SITE-SPECIFIC SEISMIC HAZARD ANALYSES

By

GONZALO ANDRES MONTALVA

A dissertation submitted in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

WASHINGTON STATE UNIVERSITY
Department of Civil and Environmental Engineering

AUGUST 2010

To the Faculty of Washington State University:

The members of the Committee appointed to examine the dissertation of GONZALO ANDRES MONTALVA find it satisfactory and recommend that it be accepted.

Adrian Rodriguez-Marek,Ph.D., Chair

Balasingam Muhunthan,Ph.D.

Donald A. Bender,Ph.D.

Steven L. Kramer,Ph.D.

ACKNOWLEDGMENT

First I would like to thank my advisor, Dr. Adrian Rodriguez-Marek for his support, guidance, and friendship. His dedication throughout these years was beyond expectation. The countless and interesting discussions we had, helped shape not only this study, but my future as an engineer, professor, scientist, and person.

I would like to thank Dr. Balasingam Muhunthan, Dr. Donald Bender, and Dr. Steven Kramer for serving on my committee, and their helpful comments. Special thanks to Dr. Balasingam Muhunthan for the continuous support I found in him from my first day at Washington State University, and for all the encouragement he gave me to become a scientist.

I would like to thank my wife, Zulia, for all she has done to help me reach my goals. Without her support, love, help, and sacrifices, I would never have arrived at this point in my journey. My three beautiful children Gonzalo, Sofia, and Amanda, gave me the extra energy and inspiration I needed. I would also like to thank my parents for instilling in me the value of excellence, the desire to learn, and for giving me every opportunity.

My gratitude also goes to my colleagues in the geo-transportation group, especially Brenton, Kal, Senthil, Farid, Brian, Nick, Joe, Kathy, Muthu and Habtamu.

Financial support by the U.S. Geological Survey and CONICYT (Chilean National Commission for Scientific and Technological Research) through the "Becas de Doctorado en el Extanjero" program is greatly appreciated.

Site-Specific Seismic Hazard Analyses

Abstract

by Gonzalo Andres Montalva, Ph.D.

Washington State University

August, 2010

Chair: Dr. Adrian Rodriguez-Marek

Current 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 outcrop ground motion and associated variability, then the ground motion is propagated to the ground surface by site response analysis.

The site response process is inherently variable. Including this uncertainty in site response analyses without modifying the input ground motion uncertainty produces double counting of the uncertainty associated with site response. In this dissertation the total uncertainty is partitioned into its several contributing components, quantifying these components, and proposing methods to perform site-specific seismic hazard analyses without double counting uncertainties.

Four random field models were developed, and an existing one was fitted to a different database. These models can be used to generate shear-wave velocity profiles for site response analyses. Two types of models are presented, using Gaussian random fields, and using Markov Chains. The first ones showed better performance, and among those a stationary Gaussian model (stationary on ρ) showed the best performance, and it is the simplest among the five models.

Three GMPE's were developed, one only from surface records, one from "at-depth" records, and a third one combining surface and "at-depth" records. The results show the

same magnitude and distance scaling for the three equations. For stations that recorded a large number of records, total uncertainty was measured by the standard deviation of the observed minus predicted, and similarly for intra-event residuals. These statistics serve as lower bounds for site-specific seismic hazard analyses, note that these standard deviations are non-ergodic. The use of a GMPE capable of predicting bedrock and surface median ground motions, allows the partition of the components of the total uncertainty at the surface into those related to the bedrock ground motion and those to site response. These components and their correlations are presented.

Measurements at a site can potentially reduce the uncertainty in the ground motion prediction from that calculated using the ergodic assumption, to that observed at single sites. This implies a reduction on the order of 25%.

Contents

Abstract **iv**

Contents **vi**

List of Figures **xi**

Table Index **xxv**

1 Introduction and Objectives **1**

1.1 Introduction 1

1.2 Objectives 3

1.3 Organization of the Dissertation 4

2 Background **5**

2.1 Site Response 5

2.1.1 Typical Site Response Analyses 6

2.2 Factors Controlling Site Response 8

2.3 Regression Models Treatment of Site Response 8

2.4 Random Fields 9

2.5 Shear Wave Velocity Measurement Uncertainty 14

2.5.1 Measurement Specific Uncertainty 15

2.5.2 Inter-method Uncertainty 16

2.5.3 Shear-wave Velocity Profile generation 16

2.6 Profile Uncertainty 19

2.7 Available Databases 19

2.8 Stochastic Site Response 20

3 Framework for site-specific PSHA **25**

3.1 Methodology 26

3.1.1 Issues with the propagation of uncertainty 27

3.1.2 Proposals for site-specific intensity estimates 28

4 Uncertainty in Site Response Analyses **36**

4.1 Random Field Model for Shear Wave Velocity 37

4.1.1 Existing Methodologies 38

4.1.2 Spatial statistics of the KiK-net database 39

4.1.2.1 Soil Statistics and Correlation Structure 39

4.1.2.2 Rock Statistics 45

4.1.3 Proposed Models 46

4.1.3.1 Modified EPRI 46

4.1.3.2 One Layer Lag Stationary Gaussian Model 52

4.1.3.3 Two Layer Lag Stationary Gaussian Model 52

4.1.3.4 Markov Chain 53

4.1.3.5 Second Order Markov Chain 56

4.1.4 Statistical Comparison 57

4.1.5 Comparison in Terms of Site Response 58

4.1.5.1 Comparison of site response between a set of measured Vs profiles and a set of artificially generated Vs profiles 58

4.2 Monte Carlo Simulation Approach 66

4.3 Conclusions 78

5 Single Site Variability of Ground Motions: estimates from the KiKnet database 79

5.1 Data Source Characteristics 80

5.2 GMPE for Surface and Borehole 86

5.2.1 Combined Model 87
5.2.2 Surface Model 94
5.2.3 Borehole Model 98
5.2.4 Model Comparisons 100
5.2.5 Analysis of Residuals 105
5.2.5.1 Inter-Event Residuals 105
5.2.5.2 Intra-Event Residuals 116
5.2.6 Comparison of Inter and Intra-Event Residual Standard Deviations . 127
5.3 Single-Station Standard Deviations 139
5.3.1 Introduction 139
5.3.2 Analysis of Residuals at Single-Stations 140
5.3.3 Study of Cross-correlation between Residuals 160
5.3.4 Study of Single-Station Standard Deviations: path, magnitude, and distance effects 165
5.4 PSHA Example 175

6 Conclusions **181**

6.1 Summary 181
6.2 Significant Findings 182

6.3 Recommendations for Further Study 185

Bibliography **187**

A Additional Plots of Ground Motion Residuals **195**

A.1 Intra-Event Residuals 195

List of Figures

2.1 Schematics of site response phenomena 6

2.2 Typical spatial covariance shape. 10

2.3 Inherent soil variability. Reproduced from Phoon and Kulhawy (1999) . . . 14

3.1 Flowchart for estimation of site-specific surface ground motion intensity . . 30

3.2 Flowchart for estimation of site-specific surface ground motion intensity . . 35

4.1 Shear-wave velocity statistics and correlation coefficients for the entire data set and Vs30 based subsets. 4.1a entire database, 4.1b site classes A and B (Vs30 > 760 m/s), 4.1c site class C (760 m/s > Vs30 > 360 m/s), and 4.1d site class D (360 m/s > Vs30 > 180 m/s). 41

4.2 Comparison of the standard deviation of shear wave velocity for the entire database and for Vs30-based subsets. 42

4.3 Shear-wave velocity correlation coefficients considering only lag distances that are one layer away from each other. Initial depths of 10, 50, and 90 meters are shown. 43

4.4 Shear-wave velocity correlation coefficients considering only lag distances that are two layers away from each other. Initial depths of 10, 50, and 90 meters are shown. 44

4.5 p-values versus depth, for the hypothesis that the Vs values fit a log-normal distribution at a given depth. At 5% significance level, all p-values greater than 0.05 are considered to approach a log-normal distribution. 47

4.6 Rate of layer transitions versus depth for the generic case. This parameter, λ, can be understood as the reciprocal of layer thickness for each depth. . . 50

4.7 Flow chart of the generation procedure for the modified EPRI model. 51

4.8 Correlation between layer depth and shear-wave velocity for the entire KiK-net database. 55

4.9 Comparison of the mean and one standard deviation band for one set of simulated Vs profiles using the Modified EPRI methodology, and a set of measured Vs profiles for site C class. 58

4.10 Comparison of the mean and one standard deviation band for one set of simulated Vs profiles using the One layer lag Stationary Gaussian methodology, and a set of measured Vs profiles for site C class. 59

4.11 Comparison of the mean and one standard deviation band for one set of simulated Vs profiles using the Two layer lag Stationary Gaussian methodology, and a set of measured Vs profiles for site C class. 60

4.12 Comparison of the mean and one standard deviation band for one set of simulated Vs profiles using the first order Markov chain methodology, and a set of measured Vs profiles for site C class. 61

4.13 Comparison of the mean and one standard deviation band for one set of simulated Vs profiles using the Second order Markov chain methodology, and a set of measured Vs profiles for site C class. 62

4.14 Empirical correlation functions for depth 10 m, and site class C. Comparison of measured profiles with sets of simulated profiles using the five proposed methodologies. 63

4.15 Empirical correlation functions for depth 50 m, and site class C. Comparison of measured profiles with sets of simulated profiles using the five proposed methodologies. 63

4.16 Comparison of the median Ratio of Response Spectra for the different Random Field generators. Also shown is the predicted Ratio of Response spectra using the measured shear wave velocity profiles for site class C. 64

4.17 Comparison of the standard deviation of Ratio of Response Spectra for the different Random Field generators. Also shown is the predicted standard deviation of Ratio of Response spectra using the measured shear wave velocity profiles for site class C. 65

4.18 Mean of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 0.2 seconds and Site C class. 68

4.19 Mean of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 0.5 seconds and Site C class. 68

4.20 Mean of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 1.0 seconds and Site C class. 69

4.21 Standard deviation of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 0.2 seconds and Site C class. . . . 69

4.22 Standard deviation of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 0.5 seconds and Site C class. . . . 70

4.23 Standard deviation of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 1.0 seconds and Site C class. . . . 70

4.24 Mean of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 0.2 seconds and Site D class. 71

4.25 Mean of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 0.5 seconds and Site D class. 72

4.26 Mean of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 1.0 seconds and Site D class. 72

4.27 Standard deviation of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 0.2 seconds and Site D class. . . . 73

4.28 Standard deviation of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 0.5 seconds and Site D class. . . . 73

4.29 Standard deviation of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 1.0 seconds and Site D class. . . . 74

4.30 Standard deviations of the amplification factor for Site C, and spectral period of 0.3 seconds. Each point is the standard deviation of the amplification factor for 600 artificially generated profiles. 74

4.31 Standard deviations of the amplification factor for Site C, and spectral period of 1.0 second. Each point is the standard deviation of the amplification factor for 600 artificially generated profiles. 75

4.32 Standard deviations of the amplification factor for Site D, and spectral period of 0.3 seconds. Each point is the standard deviation of the amplification factor for 600 artificially generated profiles. 76

4.33 Standard deviations of the amplification factor for Site D, and spectral period of 1.0 second. Each point is the standard deviation of the amplification factor for 600 artificially generated profiles. 77

5.1 Magnitude versus Distance distribution for the KiK-net database 81

5.2 Distribution of upper 30 m time-averaged shear-wave velocity (Vs_{30}) for the KiK-net database 82

5.3 Instrument depth distribution for the KiK-net database. Note that most subsurface instruments are located at 100m or 200m below the surface. . . . 83

5.4 Borehole shear-wave velocity histogram. 83

5.5 KiK-net ground motion station locations. Shown in red are the 46 station for which more that 15 records are available. Stations in blue are considered in the regression analysis. 84

5.6 KiK-net database epicenters for the recorded earthquakes included in the GMPE. Note that each of these events were recorded by surface and at depth instruments. 85

5.7 Ground motion estimates attenuation with respect to distance for magnitudes 4, 5, and 6. Left column corresponds to the estimates for borehole, and right column to surface estimates; first row shows estimates for a spectral period of 0.05 (sec), second row for T = 0.3 (sec), and third row for T = 1.0 (sec). Estimated scenario corresponds to Vs_{30} of 760 (m/s), depth to Vs equal to 800 (m/s) of 60 (m), Vs at depth 100 (m) of 3000 (m/s) 102

5.8 Comparison of the magnitude terms for the Combined, Surface, and Borehole models. 103

5.9 Median spectrum estimates for magnitudes 4, 5, and 6. Left column corresponds to the estimates for borehole, and right column to surface estimates; first row shows estimates for a distance to the fault plane of 20 (km), second row for R_{RUP}= 50 (km), and third row for R_{RUP}= 100 (km). Estimated scenario corresponds to Vs_{30} of 760 (m/s), depth to Vs equal to 800 (m/s) of 60 (m), Vs at depth bedrock of 3000 (m/s) 104

5.10 Inter-Event Residuals for spectral period of 0.03 seconds and the models for Surface, Borehole, and Combined versus Magnitude. 106

5.11 Inter-Event Residuals for spectral period of 0.03 seconds and the models for Surface, Borehole, and Combined versus Event Depth. 107

5.12 Inter-Event Residuals for spectral period of 0.2 seconds and the models for Surface, Borehole, and Combined versus Magnitude. 108

5.13 Inter-Event Residuals for spectral period of 0.2 seconds and the models for Surface, Borehole, and Combined versus Event Depth. 109

5.14 Inter-Event Residuals for spectral period of 0.6 seconds and the models for Surface, Borehole, and Combined versus Magnitude. 110

5.15 Inter-Event Residuals for spectral period of 0.6 seconds and the models for Surface, Borehole, and Combined versus Event Depth. 111

5.16 Inter-Event Residuals for spectral period of 1.0 second and the models for Surface, Borehole, and Combined versus Magnitude. 112

5.17 Inter-Event Residuals for spectral period of 1.0 second and the models for Surface, Borehole, and Combined versus Event Depth. 113

5.18 Inter-Event Residuals for spectral period of 1.4 seconds and the models for Surface, Borehole, and Combined versus Magnitude. 114

5.19 Inter-Event Residuals for spectral period of 1.4 seconds and the models for Surface, Borehole, and Combined versus Event Depth. 115

5.20 Intra-Event Residuals for spectral period of 0.03 seconds and the models for Surface, Borehole, and Combined versus Magnitude. 117

5.21 Intra-Event Residuals for spectral period of 0.03 seconds and the models for Surface, Borehole, and Combined versus distance to the fault. 118

5.22 Intra-Event Residuals for spectral period of 0.2 seconds and the models for Surface, Borehole, and Combined versus Magnitude. 119

5.23 Intra-Event Residuals for spectral period of 0.2 seconds and the models for Surface, Borehole, and Combined versus distance to the fault. 120

5.24 Intra-Event Residuals for spectral period of 0.6 seconds and the models for Surface, Borehole, and Combined versus Magnitude. 121

5.25 Intra-Event Residuals for spectral period of 0.6 seconds and the models for Surface, Borehole, and Combined versus distance to the fault. 122

5.26 Intra-Event Residuals for spectral period of 1.0 second and the models for Surface, Borehole, and Combined versus Magnitude. 123

5.27 Intra-Event Residuals for spectral period of 1.0 second and the models for Surface, Borehole, and Combined versus distance to the fault. 124

5.28 Intra-Event Residuals for spectral period of 1.4 seconds and the models for Surface, Borehole, and Combined versus Magnitude. 125

5.29 Intra-Event Residuals for spectral period of 1.4 seconds and the models for Surface, Borehole, and Combined versus distance to the fault. 126

5.30 Surface intra- and inter-event standard deviations as a function of magnitude. Intra-event standard deviation presented in left column, inter-event standard deviation in right column. 128

5.31 Borehole intra- and inter-event standard deviations as a function of magnitude. Intra-event standard deviation presented in left column, inter-event standard deviation in right column. 129

5.32 Comparison of the intra-event standard deviations -ϕ- for the Combined, Surface, and Borehole models. 130

5.33 Comparison of the inter-event standard deviations -τ- for the Combined, Surface, and Borehole models. 131

5.34 Comparison of the inter-event standard deviations -τ- for the Combined, Surface, and Borehole models allowing for oversaturation in the median estimate. 132

5.35 Correlation between surface Inter- and Intra-event residuals, for a spectral period of 0.03 seconds. 133

5.36 Correlation between surface Inter- and Intra-event residuals, for a spectral period of 1.0 second. 134

5.37 Correlation between borehole Inter- and Intra-event residuals, for a spectral period of 0.03 seconds. 135

5.38 Correlation between borehole Inter- and Intra-event residuals, for a spectral period of 1.0 second. 136

5.39 Surface Inter-Event terms versus Borehole Inter-Event terms for peak ground accelerations 137

5.40 Surface and Borehole Inter-Event terms versus the Inter-Event terms obtained from the Combined model, for peak ground accelerations 138

5.41 Magnitude-Distance distribution for the subset of stations with more than 10 records. 140

5.42 Station parameters for subset of stations with more than 10 records. 141

5.43 Intra-event residuals at stations with more than 10 records. 143

5.44 Intra-event residuals corrected for site term versus magnitude, for peak ground acceleration. 144

5.45 Intra-event residuals corrected for site term versus magnitude, for a spectral period of 0.3 seconds. 145

5.46 Intra-event residuals corrected for site term versus magnitude, for a spectral period of 1.0 second. 146

5.47 Correlation between intra-event residuals (corrected for the site term) and inter-event residuals, for peak ground acceleration. 147

5.48 Correlation between intra-event residuals (corrected for the site term) and inter-event residuals, for a spectral period of 0.3 seconds. 148

5.49 Correlation between intra-event residuals (corrected for the site term) and inter-event residuals, for a spectral period of 1.0 second. 149

5.50 Single station standard deviations for stations with more than 10 records, for peak ground acceleration 151

5.51 Total standard deviations compared for single-sites and the ergodic prediction. 152

5.52 Mean of single-station residuals versus Vs_{30} and site period (To). Left column shows the mean residuals ($\delta S2S$) and right column shows its standard deviation (ϕ_{S2S}). Results for peak ground acceleration. 153

5.53 Mean of single-station residuals versus Vs_{30} and site period (To). Left column shows the mean residuals ($\delta S2S$) and right column shows its standard deviation (ϕ_{S2S}). Results for spectral acceleration of 0.3 seconds. 154

5.54 Mean of single-station residuals versus Vs_{30} and site period (To). Left column shows the mean residuals ($\delta S2S$) and right column shows its standard deviation (ϕ_{S2S}). Results for spectral acceleration of 1.0 second. 156

5.55 Intra-event single-station residuals a versus Vs_{30} for peak ground acceleration.157

5.56 Intra-event single-station residuals a versus Vs_{30} for a spectral period of 0.3 seconds. 157

5.57 Intra-event single-station residuals a versus Vs_{30} for a spectral period of 1.0 second. 158

5.58 Standard deviation of empirical amplification factor for peak ground acceleration. 158

5.59 Standard deviation of empirical amplification factor for a spectral period of 0.3 seconds. 159

5.60 Standard deviation of empirical amplification factor for a spectral period of 1.0 second. 159

5.61 Correlation coefficient between the different components of the total ground surface residual. Correlation Coefficients are plotted against spectral period. 163

5.62 Mean residual amplification ($\delta S2S^{AMP}$) versus Vs_{30} for peak ground acceleration. Plus and minus 1 standard deviations are added to illustrate dependency 166

5.63 Mean residual amplification ($\delta S2S^{AMP}$) versus Vs_{30} for spectral period of 0.3 seconds. Plus and minus 1 standard deviations are added to illustrate dependency 167

5.64 Mean residual amplification ($\delta S2S^{AMP}$) versus Vs_{30} for spectral period of 1.0 seconds. Plus and minus 1 standard deviations are added to illustrate dependency 168

5.65 Single station standard deviations for $T = 0.3$ versus standard deviation of station-to-event azimuths. These two quantities are positively correlated ($\rho = 34\%$). Only stations with more than 15 records are considered for this plot. 171

5.66 Single station standard deviations for $T = 0.3$ of records sampled from a varying range of station-to-event azimuth. Only station with more than 15 records are considered for this plot. 172

5.67 Single station standard deviations for different magnitude ranges for $T = 0.3$. Only station with more than 10 records are considered for this plot, and only those with 6 or more records are within each magnitude bracket are plotted. 173

5.68 Single station standard deviations for $T = 0.3$ and for different ranges of distance to the rupture plane. Only station with more than 10 records are considered for this plot. Mean single station standard deviation of these records is 0.48. 174

5.69 Shear-wave velocity profile for site KMMH09 176

5.70 Total residuals at site KMMH09, for spectral period of 0.1 seconds. 176

5.71 Ergodic Hazard Curve 177

5.72 Ergodic and Single-Site Hazard Curves 177

5.73 Site-specific hazard curve using borehole ground motion estimate and site response. 178

5.74 Hazard curve using site response and partially ergodic standard deviation . 179

5.75 Short return period portion of the hazard curves. Note the misfit of the site response generated curves with respect to the ergodic curve, this reflects the reduction of the bias in the ground motion predictions. 179

5.76 Hazard curves including nonlinear site response. 180

A.1 Intra-Event Residuals for spectral period of 0.03 seconds and the models for Surface, Borehole, and Combined versus Vs_{30}. 196

A.2 Intra-Event Residuals for spectral period of 0.03 seconds and the models for Surface, Borehole, and Combined versus predominant period (To). 197

A.3 Intra-Event Residuals for spectral period of 0.03 seconds and the models for Surface, Borehole, and Combined versus depth to shear-wave velocity of 800 (m/s). 198

A.4 Intra-Event Residuals for spectral period of 0.03 seconds and the models for Surface, Borehole, and Combined versus Magnitude, considering only events closer than 20 (km). 199

A.5 Intra-Event Residuals for spectral period of 0.2 seconds and the models for Surface, Borehole, and Combined versus Vs_{30}. 200

A.6 Intra-Event Residuals for spectral period of 0.2 seconds and the models for Surface, Borehole, and Combined versus predominant period (To). 201

A.7 Intra-Event Residuals for spectral period of 0.2 seconds and the models for Surface, Borehole, and Combined versus depth to shear-wave velocity of 800 (m/s). 202

A.8 Intra-Event Residuals for spectral period of 0.2 seconds and the models for Surface, Borehole, and Combined versus Magnitude, considering only events closer than 20 (km). 203

A.9 Intra-Event Residuals for spectral period of 0.6 seconds and the models for Surface, Borehole, and Combined versus Vs_{30}. 204

A.10 Intra-Event Residuals for spectral period of 0.6 seconds and the models for Surface, Borehole, and Combined versus predominant period (To). 205

A.11 Intra-Event Residuals for spectral period of 0.6 seconds and the models for Surface, Borehole, and Combined versus depth to shear-wave velocity of 800 (m/s). 206

A.12 Intra-Event Residuals for spectral period of 0.6 seconds and the models for Surface, Borehole, and Combined versus Magnitude, considering only events closer than 20 (km). 207

A.13 Intra-Event Residuals for spectral period of 1.0 second and the models for Surface, Borehole, and Combined versus Vs_{30}. 208

A.14 Intra-Event Residuals for spectral period of 1.0 second and the models for Surface, Borehole, and Combined versus predominant period (To). 209

A.15 Intra-Event Residuals for spectral period of 1.0 second and the models for Surface, Borehole, and Combined versus depth to shear-wave velocity of 800 (m/s). 210

A.16 Intra-Event Residuals for spectral period of 1.0 second and the models for Surface, Borehole, and Combined versus Magnitude, considering only events closer than 20 (km). 211

A.17 Intra-Event Residuals for spectral period of 1.4 seconds and the models for Surface, Borehole, and Combined versus Vs_{30}. 212

A.18 Intra-Event Residuals for spectral period of 1.4 seconds and the models for Surface, Borehole, and Combined versus predominant period (To) 213

A.19 Intra-Event Residuals for spectral period of 1.4 seconds and the models for Surface, Borehole, and Combined versus depth to shear-wave velocity of 800 (m/s). 214

A.20 Intra-Event Residuals for spectral period of 1.4 seconds and the models for Surface, Borehole, and Combined versus Magnitude, considering only events closer than 20 (km) 215

Table Index

2.1 Intra-Method Variability for Vs_{30} Measurement 15

4.1 Half-space Statistics 45

4.2 Parameters for Non-stationary Correlation Function 49

4.3 Coefficients for layering model, 95% confidence interval in parenthesis 49

4.4 Correlation Coefficient for Stationary Models 52

5.1 Model Parameters for the Distance and Magnitude terms of the Combined Model 91

5.2 Model Parameters for the Site terms of the Combined Model 92

5.3 Standard Deviations of the Residuals from the Combined Model 93

5.4 Model Parameters for the Distance and Magnitude terms of the Surface Model 95

5.5 Model Parameters Site term for the Surface Model 96

5.6 Standard Deviations of the Residuals from the Surface Model 97

5.7 Model Parameters for the Distance and Magnitude terms of the Borehole Model 99

5.8 Model Parameters for the Site terms of the Borehole Model 100

5.9 Standard Deviations of the Residuals from the Borehole Model 101

5.10 Single-Station Standard Deviations for Surface and Borehole 155

5.11 Single-Station Residuals correlation Coefficients 162

5.12 Standard Deviation of Random Variables that compose the total Surface Standard Deviation (ergodic) 164

5.13 Empirical check for Equation 5.25 165

5.14 Single-Station Standard Deviations for Surface and Bracketed Azimuths . . 169

5.15 Single-station standard deviations for PGA 170

To my wife, Zulia

Chapter 1

Introduction and Objectives

1.1 Introduction

Prediction of earthquake effects is a branch of science under constant development because of the great impacts that these events have on humanity. Of special concern are surface ground motion intensities, their frequency, amplitude, and if at all possible their recurrence. These predictions are possible but they involve a significant amount of uncertainty, due to the random nature of the seismic process, and due to our limited understanding of the phenomena.

A probabilistic framework, known as probabilistic seismic hazard analysis, is used to account for these uncertainties. One of the many components in a hazard analysis is the effect of the near surface materials. In current practice, these effects are introduced into an analysis with varying degrees of sophistication. In all cases, however, the uncertainties are usually not rigorously accounted for. A reduction of the overall uncertainty in the prediction of ground motions can be achieved by accounting only for the uncertainties that are possible on a specific site.

This study focuses on the effects that the random nature of the near surface soil parameters would have in seismic wave propagation, particularly on the waves traveling vertically through soil and shallow surface deposits, a problem known as site response. The importance of site response is widely acknowledged and can lead to an amplification of up to ten times in the ground response (Boore, 2004).

Measured values of soil parameters are known to be uncertain; even the theoretical concept of a true value for a soil parameter is hard to be formulated realistically. A more sensible view would accept that a measured value is only an average of the parameters within the sample, a sample of limited dimensions that is used to represent a much larger volume. Also, a comprehensive view needs to acknowledge that the measured value is also inherently uncertain from the testing point of view.

The evident randomness of the soil parameters and their measurement affects all areas of geotechnical engineering and has been studied extensively over the last 10 years (e.g. Phoon and Kulhawy, 1999; Fenton et al., 2005; Fenton and Griffiths, 2005; Kim and Santamarina, 2008). Despite the work of these authors and many others, there is a latent need to advance the state of the art in stochastic analysis in geotechnical engineering.

The interest to appropriately characterize and quantify site response concerns all those interested in surface ground motions. Seismologists studying seismic sources or their paths need to extract the effects of the surface materials. Engineers need to assess how the frequency content, amplitude, and duration of the surface ground motions are affected by the surface materials parameters.

This research attempts to better characterize the uncertainty involved in site response analyzes, to improve seismic hazard assessments. The physics of the site response problem are complex, however often modeled with a one-dimensional approach, and although the mechanism of one-dimensional wave propagation is widely accepted, our understanding of site response is still improving. This study will use existing site response models, and mea-

sured site response to improve on our understanding of the characterization of uncertainty in site response analyses.

1.2 Objectives

The main objective of this work is to advance the knowledge in the use and assessment of site response uncertainty. The focus is on site-specific studies, for which *stochastic* modeling of the soil profile along with empirical estimates of site response are studied. It is acknowledged that the problem is complex, however to provide useful tools as end result of this study, wave propagation analysis is assumed to be one-dimensional, hence all the consequences that result of this assumption (i.e. soil layers are infinite in the horizontal dimension, no surface waves are considered, and shear-waves are the mayor contributor to seismic shaking) are also carried along. Specific objectives are:

1. Assess the site response impact on the propagation of the uncertainty from *bedrock* level to the *soil surface*.

2. Propose a new methodology to generate random site profiles that accurately replicate a selected database or a part of it (i.e. a site class). Update currently used methodologies (i.e. EPRI, 1993) to fit a new data set.

3. For a limited set of vertical arrays from the Japanese KiK-net network (Kiban-Kyoshin, 2010), assess the effects in predicted uncertainty when the total error parameter in ground motion prediction equations ($GMPE$) is divided into: (a) inter-site, (b) intra-site, (c) site response terms, and (d) inherent variability, where the first two terms are divided into inter-event, and intra-event terms as well. Study the possible partition of these uncertainty terms into path, and region specific effects.

4. Proposed a methodology to incorporate the reduced uncertainty, into site-specific probabilistic seismic hazard analysis ($PSHA$).

1.3 Organization of the Dissertation

After the introduction to the problem and proposed objectives, the dissertation describes in chapter 2 the current understanding and treatment of *site response*, followed by a description of the factors that control the problem. This description encompasses empirical and analytical methods. Chapter 3 presents the framework to include uncertainty in site response into site-specific probabilistic seismic hazard analyses (PSHA). Chapter 4 covers the development of novel random field models to generate random soil profiles that maintain the spatial dependence of measured data. A popular existing model (EPRI, 1993) is modified and tested for comparison with the newly developed models. Monte Carlo simulations are used to show the applicability of these models. Chapter 5 presents surface and borehole $GMPE$ developed for the KiK-net database. Also, a $GMPE$ constrained with data combined from borehole and surface data is presented, the appropriateness of combining these two sets of data to better constrain the overall estimates is discussed. The developed $GMPE$ models are used to estimate residuals at stations with several earthquake recordings, obtaining estimates of single site ground motion uncertainty. Finally, in chapter 6 practical applications and the potential use in practice of the methods developed are discussed.

Chapter 2

Background

2.1 Site Response

Local site conditions have an important effect on seismic ground motions, and the estimation of seismic risk at a site must take this into account. These important effects have been recognized since the early work of Idriss and Seed (1970), but it was the occurrence of the earthquakes of Michoacan (Mexico, 1985), and Loma Prieta (California, 1989) that brought attention to the devastating effects that could result from site amplification, and hence the need to better understand the phenomena (Finn, 1991).

A complete ground motion analysis would include the modeling of the fault mechanism, the propagation of body waves from the source to the top of the *bedrock*, the generation and propagation of surface waves to the site of interest, and the propagation of body waves from the *bedrock* to the *soil surface*. In practice, it is still impossible to reliably predict the mechanism of fault rupture (that depends on fault size, the slip distribution, asperity of the interface, and direction of rupture among other parameters), and the complex wave propagation from the source to *bedrock* levels (that is affected by crustal wave velocities,

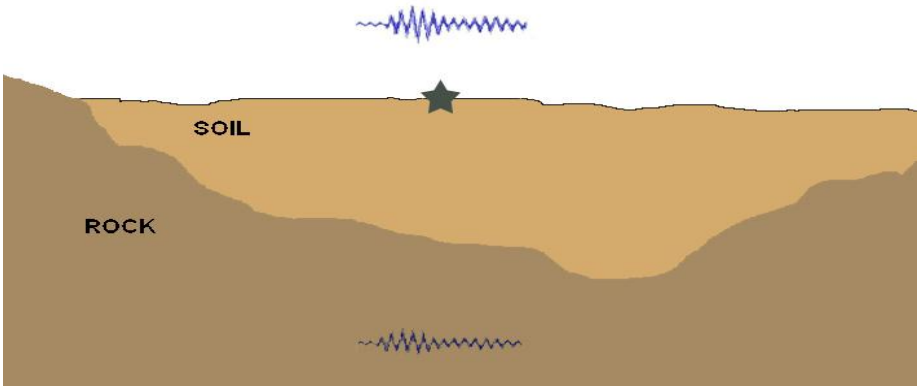


Figure 2.1. Schematics of site response phenomena

and wave scattering). Figure 2.1 illustrates a process that results in the surface ground motion, that is the modification of the bedrock motion by the near surface material, this process is referred to as site response.

The way the engineering community deals with this complex phenomenon is by using empirical methods based on observed data from previous earthquakes. Various regression models (e.g. Abrahamson et al., 2008) have been developed, these equations are based on recorded ground motions, and numerical modeling. They are then semi-empirical, and are parameterized by physical properties known to affect ground motions such as, distance to the rupture zone, magnitude, and shear-wave velocity of the near surface materials.

2.1.1 Typical Site Response Analyses

Site amplification is usually accounted for by an *amplification factor* defined as the ratio of a given ground motion intensity measure ($GMIM$), such as pseudo spectral acceleration (S_a), at the surface to that at the *bedrock* level. There are several definitions for site response (see Boore, 2004), among which a very important one is the ratio of a S_a at the *soil surface*, S_a^s to that at a rock *outcrop* reference site S_a^r.

There are two ways to compute site amplification factors, theoretically using wave propagation theory, and empirically by using data from recorded ground motions. The theoretical approach assumes that the ground response depends on the shear-wave velocity and density of the near surface materials. Empirical methods are developed by computing the ratios of recorded ground motions at two locations, generally ground surface, and bedrock or surface rock reference site.

The site amplification factor (AF) approach is generally used by site classification systems, that attempt to characterize site response with coefficients for sites with similar characteristics, this approach is employed by building codes or regulatory agencies to identify areas with similar ground motion levels. Generally, the number of site categories range from 3 to 6 main classes, and are based on shear wave-velocity, geotechnical data, or surface geology (e.g. Borcherdt, 1994; Rodriguez-Marek et al., 2001; Dobry et al., 2000; Stewart et al., 2003). One of the important conclusions of a number of verification studies that followed the work of Borcherdt (1994), and its subsequent modified adoption by NEHRP seismic design provisions, was that amplification factors have significant variability (Kramer and Stewart, 2004). This deterministic approach has been criticized by Goulet and Stewart (2009) for underestimating the surface ground motions computed probabilistically, that is the code nonlinear amplification factors are often lower than those obtained within a probabilistic seismic hazard analysis.

An alternative to using site classification is to perform site specific analyses. These type of analyses have the advantage of using site specific information, rather than relying in correlations such as the classification systems used by building codes (this is further discussed in 2.2), and in turn the site response results are not bounded to a coefficient but rather reflect the effect of the near surface material at all frequencies or times, depending on the type of analysis performed. Site specific analysis can use linear-equivalent (e.g. Schnabel et al., 1972) or fully nonlinear models (e.g. Park and Hashash, 2004; Reyes et al., 2009a,b), which is yet another advantage.

2.2 Factors Controlling Site Response

The assumption that a single dimension ($1D$) is sufficient to study site response has been confirmed (Kramer, 1996; Reyes et al., 2009b,a, to name some examples), and refuted (i.e. Thompson et al., 2009) in the literature, but it is by far the most popular approach. Under this assumption and using an equivalent linear approach (Schnabel et al., 1972), the phenomena is controlled by the soils shear-wave velocity ($\bar{V_s}$), density (ρ), damping, and their variations with depth and strain during the shaking. For the case of nonlinear analysis and the $1D$ assumption, the problem is controlled by the stress-strain behavior of the soil, note that the equivalent linear approach is an approximation to the nonlinear behavior.

Borcherdt (1994) introduced a classification system based on the upper 30 m time-averaged shear-wave velocity (Vs_{30}). This classification is the base for modern building codes treatment of site amplification, and showed that shear-wave velocity is the controlling factor in seismic site response.

2.3 Regression Models Treatment of Site Response

Ground motion prediction equations ($GMPE$) use as primary predictor variables the fault type, moment magnitude, $\mathbf{M}$, closest distance to the fault plane R_{rup} (alternatively the closest distance to the surface projection of the fault plane is also used), and Vs_{30}. Where Vs_{30} is used as proxy for near surface site effects. Eq. 2.1 is the basic functional form of $GMPE$, where y_{med} represents the median estimate of an intensity measure (i.e. S_a) in natural logarithm units, which is composed by a magnitude term, a distance term, and a site response term (see Eq. 2.2).

$$y = y_{med} + \epsilon \tag{2.1}$$

where, ϵ is a random variable that represents the error of the median prediction.

$$y_{med} = F_m(Mw) + F_d(R_{RUP}, Mw) + F_{site}(Vs_{30}) \tag{2.2}$$

The use of Vs_{30} as an index for site amplification, by $GMPE$ such as the *Next Generation of Ground-Motion Attenuation Models* (Abrahamson et al., 2008), has been questioned in the literature (e.g. Castellaro et al., 2008; Kokusho and Sato, 2008). However, it remains the most widely used proxy for site effects in $GMPE$.

2.4 Random Fields

The stochastic nature of geotechnical problems is widely acknowledged, yet often not considered. It refers not only to random material properties, but also to random loading (e.g. seismic ground motions), and random boundaries, a topic where there is little research done (Manolis, 2002). The nature of geo-materials necessarily implies a certain degree of spatial dependence.

The spatial dependence, or spatial structure, of soil properties is basically the relation between a random variable X_i, which corresponds to the value of the property at a given location (x_i, y_i, z_i), with one X_{i+h}, where X_{i+h} is the same property at a distance h. Random fields are a way to characterize the randomness at unsampled locations.

Each location i, or equivalently (x_i, y_i, z_i), represents one realization of the random field (or random process) X. If the same probability density function (*pdf*) can describe both X_i and X_{i+h}, and the joint density function $f_{X_i,X_{i+h}}$ is only a function of h, the random field is considered *stationary*. Stationarity is not a requirement for the use of random fields, but its use simplifies the analyses considerably. This stationarity condition is generally not met by natural data sets, hence a relaxation of this hypothesis is often adopted so that only

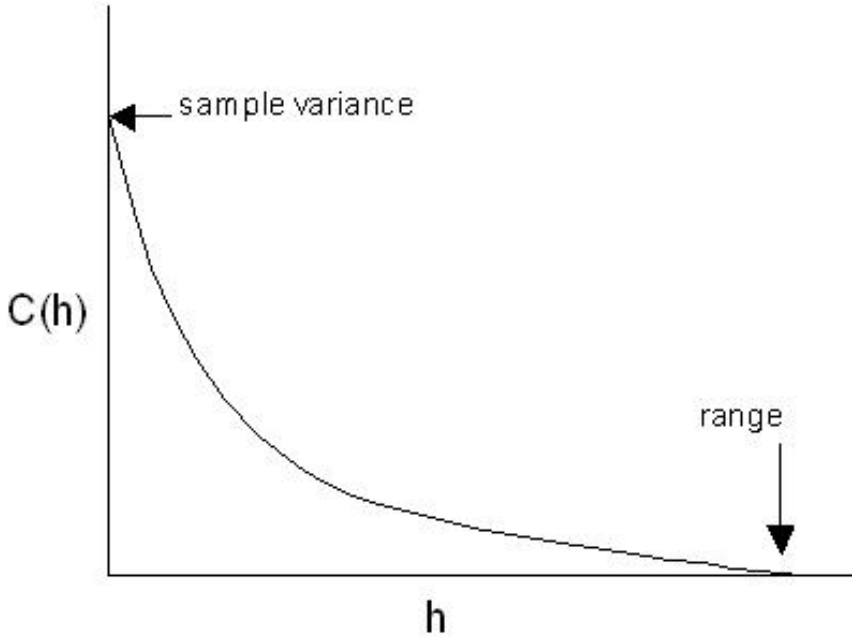


Figure 2.2. Typical spatial covariance shape.

the mean and variance are constant in space (the joint *pdf* is still independent of spatial location), this is called a *weak stationarity* condition.

Gaussian random fields are the most utilized in geotechnical engineering, largely as a result of their simple characterization. A *gaussian* random field is one whose joint *pdf* is a multivariate normally distributed random process. Under the stationarity assumption, they are completely defined by their mean, μ_X, and spatial covariance (see Equation 2.3 for general form, and Equation 2.4 for stationary form), which for $h = 0$ becomes the variance, σ_X^2. Figure 2.2 shows a typical covariance structure shape.

$$C_x(u_1, u_2) = \sigma(u_1)\sigma(u_2)\rho(u_1, u_2) \tag{2.3}$$

$$C_x(h) = E[(X_i - \bar{X}_i)(X_{i+h} - \bar{X}_{i+h})] = E[X_i X_{i+h}] - \mu_{X_i}\mu_{X_{i+h}} \tag{2.4}$$

$$\rho(h) = \frac{C_x(h)}{\sigma_x^2} \tag{2.5}$$

For the discrete case, in which only certain locations are measured, Equation 2.4 is estimated by Equation 2.6, where n_h is the number of pairs of data. There are several functions that are commonly fitted to the empirical data, usually a *correlation function* (Equation 2.5) is used to characterize how the field changes in space, this is done as a way to normalize the otherwise variance dependent values of the *spatial covariance* (see Equation 2.3). This property of the random fields can also be captured by the *spectral density function*, or the *variance function*.

$$\hat{C}_x(h) = \frac{1}{n_h} * \sum_{i=1}^{n_h} (X_i * X_{i+h}) - \bar{X}_i * \bar{X}_{i+h} \tag{2.6}$$

The *spectral density function* is shown in Equation 2.9 (see also Equations 2.7 and 2.8 leading to 2.9), where the random field $X(i)$ is first expressed as a sum of sinusoids, $S(w)$ is the two-sided spectral density function which is the inverse Fourier transform of the *spatial covariance*, also referred as *autocovariance* or *autocorrelation*, of the random field.

$$X(i) = \mu_X + \sum_{k=-N}^{N} (A_k * cos(w_k * i) + B_k * sin(w_k * i)) \tag{2.7}$$

$$C(h) = Cov[X(0), X(h)] = \int_{-\infty}^{\infty} S(w) * cos(wh)dw = \int_{0}^{\infty} G(w) * cos(wh)dw \tag{2.8}$$

$$G(w) = \frac{2}{\pi} \int_{0}^{\infty} C(h) * cos(wh)dh \tag{2.9}$$

The *variance function*, shown in Equation 2.10, gives the reduction in the variance when the random field $X(i)$ is averaged over the length H.

$$\gamma(H) = \frac{1}{H^2} \int_0^H \int_0^H \rho_X * (\xi - \eta) d\xi d\eta \tag{2.10}$$

Among the most popular correlation functions are the Markov (Eq. 2.11), and the Gaussian (Eq. 2.12) models.

$$\rho_X(h) = exp[-2h/c] \tag{2.11}$$

$$\rho_X(h) = exp[-\pi \left(\frac{h}{c}\right)^2] \tag{2.12}$$

Geostatistics

Geostatistics is a branch of applied statistics that focuses on the characterization of spatial dependence in properties that vary in value over space and the use of that dependence to predict values at unsampled locations. The notion of spatial dependence implies that two data values from nearby locations will be more alike than two values from distant locations. The technique most commonly used in geostatistics is called "kriging" (Pawlowsky-Glahn and Olea, 2004), which is used to estimate a random field between measured data.

The idea behind kriging is to estimate the random variable X_i at any location i using a weighted linear combination of the observations. This procedure allows for the appropriate treatment of clusters of data with high or low values, of the attribute of interest (i.e. soil property) with respect to the trend of the entire data set, without biasing the results for the entire data set.

One Dimensional Random Fields in Geotechnical Engineering

Phoon and Kulhawy (1999) studied a breakdown of geotechnical uncertainty into inherent soil variability, measurement error, and transformation uncertainty. The last two corresponding to epistemic uncertainty, and the first one modeled by *random fields*. Figure 2.3 shows the basic modeling of a soil property as a function of depth, where $\xi(z)$ is a soil property at depth z, $t(z)$ is the trend of the property, and $w(z)$ is the residual. The *scale of fluctuation* (Equation 2.13), also referred as *correlation length*, is an alternative to model spatial dependence mathematically defined by the area under the correlation function (see Equation 2.13), and it represents the distance within which points are strongly correlated. Equation 2.14 from Phoon and Kulhawy (1999) divides the vertically varying $\xi(z)$ into a trend $t(z)$, and $w(z)$, where the residuals $w(z)$ can be modeled as a *random field*. The advantage of removing the trend is that in the "detrended" model, the residuals, may acquire some stationary characteristics.

$$\theta = \int_{-\infty}^{\infty} \rho(h) \tag{2.13}$$

$$\xi(z) = t(z) + w(z) \tag{2.14}$$

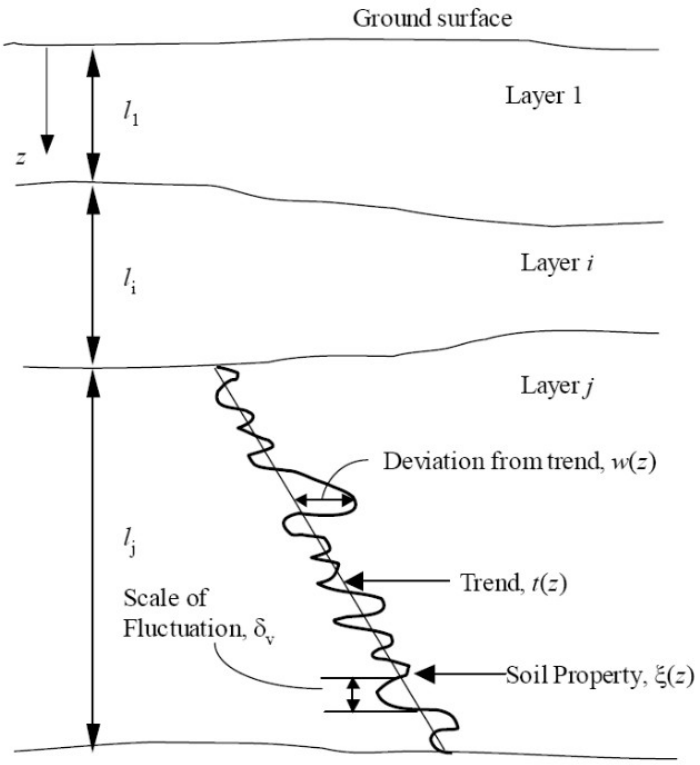


Figure 2.3. Inherent soil variability. Reproduced from Phoon and Kulhawy (1999)

2.5 Shear Wave Velocity Measurement Uncertainty

Uncertainty in shear wave velocity comes from two sources, uncertainty associated to the measurements (inter and intra-method), and inherent soil variability. The first one corresponds to epistemic uncertainty, while the second one is aleatoric. Another way to explain these two types of uncertainty is that the epistemic uncertainty is due to lack of knowledge, and thus can be reduce by a better understanding of the phenomena, in this case by better sampling techniques. The second type, aleatoric uncertainty, cannot be reduced as it corresponds to the random nature of the problem, it is also the consequence of mapping spatial

variability to point variability (*ergodicity assumption*). Although this division is widely accepted, it is arguable whether this is an aleatoric uncertainty or what is considered as such is just the result of physical processes that we are unable to account for.

Time-averaged shear-wave velocity of the upper 30 m (V_{s30}) is commonly used as an index for site period (Dobry et al., 1976), and hence for the dynamic behavior of the near surface soil. This index is only a simplified approach to parameterize site effects but its popularity has many advantages among which is the availability of data to quantify measurement errors, an essential component of the overall site response uncertainty.

2.5.1 Measurement Specific Uncertainty

Among the different methods used for shear wave velocity measurement, SASW (spectral analysis of surface waves), MASW (multi-channel analysis of surface waves), suspension logging (P-S suspension logging), and the seismic cone (SCPT) are the most commonly used in practice along with geologic-based estimates (Moss, 2008). Table 2.1 summarizes V_{s30} coefficient of variation for each of these methods.

Table 2.1. Intra-Method Variability for Vs_{30} Measurement

Method	Coefficient of Variation
MASW	1-4%[1]
SASW	5-10%[2]; 6%[3]
P-S logging and SCPT	1-3%[1,4]
Geologic Estimates	20-35%[5]

[1] Asten and Boore (2005)

[2] Marosi and Hiltunen (2004)

[3] Martin and Diehl (2004)

[4] Brown et al. (2002)

[5] Moss (2008)

SASW is different from the rest in that the measurements of phase angle and phase velocity have a COV of only about 2%. According to Marosi and Hiltunen (2004) the difference be-

tween Vs_{30} variability and the little measurement variability occurrs because the inversion process magnifies the uncertainty in the dispersion data, the epistemic uncertainty here could be separated into measurement and analysis. Geologic Estimates is, as expected, the method with greater uncertainty but it is also singular in that the $\sigma_{Vs_{30}}$ estimates do not only increase with the μ_{Vs} (i.e. maintaining a constant COV) but the COV increases linearly with V_{s30} estimates (Moss, 2008).

2.5.2 Inter-method Uncertainty

Two of the most popular methods are suspension logging and SASW. Brown et al. (2002) report a variability in site amplification factors $AF(f)$ for most frequencies, predicted by both methods, of about 15%. The same value was obtained for the comparison of SASW with downhole measurement.

