e,s}) - f_M(M_e)$ is split into random-effects δB_e and $\delta S2S_s$, and event-and-site corrected residual $\delta WS_{e,s}$. These random-effects and residuals can be further investigated to evaluate the GMPE performance and/or to identify new predictor variables that can be modelled as fixed-effects.

Figure 5-4 is the customary residual analysis performed after a GMPE regression to verify if the fixed-effects components capture well the attenuation of *PSA*s at all magnitudes and distances. If in case the fixed-effects components are biased by a particularly well-sampled magnitude-distance bin in the dataset, then we should expect to see artifacts in the random-effects and the residuals. In the top panels of Figure 5-4, we plotted δB_e versus M_w to evaluate our choice of $f_M(M_w)$. We divide the magnitude range M3.4 – M7.3 into 10 magnitude bins, and calculate the mean and 15th-85th percentile error bars on δB_e within each bin. At all periods, the mean δB_e for each bin falls very close to zero, implying no significant trend with M_w and that $f_M(M_w)$ captures the magnitude scaling of *PSA*s very well.

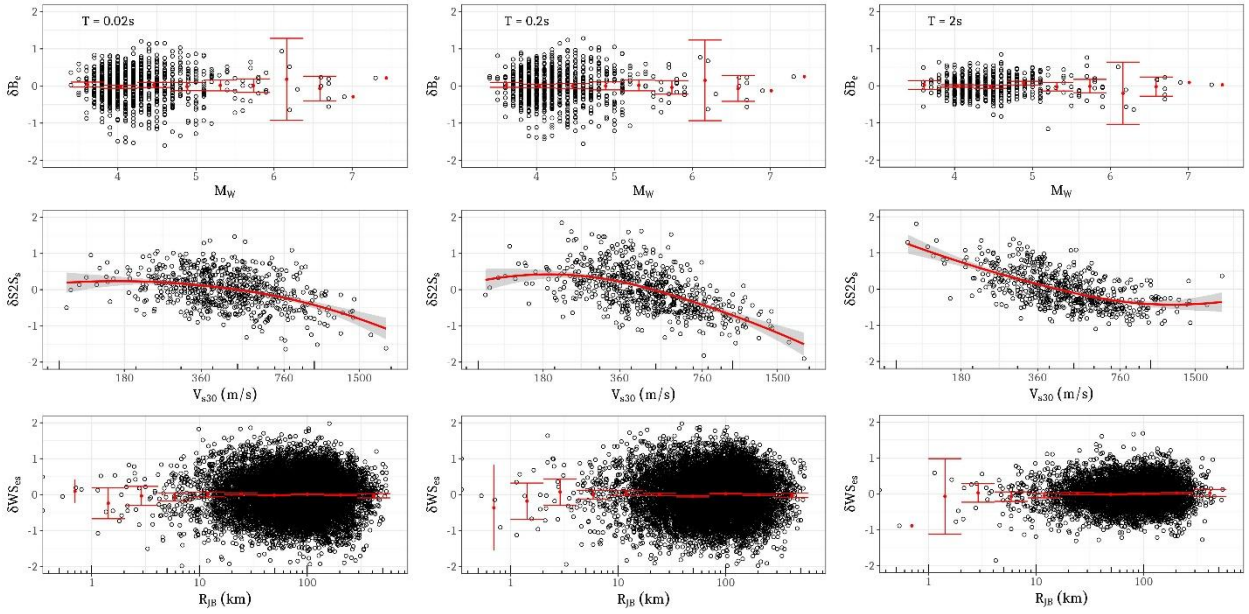


Figure 5-4: Random-effects and residual plots for GMPE evaluation at T = 0.02s, 0.2s, and 2s. In each panel, δB_e is plotted against M_w, $\delta S2S_s$ against V_{S30}, and $\delta WS_{e,s}$ against R_{JB} to check if random-effects and residuals show a systematic trend with predictor variables.

The bottom panels of Figure 5-4 show the event-and-site corrected residuals, $\delta WS_{e,s}$ versus the distance metric, R_{JB}. We recall that $f_R(M_w, R_{JB})$is regressed for data with 0km $\le R_{JB} <$ 600km, which is a considerably larger distance range than any GMPE developed for Active Shallow Crustal environments. Such modeling choice is motivated by the need to constrain the site terms with a large amount of data. Nevertheless, our $f_R(M_w, R_{JB})$ performs very well at all distances, as indicated by the zero mean $\delta WS_{e,s}$ within each distance bin.

The middle panels of Figure 5-4 showing $\delta S2S_s$ versus V_{S30} (in log-scale) is the most important plot of this section. Since a site-response component is deliberately left out of the fixed-effects in GMPE, the random-effects $\delta S2S_s$ show a trend with V_{S30}. However, at $T = 0.02$s (and $T = 0.2$s) gradient of the LOESS fit (Local regression of scatterplots by Cleveland, 1979) of $\delta S2S_s$ versus V_{S30} is close to zero for $V_{S30} < 600$m/s , implying that high frequency soil response is weakly correlated to V_{S30} (also in *Seyhan and Stewart*, 2014). For longer periods ($T = 2$s), although a steeper gradient at $V_{S30} < 200$m/s indicates better relevance of V_{S30}, it appears that low frequency response of stiffer soils may not be captured with V_{S30} alone. Our observations suggest that V_{S30} may not be an ideal proxy of linear site-response. In the later sections, we use these inferences in developing alternative approaches to empirical site-response modelling.

5.4 Results

In Figure 5-5 we plot the distance scaling fixed-effects component of the GMPE against the observed *PSA*s corrected for the between-event and between-site random-effects, i.e. $\ln(PSA_{\delta B_e+\delta S2S_s}) = \ln PSA_{e,s} - \delta B_e - \delta S2S_s$, where e and s are the indices of event and site respectively. The scatter of data around the GMPE median is the record-to-record variability $\delta WS_{e,s}$.

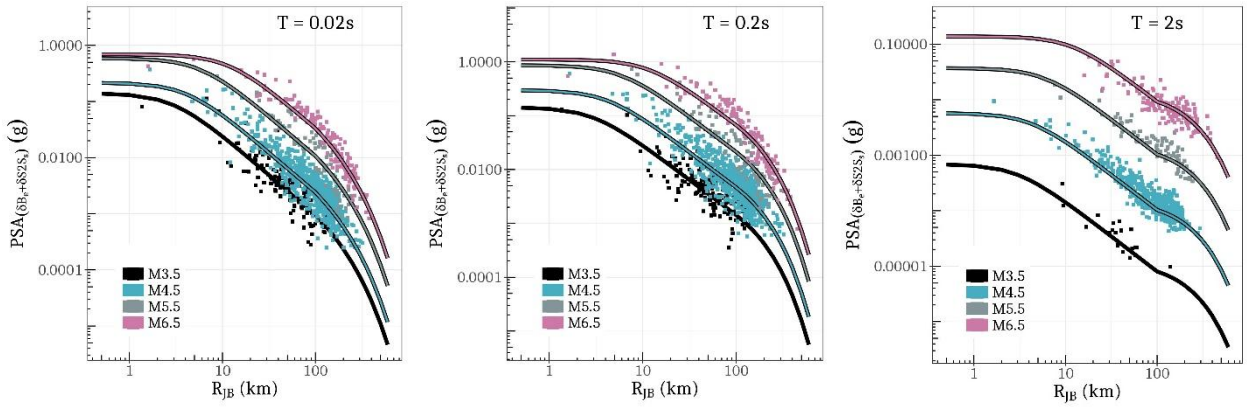


Figure 5-5: Distance scaling of the GMPE fixed-effects at T = 0.02s, 0.2s, and 2s. Note that the vertical axis shows observed PSAs minus the between-event and between-site random-effects.

At all periods, the magnitude-dependent h-parameter appears to capture well the near-source saturation of *PSA*s. Campbell (1981) observed that the near-source attenuation of high frequency *PSA*s (e.g. PGA) is weakly dependent on magnitude and distance, which is evident from the minor differences in our GMPE median predictions at $R_{JB} < 10$km for $M6.5$ and $M7.5$ events in Figure 5-5. Secondly, the fixed-effects coefficient c_3 forces an exponential decay of GMPE *PSA* predictions to mimic the anelastic attenuation of *PSA*s at far-source distances, $R_{JB} > 100$km. In the panel for $T = 2$s, the bump in the predicted *PSA*s at 100km, which is in agreement with the observations as well, indicates a possibility of post-critical reflections at crustal discontinuities and/or transition from body waves at near-source distances to surface waves at far-source distances. This phenomenon is also in agreement with the kink at about 90km observed by *Oth et al.* (2011) in the attenuation of Fourier amplitude spectra at low frequencies in Japan.

Figure 5-6 shows the magnitude scaling of GMPE fixed-effects versus the observed *PSA*s corrected for random-effects δB_e and $\delta S2S_s$. Despite two break-points, one at $M_w = 4.5$ and the other at $M_h = 5.5$-6.5 depending on the period, magnitude scaling for $T = 2$s appears to be constant for all events $M_w \leq 6.5$, and an over-saturation for $M_w > 6.5$. Although the over-saturation appears rather strong, the panels corresponding to $T = 2$s, in our non-parametric analysis of Figure 5-3 do exhibit a decreasing trend of PSA_{Rref}, while that in residual analysis of show s no bias for $M_w > 6.5$. The strong over-saturation at short distances is a combined effect of the h-parameter (Eq. 2) and the period-dependent M_h (Eq. 4).

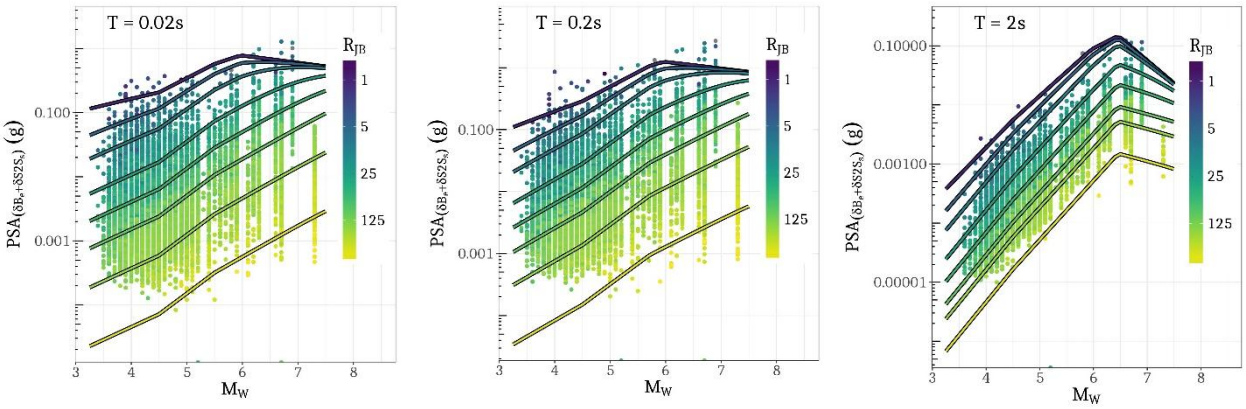


Figure 5-6: Magnitude scaling of the GMPE fixed-effects at T = 0.02s, 0.2s, and 2s. Note that the vertical axis shows observed *PSA*s minus the between-event and between-site random-effects.

Figure 5-7 shows the GMPE fixed-effects response spectra in the left panel, and random-effects standard deviations in the right panel. The GMPE standard deviations in the left panel are in natural-log scale. ϕ_{S2S} is the largest among the three components, indicating a large site-to-site variability in the dataset, and also because we chose not to include any site-response parameters (e.g. V_{S30}) in the GMPE fixed-effects component. In the following sections we introduce the clustering approach to site-response modelling, which is expected to reduce the ϕ_{S2S}, and the total aleatory variability σ to a more reasonable value.

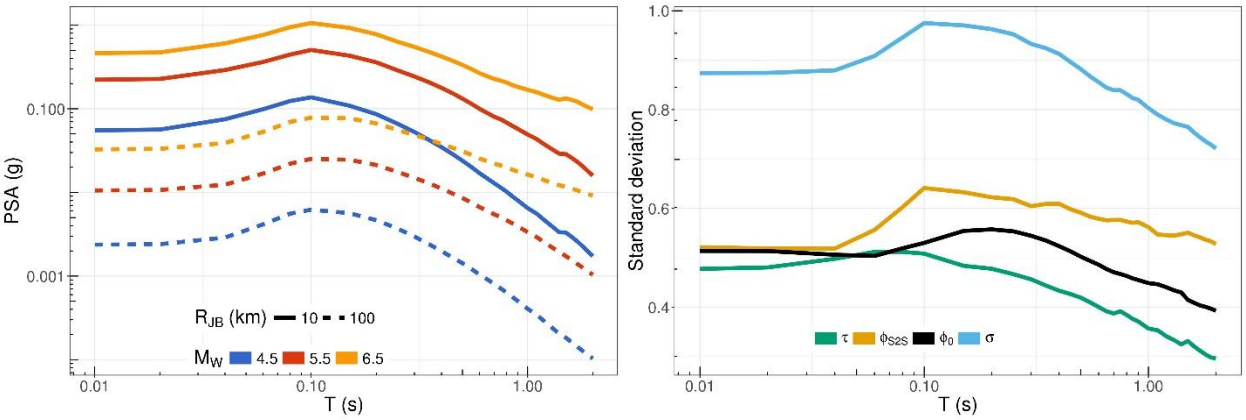


Figure 5-7: Median response spectra, between-event (τ), between-site (ϕ_{S2S}), event-and-site corrected standard deviations (ϕ_0) and the total aleatory variability (σ) of the GMPE in natural-log scale.

In our mixed-effects GMPE, the magnitude and distance scaling are captured by the fixed-effects $f_M(M_w)$ and $f_R(M_w, R_{JB})$, event-specific adjustments by random-effects δB_e, and record-to-record variability by $\delta WS_{e,s}$. Therefore, each site-specific random-effect $\delta S2S_s$ essentially captures the empirical mean site-response of a site in the GMPE regression. For every period at which the GMPE is regressed (T = 0.01s-2s), a well-recorded site in the dataset has an associated period-dependent $\delta S2S_s$. The scalar $\delta S2S_s$ at a given period can be used to adjust the generic GMPE $PSA(T)$ predictions to yield site-specific $PSA_{ss}(T)$ (e.g. *Rodriguez-Marek et al.*, 2013, *Kotha et al.*, 2017a). A vector of site-specific $\delta S2S_s$ for a range of T, notated as $\Delta S2S_s$ from hereon, resembles an empirical site amplification function (AF). $\Delta S2S_s$ can be used to adjust the GMPE PSA response spectra, Conditional Spectra (*Baker*, 2010, and *Kotha et al.*, 2017b), or even the Uniform Hazard Spectra. In this study, the $\Delta S2S_s$ vector for well-recorded sites serve as the empirical site amplification functions.

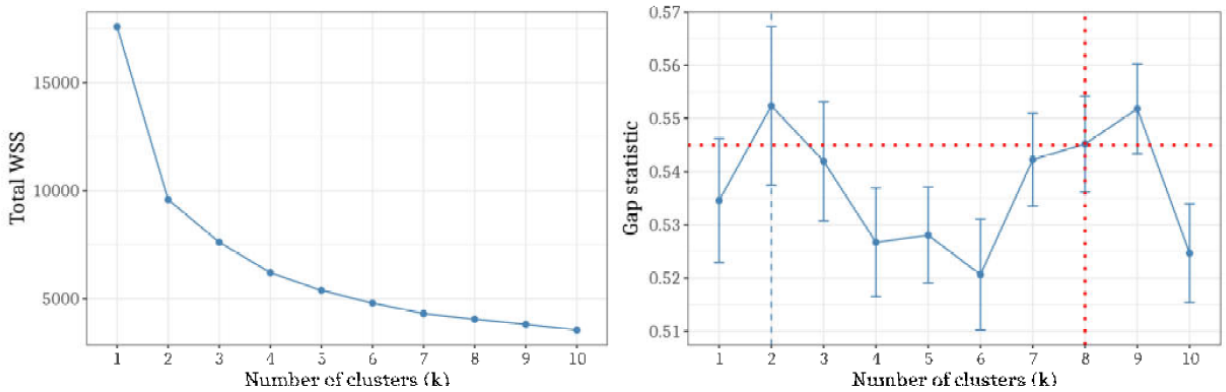


Figure 5-8: Optimal number of clusters based on Total Within Sum of Squares (WSS) and Gap statistic (GS). The WSS metric reduces with increasing number of clusters, but the optimal number of clusters is when Gap statistic is maximized – in which case the WSS is low and the inter-cluster distance is high.