Asten and Boore (2005) compared nine methods to suspension logging, obtaining differences in the value of V_{s30} of up to 20% for all the methods with exception of horizontal to vertical spectral ratio (HVSR) where the differences ranged from 30 to 60%. Other studies on inter-method variability are sumarized by Moss (2008) showing the same magnitud of variability.

The variability of invasive versus noninvasive methods could be explained in part due to soil disturbance, if this was the case it could be argued that the invasive methods are biased. This biasing would be attributed to strain softening is soft soils, and to strain hardening in stiffer soils (Moss, 2008). In the absence of significant quantity of data and a study that correlates them, this hypothesis is still to be confirmed.

2.5.3 Shear-wave Velocity Profile generation

Based on the statistics observed in data from 557 sites Toro (1995) used a gaussian random field to generate artificial shear-wave velocity profiles. His method proposes an first-order

auto regressive (also called Markovian) model to generate profiles.

The velocity model operates with a normalized quantity, Z_i (Equation 2.15), that assumes the Vs data is appropriately characterized by a log-normal distribution.

$$Z_i = \frac{ln(V_i) + ln[V_{median}(h_i)]}{\sigma_{lnVs}} \tag{2.15}$$

The log-normal velocity distribution and the correlation structure is characterized by a first-order auto-regressive model as in Equations 2.16, and 2.17; where ρ is the autocorrelation coefficient of Z, and ϵ_i are standard normal independent random variables.

$$Z_1 = \epsilon_1 \tag{2.16}$$

$$Z_i = \rho * Z_{i-1} + \sqrt{1 - \rho^2} * \epsilon_i \tag{2.17}$$

As stated by Toro (1995) " ideally, one should remove all bedrock velocities from the data prior to the analysis", that is on order to avoid mixing soil Vs with bedrock Vs. To this end, a maximum likelihood estimation (MLE) method is used to estimate model parameters.

The depth to bedrock and it's corresponding shear-wave velocity is generated separately from the shear-wave velocity profile on soil. Depth to bedrock is assumed to have a uniform distribution, and the bedrock shear-wave velocity distribution is estimated by a log-normal distribution with median = 1020 m/s and $\sigma_{lnVs} = 0.30$. The proportion of bedrock velocities p_R in the database is an unknown that is estimated as part of the MLE.

The likelihood function for layers i to n is given by Equation 2.18.

$$L_i = (1 - p_R) * f_S(lnV_i | lnV_1, lnV_2, \ldots lnV_{i-1}) + p_R * f_R(lnV_i) \tag{2.18}$$

where f_S is the probability density function of soil log-velocities, and f_R is the probability density function of bedrock log-velocities. The first-order model produced the best fit to the data, for this case f_S is given as:

$$f_S(lnV_i|lnV_{i-1}) = \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_{lnV}} exp\left[-\frac{1}{2}\left(\frac{Z_i - \rho * Z_{i-1}}{\sqrt{1-\rho^2}}\right)^2\right] \tag{2.19}$$

In Toro's analyses, information (i.e. shear-wave velocity) for depths with less than five profiles is not considered in the calculations.

The correlation coefficient, ρ, used by Toro depends on depth and the distance between the mid points of two consecutive layers, thus making the model non-stationary. Equation 2.20 shows the expression used, where $\rho_d(z)$ (see Equation 2.21) is a depth-dependent correlation, and $\rho_h(\Delta z)$ (see Equation 2.22) is a thickness-dependent correlation

$$\rho(z, \Delta z) = (1 - \rho_d(z)) * \rho_h(\Delta z) + \rho_d(z) \tag{2.20}$$

$$\rho_d(z) = \begin{cases} \rho_{200}\left[\frac{z+z_0}{200+z_0}\right]^b & \text{for } z \leq 200 \\ \rho_{200} & \text{for } z > 200 \end{cases} \tag{2.21}$$

$$\rho_h(\Delta z) = \rho_0 * exp(-\frac{\Delta z}{\Delta}) \tag{2.22}$$

In Equations 2.20, 2.21, and 2.22 z is the average of the midpoints depths of two consecutive layers, and h is the difference between these depths. ρ_{200}, z_0, b, ρ_0, and Δ are parameters to be calculated by the MLE procedure. Δz is the difference between two depths being considered and is equivalent to h in equations in section 2.4.

2.6 Profile Uncertainty

The layering system defined after shear-wave velocity testing is performed in different ways and therefore it has certain level of uncertainty. Toro (1995) proposed a model for artificial layering, which can be used to account for the testing uncertainty, the model generates layers from a non-homogeneous Poisson process with parameter $\lambda(z)$ being a function of depth z measured from the ground surface. The functional form proposed by Toro is shown in Equation 2.23, where a, b, and c are coefficients were fitted to match the observed layering system at 557 California sites.

$$\lambda(z) = a * (b + z)^{-c} \tag{2.23}$$

A maximum likelihood methodology is used by Toro (1995) to estimate the coefficients a, b, and c in equation 2.23. The calculated results, for the California database are 1.98, 10.86, and 0.89 for a, b, and c respectively.

Boore and Thompson (2007) studied the sensitivity of the process of defining the number and position of layers by a number methods, including different manual methods, and automatic picking. They concluded that for periods over 0.2 seconds site amplification was negligibly affected by the layering of the shear-wave model.

2.7 Available Databases

There are three databases that have significant amount of shear-wave velocity profiles, the Kiban-Kyoshin network (KiK-net) in Japan available at http://www.kik.bosai.go.jp/kik/ (last accessed October 2009), a compilation of shear-wave profiles primarily from California compiled by Dr. Walter Silva from Pacific Engineering and Analysis (Wills and Silva, 1998),

and a public database available through Dr. David Boore's web site http://quake.usgs.gov/ boore/.

The KiK-net data base consists of 629 vertical arrays, where each consists of a 3D accelerometer at surface level and at a base layer (Kokusho and Sato, 2008). Each of these profiles have an associated shear-wave velocity profile obtained by P-S suspension logging. This is the only database that has all measurements done by the most precise (note that this does not necessarily imply accuracy) method used in practice, and that also allows for the computation of empirical site response through a ratio of surface to borehole ground motion data. Both these attributes are important strengths of this database when compared to others.

The PE&A database consists of 1081 shear-wave velocity profiles measured by a variety of methods, where the most significant was downhole measurement, but including cross-hole, P-S suspension logging, geology estimations, and CXW (early form of surface wave dispersion). The strength of this database is the large number of profiles, and hence its statistical value. Two disadvantages of it is the large number of methods, and the inclusion of methods with large uncertainties (see 2.5.1), these two characteristics anticipate a larger uncertainty in the shear-wave velocity profiles than that from other databases.

The USGS database consists of 278 shear-wave velocity profiles measured using downhole type measurement. This database is part of the compilation made by PE&A. The fact that a single measurement technique was used and the significant number of sites, makes this database worth analyzing separately.

2.8 Stochastic Site Response

The incorporation of local site conditions into the estimation of seismic risk is achieved by means of a frequency -and amplitude (see Borcherdt, 2002)- dependent site amplification factor given by Equation 2.24, where $S_a^s(f)$ is the spectral acceleration at the soil surface,

$S_a^r(f)$ is the spectral acceleration in rock, and f is the frequency.

$$AF = \frac{S_a^s(f)}{S_a^r(f)} \tag{2.24}$$

The assessment of AF are done using empirical (Borcherdt, 1994) or theoretical (Schnabel et al., 1972) methods, this work uses both.

Site response affects all characteristics of a ground motion, but among the most significant for the engineering profession is the spectral acceleration. This will be the $GMIM$ of main interest throughout this study. In addition to the amplification factor described, the uncertainty associated with it (or with its product, the spectral acceleration of the soil motion) is also needed for risk analyses. Often, this uncertainty is modeled by the the standard deviation of the logarithm of the spectral acceleration $\sigma_{lnS_a^s}(f)$ (see Abrahamson et al., 2008) because of the assumption that S_a values distribute log-normally.

The site response problem can then be defined as the computation of both $S_a^s(f)$ and $\sigma_{lnS_a^s}(f)$ for given values of $S_a^r(f)$ and $\sigma_{lnS_a^r}(f)$ (Bazzurro and Cornell, 2004a). The inputs to a site response analysis are the soil profile (i.e. layering), the soil properties, and the input ground motion in terms of its spectral acceleration. It is reasonable to assume that both the median values and the uncertainties in the input parameters have an influence on the ground motion at the soil surface.

The variation of the soils shear strength and damping due to cyclic loading has an important effect on the site amplification at a site. Darendeli (2001) used dynamic laboratory tests on "intact" soil samples from 20 sites to develop mean values for both normalized modulus reduction curves, and material damping curves along with their associated uncertainty. The ability to handle uncertainty within this curves, allows for their proper use in probabilistic seismic hazard analysis.

The implementation of site effects into $PSHA$ has been formalized by Cramer (2003), and can be performed by the procedures presented by Bazzurro and Cornell (2004b). Bazzurro and Cornell (2004a) showed that a relatively small number of input motions are needed to obtain stable estimates of the mean and standard deviation of the amplification factor. The problem with this approach is that when it is used in PSHA to obtain hazard curves for soils, the exceedance rates are unknown, inconsistent across frequencies, and generally non-conservative (Bazzurro and Cornell, 2004b). In summary, the application of deterministic site response analyses superposed on ground motion predictions on rock results in ground motion estimates that lack an appropriate statistic characterization.

Bazzurro and Cornell (2004b,a) proposed a methodology for obtaining AF that are applicable to seismic hazard analyses. Bazzurro and Cornell's methodology provides a mapping between $S_a^s(f)$, $S_a^r(f)$, $\sigma_{lnS_a^s}(f)$, and $\sigma_{lnS_a^r}(f)$ that is statistically rigorous based on a mixture of site response analyses and statistical regression.

Bazzurro and Cornell (2004b) proposed two alternatives for the AF and $PSHA$ integration. The first method is to convolve the AF, obtained from a site response analysis, with a site hazard curve for a reference site (i.e. rock outcrop or bedrock). The second procedure modifies the statistical moments of the estimated ground motion intensity at a rock level to account for site effects. The Bazurro and Cornell methodology constitutes the bases of the methodology proposed in this study and is discussed further in Chapter 3. Baturay and Stewart (2003) present an alternative *empirical* methodology to integrate site response and hazard analyses.

Monte Carlo Simulation

Monte Carlo simulations (MCS) are used in many fields, and although computation intensive are considered to provide the exact solutions for problems involving randomness of its parameters. These methods were considered computationally too expensive in the past, and depending on the application (e.g. including site response in each iteration of a

probabilistic seismic hazard analysis) they still are, but the increasing processing capacity of modern computers is making them more readily available.

In site response analysis, the use of MCS has been used to study the effects of stochastic soil and input ground motion properties on surface ground motions (Rodriguez-Marek et al., 2010). Rahman and Yeh (1999) used a stochastic finite element method to conduct a parametric study on the effects of stochastic ground motion and soil stiffness on site response, concluding that ground motion variability has a greater effect, the same conclusion has been reached in other studies (see Bazzurro and Cornell, 2004a).

Recently, Douglas et al. (2009) presented a method, based on MCS, to use available information on the site of interest to reduce the uncertainty associated to site response, the method assumes that all conceivable variables that affect site response are independent from each other. These variables include geological information, standard penetration test (SPT), profile depth (to bedrock), crustal structure, near surface V_s, and topography. While the idea of reducing uncertainty is very important (the same motivation for the proposed study), the independence assumption may not be warranted and could lead to errors.

Random Vibration Theory

Random vibration theory (RVT) has also been used to perform site response analysis. Initially used in seismology to predict ground motion parameters as a function of source distance and magnitude (e.g. Boore, 1983), the method uses a Fourier amplitude spectrum (FAS) to characterize a ground motion, and uses RVT compute response spectra, PGA, and PGV. The inclusion of site response (Rathje and Ozbey, 2006) implies that the FAS corresponding to bedrock is modified by the near surface soil resulting in a surface FAS from where the same ground motion parameters can be obtained.

The RVT method has the advantage of producing median response spectra with a single

analysis, and if the input FAS is appropriately selected this results match the median value of a Monte Carlo simulation. One of the limitations of the RVT site response procedure is that no time history is produced. The advantage of obtaining the median response spectra in one step is also viewed as a disadvantage when modeling uncertainty due to the near surface soil, because of its ability to compute site response for the equivalent of a set of input ground motions in one step, the uncertainty associated solely to the soil variability, and its nonlinear behavior, the surface uncertainty in response spectra would be artificially reduced.

Chapter 3

Framework for site-specific PSHA

In current practice, structures are designed using pseudo-acceleration spectra, $S_a(f)$, as an index for their seismic demand. These spectra are intended to have equal probability of exceedance over a set time interval at all periods. This is called a uniform hazard spectrum (UHS).

However, code-prescribed design spectra are provided only for Vs_{30}-based site classes, which are obtained by scaling a uniform hazard spectrum for a reference site by an amplification factor. The resulting spectra loses the uniform hazard characteristics. An alternative approach is to scale a uniform hazard spectrum for a reference site using site response analyses. Various authors have proposed methodologies to rigorously incorporate site response in a $PSHA$ such that the uncertainty is properly propagated from the input to the output ground motion, and hence the resulting spectra maintains its uniform hazard characteristics (e.g. Cramer, 2003; Bazzurro and Cornell, 2004b).

The uncertainty associated with surface ground motion intensity estimates, and hence with the propagation of the uncertainty in ground motion intensities from a reference site to a specific site of interest, is of great concern within the framework $PSHA$.

This chapter describes the framework for two alternative methodologies to compute site-specific spectra that incorporate measurement uncertainty.

3.1 Methodology

The Bazzurro and Cornell (2004a,b) framework for including site response into $PSHA$ is herein presented and used throughout this study to introduce non-linear site response effects into site-specific hazard assessment. The amplification factor, AF was given in Equation 2.24, is mapped as a function of $S_a^r(f)$ in a log-log space as:

$$lnAF(f) \approx c_0 + c_1 * (\overline{lnS_a^r(f)} + \epsilon_{lnS_a^r(f)} * \sigma_{lnS_a^r(f)}) + \epsilon_{lnAF(f)} * \sigma_{lnAF(f)} \tag{3.1}$$

where c_0, and c_1 are model coefficients, $\overline{lnS_a^r(f)}$ is the median value of the logarithm of the spectral acceleration on a bedrock site (Bazzurro and Cornell, 2004b, assumed this to be reference bedrock site, in this dissertation a borehole-bedrock will be used instead), $\epsilon_{lnS_a^r(f)}$ and $\epsilon_{lnAF(f)}$ are standard normal variables (e.g. $\epsilon_{lnS_a^r(f)}$ $N(0,1)$), $\sigma_{lnS_a^r(f)}$ is the standard deviation of the logarithm of the spectral acceleration in rock, and $\sigma_{lnAF(f)}$ is the standard deviation of the AF. However, as will be discussed in Chapter 5, there are alternative ways to compute the $\sigma_{lnS_a^r(f)}$ value that better account for input uncertainty in site-specific analyses.

The model coefficients c_0, and c_1 are obtained from a fit in log-log space of results of site response analyses (Bazzurro and Cornell, 2004a). The methodology is not constrained to any single site amplification prediction code.

Combining Equations 2.24 and 3.1, the median estimate of the log-spectral acceleration in soil, $S_a^s(f)$, can be estimated as:

$$\overline{lnS_a^s(f)} \approx c_0 + (c_1 + 1) * \overline{lnS_a^r(f)} \tag{3.2}$$

The standard deviation of $S_a^s(f)$ is expressed as:

$$\sigma_{lnS_a^s(f)} \approx \sqrt{(c_1 + 1)^2 * \sigma^2_{lnS_a^r(f)} + \sigma^2_{lnAF(f)}} \tag{3.3}$$

In the next section, some of the assumptions inherent to this approach are discussed.

3.1.1 Issues with the propagation of uncertainty

Bazzurro and Cornell's methodology to propagate uncertainty from a rock site to a soil site, accounting for the site amplification effect and its associated uncertainty, assume that the standard deviation of the logarithm of the amplification factor, $\sigma_{lnAF(f)}$, is independent of both the logarithm of the rock pseudo-acceleration, $lnS_a^r(f)$, and its standard deviation, $\sigma_{lnS_a^r(f)}$. Both these assumptions will be tested in Chapters 4 and 5.

Estimation of $lnS_a^r(f)$, and $\sigma_{lnS_a^r(f)}$ from currently available $GMPE$ (e.g. Boore and Atkinson, 2008) is not appropriate for site specific cases in which Equation 3.3 is applied, because in such case the site-to-site variability embedded in the estimation of $lnS_a^r(f)$ would be erroneously convolved with the uncertainty in the AF. Most $GMPE$ are the result of regressing data from earthquakes that occurred in a wide range of locations and recorded at various sites. By using these equations in seismic hazard analyses we are assuming that the variability observed in the combined data set (parameterized as described in Section 2.3) is the same as that in a single site, this is known as the *ergodic assumption* (Anderson and Brune, 1999). As shown in Chapter 5, the variability of intensity estimates for site-specific cases are much lower than those invoking the *ergodic assumption*.

3.1.2 Proposals for site-specific intensity estimates

The discussion above evidences the need for different methodologies for site-specific ground motion intensity estimates. Two alternatives are herein proposed to this end.

Method 1 The first proposal for site-specific estimation of ground motion intensities is shown in Figure 3.1. The first step of the proposed method, makes use of the $GMPE$ developed in Chapter 5. Note that the $GMPE$ used to select $lnS_a^r(f)$ and $\sigma_{lnS_a^r(f)}$ is for a single site, that is the median model is parameterized as described in Section 2.3 *and* includes the repeatable site effect, which reduces the value of $\sigma_{lnS_a^r(f)}$ and eliminates part of the bias from the median estimate.

In order to be able to estimate a repeatable site effect, an instrument must be located at a bedrock site over a sufficiently long period of time such that it would permit a stable estimate of both the median value of the ground motion estimated at the site, and its standard deviation (that is, the single-station standard deviation).

For the case where the available recordings are on bedrock and the hazard analysis is being conducted on soil deposit, site amplification analyses must be conducted. The method of Bazurro and Cornell can be used to capture nonlinearity. The selection of compatible earthquake records for the site response analyses can be achieved by a number of available methods (e.g. Kottke and Rathje, 2008; Baker and Cornell, 2006), and typically involves a suite of about 7 to 10 records (Rathje et al., 2010), the selected records are linearly scaled and/or spectrally matched, the objective is to reproduce the target spectrum median and standard deviation.

Step three performs the site response analyses. The methodology of Bazurro and Cornell can be used to introduce nonlinear effects. The random field models presented in Chapter 4 can be used to introduced uncertainty in the measured shear wave velocity profiles. The

uncertainty has to be compatible with the anticipated coefficient of variation for the method used to measure the shear-wave velocity profile (see Table 2.1). Site response analyses can be performed with any available method. Epistemic uncertainty on the site response model can be included by considering different site response models within a logic three approach.

The method proposed is theoretically rigorous but of limited use because there are very few sites for which sufficient data is available to constrain $GMPE$ that include the repeatable site effect. It is not current practice to rely on measurements at a given site for predicting seismic hazard. However, considering the length of time that is involved in the planning and design of critical facilities, it is not unreasonable to consider the option of installing instruments to measure repeatable site effects at a design site. Moreover, since site response can be predicted, a single, well placed bedrock instrument over a region can serve as a reference site for input to site response analyses.

It is important to note that should a station be located at the site of interest, then the repeatable site effect would be captured at the site and site response analyses would be needed only to account for deviations from the median due to non-linear effects. In such a case, the measured standard deviation at the site (the single-site standard deviation at the design site) would constitute a lower bound for the standard deviation that must be used in a $PSHA$. Values of the single-site standard deviation are studied in Chapter 5.

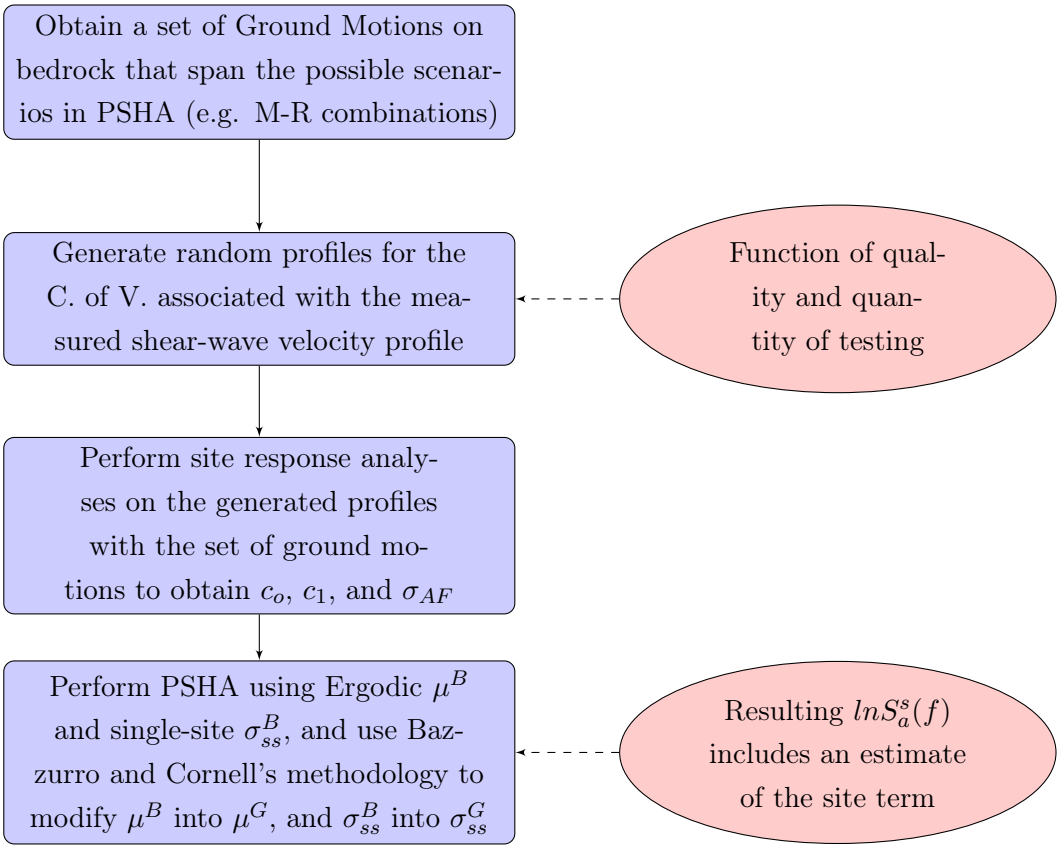


Figure 3.1. Flowchart for estimation of site-specific surface ground motion intensity

Method 2 The second method proposed for site-specific $PSHA$ considers the more general case in which no recordings are available at the site of interests. However, since the analysis is being conducted at a specific site, the ergodic assumption implicit in $GMPE$ derived from regional data sets should be removed, at least for the components that are predictable. In this method, it is assumed that site response is predictable through analytic means (e.g. a site response analysis). Prior to describing the method, it is necessary to describe a breakdown of the uncertainty in GMPE. A more rigorous presentation of this breakdown is described in 5, but for the interest of clarity, the breakdown is summarized herein.

It will be assumed that it is possible to predict ground motion at the soil-rock interface. Chapter 5 uses the KiK-net database to develop such predictions. Therefore, we express the ground motion prediction at the borehole as

$$y^B = \mu^B + \delta W^B + \delta B^B \tag{3.4}$$

where, y^B is the natural logarithm of the pseudo-spectral acceleration at the borehole, μ^B is the median estimate of y^B, δW^B is the intra-event residual, and δB^B is the inter-event residual. Capital letter W is used to indicate that the residuals are measured within a given event, capital letter B is used to indicate that the residuals are measured between different events. The standard deviation of the random variables that vary within the same event (i.e. earthquake) will be called ϕ, the standard deviations of the random variables that vary between events will be called τ, and the letter σ is reserved for total uncertainty.

The ground motion at the ground surface can then be predicted by

$$y^G = \mu^B + \delta W^B + \delta B^B + \mu^{AMP} + \Delta AMP \tag{3.5}$$

In this equation, the term $\mu^B + \mu^{AMP}$ constitutes the median ground motion at the ground surface and can be obtained from regression analysis of surface ground motions using a parameterization for site conditions (e.g. Vs_{30} and $h800$). The term ΔAMP is the remaining variability in the amplification factor and can be broken down into

$$\Delta AMP = \delta S2S^{AMP} + \delta AMP \tag{3.6}$$

Where delta $\delta S2S^{AMP}$ is a random variable that represents the variability between various sites that share the same parameterization, and δAMP is a random variable that represents the residual variability at a single site. The nature of this residual variability is important to the proposed method. The residual variability can be due to:

1. differences in amplification due to differences in phasing of the input motion
2. soil nonlinearity
3. site response differences due to non-vertically propagating waves.
4. 2D or 3D effects, such as basin response or topographic amplification

Items 1 and 2 are due to one-dimensional effects and can be captured analytically. Items 3 and 4 are not predictable through 1D analysis.

When a site-specific $PSHA$ is conducted, the $\delta S2S^{AMP}$ random variable takes a given value (i.e. a deterministic value). In absence of additional information at the site (e.g., if only Vs_{30} is given), it is not possible to constrain this deterministic value and delta $\delta S2S^{AMP}$ has to be considered a random variable. However, if additional information at the site is available, such as additional details on the Vs profile that allow for a site response analysis, then the value of $\delta S2S^{AMP}$ can be predicted, with a possible uncertainty on the prediction. With this consideration, Equation 3.5 can then be rewritten as

$$y^G = \mu^B + \left[\mu^{AMP} + \widehat{\delta S2S}^{AMP}\right] + D^{AMP} + \delta W^B + \delta B^B + \delta AMP^* \qquad (3.7)$$

The terms $\widehat{\delta S2S}^{AMP}$ and D^{AMP} represent analytical estimates that correspond to the deviation of the predicted amplification from the mean amplification (μ^{AMP}) and its uncertainty. D^{AMP} is a random variable with zero mean and standard deviation $\phi_{D^{AMP}}$ The remaining uncertainty related to the residuals δAMP^* is not known a priori but it is bounded to be lower than the uncertainty from δAMP because part of the effects captured by δAMP go into the D^{AMP} term. The term, however, is capped by the δAMP measured at single stations (Chapter 5) and limited below by zero, if the site is devoid of non 1D effects. Preliminary studies (figures 4.5 and 4.6 from Rodriguez-Marek and Montalva, 2010) indicate that site response analysis do capture most of the δAMP, hence the zero bound is realistic for some sites. In a PSHA analysis, this choice becomes an issue of epistemic uncertainty.

The $GMPE$ developed in Chapter 5 assumes linear site response. To introduce nonlinear site response, the methodology of Bazzurro and Cornell (2004b) can be used. In this case, however, the uncertainty of the input motion has to be modified to reflect the breakdown of uncertainty shown Equation 3.7. A switch from the notation used in this dissertation to that used by Bazurro and Cornell is given below.

- $lnS_a^s(f) \rightarrow y^G$
- $lnS_a^r(f) \rightarrow y^B$
- $\sigma_{AF} \rightarrow \phi_{S2S}$
- $\sigma_{lnS_a^r(f)} \rightarrow \sigma^B$

Note that $\phi_{\delta AMP^*}$ has to be added to the final uncertainty. Its value must be chosen arbitrarily or made to vary within a logic tree analysis. The uncertainty in the input

motion ($\sigma_{lnS_a^r(f)}$) has to be the ergodic uncertainty at bedrock (σ^B) because the bedrock term ($\delta B2B$) can not be estimated a priori in the analysis. Nevertheless, (σ^B) is lower than its counterpart at the surface (σ^G, see Figure 5.51), this is because an important part of the site response variability is included in σ^G.

The proposed methodology is shown in Figure 3.2. As indicated for Method 1, the choice of input motion set for the site response analyses has to reflect the statistics of the input motion, y^B, as predicted by a $GMPE$ for bedrock. Again, the median value can be estimated with only a few records (the recommendation of Rathje et al. (2010) is 7 records, of Bazzurro and Cornell (2004b) is 5 to 10). The uncertainty of the amplification factor necessitates a larger number of records. In the example given in Rodriguez-Marek and Montalva (2010) 14 records were used to obtain a stable estimate of the standard deviation of the amplification factor, that is the AF does not change with increasing number of records.

The significance of the proposed method is that it allows for two important contributions:

- Replacing the prediction of the amplification term given by the GMPE with a site-specific prediction (e.g., bias reduction)
- A possible reduction in uncertainty that would result if $\phi_{D^{AMP}} + \phi_{\delta AMP^*}$ is lower than $\phi_{\delta AMP} + \phi_{S2S^{AMP}}$. Chapter 5 looks at the breakout of residuals and considers the contribution of each residual.

A more subtle contribution of the proposed method is that the variability in site response computed analytically, D^{AMP} may be bounded by physical considerations of site response (e.g. soil nonlinearity), while there are no statistical basis to propose such bound on the random variables computed from the ground motion data.

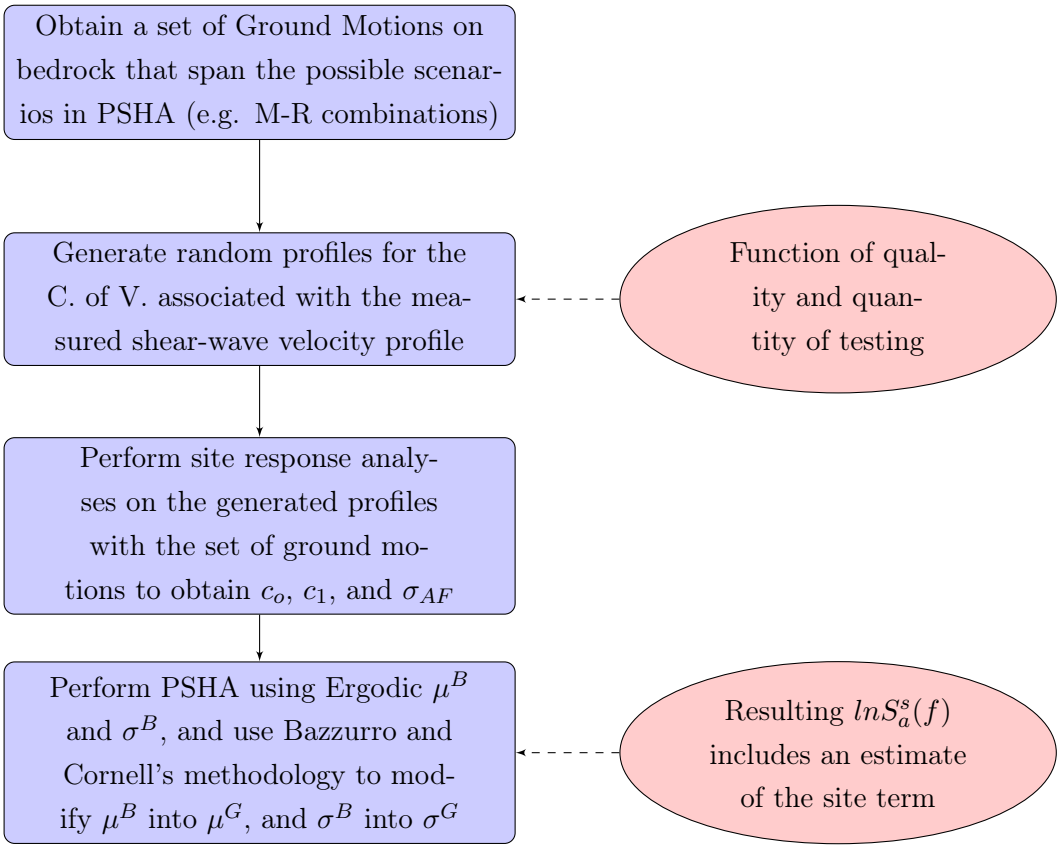


Figure 3.2. Flowchart for estimation of site-specific surface ground motion intensity

Chapter 4

Uncertainty in Site Response Analyses

The significance of site response was described in section 2.1. Site response analyses can have varying degrees of sophistication. For routine projects, code-defined factors that are based on site classification schemes can be used to modify mapped ground motion values. Alternatively, as described in section 2.3 site effects can be captured using ground motion prediction equations ($GMPE$) that incorporate site response through simple parametrization of site conditions (Abrahamson et al., 2008). The latter approach has the advantage that GMPE provide estimates of both the median value of a ground motion parameter and a quantification of its aleatoric uncertainty (e.g., the standard deviation value of the ground motion parameter).

When the shear-wave velocity (Vs) profile at a site is known or can be estimated, site-specific estimate of ground motions can be computed using site response analysis (e.g. Cramer, 2003). Such analyses have been performed routinely in engineering practice. When ground motions estimates are needed for PSHA, it becomes necessary to estimate both the median

and the uncertainty of ground motions at the surface of a given site. This implies that site response analyses must be conducted in a stochastic fashion, and both epistemic and aleatoric uncertainty of input parameters must be computed. For this purpose it becomes imperative to have a model that quantifies the uncertainty of Vs profiles (e.g. Asten and Boore, 2005; Brown et al., 2002; Marosi and Hiltunen, 2004; Martin and Diehl, 2004; Moss, 2008).

Uncertainty in the shear-wave velocity of the soil profile can be quantified by a random field model (e.g. Toro, 1995). This chapter describes the development of a random field model using the KiK-net database to constrain the parameters of the model. Once the input variability is constrained, a Monte Carlo simulation approach can be used to compute how input variability propagates to the estimates of site response.

4.1 Random Field Model for Shear Wave Velocity

Point-estimates of variability are not sufficient to quantify the statistical distribution of a spatially varying quantity. Spatial variability models for geotechnical properties have been presented by various authors (e.g. Fenton, 1994; Phoon and Kulhawy, 1999). EPRI (1993) presented a model for generating Vs profiles based on statistical data of recorded profiles. This type of model can be used to generate artificial Vs profiles. Random profile generators can be used in conjunction with site response analyses to obtain the epistemic uncertainty of site response when site profile data is limited.

The objective of this section is to present five models for the artificial generation of Vs profiles. The first model is based on the EPRI (1993) model. The second and third models are a simplification of the EPRI model. The last two models are based on the concept of Markov Chains. The model parameters are constrained using the KiK-net database (Kiban-Kyoshin, 2010), a large database of Vs profiles from Japan. The appropriateness

of the models is judged by comparing the spatial variability of the measured shear-wave velocities with the spatial variability of sets of artificially generated Vs profiles. In addition, the appropriateness of the proposed models is evaluated in terms of their ability to simulate the one-dimensional site response of the original data set. For this purpose, site response statistics are computed using subsets of the measured Vs and artificially-generated sets of profiles.

4.1.1 Existing Methodologies

The most widely used random field model for Vs profiles is the model presented in EPRI (1993), and extended by Toro (1995). The model defines independently the properties of the bedrock, the profile layering, and the Vs variation with depth. The bedrock depth is defined by a uniform distribution while the bedrock Vs is defined by a log-normal distribution that is independent of the Vs of the soil overlying the bedrock. The profile layering is defined by a non-homogeneous Poisson process (e.g. the process' parameters are depth dependent). Finally, the Vs model uses a first order non-stationary auto-regressive Gaussian process to reproduce the spatial statistics of the Vs at the mid point of each layer. Details of this model are given later in the paper.

Only few other models to reproduce Vs profiles have been proposed. Notably, Douglas et al. (2009) proposed a model that is based on the quarter-wavelength method and uses velocity gradients to generate Vs profiles. This model is focused on shear-wave velocities at large depths (of interest to seismologists). Since the focus of this study is on the shallow Vs profile, the Douglas et al. model will not be included in comparisons with the proposed models.

4.1.2 Spatial statistics of the KiK-net database

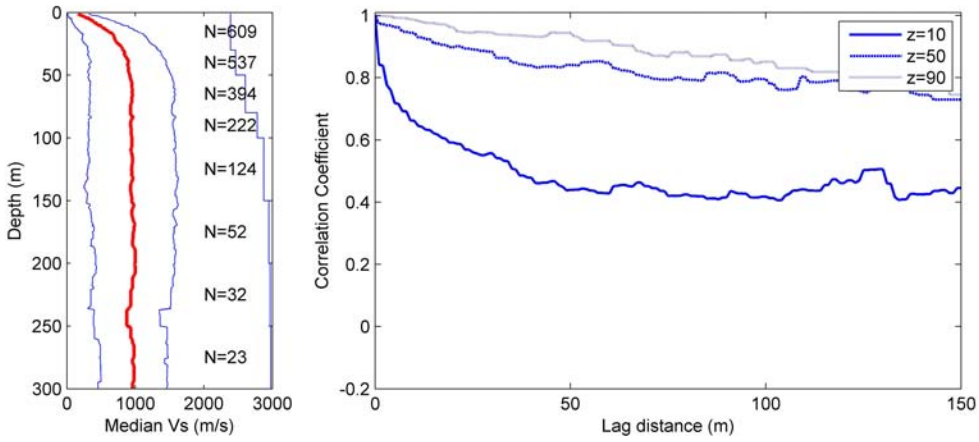
The KiK-net database is a ground motion database located in Japan, consisting on more than 600 stations. Each station in this network has two 3-component accelerometers, one at the surface and another at depth (most sensors are located at a depth of -100m or -200m). The instruments have a 24 bit analog-to-digital converter with a sampling frequency of 200 Hz (Fujiwara et al., 2004). Vs profiles at each of the ground motion stations were obtained using PS-logging Vs measurements.

Shear-wave velocity distribution of the measured profiles are herein analyzed differentiating it from the shear-wave velocity at the instrument depth, statistics for the latter are given in section 4.1.2.2.

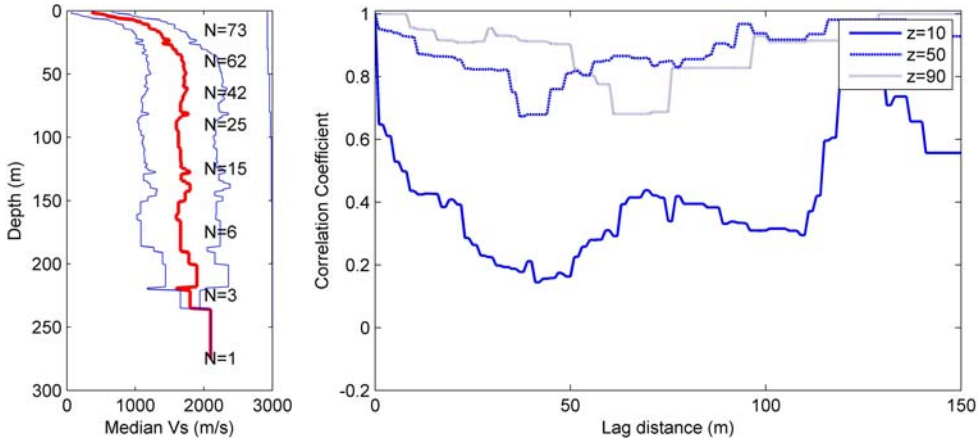
4.1.2.1 Soil Statistics and Correlation Structure

The statistics of the Vs profiles of the KiK-net database are shown in Figure 4.1 for Vs_{30}-based subsets of the KiK-net database. These subsets are selected because Vs_{30}-based site classifications have been adopted by various national codes (e.g. IBC, 2006; CEN, 2003). The plots on the left show the median Vs value with a one-standard deviation band. Spatial correlation was calculated for all depths but is shown only for three arbitrary depths (10, 50, and 90 m). The correlation function shown in Figure 4.1 is the Pearson's correlation coefficient between all the natural logarithm of Vs values at the depth of interest (i.e. z = 10, 50, and 90 m) and the natural logarithm of Vs at a lag distance Δz (e.g. at a depth of z + Δz).

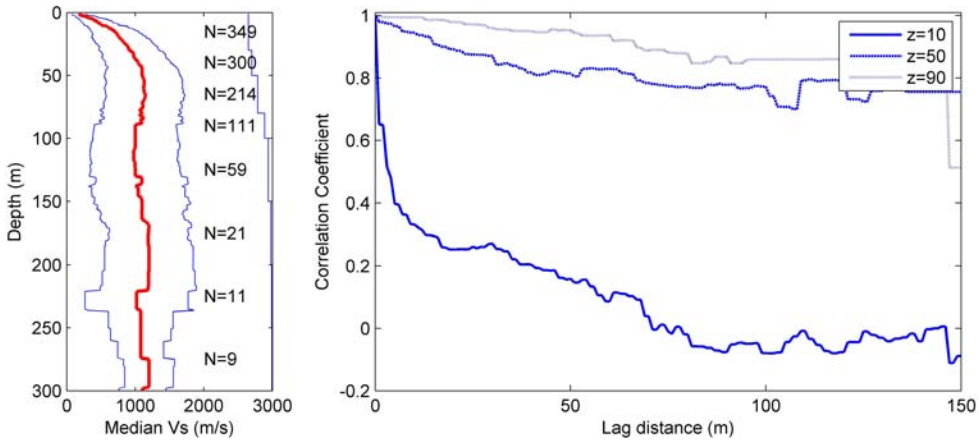
The standard deviations of the natural logarithm of Vs for the entire database and Vs30 based subsets are compared in Figure 4.2. As expected, the statistics of the entire database have more variability than the statistics of subsets based on Vs30 based site classes, in spite of the fewer data on each class category. Differences are also observed on the spatial



(a)

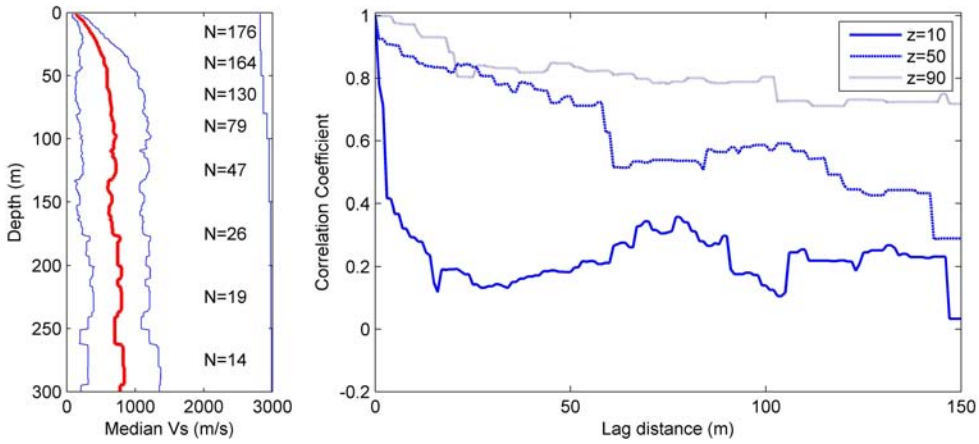


(b)



(c)

Figure 4.1



(d)

Figure 4.1. Shear-wave velocity statistics and correlation coefficients for the entire data set and Vs30 based subsets. 4.1a entire database, 4.1b site classes A and B (Vs30 $>$ 760 m/s), 4.1c site class C (760 m/s $>$ Vs30 $>$ 360 m/s), and 4.1d site class D (360 m/s $>$ Vs30 $>$ 180 m/s).

correlation (Figure 4.1). Note that the spatial correlation of the site classes (Figures 4.1b to 4.1d) is "noisier" than the spatial correlation of the entire data set; however, all subsets show strong spatial dependence, suggesting that an auto-regressive model (e.g. a model where properties at depth i are a function of properties at depth i-1) should be appropriate for random profile generation. It is interesting to note that shallow shear-wave velocities (e.g. at z $\leq$ 10 m) are weakly correlated with deeper layers even for relatively short lag distances.

The lower standard deviations observed in the Vs30-based site classes, along with strong correlation coefficients between different profile depths indicate that Vs30-based site classification is an adequate proxy for Vs variability at all depths in the analyzed database. A similar observation was made by Toro (1995), who compared the statistics of Vs profiles for different site classes from a California data set.

Figure 4.3 shows the correlation coefficient of shear-wave velocities between two depths separated by lag-distance Δz and considering only profiles where those two depths are in

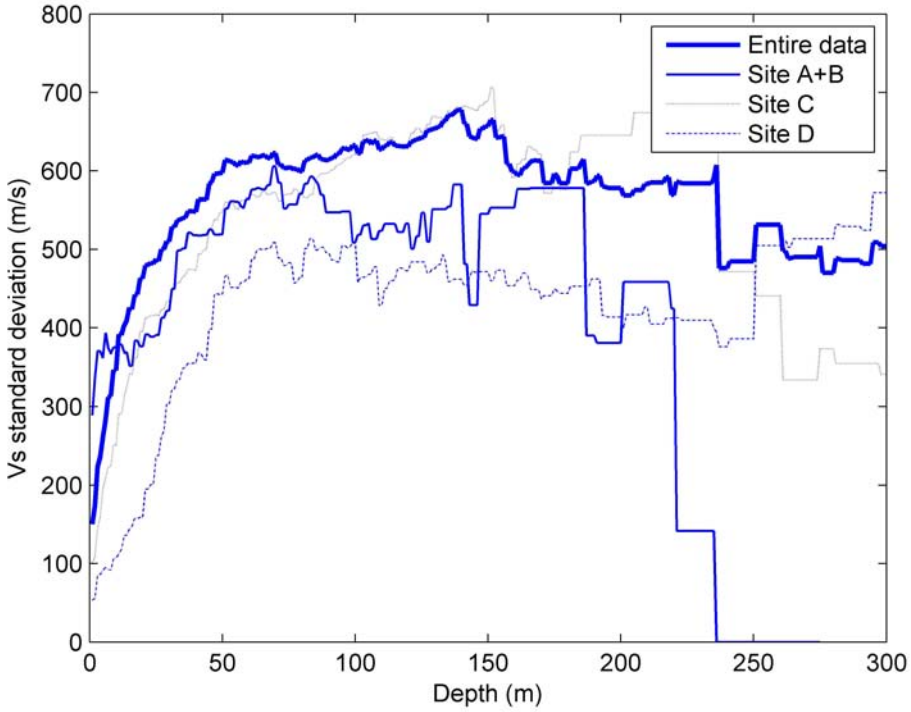


Figure 4.2. Comparison of the standard deviation of shear wave velocity for the entire database and for Vs30-based subsets.

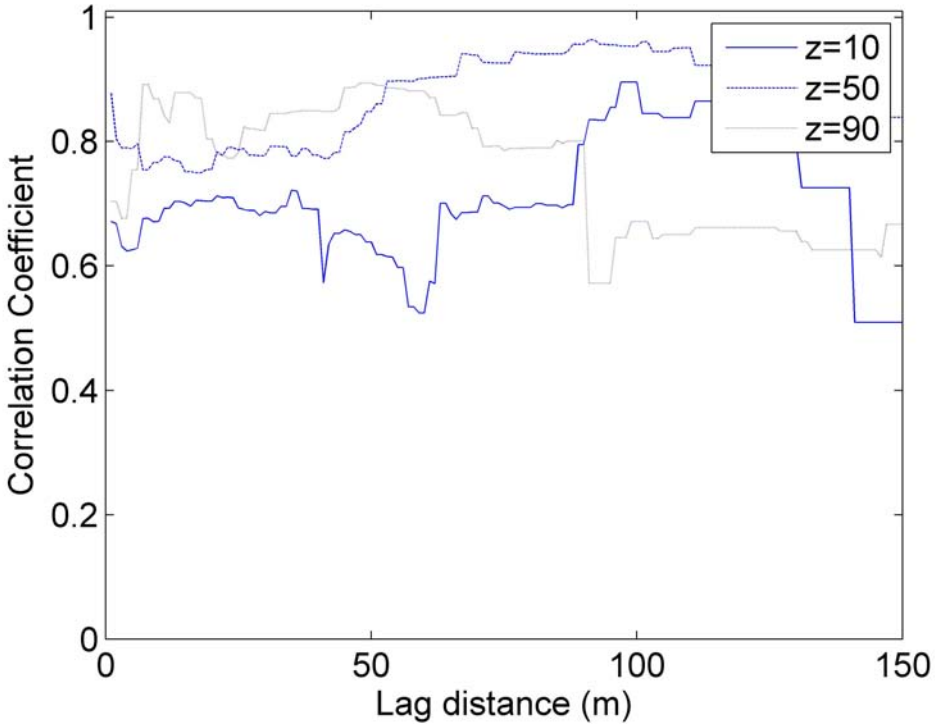


Figure 4.3. Shear-wave velocity correlation coefficients considering only lag distances that are one layer away from each other. Initial depths of 10, 50, and 90 meters are shown.

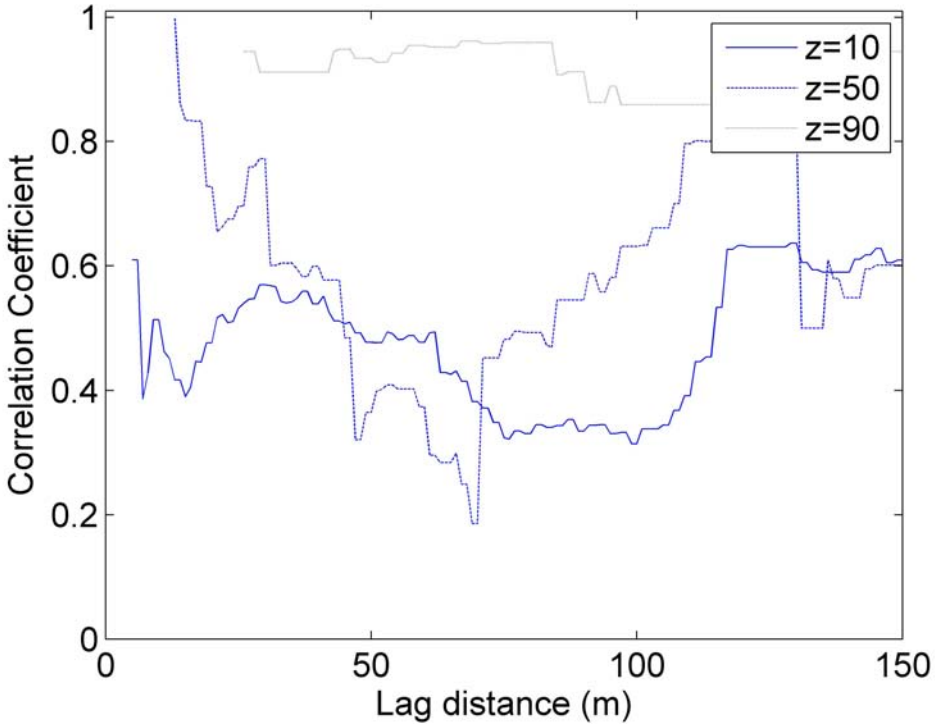


Figure 4.4. Shear-wave velocity correlation coefficients considering only lag distances that are two layers away from each other. Initial depths of 10, 50, and 90 meters are shown.

Table 4.1. Half-space Statistics

Site Class	Rock depth log-normal parameters		Shear wave velocity (m/s)		Vs_R[1] beta distribution parameters		Correlation coefficient between Vs_R and indicated parameters		
	μ	σ	μ	σ	α	β	Depth	Vs_{30}	Vs_{last}[2]
Generic	4.33	0.66	1626	704	1.6	2	0.01	0.44	0.77
Sites A + B	4.2	0.67	2196	562	1.61	1.51	-0.09	0.27	0.63
Sites C	4.27	0.62	1691	629	1.56	2.07	-0.05	0.26	0.77
Sites D	4.53	0.69	1310	694	1.04	2.18	0.29	0.02	0.75

[1] Shear-wave velocity of the bedrock layer

[2] Shear-wave velocity of the deepest soil layer

contiguous layers (i.e., $\Delta_{Layer} = 1$). Observe that the correlations are strong. Figure 4.4 shows the same correlations but only considering cases where there is a two-layer separation ($\Delta_{Layer} = 2$). Note that the correlation significantly decreases when considering two layers separation. Note that while correlation between two depths decreases with lag distance (Figure 4.1), correlation seems to be constant when layer separation is considered (Figures 4.3 and 4.4). A functional form to model this dependence will be shown below.

4.1.2.2 Rock Statistics

Site response analyses require a characterization of the material below the soil column. This material is commonly assumed to constitute an elastic half space (elastic material that continues infinitely below a given depth) in site response modeling. Based on an analysis of the KiK-net data, it appears convenient to model the half space separately from the soil column. Table 4.1 shows the statistics of the half-space. Observe that Vs at and below the instrument depth (which corresponds generally to bedrock, but in some cases is an arbitrary depth at which profiling was stopped), is well correlated with Vs at the deepest measured layer, but very poorly correlated with depth.

It is commonly assumed that shear-wave velocities at a given depth can be modeled by a log-normal distribution. This is the case of the studies by Toro (1995, 2006); EPRI (1993); Andrade and Borja (2006). This hypothesis was tested using a χ^2 goodness-of-fit test at 5% significance level. Site classes A and B were combined for the analysis due to the lack of sufficient data for individual analyses, the results show that a log-normal distribution fits the data but only up to a depth of about 30 meters; between 30 and 60 meters a normal distribution better fits the available data; beyond 60 meters either normal or uniform distributions are appropriate, but the amount of data is not sufficient to make such fits reliable.

Figure 4.5 shows a formal assessment of the goodness-of-fit of the log-normal distribution to the Vs at each depth. The p-values were calculated using a χ^2 test for most of the cases and a Lilliefors test for small samples (Conover, 1980). Observe that the log-normal distribution is poorly suited to the data for depths between 30 and 70 meters for sites classes C and D. Since the overall data set is controlled by these site classes, the lumped data set also has a poor fit between 30 and 70 meters. However, an examination of the cumulative distribution functions (CDF's) for that depth range reveals a close match between the log-normal distribution and the data in all regions except the right tail.

4.1.3 Proposed Models

4.1.3.1 Modified EPRI

The EPRI (1993) model has been modified to match the characteristics of the KiK-net database. Let the natural logarithm of Vs as a function of depth be defined by a random field V(z) where z is depth. Assume that

$$V(z) = t(z) + \epsilon(z) \tag{4.1}$$

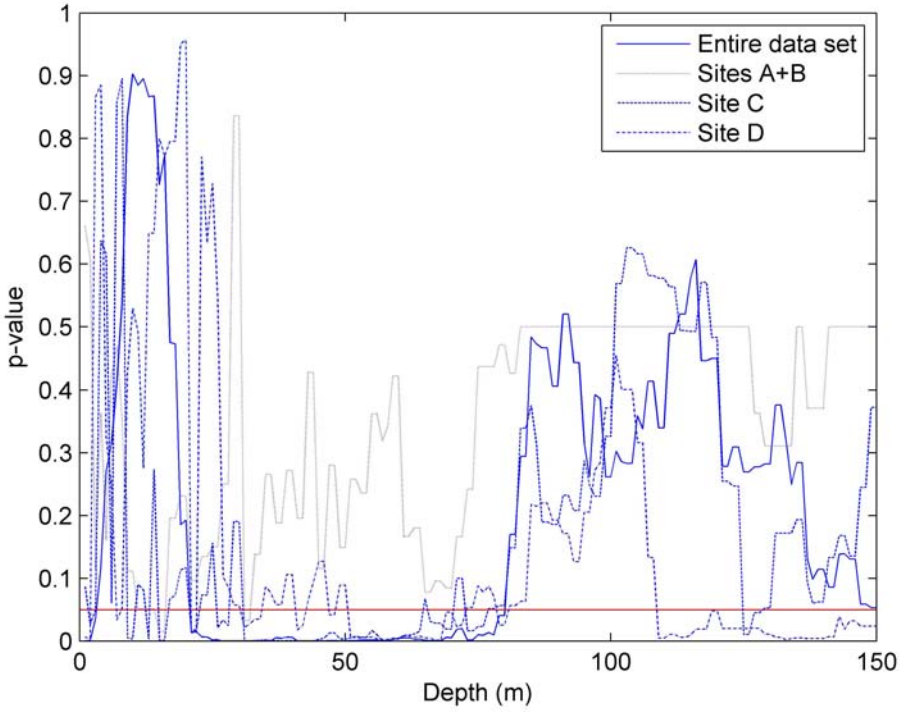


Figure 4.5. p-values versus depth, for the hypothesis that the Vs values fit a log-normal distribution at a given depth. At 5% significance level, all p-values greater than 0.05 are considered to approach a log-normal distribution.

where $t(z)$ is the median value of the logarithm of Vs, and $\epsilon(z)$ is a Gaussian random process with zero mean, standard deviation $\sigma_\epsilon(z)$ and auto-correlation $\rho(\Delta z,\ z)$ given by:

$$\rho(z, \Delta z) = (1 - \rho_S(z)) * exp(-\Delta z/\kappa) + \rho_S(z) \tag{4.2}$$

where $\rho_S(z)$ and κ are given by

$$\rho_S(z) = \rho_0 - exp(-(z - z_0)/b) \geq 0 \tag{4.3}$$

$$\kappa = \kappa_0 + (\kappa_{max} - \kappa_0) * (z/z_{max}) \geq 0 \tag{4.4}$$

where ρ_0, z_0, b, κ_0, κ_{max}, and z_{max} are fitting parameters, and where κ is limited to κ_{max} ($\kappa \leq \kappa_{max}$). The parameters obtained from a fit to the entire KiK-net data set and from separate fits to subsets based on Vs30 site categories are shown in Table 4.2. The parameters were obtained using the Levenberg-Marquardt algorithm for nonlinear regression method implemented into Matlab.

The geostatistical model above can be used in conjunction with a routine to generate artificial Vs profiles that replicate the statistics of the underlying data. The first step for generating artificial profiles is to determine the depth to the bottom of the profile. By definition, the bottom of the profile corresponds to the location of the elastic half space in site-response analyses. For the KiK-net database, the depth to the bottom of the profile follows a log-normal distribution (see Section 4.1.2.2). Parameters of the log-normal model for the KiK-net database are shown in Table 4.2.

After the depth to the bottom of the profile is defined, the next step consists in the generation of the layering for each profile. This process is identical to that presented in 2.6,

Table 4.2. Parameters for Non-stationary Correlation Function

	ρ_0	z_0	b	κ_0	κ_{max}	z_{max}
Generic	0.5532	-5.0031	5.4041	-19.00	83.9813	35
Sites A & B	0.8860	1.6252	9.5663	3.16	5.2920	43
Site C	0.5756	7.3486	4.2589	-8.1	77.5863	50
Site D	0.8143	38	12.3571	-105.00	58.4465	16.6

Table 4.3. Coefficients for layering model, 95% confidence interval in parenthesis

	a	b	c
Generic	0.34 (-0.08; 0.75)	2.39 (-4.08; 8.87)	0.58 (0.27; 0.89)
Sites A & B	0.1 (0.001; 0.2)	-1 (-1.57; -0.43)	0.01 (-0.25; 0.27)
Site C	0.24 (-0.08; 0.57)	1.27 (-5.69; 8.22)	0.46 (0.1; 0.81)
Site D	0.18 (-0.005; 0.36)	1.4 (-7.24; 10.11)	0.33 (0.08; 0.58)

briefly repeated herein for clarity. This is accomplished by using a non-stationary Poisson model with parameter λ^{-1}, where λ is a model parameter that represents the frequency of layer transitions at each depth. Figure 4.6 plots the number of layer transitions as a function of depth. A relationship for λ as a function of depth that fits the data shown in the figure is given by:

$$\lambda(z) = a * (b + z)^{-c} \tag{4.5}$$

Table 4.3 shows the regression results for the coefficients of Equation 4.5 on each category, and Figure 4.6 shows the fit of Eq. 4.5 for the generic case (i.e. the entire database). Once the layering is generated for each profile using Eq. 4.5, the Vs values are generated using the statistical model given in Eqs. 4.1 through 4.4. Rock Vs values are generated using the mean and standard deviations from Table 4.1, and are correlated to the Vs value of the last layer of the profile with the corresponding correlation coefficients given in Table 4.1. Figure 4.7 summarizes the process.

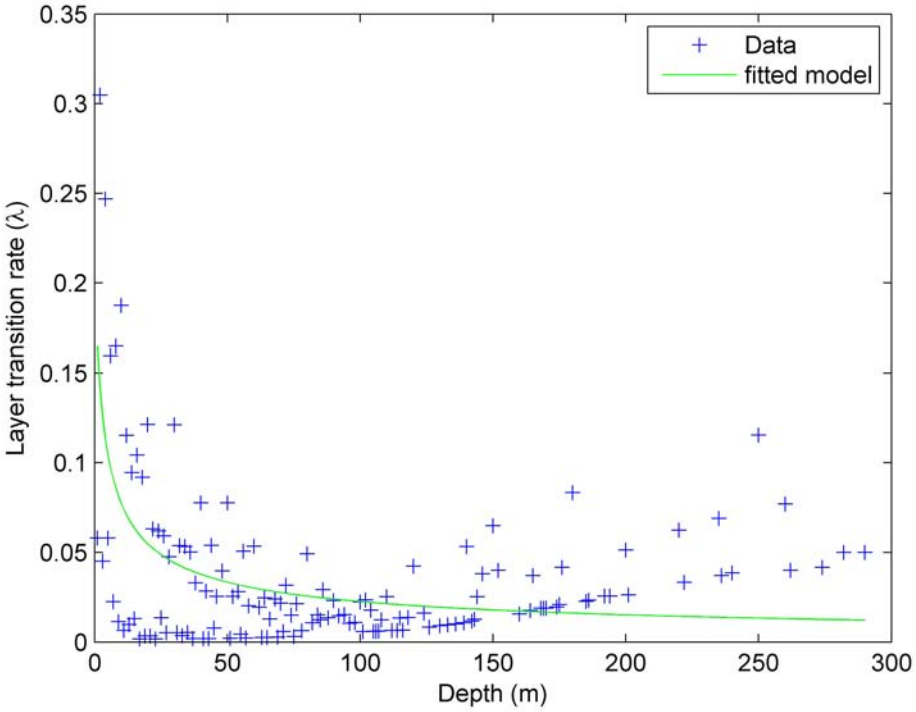


Figure 4.6. Rate of layer transitions versus depth for the generic case. This parameter, λ, can be understood as the reciprocal of layer thickness for each depth.

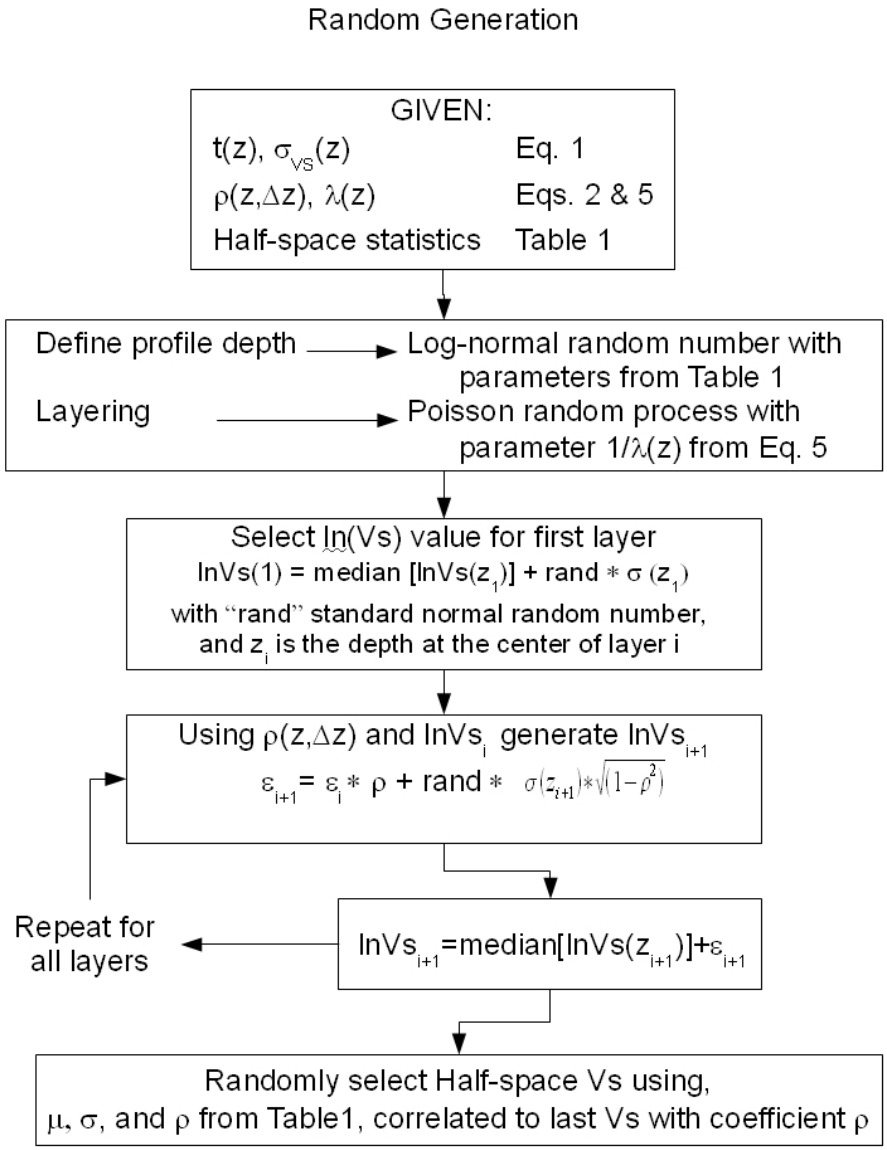


Figure 4.7. Flow chart of the generation procedure for the modified EPRI model.

Table 4.4. Correlation Coefficient for Stationary Models

Category	$\rho(\Delta L = 1)$	$\rho(\Delta L = 2)$
Generic	0.85	0.69
Sites A & B	0.67	0.88
Site C	0.83	0.65
Site D	0.71	0.53

4.1.3.2 One Layer Lag Stationary Gaussian Model

The strong correlation of Vs values between adjacent layers (Figure 4.3) indicates the suitability of a model that uses only this correlation as its correlation structure (e.g., a stationary model).

Table 4.4 shows the depth-independent correlation values for each category. Profile generation follows the same procedure as the Modified EPRI model.

4.1.3.3 Two Layer Lag Stationary Gaussian Model

Figure 4.4 shows the correlation structure resulting from considering only layers that are separated by one intermediate layer (two layer lag). Note that the correlation coefficients are significantly lower than those for a one layer lag (Figure 4.3). The correlation structure could be also modeled using a constant (e.g. depth-independent) correlation coefficient.

Using a normal conditional probability density function (*pdf*), random profiles can be generated such that the correlation between two adjacent layers, and layers that are separated by one layer is always constant, $\rho_{\Delta L=1}$ and $\rho_{\Delta L=2}$ respectively. For any given ln-Vs value can be generated using the covariance matrix in Eq. 4.6 and a normal *pdf* with mean and standard deviations given by Eqs. 4.7 and 4.8.

$$C = \begin{pmatrix} \sigma(i)^2 & \rho_{\Delta L=1} * (\sigma(i) * \sigma(i-1)) & \rho_{\Delta L=2} * (\sigma(i) * \sigma(i-2)) \\ \rho_{\Delta L=1} * (\sigma(i) * \sigma(i-1)) & \sigma(i-1)^2 & \rho_{\Delta L=1} * (\sigma(i-2) * \sigma(i-1)) \\ \rho_{\Delta L=2} * (\sigma(i-2) * \sigma(i)) & \rho_{\Delta L=1} * (\sigma(i-1) * \sigma(i-2)) & \sigma(i-2)^2 \end{pmatrix}. \quad (4.6)$$

$$\mu^* = \mu_i + C(1, 2:3) * C(2:3, 2:3)^{-1} * ([lnVs(i-1)lnVs(i-2)] - [\mu_{i-1}\mu_{i-2}])^T \quad (4.7)$$

$$\sigma^* = C(1,1) - C(1, 2:3) * C(2:3, 2:3)^{-1} * C(2:3, 1) \quad (4.8)$$

This procedure can be used iteratively to generate all the needed shear-wave velocity values.

4.1.3.4 Markov Chain

A Markov Chain based model is proposed for the generation of random profiles. The model is constrained by the statistics and correlation structure of the underlying database (the KiK-net database in this case). The first step is the generation of Vs values adequately correlated with layer depth. The second step is the assemblage of the correlated pairs into random Vs profiles. In this study, the layer depth was chosen to represent the depth to the bottom of the layer.

This model uses the fact that shear-wave of consecutive layers in the KiK-net database are strongly correlated. The generation of the correlated Vs values and their corresponding layer depths is done by means of rank correlation coefficients (RCC), as opposed to simply using a linear correlation coefficient (ρ). The selected RCC for this study was Kendall's τ. The procedure for the generation of the correlated variables consists of the following steps:

1. Compute the RCC (i.e. Kendall's τ) for the data set.

2. Transform that coefficient to Pearson's ρ using

$$\rho = sin(\tau * \pi/2) \tag{4.9}$$

3. Generate correlated Gaussian random variables using the correlation coefficient computed in (2).

4. Obtain the CDF for each variable, which by definition will lead to two uniform distributions.

The CDF's obtained in (4) can be inverted to obtain the marginal distributions of layer depth and Vs constrained by the empirical Vs and layer depth distributions. The marginal distribution for shear-wave velocities in the KiK-net database can be modeled by a log-normal distribution, with mean and standard deviation of the associated normal distribution of 6.2 and 0.84 respectively. The marginal distribution for layer depth can be modeled by a log-normal distribution, with mean and standard deviation of the associated normal distribution of 3 and 1.3 respectively. Goodness of fit tests confirmed this fit to be appropriate. The result of this approach is a set of randomly generated pairs of Vs and layer depth that follow the same correlation structure as the original database (Figure 4.8)

Once pairs of layer depth and shear wave velocity are generated, these pairs can be assembled into realistic Vs profiles. This is achieved by first creating depth and Vs bins that span the entire range of possible Vs and depth values. A *state* is then defined as a combination of a Vs bin and depth bin. As a simple example, we could define depth and Vs bins as:

depth bin 1 (d1) = [0, 50]

depth bin 2 (d2) = [50, 150]

Vs bin 1 (V1) = [360, 560]

Vs bin 2 (V2) = [560, 760]

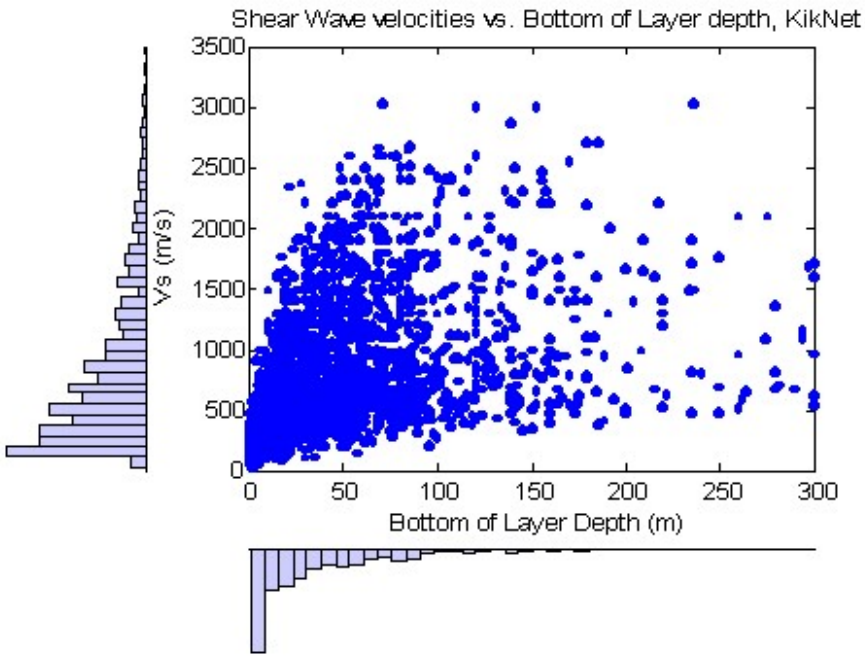


Figure 4.8. Correlation between layer depth and shear-wave velocity for the entire KiK-net database.

All possible *states* (i.e. d1V1, d1V2, d2V1, and d2V2) define the s*tate vector* (Pi), which in this simple example would be a 4 by 1 vector. For a given layer (characterized by a given Vs and layer depth), the state vector takes a given value. In this example, the state vector for layer i, denoted by Pi, would have 3 zeros and a 1 in the position that corresponds to the Vs-depth combination of layer i. A transition matrix can then be defined to transition between the state for one layer to the state of the adjacent layer:

$$P_i = T * P_{i-1} \tag{4.10}$$

where P is the *state vector* and T the transition matrix. In matrix form Eq. 4.10 takes the form:

$$\underbrace{\begin{pmatrix} state_1 \\ state_2 \\ \vdots \\ state_n \end{pmatrix}}_{P_i} = \underbrace{\begin{pmatrix} P(state_1|state_1) & P(state_1|state_2) & \cdots & P(state_1|state_n) \\ P(state_2|state_1) & P(state_2|state_2) & \cdots & P(state_2|state_n) \\ \vdots & \vdots & \ddots & \vdots \\ P(state_n|state_1) & P(state_n|state_2) & \cdots & P(state_n|state_n) \end{pmatrix}}_{T} * \underbrace{\begin{pmatrix} state_1 \\ state_2 \\ \vdots \\ state_n \end{pmatrix}}_{P_{i-1}} \tag{4.11}$$

The transition matrix corresponds to the conditional probability that the Vs-layer depth pair in layer i be any state given a current *state* (represented by P_{i-1}). The first step in constructing the transition matrix is defining the desired accuracy, which is directly related with the size of the transition matrix. In this study Vs and layer depths were divided into 6 groups each, giving 36 possible states. One additional case was added to include the probability that given P_{i-1} the profile ends at that depth, thus the transition matrix is a square matrix of size 37 and the *state vector* is 37 by 1. The process is initiated by selecting an initial state using the surface Vs distribution. Eq. 4.11 is then used to generate a *state vector* where each element corresponds to the probability of each state (i.e. Vs-depth bin) occurring. A random number generator is then used to generate a realization of this state vector that has only one nonzero state (i.e. *state k =1*). This process is repeated iteratively until the nonzero state in the *state vector* P_i corresponds to the case in which the profile ends (the last state in the state vector). The specific Vs and depth values of each layer are sampled from the Vs-depth pairs computed in step one within the limits set by the corresponding state.

4.1.3.5 Second Order Markov Chain

Using the same principle, but taking into consideration two layers of dependence $f(\epsilon_3|\epsilon_1, \epsilon_2)$, a second order Markov chain model is proposed. This model is identical to the previous model (Section 4.1.3.4), with a notable exception on the dimension of the transition matrix,

T. The transition matrix is not square because it considers the same number of states (i.e. the size of the state vector) and all the possible combination of states. This results on a transition matrix with the number of rows equal to the length of the state vector, and the number of columns equal to the square of that number, plus the end of profile case in both dimensions. Each component of T is now the conditional probability that the Vs-layer depth pair in layer i be any state given a current state (represented by P_{i-1}) and a previous one (represented by P_{i-2}). An auxiliary state vector is used to consider all possible state combinations. Eq. 4.12 shows the dimensions of the matrices, where n is the length of the *state vector*.

$$[n * 1] = [n * n^2] * [n^2 * 1] \tag{4.12}$$

4.1.4 Statistical Comparison

The ability of each of the proposed models to reproduce the statistics of an underlying ground motion database is tested by generating sets of artificial profiles. Figures 4.9 through 4.13 show the median plus and minus one standard deviation for the five sets of artificially generated profiles, one for each of the proposed models, compared with the statistics from the KiK-net database. In both the generated profiles and the KiK-net database cases, the compared profiles correspond to IBC site C class.

All five models show reasonable ability to reproduce the median and standard deviations, although it is worth noting that the Gaussian models are smoother. The statistics for the five models are calculated from 600 profiles each, which was selected based on the number of profiles available from the KiK-net database, the number of generated profiles will have a direct impact on the smoothness of the calculated statistics.

Figures 4.14 and 4.15 show a comparison of the empirical correlation functions calculated

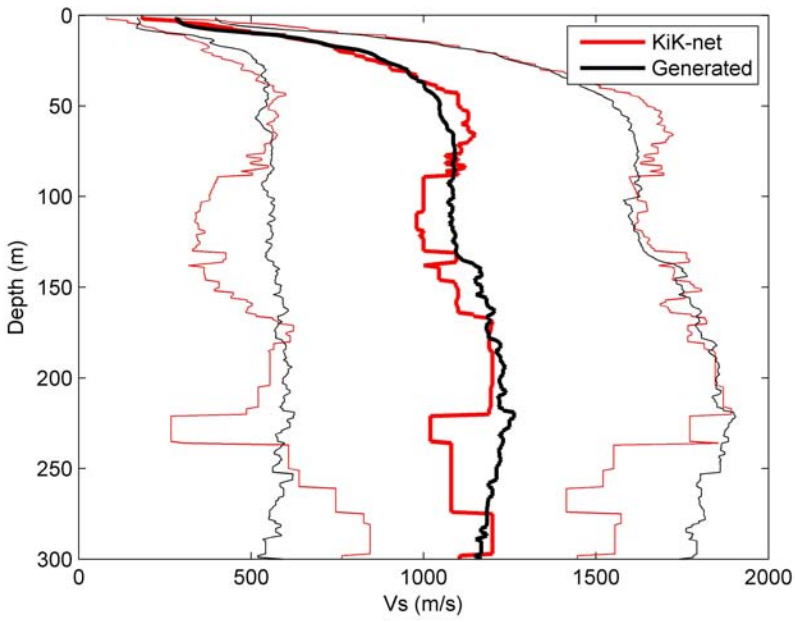


Figure 4.9. Comparison of the mean and one standard deviation band for one set of simulated Vs profiles using the Modified EPRI methodology, and a set of measured Vs profiles for site C class.

for the five models and the KiK-net database. Here the best fitting model is the two layer lag stationary Gaussian (lines in thick blue).

4.1.5 Comparison in Terms of Site Response

4.1.5.1 Comparison of site response between a set of measured Vs profiles and a set of artificially generated Vs profiles

To compare the impact of the randomly generated profiles in a site response exercise, we compared the results of equivalent linear site response analyses using the measured KiK-net profiles and those generated using the proposed models. Results are compared in terms of the ratio of the response spectra at the surface to the input response spectra (Ratio of Response Spectra). As input motions for the analyses a sample of 100 ground motions

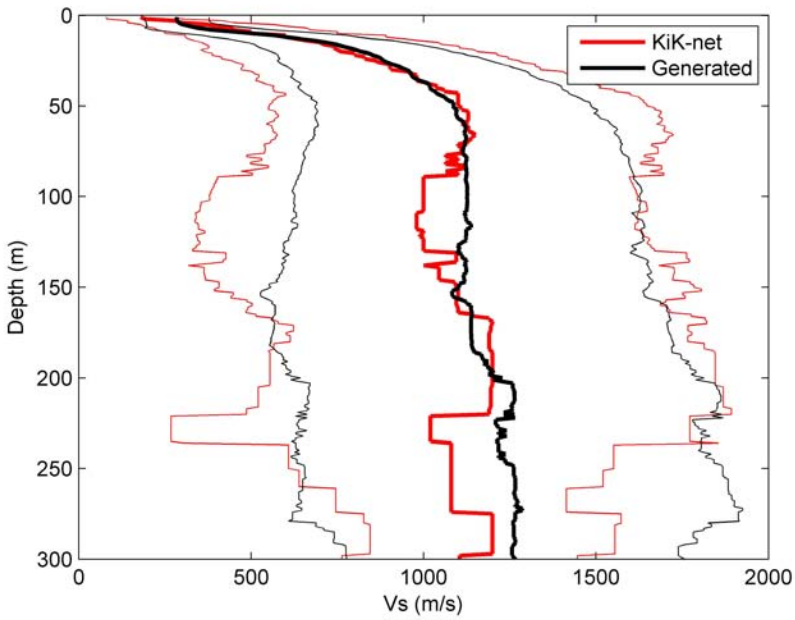


Figure 4.10. Comparison of the mean and one standard deviation band for one set of simulated Vs profiles using the One layer lag Stationary Gaussian methodology, and a set of measured Vs profiles for site C class.

from the NGA database (PEER, 2010) was selected. These ground motions satisfy certain general criteria on magnitude and distance form the fault ($6 \leq$ Mw ≤ 8; Rjb ≤ 100 km, and Vs30 ≥ 650 m/s).

Figure 4.16 shows the comparison of median ratio of response spectra (RRS). Note that although the Markov Chain models are able to reproduce reasonably well statistics and spatial correlation of the Vs profiles (Figures 4.12 to 4.15), they fail to reproduce site-response behavior. This shows the sensitivity of the models to the rock or half-space depth. In the three Gaussian models the depth to rock is randomly generated following the distribution of the measured depths, while in the Markov models this depth is automatically generated within the selected bins (e.g. accuracy level), hence the response of the Markov models can be improved by selecting closely spaced bins near the mean depth of the measured profiles.

The variability obtained on the ratio of response spectra using the five different models

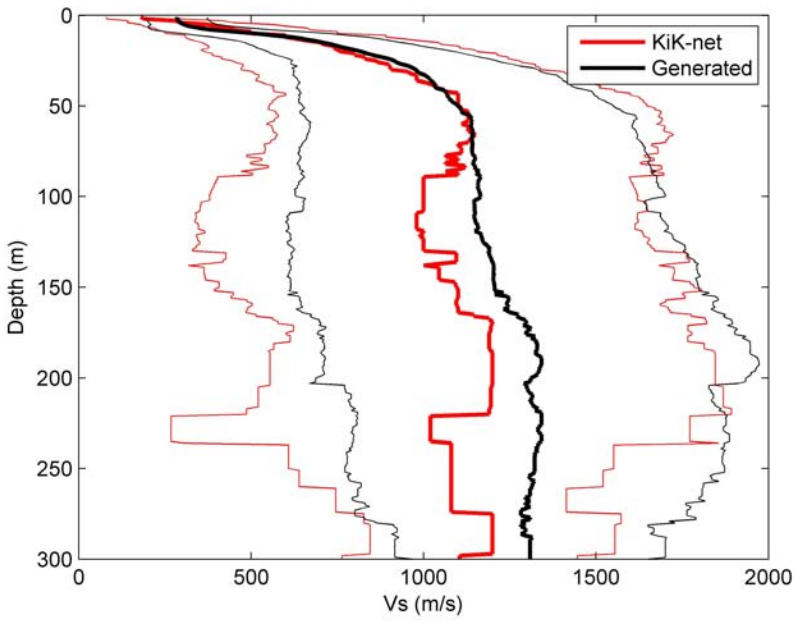


Figure 4.11. Comparison of the mean and one standard deviation band for one set of simulated Vs profiles using the Two layer lag Stationary Gaussian methodology, and a set of measured Vs profiles for site C class.

are shown in Figure 4.17, the results show the Gaussian models better approximate the variability in ratio of response spectra computed using the measured profiles.

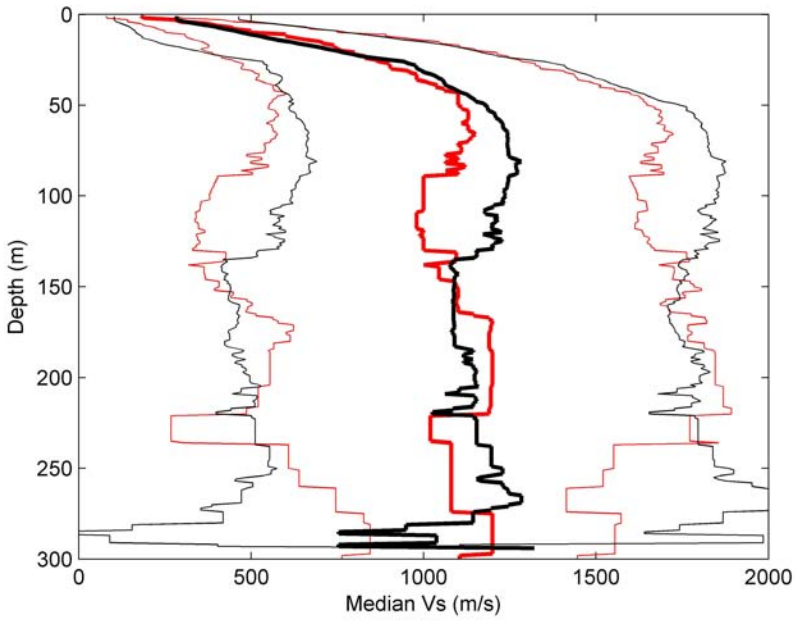


Figure 4.12. Comparison of the mean and one standard deviation band for one set of simulated Vs profiles using the first order Markov chain methodology, and a set of measured Vs profiles for site C class.

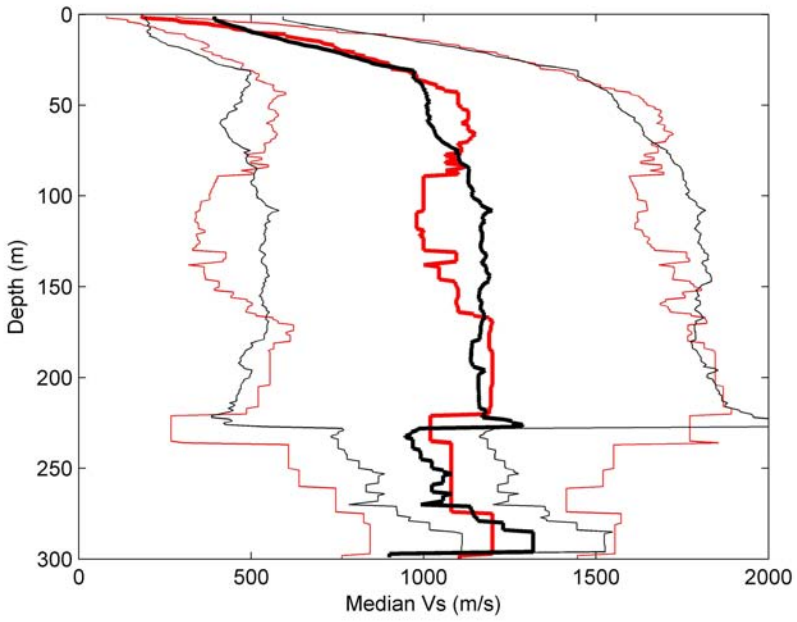


Figure 4.13. Comparison of the mean and one standard deviation band for one set of simulated Vs profiles using the Second order Markov chain methodology, and a set of measured Vs profiles for site C class.

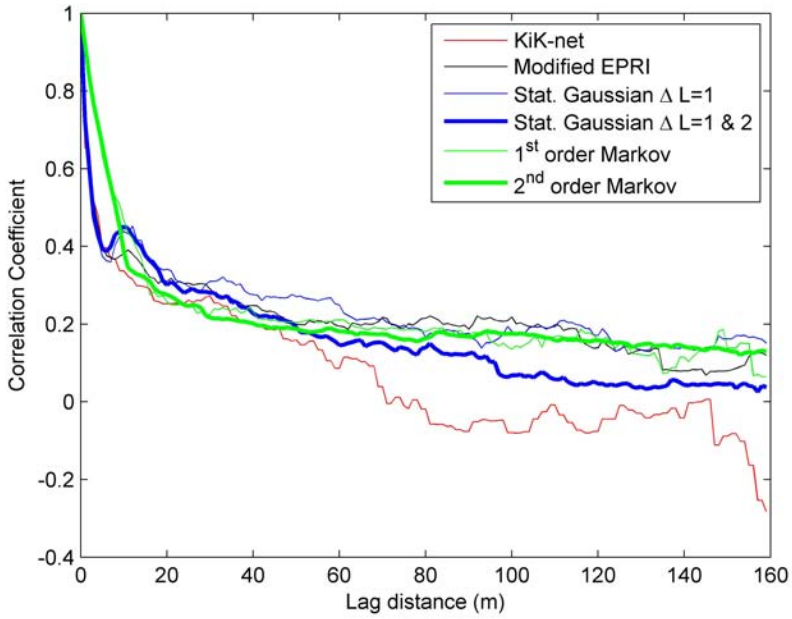


Figure 4.14. Empirical correlation functions for depth 10 m, and site class C. Comparison of measured profiles with sets of simulated profiles using the five proposed methodologies.

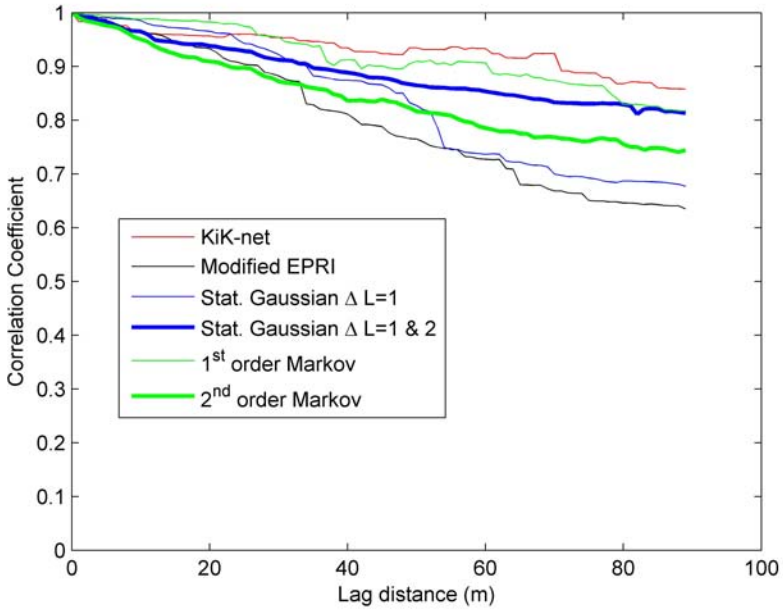


Figure 4.15. Empirical correlation functions for depth 50 m, and site class C. Comparison of measured profiles with sets of simulated profiles using the five proposed methodologies.

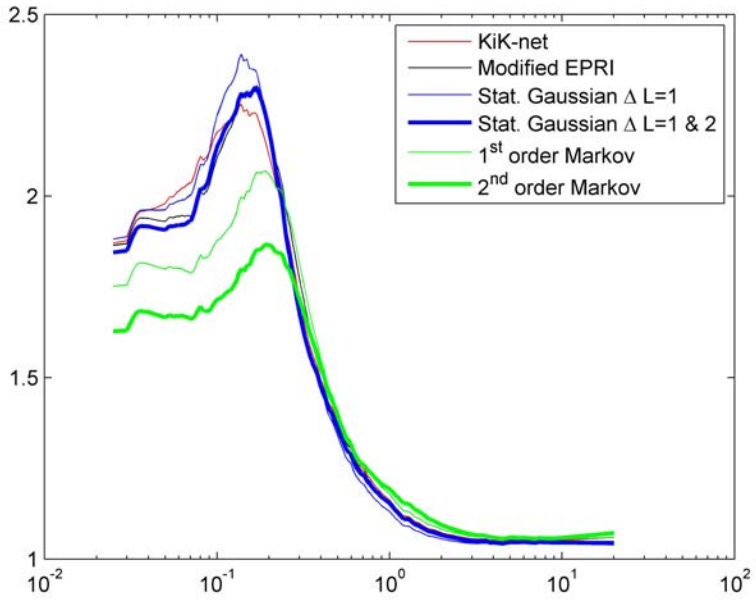


Figure 4.16. Comparison of the median Ratio of Response Spectra for the different Random Field generators. Also shown is the predicted Ratio of Response spectra using the measured shear wave velocity profiles for site class C.

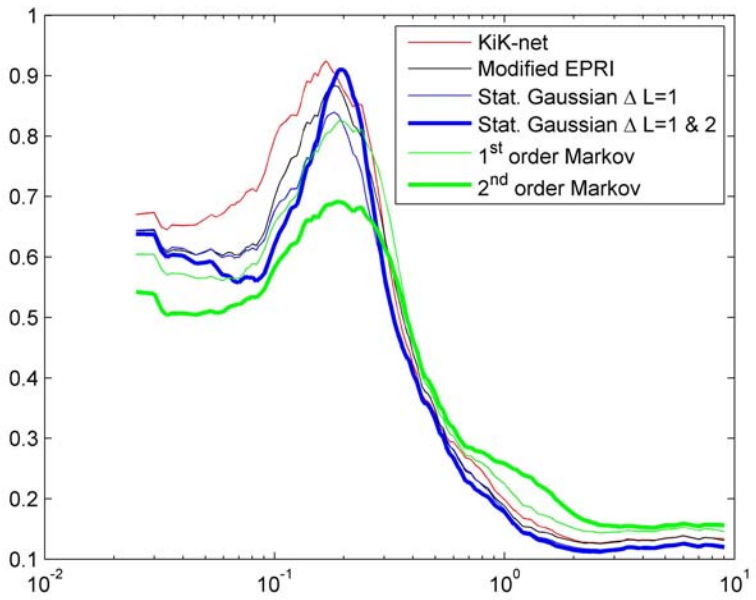


Figure 4.17. Comparison of the standard deviation of Ratio of Response Spectra for the different Random Field generators. Also shown is the predicted standard deviation of Ratio of Response spectra using the measured shear wave velocity profiles for site class C.

4.2 Monte Carlo Simulation Approach

The proposed random profile generators allow for the study of the contribution of profile uncertainty to the uncertainty in ground motions computed from site response analyses. Observe that the site response uncertainty is an epistemic variable in probabilistic seismic hazard analyses; hence it can be potentially lowered if Vs profiles are better characterized. To evaluate how the uncertainty in the site profile affects the predicted site response variability, a Monte Carlo Simulation experiment was conducted.

Using the one layer lag model, 100,000 profiles were generated for sites corresponding to class C (360 m/s $\leq Vs_{30} \leq$ 760 m/s), and D (180 m/s $\leq Vs_{30} \leq$ 360 m/s). For each site class, six sets of 100 profiles out of the 100,000 profiles were selected such that the standard deviation of their Vs_{30} matched preset target values. One of the preset values is the standard deviation of the measured KiK-net profiles for each site class (C and D). The remaining preset values were selected to provide a range of Vs_{30} standard deviations as wide as possible. All 6 sets have equal mean Vs_{30}, 560 m/s for the class C sites, and 270 m/s for the class D sites.

The sampling was done using a beta distribution that (a) has the bounds of the site category (C or D), and (b) has the desired mean and standard deviation. The distribution is fully characterized by parameters "p" and "q" (see Eqs. 4.13 and 4.14), where a and b are the Vs_{30} bounds. With this characterization is possible to fix the mean (μ), and then solve for the values of p and q for a desired standard deviation value (σ).

$$p = \hat{\mu} * \frac{\hat{\mu} * (1 - \hat{\mu})}{\hat{\sigma}^2} - 1 \tag{4.13}$$

$$q = (1 - \hat{\mu}) * \frac{\hat{\mu} * (1 - \hat{\mu})}{\hat{\sigma}^2} - 1 \tag{4.14}$$

with,

$$\hat{\mu} = \frac{\bar{x} - a}{b - a}$$

and

$$\hat{\sigma}^2 = \frac{s^2}{(b - a)^2}$$

where $\bar{x}$ and s are the sample mean and sample standard deviation, respectively. Latin Hypercube Sampling was then used for the selection of 100 sites from the 100,000 sites samples. Using the distribution defined by p and q with the preset standard deviations, the 12 groups (6 for each site class) were created. A subset of the 100 ground motions described in 4.1.5.1 was used to obtain ratios of response spectra for each of the groups. The subset was selected to maintain the same mean and standard deviation of the input response spectra.

Figures 4.18 to 4.20 show the results of varying Vs_{30} standard deviations on the mean of the ratio of response spectra, for class C sites and periods of 0.2, 0.5, and 1.0 seconds. Plus and minus one standard deviation bands are included for reference. Note that the mean values of the calculated ratio of response spectra are relatively insensitive to the variability in Vs_{30}.

Figures 4.21 to 4.23 show the effects of varying standard deviations in Vs_{30} over the resulting ratio of response spectra, for site C class sites. Note that for high frequencies the effect of variability in Vs_{30} is negligible, and that for lower frequencies, there is a direct relation among the variability in site profiles and the calculated site amplification variability.

The results shown and discussed for site C class in the paragraphs above, are shown in Figures 4.24 to 4.29 for the case of site D class. The observation made on the effect of Vs_{30} variability on the ratio of response spectra for class C sites, is also true for class D sites,

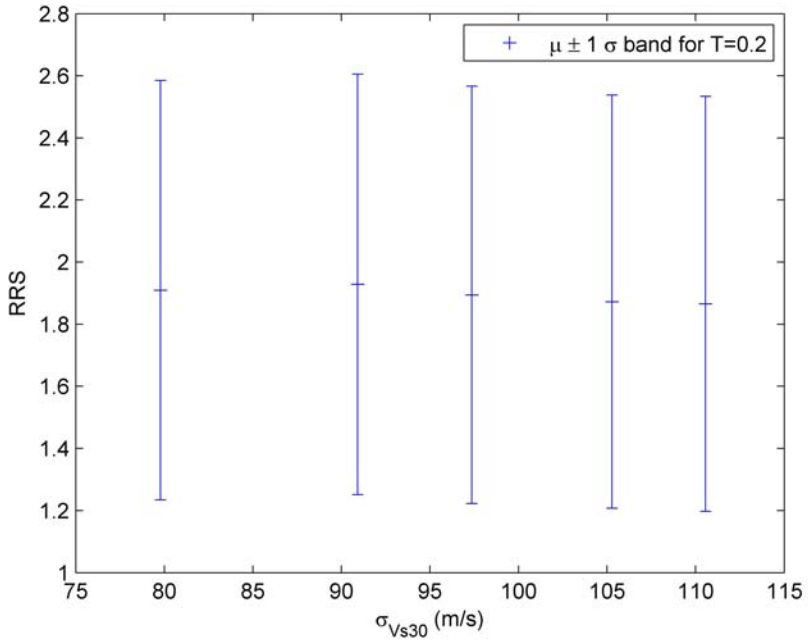


Figure 4.18. Mean of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 0.2 seconds and Site C class.