5.5 Spectral clustering analysis

With increasing amount of strong motion data, ground-motion modelers are adapting advanced statistical tools to analyze large amounts of ground-motion data to identify repeatable effects, evaluate parametrization, and quantify better the uncertainties in prediction (e.g. *Derras et al.*, 2012; *Ullah et al.*, 2013; *Luzi et al.*, 2011). Among the various tools, spectral clustering analysis, a type of unsupervised machine learning, refer to a variety of statistical techniques aimed at extracting hidden patterns/structures from large amounts of unlabeled multidimensional data. Constituted by n scalar $\delta S2S_s$ value for n spectral periods (T = 0.01s-2s), the $\Delta S2S_s$ vectors of length n are our multidimensional data points. The steps involved in spectral clustering are the following:

1. *Preparing the data:* $\Delta S2S_s$ vectors of all the sites to be clustered must be of equal length, therefore we only select the 588 sites (of the 641 sites with measured V_{S30}) with $\delta S2S_s$ available at all periods in the range T = 0.01s-2s. $\Delta S2S_s$ vectors of the 588 sites are then normalized with the period-dependent ϕ_{S2S}.
   It is possible to extend the period range of $\Delta S2S_s$ vectors, but the number of sites with a reliable estimate of $\delta S2S_s$ falls rapidly beyond $T > 2$s, which in turn is controlled by the maximum usable period of T_{max} of a record (in our record selection step). In fact, including long period $\delta S2S_s$ (up to T = 5s) in our analysis resulted in better separation of clusters, but with very few sites in each cluster.
2. *Choice of clustering technique:* There are several advanced machine learning techniques depending on the amount of supervision (a priori information) that is input and the knowledge that is being queried. For our purpose, we chose a basic partitioning algorithm: the k-means clustering technique (*MacQueen*, 1967). k-means technique splits the $\Delta S2S_s$ vectors into k groups (clusters), where k must be specified in advance. The clustering algorithm (available in the R library ClusterR by *Lampros Mouselimis* (2017) and factoextra by *Kassambara and Mundt* (2016)) iteratively partitions the data until the Total Within Sum of Squares (WSS) is minimized.
3. *Selection of the number of clusters k:* We use two indices to guide the selection of optimal k: Total Within Sum of Squares (WSS) and the Gap statistic that compares the WSS change with that expected under an appropriate null reference distribution of the data (see *Tibshirani et al.*, 2001, for more details on this statistic). After testing different selections for the number of clusters, we found that k = 8 provides an acceptable WSS reduction without introducing large overlaps among the clusters (Figure 5-8).

5.6 Site classification from cluster analysis

In a way our approach is inverse of the current practice, where the number of site classes (e.g. 5 soil classes - A, B, C, D, E in Eurocode 8 (*Code*, 2005)), preferred site-response proxy (e.g. V_{S30}), and parametric ranges of selected proxy (e.g. sites with $V_{S30} > 800$m/s as EC8-A) are fixed a priori – and then, the available strong motion data is grouped and processed within each class to derive empirical site amplification functions. In our approach, we first derived the empirical site amplification functions ($\Delta S2S_s$) of the 588 sites, and then classified them into 8 k-means clusters. We now present the site clusters and their mean amplification functions. Later, we investigate and identify site-response proxies that can effectively characterize these eight site classes.

5.7 Site clusters

The eight site clusters partitioning the 588 sites in our dataset are visualized in Figure 5-9, and the number of sites in each cluster along with within-cluster sum of square (WCSS) are provided in Table 5-1. In the left panel is the 2D visualization of the k-mean clusters. Regarding the two dimensions, the visualization algorithm performs a principal component analysis (PCA) in which the higher dimensional $\Delta S2S_s$ vectors are reduced to two principal dimensions (Kassambara and Mundt, 2016). The distance along each dimension can be interpreted as how similar or dissimilar are any two cluster means. For instance, cluster 6 is farthest from cluster 8 along Dim1, and is closest to cluster 7. To interpret this separation, we refer to the more familiar plot in the right panel of Figure 5-9.

Table 5-1: Number of stations within each cluster and within-cluster sum of squares (WCSS)

	Cluster 1	**Cluster 2**	**Cluster 3**	**Cluster 4**	**Cluster 5**	**Cluster 6**	**Cluster 7**	**Cluster 8**
No. of Sites	78	68	101	69	66	45	95	66
WCSS (%)	12	12	13	10	11	12	19	12

Normalized $\Delta S2S_s$ vectors of the 588 sites in our dataset are plotted in the right panel of Figure 5-9. Each thin translucent lines corresponds to a single site, while the thick overlaid lines represent the cluster-specific mean normalized $\Delta S2S_s$ vectors, for the period range $T = 0.01$s-2s. These are used to develop to our empirical site amplification functions associated with the site clusters/classes derived in this study. Observing the two plots in Figure 5-9: the mean normalized $\Delta S2S_s$ for cluster 8 is well below zero for the entire period range. While cluster 7, which is diagonally the farthest from cluster 8 in the left panel of Figure 5-9, shows the opposite behavior. The same logic can be applied to cluster 1 and 5, and so on. Since the $\delta S2S_s$ distributions at each period are normally distributed with zero mean, we expect to see such symmetric classification from our procedure. Eventually, these clusters will be validated with geotechnical site-response parameters, based on which we assert that the spectral clustering procedure yields physically meaningful site classification. The following sections presents the practicality of our clusters as site classes.

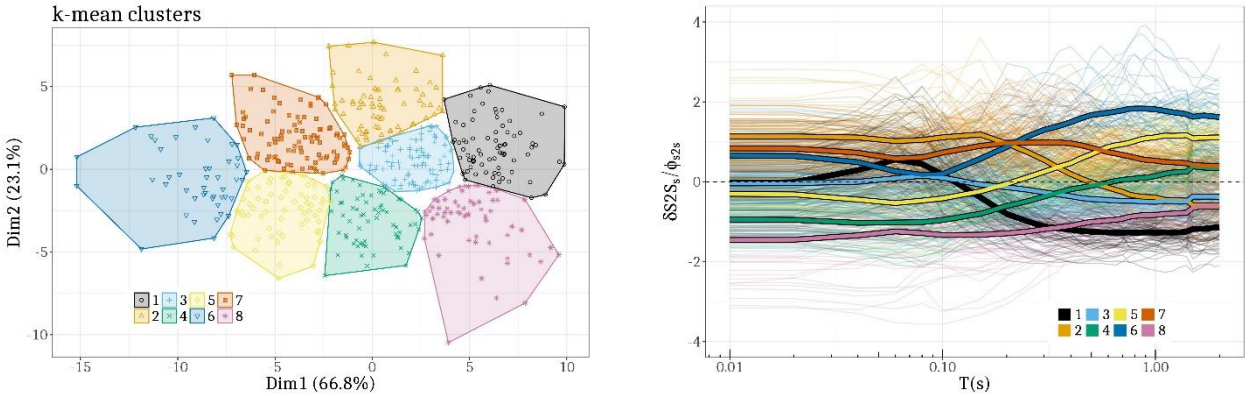


Figure 5-9: 2D visualization of k-mean clusters and the k-means amplification functions: (Left panel) Visualization of *k*-mean clustering, where each polygon is a cluster and each point within is a site ($\Delta S2S_s$). Dim1 and Dim2 are variables derived from a Principal Component Analysis of $\Delta S2S_s$ vectors, which together describe 89.9% of data variability. (Right panel) Normalized $\Delta S2S_s$ of 588 sites in thin lines, and cluster-specific normalized mean $\Delta S2S_s$ overlaid as thick lines – color coded according to cluster number.

5.7.1 Site amplification functions: Mean and variability

It is customary to present site amplification functions for different site classes with respect to the reference site conditions. For example, in EC8 the reference site conditions are characterized as outcropping rock sites with $V_{S30} = 800$m/s. Probabilistic seismic hazard assessment (PSHA) in a region and associated hazard estimates such as hazard curves and hazard maps are first produced for such reference site conditions, and then adjusted to site-specific estimates using the amplification functions. Our first point of interest is then to identify reference site conditions derived from the site clusters.

In this study, we select the reference site cluster as the one with a relatively low and flat mean $\Delta S2S_s$ vector. In the right panel of Figure 9, cluster 8 shows $\Delta S2S_s$ ideal to be qualified as reference site conditions, since it shows no selective amplification of any period ranges with respect to other sites in the dataset. Note that, until this point we set no a priori criterion on reference site geotechnical conditions (in terms of V_{S30} or other parameters). Our selection of reference site cluster is solely based on its relatively flat empirical mean amplification function. In the left panel of Figure 5-10, we show the empirical site amplification functions of the other seven non-reference site conditions with respect to cluster 8. The amplification functions in this plot are estimated from following steps:

1. The normalized $\Delta S2S_s$ of Figure 9 are scaled back to their original random-effect estimates by multiplying them with period-dependent between-site standard deviations φS2S of Figure 5-7.
2. The de-normalized $\Delta S2S_s$ vectors are scaled with respect to the reference cluster 8, so that the reference cluster 8 $\Delta S2S_s$ vector is now a null vector.
3. Since our $\Delta S2S_s$ are additive random-effects of a mixed-effects GMPE in natural-log scale, the multiplicative amplification functions would be $AF = e^{\Delta S2S_s}$. In this step, the amplification function of reference cluster 8 becomes a unit vector. For example, if the GMPE predicted PSA(1s) for the reference cluster 8 is 0.1g, and the (multiplicative) amplification factors for cluster 1 through 7 are [0.75, 1.25, 1.25, 1.75, 3.00, 4.50, 2.00], the scaled ground-motions would be [0.08g, 0.13g, 0.13g, 0.18g, 0.3g, 0.45g, 0.2g] respectively.
4. The right panel of Figure 5-10 compares the within-cluster site-to-site variability ($\phi_{S2S,c}$) against the pre-clustered between-site (ϕ_{S2S}) variability of the dataset (for our GMPE). The average reduction in site-to-site variability is approximately 50% with respect to the dataset value, while the reduction for a few clusters in at longer period ranges is larger (up to 70%). Such reduction in variability has a dramatic effect on total standard deviation (σ in right panel of Figure 5-7).

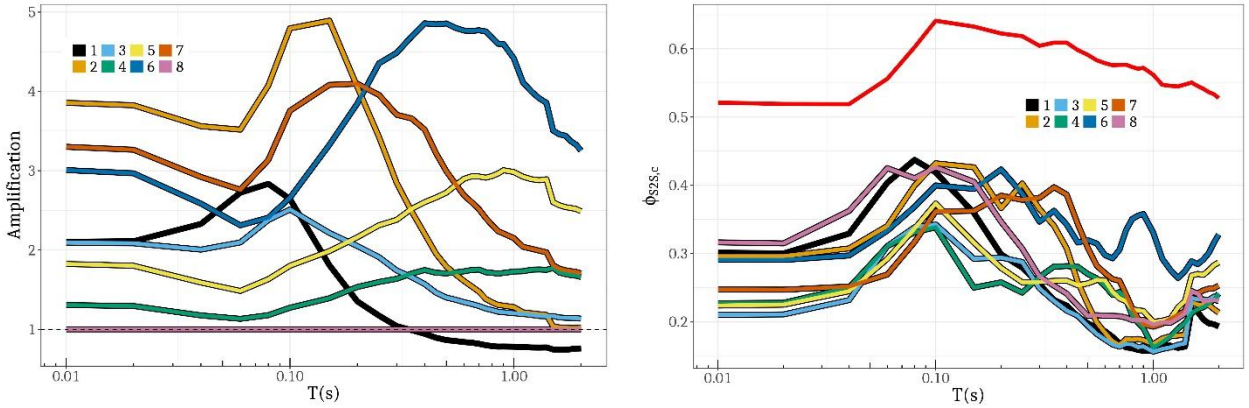


Figure 5-10: Site amplification functions and within-cluster site-to-site variability. (Left panel) Site amplification functions of each cluster scaled with respect to the reference site conditions: cluster 8. (Right panel) The within-cluster site-to-site variability ϕ_{S2S} of the 8 clusters compared to the overall GMPE ϕ_{S2S} prior to clustering (red curve). Note that the ϕ_{S2S} (red curve) is the same as the ϕ_{S2S} (yellow curve) in Figure 5-7.

Looking at the amplification functions in Figure 5-10, it is rather clear that the spectral clustering technique distinguishes sites based on their peak amplification period (T_G, analogue to H/V spectral ratio based predominant period of Zhao et al., 2006) and amplification level at T_G. However, the within-cluster between-site variability $\phi_{S2S,c}$, also reaches its maximum value around its T_G, indicating a large variability in its amplification. For example, cluster 1 shows a peak amplification at its $T_G = 0.08$s ($AF = 2.8$ in left panel of Figure 5-10), which also coincides with the period where its $\phi_{S2S,1}$ is highest (0.45 in right panel of Figure 5-10). Such correspondence between peak amplification and peak variability through T_G was reported in Zhao et al. (2006) for K-Net sites, and for EC8 classification in Cauzzi and Faccioli (2017). The cause for such a parallel is the primary limitation of discrete site classification, where sites with similar T_G but very different AF at T_G are grouped together, resulting in a large site-to-site variability of amplification. In addition, a generic high variability is observed at $T = 0.1$s, which can be partially attributed to the highly non-linear transformation from Fourier spectra (frequency domain) to response spectra via convolution with a single-degree-freedom oscillator transfer function (discussed in Stafford et al., 2017, and Kotha et al., 2017b). Decreasing ϕ_{S2S} trend is observed on either sides of $T = 0.1$s, but this can be resolved only in Fourier domain, and not with the response spectra used in this study.

5.7.2 Site-response proxies

Using the empirical site amplification functions and cluster-specific ϕ_{S2S}, the GMPE can be adjusted to predict site class dependent ground-motions in hazard assessment; the missing link is sufficient and efficient site response proxies to classify new sites in a PSHA. *Dawood et al.* (2016) provided the time-averaged shear wave velocity at depth z (m), $V_{s,z}$ for $z = 0$, 5, 10, 20, 30, 50, 100, borehole depth, and H_{800}– depth to seismic bedrock with $V_S = 800$m/s. In this study, we attempted characterizing the cluster amplification functions of Figure 5-10 using only the geotechnical site-response parameters available in the dataset. In process of developing a new site classification scheme, we first evaluated our eight site clusters against the site classification scheme defined in *Zhao et al.* (2006), *Association* (1980), and *Association* (1990). For a similar evaluation against the Eurocode 8 site classification, we refer the readers to *Kotha et al., (2018).*

5.7.2.1 Predominant period of the soil column: T_G

Current EC8 categorizes five site classes using only V_{S30} ranges, while *Zhao et al.* (2006) classified the K-Net stations into four sites classes (SC1 through SC4) based on their H/V spectral ratios. Each of these 4 classes is attributed a V_{S30} range, a characteristic range of predominant period T_G, and corresponding NEHRP class (*Council*, 2000). For reference, we provide the site classification scheme of *Zhao et al.* (2006) in Table 5-2. In this study, we assume that the period at which the cluster amplification functions attain their peak values in Figure 5-10, are analogues to H/V spectral ratio based T_G of *Zhao et al.* (2006). Under

this assumption, we categorize our eight site clusters based on their inferred T_G (Figure 5-10) in the right most column of Table 5-2.

Table 5-2: Site classification based on Zhao et al. (2006).

Zhao et al. (2016)	T_G (s)	V_{S30} (m/s)	NEHRP	Clusters (this study)
SC1	T_G < 0.2s	V_{S30} > 600m/s	A + B	C1 + C2 + C3
SC2	0.2s ≤ T_G < 0.4s	300m/s ≤ V_{S30} < 600m/s	C	C7
SC3	0.4s ≤ T_G < 0.6s	200m/s ≤ V_{S30} < 300m/s	D	C6
SC4	0.6s ≤ T_G	V_{S30} < 200m/s	E + F	C4 + C5

From this comparison of site classification criteria:

1. Cluster 8 does not fit in any of the *Zhao et al.* (2006) site classes because it showed no clear T_G, which is also the reason it was chosen as a reference site cluster.
2. Cluster 7 and 6 show clear peak amplification, and are in accordance with *Zhao et al.* (2006) proposed classes SC2 and SC3 respectively.
3. Cluster 4 and 5 show a broad amplification plateau beyond $T > 0.6$s, hence grouped into SC4. However, the amplification levels for these two clusters are significantly different, where also the within-cluster variability $\phi_{S2S,c}$ is relatively very small (~0.2 for $T > 0.6$s in Figure 5-10). Based on their low within-cluster variabilities, we chose not merge cluster 4 and 5 into a single cluster.
4. Clusters 1, 2 and 3, all show a clearly defined peak amplification at $T_G < 0.2$s. Figure 5-10 suggests the mean amplification levels for these clusters to be very different in the vicinity of T_G. Reducing to sub-optimal number of clusters (in k-mean clustering) does not merge these clusters, suggesting the need to divide SC1 into three sub-class: C1, C2, and C3 – against merging them into SC1.

5.7.2.2 Time averaged shear-wave velocity of soil column: V_{s10}, V_{s30} and H_{800}

Table 5-2 shows the compatibility of our clusters with *Zhao et al.* (2006) classification based on their T_G, while Figure 5-11 categorizes them further based on their distribution of H_{800}, V_{S10} and V_{S30}. Note that the clusters in Figure 5-11 are rearranged based on their median H_{800} (depth to engineering bedrock with $V_S = 800$m/s). For instance, clusters 4, 5, 6 and 7 in the top row have the deepest soil column with median $H_{800} > 50$m. Clusters 1, 2, 3 and 8 on the other hand are characterized by shallow soil profiles with median H_{800} around 10m. For visual clarity, both panels provide guiding lines at $H_{800} = 10$, 30, and 100m. In the left panel, the x-axis marks the bounding V_{S30} values of *Zhao et al.* (2006) site classification at 200, 300 and 600m/s, while in the right panel, the x-axis is divided at $V_{S10} = 200$, 300 and 400m/s. We used Table 5-2 and Figure 5-11 to evaluate the physical meaning of our clusters, and also to define new site classes in Table 5-3.

The colored contours of Figure 5-11 represent the 2D kernel density (bivariate normal distribution) of V_{S30}–H_{800} and V_{S10}–H_{800} of each cluster. The brightly colored regions bound the high density of points (sites), which diffuses into outer counters covering low density regions. We use the high density contours (green patches in Figure 5-11) of the kernel density plot to identify representative ranges of V_{S30}, V_{S10}, and H_{800} combination that characterize our new site classification scheme in Table 5-3. In addition, we provided a revision of T_G ranges inferred from the peaks in amplification functions of Figure 5-10.