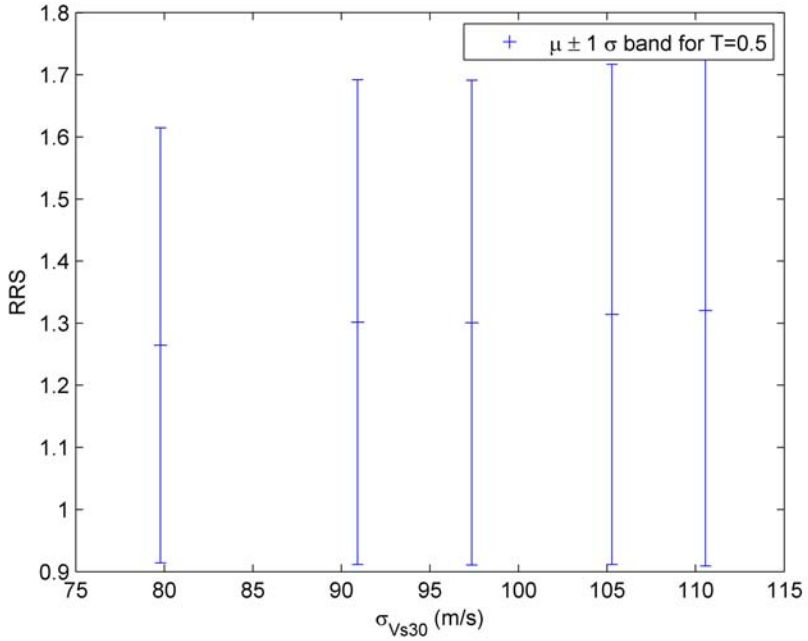


Figure 4.19. Mean of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 0.5 seconds and Site C class.

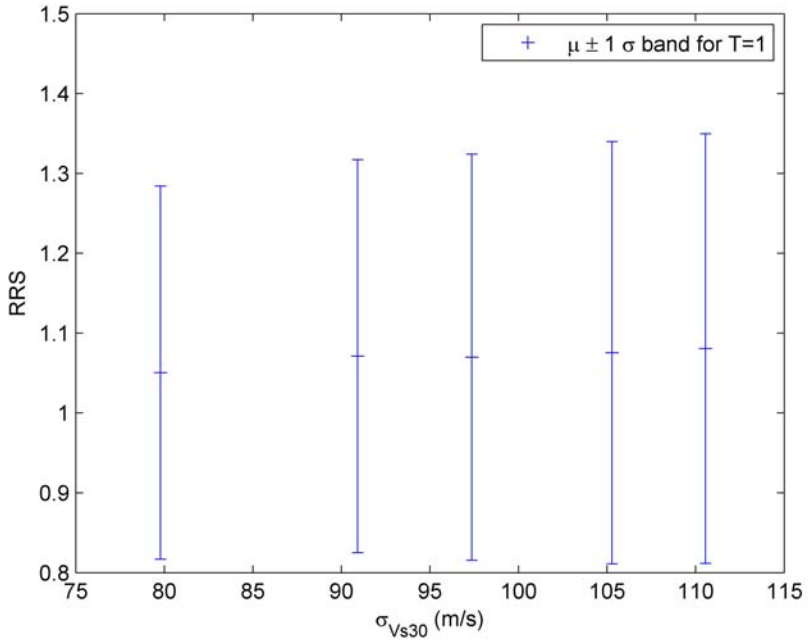


Figure 4.20. Mean of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 1.0 seconds and Site C class.

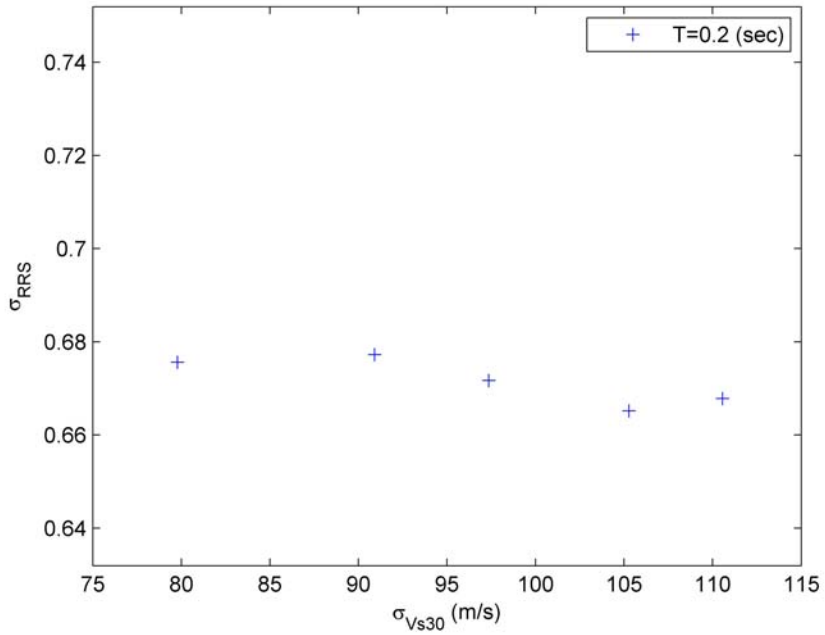


Figure 4.21. Standard deviation of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 0.2 seconds and Site C class.

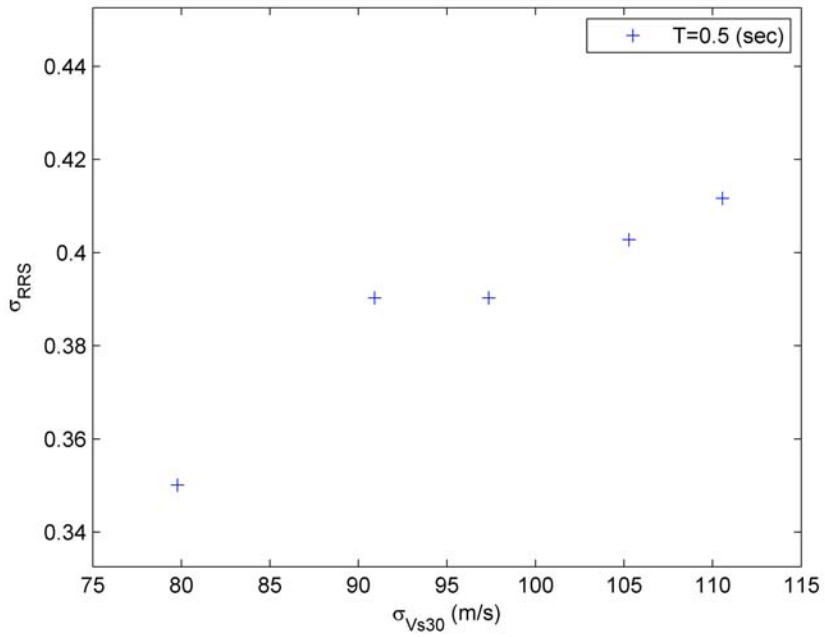


Figure 4.22. Standard deviation of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 0.5 seconds and Site C class.

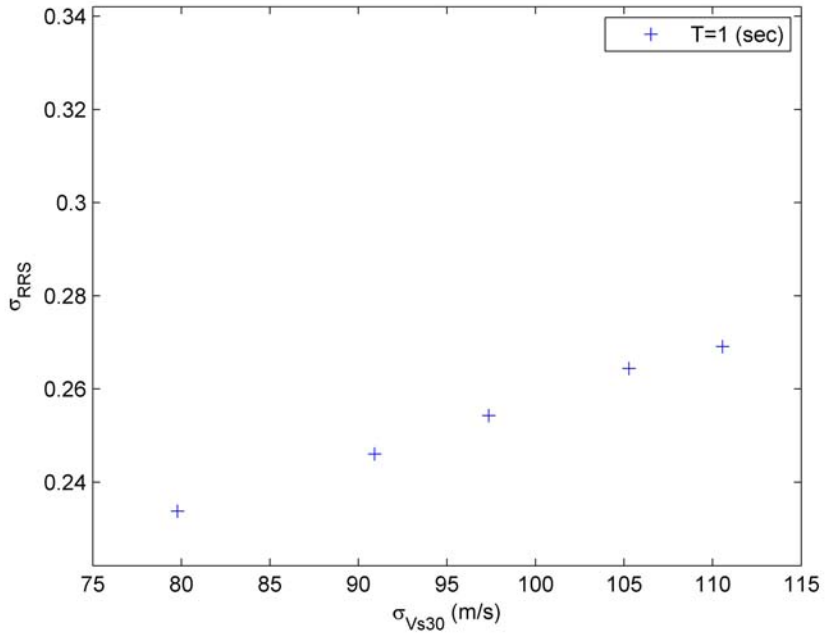


Figure 4.23. Standard deviation of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 1.0 seconds and Site C class.

where the mean value of the ratio of response spectra, a measure of site response, do not show significant variation with varying the standard deviation of Vs_{30}. This observation is contrary to past observations that indicate that uncertainty in Vs profiles reduces mean estimates of surface response spectra when the only uncertainty is variability in Vs values around a baseline Vs profile and stratigraphy is fixed (Rodriguez-Marek et al., 2010; Rathje et al., 2010).

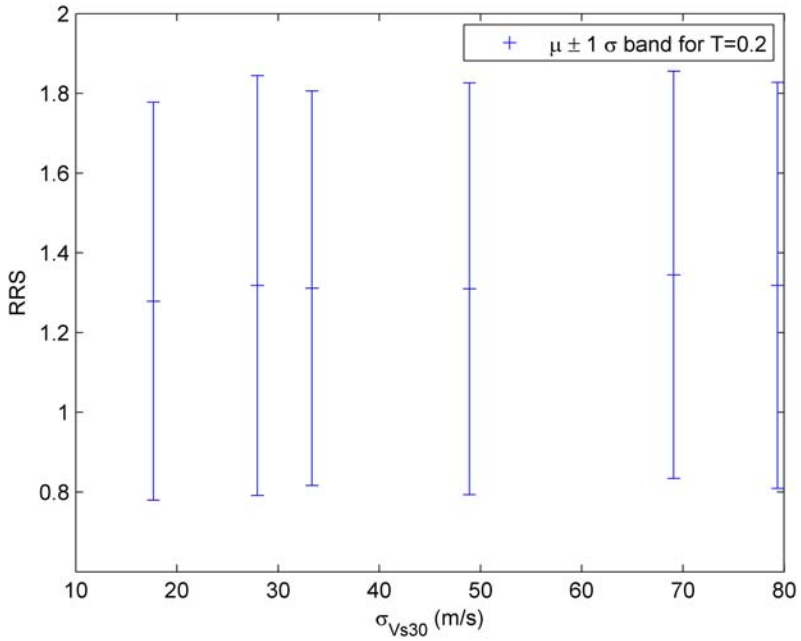


Figure 4.24. Mean of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 0.2 seconds and Site D class.

The simulation performed allows for the study of the Bazzurro and Cornell (2004b) affirmations. The dependence of the values of σ_{lnAF} with lnS_a^r is shown in Figures 4.30 to 4.33 for spectral periods of 0.3 and 1.0 seconds, and sites classes C and D.

Figures 4.30 to 4.33 above show a clear, and often strong, correlation between σ_{lnAF} and lnS_a^r. Bazzurro and Cornell (2004b) assumption is that they are independent.

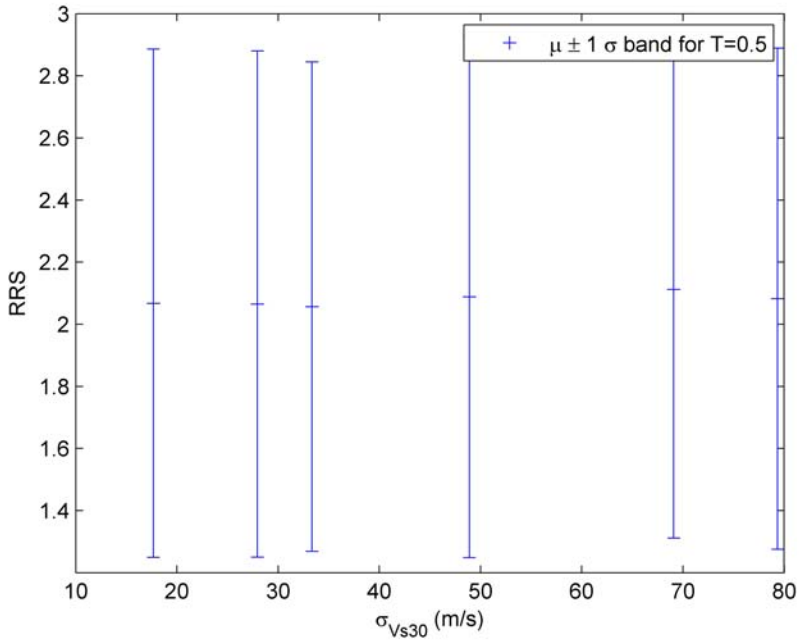


Figure 4.25. Mean of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 0.5 seconds and Site D class.

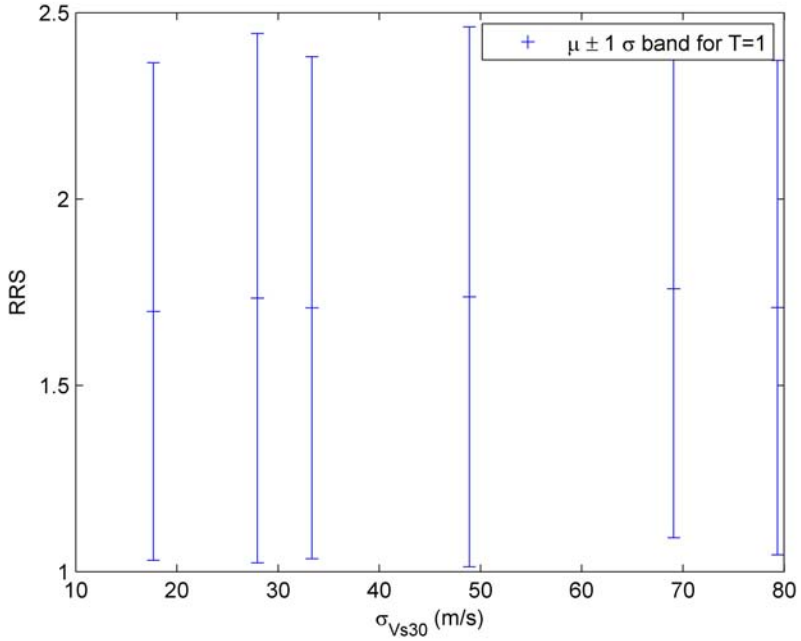


Figure 4.26. Mean of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 1.0 seconds and Site D class.

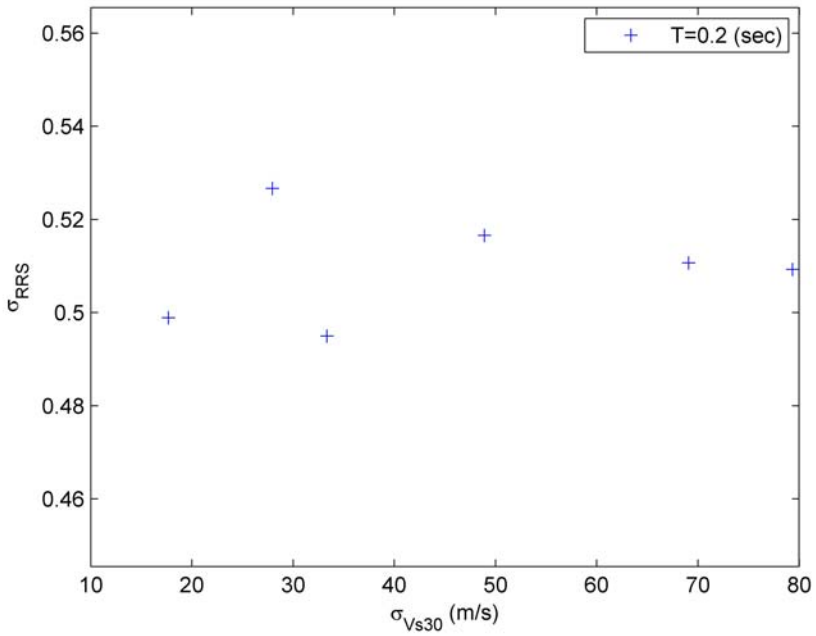


Figure 4.27. Standard deviation of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 0.2 seconds and Site D class.

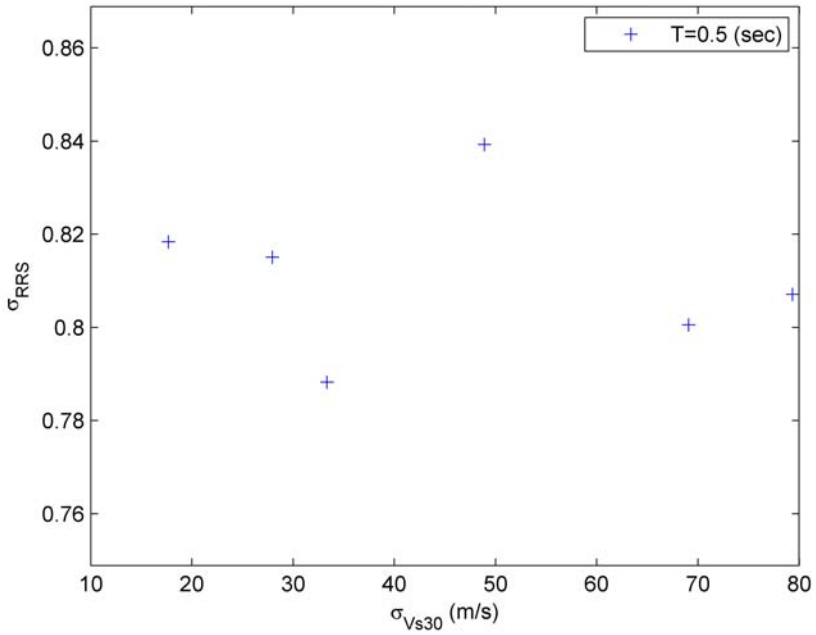


Figure 4.28. Standard deviation of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 0.5 seconds and Site D class.

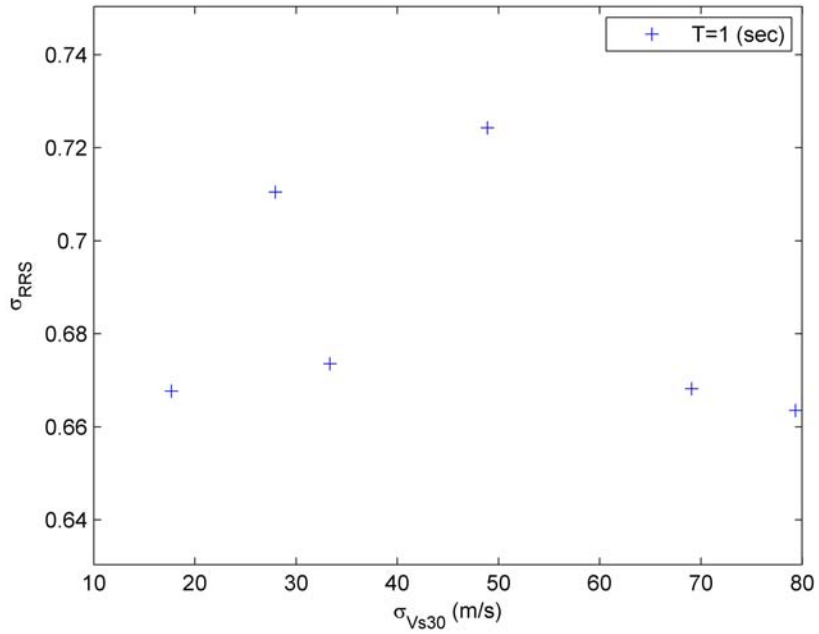


Figure 4.29. Standard deviation of the Ratio of Response Spectra versus standard deviation of Vs_{30}. Results shown for period of 1.0 seconds and Site D class.

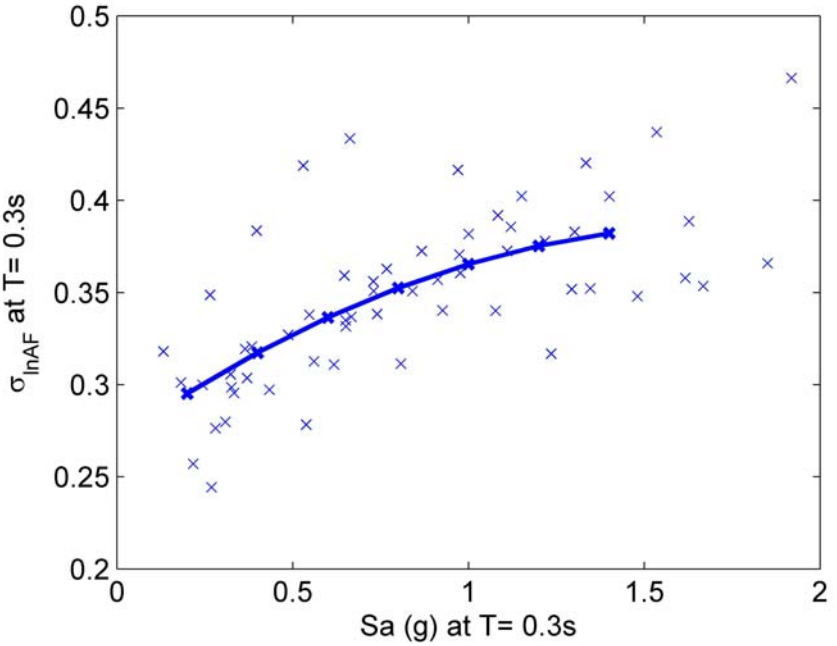


Figure 4.30. Standard deviations of the amplification factor for Site C, and spectral period of 0.3 seconds. Each point is the standard deviation of the amplification factor for 600 artificially generated profiles.

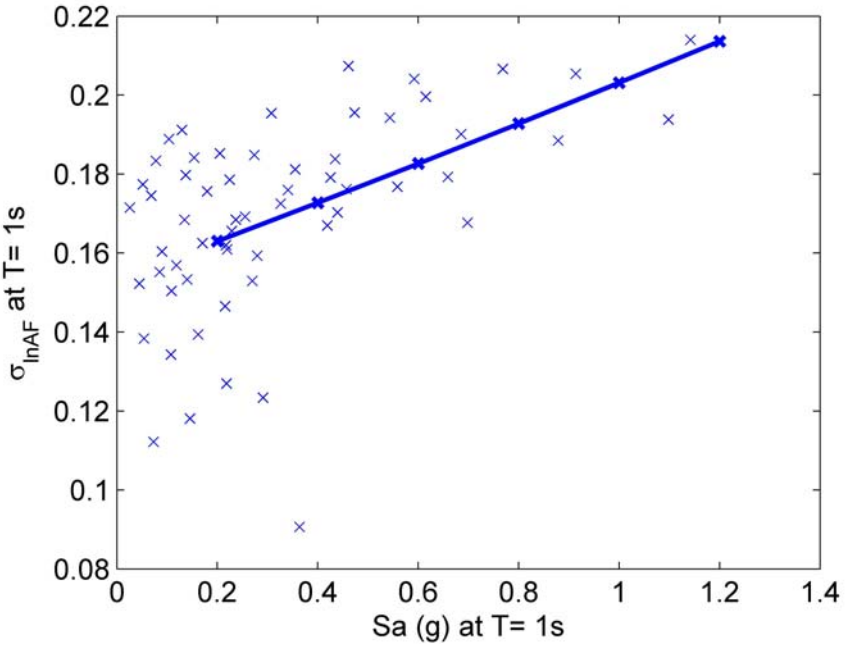


Figure 4.31. Standard deviations of the amplification factor for Site C, and spectral period of 1.0 second. Each point is the standard deviation of the amplification factor for 600 artificially generated profiles.

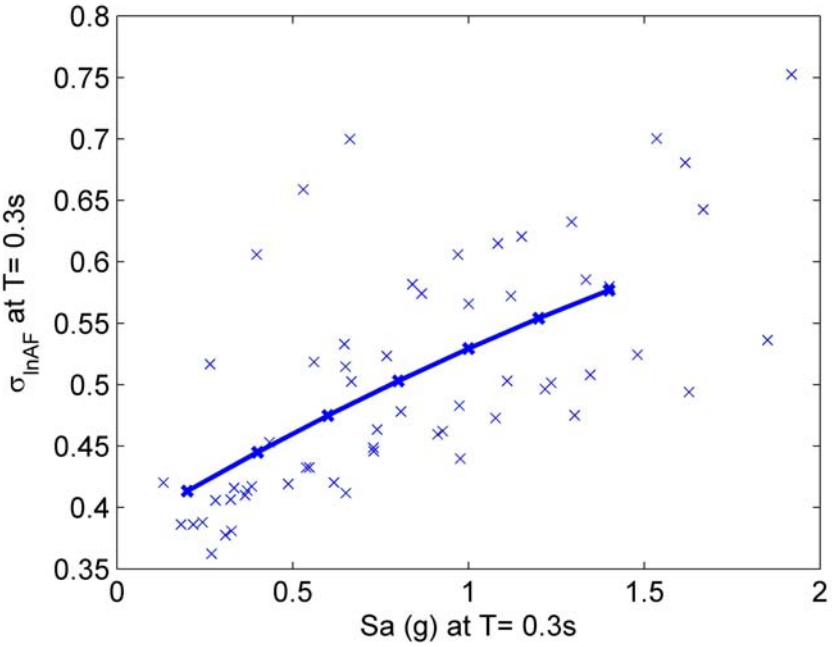


Figure 4.32. Standard deviations of the amplification factor for Site D, and spectral period of 0.3 seconds. Each point is the standard deviation of the amplification factor for 600 artificially generated profiles.

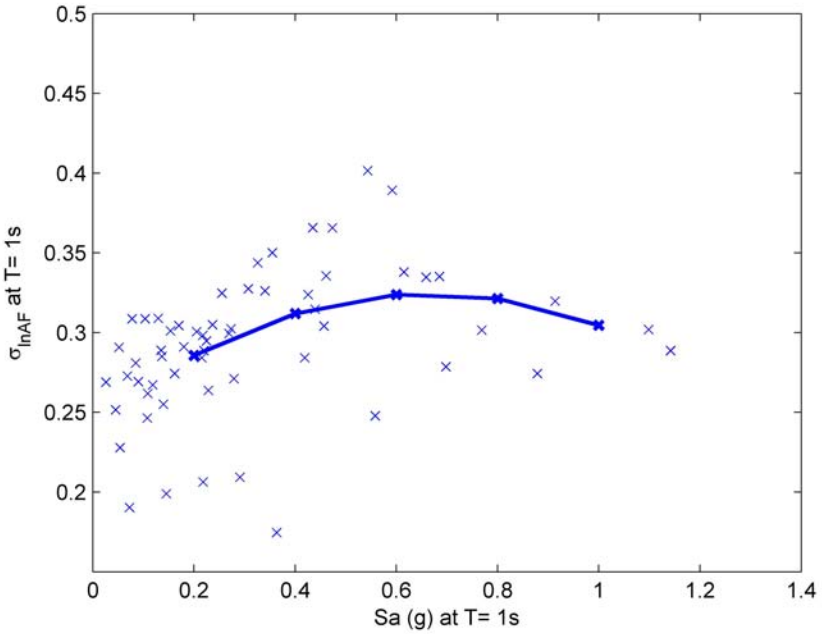


Figure 4.33. Standard deviations of the amplification factor for Site D, and spectral period of 1.0 second. Each point is the standard deviation of the amplification factor for 600 artificially generated profiles.

4.3 Conclusions

This chapter presented a comparison of five models for the generation of shear-wave velocity profiles. The models were calibrated to the measured shear-wave velocity profiles of the KiKnet ground motion database. These models can be used to generate shear wave velocity profiles for site-response analyses. Two types of models were presented, one using Gaussian random fields, and the other set using Markov Chains. Of the Gaussian models, one model is an update of a non-stationary model previously developed by EPRI (1993). The other two Gaussian models make use of the stationary characteristic of the spatial correlation function, when look at as a concatenation of layers rather than a continuous. The two-layer lag stationary Gaussian is a simple model, easy to implement, and shows the best fit of the five proposed models. The results obtained by the stationary Gaussian models can be easily modified to fit other data sets, subsets of data that have particularities, or to characterize the variability in site-specific site response. Profile variability in a site-specific site response analysis would depend on the quality and extent of the site characterization.

While the two-layer lag stationary Gaussian model herein proposed has the limitation of being one dimensional, it gives the engineer the possibility to easily generate random profiles which can be used in forward analyses.

The proposed models are foreseen to be applicable to the modeling of other vertically varying parameters, such as results of Standard Penetration Tests.

Chapter 5

Single Site Variability of Ground Motions: estimates from the KiKnet database

To obtain single-site variability of ground motions, median ground motion estimates and their associated uncertainty are needed, which can be computed using a ground motion prediction equation (GMPE) that uses the ergodic assumption. The ergodic assumption implies the assumption that two different sites with the same parameterization (e.g. same Vs_{30}) have the same median ground motion for an equally parameterized earthquake (e.g. Distance to Fault, Magnitude), and that the variability in the entire database is the same as the variability for a single site. As discussed in Chapter 3 this assumption results in an overestimation of the single-site ground motion variability. This chapter shows this is the case; single-site residuals are calculated for 131 stations that recorded more that 10 events, and estimates of their standard deviations are presented. In addition a further break down of ground motion residuals, as presented previously in Chapter 3 is also studied.

5.1 Data Source Characteristics

The source of strong ground motion data is the KiK-net database as processed by Pousse et al. (2005) and Cotton et al. (2008), and described in Section 2.7. Each station in the KiK-net network has two 3-component accelerographs, one at the surface and another at depth. The instruments have a 24 bit analog-to-digital converter with a sampling frequency of 200 Hz (Fujiwara et al., 2004). All records between 1996 and October 2004 with $M_{JMA} > 4$ have been downloaded. As a preliminary check to avoid subduction related records, only events with depth less than 25 km were analyzed. The M_{JMA} magnitude was converted to seismic moment magnitude using the Fukushima (1996) relationship (Cotton et al., 2008). A visual inspection was performed on ground motion data to check for errors and to keep only the main event if multiple events were recorded in the time series. The signals were band-pass filtered between 0.25 and 25 Hz with a Butterworth filter. This filtering was performed in the time domain with four poles and two passes using SAC2000 routines (Goldstein et al., 2003). Closest distance to the rupture was computed for all recordings. This rupture is assumed to correspond to the hypocentral distance for small to moderate earthquakes or when the source dimensions remain unknown. The magnitude-distance sampling of the database is shown in Figure 5.1. Cotton et al. (2008) state that the longest usable period for the database is 3.0 s. However, some of the spectral accelerations at long periods are lower than the number of decimals used in the database. For that reason, this work is performed only for spectral accelerations less than 1.3 s. For a more detailed description of the data processing, please refer to Pousse et al. (2005) and Cotton et al. (2008).

The magnitude-distance distribution of the data (Figure 5.1) shows that the majority of the records are from earthquakes with magnitudes equal or lower than 6.1, this will have an effect on the regression as it is discussed further in this Chapter. Figure 5.2 shows the time-averaged upper 30 meters shear-wave velocity (Vs_{30}) distribution.

Figure 5.3 shows the distribution of the subsurface instrument depths. Instruments are generally installed on bedrock, sometimes at the soil-bedrock interface, but often are installed at an arbitrary depth. The histogram shows that most of the instruments are located either at 100 meters or 200 meters below the ground surface, for that reason it was chosen to include flags to separate the data recorded at the surface, at 150 meters or shallower depths, and at depths greater than 150 meters. The same approach was taken by Cotton et al. (2008). Figure 5.4 shows the histogram of the shear-wave velocity at the borehole (i.e., on the layer where the borehole instrument is installed). Note that most of the instruments are installed at layers with shear-wave velocities higher than 800 m/s.

Figures 5.5 and 5.6 show the spatial distribution of the stations and epicenters of the recorded earthquakes. Note that there are clusters of events that could enable future studies to separate source, and path effects from the prediction of ground motion intensities and their uncertainties. This approach is not included in this study, but would produce a

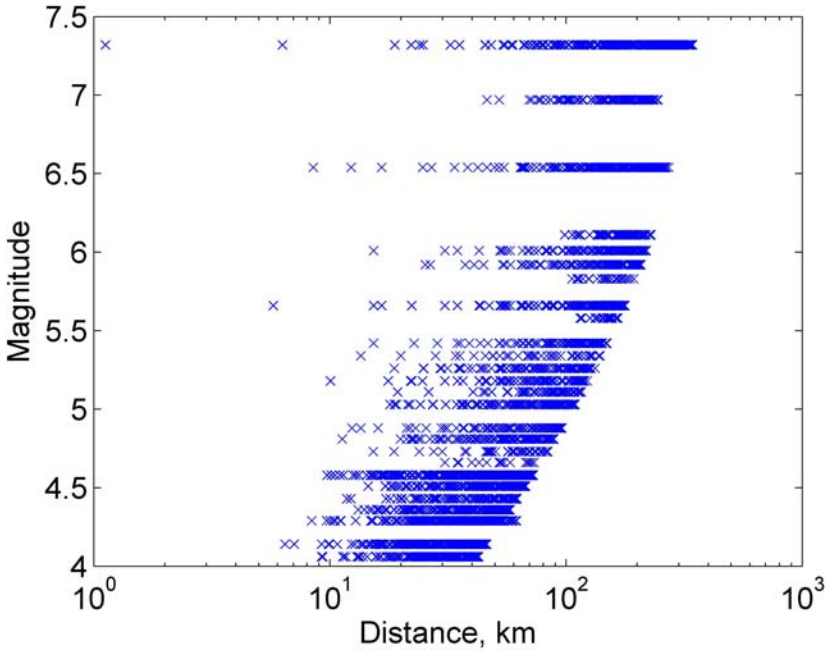


Figure 5.1. Magnitude versus Distance distribution for the KiK-net database

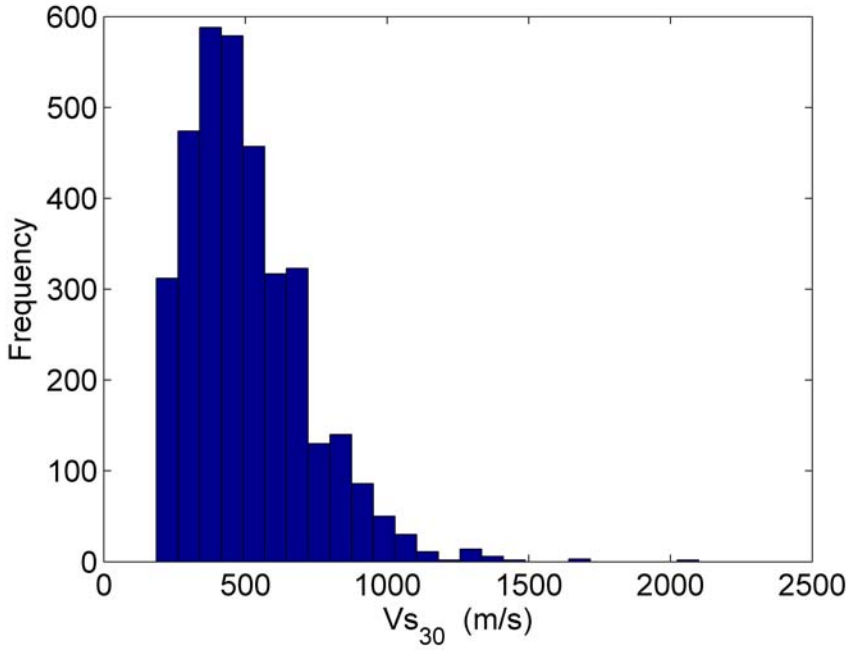


Figure 5.2. Distribution of upper 30 m time-averaged shear-wave velocity (Vs_{30}) for the KiK-net database

reduction in the uncertainty due to the elimination of the source-to-source variability, and the path-to-path variability.

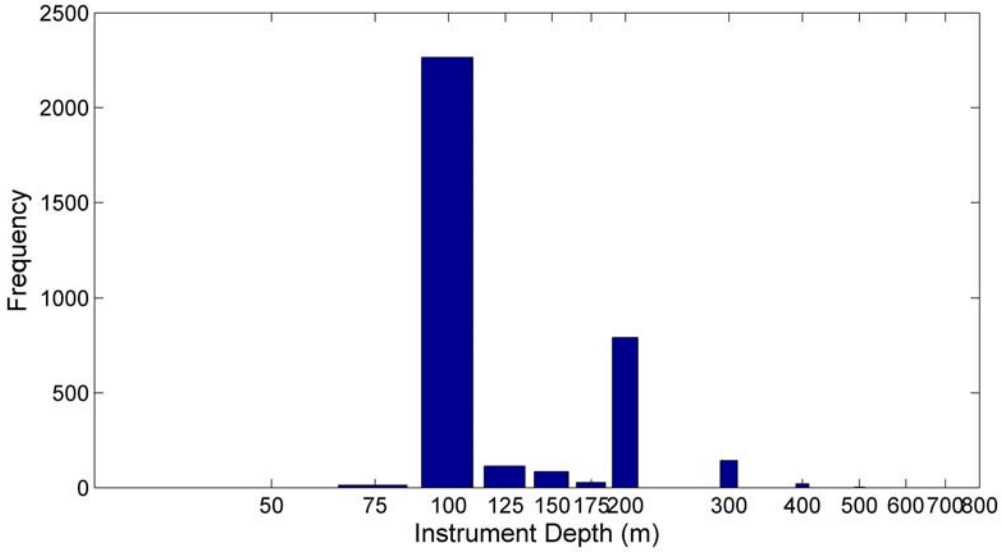


Figure 5.3. Instrument depth distribution for the KiK-net database. Note that most subsurface instruments are located at 100m or 200m below the surface.

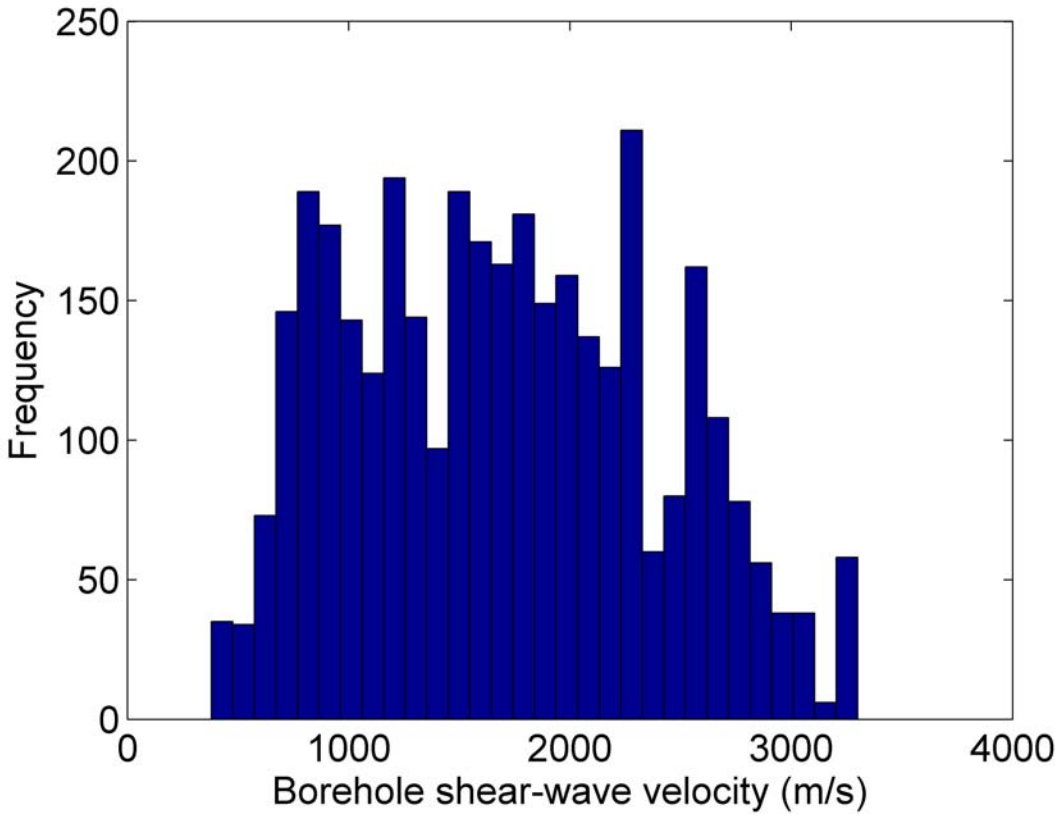


Figure 5.4. Borehole shear-wave velocity histogram.

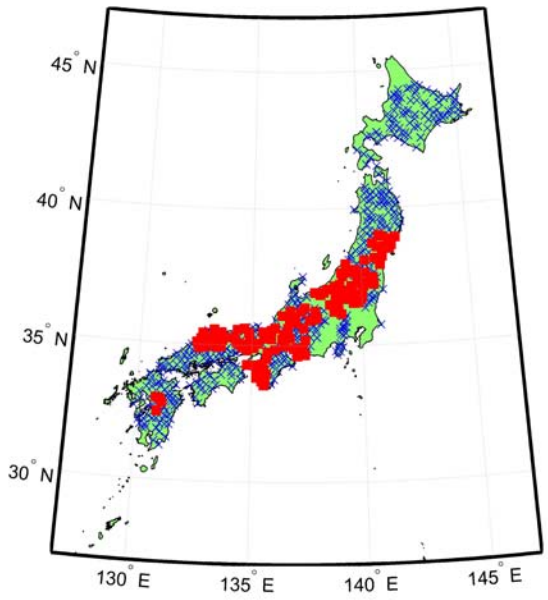


Figure 5.5. KiK-net ground motion station locations. Shown in red are the 46 station for which more that 15 records are available. Stations in blue are considered in the regression analysis.

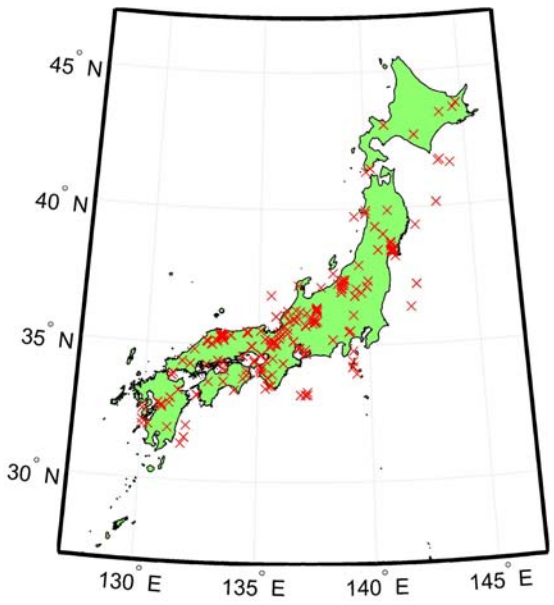


Figure 5.6. KiK-net database epicenters for the recorded earthquakes included in the GMPE. Note that each of these events were recorded by surface and at depth instruments.

5.2 GMPE for Surface and Borehole

The ground motion prediction models herein developed correspond to the median estimates, and their variability, of the natural logarithm of the pseudo-spectral acceleration at the ground surface (denoted by superscript G), at the bedrock (denoted by a superscript B), and a combined model that is constrained using simultaneously the data for surface and borehole.

All three models were developed following the functional form used by Boore and Atkinson (2008). It was desired to use one of the Next Generation Attenuation (NGA), PEER (2010), models as a basis to the functional form of the GMPE to reflect the state of the art in ground motion prediction. The selection of the functional form was made on the basis of its ability to be constrained by data (other models require intensive seismological modeling, along with recorded data, to constrain their model parameters). This is reflected in the simplicity of the Boore and Atkinson (2008) compared to most of the other NGA models.

The variables to be predicted by the model (i.e. the response variables) correspond to the peak ground acceleration, PGA, and pseudo-spectral accelerations at 5% damping. These variables are taken as the geometric mean of the horizontal components of the recorded ground motions.

The predictor variables of the model (i.e. the independent variables) are moment magnitude, denoted by Mw, closest distance to the fault rupture, denoted by R_{RUP}, time-averaged upper 30 meters shear-wave velocity, denoted by Vs_{30}, depth for which shear-wave velocity of 800 m/s is reached, denoted by $h800$, and the shear-wave velocity at the bedrock, denoted by $Vshole$. The shear-wave velocity data used to compute Vs_{30} and $Vshole$ is described in Chapter 4.

The general form of the model, similar to most available ground motion prediction models, is given by

$$y = \mu + \delta W_{es} + \delta B_e \tag{5.1}$$

where y is the natural logarithm of the estimated ground motion parameter or response variable, μ is the median ground motion model, which is a function of the predictor variables Vs_{30}, R_{RUP}, Vs_{30}, $h800$, $Vshole$, where y is being predicted (i.e. surface, borehole located shallower than 150 meters, or borehole located deeper than 150 meters). δB_e is the between-event residual, or the inter-event residual, which is defined as the misfit of the mean of the observation for one particular earthquake (or event) from the median ground motion model. δW_{es} is the intra-event residual, that is the difference between an individual observation and the event corrected median estimate (i.e. $\delta W_{es} = y - \mu - \delta B_e$).

5.2.1 Combined Model

The ground motion model for the combined data set is the most general of the three models as it allows to predict the ground motion intensities at the ground surface and at the borehole with a single predictive equation. The general form for the median estimate of this model is given by

$$\mu^A_{med} = F_m + F_d + F_{site} * Surf_{flag} + F_{100} * S100_{flag} + F_{200} * S200_{flag} \tag{5.2}$$

where,

$$
\begin{aligned}
F_m &= e_1 + e_5 * (Mw - M_h) + e_6 * (Mw - M_h)^2 \quad \text{for } Mw \leq M_h \\
F_m &= e_1 + e_7 * (Mw - M_h) \qquad\qquad\qquad\qquad\quad \text{for } Mw \geq M_h
\end{aligned}
\tag{5.3}
$$

$$
\begin{aligned}
&F_d = (c_1 + c_2 * (Mw - M_{ref})) * log(R/R_{ref}) + c_3 * (R - R_{ref}); \\
&R = \sqrt{R_{RUP}^2 + h^2}
\end{aligned}
\tag{5.4}
$$

$$
F_{site} = blin * ln(Vs_{30}/V_{ref}) + bh800 * ln(h800/h_{ref}) \tag{5.5}
$$

$$
F_{100} = a_{100} + b_{100} * ln(Vs_{30}/V_{ref}) + c_{100} * ln(Vshole/Vshole_{ref}) \tag{5.6}
$$

$$
F_{200} = a_{200} + b_{200} * ln(Vs_{30}/V_{ref}) + c_{200} * ln(Vshole/Vshole_{ref}) \tag{5.7}
$$

where, M_{ref} is a reference value for Mw equal to 4.5, R_{ref} is a reference distance to the rupture plane equal to 1 kilometer, V_{ref} is a reference value of Vs_{30} equal to 760 (m/s), h_{ref} is a reference value for the depth at which a profile reaches a shear-wave velocity of 800 (m/s) equal to 60 (m), and $Vshole_{ref}$ is a reference value for the shear-wave velocity at bedrock equal to 3000 (m/s). These reference values do not affect the prediction performance of the model, the reason to use them is to provide a reference for which the model becomes a constant. The constants c_1, c_2, c_3, e_1, e_5, e_6, e_7, $blin$, $bh800$, a_{100}, b_{100}, c_{100}, a_{200}, b_{200}, c_{200}, h, and M_h are model parameters to be determined during the regression analysis.

Note that the subscript A in Equation 5.2 is used to denote that the median predictor uses simultaneously the borehole and surface data. However, the model predicts ground motions either at the ground surface (G) or at the borehole (B) for the same location.

Modeling of the Residuals The inter (δB_e) and intra-event residuals (δW_{es}) are modeled by a normal distribution with mean zero and variance τ^2, and ϕ^2, respectively. In the regression analysis ϕ^2, and τ^2 were allowed to be a function of magnitude. The functional form was restricted to a constant value for magnitudes less than 5, another constant value for magnitudes greater than 6.5, and linear interpolation between magnitudes 5 and 6.5. To allow for an adequate constrain of the inter-event residuals (δB_e), only earthquakes that were recorded at 5 or more stations were considered.

Regression Analysis Methodology The determination of the parameters begins with the determination of c_3, the term that controls the curvature of the distance attenuation term, F_d (see Equation 5.4), at large distances. The data set used to constrain c_3 was restricted to those events with more than 100 records. First the slope of F_d, c_1, was fixed to a value between -0.2, and -1.1 (the slope term c_1 has to be negative for the model to produce lower intensities with greater distances, which is the logical outcome), c_2 was fixed to zero (i.e. making F_d magnitude independent). Then c_3 along with the pseudo-depth coefficient h, were obtained by maximum likelihood estimation method. The slope term c_1 was varied within the specified range, and the value that maximized the likelihood function was used to select the final coefficients for c_3, and h. The process was repeated for each spectral period.

A "magnitude hinge" term, M_h, that is used as a flag in the magnitude attenuation term, F_m, was defined to separate the magnitude range where magnitude scaling changes from quadratic to linear (Equation 5.3). A negative value for the linear term for Mw $\geq M_h$ (e_7) would indicate a reduction of ground motion with an increase in magnitude. The reasons behind the observation of greater magnitude earthquakes producing less pseudo-spectral acceleration than lower magnitude ones, known as oversaturation, are not clear. Anderson (2010) speculates that a possible explanation is that large earthquakes occur along the same surfaces, hence having smoother interfaces. For the KiK-net database oversaturation

is observed at all periods, but because the amount of data from high magnitude earthquakes is limited, and lack of knowledge, the linear term e_7 was not allowed to take negative values during the regression. Boore and Atkinson (2008) followed a similar approach in limiting oversaturation. The value for M_h was chosen by inspection for all spectral periods.