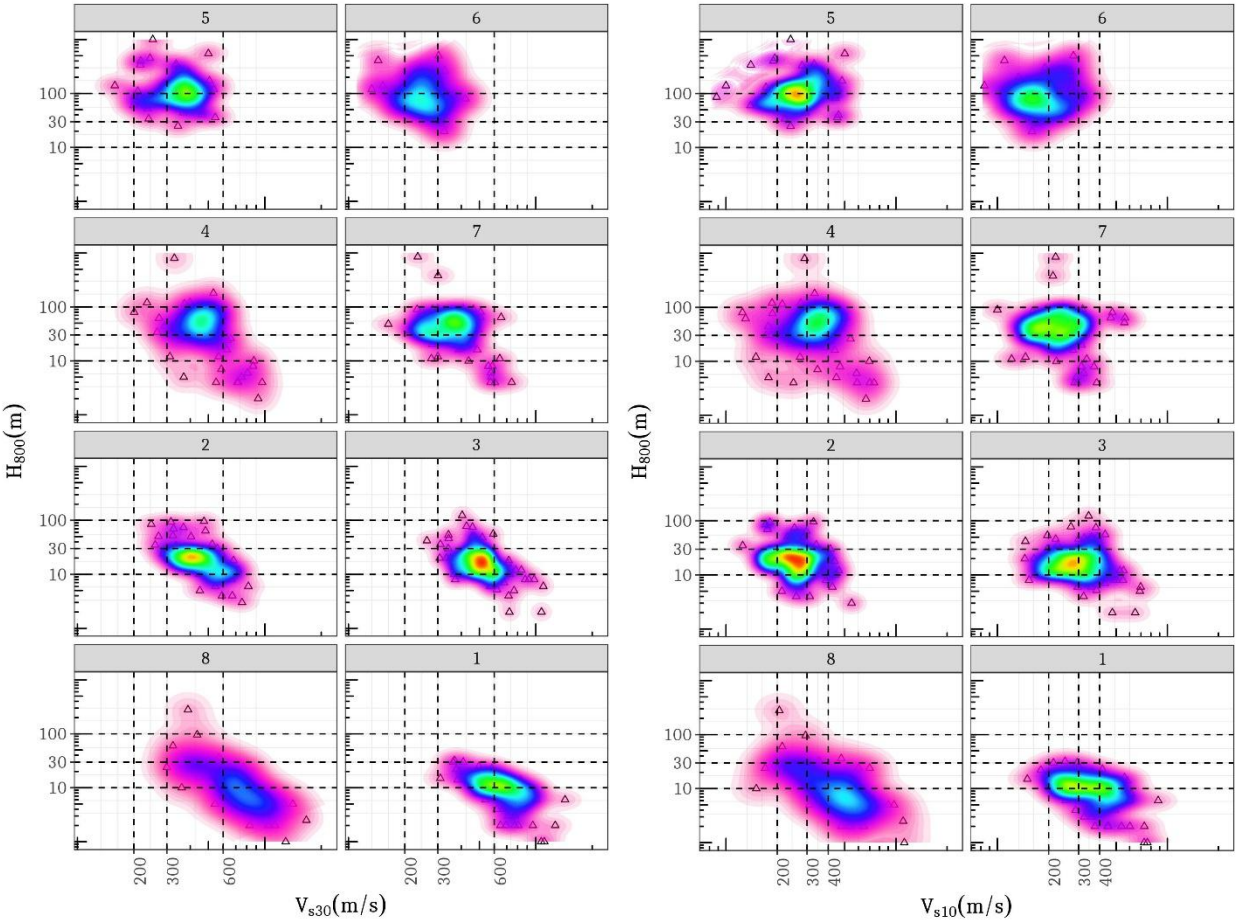


Figure 5-11: Evaluation of site response proxies in characterizing site clusters: In the left panel is the combination of H_{800} and V_{S30}, and in the right panel, H_{800} and V_{S10}. The contour colors represent the 2D Kernel (bivariate normal) distribution of geotechnical parameters of sites within each cluster. Warmer colors (Red - Green) imply high density, colder colors (Blue – Purple) mark low density boundaries.

The purpose of this exercise was to identify a combination of geotechnical parameters to classify the sites which were otherwise grouped into the same site classes of *Zhao et al*. (2006). From Figure 5-11 and Table 5-3, we make the following observations:

- Cluster 4 and 5 are constituted of KiK-net sites with 300m/s < V_{S30} < 600m/s and 150m/s < V_{S10} < 400m/s, showing a broad amplification plateau, increasing towards longer periods. These two clusters can be set apart from other clusters based on their large T_G. Between cluster 4 and 5, the distinction can be made based on their V_{S10}, V_{S30} and H_{800}. Note that these clusters resembled SC4 of *Zhao et al.* (2006) based on their T_G (> 0.6s), and were expected to also show a V_{S30} < 200m/s. However, our data consists of very few KiK-net sites with V_{S30} < 200m/s.
- Cluster 6 and 7 can be distinguished from other clusters based on their well-defined T_G ranges. Cluster 6 shows a much higher amplification with respect to cluster 5, despite similar H_{800}, due to its systematically lower V_{S30} and V_{S10} ranges. Similarly, cluster 7 despite having the same H_{800} range as cluster 4, shows much higher amplification at a lower T_G due to its softer soil profile – characterized by lower V_{S30} and V_{S10} ranges. Interestingly, cluster 6 and 7 are hard to be distinguished based on their V_{S30} ranges alone. In which case, T_G, V_{S10} and H_{800} as site-response proxies perform significantly better, proving a case against V_{S30} as a standalone proxy.
- Cluster 8 serves as our reference site cluster, with the highest values of V_{S30} and V_{S10}. Sites in this cluster showed no clear peak amplification (left panel of Figure 5-10), hence could not be compared with *Zhao et al.* (2006) classes.

- Cluster 1 from 2 and 3, which were nested in SC1 of *Zhao et al.* (2006), can be resolved based on their T_G values and H_{800} ranges. Cluster 1 with $T_G < 0.1$s (left panel of Figure 5-10), is the only one showing a strong amplification at high frequency and de-amplification towards longer periods, with respect to reference site cluster 8. Cluster 1 and 8 can be distinguished based on their H_{800} and V_{S10} ranges, but not V_{S30} – suggesting V_{S10} as a better proxy of high frequency site-response.
- Cluster 2 and 3 are separated from the closest resembling cluster 1, based on their longer T_G and larger H800 values. However, these two clusters do not appear to differ in their V_{S10}–V_{S30}–H_{800} ranges, as much as they do from other clusters. Given their identical T_G ranges, but radically different amplification levels, we suspect these clusters to differ in the velocity contrast of their soil profiles. A higher impedance contrast results in significantly higher amplification at T_G (see Figure 5-10), which appears to be the case considering the relatively lower V_{S30} and V_{S10} ranges of cluster 2 against cluster 3. Shear-wave velocity profiles, and additional geotechnical parameters might help in better characterizing the differences among cluster 2 and 3, and cluster 1 as well.

Table 5-3: Site cluster characterization based on V_{S10}-V_{S30}-H_{800} ranges.

Site cluster	TG (s)	V_{S30} (m/s)	V_{S10} (m/s)	H_{800} (m)
C5	> 1s	300 – 450m/s	200 – 300m/s	> 50m
C4		400 – 600m/s	300 – 400m/s	30 – 100m
C6	0.4 – 1s	200 – 300m/s	< 200m/s	> 50m
C7	0.2 – 0.4s	200 – 450m/s	200 – 400/s	30 – 100m
C3	0.1 – 0.2s	450 – 600m/s	200 – 400m/s	10 – 30m
C2		300 – 600m/s	150 – 350m/s	
C1	< 0.1s	450 – 600 m/s	200 – 600m/s	5 – 20m
C8	-	> 600m/s	> 600m/s	< 5m

5.8 Discussion and conclusions

In this study we introduce an approach to site classification derived from cluster analysis of empirical site amplification functions. The resulting site classification is aimed to be simple, robust, and data-driven with minimal a priori constrains in terms of relevant site-response parameters. The fundamental requirement for such classification was to derive statistically well-constrained empirical site adjustment functions ($\Delta S2S_s$ vectors). As a first step, we selected a rich dataset featuring several well-characterized sites recording many earthquakes in a region; the KiK-net dataset by *Dawood et al.* (2016). The next step was to fit a mixed-effects GMPE, whose site-specific random-effects ($\delta S2S_s$) for periods $T = 0.01$s–2s constitute the $\Delta S2S_s$ vectors. Given the critical importance of GMPE in our approach, it was necessary that its magnitude and distance scaling fixed-effects components are calibrated very well for a wide magnitude range $3.4 \leq M_w \leq 7.3$, and large distance range $0 \leq R_{JB} < 600$km. It IS necessary for a variety of reasons:

1. We need as many records as possible for each site, so that the estimated $\Delta S2S_s$ vectors have a low estimation errors.
2. It is important to include a large number of small events and long distance recordings, as opposed to the more important (in a hazard perspective) large magnitudes and short distances, to ensure that the $\Delta S2S_s$ vectors capture predominantly linear site response – and not biased by nonlinear response on very soft soils.
3. Given that large magnitude events at near source distances may trigger nonlinear site response, and that these scenarios are also subject to other phenomena such as: saturation of ground-motions beyond $M > 6.5$ and at near-source distances $R_{JB} < 20$km, we were required to model carefully the

near-source magnitude and distance scaling. In doing so, we used a period-dependent effective-depth, or the h-parameter, proposed by *Yenier and Atkinson* (2015). For far-source distances, we could capture the apparent anelastic attenuation phenomenon and, what appears to be a combination of post-critical reflections at crustal discontinuities (e.g. Moho reflections) and transition from body waves at near-source to surface waves at far-source distances from the event. We observed that our distance scaling works reasonably well for the large distance range we chose.

4. Finally, we observed that the $\delta S2S_s$ show a weak correlation relation with V_{S30}. For this dataset, the site-response: 1) is not efficiently captured by V_{S30} particularly at short – moderate periods, and 2) highly variable even for sites with identical V_{S30}. We therefore attempted our alternative approach to classifying sites based on V_{S30}.

Our need to derive a GMPE is to demonstrate that the magnitude, and distance scaling of ground-motions (here PSA) are strongly period dependent. Following our exercise, we suggest caution against deriving empirical site amplification functions from response spectra normalized by Peak Ground Acceleration (PGA). The $\Delta S2S_s$ vectors on the other hand, are site-specific terms filtered for event and path effects with a robust GMPE median. In this sense, we consider $\Delta S2S_s$ vectors as more suitable in developing site classification schemes and amplification functions. We chose the k-mean clustering technique to reduce the higher dimensional $\Delta S2S_s$ vectors of 588 sites into 8 clusters (of sites) with similar linear soil response under seismic action (in terms of their $\Delta S2S_s$ vectors). These 8 clusters serve as the site classes in our new classification scheme.

Site amplification functions are usually presented as scaling functions with respect to the reference site conditions. Traditionally, outcropping rock sites with $V_{S30} > 800$m/s are considered as a reference sites. Hazard estimates are made on reference site conditions and then scaled using the soil-site amplification functions. The $\Delta S2S_s$ vectors do not presume any reference site conditions, but instead are additive random-effects (scalar adjustments) to the GMPE fixed-effects median. Technically, the $\Delta S2S_s$ vectors are site-specific deviations from a *hypothetical* reference site, whose response is an average of all sites in the dataset, i.e. a site with null $\Delta S2S_s$ vector. However, for engineering purposes, it is necessary to characterize real reference site conditions. We therefore select the cluster of sites whose mean $\Delta S2S_s$ vector is *low and flat*. Meaning, sites in this cluster show a systematic de-amplification over the entire period range, with respect to other sites in the dataset. This unique cluster (cluster 8) represents the reference site conditions in our approach. Essentially, we identified a reference site cluster with no amplification, and seven other clusters with unique non-zero site amplification functions. Additional benefit of clustering technique is seen as the significantly smaller within-cluster site-to-site variability, which is on an average $\sim$50% smaller than the pre-clustered, overall site-to-site response variability of the dataset. This in our opinion is a significant improvement in the context of seismic hazard assessment.

For site amplification functions to be applicable at new sites, we need to develop site-response proxies based on which the new sites can be classified. From this point of view, we are limited by the available geotechnical information at the sites. Among the most prevalently used site parameters in the site-response characterization are the predominant period (T_G), time-averaged shear-wave velocity up to 10m (V_{S10}) and 30m (V_{S30}), and the depth to engineering bedrock with shear-wave velocity $V_S = 800$m/s (H_{800}). The inferences from this part of the study are enumerated below:

1. Multiple clusters show significantly different site amplifications but similar V_{S30} ranges, suggesting that V_{S30} is insufficient as a standalone proxy in site-response classification.
2. Classification based on T_G works well in classifying sites at first order. However, it is insufficient in distinguishing sites with identical ranges of T_G, but different amplification levels at T_G.
3. For site clusters with $H_{800} > 30$m and $T_G > 0.2$s, both V_{S30} and V_{S10} perform well in distinguishing the four clusters with moderate – long period amplifications. Clusters (5 and 6) with deepest soil profiles ($H_{800} > 50$m) can be distinguished with their T_G, V_{S10} and V_{S30}, where lower soil stiffness (of

cluster 5) translated into lower T_G and a much higher amplification. Similar is the case for clusters (4 and 7) with shallower soil profiles (30m $< H_{800} <$ 100m).

4. For sites with 10m $< H_{800} <$ 30m and 0.1s $< T_G <$ 0.2s, we identified two clusters with very similar V_{S30} and V_{S10} distribution, but significantly different amplification levels. These clusters cannot be distinguished with the available geotechnical information. A detailed investigation of their shear-wave velocity profiles may help better distinction of these clusters.
5. We identified two clusters with $V_{S30} >$ 600m/s that can be separated based on their T_G, V_{S10} and H_{800}. Cluster (1) with lower V_{S10} and a higher H_{800} shows a strong amplification at its $T_G <$ 0.1s, while the one with higher V_{S10} and lower H_{800} shows a flat response (cluster 8). Evidently, V_{S30} based classification groups these two very different site types into a unique site class. In our opinion, such misclassification leads to a significant bias and a large variability in response of the so-called reference site class (e.g. $V_{S30} >$ 800m/s in EC8 classification). We suggest using at least the V_{S10}, or even better - V_S profiles, to characterize reference site conditions.

Our approach is beneficial in identifying hidden site classes, resolving site-to-site variability, and developing efficient site classes from a rich dataset. It can be extended to the point where the clusters can be hierarchically divided or merged depending on the available site parametrization in a region. However, we note that site types, sparsely represented or not represented at all in the dataset, cannot be identified with data-driven techniques as ours. A more flexible, predictive method is in development for application to pan-European dataset (*Lanzano et al.*, 2017). The number of clusters, the mean and variability of empirical site amplification functions, and even the relevant site-response proxies may depend on the spatial coverage of regional datasets.

Acknowledgements

We appreciate the efforts of Dr. Julian Bommer, and the anonymous reviewer, for their impressively detailed review. We would like to thank Prof. John Anderson for his valuable insights in interpreting the results. This research is funded by the SIGMA2 project (EDF, CEA, PG&E, SwissNuclear, Areva, CEZ, CRIEPI) – 2017 - 2021 (http://www.sigma-2.net/)

6
Site-corrected Magnitude and Region Dependent Correlations of Horizontal Peak Spectral Amplitudes

S. R. Kotha[1, 2], D. Bindi[1], F. Cotton[1, 2]

[1]German Research Centre for Geosciences GFZ, Potsdam, Germany

[2] University of Potsdam, Potsdam, Germany

Abstract

Empirical correlations of horizontal peak spectral amplitudes (PSA) are modelled using the total-residuals obtained in a Ground-Motion Prediction Equation (GMPE) regression. Recent GMPEs moved towards partially non-ergodic region- and site-specific predictions, while the residual correlation models remained largely ergodic. Using mixed-effects regression, we decompose the total-residuals of a pan-European GMPE into between-event, between-site, and event-and-site corrected residuals to investigate the ergodicity in empirical PSA correlations. We first observed that the between-event correlations are magnitude dependent, partially due to the differences in source spectra, and influence of stress-drop parameter on small and large events. Next, removing the between-site residuals from within-event residuals yields the event-and-site corrected residuals which are found to be region dependent, possibly due to the regional differences in distance-decay of short period SAs. Using our site-corrected magnitude and region dependent correlations, and the between-site residuals as empirical site-specific ground-motion adjustments, we compute partially non-ergodic conditional mean spectra at four well-recorded sites in Europe and Middle-Eastern regions.

Keywords: Site-specific Ground-Motion Prediction Equations, Correlations, Conditional Spectra, Spectral Accelerations

6.1 Introduction

Ground-Motion prediction equations (GMPE) describe the probabilistic distribution of the peak spectral amplitudes (PSA) of single degree of freedom oscillators with different periods (T), as a function of event magnitude, distance to the source, and proxies for site effects (e.g., V_{S30} defined as the shear-wave velocity averaged over the uppermost 30m of soil). At each spectral period, the standard deviation (σ) of the total-residual (ε) i.e., distribution of observed minus predicted spectral values, is used to quantify the aleatory variability of the ground-motion.

The total-residuals obtained from a Ground-Motion Prediction Equation (GMPE) regression can be used to derive empirical correlations of horizontal peak spectral amplitudes. These empirical correlation models are widely used in engineering applications, e.g. simulation of scenario response spectra (*Baker and Cornell*, 2006), vector-valued probabilistic seismic hazard assessment (VPSHA; *Bazzurro and Cornell*, 2002), and conditional spectra (CS) for an effective selection of time histories matching a design scenario (*Baker*, 2011). For this purpose, several correlation models were developed using strong motion datasets from Europe-Middle East (e.g. *Cimellaro*, 2013; *Akkar et al.*, 2014a), regional datasets (e.g. *Jayaram et al.*, 2011; *Itoi and Takada*, 2012), and global datasets (e.g. *Baker and Cornell*, 2006; *Baker and Jayaram*, 2008; *Azarbakht et al.*, 2014; *Abrahamson et al.*, 2014; *Baker and Bradley*, 2017). Based on the findings of *Baker and Cornell* (2005), the subsequent total-residual correlation models reported no significant magnitude and distance dependence. The underlying datasets then, such as NGA-West (*Chiou et al.*, 2008), did not allow investigating such dependencies because of too few records to quantify any statistically significant dependencies. *Azarbakht et al.* (2014) reported a magnitude and distance dependence of correlations using NGA-West dataset, which was later refuted by *Baker and Bradley* (2017) on the basis of statistical inconsistencies and a lack of physical reasoning. Using the NGA-West2 dataset (*Ancheta et al.*, 2014), *Campbell and Bozorgnia* (2014) observed a magnitude dependence of correlations, but only provided magnitude independent correlation models for events larger than $M5$.

In the more recent strong motion datasets, such as RESORCE for pan-European region (*Akkar et al.*, 2014b) and the NGA-West2 global dataset the number of records from several active shallow crustal regions increased considerably. Consequently, recent GMPEs tended towards resolving the regional ergodicity in ground-motion prediction (e.g. *Boore et al.*, 2014; *Campbell and Bozorgnia*, 2014; *Kale et al.*, 2015; *Kotha et al.*, 2016; etc.). However, the empirical correlation models remained largely ergodic, in the sense that, the magnitude, distance, region, and site dependence of PSA residual correlations was not investigated. The improvement in the data availability motivated us to investigate and quantify such dependencies of correlation models, thus moving towards partially non-ergodic correlation models alongside the partially non-ergodic GMPEs.

This study describes the development of new site-corrected magnitude and region dependent correlation models corrected for site-to-site variability. We first explain the need to decompose the total-residuals into between-event, between-site, and event-and-site corrected residuals, in order to investigate the event, site and regional ergodicities respectively. Based on our investigation with the RESORCE dataset, supported by statistical checks and physical reasoning, we propose new partially non-ergodic horizontal PSA correlation models. Additionally, we demonstrate an application of our correlation models in computation of partially non-ergodic conditional spectra at four well-recorded sites in the pan-European region.

6.2 A new method to compute residual correlations taking into account repeatable site effects

Using the *Abrahamson and Youngs* (1992) mixed-effects algorithm, the record-specific total-residuals in Eq. 1 (ε) can be decomposed as in Eq. 2 into between-event (δB_e) and within-event (δW_{es}) components (*Al-Atik et al.*, 2010):

$$\log(Y_{observed}) = \log(Y_{median}) + \varepsilon \qquad (1)$$

$$\varepsilon = \delta B_e + \delta W_{es} \quad (2)$$

Since the between-event and within-event residuals follow mutually independent normal distributions i.e., $\delta B_e = \mathcal{N}(0, \tau^2)$ and $\delta W_{es} = \mathcal{N}(0, \phi^2)$, the respective correlation models computed independently can be combined to estimate the total-residual correlation model using Eq. (3), where $\rho(T_1, T_2)$ is the correlation coefficient for residuals at two response spectral periods (T_1, T_2); ϕ, τ and σ are the period dependent within-event, between-event, and total-residual standard deviations respectively:

$$\rho(T_1, T_2)_{total} = \frac{\rho_{within-event}\ \phi(T_1)\,\phi(T_2) + \rho_{between-event}\ \tau(T_1)\,\tau(T_2)}{\sigma(T_1)\sigma(T_2)} \quad (3)$$

The within-event residuals account for path and site variability, and its variance (ϕ^2) is generally larger than the between-event variance (τ^2). Consequently, when the total-residual correlations are computed using Eq. (3), the contribution of the between-event correlations is often negligible, resulting in the usually observed magnitude-independence of the total-residual correlations (e.g. *Baker and Cornell*, 2005). To overcome this limitation, we treat the repeatable site effects as empirical site-specific adjustments to the median of GMPE, rather than as a component of the aleatory within-event variability (e.g. *Rodriguez-Marek et al.*, 2013). The computation of repeatable site effects is nowadays feasible due to the increase of data at individual sites. We decompose the within-event residuals into between-site residuals ($\delta S2S_s$) and event-and-site corrected residuals (δWS_{es} in Eq. 4):

$$\delta W_{es} = \delta S2S_s + \delta WS_{es} \quad (4)$$

$\delta S2S_s$ can be estimated in a post-processing step of GMPE regression as a mean of within-event residuals at a site (*Rodriguez-Marek et al.*, 2013), as random-intercept using Eq. (12) of *Kotha et al.* (2017) or during the GMPE regression itself using a mixed-effects algorithm (e.g. lmer, *Bates et al.*, 2014). In this study, we used the later approach to minimize any bias in the GMPE median and $\delta S2S_s$ estimates (further detail in *Stafford*, 2014, and *Kotha et al.*, 2017). The standard deviation of between-site residuals is the site-to-site variability denoted by ϕ_{S2S}, and that of normally distributed event-and-site corrected residuals is ϕ_0, i.e. $\delta S2S = \mathcal{N}(0, {\phi_{S2S}}^2)$ and $\delta WS_{es} = \mathcal{N}(0, {\phi_0}^2)$. The between-site variance (${\phi_{S2S}}^2$), when deducted from the within-event variance (ϕ^2), results in the event-and-site corrected variance (${\phi_0}^2$). The event-and-site corrected variance together with the between-event variance (τ^2) yields the so-called single-site total variance (${\sigma_0}^2$). Replacing the within-event with event-and-site corrected correlations and standard deviations, and total standard deviation (σ in Eq. 3) with the site-corrected standard deviation (σ_0), gives us the site-corrected residual correlations (Eq. 5).

$$\rho_{site-corrected}(T_1, T_2) = \frac{\rho_{event-site}\phi_0(T_1)\phi_0(T_2) + \rho_{between-event}\tau(T_1)\tau(T_2)}{\sigma_0(T_1)\sigma_0(T_2)} \quad (5)$$

The benefits of site-corrected residual correlation formulation are manifold:

1) The contribution of the between-event correlations to site-corrected residual correlation in Eq. (5) becomes noticeable due to the comparable sizes of ϕ_0 and τ. Meaning, if the magnitude dependence of between-event correlations is indeed significant, then it is less likely to be concealed in Eq. (5) than in Eq. (3).
2) The within-event residuals and correlations maybe biased by a well-recorded site, which could bias the within-event correlations in favor of representing the correlations at the well-recorded site. By discounting the between-site residuals, the resulting event-and-site corrected residuals (δWS_{es}) can be investigated for distance and region dependencies of their correlations.
3) For any site with strong motion recordings, the corresponding period dependent between-site residuals ($\delta S2S_s$) can be estimated with respect to a chosen GMPE (e.g. *Rodriguez-Marek et al.*, 2013). These residuals can be used as empirical site-specific adjustment factors to scale the GMPE predicted spectral values (*Kotha et al.*, 2017). Since $\delta S2S_s$ is estimated using a small sample of within-event re-

siduals, an appropriate epistemic uncertainty in terms of estimation error of $\delta S2S_s$ needs to be considered. For sites with no strong motion recordings, the $\delta S2S_s$ can be assumed to be zero with a standard error equal to ϕ_{S2S} (*Abrahamson and Hollenback*, 2012).

For the between-event and event-and-site corrected residuals, we estimate the Pearson product-moment correlation coefficients using Eq. (6), where X and Y are random variables (residuals in this context) at period T_1 and T_2 respectively:

$$\rho_{X,Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \cdot \sigma_Y} \quad (6)$$

Since the residuals are sampled from a normal distribution, the standard-error on the correlation estimates can be approximated using Eq. (7) (*Hotelling*, 1953), n being the sample size. Note that according to Eq. (7), the standard error in the point estimate of correlation is larger for smaller $\rho_{X,Y}$ values (*Baker and Bradley*, 2016):

$$SE(\rho_{X,Y}) = \frac{(1-{\rho_{X,Y}}^2)}{\sqrt{n-1}} \quad (7)$$

6.3 Datasets

6.3.1 NGA-West2 dataset

The NGA-West2 project dataset includes ground-motion data and metadata from earthquakes recorded worldwide, in particular from California, Japan and Taiwan. The dataset includes more than 20000 recordings from about 600 shallow crustal earthquakes in the magnitude range 3 to 7.9. NGA-West 2 extends the earlier NGA-West dataset (*Chiou et al.*, 2008) by adding several moderate-to-large earthquakes between 2003 and 2011. A large number of small-to-moderate earthquakes occurring in California were included as well. NGA-West2 dataset based GMPEs superseded those previously derived from NGA-West. In the following sections, we consider an NGA-West and an NGA-West2 correlation model derived for RotD50 component of horizontal peak spectral accelerations, *Baker and Jayaram* (2008) and *Abrahamson et al.* (2014) respectively, to investigate differences between correlation models.

6.3.2 RESORCE dataset

The RESORCE project dataset (*Akkar et al.*, 2014b) includes about 6000 recordings from 1814 earthquakes in the magnitude range from 3 to 7.6 that occurred in Europe and Middle-Eastern regions. RESORCE expands the earlier European strong motion data set ISESD (*Ambraseys et al.*, 2004) including the outcomes of several national strong motion projects in Europe, such as ITACA in Italy (*Luzi et al.*, 2008), HEAD in Greece (*Theodulidis et al.*, 2004), and TSNMP in Turkey (*Akkar et al.*, 2010). The dataset was used to derive a set of GMPEs (*Douglas et al.*, 2014) following both parametric and data-driven approaches. In addition a correlation model based on RESORCE was provided by *Akkar et al.* (2014a), which superseded models derived from ISESD, e.g. *Cimellaro* (2013).

The *Bindi et al.* (2014) GMPE based on RESORCE dataset is derived for geometric mean (GM) of 5% damped horizontal peak spectral acceleration (PSA) at 17 spectral periods from PGA to 4s, using strong motion data with $M_w \geq 4$, focal depth $\leq$ 35km, and $R_{JB} \leq$ 300km. In the published version of the GMPE, the total-residuals were split into two components: between-event and within-event residuals as in Eq. (2). In this study however, our aim is to investigate site-corrected correlation models, where the within-event correlations are replaced with event-and-site corrected correlations for reasons listed previously. For this purpose, using a non-linear mixed effects regression algorithm (NLMER; *Bates et al.*, 2015) we split the *Bindi et al.* (2014) within-event residuals into between-site and event-and-site corrected residuals as in Eq. (4). Figure 6-1 compares standard deviations of the resulting residual distributions: within-event (ϕ) and between-event (τ) from *Bindi et al.* (2014), between-site (ϕ_{S2S}) and event-and-site corrected (ϕ_0) from splitting the within-event residuals. Notice that ϕ_0 is comparable in size to τ, while ϕ is

significantly larger. The relatively stronger influence of between-event correlations on site-corrected residual correlations motivates us to investigate their magnitude dependence.

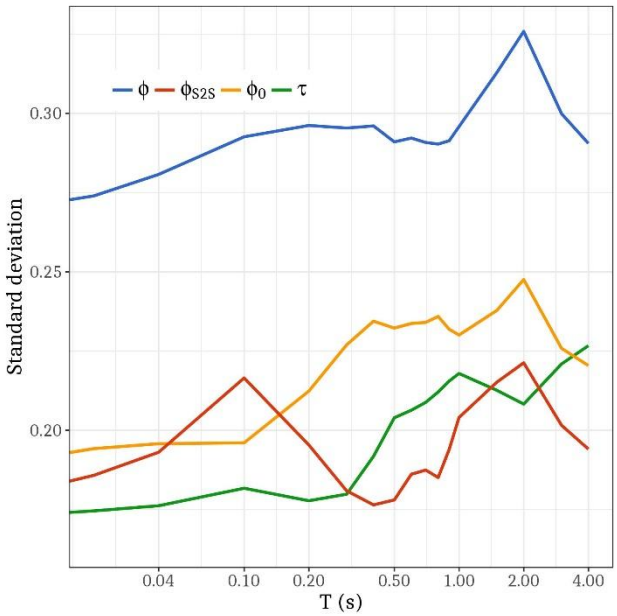


Figure 6-1: Reduction of within-event standard deviation (ϕ) to event-and-site corrected standard deviation (ϕ_{ss}) by removing between-site variability (ϕ_{S2S}), considering the RESORCE dataset. The between-event variability (τ) is shown as well.

6.4 Dataset dependence of total-residual correlations

Before proceeding to investigating non-ergodicity of correlation models, we first examine the dependence of correlation models on the dataset. We used the peer-reviewed models of *Abrahamson et al.* (2014) (ASK14) from NGA-West2, *Baker and Jayaram* (2008) from NGA-West, and *Akkar et al.* (2014a) (A14) from RESORCE datasets. We also included in the comparison a correlation model developed in this study using total-residuals from *Bindi et al.* (2014) (B14) GMPE. Figure 6-2 shows a comparison between the 4 models: those derived from RESORCE in panel (a), and those derived from NGA-West and NGA-West2 datasets in panel (b). In Figure 6-2, and in all subsequent correlation plots, the correlation values $\rho(T_1, T_2)$ are shown as a function of T_2 for a set of fixed primary period T_1. For all models, the correlation is equal to one when $T_1 = T_2$ and decreases as the separation between T_1 and T_2 increases. For visual clarity, we show the correlation curves only for three selected T_1 values. In particular, we select the periods $T_1 =$ 0.02s, 0.5s, and 2s to discuss the correlation of short, moderate and long periods, respectively.

At a first glance, Figure 6-2 suggests that the correlation models developed on the RESORCE dataset differ from the NGA-West and NGA-West2 models. Particularities of these differences and their possible physical reasons are as follows:

- In Figure 6-2a, A14 and B14 models are in good agreement since they are derived over a similar record selection from the same data set. The only observable differences are in the correlations between short period SAs and moderate – long period SAs. A major difference between the two models is that the underlying GMPE of A14 includes a non-linear site-response component, which predicts high frequency SAs (for $M > 6$) significantly different from the linear-only site-response component of the *Bindi et al.* (2014) GMPE (*Douglas et al.,* 2014).
- Figure 6-2b emphasizes the difference between ASK14 and BJ08 models derived from NGA-West2 and NGA-West datasets respectively. Referring to the curves $T_1 = 0.02$s and $T_1 = 0.5$s, the correlations between short and moderate period SAs appear to be similar among the two models. The long period SAs ($T_1 = 2$s) are less correlated to the moderate frequency SAs (e.g. $0.1\text{s} < T_2 < 1\text{s}$) in case of BJ08. A clear physical reason for such differences is difficult to assert, given that: (1) NGA-West2 is a substantial update of NGA-West from inclusion of several M3-M4 events from California, (2) *Baker and Jayaram* (2008) correlations shown here are from an analytical function, and (3) record selection

and choice of functional form by *Chiou and Youngs* (2008) and *Abrahamson et al.* (2014) are not similar. Similar comparison between the NGA-West based BJ08, and NGA-West2 based *Baker and Bradley* (2017) were recently reported, wherein the models were assessed to be identical for practical purposes.

- Comparing the models B14 and ASK14 across the datasets, pan-European RESORCE and global NGA-West2 respectively, the only significant differences observed are the correlation between short and moderate periods with long period SAs. For instance, the RESORCE models show stronger correlations between $T_1 = 0.5$s and $T_2 = 0.2$s. The evident reason being that the NGA-West2 dataset based GMPEs were derived over much larger magnitude and distance ranges, from events and sites across several regions (detail in the Datasets section). In comparison, the pan-European GMPEs considered only events larger than $M4$, from a dataset dominated by $M5$ to $M6$ events in Italy and Turkey. The diversity of NGA-West2 data exhibits higher record-to-record variability at short period SAs with respect to the RESORCE dataset, thus the relatively lower correlations with moderate – long period SAs.

It is important to note that the NGA-West2 GMPEs are derived on RotD50, while the RESORCE GMPEs are derived on Geometric Mean (GM) of horizontal SAs. *Baker and Jayaram* (2008) reported that even though the GMPE residual variances for different PSA component definitions are significantly different, the residual correlation models are not. Assuming the same and based on our observations, we then consider the correlation models to be dataset dependent or, more generally, region dependent. Given the recent shift towards partially non-ergodic regionalized GMPEs, we suggest the correlation models to be regionalized as well. Such working hypothesis and the physical reasoning, i.e. the magnitude and distance ranges, and regional diversity of the datasets, motivate us to inspect the various within-dataset dependencies of correlation models. For this purpose, we choose the RESORCE dataset based *Bindi et al.* (2014) GMPE and its correlation model (B14) for further investigations in the following sections.

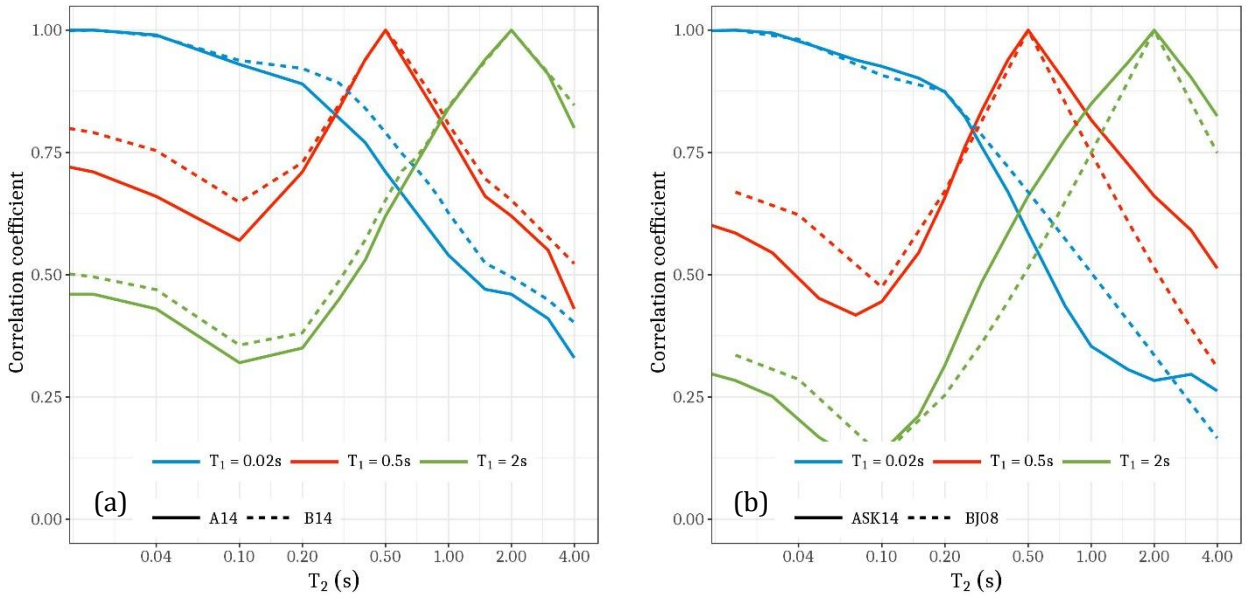


Figure 6-2: Comparison between total-residual correlation models for (a) Akkar et al. (2014) (A14) and Bindi et al. (2014) (B14) derived from RESORCE dataset (b) Abrahamson et al. (2014) (ASK14) and Baker and Jayaram (2008) (BJ08) derived from NGA-West2 and NGA-West datasets, respectively.