The rest of the parameters were obtained using random effects method (Searle, 1971; Abrahamson and Youngs, 1992). The method delivers both the median model, and the statistics for the inter- and intra-event residuals. The steps of the iterative process to obtain, ϕ^2, τ^2, and the model parameters θ can be summarized as:

1. Calculate model parameters, θ, using maximum likelihood. Set the inter-event residuals δB_e to zero.

2. With the model parameters, maximize the random effects likelihood function (see Searle, 1971) to obtain ϕ^2 and τ^2.

3. Compute the inter-event residuals for each event as given by

$$\delta B_e = \frac{\tau^2 * \sum_{s=1}^{n_e} y_{es} - \mu_{es}}{\tau^2 * n_e + \phi^2} \tag{5.8}$$

4. Subtract the estimate of each inter-event residual from the observed ground motion. This results in ground motion corrected for event-to-event variability.

5. Calculate model parameters, θ, using maximum likelihood for the corrected ground motion residuals. Repeat steps 2 to 5 until the likelihood obtained in step 2 is maximized.

The random effects regression computes the model parameters (θ) and 4 parameters for the magnitude dependent standard deviations of the within and between residuals. Model parameters, for the distance and magnitude attenuation terms (F_d and F_m) of the combined

model, are summarized in Table 5.1, and the parameters for the site terms (F_{site},F_{100}, and F_{200}) are shown in Table 5.2. The parameter e_7 is not included in the table because it took a value of zero during the regression.

Table 5.1. Model Parameters for the Distance and Magnitude terms of the Combined Model

Period	c_1	c_2	c_3	e_1	e_5	e_6
PGA	-1.2534	0.4271	-0.0140	-0.0663	-0.5997	-0.5012
0.0384	-1.2928	0.4089	-0.0140	0.3986	-0.4943	-0.4423
0.0484	-1.3005	0.3719	-0.0140	0.9018	-0.2267	-0.3538
0.0582	-1.2725	0.3373	-0.0140	1.2277	-0.0006	-0.2963
0.0769	-1.2423	0.3270	-0.0140	1.4185	0.0608	-0.3109
0.0844	-1.2267	0.3221	-0.0140	1.4818	0.2850	-0.2211
0.0970	-1.2053	0.3169	-0.0140	1.4496	0.3325	-0.2346
0.1167	-1.1925	0.3135	-0.0138	1.3627	0.2952	-0.3234
0.1472	-1.1937	0.3101	-0.0131	1.1992	0.3159	-0.3877
0.1691	-1.2169	0.3181	-0.0126	1.1069	0.2759	-0.4302
0.2036	-1.2454	0.3343	-0.0119	0.8915	0.1408	-0.5095
0.2340	-1.2639	0.3267	-0.0113	0.8740	0.2653	-0.5054
0.3090	-1.2822	0.3148	-0.0100	0.6841	0.3889	-0.5591
0.3551	-1.2927	0.3136	-0.0092	0.5551	0.0850	-0.6291
0.3896	-1.3016	0.3135	-0.0087	0.3939	-0.3426	-0.7271
0.4274	-1.3038	0.3088	-0.0082	0.3443	-0.1975	-0.6910
0.4690	-1.3064	0.3010	-0.0076	0.3107	-0.0902	-0.6799
0.5913	-1.3224	0.2829	-0.0062	0.1837	0.0808	-0.6842
0.7456	-1.3606	0.2628	-0.0049	0.1503	0.2970	-0.6804
0.8180	-1.3813	0.2577	-0.0043	0.1341	0.3733	-0.6725
0.9401	-1.3975	0.2424	-0.0036	0.1483	0.6130	-0.6271
1.3622	-1.4495	0.2370	-0.0020	-0.2017	0.8621	-0.5687

The intra-event (ϕ) and inter-event (τ) standard deviations resulting from the random effects regression are shown in Table 5.3. The inter-event standard deviation (τ) has unusually large values for large magnitudes. This results from a consistent bias of the predictive model as a result of limiting oversturation. The last column of Table 5.3 shows an estimate of (τ) computed using the inter-event residuals (δ_{B_e}) corrected for the observed bias using a linear term. This is equivalent to allowing for a negative e_7 term, except that the interdependency with the intra-event residuals is not captured. The values of (τ) computed by

Table 5.2. Model Parameters for the Site terms of the Combined Model

Period	$blin$	$bh800$	a_{100}	a_{200}	b_{100}	b_{200}	c_{100}	c_{200}
PGA	-0.4665	-0.1801	-1.4372	-1.6518	-0.0269	-0.1884	-0.2666	-0.3793
0.0384	-0.3756	-0.2219	-1.3309	-1.5398	-0.0184	-0.2502	-0.1435	-0.2053
0.0484	-0.2871	-0.2420	-1.2680	-1.5018	-0.0301	-0.2628	-0.0716	-0.1511
0.0582	-0.2124	-0.2303	-1.3806	-1.6006	-0.0323	-0.2229	-0.0943	-0.1675
0.0769	-0.2658	-0.2060	-1.5810	-1.7920	0.0391	-0.1961	-0.2066	-0.2871
0.0844	-0.3207	-0.2000	-1.6083	-1.8197	0.0520	-0.1924	-0.2591	-0.3398
0.0970	-0.4062	-0.1844	-1.6143	-1.8366	0.0692	-0.1898	-0.3091	-0.4050
0.1167	-0.5213	-0.1767	-1.6022	-1.7851	0.0508	-0.1513	-0.3675	-0.4637
0.1472	-0.6892	-0.1256	-1.5497	-1.6843	-0.0259	-0.0765	-0.4355	-0.5998
0.1691	-0.7431	-0.0964	-1.4863	-1.6025	-0.0206	0.0000	-0.4488	-0.6506
0.2036	-0.8023	-0.0563	-1.3907	-1.5946	-0.0235	-0.0753	-0.4565	-0.6952
0.2340	-0.8370	-0.0392	-1.3202	-1.5249	-0.0543	-0.0553	-0.4613	-0.7024
0.3090	-0.8629	0.0427	-1.1687	-1.4151	-0.0567	-0.0034	-0.4593	-0.7573
0.3551	-0.8394	0.0765	-1.0886	-1.3765	-0.0900	0.0059	-0.4044	-0.7769
0.3896	-0.8161	0.0972	-1.0337	-1.3443	-0.1057	-0.0095	-0.3754	-0.7479
0.4274	-0.7922	0.1268	-0.9834	-1.3163	-0.1310	-0.0220	-0.3528	-0.7326
0.4690	-0.7696	0.1623	-0.9402	-1.3116	-0.1451	-0.0214	-0.3395	-0.7804
0.5913	-0.6798	0.2291	-0.8595	-1.2244	-0.1296	-0.0181	-0.3465	-0.7723
0.7456	-0.6337	0.2592	-0.8001	-1.0314	-0.1058	0.0533	-0.3996	-0.6770
0.8180	-0.6028	0.2797	-0.7991	-1.0049	-0.0847	0.0645	-0.4442	-0.6675
0.9401	-0.5682	0.3004	-0.7968	-0.9580	-0.0778	0.0604	-0.4948	-0.6448
1.3622	-0.5063	0.3041	-0.8443	-0.8703	-0.1381	0.0310	-0.6149	-0.6923

reducing the bias are significantly lower than those that ignore it, in particular for short periods.

Table 5.3. Standard Deviations of the Residuals from the Combined Model

Period	ϕ for $Mw < 5$	ϕ for $Mw > 6.5$	τ for $Mw < 5$	τ for $Mw > 6.5$	τ^* for $Mw > 6.5$
PGA	0.6293	0.6202	0.4929	0.9164	0.4981
0.0384	0.6252	0.6223	0.5018	0.8963	0.5428
0.0484	0.6321	0.6315	0.5188	0.8533	0.6152
0.0582	0.6469	0.6457	0.5193	0.8599	0.6654
0.0769	0.6623	0.6658	0.5063	0.9332	0.6940
0.0844	0.6697	0.6706	0.5052	0.9433	0.6991
0.0970	0.6846	0.6790	0.5104	0.9274	0.6940
0.1167	0.6878	0.6801	0.5084	0.9242	0.6676
0.1472	0.6893	0.6812	0.4917	0.9348	0.6204
0.1691	0.6847	0.6751	0.4879	0.9446	0.5883
0.2036	0.6772	0.6680	0.5044	0.8520	0.5420
0.2340	0.6648	0.6563	0.5138	0.7751	0.5083
0.3090	0.6436	0.6332	0.5389	0.6936	0.4468
0.3551	0.6352	0.6262	0.5431	0.6914	0.4123
0.3896	0.6272	0.6228	0.5476	0.6963	0.3841
0.4274	0.6170	0.6172	0.5516	0.6743	0.3744
0.4690	0.6049	0.6067	0.5500	0.6638	0.3696
0.5913	0.5778	0.5878	0.5695	0.5611	0.3546
0.7456	0.5535	0.5668	0.5821	0.4478	0.3034
0.8180	0.5478	0.5666	0.5773	0.4244	0.2823
0.9401	0.5421	0.5673	0.5750	0.3840	0.2498
1.3622	0.5374	0.5614	0.5886	0.3711	0.2097

[*]Allowing for oversaturation at large magnitudes

As a closing for the presentation of the combined model it is important to point out that the model considers single magnitude and distance terms both for the surface and the borehole data. This characteristic is desired from a phenomenological point of view because source- and path- terms should be independent of near-surface layering. Also implicit in the methodology is that inter-event residuals (δB_e), are also equal for surface and borehole. This is an important property of the model as it allows to separate site response effects. Therefore, all the difference between surface and borehole is captured by the site terms (F_{site}, F_{100}, and F_{200}).

5.2.2 Surface Model

The model for the surface follows the same philosophy adopted for the combined model. The main difference is that it is constrained using purely ground surface records, which is what most GMPE are based upon. The functional form for the median model is given by Equation 5.9.

$$\mu^G = F_m + F_d + F_{site} \tag{5.9}$$

where, the three terms F_m, F_d, and F_{site} are the equal to those described for the combined model. Note that $Surf_{flag}$ is not needed in this model as all the dependent variables (i.e. the data) comes from ground surface records. Table 5.4 summarizes the results of the regression for the magnitude and distance terms. Table 5.5 summarizes the site term parameters for the surface model.

Table 5.6 shows the standard deviations for the inter- and intra-event residuals. Note that, similarly to what was observed in the combined model, the elimination of the bias due to oversaturation greatly reduces the variability in large magnitudes of the inter-event random variable (i.e. τ), and that although it affects all spectral periods the greater effect is observed at low periods.

Table 5.4. Model Parameters for the Distance and Magnitude terms of the Surface Model

Period	c_1	c_2	c_3	e_1	e_5	e_6
PGA	-1.1630	0.3493	-0.0140	0.1103	0.0390	-0.3150
0.0384	-1.1766	0.3333	-0.0140	0.4665	0.1189	-0.2666
0.0484	-1.1855	0.3101	-0.0140	0.8874	0.3145	-0.1819
0.0582	-1.1453	0.2785	-0.0140	1.1572	0.5749	-0.0895
0.0769	-1.1246	0.2649	-0.0140	1.3796	0.5904	-0.1321
0.0844	-1.1175	0.2644	-0.0140	1.4104	0.6435	-0.1290
0.0970	-1.1021	0.2647	-0.0140	1.3724	0.6425	-0.1647
0.1167	-1.0987	0.2614	-0.0138	1.3265	0.6495	-0.2286
0.1472	-1.1454	0.2740	-0.0131	1.2530	0.6277	-0.2831
0.1691	-1.1919	0.2901	-0.0126	1.1896	0.5265	-0.3448
0.2036	-1.2486	0.3144	-0.0119	1.0459	0.4153	-0.3973
0.2340	-1.2803	0.3172	-0.0113	1.0110	0.4559	-0.4235
0.3090	-1.3074	0.3111	-0.0100	0.7873	0.5632	-0.4542
0.3551	-1.3239	0.3109	-0.0092	0.6854	0.2588	-0.5355
0.3896	-1.3424	0.3113	-0.0087	0.6084	-0.0382	-0.5837
0.4274	-1.3554	0.3119	-0.0082	0.5037	-0.0969	-0.6259
0.4690	-1.3733	0.3088	-0.0076	0.4889	-0.0486	-0.6315
0.5913	-1.3841	0.2827	-0.0062	0.4180	0.2373	-0.5873
0.7456	-1.4126	0.2628	-0.0049	0.3764	0.5342	-0.5361
0.8180	-1.4287	0.2545	-0.0043	0.3652	0.6589	-0.5074
0.9401	-1.4477	0.2392	-0.0036	0.3623	0.8620	-0.4711
1.3622	-1.5023	0.2319	-0.0020	0.0091	1.1450	-0.3811

Table 5.5. Model Parameters Site term for the Surface Model

Period	$blin$	$bh800$
PGA	-0.4804	-0.1654
0.0384	-0.3701	-0.2024
0.0484	-0.2796	-0.2286
0.0582	-0.2058	-0.2206
0.0769	-0.2615	-0.1961
0.0844	-0.3163	-0.1853
0.0970	-0.4045	-0.1668
0.1167	-0.5276	-0.1609
0.1472	-0.7093	-0.1146
0.1691	-0.7812	-0.0855
0.2036	-0.8648	-0.0604
0.2340	-0.9034	-0.0491
0.3090	-0.9553	0.0289
0.3551	-0.9197	0.0677
0.3896	-0.8900	0.0902
0.4274	-0.8507	0.1239
0.4690	-0.8123	0.1609
0.5913	-0.6918	0.2333
0.7456	-0.6203	0.2671
0.8180	-0.5988	0.2823
0.9401	-0.5746	0.2987
1.3622	-0.5136	0.2956

Table 5.6. Standard Deviations of the Residuals from the Surface Model

Period	ϕ for $Mw < 5$	ϕ for $Mw > 6.5$	τ for $Mw < 5$	τ for $Mw > 6.5$	τ^* for $Mw > 6.5$
PGA	0.6858	0.6809	0.4391	0.7987	0.5582
0.0384	0.6778	0.6853	0.4325	0.7954	0.6014
0.0484	0.6722	0.6897	0.4362	0.7773	0.6708
0.0582	0.6869	0.7046	0.4369	0.7722	0.7124
0.0769	0.7350	0.7528	0.4429	0.7968	0.7295
0.0844	0.7473	0.7603	0.4515	0.7891	0.7329
0.0970	0.7715	0.7752	0.4714	0.7712	0.7313
0.1167	0.7718	0.7739	0.4819	0.7623	0.6975
0.1472	0.7727	0.7772	0.4639	0.7715	0.6445
0.1691	0.7617	0.7623	0.4725	0.8026	0.6090
0.2036	0.7489	0.7492	0.4887	0.7846	0.5664
0.2340	0.7341	0.7312	0.4974	0.7596	0.5198
0.3090	0.7008	0.6907	0.5170	0.6857	0.4503
0.3551	0.6876	0.6799	0.5274	0.6863	0.4144
0.3896	0.6808	0.6770	0.5337	0.7059	0.3856
0.4274	0.6677	0.6691	0.5374	0.6723	0.3770
0.4690	0.6530	0.6560	0.5311	0.6506	0.3703
0.5913	0.6237	0.6311	0.5289	0.5745	0.3499
0.7456	0.6005	0.6073	0.5321	0.4904	0.3039
0.8180	0.5958	0.6084	0.5240	0.4633	0.2871
0.9401	0.5850	0.6055	0.5187	0.4339	0.2569
1.3622	0.5666	0.5827	0.5173	0.3949	0.2212

[*]Allowing for oversaturation at large magnitudes

5.2.3 Borehole Model

The model for borehole is constrained with data from at-depth instruments, and although the instruments are typically on bedrock occasionally they are at an arbitrary depth. Hence, it is herein called "borehole model" rather than "bedrock model". The regression process is analogous to the process for the combined and surface models. The parameters for the distance attenuation term c_3, and h were taken as the same as the ones calculated for surface data, as was the parameter M_h for the magnitude attenuation term. This does not mean that the distance and magnitude terms are equal, as the remaining parameters were obtained using solely from borehole data.

The functional form for the median of the borehole ground motion prediction equation is give by Equation 5.10.

$$\mu^B = F_m + F_d + F_{100} + F_{200} \tag{5.10}$$

where, all the terms are equal to those proposed for the combined model, with the difference that this model does not include an F_{site} term, which means that all site effects are captured by F_{100} and F_{200}, to reflect the fact that site response occurs in the near surface region. The parameter e_1 is a constant that shifts the magnitude attenuation model, given that the model has other constant parameters, the necessary shift can be captured by other parameters. The value of e_1 was fixed to the value obtained from the surface regression in order to facilitate comparison of the magnitude attenuation terms. The parameters for the magnitude and distance attenuation terms for borehole model are given in Table 5.7, and the site terms in Table 5.8.

Standard deviation for the inter- and intra-event residuals are shown in Table 5.9, along with an estimate for τ at magnitudes greater than 6.5 to check what would be the standard

Table 5.7. Model Parameters for the Distance and Magnitude terms of the Borehole Model

Period	c_1	c_2	c_3	e_5	e_6
PGA	-1.3569	0.5226	-0.0140	-1.0677	-0.5605
0.0384	-1.4232	0.5059	-0.0140	-0.9255	-0.4878
0.0484	-1.4148	0.4354	-0.0140	-0.5708	-0.4189
0.0582	-1.3897	0.3857	-0.0140	0.0004	-0.2156
0.0769	-1.3553	0.3796	-0.0140	-0.0038	-0.2532
0.0844	-1.3376	0.3791	-0.0140	-0.0125	-0.2677
0.0970	-1.3134	0.3716	-0.0140	0.0278	-0.2929
0.1167	-1.2824	0.3656	-0.0138	0.1160	-0.2908
0.1472	-1.2364	0.3424	-0.0131	0.3216	-0.2823
0.1691	-1.2402	0.3430	-0.0126	0.3790	-0.2957
0.2036	-1.2423	0.3504	-0.0119	0.3143	-0.3587
0.2340	-1.2481	0.3353	-0.0113	0.4962	-0.3398
0.3090	-1.2535	0.3175	-0.0100	0.7236	-0.3537
0.3551	-1.2529	0.3092	-0.0092	0.4604	-0.4529
0.3896	-1.2559	0.3128	-0.0087	0.0107	-0.5722
0.4274	-1.2469	0.3035	-0.0082	0.1362	-0.5578
0.4690	-1.2238	0.2837	-0.0076	0.3243	-0.5324
0.5913	-1.2341	0.2644	-0.0062	0.4412	-0.5648
0.7456	-1.2794	0.2452	-0.0049	0.6268	-0.5794
0.8180	-1.3027	0.2384	-0.0043	0.6740	-0.5965
0.9401	-1.3213	0.2260	-0.0036	0.8908	-0.5624
1.3622	-1.3603	0.2124	-0.0020	1.2335	-0.5013

Table 5.8. Model Parameters for the Site terms of the Borehole Model

Period	a_100	a_200	b_100	b_200	c_100	c_200
PGA	-1.6846	-1.9046	0.0306	-0.1359	-0.3047	-0.4115
0.0384	-1.3288	-1.5572	0.0391	-0.2036	-0.1685	-0.2443
0.0484	-1.1233	-1.3833	0.0382	-0.2157	-0.1047	-0.2086
0.0582	-0.9882	-1.2369	0.0284	-0.1903	-0.1257	-0.2231
0.0769	-1.2645	-1.4851	0.0933	-0.1577	-0.2253	-0.3181
0.0844	-1.3753	-1.5925	0.0950	-0.1495	-0.2751	-0.3694
0.0970	-1.3770	-1.5972	0.0943	-0.1450	-0.3220	-0.4262
0.1167	-1.4773	-1.6442	0.0737	-0.0839	-0.3967	-0.4883
0.1472	-1.5884	-1.6977	0.0023	-0.0001	-0.4717	-0.6178
0.1691	-1.5601	-1.6901	0.0030	0.0000	-0.4859	-0.6454
0.2036	-1.5845	-1.7787	0.0000	-0.0000	-0.4940	-0.7212
0.2340	-1.5052	-1.7053	-0.0029	0.0077	-0.5175	-0.7264
0.3090	-1.3217	-1.5619	-0.0079	0.0665	-0.5180	-0.7972
0.3551	-1.2451	-1.5238	-0.0297	0.0667	-0.4676	-0.8085
0.3896	-1.3256	-1.6308	-0.0347	0.0464	-0.4445	-0.7870
0.4274	-1.2323	-1.5656	-0.0526	0.0326	-0.4194	-0.7769
0.4690	-1.2130	-1.5903	-0.0636	0.0250	-0.4052	-0.8231
0.5913	-1.2298	-1.6056	-0.0509	0.0075	-0.4148	-0.8031
0.7456	-1.1376	-1.3802	-0.0340	0.0770	-0.4538	-0.6985
0.8180	-1.1180	-1.3333	-0.0145	0.0891	-0.4971	-0.6868
0.9401	-1.1070	-1.2754	-0.0224	0.0818	-0.5412	-0.6630
1.3622	-1.0976	-1.1312	-0.0729	0.0602	-0.6703	-0.7195

deviation of the inter-event residuals if the model allowed for oversaturation.

5.2.4 Model Comparisons

Figure 5.7 shows the comparison of the distance attenuation for the Borehole and Surface models. As expected the attenuation at higher frequencies is faster than at lower frequencies, and it can be concluded that both models attenuate at a similar rate, while the difference is a constant shift in the vertical axis (the median prediction) due to the effect of the shallow surface deposits.

Figure 5.8 shows the magnitude attenuation terms for the three models. The fixed value of e_1 allows the comparison of F_m for the ground surface and borehole. While for the

Table 5.9. Standard Deviations of the Residuals from the Borehole Model

Period	ϕ for $Mw < 5$	ϕ for $Mw > 6.5$	τ for $Mw < 5$	τ for $Mw > 6.5$	τ^* for $Mw > 6.5$
PGA	0.5782	0.5400	0.4590	1.0701	0.4407
0.0384	0.5787	0.5465	0.4749	1.0331	0.4836
0.0484	0.5936	0.5636	0.5082	0.8667	0.5611
0.0582	0.6127	0.5791	0.5045	0.7940	0.6207
0.0769	0.5973	0.5692	0.4696	0.8376	0.6562
0.0844	0.5985	0.5703	0.4586	0.8480	0.6599
0.0970	0.6023	0.5732	0.4538	0.8183	0.6476
0.1167	0.6067	0.5800	0.4586	0.7953	0.6262
0.1472	0.6091	0.5797	0.4407	0.7710	0.5810
0.1691	0.6117	0.5816	0.4286	0.7891	0.5530
0.2036	0.6106	0.5798	0.4377	0.7516	0.5113
0.2340	0.5979	0.5757	0.4482	0.6882	0.4890
0.3090	0.5849	0.5736	0.4609	0.6341	0.4429
0.3551	0.5790	0.5696	0.4677	0.6274	0.4093
0.3896	0.5700	0.5661	0.4754	0.6597	0.3823
0.4274	0.5627	0.5618	0.4787	0.6408	0.3692
0.4690	0.5552	0.5540	0.4865	0.6031	0.3659
0.5913	0.5289	0.5380	0.5185	0.5106	0.3501
0.7456	0.4994	0.5179	0.5467	0.4291	0.2983
0.8180	0.4932	0.5146	0.5473	0.3885	0.2761
0.9401	0.4882	0.5157	0.5563	0.3313	0.2420
1.3622	0.4811	0.5196	0.6030	0.2976	0.2052

[*]Allowing for oversaturation at large magnitudes

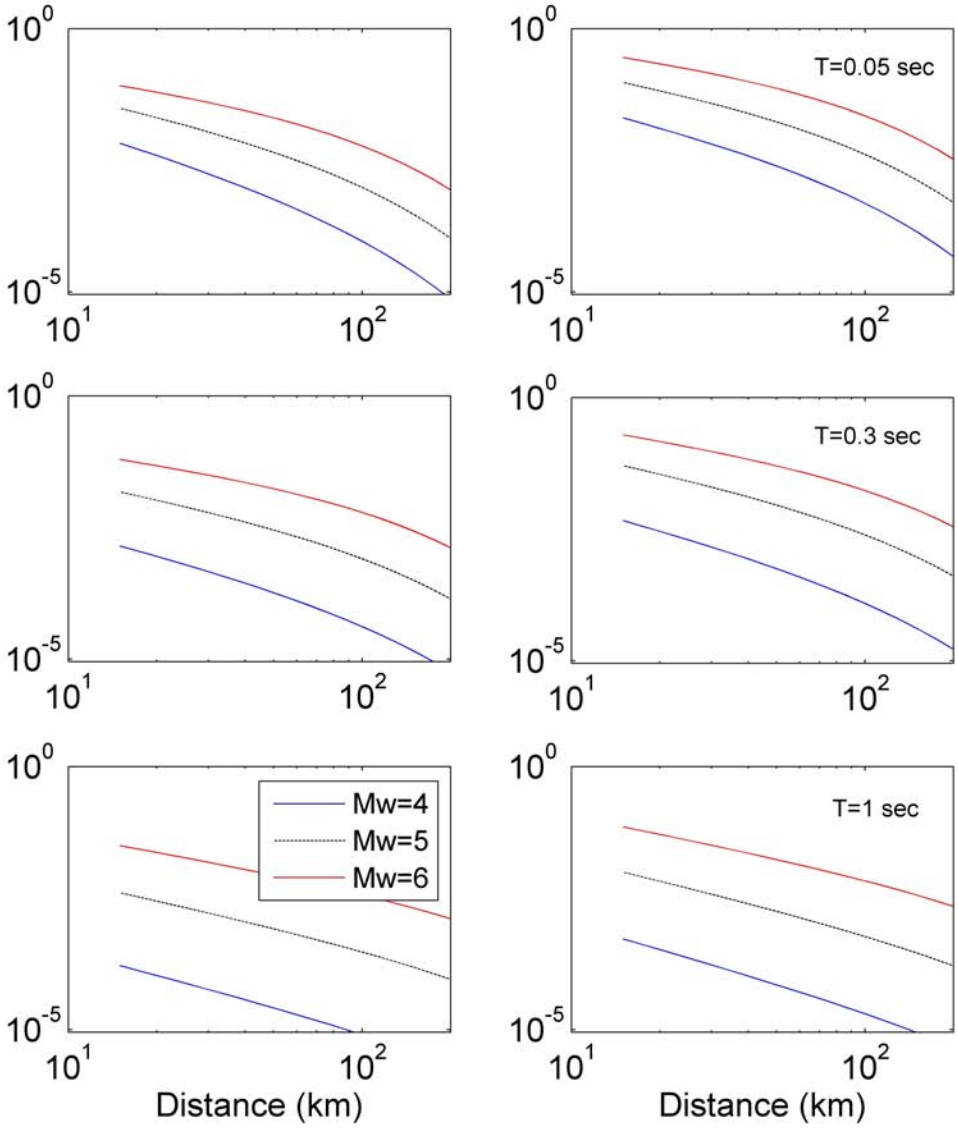


Figure 5.7. Ground motion estimates attenuation with respect to distance for magnitudes 4, 5, and 6. Left column corresponds to the estimates for borehole, and right column to surface estimates; first row shows estimates for a spectral period of 0.05 (sec), second row for T = 0.3 (sec), and third row for T = 1.0 (sec). Estimated scenario corresponds to Vs_{30} of 760 (m/s), depth to Vs equal to 800 (m/s) of 60 (m), Vs at depth 100 (m) of 3000 (m/s)

combined model the constant value e_1 was not fixed, it can be observed that the shape and magnitude of the magnitude dependence is in line with the surface and borehole models.

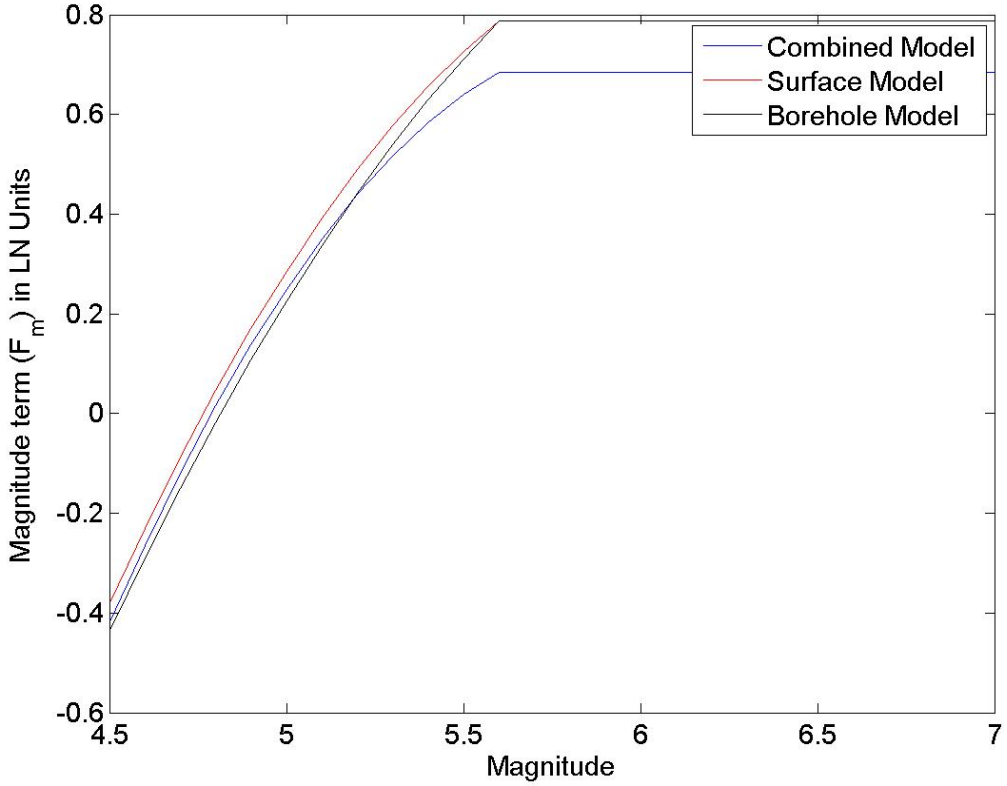


Figure 5.8. Comparison of the magnitude terms for the Combined, Surface, and Borehole models.

Response spectra are plotted in Figure 5.9 comparing the predictions for borehole and surface for magnitudes 4, 5, and 6 and distance to the fault rupture of 20, 50, and 100 kilometers. The rest of the predictor variables were set to reference values. The difference in predicted amplitude of the ground motion reflects the site response effect on the ground motion.

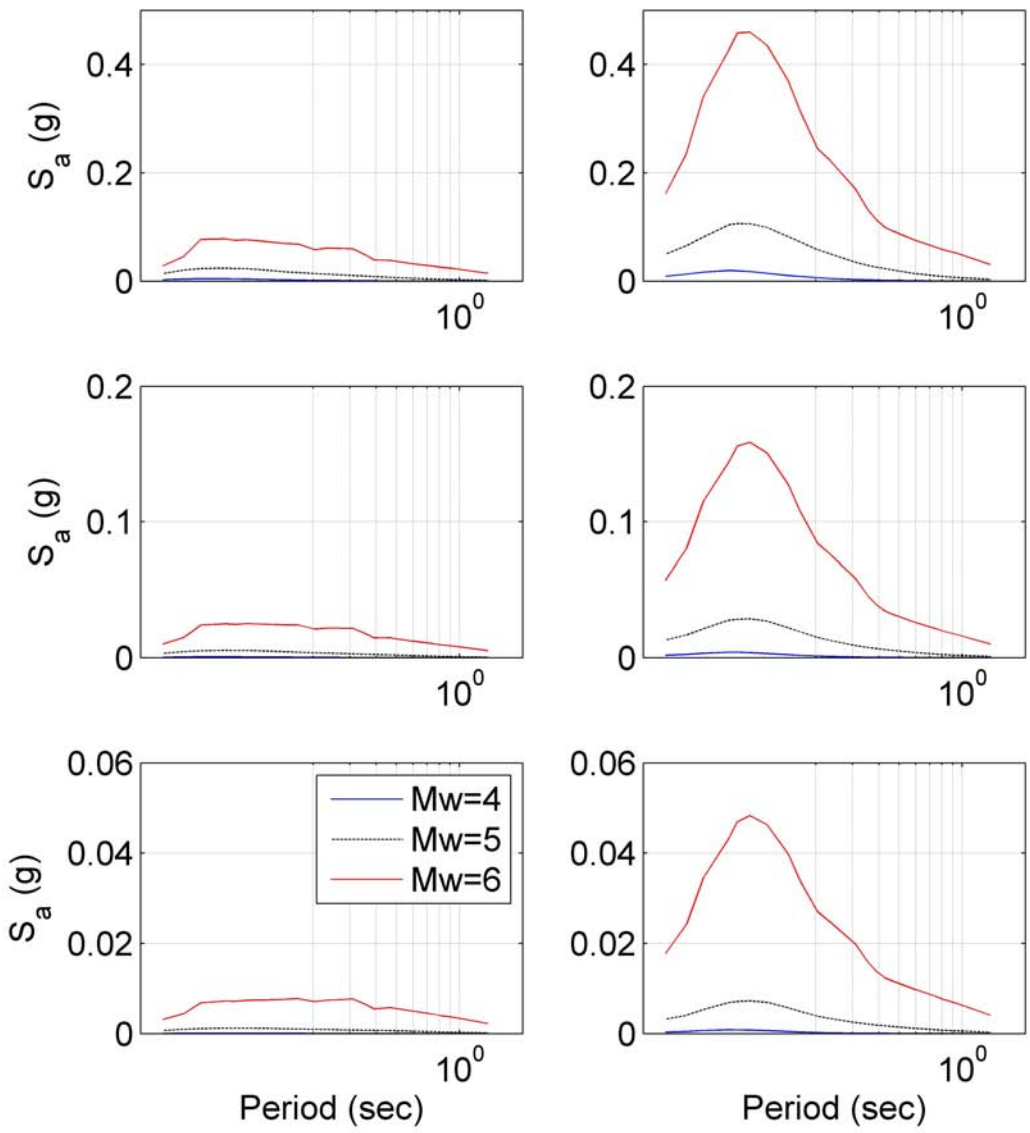


Figure 5.9. Median spectrum estimates for magnitudes 4, 5, and 6. Left column corresponds to the estimates for borehole, and right column to surface estimates; first row shows estimates for a distance to the fault plane of 20 (km), second row for R_{RUP}= 50 (km), and third row for R_{RUP}= 100 (km). Estimated scenario corresponds to Vs_{30} of 760 (m/s), depth to Vs equal to 800 (m/s) of 60 (m), Vs at depth bedrock of 3000 (m/s)

5.2.5 Analysis of Residuals

This section presents the residuals from the regression analyses as functions of Magnitude and Event depth. The residuals from the regression analyses as functions of other main independent parameters is presented in the appendix. This allows the assessment of the models' performance. Results are presented for the three models developed, for *Surface*, *Borehole*, and *Combined* for spectral periods of 0.03, 0.2, 0.6, 1.0, and 1.4 seconds.

5.2.5.1 Inter-Event Residuals

Figures 5.10 through 5.19 show the inter-event residuals for the surface, borehole, and combined models, for the spectral periods and parameters previously mentioned. The magnitude dependence figures show the median prediction overestimating the data for magnitudes over 6.5, this trend was expected as over-saturation was not allowed during the regression. Note also that there is a clear trend of the residuals with respect to event depths, this shows that for deeper events the median under-estimates the measured values. This has a fundamental reason in that deeper events, parameterized by distance-to-rupture (R_{RUP}), will be less attenuated than shallower events with similar distance parameterization because deeper events' paths include a larger portion of hard rock. In addition, events at larger depths may have different stress drop than events at lower depths.

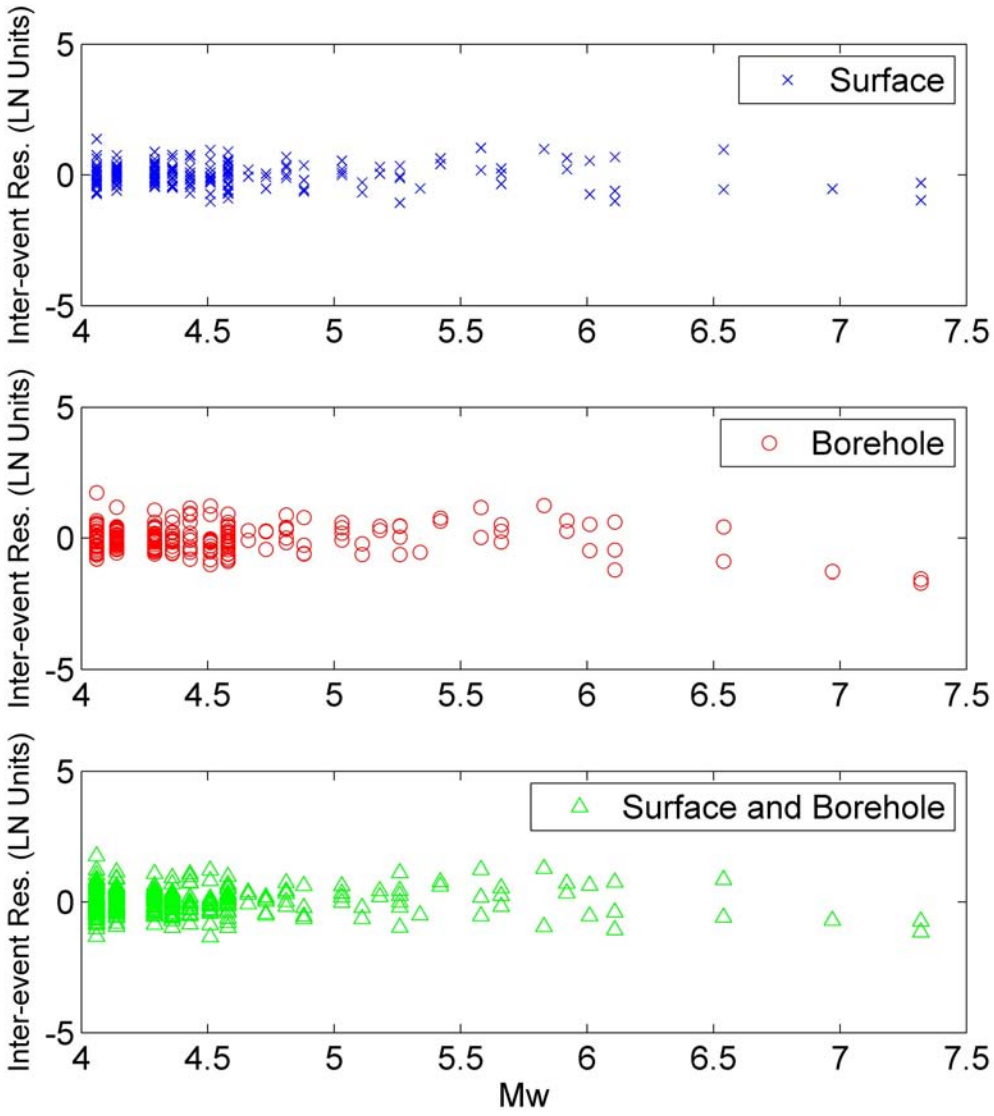


Figure 5.10. Inter-Event Residuals for spectral period of 0.03 seconds and the models for Surface, Borehole, and Combined versus Magnitude.

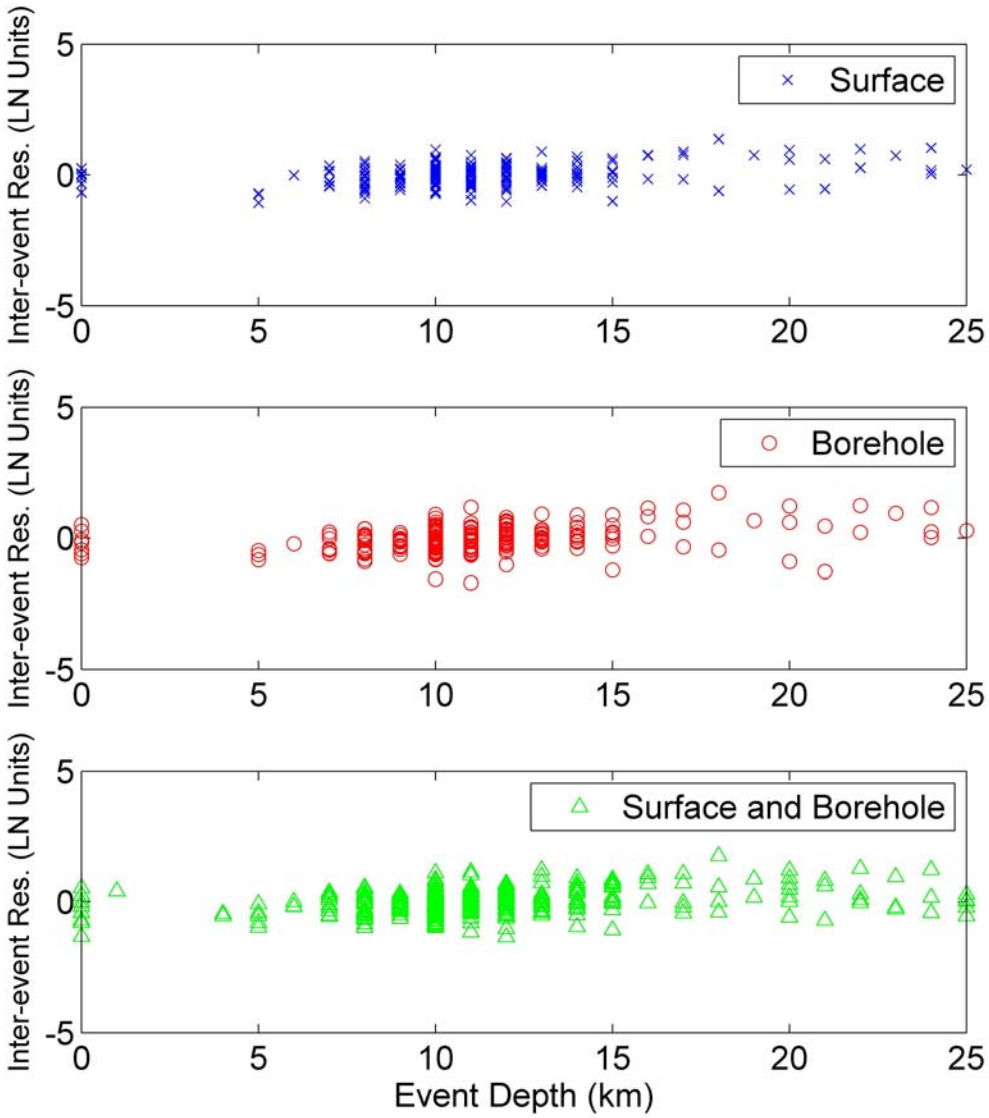


Figure 5.11. Inter-Event Residuals for spectral period of 0.03 seconds and the models for Surface, Borehole, and Combined versus Event Depth.

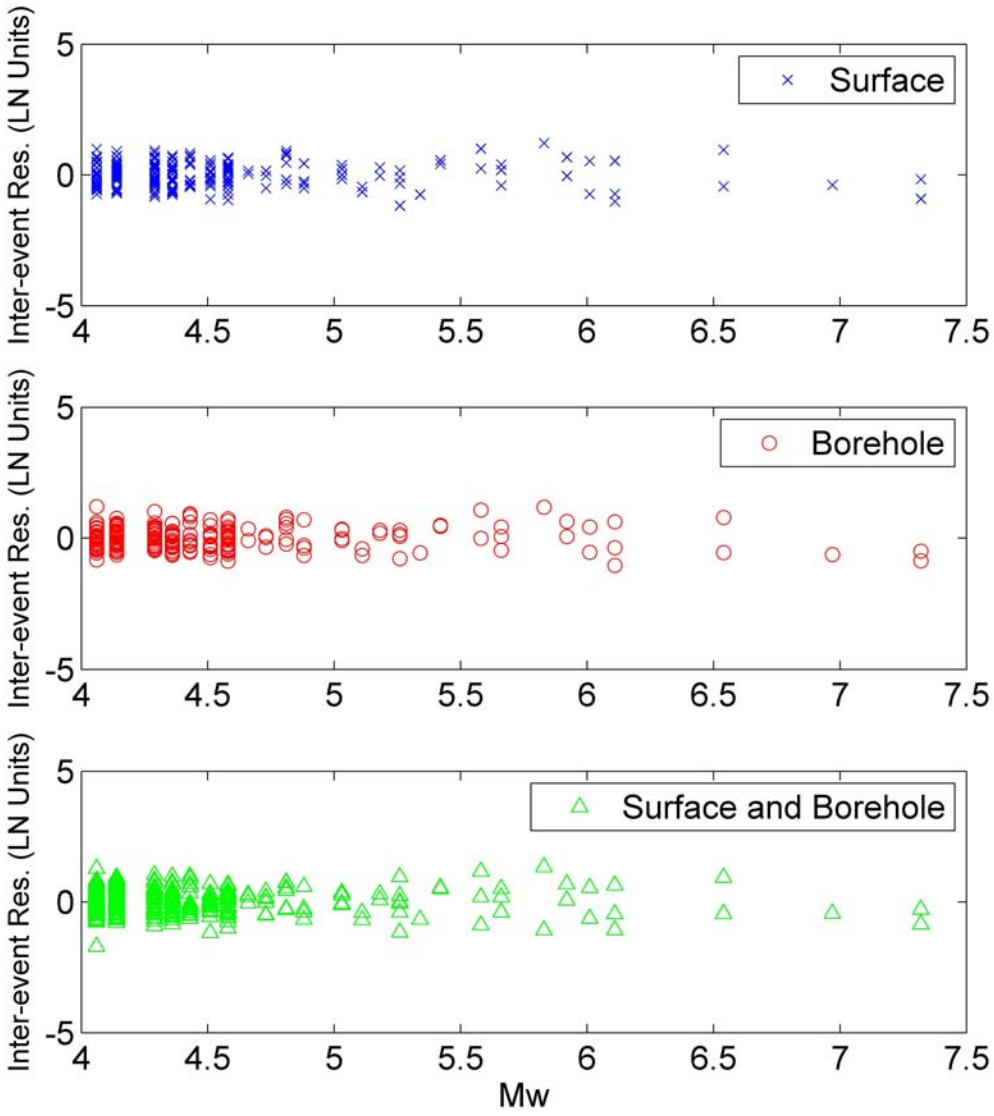


Figure 5.12. Inter-Event Residuals for spectral period of 0.2 seconds and the models for Surface, Borehole, and Combined versus Magnitude.

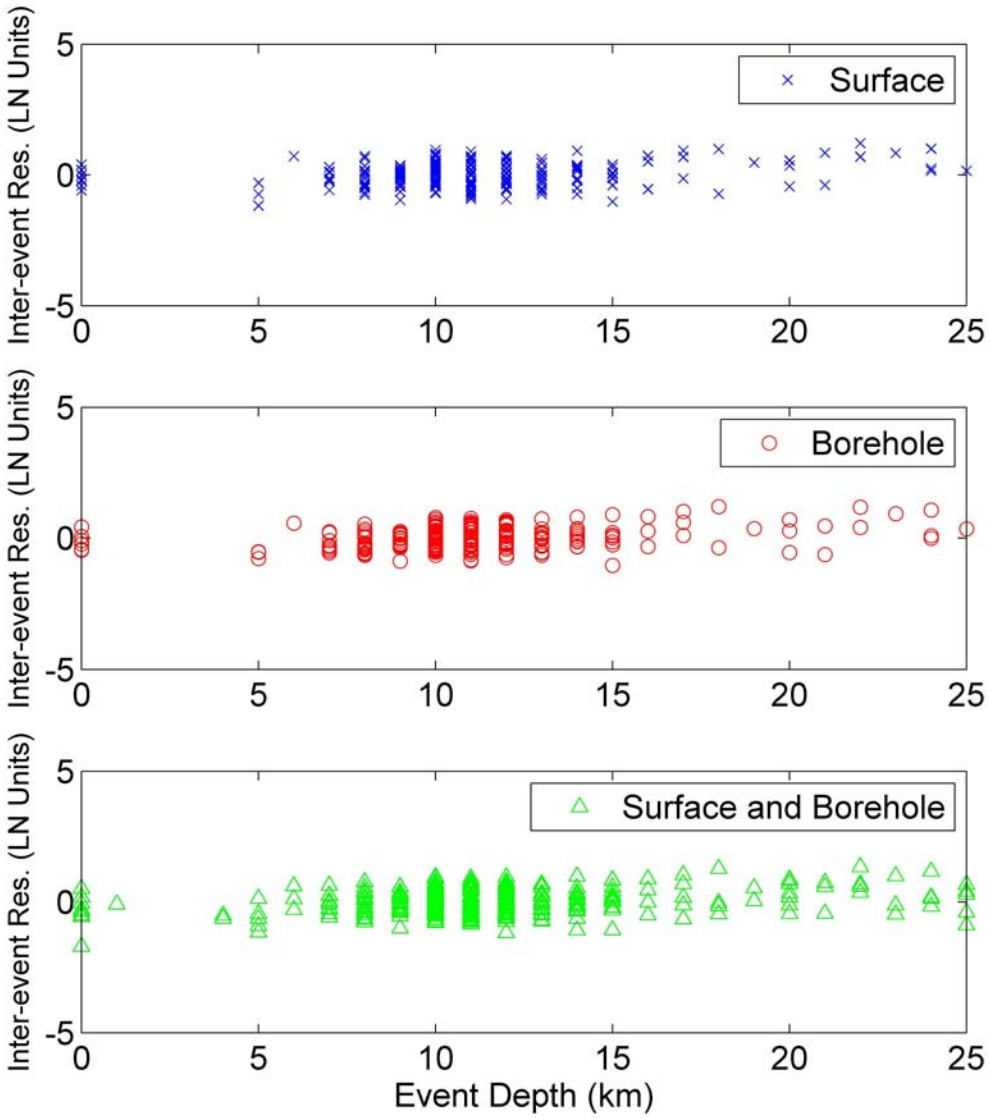


Figure 5.13. Inter-Event Residuals for spectral period of 0.2 seconds and the models for Surface, Borehole, and Combined versus Event Depth.

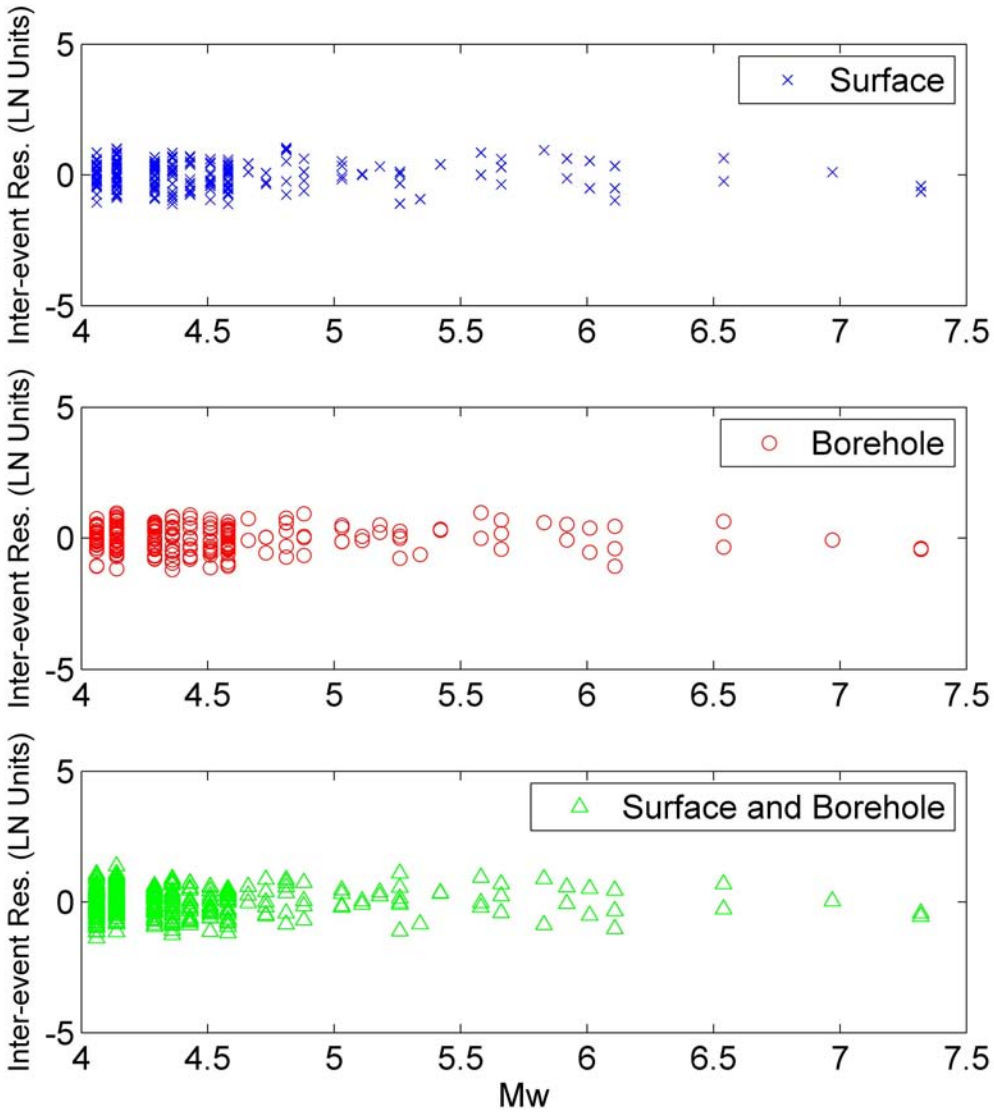


Figure 5.14. Inter-Event Residuals for spectral period of 0.6 seconds and the models for Surface, Borehole, and Combined versus Magnitude.

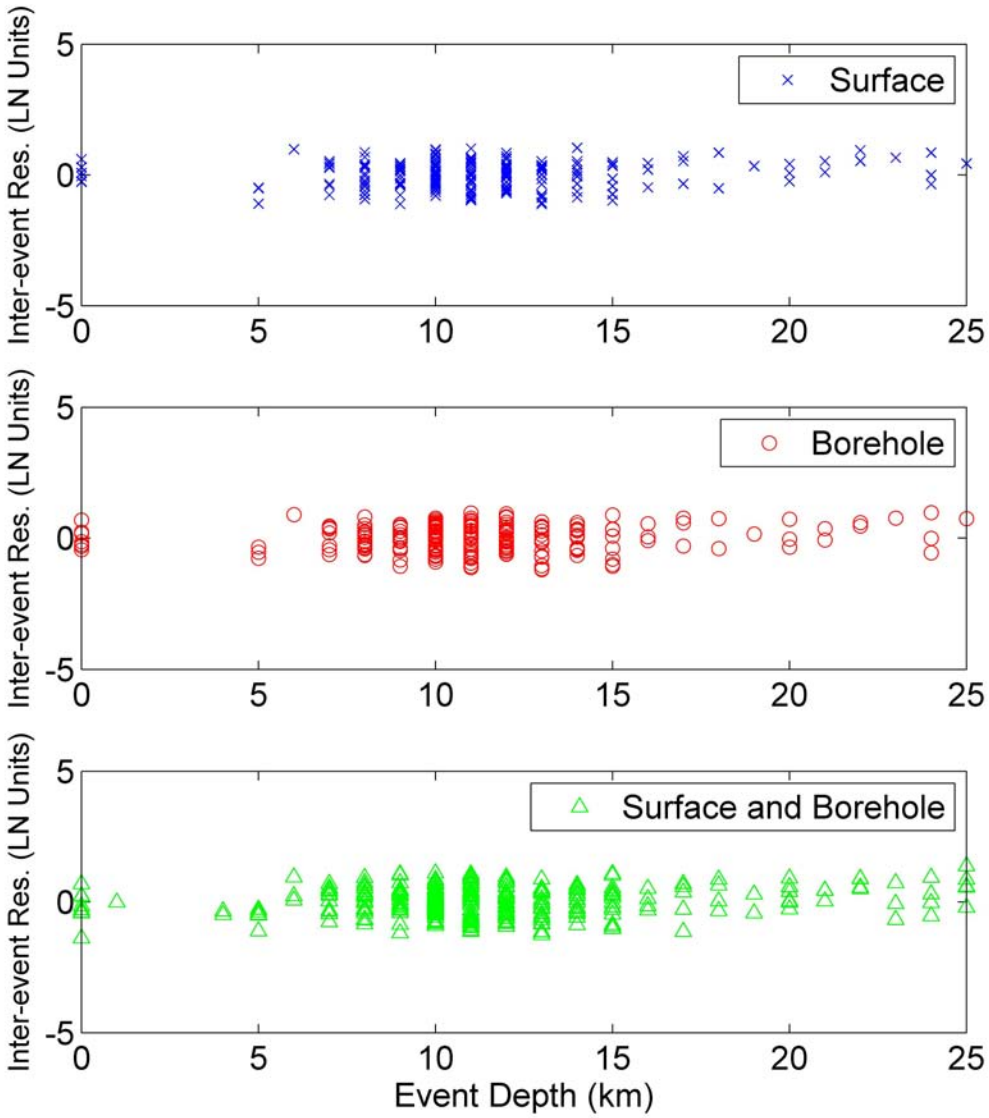


Figure 5.15. Inter-Event Residuals for spectral period of 0.6 seconds and the models for Surface, Borehole, and Combined versus Event Depth.

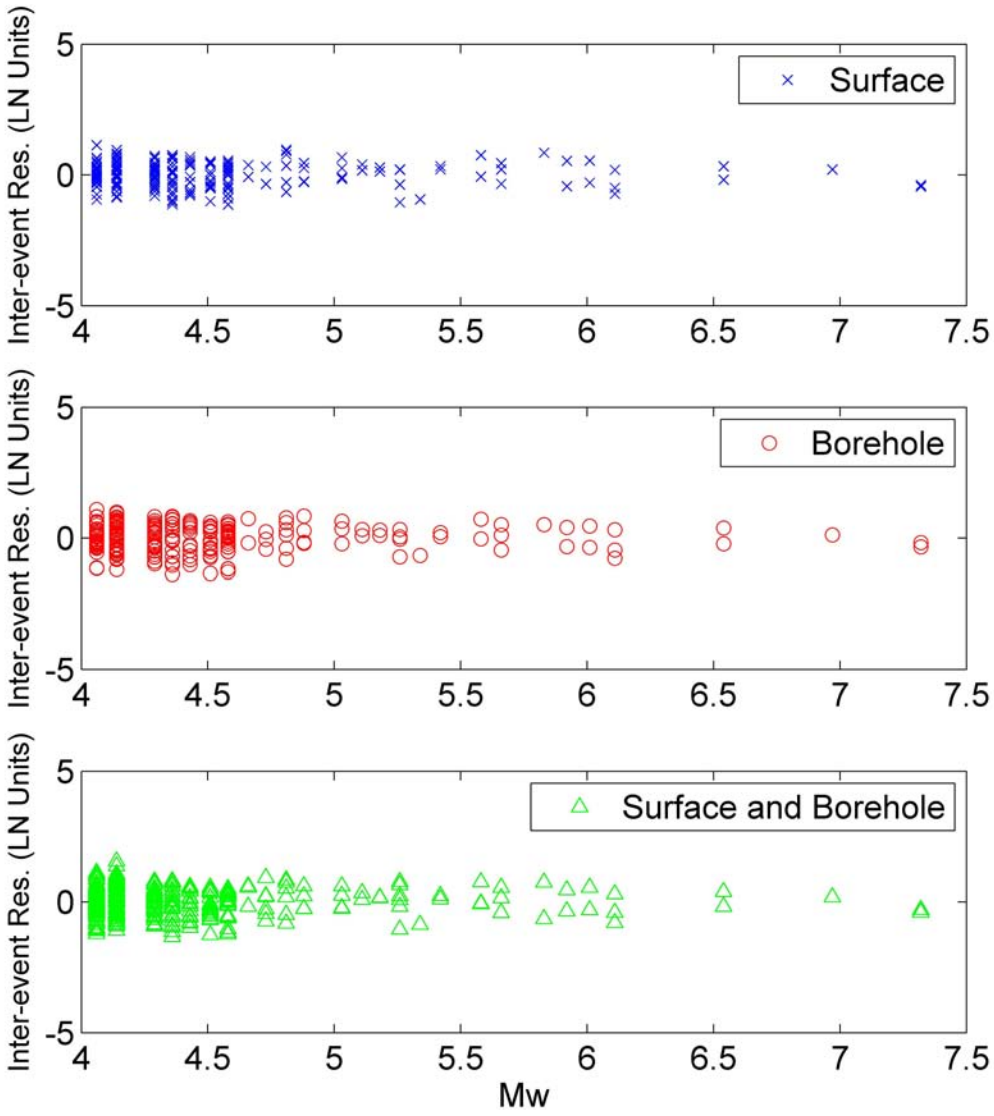


Figure 5.16. Inter-Event Residuals for spectral period of 1.0 second and the models for Surface, Borehole, and Combined versus Magnitude.

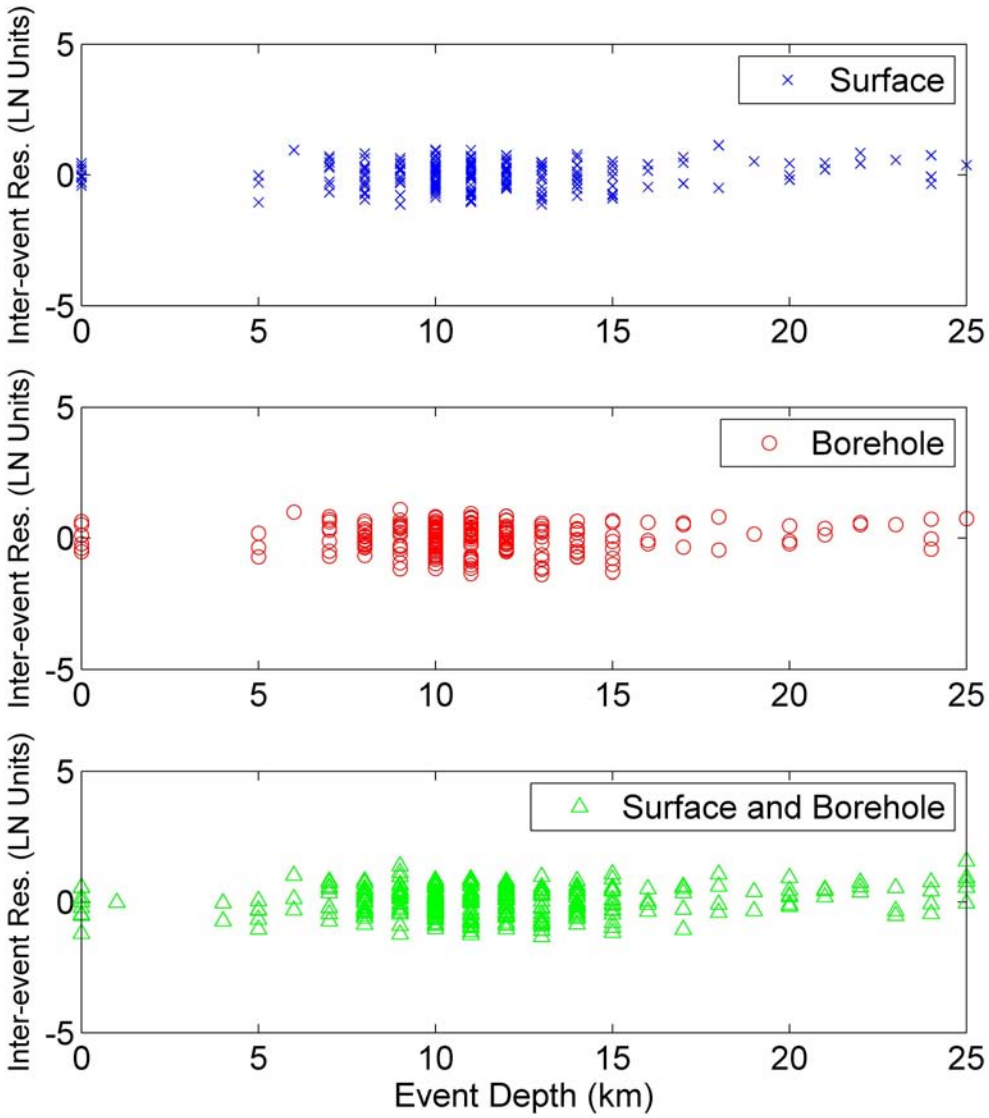


Figure 5.17. Inter-Event Residuals for spectral period of 1.0 second and the models for Surface, Borehole, and Combined versus Event Depth.

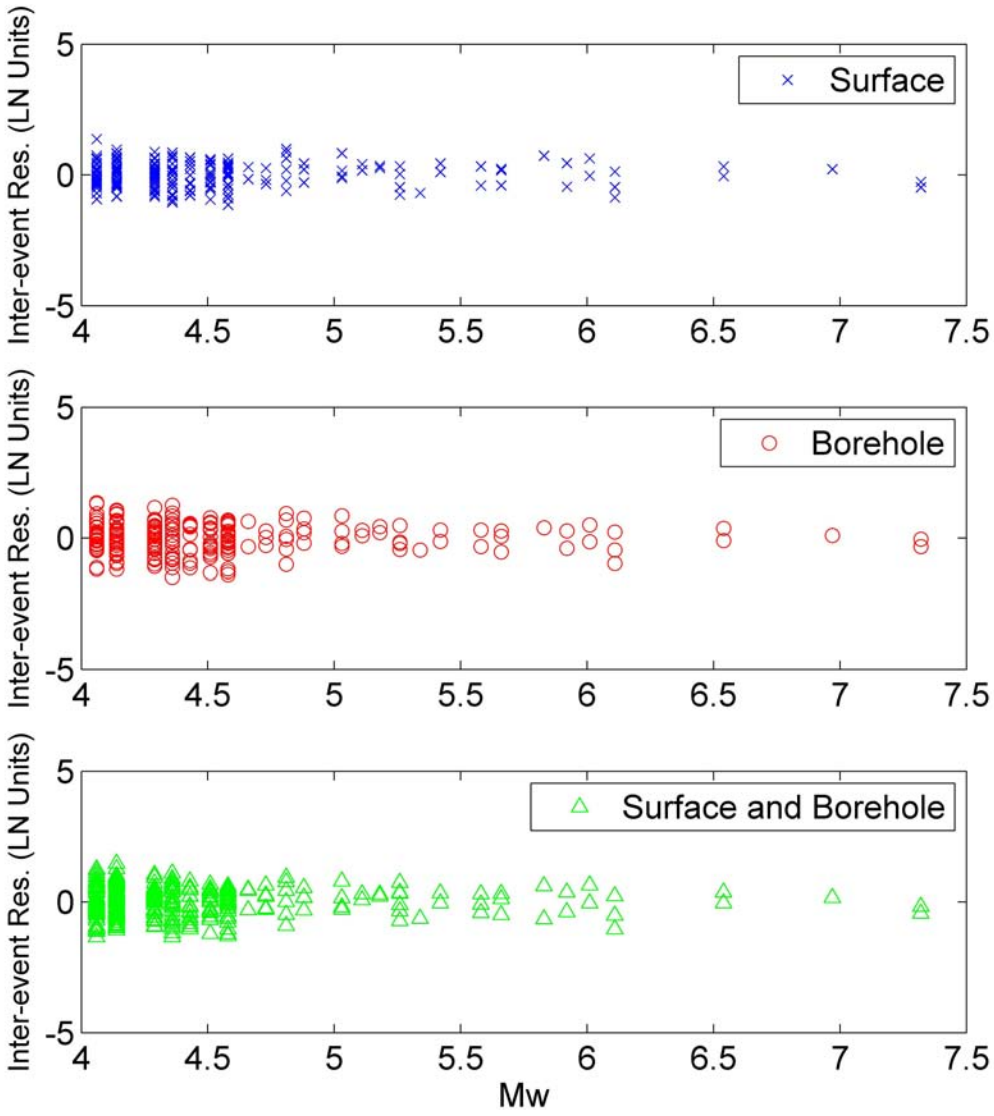


Figure 5.18. Inter-Event Residuals for spectral period of 1.4 seconds and the models for Surface, Borehole, and Combined versus Magnitude.

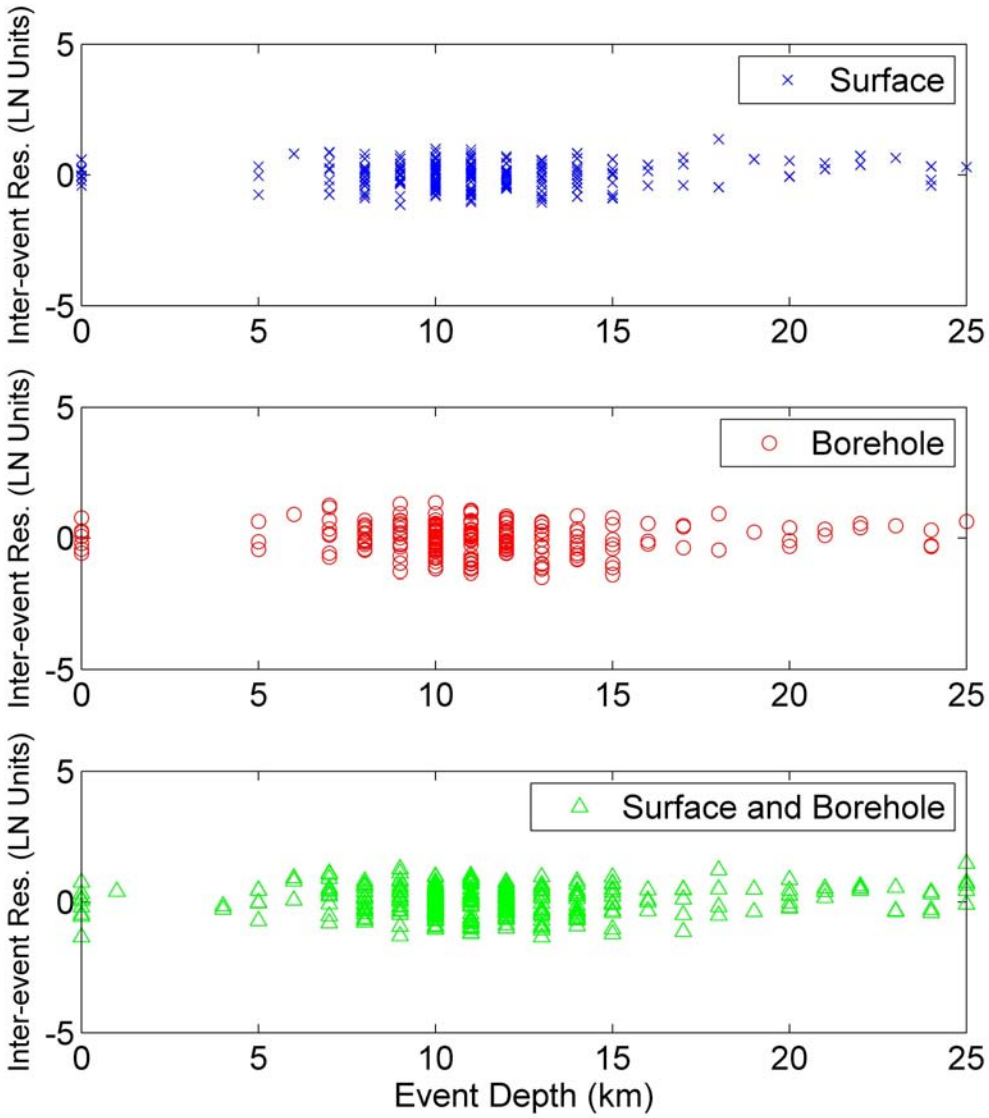


Figure 5.19. Inter-Event Residuals for spectral period of 1.4 seconds and the models for Surface, Borehole, and Combined versus Event Depth.

5.2.5.2 Intra-Event Residuals

Figures 5.20 to 5.29 show the intra-event residuals for the surface, borehole, and combined models, for the spectral periods previously mentioned. Residuals are plotted versus Magnitude and Closest distance to the fault rupture. There is no trend with any of the regressed variables (see also appendix A). However a trend was found when looking at close events (e.g. recorded at distances to the fault lower than 20 km), this is because the model does not consider any specific term to account for near fault effects. Moreover, soil nonlinearity is not accounted for in the model and these effects are more likely to be observed at close distances. The fact that the trend is, in general, observed in the surface residuals more than in the borehole residuals concurs with nonlinear soil behavior being a factor. Note also, that the trend is only noticeable at periods lower than 0.6 seconds.

A noticeable trend of decreasing variation of the intra-event residuals is noticed for distances larger than 200 km. This effect is not reflected by recent attenuation models (e.g., the NGA attenuation models), because because these models did not include records at very large distances in their regressed data.

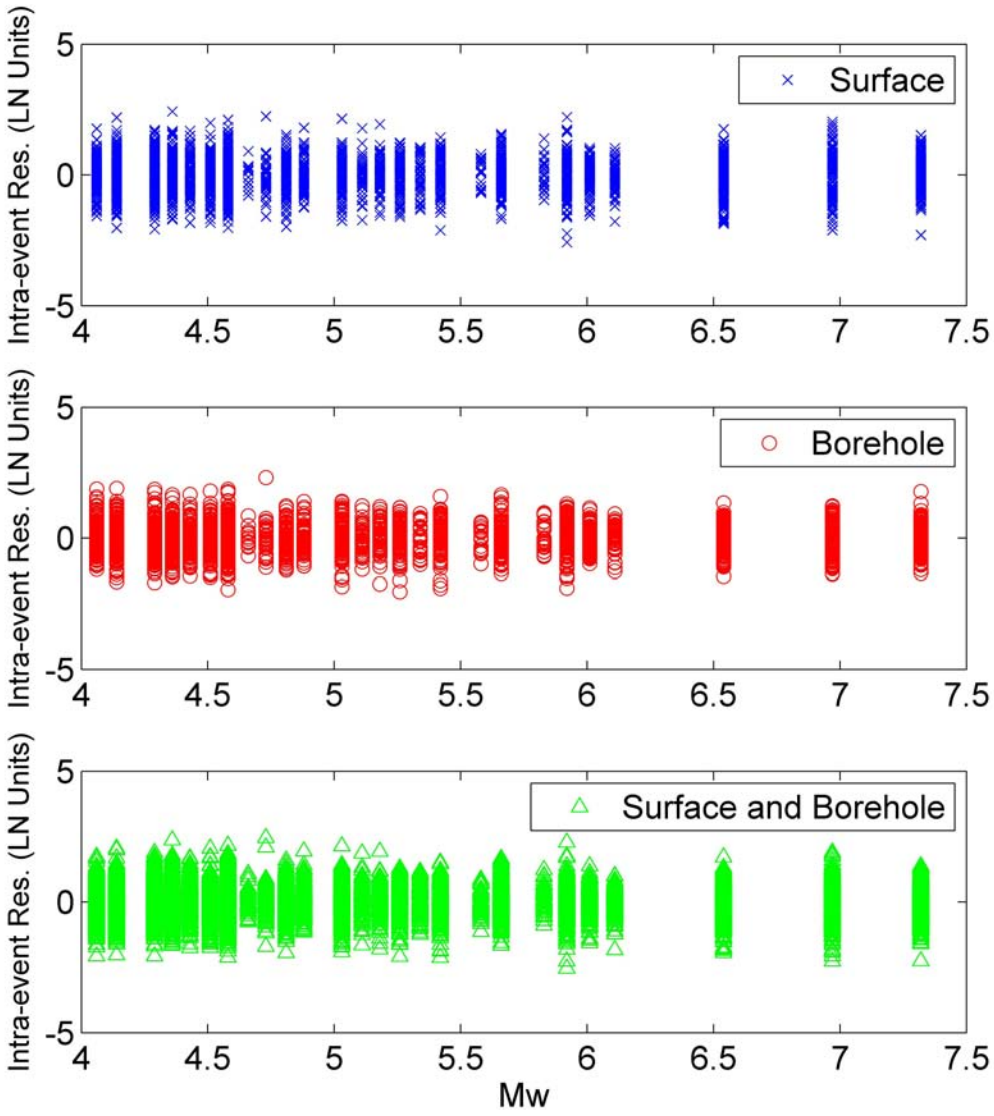


Figure 5.20. Intra-Event Residuals for spectral period of 0.03 seconds and the models for Surface, Borehole, and Combined versus Magnitude.

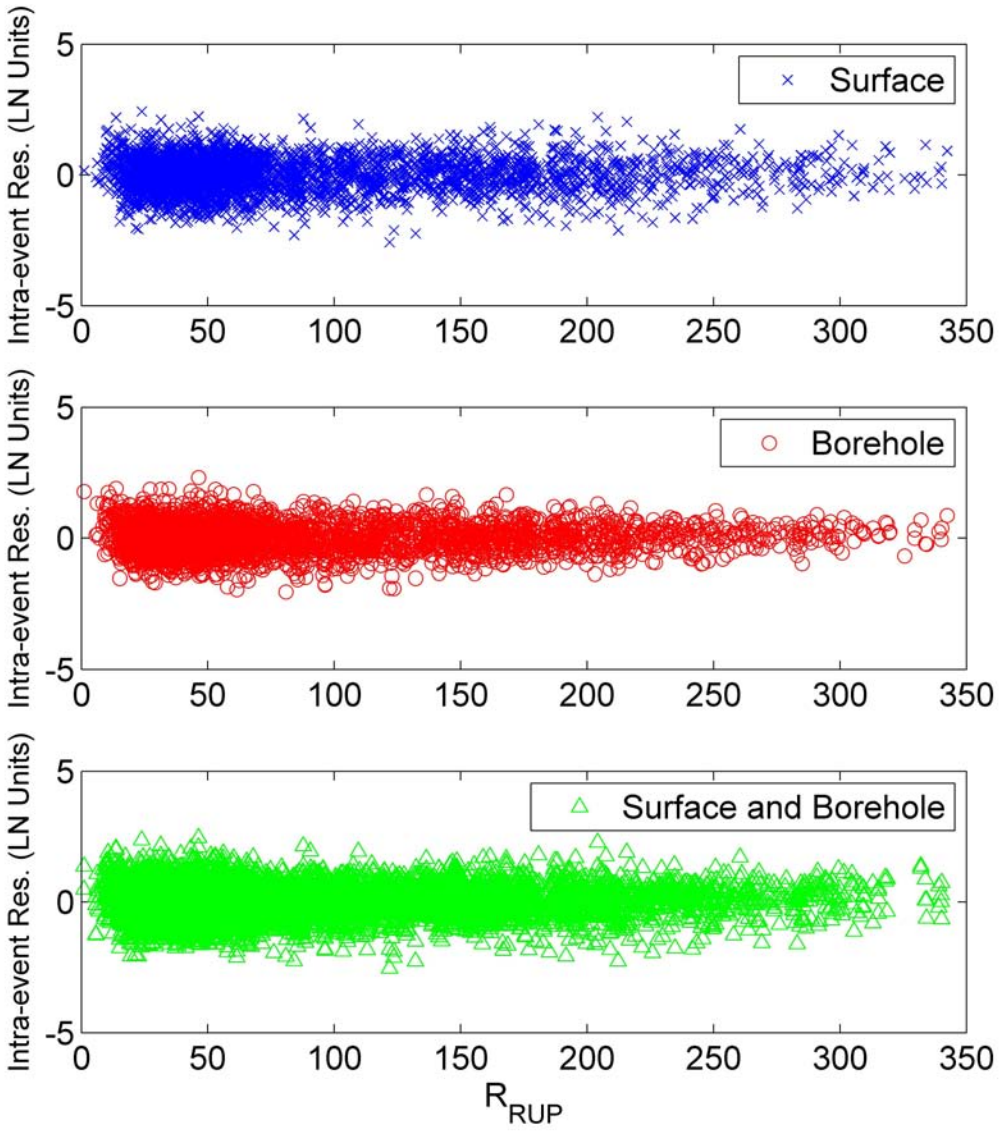


Figure 5.21. Intra-Event Residuals for spectral period of 0.03 seconds and the models for Surface, Borehole, and Combined versus distance to the fault.

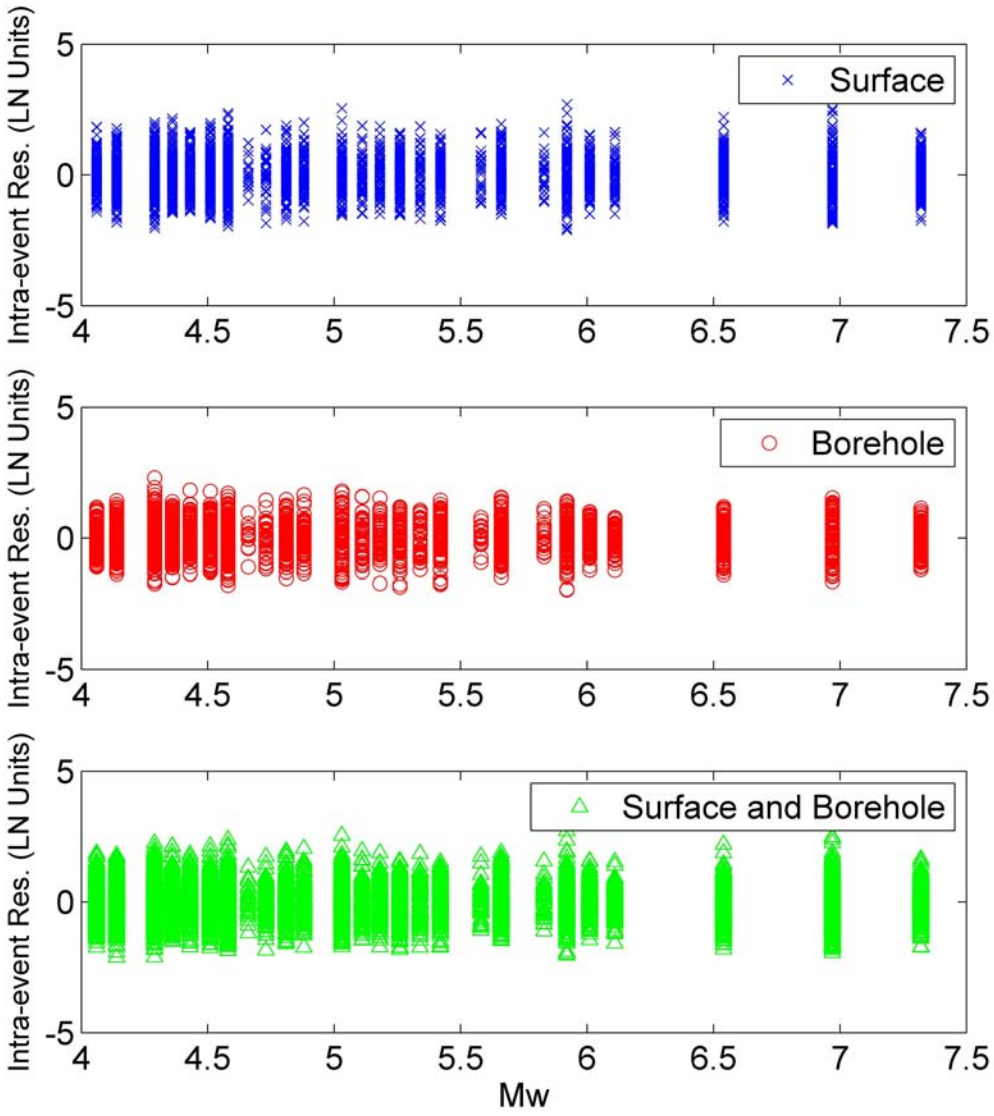


Figure 5.22. Intra-Event Residuals for spectral period of 0.2 seconds and the models for Surface, Borehole, and Combined versus Magnitude.

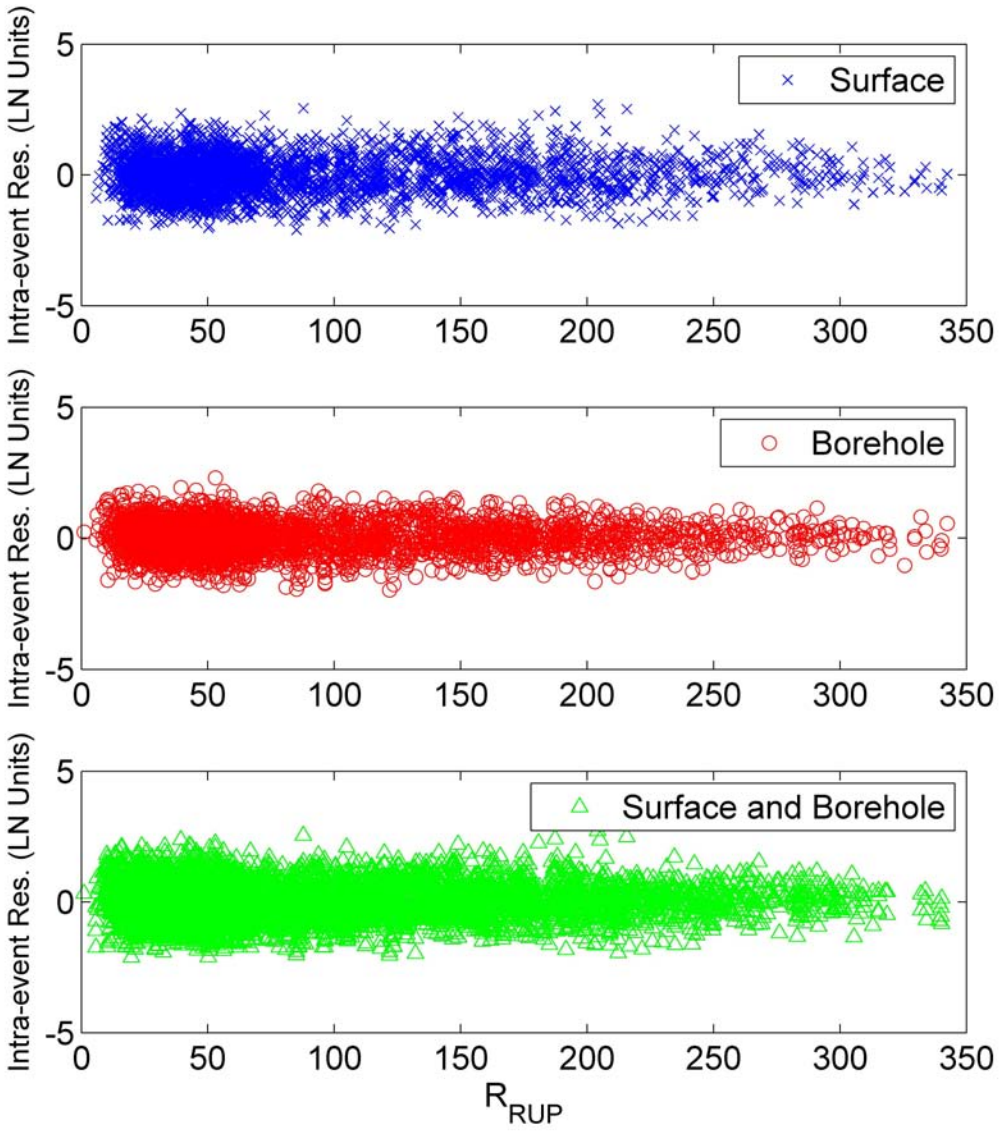


Figure 5.23. Intra-Event Residuals for spectral period of 0.2 seconds and the models for Surface, Borehole, and Combined versus distance to the fault.

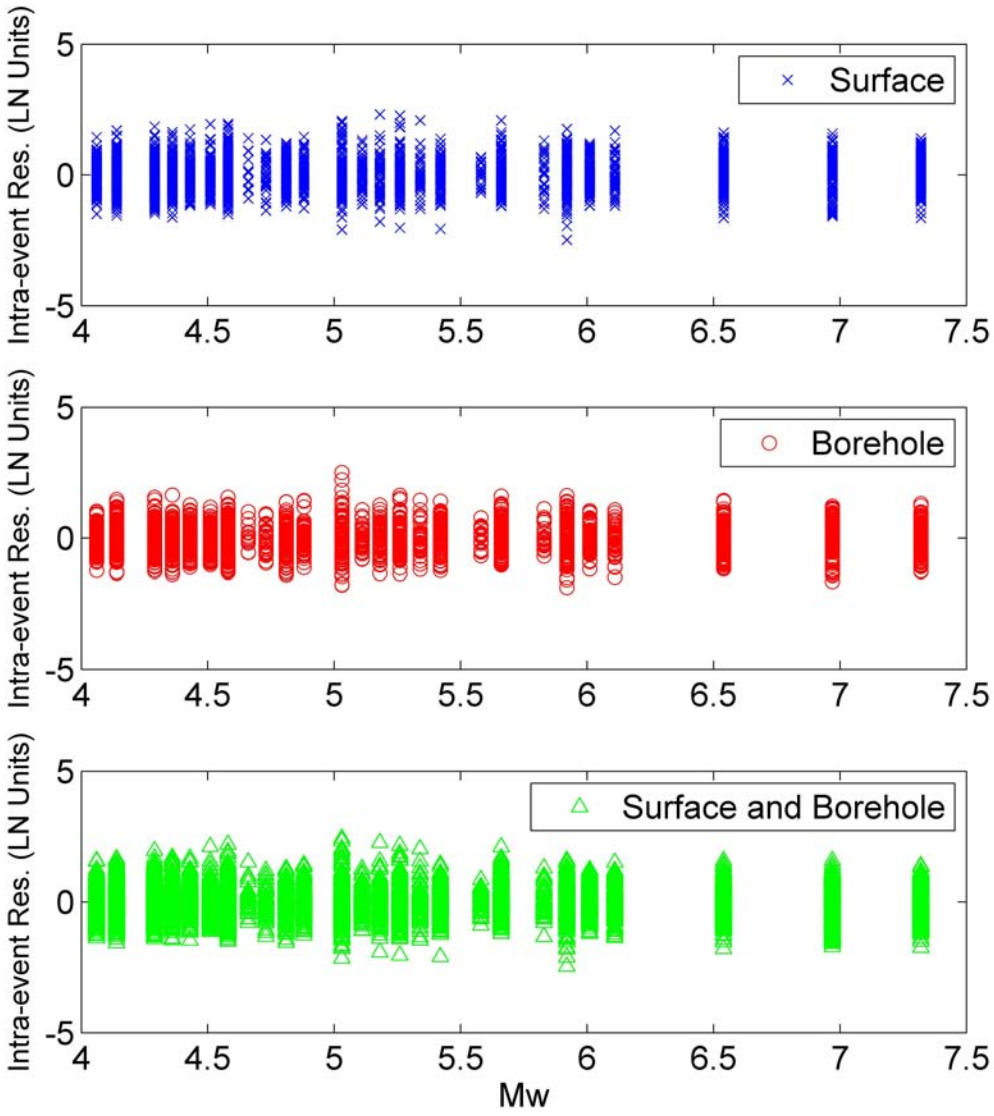


Figure 5.24. Intra-Event Residuals for spectral period of 0.6 seconds and the models for Surface, Borehole, and Combined versus Magnitude.

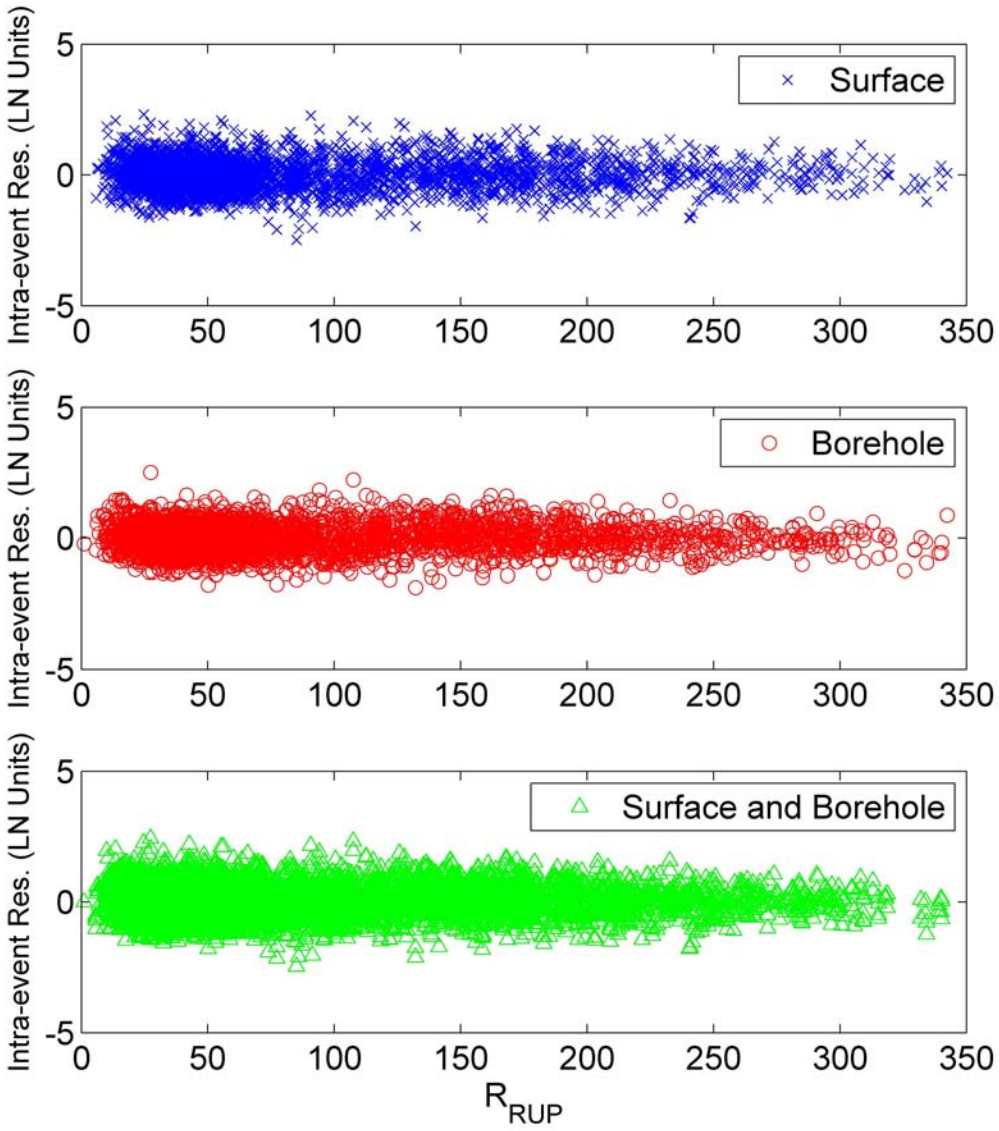


Figure 5.25. Intra-Event Residuals for spectral period of 0.6 seconds and the models for Surface, Borehole, and Combined versus distance to the fault.

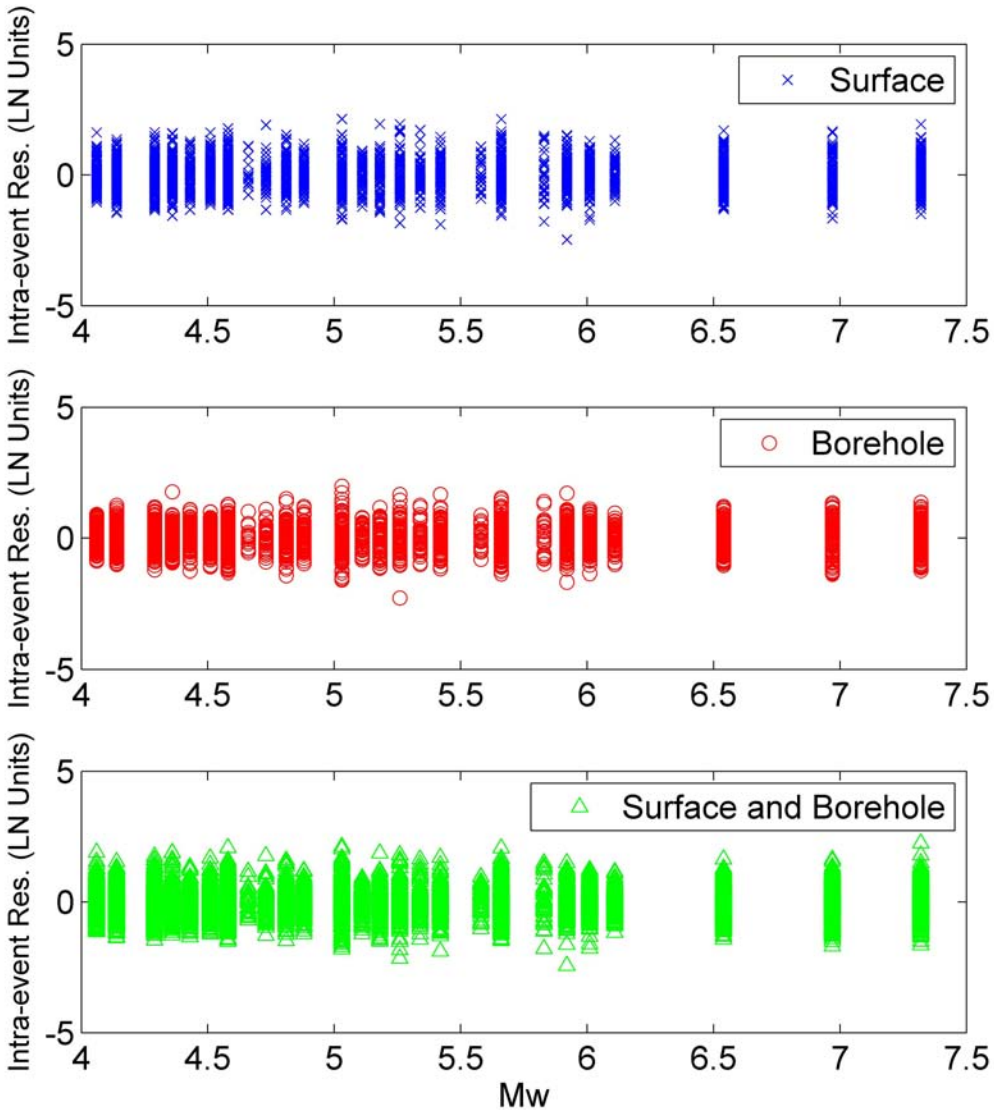


Figure 5.26. Intra-Event Residuals for spectral period of 1.0 second and the models for Surface, Borehole, and Combined versus Magnitude.