6.5 Magnitude dependence of between-event correlations

To investigate the magnitude dependence of between-event correlations, we divided the RESORCE events into small and large magnitude events. *Baker and Cornell* (2006) and *Cimellaro* (2013), in deriving their correlation models from NGA-West and European dataset ISESD respectively, assumed $M5.5$ as the lower magnitude limit for earthquakes of engineering relevance. In terms of ground-motion variability, *Boore et al.* (2014) chose $M5.5$ to separate small from moderate – large events for their heteroscedastic model

of between-event variability. On similar lines, we divide the 176 between-event residuals from *Bindi et al.* (2014) GMPE into 97 small events with $M < 5.5$, and 79 large events with $M \geq 5.5$.

The between-event correlation models computed using between-event residuals from $M < 5.5$ and $M \geq 5.5$ magnitude ranges (Figure 6-3a) show clear differences depending on T_1 and T_2. To check if the differences were not contingent of limited sample sizes, we use Eq. (7) to compute the standard error on the correlation coefficients. For example, for $M < 5.5$ when $T_1 = 0.5$s and $T_2 = 0.1$s $\rho(T_1, T_2) = 0.5$, and for $M \geq 5.5$ it is 0.81. Using Eq. (7) the standard error on a correlation value 0.81 estimated using 79 data points is 0.04, while for the correlation value 0.5 obtained using 97 data points is 0.08. The small standard errors suggest the differences to be statistically significant.

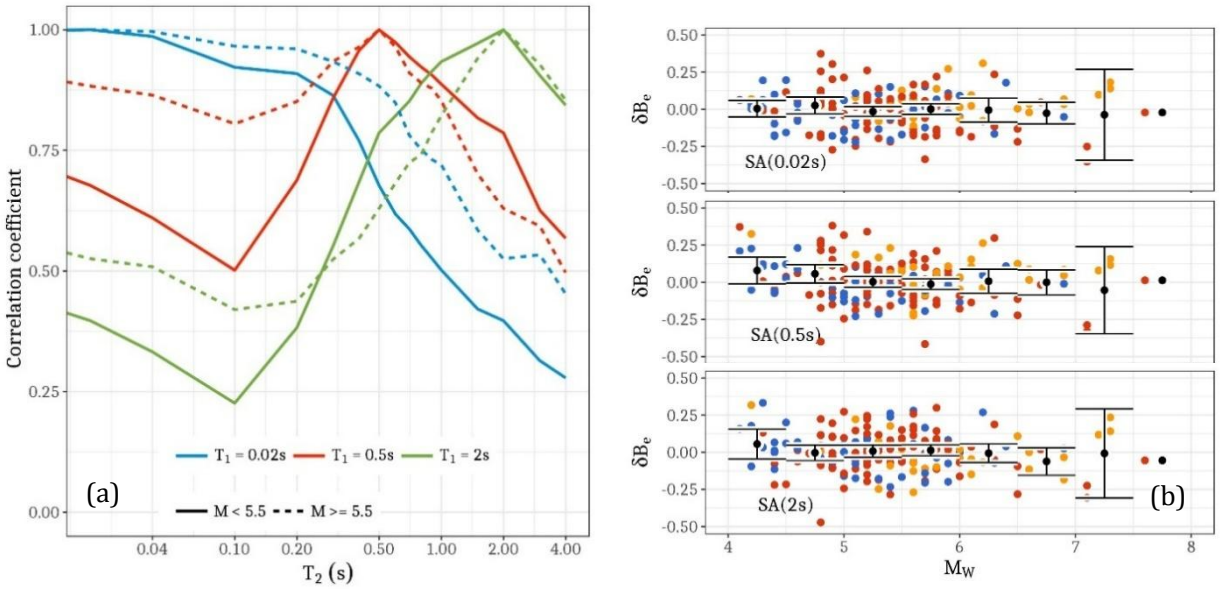


Figure 6-3: Between-event correlations for $M < 5.5$ events and $M \geq 5.5$ obtained in this study using (a) RESORCE dataset (b) Between-event residuals trend and bias within magnitude bins, color coded for the region of event.

The magnitude dependence of between-event correlations, although statistically significant, can also be an artifact of the magnitude scaling component of the GMPE. For instance, the magnitude scaling could be biased by a well-sampled magnitude range resulting in systematically biased between-event residuals in other magnitude ranges. In RESORCE dataset, the most well-sampled magnitude range is between $M5$ and $M6$. To check for any systematic biases in lower and higher magnitude ranges, in Figure 6-3b we plot the residuals against the event magnitude, along with the mean of residuals and standard error within $\Delta M = 0.5$ bins. Figure 6-3b shows that the quadratic magnitude scaling of *Bindi et al.* (2014) GMPE does not introduce a statistically significant bias in any of the magnitude residual bins. We can assert that the magnitude dependence of the (B14) between-event correlations is not an artifact of *Bindi et al.* (2014) GMPE magnitude scaling.

A physical explanation for the magnitude dependence of between-event correlations can be obtained considering a standard omega-square source model by *Brune* (1970). For a given moment magnitude (seismic moment), increase in the static stress drop lowers the corner frequency (f_c), and increases the amplitudes at frequencies higher than the f_c; while keeping the amplitudes at frequencies lower than f_c unaffected. For illustration, we consider as primary frequency $f_c = 1/T_1$ and secondary frequency $f_2 = 1/T_2$. When $(f_1, f_2) < f_c$, neither f_1 nor f_2 is effected by the variability in stress drop, hence their correlations are strong. When $(f_1, f_2) > f_c$, amplitudes at both f_1 and f_2 change monotonically with the static stress drop, hence the correlation of their amplitudes is strong as well. Based on this assertion, for large magnitude events with a low f_c, one would expect stronger correlations over a wide range of $(f_1, f_2) > f_c$. In Figure 6-3a, curves $T_1 = 0.02$s and 0.5s capture this phenomenon as stronger and more gradually decreasing correlations in the range $0.01\text{s} < T_2 < 0.5\text{s}$ for $M \geq 5.5$ events. On the other hand, for small magnitude

events with a high f_c, stronger correlations are expected over a wide range of $(f_1, f_2) > f_c$. In Figure 6-3a, curves $T_1 = 0.5$s and 2s capture this phenomenon as stronger correlations in the range $0.5\text{s} < T_2 < 4\text{s}$ for $M < 5.5$ events.

Magnitude dependence of ground-motion variability (as in *Boore et al.*, 2014 and other NGA-West2 models) supports a higher variability of ground-motion among smaller earthquakes which tend to exhibit more variability than larger events, e.g. in terms of source depth, stress drop, etc. The large number of $M3$-$M4$ events in NGA-West2 dataset may be the probable cause behind their weaker correlations seen in Figure 6-2. In any case, the magnitude dependence of between-event correlations would not be noticeable in the total-residual correlations for the reason that within-event correlations dominate Eq. (3). In the following sections, we first demonstrate the benefit of replacing within-event correlations with event-and-site corrected correlations, and then proceed to exploring regional dependence of correlations.

6.6 Within-event and event-and-site corrected residuals

Using datasets compiled from different regions (e.g. NGA-West2 and RESORCE), the GMPEs developed under ergodic assumption are likely to be biased by a few well-sampled regions, source – site paths, and local site conditions. For instance, if indeed there exist period dependent regional biases in within-event residuals, subsequent ergodic within-event correlation models would be biased as well. In the left panel of Figure 6-4, we show the region dependent distance trend of *Bindi et al.* (2014) within-event residuals. In the right panel of Figure 6-4, we show that event-and-site corrected residuals, resulting from removal of site-residuals (between-site terms) from within-event residuals, are unbiased at near source distances R_{JB}<80km. At far source distances (R_{JB}≥80km), the distance dependent trends persist despite filtering out the between-site residuals, possibly due to regional differences in distance scaling of high frequency SAs (*Kotha et al.*, 2016).

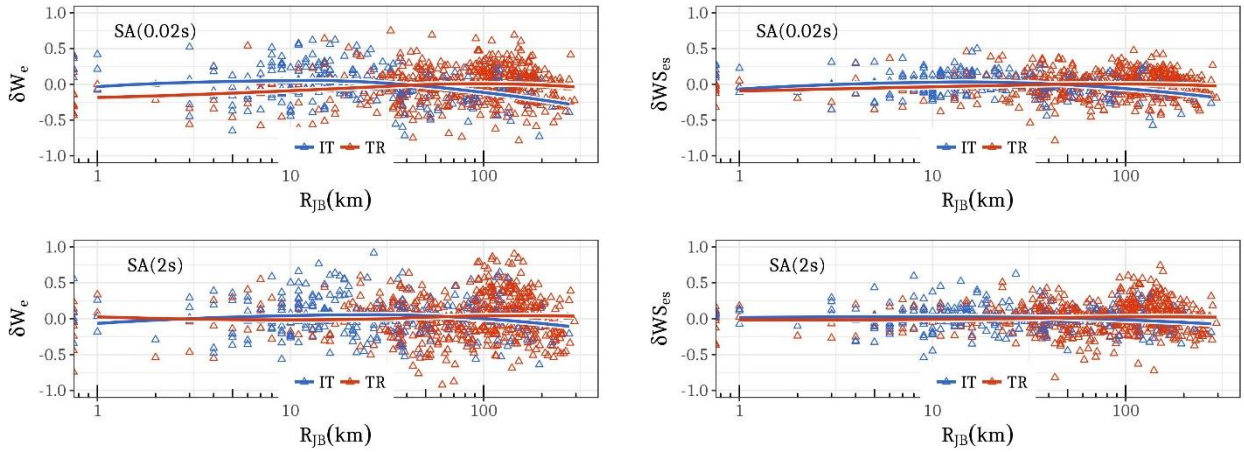


Figure 6-4: Distance dependence of within-event residuals and event-and-site corrected residuals for two spectral periods (0.02s, 2s). Colors indicate the regions: Italy (blue), Turkey (red). The solid lines represent the residual trend with Joyner-Boore distance (km).

Figure 6-5 extends the investigation in Figure 6-4 by comparing the within-event and event-and-site corrected correlations computed using 536 residuals from near source recordings ($0 \leq R_{JB} < 80$km). In this distance range, the within-event residual correlations between short period SAs ($T_1 = 0.02$s) and the moderate – long period SAs ($T_2 > 0.5$s) are significantly weaker than the event-and-site corrected correlations. In a similar exercise using K-NET and KiK-net strong motion data, *Itoi and Takada* (2012) observed the contrary i.e., the within-event correlations are stronger than the event-and-site corrected correlations. *Itoi and Takada* (2012) used strong motion data from 19 aftershocks ($M_J > 5$) following the 2004 Niigata-ken Chuetsu earthquake, with peak horizontal acceleration less than $0.2g$ and hypocentral distances less than 100km. The stations selected by Itoi and *Takada* (2012) are located such that the source – site azimuth (path) is very similar for the 19 aftershocks. In such datasets, the observed within-

event record-to-record variability is small due to similarity of propagation path and site response (from considering only recordings with PSA<0.2g). This may explain why, in their study, removal of between-site residuals renders the event-and-site corrected correlations weaker than the within-event correlations. On the other hand, the RESORCE strong motion used in our study exhibits a much larger record-to-record variability due to the diversity of propagation path and site conditions; consequently the within-event correlations appear weaker than event-and-site corrected correlations (Figure 6-5). We infer from these two studies that, the differences between within-event and event-and-site corrected correlations depend on the initial strong motion data selection. Interestingly, we also noticed that the event-and-site corrected correlations by *Itoi and Takada* (2012) are remarkably close to the RESORCE based models of this study.

Removal of the between-site residuals from the within-event residuals reduces the scatter of event-and-site corrected residuals as seen in Figure 6-4. Figure 6-1 depicts the reduced scatter in terms of a smaller event-and-site corrected variability ϕ_0 compared to within-event variability ϕ. The smaller variance and reduced regional (and site) bias suggest that the event-and-site corrected residuals are more suitable for developing non-ergodic correlation models in combination with our magnitude dependent between-event correlations.

6.7 Regional dependence of event-and-site corrected residuals

Event-and-site corrected correlations can be analyzed for regional variations not accounted in the *Bindi et al.* (2014) GMPE. Of the 536 residuals in the near source distance range $0 \leq R_{JB} < 80$km, 217 are from Italy, 225 from Turkey, and 94 from other regions. The other 350 residuals from far source recordings ($80 \leq R_{JB} < 300$km) are strongly imbalanced in terms of regional contributions: 295 from Turkey, 42 from Italy, and 13 from other pan-European regions. In this study, we investigate region dependence of only the near source event-and-site corrected correlations.

Two correlation models derived using residuals from Italy and Turkey separately, and one using near source residuals from all regions are compared in Figure 6-6. Correlations between moderate and long period SAs ($T > 0.5$s) for the two regions appear to be identical, while the short period SAs ($T < 0.2$s) in Italy are weakly correlated to moderate – long period SAs with respect to Turkey. For example, with $T_1 = 0.5$s and $T_2 = 0.1$s, $\rho(T_1, T_2)$ for Turkey is (0.7 ± 0.03) while for Italy it is (0.47 ± 0.05). The small standard errors imply the differences are not due to limited sample sizes, but due to possible regional differences in propagation paths. There could be two possible reasons for the regional differences:

- In RESORCE, a large fraction of the Italian records ($0 \leq R_{JB} < 80$km) originate from the Central Apennine region, and the L'Aquila sequence. Following the recent 2016 Central Italy sequence (*Lanzano et al.*, 2016), the event-and-site corrected residual analysis of the strong motion data showed significant differences in attenuation of high frequency ground-motions between North-East and South-West regions around the Apennine range. Such local anisotropy of high frequency attenuation may increase the record-to-record variability of short period SAs (e.g. $T = 0.02$s). The lower correlations between short and moderate – long period SAs in Italy could be a consequence of such local phenomenon.
- The focal mechanism of Italian earthquakes is predominantly Normal, while in Turkey it is Strike-Slip. We observed a focal mechanism dependence of the correlations very similar to those in Figure 6-6, i.e. better moderate and long period PSA correlations among Strike-Slip faulting (as with Turkey) compared to the Normal faulting (as with Italy).

It is unclear whether the statistically significant regional differences in correlations are due to differences in the attenuation or due to the differences in source mechanism. Our observations however suggest that correlation models (as with the GMPEs) developed from a compendium dataset (global or pan-European) may not be reliably used across regions discounting the regional differences in predominant rupture mechanisms and high frequency ground-motion attenuation. Hence, we provided region-specific event-

and-site corrected correlations for Italy, Turkey, and in addition, an ergodic region independent model from RESORCE dataset.

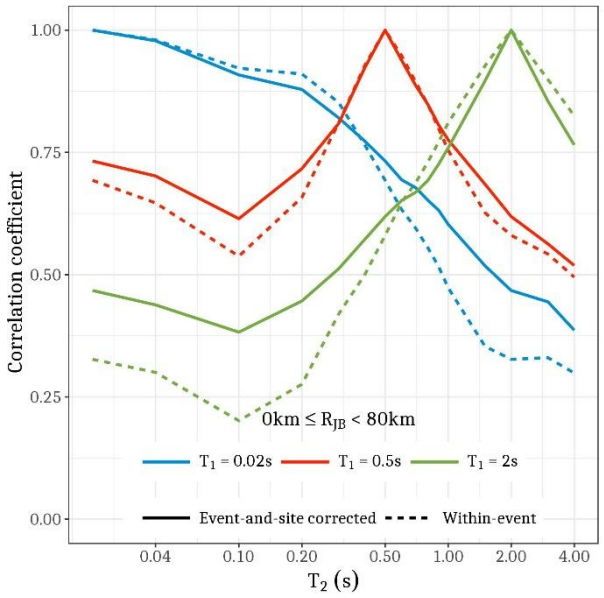


Figure 6-5: Comparison of within-event and event-and-site corrected residual correlations at near source records with $0 \leq R_{JB} < 80$km.

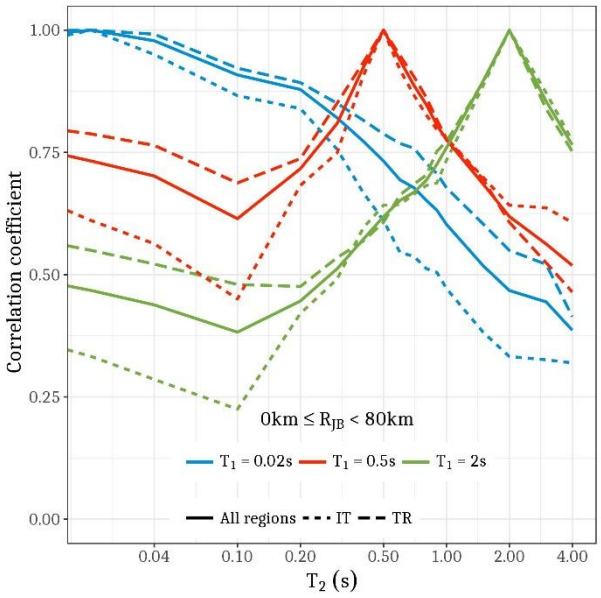


Figure 6-6: Region dependence of event-and-site corrected residual correlations in near source distance range of residuals $0 \leq R_{JB} < 80$km.