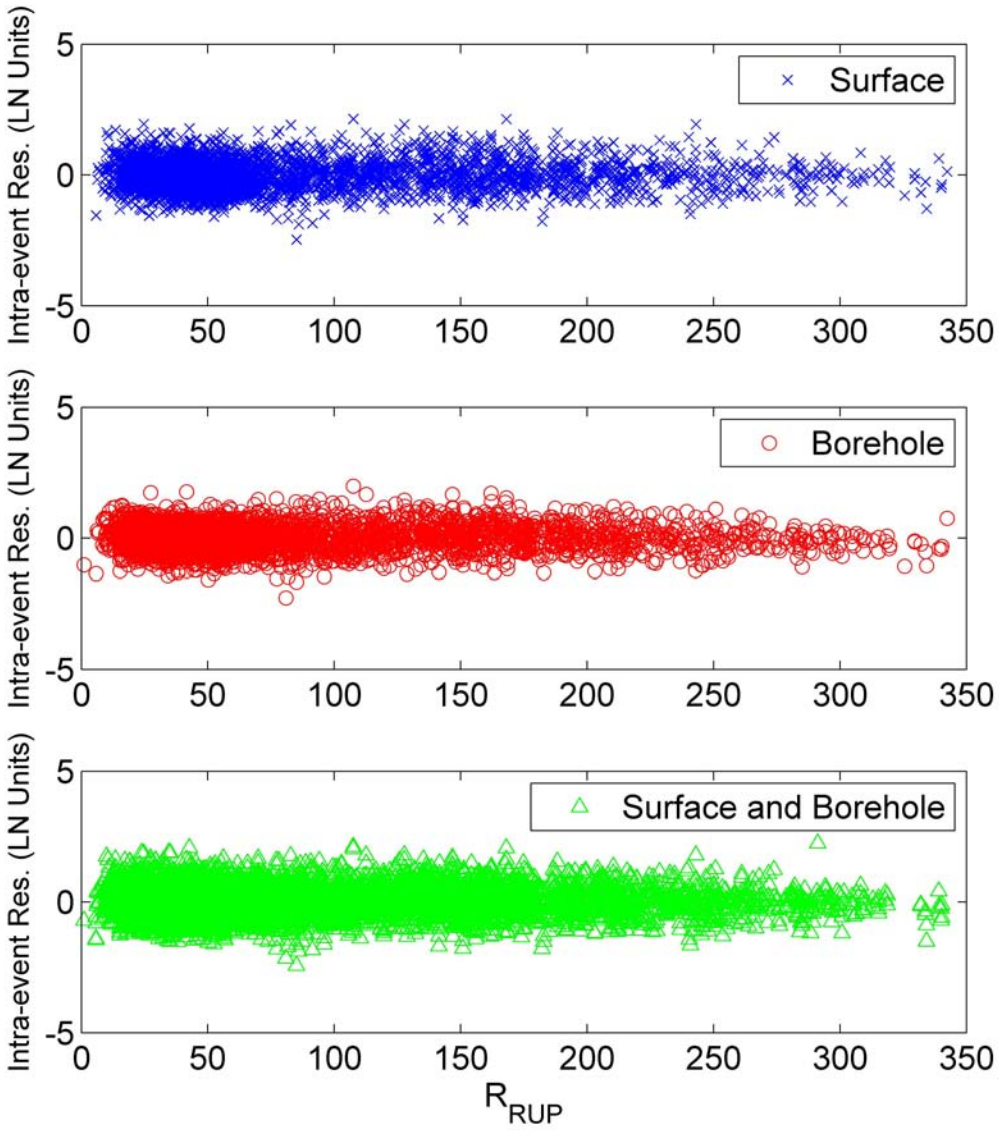


Figure 5.27. Intra-Event Residuals for spectral period of 1.0 second and the models for Surface, Borehole, and Combined versus distance to the fault.

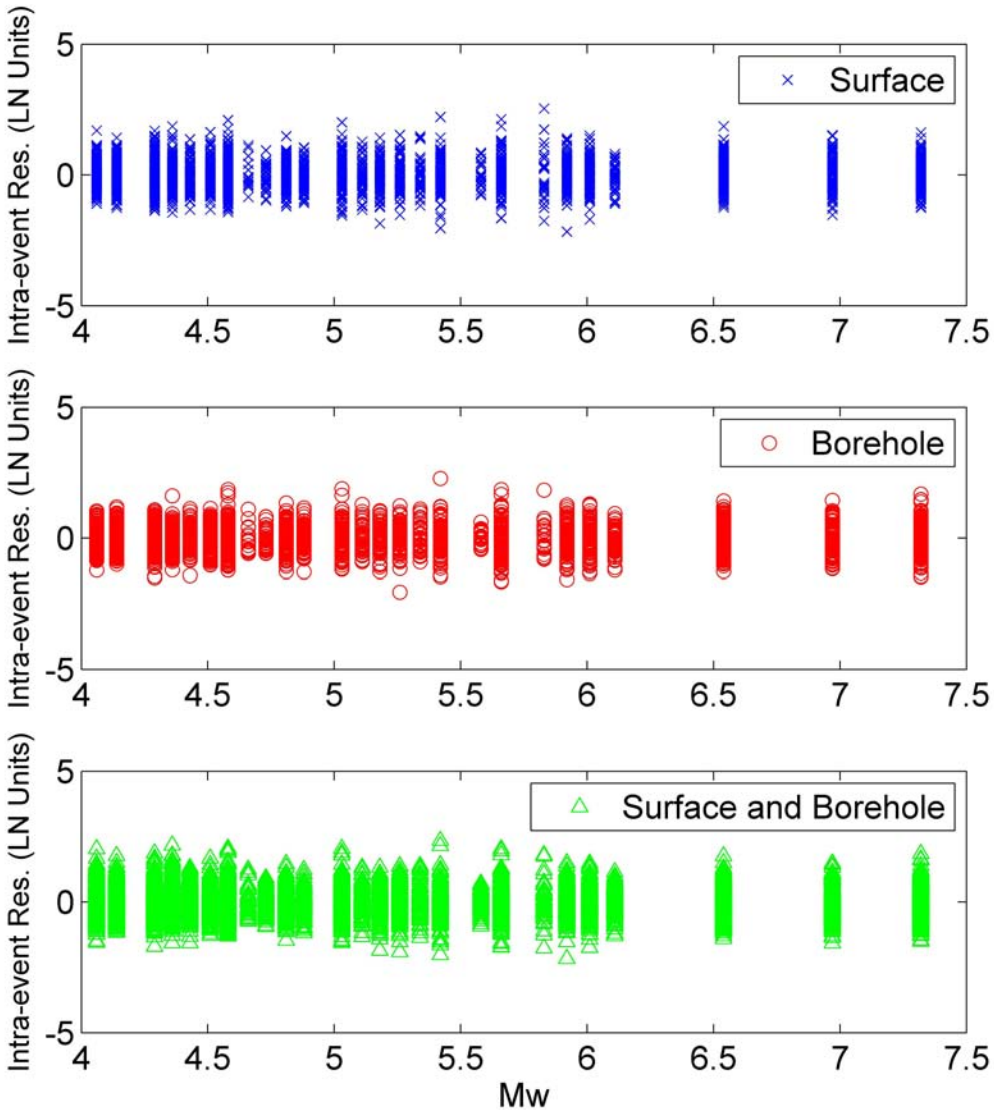


Figure 5.28. Intra-Event Residuals for spectral period of 1.4 seconds and the models for Surface, Borehole, and Combined versus Magnitude.

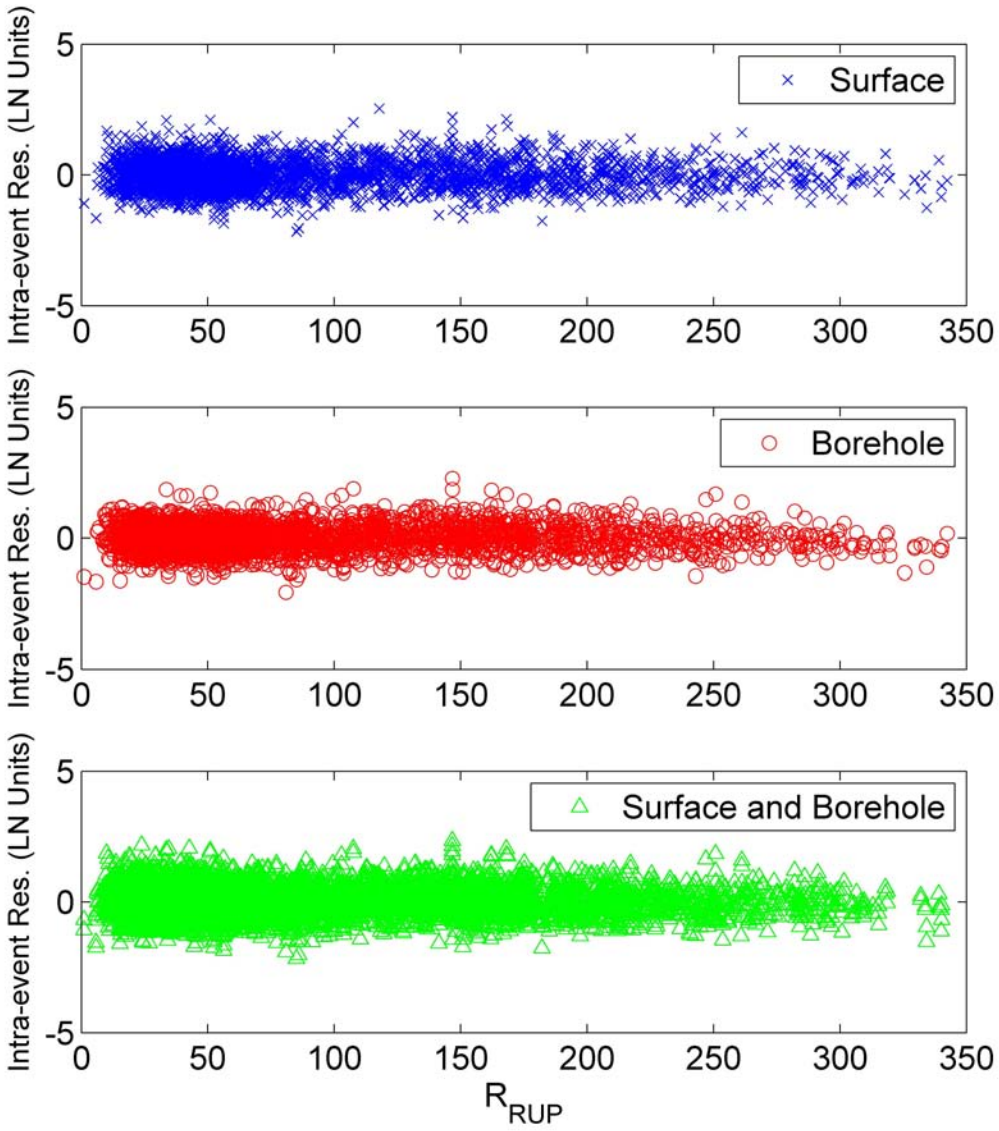


Figure 5.29. Intra-Event Residuals for spectral period of 1.4 seconds and the models for Surface, Borehole, and Combined versus distance to the fault.

5.2.6 Comparison of Inter and Intra-Event Residual Standard Deviations

Inter- and intra-event residuals, as explained above, are normally distributed random variables with zero mean and standard deviations τ and ϕ respectively. The study of these standard deviations is relevant because they dictate how much uncertainty is embedded in the GMPE. Figures 5.30 and 5.31 show the magnitude dependence of these two standard deviations. Note that intra-event standard deviation, ϕ, is almost independent of magnitude, while inter-event standard deviation,τ, are positively correlated with magnitude for short spectral periods, and negatively correlated for higher spectral periods. It is important to note that, as mentioned before, the τ values are over-estimated for large magnitudes due to the bias resulting from preventing oversaturation.

Further comparison of these measures of uncertainty, can be observed in Figures 5.32 to 5.34. Note first that intra-event standard deviation for the combined model falls between ϕ^B and ϕ^G, as expected. Figures 5.33 and 5.34 show that inter-event standard deviations for the three models are not very different. This fact reinforces the suitability of using the Combined model which forces the inter-event average misfits to be equal (there is only one δB_e in the combine model).

Ground motion prediction equations generally assume that inter- and intra-event residuals are independent, this assumption is implicit in that no correlation is included in the equations. Correlation values, however, are seldom published.

To test whether these two random variables are correlated or not, the correlation coefficient between them was calculated. Because the regression process allowed for the inter- and intra-event standard deviations to be magnitude dependent, the values for each of them were normalized to make them standard normal random variables (i.e. $N\left(0,1\right)$) by dividing them by the corresponding standard deviation (τ for the inter-event, and ϕ for the intra-

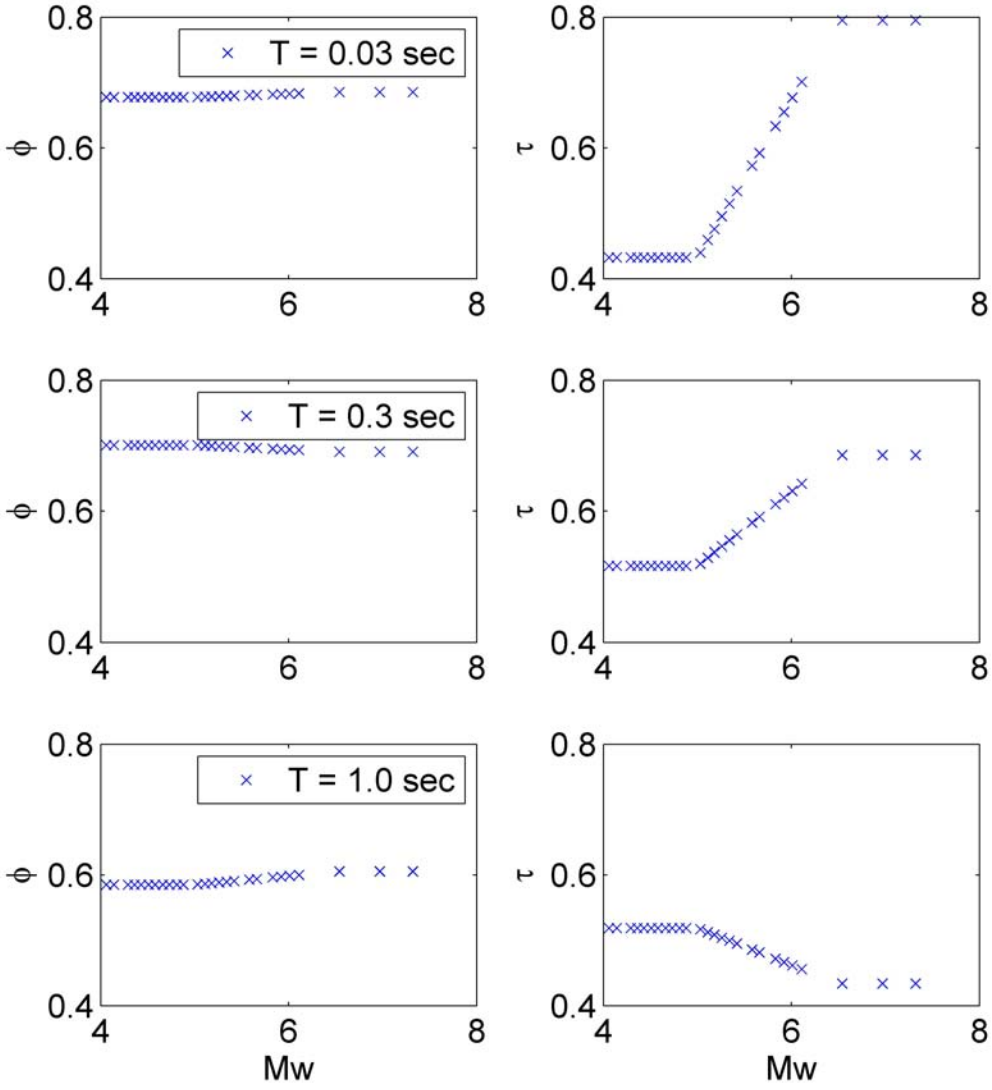


Figure 5.30. Surface intra- and inter-event standard deviations as a function of magnitude. Intra-event standard deviation presented in left column, inter-event standard deviation in right column.

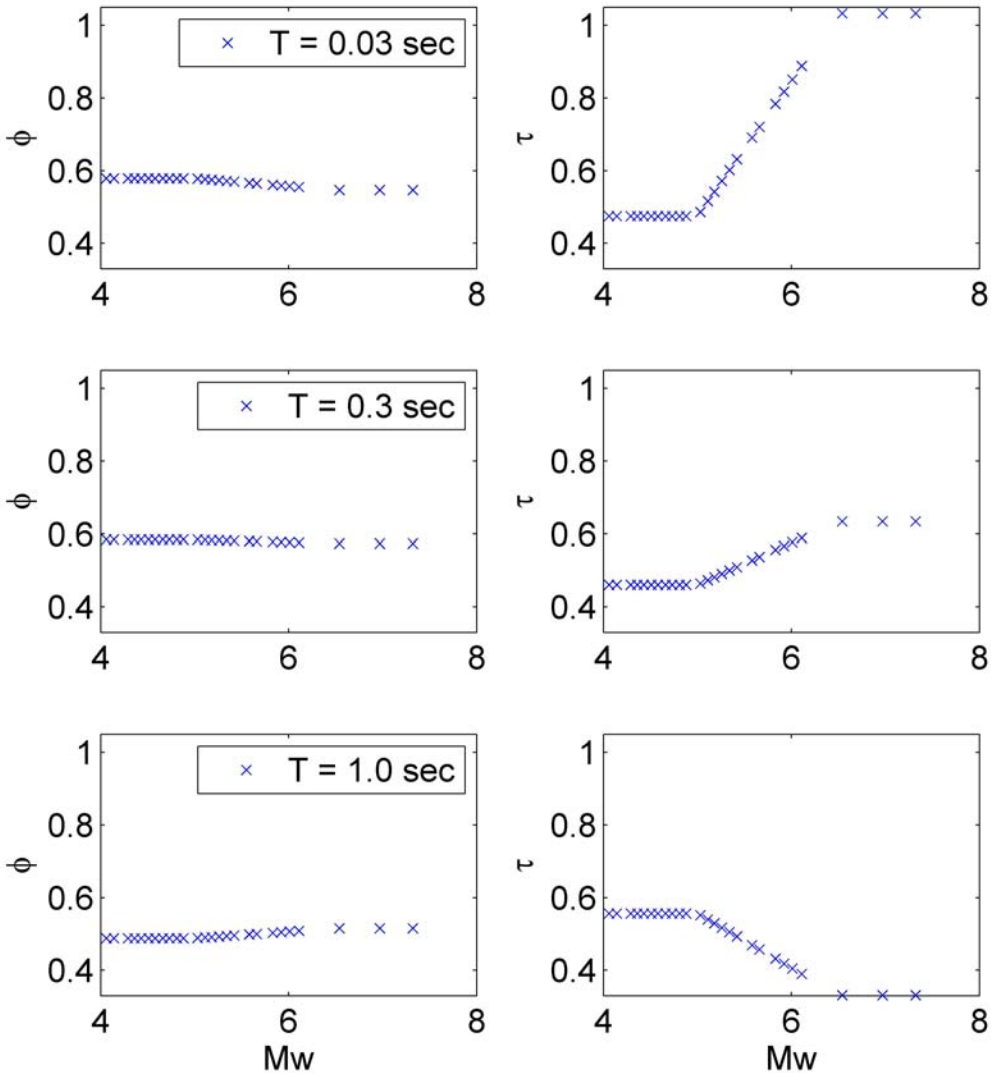


Figure 5.31. Borehole intra- and inter-event standard deviations as a function of magnitude. Intra-event standard deviation presented in left column, inter-event standard deviation in right column.

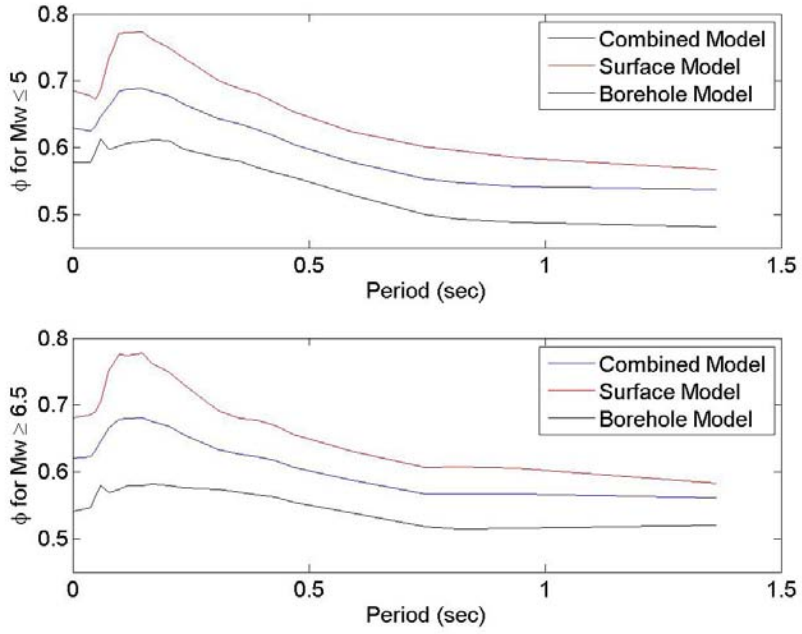


Figure 5.32. Comparison of the intra-event standard deviations -ϕ- for the Combined, Surface, and Borehole models.

event residuals). Figures 5.35 to 5.38 show the correlation between δB_e and δW_{es} for spectral periods of 0.03 and 1.0 seconds for the borehole and surface models. From the analysis it is concluded that the independence assumption is valid.

Figures 5.39 and 5.40 show the close match between Surface and Borehole inter-event residuals. The close agreement between the inter-event residuals obtained from the Combined model, and those from the Surface and Borehole models indicates that the often made assumption that borehole and surface ground motions have the same event residual is correct. Moreover, it also indicates that the Combined model gives about the same predictions as those given by the Surface and Borehole models alone.

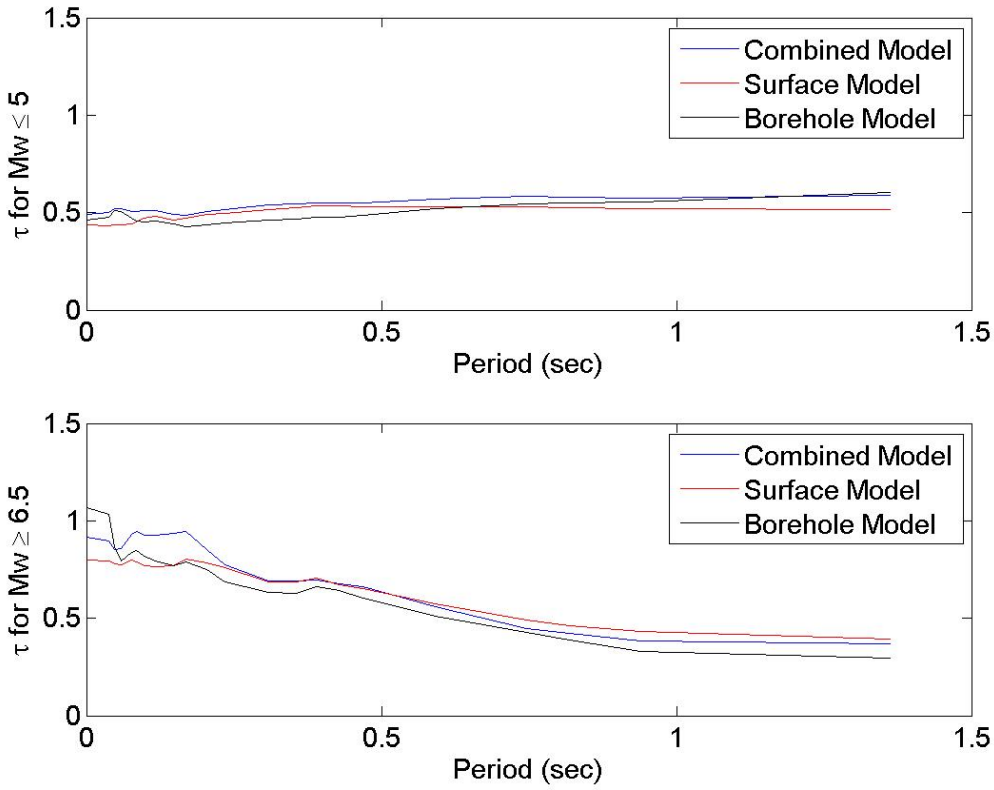


Figure 5.33. Comparison of the inter-event standard deviations -τ- for the Combined, Surface, and Borehole models.

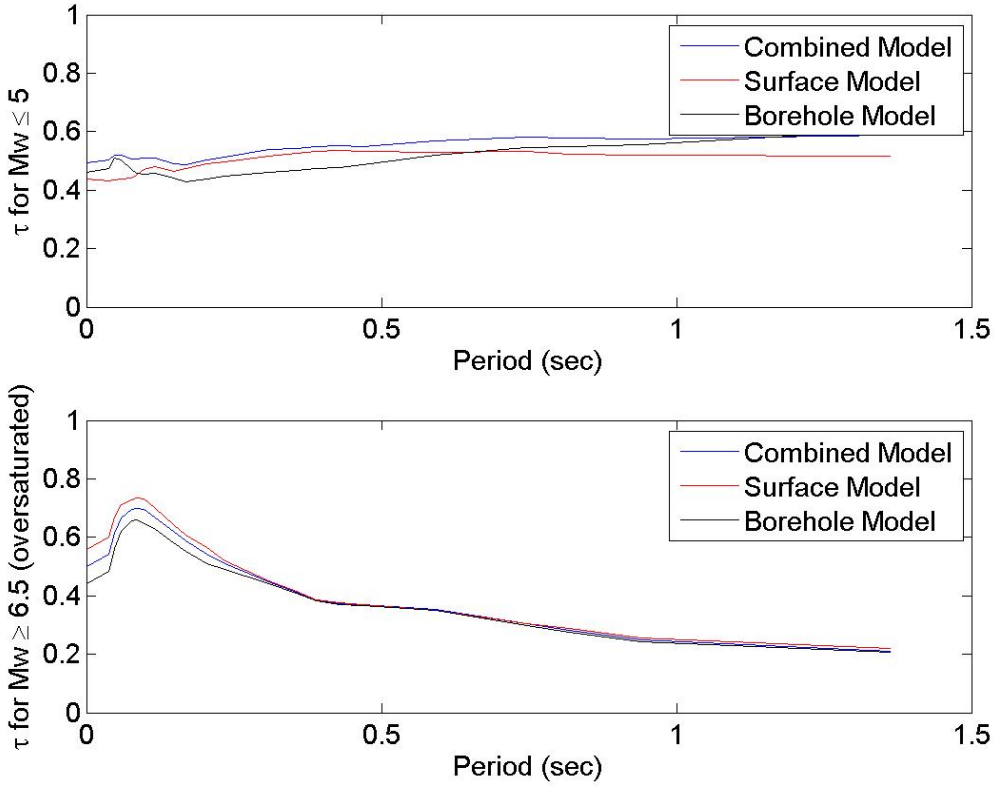


Figure 5.34. Comparison of the inter-event standard deviations -τ- for the Combined, Surface, and Borehole models allowing for oversaturation in the median estimate.

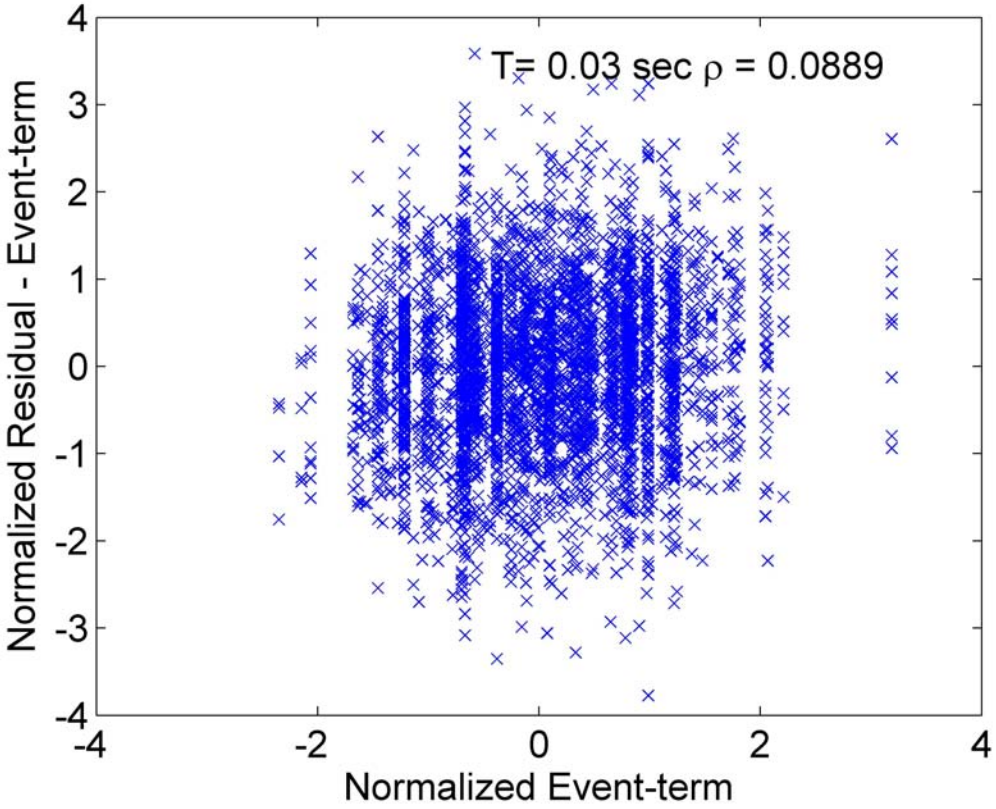


Figure 5.35. Correlation between surface Inter- and Intra-event residuals, for a spectral period of 0.03 seconds.

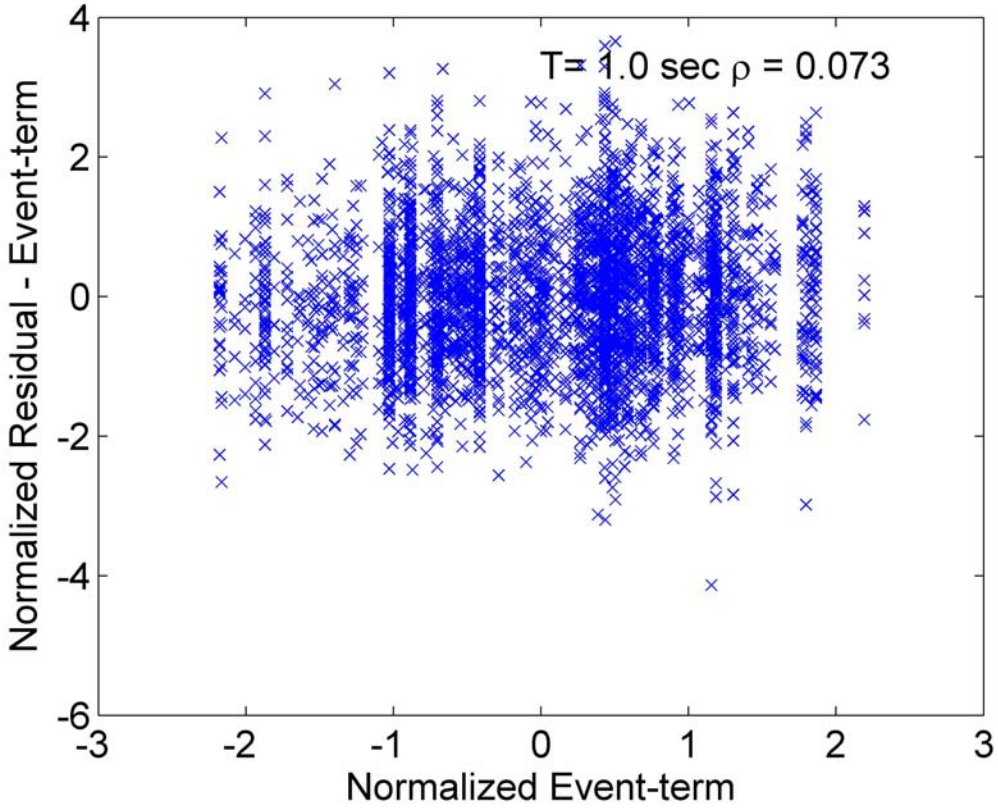


Figure 5.36. Correlation between surface Inter- and Intra-event residuals, for a spectral period of 1.0 second.

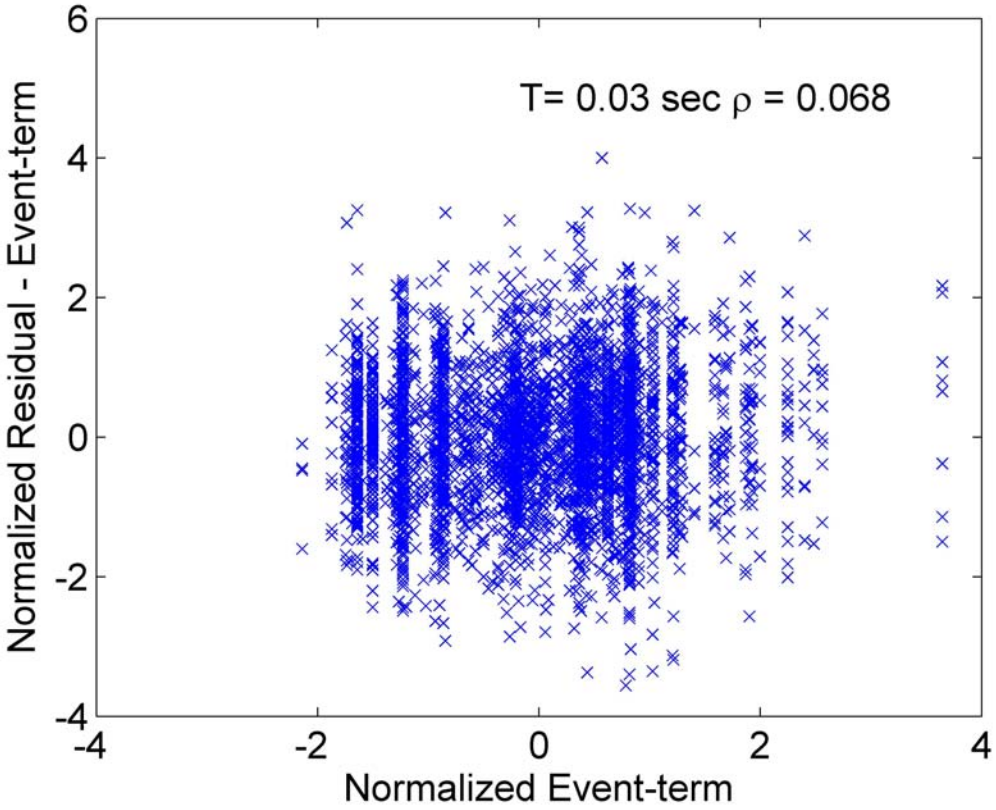


Figure 5.37. Correlation between borehole Inter- and Intra-event residuals, for a spectral period of 0.03 seconds.

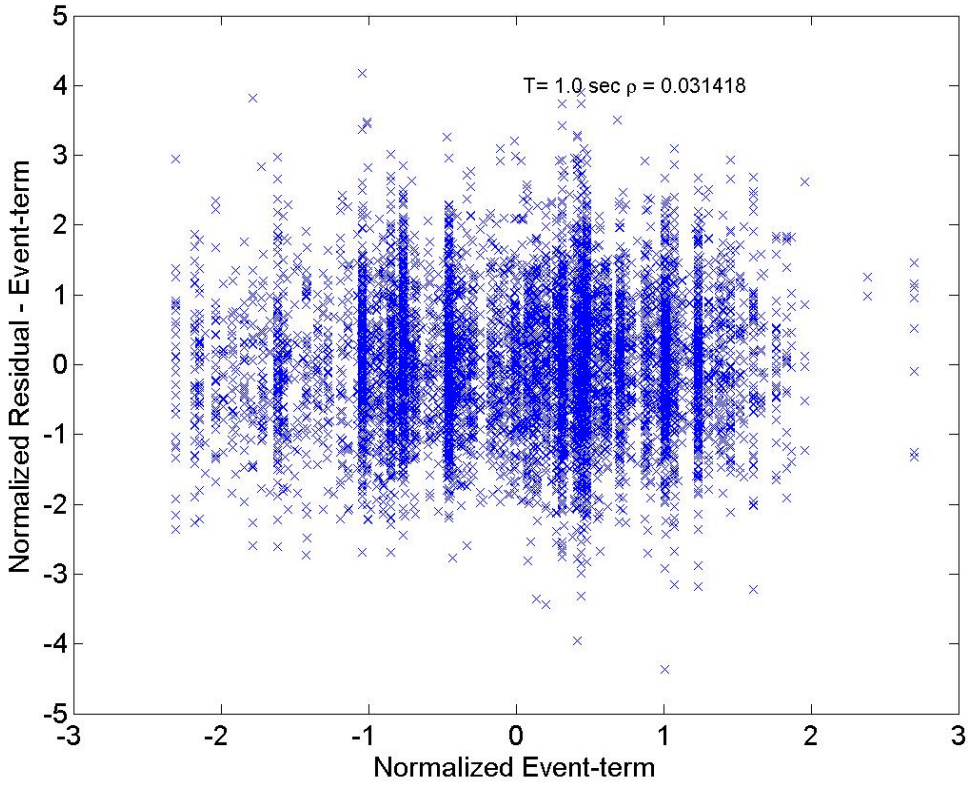


Figure 5.38. Correlation between borehole Inter- and Intra-event residuals, for a spectral period of 1.0 second.

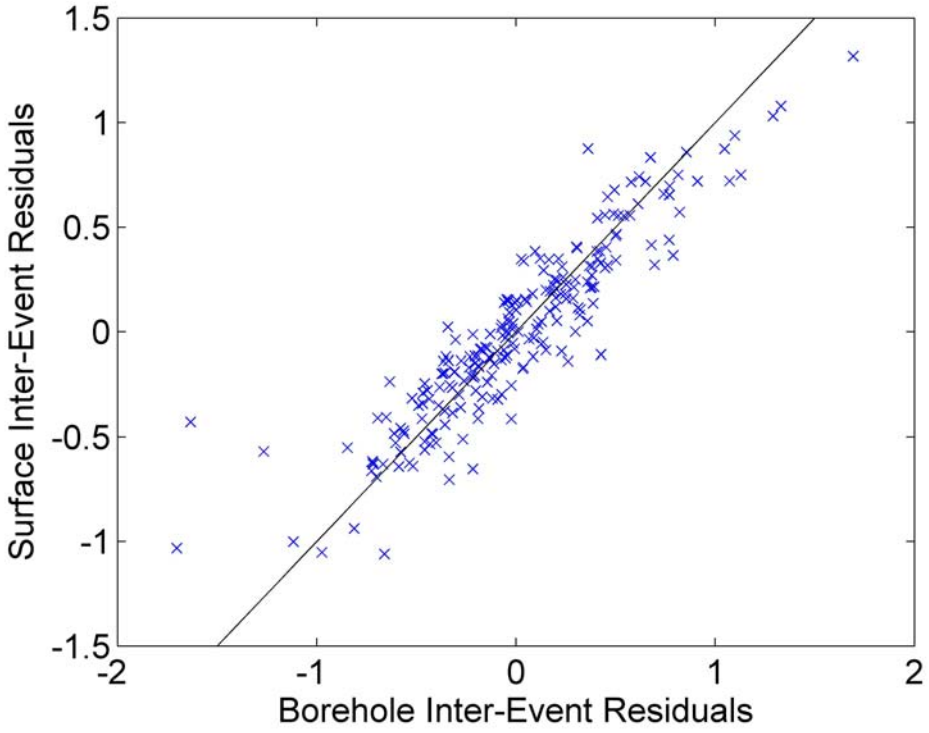


Figure 5.39. Surface Inter-Event terms versus Borehole Inter-Event terms for peak ground accelerations

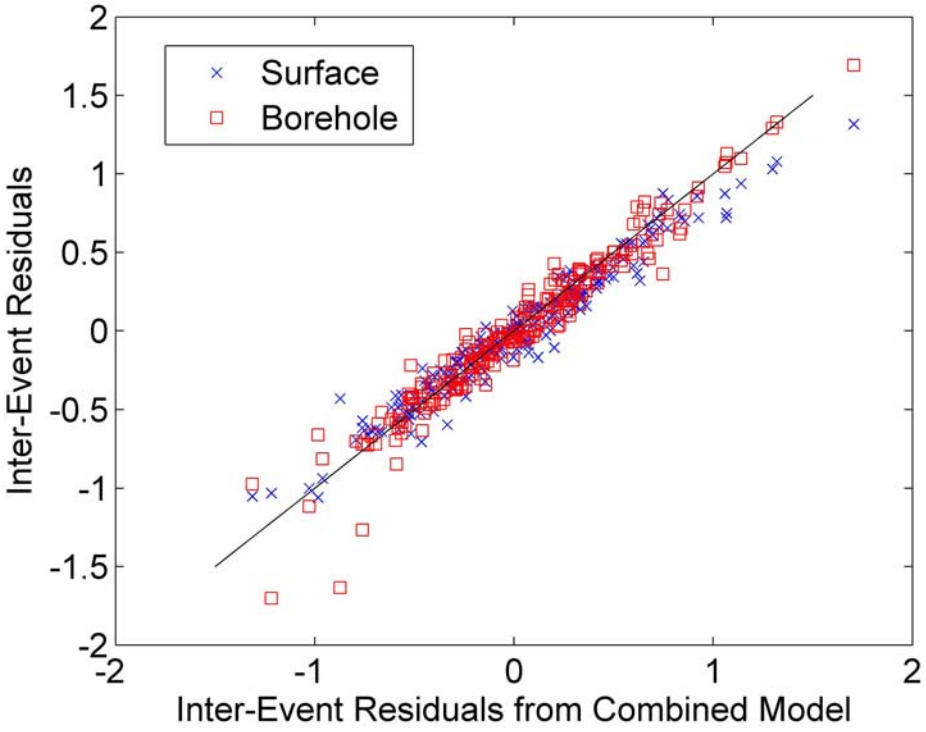


Figure 5.40. Surface and Borehole Inter-Event terms versus the Inter-Event terms obtained from the Combined model, for peak ground accelerations

5.3 Single-Station Standard Deviations

5.3.1 Introduction

Much recent work has been devoted to compute the standard deviation of recorded ground motions at a single station (e.g. Lin et al., 2010; Al-Atik et al., 2010). This work has been driven by the need to develop better estimates of standard deviation for non-ergodic $PSHA$ analyses (Anderson and Brune, 1999). In traditional (e.g., ergodic) $PSHA$, the underlying assumption is that the variability computed from an entire data set (e.g., including various sites and various sources for various events) is the same as the variability at a single site for various events. Removing the ergodicity assumption implies that the median ground motion at a given site can be estimated using independent means, and hence the standard deviation must also be reduced to reflect this additional information. As indicated in Chapter 3, single station standard deviations can be used as a lower bound value of the standard deviation that would be used in a PSHA analysis that removes the ergodic assumption on site response (Al-Atik et al., 2010). This Section presents single station standard deviations for the KiK-net database and compares the results with similar studies for other geographical regions. The results presented herein have the originality of being applicable both to the surface and to the borehole, since the KiK-net database has records at both the surface and at depth.

This section presents an analysis of single-station standard deviations using the KiK-net database. The approach taken was to develop a ground motion prediction equation specific to the KiK-net data (Section 5.2). A subset of stations that have recorded more than 10 records is selected and single station standard deviations are computed for these stations. These results are presented in Section 5.3.2. Section 5.3.3 looks at correlations between single station residuals and other subsets of residuals. An additional breakdown of residuals is studied in Section 5.3.4. Finally, the computed single-station standard deviations are

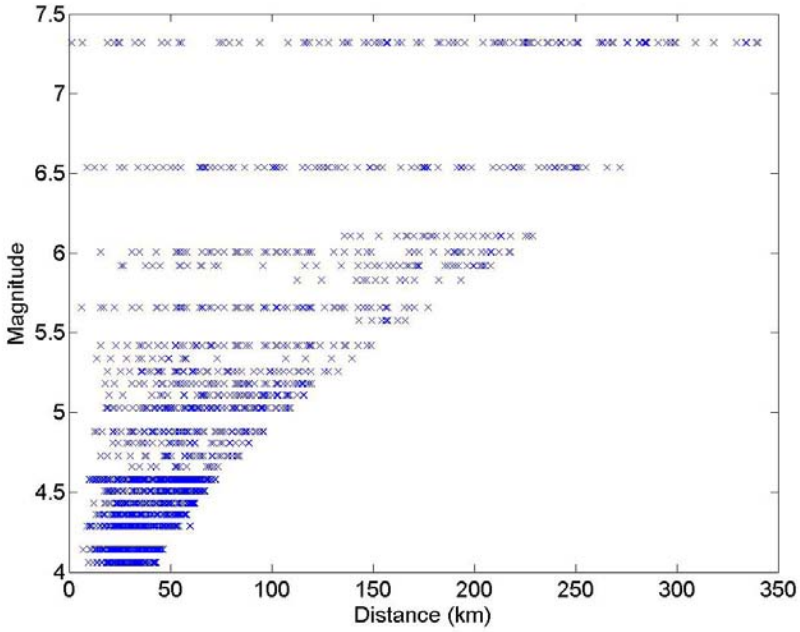


Figure 5.41. Magnitude-Distance distribution for the subset of stations with more than 10 records.

compared to those computed for other geographical regions.

5.3.2 Analysis of Residuals at Single-Stations

A subset of ground motions from the KiK-net database that corresponds to stations that recorded at least 10 events is selected. The magnitude-distance distributions of earthquakes recorded at these stations is shown in Figure 5.41. The station parameters (Vs30, Vshole, H800) of these stations are shown in Figure 5.42.

In order to compute the single station standard deviation, the intra-event residuals computed from the GMPE (Section 5.2) for each station are used to define the site term for each of the stations as given by

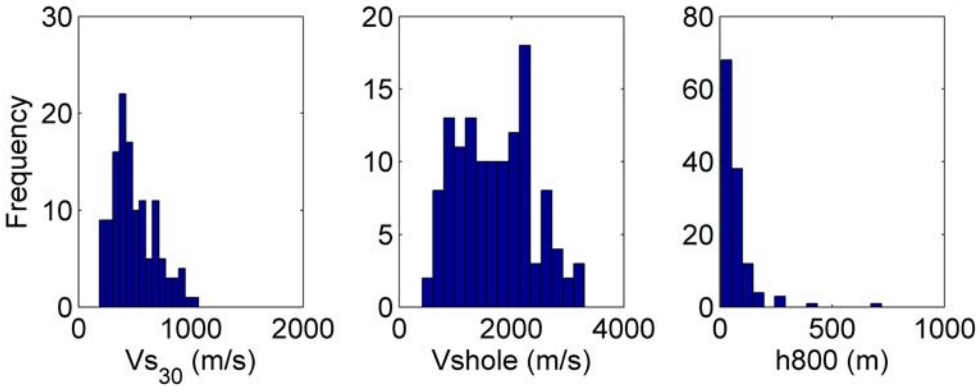


Figure 5.42. Station parameters for subset of stations with more than 10 records.

$$\delta S2S_s = \frac{1}{NE_s} \sum_{e=1}^{NE_s} \delta W_{es} \tag{5.11}$$

where, $\delta S2S_s$ is a random variable that represents the average within-event residual at each station and is hereby referred to as the site term. This is the average intra-event residual at site "s". Assuming that there is no bias in the subset of single-station records, this random variable is a zero mean random variable, and its standard deviation is denoted by ϕ_{S2S}. This standard deviation quantifies the site-to-site variability that is not explained by the ground motion prediction equation. The introduction of the site term permits a decomposition of the ground motion residuals as follows

$$\Delta = \delta W_{es} + \delta B_e = \delta S2S + \delta W_o + \delta B_e \tag{5.12}$$

Note that at a single station $\delta S2S_s$ takes a given, deterministic value, hence the single station ground motion variability is computed by eliminating this term from the residuals. δW_o is a residual variability. The single-station standard deviation can then be computed for the entire subset (of stations that recorded more than 10 earthquakes) using:

$$\phi_{SS} = \sqrt{\frac{\sum_{s=1}^{NS} \sum_{e=1}^{NE_s} (\delta W_{es} - \delta S2S_s)^2}{(\sum_{s=1}^{NS} NE_s) - 1}} \tag{5.13}$$

Equation 5.13 which assumes a homoscedastic model, is equivalent to a weighted average of the individual single-site intra-event standard deviation. ϕ_{SS} can then be allowed to vary with station or event parameters in order to capture any dependence on single station sigma. Alternatively, the single-station standard deviation can be computed for each station using Equation 5.14.

$$\phi_{SS_s} = \sqrt{\frac{\sum_{e=1}^{NE_s} (\delta W_{es} - \delta S2S_s)^2}{NE_s - 1}} \tag{5.14}$$

Figure 5.43 shows the within event residual at each of the N stations with more than 10 records for the PGA. The square symbols represent the mean residual at each station ($\delta S2S_s$) and the error bars correspond to a one standard deviation range (ϕss_s) for each station. The stations are placed in order of increasing number of records. The residuals corrected for the site term $\delta S2S_s$ are shown in Figures 5.44 to 5.46, both for the surface and the borehole, and for selected frequencies. These residuals are plotted versus magnitude. Observe that there is only a slight dependency on magnitude of ϕ_{ss}.