6.8 Application: Partially non-ergodic Conditional Spectra

Step-wise procedures to estimate conditional spectra (CS), VPSHA, time history selection, generation of synthetic time history are available in the peer-reviewed literature. In all these applications the residual correlation models are required. Our between-event and event-and-site corrected correlations are combined using Eq. (3) to obtain site-corrected magnitude- and region-dependent correlations. The site corrections, i.e. the between-site residuals ($\delta S2S$ in Eq. 4), can be used as site-specific adjustment factors to scale the GMPE predicted ground-motions to obtain a partially non-ergodic conditional mean spectra (CMS) and CS.

Four sites with at least 10 recordings at all spectral periods from 0.02s to 4s are selected from RESORCE. Sites 1604 and 5401 in Figure 6-7 are from Turkey, with a measured V_{S30} of 457 and 416m/s respectively. Sites AQV and GSA in are from Italy with a measured V_{S30} of 474 and 488m/s respectively. These empirical site-specific adjustment factors (AF equal to $10^{\delta S2S}$) scale the linear site-response prediction of *Bindi et al.* (2014) GMPE. For new sites with new strong motion data, the equations presented in *Rodriguez-Marek et al.* (2013) can be used to estimate the $\delta S2S_s$.

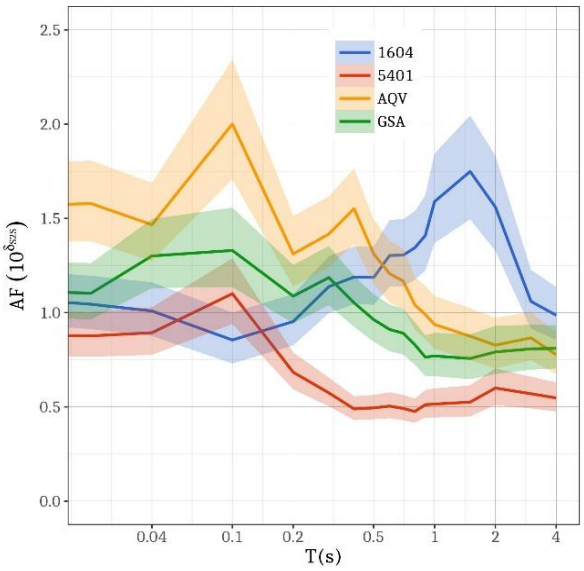


Figure 6-7: Adjustment factor ($10^{\delta S2S}$) for four sites with $V_{S30} = 450\,m/s$ and 10 records in RESORCE dataset. Sites 1604 and 5401 are located in Turkey, and sites AQV and GSA in Italy. The adjustment factors at each spectral period scale the V_{S30} based linear ground-motion response predicted by the Bindi et al. (2014) GMPE to obtain a site-specific ground-motion. The color ribbon indicates the uncertainty in the estimation of $\delta S2S$ due to the limited number of records at a site.

We use the correlation model based on *Bindi et al.* (2014) GMPE residuals, B14 (in Figure 6-3a), as the ergodic (magnitude and region independent) correlation model. The site-corrected magnitude and region dependent models presented in this study are the partially non-ergodic correlation models. For estimation of CMS, we choose PSA at $T_1 = 0.5$s as the primary intensity measure, on which the other secondary intensity measures are conditioned. In a strict sense, the design scenarios are site-specific and are identified via disaggregation of hazard at the site of interest (*Bazzurro and Cornell*, 1999). For a simple demonstration, we chose two design earthquakes characterized by (M, R_{JB}, V_{S30}, η), where η indicates the number of standard deviations. As a small earthquake scenario, we chose a scenario with (M, R_{JB}, V_{S30}, η) = (5, 20, 450, 1). As a large earthquake scenario, we chose (M, R_{JB}, V_{S30}, η) = (6, 40, 450, 1). In Figure 6-8, the ergodic CMS (identified with a solid black line) is common for all the four sites, irrespective of the region and site. The partially non-ergodic CMS (colored lines in Figure 6-8) are unique to each magnitude, region and site combination. For example, in Figure 6-8 the selective amplification of SA(0.1s) and de-amplification at SA(0.2s) at site AQV is in accordance with its AF in Figure 6-7. Such behavior is observed across all sites in this example, suggesting the partially non-ergodic conditional mean spectrum as a better tool for site-specific hazard and risk assessment.

6.9 Conclusions

Correlations of horizontal peak spectral amplitudes (PSA) are widely used in calculating the conditional spectra, in vector-valued probabilistic seismic hazard assessment (PSHA), time history selection for dynamic analysis, etc. We first showed that empirical correlation models are dataset dependent, which can be attributed to the differences in magnitude range, and regional diversity of the strong motion dataset. For practical applications in pan-European region, we suggest using the correlation models developed in conjunction with the pan-European Ground-Motion Prediction Equations (GMPEs).

Using the pan-European events in RESORCE dataset we observed a significant magnitude dependence of the between-event correlations. The magnitude dependence of between-event correlations is a product of differences in the source spectrum between large and small magnitude events, particularly in terms of stress-drop parameter and its influence on short and moderate period SAs. We provided between-event correlation models for small events ($M < 5.5$) and large events ($M \geq 5.5$) separately.

In case of RESORCE dataset and *Bindi et al.* (2014) GMPE, we observed a regional bias of within-event residuals with distance. Such biases in residuals may bias the subsequent residual correlation models. Removing the site-to-site variability, in terms of between-site residuals, from within-event residuals yields the event-and-site corrected residuals with reduced regional biases. We recommend replacing the

potentially biased within-event correlations with event-and-site corrected correlations. In further analysis, we observed that the near source event-and-site corrected residual correlations are region dependent. The near source records from Turkey show a stronger correlation between short and moderate period SAs compared to Italy. We offer two physical reasons for the region dependence: (1) the difference in attenuation of high frequency SAs between North-East and South-West of Apennine ranges, where most of the Italian data is concentrated, could result in higher variability of short period PSA residuals, and thus lower their correlation with moderate – long period SAs. (2) The regional differences could in fact be due to the differences in predominant focal mechanisms of Turkey and Italy; Strike-slip in Turkey, and Normal faulting in Italy. However, in this study, we hypothesize the regional differences to be a combination of both crustal effects and focal mechanisms.

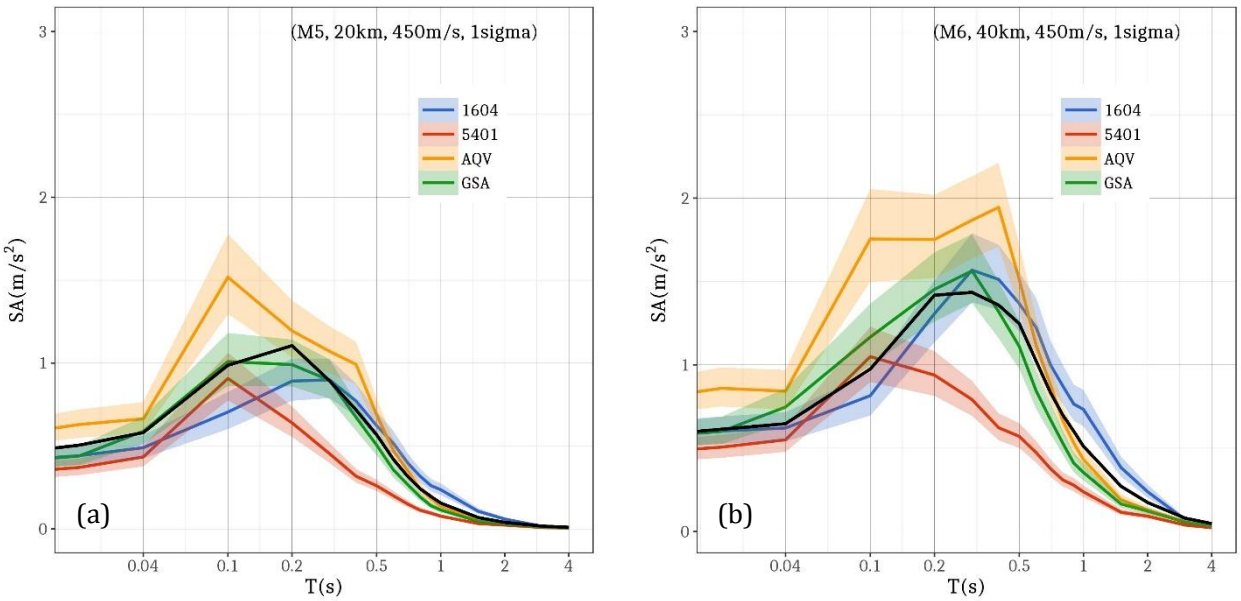


Figure 6-8: Comparison of ergodic CMS and partially non-ergodic CMS: Ergodic CMS (black) and non-ergodic CMS(colored) for (a) small magnitude events (b) large magnitude events in Italy and Turkey. The colored ribbons represent the epistemic uncertainty on site-specific adjustment factors (AF).

The magnitude dependent between-event correlations and region dependent event-and-site corrected correlations in combination yield the partially non-ergodic site-corrected correlation models. Using our partially non-ergodic correlations and the site-specific adjustment factors, we demonstrated the calculation of partially non-ergodic conditional mean spectra (CMS) at two sites from Italy, and two from Turkey with similar V_{S30}. The non-ergodic CMS better represents the linear site-specific response compared to the traditional ergodic CMS.

Acknowledgments

We thank Carlo Cauzzi, Graeme Weatherill and Peter Stafford for their insightful and constructive reviews which have significantly improved the manuscript. The authors are also grateful to Luis Dalguer and Philippe Renault (Swiss Nuclear) who have triggered our interest in this topic. The initial phase of this work has been developed within the framework of the STREST European project (FP7/2007-2013 under grant agreement no. 60338) and the later parts within the framework of the EPOS-IP European GMPE service (EU-H2020 project n° 676564). We also wish to thank Frank Scherbaum, Sanjay Bora and our colleagues at section 2.6 of GFZ German Research Centre for Geosciences for the fruitful discussions.

7
Conclusions and Outlook

7.1 Conclusions

Goal of this study was to quantify and mitigate the prediction uncertainties of ground-motion models (Ground-Motion Predictive Equations (GMPEs)) and improve the reliability of Probabilistic Seismic Hazard Assessments (PSHA) estimations. In process, we worked with global and regional ground-motion datasets, developed statistical tools to quantify the various contributors to GMPE variances, and suggested methodological refinements to the development of region- and site-specific GMPEs for use in empirical site-specific PSHA. Chapter 1 provides an overview of the origin of GMPE uncertainties, their technical (statistical) definition, and steps to resolve them. Subsequent chapters constitute the series of steps in achieving the goal.

Chapter 2 discusses the very fundamental GMPE uncertainty – the statistical uncertainty on regression fits. Since GMPEs are essentially regressed formulations of observed ground-motion scaling with event-path-site parameters, their regression coefficients are well-constrained (or not) depending on the density of recorded ground-motion data in different ranges of event-path-site parameters. The asymptotic variance σ_μ^2, quantifying this uncertainty of regression fit, is large over poorly sampled event-path-site parametric ranges – making predictions for under-sampled scenarios unreliable.

In this chapter, we developed a GMPE from the NGA-West2 dataset, for application in PSHA of non-cratonic moderate seismicity regions (e.g. Germany). The high quality of event-path-site metadata in NGA-West2 ensures negligible contributions to GMPE variances arising from incorrect metadata. However, absence of sufficient data from large earthquakes ($M_w \geq 7.5$) and at short distances ($R_{JB} < 10$km) results in large σ_μ (prediction uncertainty) for such scenarios. For instance, for an $M7.5$ rupture occurring at 25km from a site, $\sigma_\mu = 0.25$ implies the GMPE prediction of SA(2s) is uncertain by approximately 30%. For less frequent scenarios, the uncertainties are even larger.

Our analysis corroborates the need to collect more data from near-source distances in the vicinity of active fault systems, so as to reduce the prediction uncertainty of GMPEs. Although not emphasized in this chapter, I observe it is crucial to sample ground-motion data over a wide range of site parameters as well. Typical PSHA practice requires GMPEs to predict accurate ground-motions on generic rock sites ($V_{S30} \geq$ 800m/s). But most ground-motion datasets, global and regional, lack sufficient data from such reference rock sites, which result in poorly constrained GMPEs and unreliable hazard estimates.

Chapter 3 discusses the two contributors to GMPE aleatory variability σ – incorrect metadata and the ergodic assumption. The statistical uncertainty of regression fit, quantified in σ_μ, can be minimized by collecting more ground-motion data in a region, for an extended period of time. An alternative to temporal sampling of event-path-site parameters (in a region) is to aggregate spatially diverse samples from several, tectonically similar, regions – the ergodic assumptions in GMPE development. Along with the ergodic assumptions come the issues of errors (and inhomogeneity) of event-path-site metadata and the inefficiency of parametrization. In this chapter, I worked on developing a regionalized GMPE for pan-European region, using only the data with reliably accurate metadata.

Our first finding was that the incorrect event metadata of the many $M_w < 5$ earthquakes from Italy and Turkey inflated the between-event variability τ by 30%. Discarding data from such event in the GMPE regressions reduces the total GMPE standard deviation σ by 10%, which has significant impact on PSHA at pan-European sites. In fact, these events were originally parametrized on a magnitude scale different from M_w; e.g. region-specific magnitude scales such M_L, M_s, etc. Using region-specific empirical relation-

ships between $M_w - M_L$ and $M_w - M_s$, these magnitude estimates were then converted into M_w to maintain magnitude scale homogeneity in the RESORCE dataset. However, the conversions being region-specific, introduced a regional ergodicity into magnitude scaling component of the GMPE, and led to an inflated estimate of between-event variability. Discarding such data reduces the bias and variance of the GMPE, but a more desirable remedy would be to homogenize the magnitude scale across the several active tectonic regions in pan-Europe.

The second, more interesting, outcome of our GMPE regionalization is the regional variability of ground-motion scaling with path and site parameters. The mixed-effects regression I employed could quantify statistically significant differences in apparent anelastic attenuation of high frequency ground-motions and (linear) scaling of ground-motion with V_{S30}, between Italy and Turkey. Allowing the GMPE to account regional variability of ground-motion scaling did not necessarily decrease the σ of GMPE, but improved its explanatory and predictive power (measured in terms of AIC and BIC). These findings were later affirmed by *Bora et al.* (2017), in terms of regional differences in crustal quality factor (Q) between Italy and Turkey. In regionalizing the GMPE, I could resolve partially the region-level ergodicity in the GMPE and remove regional biases in its prediction.

Chapter 4 is primarily focused on furthering the region-specific GMPE into a region- and site-specific GMPE, and refining the empirical site-specific PSHA procedure. In addition, I illustrated the significance of resolving region and site ergodicities in GMPEs by comparing the ergodic PSHA with region- and site-specific PSHA estimates at 225 sites in pan-European region.

In this chapter, I further exploited the advantages of mixed-effects GMPE regression techniques in resolving ergodicity in GMPEs. I reaffirmed *Stafford* (2014) that the existing method of estimating site-specific terms from the GMPE residuals (*Rodriguez-Marek et al.*, 2013) is an approximate method, by demonstrating its misconfiguration of GMPE median and variances. Mixed-effects GMPE regression as I performed in this chapter, isolates site-specific non-ergodic biases more accurately, and also appears to provide a better constrain on the regression. I note once again that the PSHA estimates are very sensitive to GMPE median and variances, and even more so to the region- and site-specific adjustments.

The key outcome of this chapter is an illustration of the impact on PSHA from resolving region and site ergodicity in the GMPEs. I demonstrated that at several (well-sampled) sites in pan-European region, the region- and site-specific hazard estimates vary from ergodic estimates by 50% at long return periods. I see our results as a motivation to shift towards region- and site-specific PSHA. However, the reliability of non-ergodic GMPEs depends on the amount of data used in constraining the non-ergodic biases; emphasizing the need to collect more data at several sites.

Chapter 5 explores an engineering solution to the insufficiency of site-specific ground-motion data to estimate site-terms. In absence of site-specific ground-motion data, a common practice is to attribute approximate site-terms based on the best representative site-class. Traditionally, site-classes are identified by their characteristic amplification factors, and defined by characteristic ranges of site parameters, or site-response proxies (e.g. V_{S30}). However, the choice of parametrization and predefined parametric ranges (of proxies) introduce subjectivity in site-classification, which appears to reduce their efficiency in predicting site-response (through GMPEs). Therefore, I sought an alternative data-driven site-classification scheme to complement our partially non-ergodic region- and site-specific GMPEs.

With increasing ground-motion data, it became possible to reliably estimate (predominantly linear) site-terms at several sites in active shallow crustal regions, e.g. Japan. I applied the mixed-effects GMPE regression on KiK-Net dataset to estimate site-terms at ~600 sites in Japan. These ~600 site-terms were then processed with spectral clustering techniques to identify sites with similar site-response – optimal number of clusters being the only subjective input.

Our approach could identify 8 distinct site clusters in KiK-Net ground-motion network. These 8 clusters could serve as the new site-classes, provided they can be characterized by a combination of site-response

proxy parameters. Our primary inference from this chapter is the inefficiency of preferred site-response proxy for site classification V_{S30}. Within the existing Eurocode8 reference rock site-class of soil-type A, I could identify two sub-classes with radically different high frequency amplification. These sub-classes cannot be characterized based on their V_{S30} ranges alone (as in EC8), but a combination site-proxy (V_{S10}, H_{800}) was verified to perform better. In addition, the spectral clustering based site-classification being more efficient and sufficient in capturing the diversity in site-response, reduced the ϕ_{S2S} by ~50%, which is roughly a 15% reduction in σ of GMPE. With the improvement in accuracy and reduction in variance of site-response prediction, our approach could bridge the gap between ergodic and partially non-ergodic site-specific GMPEs.

Chapter 6 is aimed at predicting a site-specific magnitude and region-dependent Conditional Spectra, using empirically modeled correlation of peak spectral amplitudes. Existing correlation models were developed from total residuals, assuming no event-path-site dependencies in ground-motion. Indeed the total residuals (ε) and their correlations show no event-path-site specificities. But when they were dissociated into event, site and path components, I observed significant magnitude (event) dependence of between-event (δB_e) correlations and region (path) dependence of event-and-site corrected (δWS_{es}) correlation coefficients over the period range $T = 0.01$-2s. Using our magnitude and region dependent correlation models, along with the site-terms, I proposed refinements to the computation of Conditional Spectra. Our new procedure accounts for differences in source spectra of small and large events, regional differences in attenuation with distance, and site-specific amplification of ground-motions.

7.2 Outlook

Our study is aimed at characterizing and reducing prediction uncertainties of GMPE, in order to make PSHA more reliable. In process I developed non-ergodic GMPEs and demonstrated their usefulness. Of course, the strongest inference is that we need to collect more ground-motion data if we were to develop better predictive models. Even with less than optimal amount of data, I find the mixed-effects regression and machine learning techniques to have a remarkable potential in quantifying uncertainties and refining the PSHA framework. To summarize the prospects of our study:

- Asymptotic variance is usually an unreported statistical uncertainty in GMPEs, whose disclosure can foster better selection of GMPEs in PSHA, depending on the seismicity of a region and the predominant geological conditions
- Mixed-effects regression based region- and site-specific GMPEs are unbiased by ergodicity, and may serve as regionally adaptable backbone models for regions with too little ground-motion data
- Considering the alternative of discarding large amounts of usable but incompatibly parametrized ground-motion data from many pan-European regions, homogenization of magnitude scale and re-evaluation (and parametrization) of event-path-site metadata is a more desirable proposal
- Spectral clustering techniques could identify hidden site classes – it would be interesting to extend this application to investigate, say, event classes or to localize path effects
- Removal of the highly variable site-term from the residuals allowed us to quantify a variety of known physical effects in correlation models, which were otherwise masked by dominance of site-to-site variability. It is worth investigating the residuals for more repeatable patterns, introduce them into predictive models, and improve our probabilistic assessments
- Given the immense potential of site-terms in performing empirical site-specific PSHA, I am strongly motivated to develop methods for their rapid estimation. Our immediate focus is therefore to use weak-motion data along with teleseismic data, investigate scenario dependence site-response, and propose a framework to predict site-terms in low seismicity regions of Europe

I firmly believe in the need for denser seismological instrumentation to collect more ground-motion data. At the same time, I am optimistic that appropriate statistical tools can accelerate our progress towards fully non-ergodic ground-motion prediction.

8
References

Abrahamson N, Youngs RR (1992) A stable algorithm for regression analyses using the random effects model. Bulletin of the Seismological Society of America 82(1), 505–510.

Abrahamson NA, Silva WJ (1997) Empirical response spectral attenuation relations for shallow crustal earthquakes. Seismological Research Letters 68(1), 94-127.

Abrahamson N, Silva WJ (2008) Summary of the Abrahamson & Silva NGA ground-motion relations. Earthquake Spectra 24(1), 67-97.

Abrahamson NA, Hollenback JC (2012) Application of single-site sigma ground motion prediction equations in practice, Paper No. 2536, In: Proceedings, 15th World Conference on Earthquake Engineering, 24-28 September 2012, Lisbon, Portugal.

Abrahamson NA, Silva WJ, Kamai R. (2014) Summary of the ASK14 ground motion relation for active crustal regions. Earthquake Spectra 30(3), 1025-1055.

Association JR (1980) Specifications for road bridges. Vol. IV (in Japanese).

Association JR (1990) Specification for highway bridges. Part V: seismic design, Tokyo, Japan.

Afshari K, Stewart JP (2015) Uncertainty of site amplification derived from ground response analysis. UCLA: 1304269. Retrieved from: http://escholarship.org/uc/item/44z6z996

Akaike H (1973) Information theory and an extension of the maximum likelihood principle. In: Petrov BN, Csáki F (eds) 2nd International Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971, Budapest: Akadémiai Kiadó, pp 267-281.

Akkar S, Çağnan Z, Yenier E, Erdoğan Ö, Sandıkkaya MA, Gülkan P (2010) The recently compiled Turkish strong motion dataset: preliminary investigation for seismological parameters. Journal of Seismology 14(3), 457-479.

Akkar S, Sandıkkaya MA, Ay BÖ (2014) Compatible ground-motion prediction equations for damping scaling factors and vertical-to-horizontal spectral amplitude ratios for the broader Europe region. Bulletin of Earthquake Engineering 12(1), 517-547.

Akkar S, Sandıkkaya MA, Bommer JJ (2014) Empirical ground-motion models for point-and extended-source crustal earthquake scenarios in Europe and the Middle East. Bulletin of Earthquake Engineering 12(1), 359-387.

Akkar S, Sandıkkaya MA, Şenyurt M, Sisi AA, Ay BÖ, Traversa P, Douglas J, Cotton F, Luzi L, Hernandez B, Godey S (2014) Reference database for seismic ground-motion in Europe (RESORCE). Bulletin of Earthquake Engineering 12(1), 311–339.

Al Atik L, Abrahamson N, Bommer JJ, Scherbaum F, Cotton F, Kuehn N (2010) The variability of ground-motion prediction models and its components. Seismological Research Letters 81(5), 794-801.

Al Atik L, Youngs RR (2014) Epistemic Uncertainty for NGA-West2 Models. Earthquake Spectra 30(3), 1301-1318.

Allen TI, Wald DJ (2009) On the use of high-resolution topographic data as a proxy for seismic site conditions (V_{S30}). Bulletin of the Seismological Society of America 99(2A), 935-943.

Ambraseys N, Smit P, Douglas J, Margaris B, Sigbjörnsson R, Olafsson S, Suhadolc P, Costa G (2004) Internet site for European strong motion data. Bollettino di Geofisica Teorica ed Applicata 45(3), 113-129.

Ancheta TD, Darragh RB, Stewart JP, Seyhan E, Silva WJ, Chiou BS-J, Wooddell KE, Graves RW, Kottke AR, Boore DM, Kishida T, Donahue JL (2014) NGA-West2 database. Earthquake Spectra 30(3), 989-1005.

Anderson JG, Brune JN (1999) Probabilistic seismic hazard analysis without the ergodic assumption. Seismological Research Letters 70(1), 19-28.

Atkinson GM, Casey R (2003) A comparison of ground motions from the 2001 M 6.8 in-slab earthquakes in Cascadia and Japan. Bulletin of the Seismological Society of America 93(4), 1823-1831.

Atkinson GM (2006) Single-station sigma. Bulletin of the Seismological Society of America 96(2), 446-455.

Azarbakht A, Mousavi M, Nourizadeh M, Shahri M (2014) Dependence of correlations between spectral accelerations at multiple periods on magnitude and distance. Earthquake Engineering & Structural Dynamics 43(8), 1193-1204.

Baker JW, Cornell CA (2005) A vector-valued ground motion intensity measure consisting of spectral acceleration and epsilon. Earthquake Engineering & Structural Dynamics 34(10), 1193-1217.

Baker JW, Cornell CA (2006) Correlation of response spectral values for multicomponent ground motions. Bulletin of the Seismological Society of America 96(1), 215-227.

Baker JW, Jayaram N (2008) Correlation of spectral acceleration values from NGA ground motion models. Earthquake Spectra 24(1), 299-317.

Baker JW (2010) Conditional mean spectrum: Tool for ground-motion selection. Journal of Structural Engineering 137(3), 322-331.

Baker JW, Bradley BA (2017) Intensity Measure Correlations Observed in the NGA-West2 Database, and Dependence of Correlations on Rupture and Site Parameters. Earthquake Spectra 33(1), 145-156.

Baltay A, Hanks T, Abrahamson N (2017) Uncertainty, Variability, and Earthquake Physics in Ground-Motion Prediction Equations. Bulletin of the Seismological Society of America 107(4), 1754-1772.

Bates D, Mächler M, Bolker B, Walker S (2014) lme4: Linear mixed-effects models using Eigen and S4. R package version, 1(7).

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67(1), 1-48.

Baturay MB, Stewart JP (2003) Uncertainty and bias in ground-motion estimates from ground response analyses. Bulletin of the Seismological Society of America 93(5), 2025-2042.

Bazzurro P, Cornell CA (1999) Disaggregation of seismic hazard. Bulletin of the Seismological Society of America 89(2), 501-520.

Bazzurro P, Cornell CA (2002) Vector-valued probabilistic seismic hazard analysis (VPSHA). In: Proceedings of the 7th US national conference on earthquake engineering (pp. 21-25).

Bazzurro P, Cornell CA (2004) Nonlinear soil-site effects in probabilistic seismic-hazard analysis. Bulletin of the Seismological Society of America 94(6), 2110-2123.

Beauval C, Tasan H, Laurendeau A, Delavaud E, Cotton F, Gueguen P, Kuehn N (2012) On the Testing of Ground-Motion Prediction Equations against Small-Magnitude Data. Bulletin of the Seismological Society of America 102(5), 1994-2007.

Bindi D, Massa M, Luzi L, Ameri G, Pacor F, Puglia R, Augliera P (2014) Pan-European ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods up to 3.0s using the RESORCE dataset. Bulletin of Earthquake Engineering 12(1), 391-430.

Bindi D (2017) The Predictive Power of Ground-Motion Prediction Equations. Bulletin of the Seismological Society of America 107(2), 1005-1011.

Bindi D, Cotton F, Kotha SR, Bosse C, Stromeyer D, Grünthal G (2017) Application-driven ground motion prediction equation for seismic hazard assessments in non-cratonic moderate-seismicity areas. Journal of Seismology 21(5), 1201-1218.

Bommer JJ, Abrahamson NA (2006) Why do modern probabilistic seismic-hazard analyses often lead to increased hazard estimates? Bulletin of the Seismological Society of America 96(6), 1967-1977.

Bommer JJ, Akkar S (2012) Consistent Source-to-Site Distance Metrics in Ground-Motion Prediction Equations and Seismic Source Models for PSHA. Earthquake Spectra 28(1), 1-15.

Boore DM (2004) Can site response be predicted? Journal of earthquake Engineering 8(spec01), 1-41.

Boore DM, Atkinson GM (2008) Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 and 10s. Earthquake Spectra 24(1), 99–138.

Boore DM (2009) Comparing stochastic point-source and finite-source ground-motion simulations: SMSIM and EXSIM. Bulletin of the Seismological Society of America 99(6), 3202-3216.

Boore DM, Thompson EM, Cadet H (2011) Regional correlations of V_{S30} and velocities averaged over depths less than and greater than 30 meters. Bulletin of the Seismological Society of America 101(6), 3046-3059.

Boore DM, Stewart JP, Seyhan E, Atkinson GM (2014) NGA-West2 equations for predicting PGA, PGV, and 5% damped PSA for shallow crustal earthquakes. Earthquake Spectra 30(3), 1057-1085.

Bora SS, Scherbaum F, Kuehn N, Stafford P (2014) Fourier spectral-and duration models for the generation of response spectra adjustable to different source-, propagation-, and site conditions. Bulletin of Earthquake Engineering 12(1), 467-493.

Bora SS, Scherbaum F, Kuehn N, Stafford P (2016) On the Relationship between Fourier and Response Spectra: Implications for the Adjustment of Empirical Ground-Motion Prediction Equations (GMPEs). Bulletin of the Seismological Society of America 106(3), 1235-1253.

Bora SS, Cotton F, Scherbaum F, Edwards B, Traversa P (2017) Stochastic source, path and site attenuation parameters and associated variabilities for shallow crustal European earthquakes. Bulletin of Earthquake Engineering. 15(11), 4531-4561.

Borcherdt RD, Glassmoyer G (1992) On the characteristics of local geology and their influence on ground motions generated by the Loma Prieta earthquake in the San Francisco Bay region, California. Bulletin of the Seismological Society of America 82(2), 603-641.

Borcherdt RD (1994) Estimates of site-dependent response spectra for design (methodology and justification). Earthquake Spectra 10(4), 617-653.

Bradley BA (2013) Systematic Ground Motion Observations in the Canterbury Earthquakes and Region-Specific Non-Ergodic Empirical Ground Motion Modelling, UC Research Report 2013-03, Department of Civil and Natural Resources Engineering, University of Canterbury, Christchurch, New Zealand, 74 pp.

Brillinger DR, Preisler HK (1984) An exploratory analysis of the Joyner-Boore attenuation data. Bulletin of the Seismological Society of America 74(4), 1441-1450.

Brillinger DR, Preisler HK (1985) Further analysis of the Joyner-Boore attenuation data. Bulletin of the Seismological Society of America 75(2), 611-614.

Brune JN (1970) Tectonic stress and the spectra of seismic shear waves from earthquakes. Journal of Geophysical Research 75(26), 4997-5009.

Cadet H, Bard PY, Duval AM (2008) A new proposal for site classification based on ambient vibration measurements and the Kiknet strong motion data set. In: Proceedings of the 14th World Conference on Earthquake Engineering, Beijing (pp. 12-17).

Campbell KW (1981) Near-source attenuation of peak horizontal acceleration. Bulletin of the Seismological Society of America 71(6), 2039-2070.

Campbell KW, Bozorgnia Y (2014) NGA-West2 Ground Motion Model for the Average Horizontal Components of PGA, PGV, and 5% Damped Linear Acceleration Response Spectra. Earthquake Spectra 30(3), 1087–1115.

Castellaro S, Mulargia F, Rossi PL (2008) VS30: Proxy for seismic amplification? Seismological Research Letters 79(4), 540-543.

Cauzzi C, Faccioli E (2017) Anatomy of sigma of a global predictive model for ground motions and response spectra. Bulletin of Earthquake Engineering, online first. doi: 10.1007/ s10518-017-0278-4.

Cauzzi C, Edwards B, Fäh D, Clinton J, Wiemer S, Kästli K, Cua G, Giardini D (2015). New predictive equations and site amplification estimates for the next-generation Swiss ShakeMaps, Geophysical Journal International 200(1), 421-438.

Cauzzi C, Faccioli E, Vanini M, Bianchini A (2015). Updated predictive equations for broadband (0.01-10s) horizontal response spectra and peak ground motions, based on a global dataset of digital acceleration records. Bulletin of Earthquake Engineering 13(6), 1587–1612.

Chambers JM, Hastie TJ (1991) Statistical models in S. CRC Press, Inc.

Chiou BS-J, Darragh R, Gregor N, Silva W (2008) NGA project strong motion dataset. Earthquake Spectra 24(1), 23-44.

Chiou BS-J, Youngs RR (2008) An NGA model for the average horizontal component of peak ground motion and response spectra. Earthquake Spectra 24(1), 173-215.

Chiou BS-J (2012) Updating the Chiou and Youngs NGA model: Regionalization of anelastic attenuation. In: Proceedings, 15th World Conference on Earthquake Engineering.

Chiou BS-J, Youngs RR (2014) Update of the Chiou and Youngs NGA model for the average horizontal component of peak ground motion and response spectra. Earthquake Spectra 30(3), 1117-1153.

Cimellaro GP (2013) Correlation in spectral accelerations for earthquakes in Europe. Earthquake Engineering & Structural Dynamics 42(4), 623-633.

Code P (2005) Eurocode 8 Design of structures for earthquake resistance-part 1: general rules, seismic actions and rules for buildings.

Cotton F, Scherbaum F, Bommer JJ, Bungum H (2006). Criteria for selecting and adjusting ground-motion models for specific target regions: Application to central Europe and rock sites. Journal of Seismology 10(2), 137-156.

Council BSS (2000) The 2000 NEHRP Recommended Provisions for New Buildings and Other Structures, Part I (Provisions) and Part II (Commentary). FEMA, 368: p. 369.

Crowley H, Monelli D, Pagani M, Silva V, Weatherill G, Rao A (2015) The OpenQuake-engine User Manual. Global Earthquake Model (GEM) Technical Report 2015-12. doi: 10.13117/ GEM.OPENQUAKE.MAN.ENGINE.1.6/01, 152 pages

Dawood HM, Rodriguez-Marek A (2013) A method for including path effects in ground-motion prediction equations: An example using the M_w 9.0 Tohoku earthquake aftershocks. Bulletin of the Seismological Society of America 103(2B), 1360-1372.

Dawood HM, Rodriguez-Marek A, Bayless J, Goulet C, Thompson E (2016) A Flatfile for the KiK-net Database Processed Using an Automated Protocol. Earthquake Spectra 32(2), 1281-1302.

Delavaud E, Cotton F, Akkar S, Scherbaum F, Danciu L, Beauval C, Drouet S, Douglas J, Basili R, Sandikkaya MA, Segou M, Faccioli E, Theodoulidis N (2012) Toward a ground-motion logic tree for probabilistic seismic hazard assessment in Europe. Journal of Seismology 16(3), 451-473.

Derras B, Bard PY, Cotton F, Bekkouche A (2012) Adapting the neural network approach to PGA prediction: an example based on the KiK-net data. Bulletin of the Seismological Society of America 102(4), 1446-1461.

Derras B, Bard P-Y, Cotton F (2016) Site-Condition Proxies, Ground Motion Variability, and Data-Driven GMPEs: Insights from the NGA-West2 and RESORCE Data Sets. Earthquake Spectra 32(4), 2027-2056.

Derras B, Bard P-Y, Cotton F (2017) V_{S30}, slope, H_{800} and f_0: performance of various site-condition proxies in reducing ground-motion aleatory variability and predicting nonlinear site response. Earth, Planets and Space 69(1), 133.

Douglas J (2004) An investigation of analysis of variance as a tool for exploring regional differences in strong ground motions. Journal of Seismology 8(4), 485-496.

Douglas J (2007) On the regional dependence of earthquake response spectra. ISET Journal of Earthquake Technology 44(1), 71-99.

Douglas J, Akkar S, Ameri G, Bard PY, Bindi D, Bommer JJ, Bora SS, Cotton F, Derras B, Hermkes M, Kuehn NM, Luzi L, Massa M, Pacor F, Riggelsen C, Sandıkkaya MA, Scherbaum F, Stafford PJ, Traversa P (2014) Comparisons among the five ground-motion models developed using RESORCE for the prediction of response spectral accelerations due to earthquakes in Europe and the Middle East. Bulletin of Earthquake Engineering 12(1), 341-358.

Douglas J, Aochi H (2016) Assessing Components of Ground-Motion Variability from Simulations for the Marmara Sea Region (Turkey). Bulletin of the Seismological Society of America 106(1), 300-306.

Drouet S, Cotton F (2015) Regional Stochastic GMPEs in Low Seismicity Areas: Scaling and Aleatory Variability Analysis - Application to the French Alps. Bulletin of the Seismological Society of America 105(4), 1883-1902.

Edwards B, Fäh D (2013) A stochastic ground-motion model for Switzerland. Bulletin of the Seismological Society of America 103(1), 78–98.

Edwards B, Cauzzi C, Danciu L, Fäh D (2016). Region-Specific Assessment, Adjustment, and Weighting of Ground-Motion Prediction Models: Application to the 2015 Swiss Seismic-Hazard Map. Bulletin of the Seismological Society of America 106(4), 1840-1857.

Faccioli E, Paolucci R, Vanini M (2015) Evaluation of Probabilistic Site-Specific seismic-hazard methods and associated uncertainties, with applications in the Po Plain, Northern Italy. Bulletin of the Seismological Society of America 105(5), 2787-2807.

Gallipoli MR, Mucciarelli M (2009) Comparison of site classification from VS30, VS10, and HVSR in Italy. Bulletin of the Seismological Society of America 99(1), 340-351.

Garcia D, Wald DJ, Hearne M (2012) A global earthquake discrimination scheme to optimize ground-motion prediction equation selection. Bulletin of the Seismological Society of America 102(1), 185-203.

Ghofrani H, Atkinson GM, Goda K (2013). Implications of the 2011 M9. 0 Tohoku Japan earthquake for the treatment of site effects in large earthquakes. Bulletin of Earthquake Engineering 11(1), 171-203.

Gianniotis N, Kuehn N, Scherbaum F (2014) Manifold aligned ground motion prediction equations for regional datasets. Computers & Geosciences 69, 72-77.

Goulet CA, Stewart JP (2009) Pitfalls of deterministic application of nonlinear site factors in probabilistic assessment of ground motions. Earthquake Spectra 25(3), 541-555.

Grünthal G, Bosse C, Stromeyer D, Cotton F, Bindi D (2017) The Probabilistic Seismic Hazard Assessment of Germany, Version 2016 - Considering the Ranges of Epistemic Uncertainties and Aleatory Variabilities. Submitted to Bulletin of Earthquake Engineering.

Héloïse C, Bard P-Y, Duval A-M, Bertrand E (2012) Site effect assessment using KiK-net data: part 2 - site amplification prediction equation based on f_0 and V_{sz}. Bulletin of Earthquake Engineering 10(2), 451-489.

Hermkes M, Kuehn NM, Riggelsen C (2014) Simultaneous quantification of epistemic and aleatory uncertainty in GMPEs using Gaussian process regression. Bulletin of Earthquake Engineering 12(1), 449-466.

Hotelling H (1953) New light on the correlation coefficient and its transforms. Journal of the Royal Statistical Society, Series B (Methodological) 15(2), 193-232.

Idriss IM (2014). An NGA-West2 Empirical Model for Estimating the Horizontal Spectral Values Generated by Shallow Crustal Earthquakes. Earthquake Spectra. Vol 30(3), 1155-1177.

Iervolino I (2016) Soil-Invariant Seismic Hazard and Disaggregation. Bulletin of the Seismological Society of America 106(4), 1900-1907.

Itoi T, Takada T (2012) Inter-Period Correlation of Acceleration Response Spectra in Observed and Simulated Ground Motions. In: Proceedings of 15th World Conference on Earthquake Engineering, Lisbon, Portugal.

Jayaram N, Baker JW, Okano H, Ishida H, McCann Jr MW, Mihara Y (2011) Correlation of response spectral values in Japanese ground motions. Earthquake and Structures 2(4), 357-376.

Joyner WB, Boore DM (1981) Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake. Bulletin of the Seismological Society of America 71(6), 2011-2038.

Joyner WB, Boore DM (1993) Methods for regression analysis of strong-motion data. Bulletin of the Seismological Society of America 83(2), 469-487.

Kale Ö, Akkar S, Ansari A, Hamzehloo H (2015) A Ground-Motion Predictive Model for Iran and Turkey for Horizontal PGA, PGV, and 5% Damped Response Spectrum: Investigation of Possible Regional Effects. Bulletin of the Seismological Society of America 105(2A), 963-980.

Kassambara A, Mundt F (2016) Package 'factoextra': Extract and Visualize the Results of Multivariate Data Analyses.

Kokusho T, Sato K (2008) Surface-to-base amplification evaluated from KiK-net vertical array strong motion records. Soil Dynamics and Earthquake Engineering 28(9), 707-716.

Kotha SR, Bindi D, Cotton F (2016) Partially non-ergodic region specific GMPE for Europe and Middle-East. Bulletin of Earthquake Engineering 14(4), 1245-1263.

Kotha SR, Bindi D, Cotton F (2017) From ergodic to region- and site-specific Probabilistic Seismic Hazard Assessment: Method development and application at European and Middle-Eastern sites. Earthquake Spectra 33(4), 1433-1453.

Kotha SR, Bindi D, Cotton F (2017) Site-corrected magnitude and region dependent correlations of horizontal peak spectral amplitudes. Earthquake spectra 33(4), 1415-1432.

Kotha SR, Cotton F, Bindi D (2018) A new approach to site classification: Mixed-effects Ground-Motion Prediction Equation with spectral clustering of site amplification functions. Soil Dynamics and Earthquake Engineering.

Kotha SR, Cotton F, Bindi D (2018) Site Classification Derived From Spectral Clustering of Empirical Site Amplification Functions. In-proceedings of 16th European Conference on Earthquake Engineering, Thessaloniki, Greece, 2018.

Koufoudi E, Ktenidou O-J, Cotton F, Dufour F, Grange S (2015). Empirical ground-motion models adapted to the intensity measure ASA_{40}, Bulletin of Earthquake Engineering 13(12), 3625-3643.

Ktenidou O-J, Roumelioti Z, Abrahamson N, Cotton F, Pitilakis K, Hollender F (2017) Understanding single-station ground motion variability and uncertainty (sigma): lessons learnt from EUROSEISTEST. Bulletin of Earthquake Engineering, online first. doi: 10.1007/s10518-017-0098-6.

Lampros Mouselimis (2017) ClusterR: Gaussian Mixture Models, K-Means, Mini-Batch-Kmeans and K-Medoids Clustering (Version R package version 1.0.6). Retrieved from https://CRAN.R585 project.org/package=ClusterR

Landwehr N, Kuehn NM, Scheffer T, Abrahamson N (2016) A Non-ergodic Ground-Motion Model for California with Spatially Varying Coefficients. Bulletin of the Seismological Society of America 106(6), 2574-2583.

Lanzano G, Luzi L, Pacor F, Puglia R, D'Amico M, Felicetta C, Russo E (2016) Preliminary analysis of the accelerometric recordings of the August 24th, 2016 MW 6.0 Amatrice earthquake. Annals of Geophysics 59, Fast Track 5, DOI: 10.4401/ag-7201

Lanzano G, Puglia R, Russo E, Luzi L, Bindi D, Cotton F, D'Amico MC, Felicetta C, Pacor F, O. WG5 (2017) ESM strong-motion flat-file 2017.

Lee VW, Trifunac MD (2010) Should average shear-wave velocity in the top 30m of soil be used to describe seismic amplification? Soil Dynamics and Earthquake Engineering 30(11), 1250-1258.

Lin P-S, Chiou B, Abrahamson N, Walling M, Lee C-T, Cheng C-T (2011) Repeatable source, site, and path effects on the standard deviation for empirical ground-motion prediction models. Bulletin of the Seismological Society of America 101(5), 2281-2295.

Luzi L, Bindi D, Puglia R, Pacor F, Oth A (2014) Single-Station Sigma for Italian Strong-Motion Stations. Bulletin of the Seismological Society of America 104(1), 467-483.

Luzi L, Hailemikael S, Bindi D, Pacor F, Mele F, Sabetta F (2008) ITACA (ITalian ACcelerometric Archive): a web portal for the dissemination of Italian strong motion data. Seismological Research Letters 79(5), 716-722.

Luzi L, Puglia R, Pacor F, Gallipoli MR, Bindi D, Mucciarelli M (2011). Proposal for a soil classification based on parameters alternative or complementary to V_{S30}. Bulletin of Earthquake Engineering 9(6), 1877-1898.

MacQueen J (1967) Some methods for classification and analysis of multivariate observations. In: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. 1967. Oakland, CA, USA.

McNamara DE, Gee L, Benz HM, Chapman M (2014). Frequency-dependent seismic attenuation in the eastern United States as observed from the 2011 central Virginia earthquake and aftershock sequence: Bulletin of the Seismological Society of America 104(1), 55-72.

Okada Y, Kasahara K, Hori S, Obara K, Sekiguchi S, Fujiwara H, Yamamoto A (2004) Recent progress of seismic observation networks in Japan—Hi-net, F-net, K-NET and KiK-net—. Earth, Planets and Space 56(8), xv-xxviii.

Oth A, Bindi D, Parolai S, Di Giacomo D (2011) Spectral analysis of K-NET and KiK-net data in Japan, Part II: On attenuation characteristics, source spectra, and site response of borehole and surface stations. Bulletin of the Seismological Society of America 101(2), 667-687.

Oth A (2013) On the characteristics of earthquake stress release variations in Japan. Earth and Planetary Science Letters 377, 132-141.

Pitilakis K, Riga E, Anastasiadis A (2013) New code site classification, amplification factors and normalized response spectra based on a worldwide ground-motion database. Bulletin of Earthquake Engineering 11(4), 925-966.

Puglia R, Albarello D, Luzi L, Bindi D, Gallipoli MR, Mucciarelli M, Naso G, Pacor F, Peronace E (2015) On the Performances of Site Parameters for Soil Classification. In: Lollino G, Manconi A, Guzzetti F, Culshaw M, Bobrowsky P, Luino F (eds) Engineering Geology for Society and Territory, Volume 5, Springer, 1149-1152.

Team R.C (2013), R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Rathje EM, Kottke AR, Trent WL (2010) Influence of input motion and site property variabilities on seismic site response analysis. Journal of Geotechnical and Geoenvironmental Engineering 136(4), 607-619.

Rey J, Faccioli E, Bommer JJ (2002), Derivation of design soil coefficients (S) and response spectral shapes for Eurocode 8 using the European Strong-Motion Database, Journal of Seismology 6(4), 54-555

Rodriguez-Marek A, Cotton F, Abrahamson NA, Akkar S, Al Atik L, Edwards B, Montalva GA, Dawood HM (2013) A model for single-site standard deviation using data from various tectonic regions. Bulletin of the Seismological Society of America 103(6), 3149-3163.

Rodriguez-Marek A, Montalva GA, Cotton F, Bonilla F (2011) Analysis of single-site standard deviation using the KiK-net data. Bulletin of the Seismological Society of America 101(3), 1242-1258.

Rodriguez-Marek A, Rathje EM, Bommer JJ, Scherbaum F, Stafford PJ (2014) Application of single-site sigma and site-response characterization in a probabilistic seismic-hazard analysis for a new nuclear site. Bulletin of the Seismological Society of America 104(4), 1601-1619.

Royston P (1992) Approximating the Shapiro-Wilk W-Test for non-normality. Statistics and Computing 2(3), 117-119.

Scasserra G, Stewart JP, Kayen RE, Lanzo G (2009) Database for earthquake strong motion studies in Italy. Journal of Earthquake Engineering 13(6), 852-881.

Scherbaum F, Kuehn NM, Ohrnberger M, Koehler A (2010). Exploring the Proximity of Ground-Motion Models Using High-Dimensional Visualization Techniques. Earthquake Spectra 26(4), 1117-1138.

Scherbaum F, Schmedes J, Cotton F (2004) On the conversion of source-to-site distance measures for extended earthquake source models. Bulletin of the Seismological Society of America 94(3), 1053-1069.

Schmedes J, Archuleta RJ (2008) Near-source ground motion along strike-slip faults: Insights into magnitude saturation of PGV and PGA. Bulletin of the Seismological Society of America 98(5), 2278-2290.

Seyhan E, Stewart JP (2014) Semi-empirical nonlinear site amplification from NGA-West2 data and simulations. Earthquake Spectra 30(3), 1241-1256.

Shmueli G (2010) To explain or to predict? Statistical Science 25(3), 289-310.

Spudich P, Bayless JR, Baker JW, Chiou BS-J, Rowshandel B, Shahi SK, Somerville P (2013) Final Report of the NGA-West2 Directivity Working Group Rep. 2013/09, 129 pp, Pacific Earthquake Engineering Research Center, University of California, Berkley.

Stafford PJ, Strasser FO, Bommer JJ (2008) An evaluation of the applicability of the NGA models to ground-motion prediction in the Euro-Mediterranean region. Bulletin of Earthquake Engineering 6(2), 149-177.

Stafford PJ (2012) Evaluation of structural performance in the immediate aftermath of an earthquake: a case study of the 2011 Christchurch earthquake. International Journal of Forensic Engineering 1(1), 58-77.

Stafford PJ (2014) Crossed and Nested Mixed-Effects Approaches for Enhanced Model Development and Removal of the Ergodic Assumption in Empirical Ground-Motion Models. Bulletin of the Seismological Society of America 104(2), 702-719.

Stafford PJ, Rodriguez-Marek A, Edwards B, Kruiver PP, Bommer JJ (2017) Scenario Dependence of Linear Site-Effect Factors for Short-Period Response Spectral Ordinates. Bulletin of the Seismological Society of America 107(6), 2859-2872.

Stucchi M, Rovida A, Capera AG, Alexandre P, Camelbeeck T, Demircioglu M, Gasperini P, Kouskouna V, Musson R, Radulian M (2013) The SHARE European earthquake catalogue (SHEEC) 1000–1899. Journal of Seismology 17(2), 523-544.

Theodulidis N, Kalogeras I, Papazachos C, Karastathis V, Margaris B, Papaioannou C, Skarlatoudis A (2004) HEAD 1.0: a unified HEllenic accelerogram dataset. Seismological Research Letters 75(1), 36-45.

Thompson EM, Baise LG, Kayen RE, Guzina BB (2009) Impediments to predicting site response: Seismic property estimation and modeling simplifications. Bulletin of the Seismological Society of America 99(5), 2927-2949.

Thompson EM, Baise LG, Tanaka Y, Kayen RE (2012) A taxonomy of site response complexity. Soil Dynamics and Earthquake Engineering 41, 32-43.

Thompson EM, Wald DJ (2016) Uncertainty in V_{S30}-based site response. Bulletin of the Seismological Society of America 106(2), 453-463.

Tibshirani R, Walther G, Hastie T (2001) Estimating the number of clusters in a data set via the gap statistic. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 63(2), 411-423.

Ullah S, Bindi D, Pittore M, Pilz M, Orunbaev S, Moldobekov B, Parolai S (2013) Improving the spatial resolution of ground motion variability using earthquake and seismic noise data: The example of Bishkek (Kyrgyzstan). Bulletin of Earthquake Engineering 11(2), 385-399.

Villani M, Abrahamson NA (2015) Repeatable site and path effects on the ground-motion sigma based on empirical data from southern California and simulated waveforms from the CyberShake platform. Bulletin of the Seismological Society of America 105(5), 2681-2695.

Woessner J, Laurentiu D, Giardini D, Crowley H, Cotton F, Grünthal G, Valensise G, Arvidsson R, Basili R, Demircioglu MB, Hiemer S, Meletti C, Musson RMW, Rovida AN, Sesetyan K, Stucchi M (2015) The 2013 European Seismic Hazard Model: key components and results. Bulletin of Earthquake Engineering 13(12), 3553-3596.

Yenier E, Atkinson GM (2015) An equivalent point-source model for stochastic simulation of earthquake ground motions in California. Bulletin of the Seismological Society of America 105(3), 1435-1455.

Zhao JX, Irikura K, Zhang J, Fukushima Y, Somerville PG, Asano A, Ohno Y, Oouchi T, Takahashi T, Ogawa H (2006) An empirical site-classification method for strong-motion stations in Japan using H/V response spectral ratio. Bulletin of the Seismological Society of America 96(3), 914-925.

Zhao JX, Zhou S, Zhou J, Zhao C, Zhang H, Zhang Y, Gao P, Lan X, Rhoades D, Fukushima Y, Somerville PG, Irikura K (2016) Ground-motion prediction equations for shallow crustal and upper-mantle earthquakes in Japan using site class and simple geometric attenuation functions. Bulletin of the Seismological Society of America 106 (4), 1552-1569.