Figures 5.47 to 5.49 show the correlation between the intra-event residuals corrected for site-term ($\delta W_{es} - \delta S2S_s$) and the inter-event residuals or event terms (δB_e). Statistical independence between these random variables can be safely assumed from the low correlation coefficients at all periods. Hence, the total single-station standard deviation can then be computed as

$$\sigma_{ss} = \sqrt{\phi_{ss}^2 + \tau^2} \tag{5.15}$$

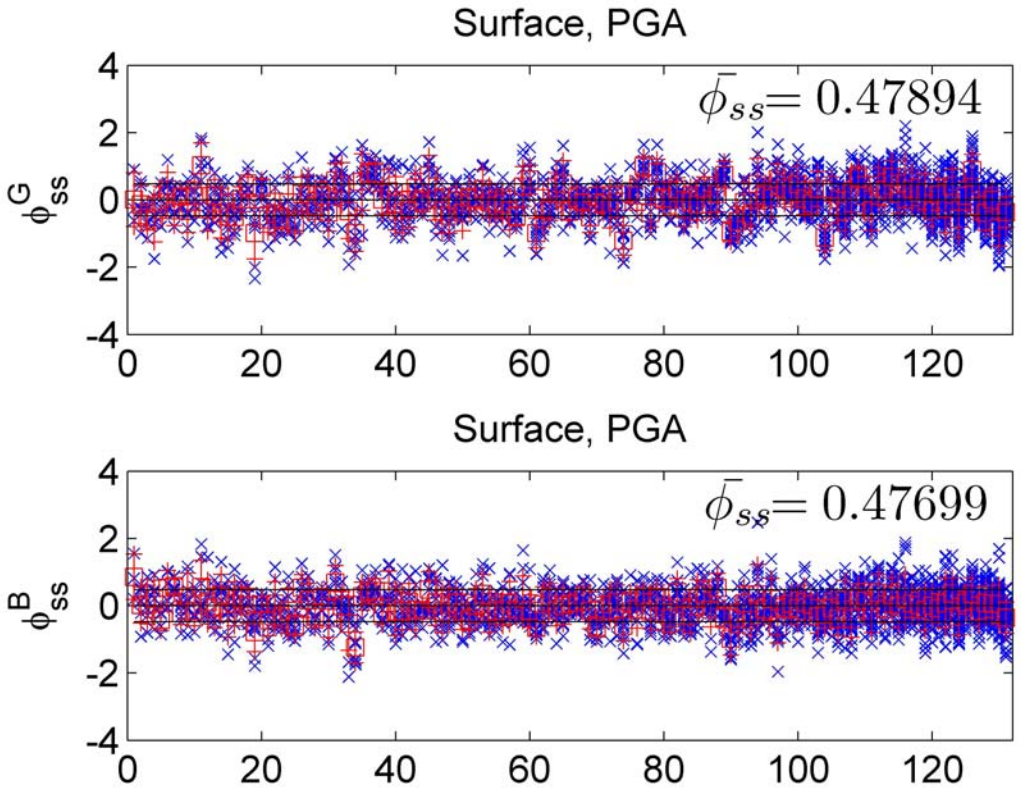


Figure 5.43. Intra-event residuals at stations with more than 10 records.

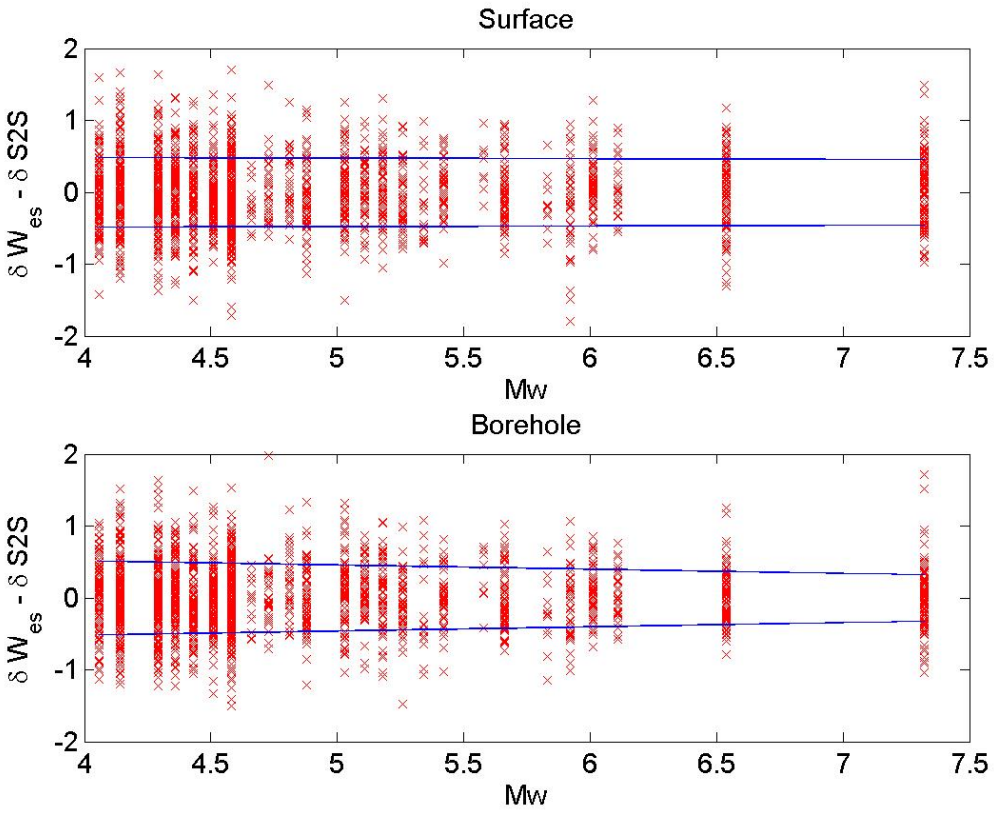


Figure 5.44. Intra-event residuals corrected for site term versus magnitude, for peak ground acceleration.

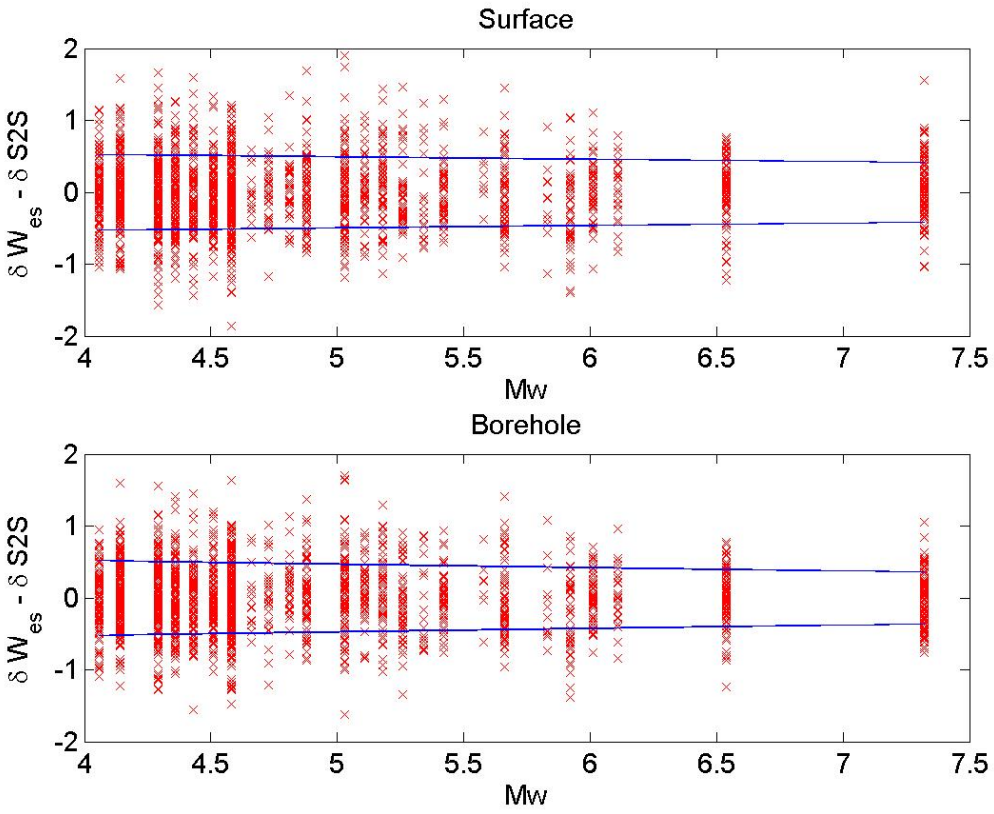


Figure 5.45. Intra-event residuals corrected for site term versus magnitude, for a spectral period of 0.3 seconds.

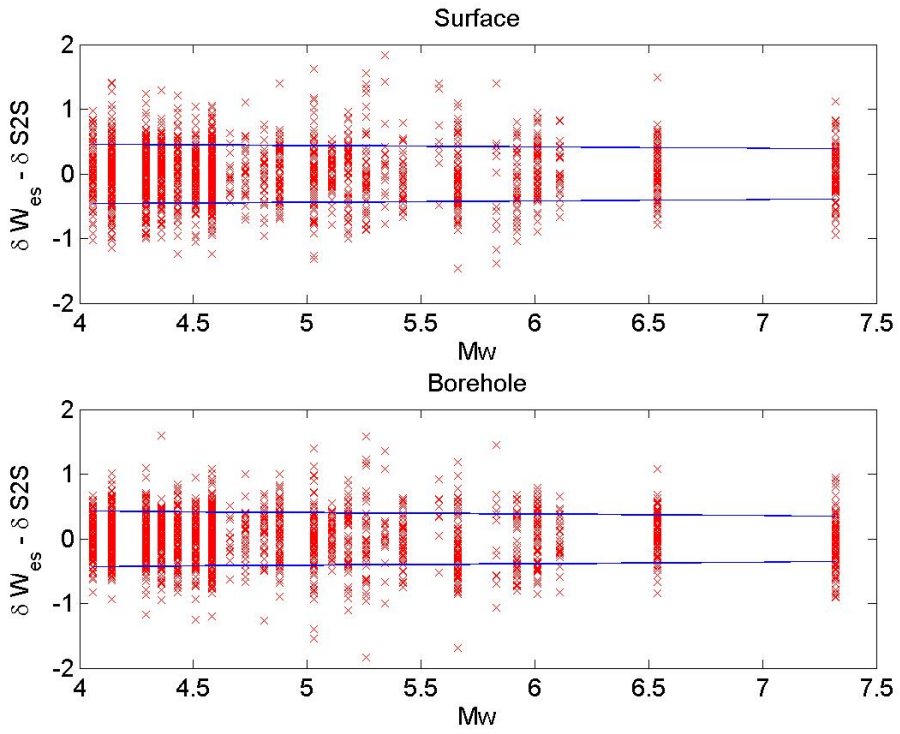


Figure 5.46. Intra-event residuals corrected for site term versus magnitude, for a spectral period of 1.0 second.

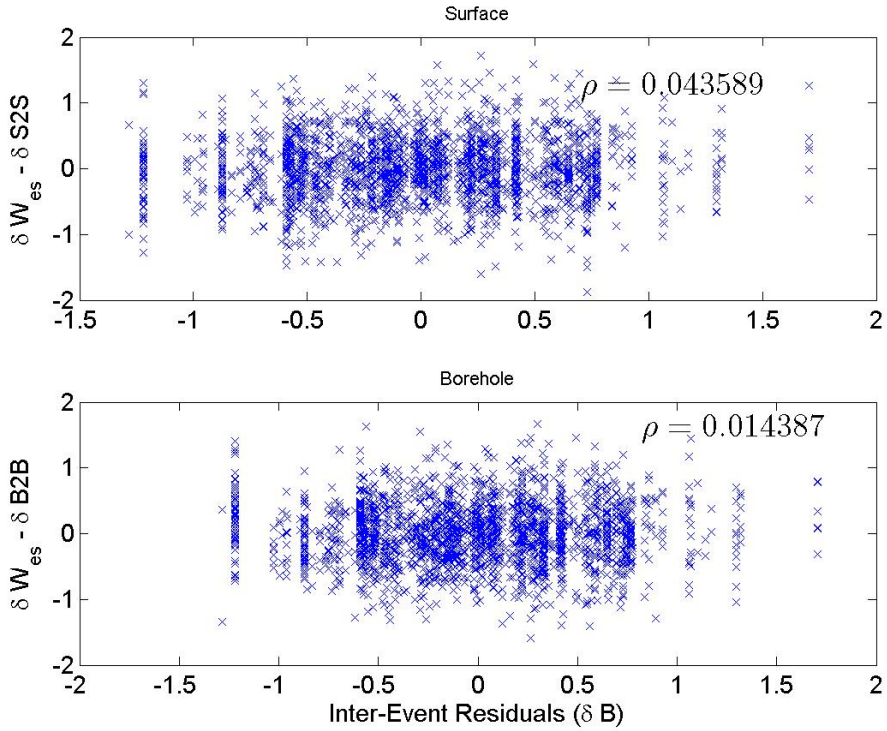


Figure 5.47. Correlation between intra-event residuals (corrected for the site term) and inter-event residuals, for peak ground acceleration.

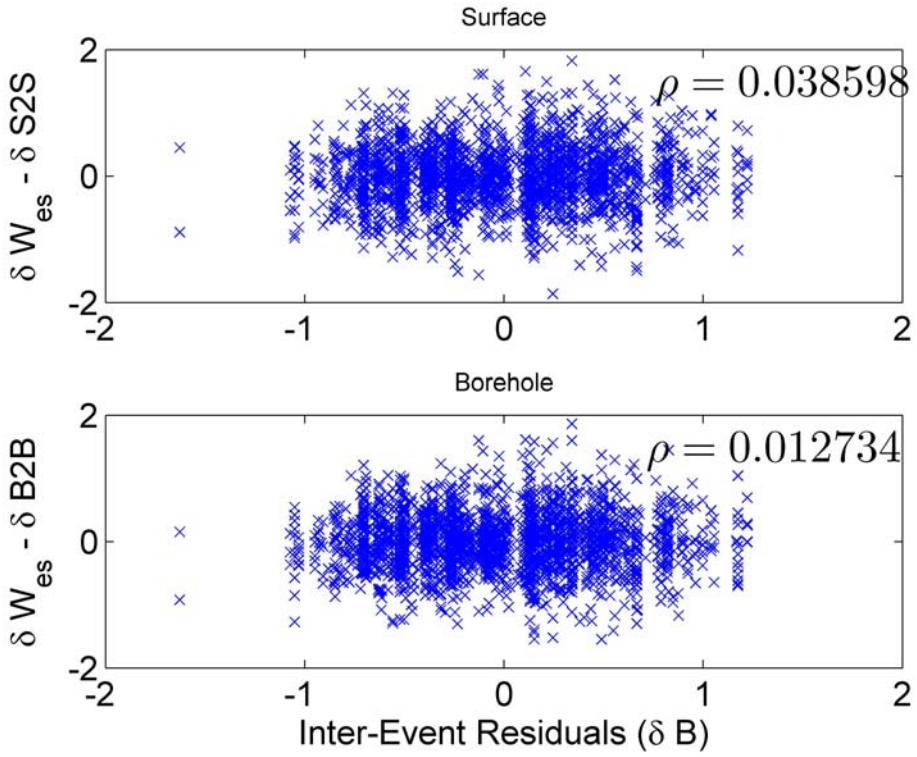


Figure 5.48. Correlation between intra-event residuals (corrected for the site term) and inter-event residuals, for a spectral period of 0.3 seconds.

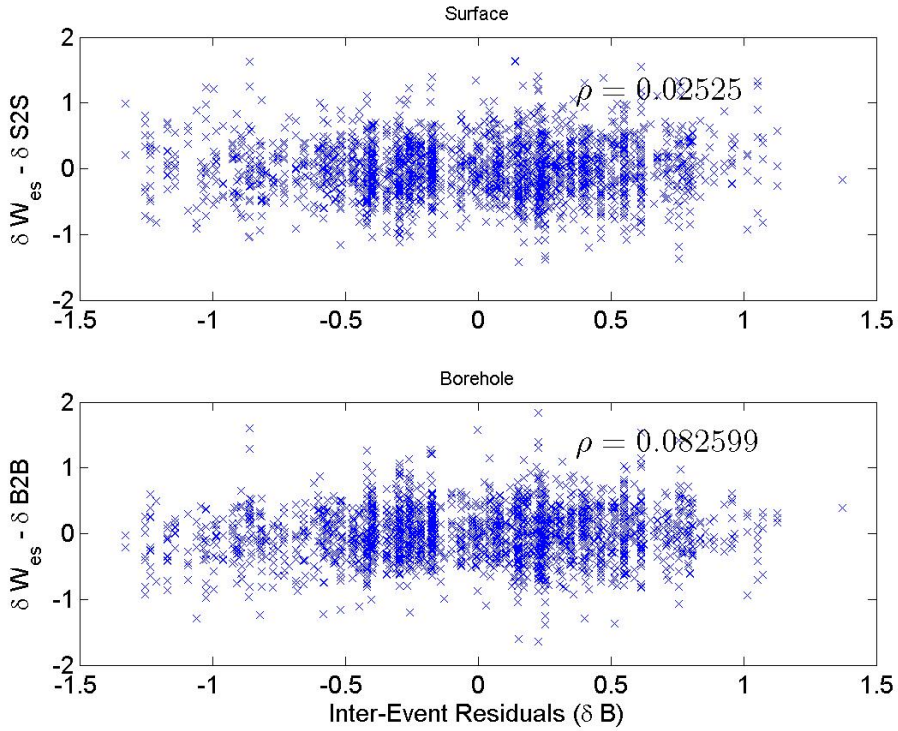


Figure 5.49. Correlation between intra-event residuals (corrected for the site term) and inter-event residuals, for a spectral period of 1.0 second.

The single station standard deviations that include the event term are shown in Figure 5.50. Recall that the magnitude scaling of the GMPE was not allowed to decrease with increasing magnitude (Section 5.2). As a result, there is a significant bias at high frequencies for earthquakes with magnitudes larger than 6.0. This bias, in turn, results in an overprediction of the between event variability at large magnitudes (τ). To limit the effect of this bias, the bias in the event terms for earthquakes with magnitude greater than 5.6 (Mh) was removed using a linear fit. The reduced values of τ are used to compute the σ_{ss} shown in Figure 5.50. The difference between single-station standard deviation of surface and borehole is consistently lower than the difference seen between standard deviation in the GMPEs (e.g. Table 5.3). The single station standard deviations for all periods are shown in Figure 5.51 both for the surface and the borehole records. Table 5.10 shows the single-station standard deviations.

The approach used to compute single station standard deviations presented above is the approach that was adopted by the PEGASOS project Renault et al. (2010).

Figures 5.52 to 5.54 show the mean of single station residuals ($\delta S2S_s$) versus Vs_{30} and site period (To). It is interesting to note that for long site period (e.g., soft soil sites), there is a negative bias in the residuals for the surface, which indicates that the GMPE is overpredicting ground motions for these stations (e.g. Figure 5.52). This is likely due to nonlinear soil behavior for these stations. Figure 5.52 to 5.54 also show the intra-event single-station standard deviation (ϕ_{ss_s}) as a function of site parameters. Observe that the standard deviation of the $\delta S2S_s$ term for the surface appear to decrease slightly for increasing values of To, this is observed at all frequencies.

Figures 5.55 to 5.57 show the dependency of intra-event single-station residuals ($\delta W - \delta S2S$) versus site conditions for PGA, $T = 0.3$, and $T = 1.0$. The dependency of intra-event single-station residuals on Vs_{30} shows no trend until large Vs_{30} values. For stations with large Vs_{30} values (i.e. $Vs_{30} > 900$ m/s), the standard deviation of these residuals (ϕ_{ss}) is

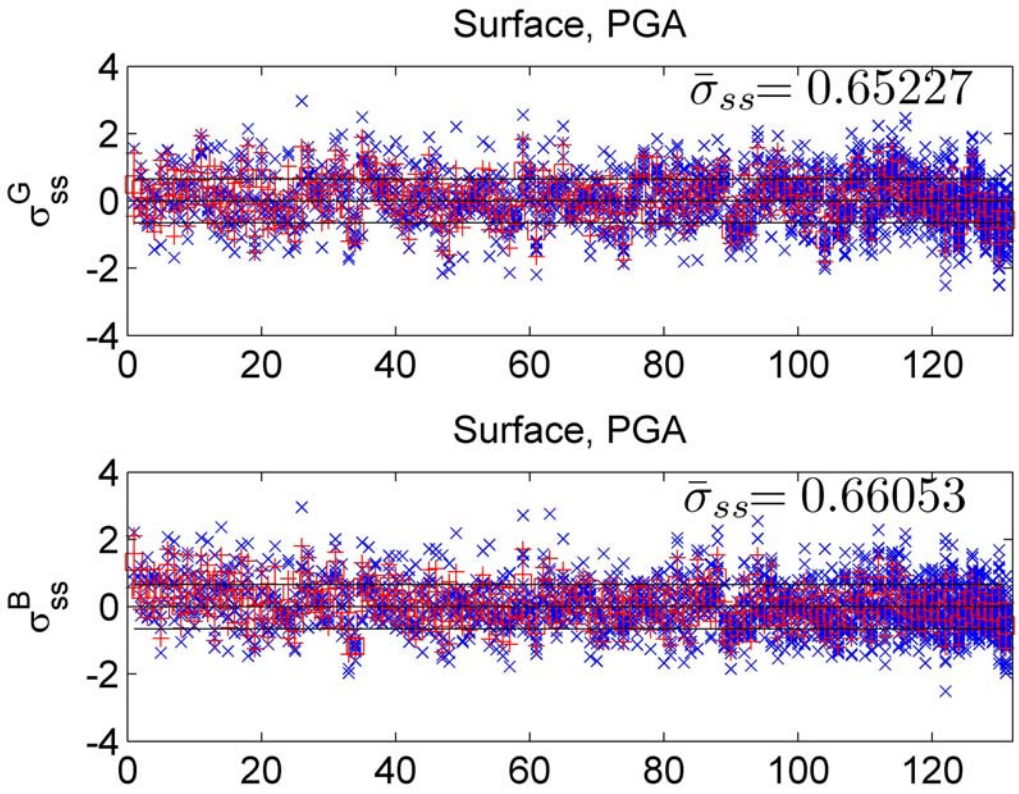


Figure 5.50. Single station standard deviations for stations with more than 10 records, for peak ground acceleration

equal to 0.44 (log units, at borehole for PGA), lower than the 0.51 value for lower Vs_{30} values. If we consider only the Vs_{30} values that correspond to engineering bedrock (Vs_{30} $>$ 760 m/s), the ϕ_{ss} value is 0.49. Note that this findings are true for borehole but not for surface.

Figures 5.58 to 5.60 show the standard deviation of the empirical amplification factor (in log-space) computed by dividing the ground motion at the surface over the ground motion at depth for each of the 131 stations with more than 10 recordings. Using the notation described above, this amplification factor corresponds to:

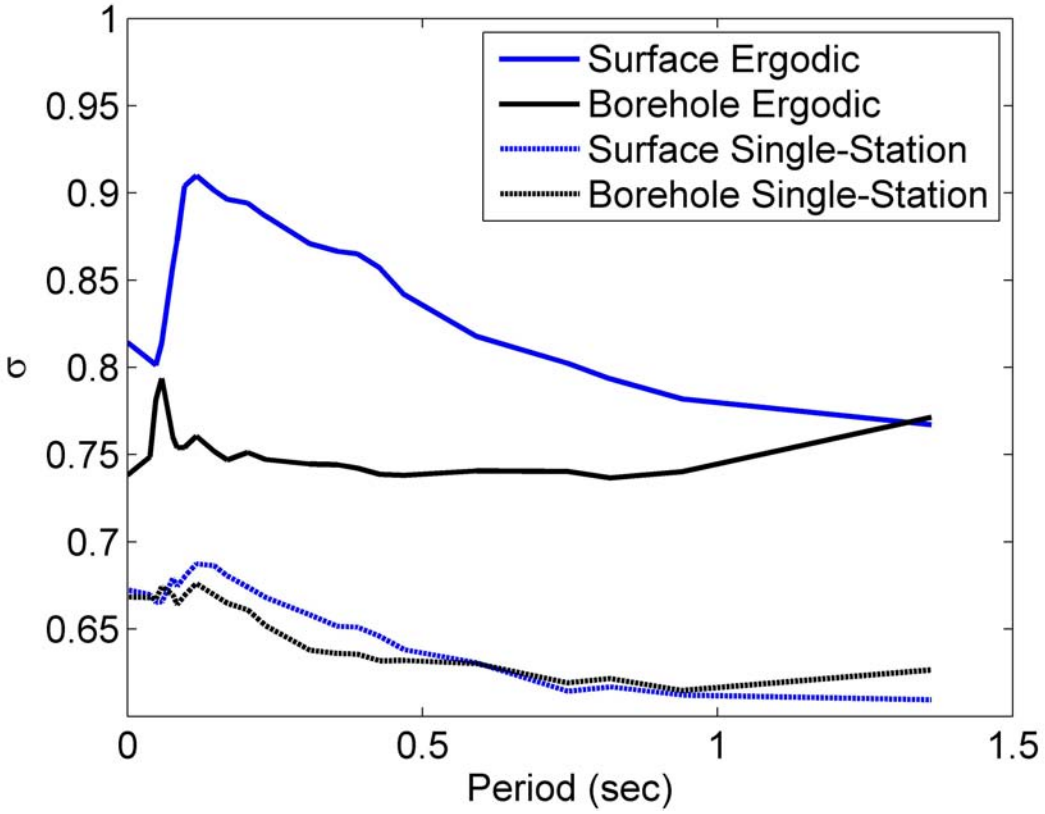


Figure 5.51. Total standard deviations compared for single-sites and the ergodic prediction.

$$AF = y^G - y^B \tag{5.16}$$

Note from Figures 5.58 to 5.60 that the standard deviation of the amplification factor tend to decrease for high Vs_{30}. This observation, combined with the previous one that Vs_{30} does not influence the observed residuals, could indicate that the GMPE (i.e. the median model) is strongly influenced by the sites where most of the data is (i.e. low Vs_{30}) inducing a slight bias for high Vs_{30}.

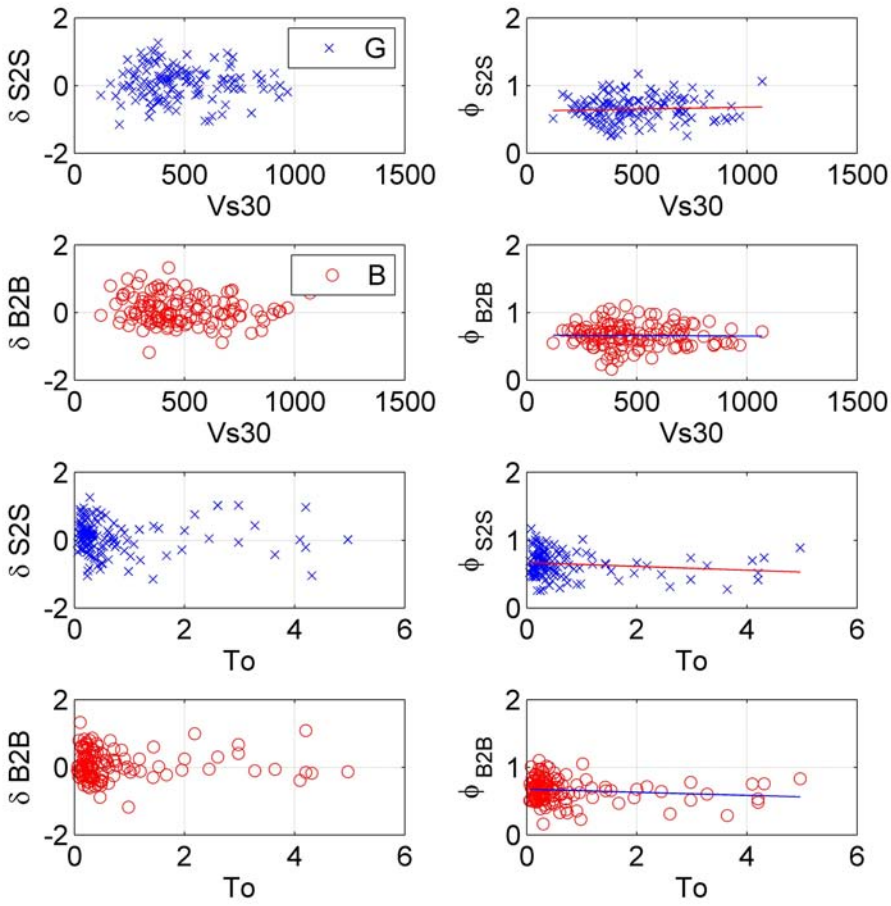


Figure 5.52. Mean of single-station residuals versus Vs_{30} and site period (To). Left column shows the mean residuals ($\delta S2S$) and right column shows its standard deviation (ϕ_{S2S}). Results for peak ground acceleration.

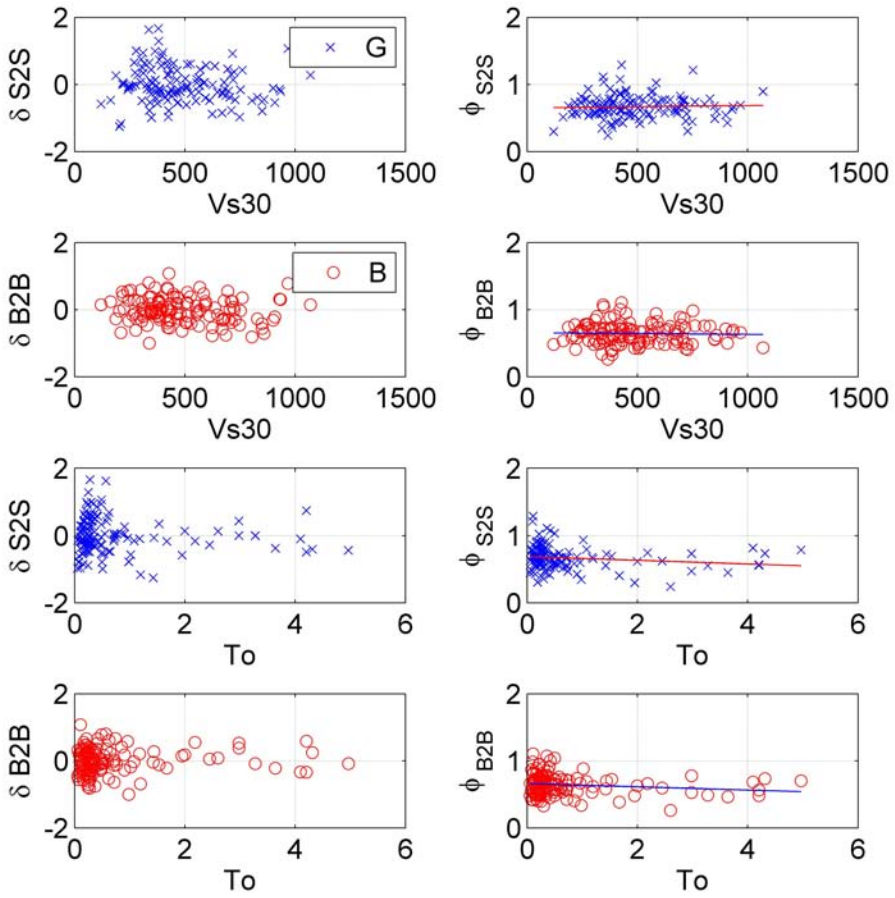


Figure 5.53. Mean of single-station residuals versus Vs_{30} and site period (To). Left column shows the mean residuals ($\delta S2S$) and right column shows its standard deviation (ϕ_{S2S}). Results for spectral acceleration of 0.3 seconds.

Table 5.10. Single-Station Standard Deviations for Surface and Borehole

Period	Surf ϕ	Borehole ϕ	Surf σ	Borehole σ
PGA	0.4967	0.5060	0.6725	0.6684
0.0384	0.4857	0.4849	0.6698	0.6681
0.0484	0.4725	0.4670	0.6653	0.6672
0.0582	0.4660	0.4722	0.6655	0.6744
0.0769	0.4774	0.4725	0.6791	0.6697
0.0844	0.4837	0.4819	0.6746	0.6641
0.0970	0.4912	0.4945	0.6800	0.6693
0.1167	0.5028	0.4999	0.6874	0.6761
0.1472	0.5149	0.5024	0.6864	0.6696
0.1691	0.5124	0.5049	0.6805	0.6648
0.2036	0.5067	0.4994	0.6741	0.6610
0.2340	0.5010	0.4865	0.6682	0.6520
0.3090	0.4812	0.4664	0.6582	0.6379
0.3551	0.4777	0.4573	0.6517	0.6361
0.3896	0.4653	0.4459	0.6511	0.6357
0.4274	0.4583	0.4427	0.6459	0.6319
0.4690	0.4531	0.4410	0.6382	0.6320
0.5913	0.4442	0.4350	0.6307	0.6302
0.7456	0.4258	0.4120	0.6144	0.6192
0.8180	0.4264	0.4086	0.6168	0.6216
0.9401	0.4258	0.4034	0.6122	0.6148
1.3622	0.4178	0.3870	0.6095	0.6266

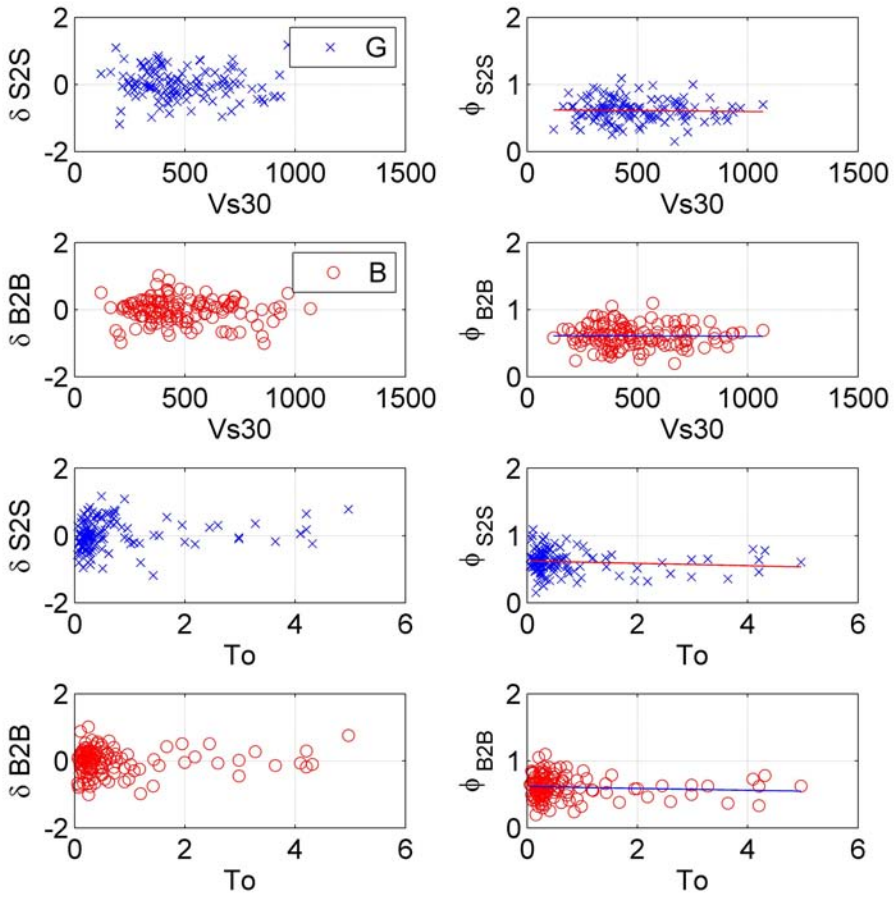


Figure 5.54. Mean of single-station residuals versus Vs_{30} and site period (To). Left column shows the mean residuals ($\delta S2S$) and right column shows its standard deviation (ϕ_{S2S}). Results for spectral acceleration of 1.0 second.

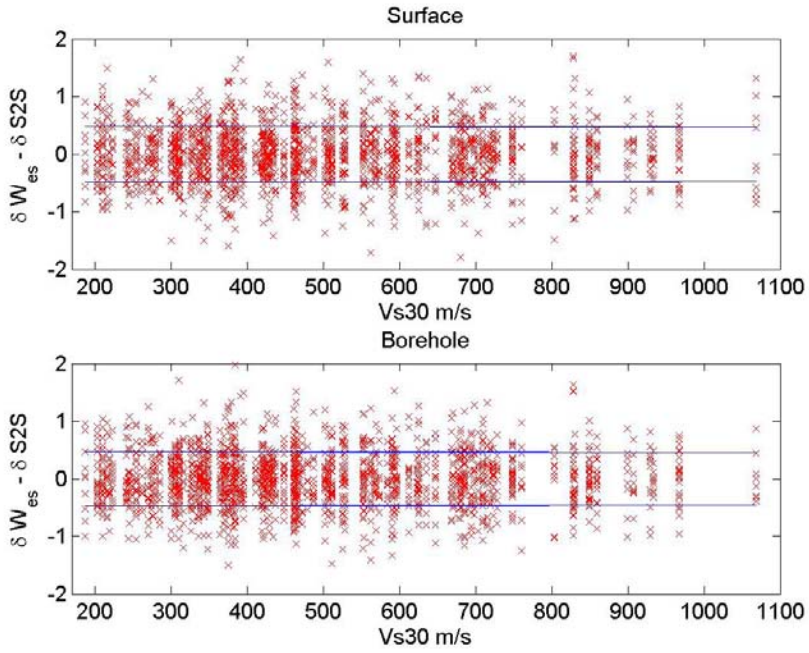


Figure 5.55. Intra-event single-station residuals a versus Vs_{30} for peak ground acceleration.

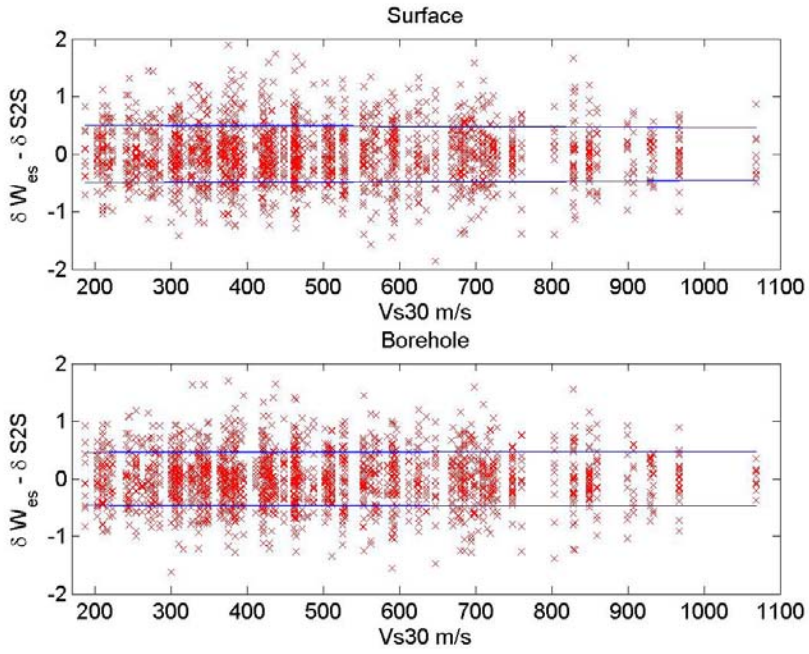


Figure 5.56. Intra-event single-station residuals a versus Vs_{30} for a spectral period of 0.3 seconds.

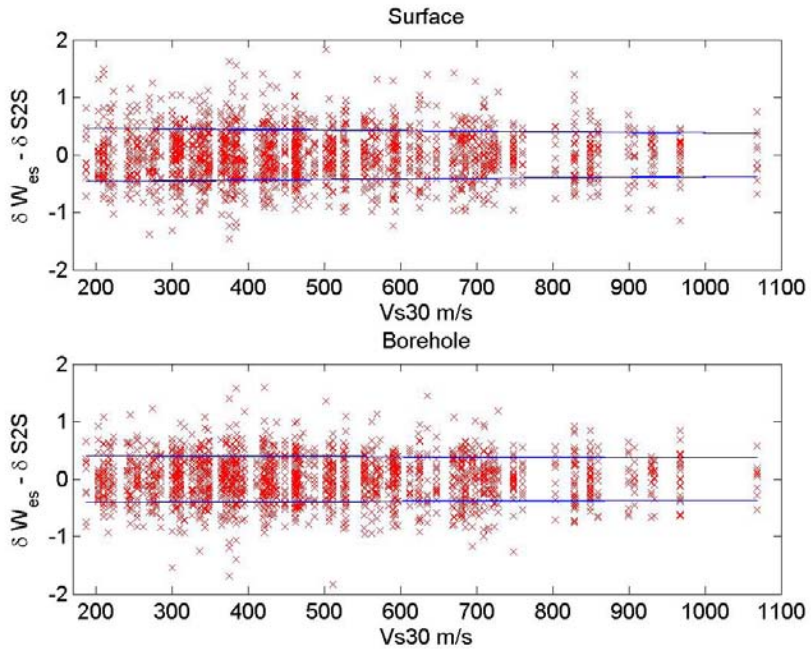


Figure 5.57. Intra-event single-station residuals a versus Vs_{30} for a spectral period of 1.0 second.

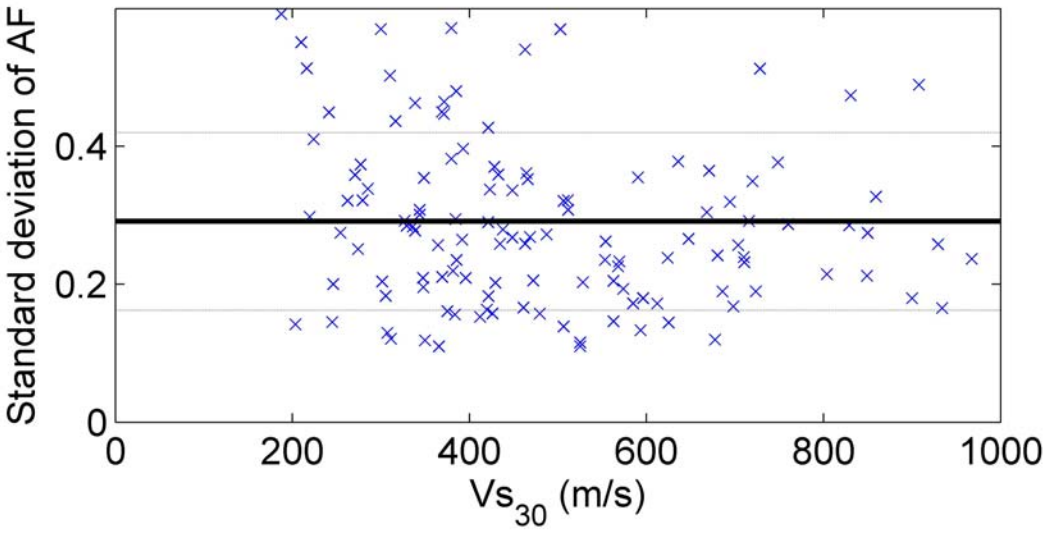


Figure 5.58. Standard deviation of empirical amplification factor for peak ground acceleration.

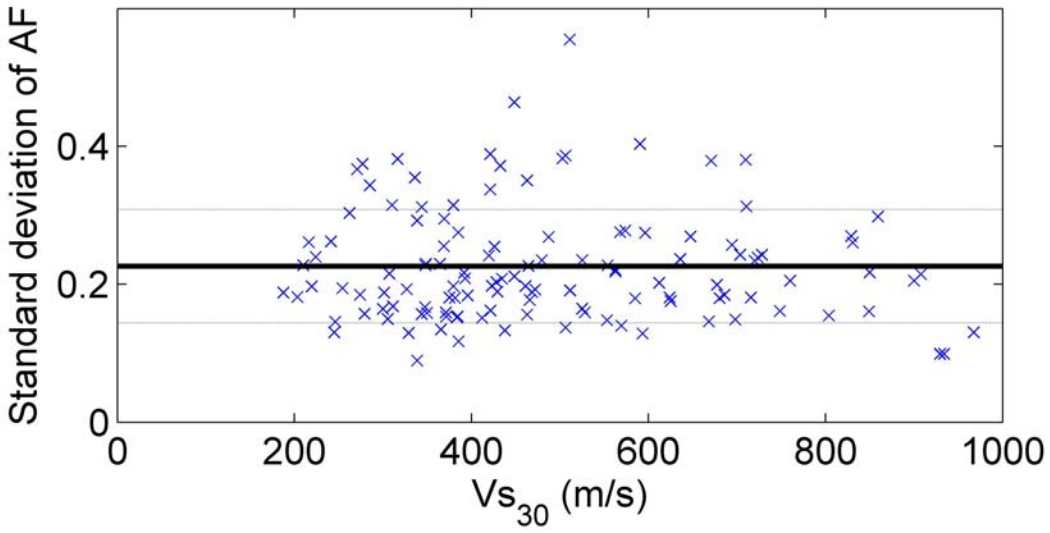


Figure 5.59. Standard deviation of empirical amplification factor for a spectral period of 0.3 seconds.

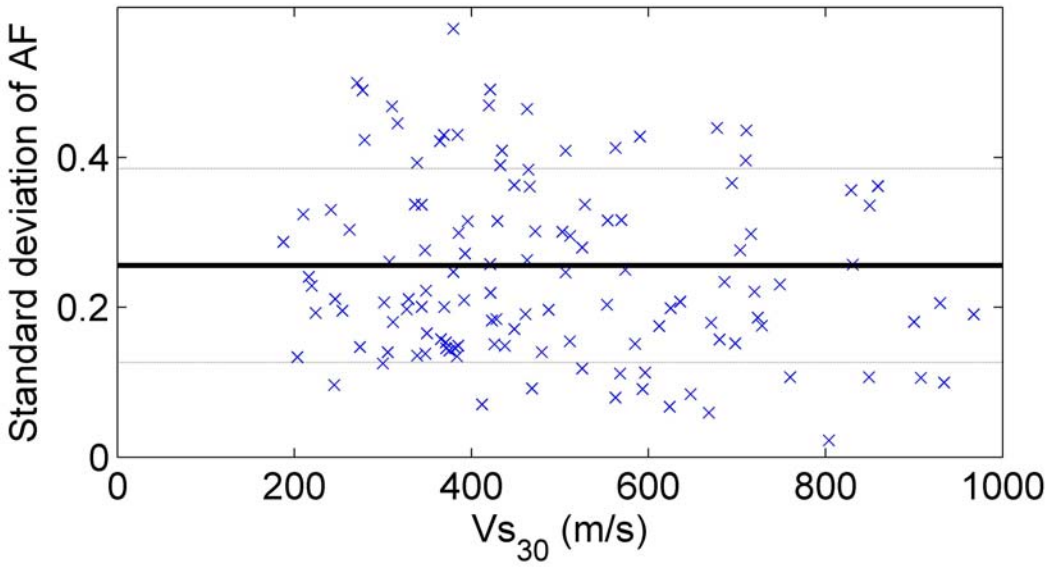


Figure 5.60. Standard deviation of empirical amplification factor for a spectral period of 1.0 second.

5.3.3 Study of Cross-correlation between Residuals

When dealing with borehole and surface records, surface ground motion can be written as,

$$y^G = \mu^B + \mu^{AMP} + \delta W^B + \delta S2S^{AMP} + \delta AMP + \delta B^B \tag{5.17}$$

where, μ^B is the median prediction for the borehole, μ_{AMP} is the median amplification factor predicted by the GMPE, function of Vs_{30} and/or other site parameters (see Equation 5.5), δW^B is the intra-event residual at borehole, $\delta S2S^{AMP}$ is the mean residual among sites with the same site parameterization (e.g. same Vs_{30} and $h800$), δAMP is the residual in site amplification within each site, and δB^B is the inter-event term at borehole. The sum of the terms $\delta S2S^{AMP}$ and δAMP represents the residual of the empirical amplification factor. If δB^G is assumed to be equal to δB^B, which is consistent with the results presented in Section 5.2 (see Figure 5.34) and it is imposed in the Combined model, Equation 5.17 can be rewritten as,

$$y^G = \mu^B + \mu^{AMP} + \delta W^B + \delta S2S^{AMP} + \delta AMP + \delta B^G \tag{5.18}$$

or equivalently,

$$y^G = \mu^G + \delta W^B + \delta S2S^{AMP} + \delta AMP + \delta B^G \tag{5.19}$$

This implies that,

$$\sigma^G = \sqrt{(\phi^B)^2 + (\phi_{S2S^{AMP}})^2 + \phi^2_{AMP} + \tau^2} \tag{5.20}$$

For this to be true, δW^B, $\delta S2S^{AMP}$, δAMP and δB^G must be independent random variables.

Computation of $\delta S2S^{AMP}$ The median estimate at borehole is given by,

$$y^B = \mu^B + \delta W^B + \delta B^B \tag{5.21}$$

and at the ground surface by,

$$y^G = \mu^G + \delta W^G + \delta B^G \tag{5.22}$$

subtracting them, the empirical amplification factor (AF) is obtained. Recalling that the intensities y^B and y^G are in logarithmic units the equation for AF is then,

$$AF = y^G - y^B = \underbrace{\mu^G - \mu^B}_{\mu^{AMP}} + \underbrace{(\delta W^G - \delta W^B)}_{\Delta AMP} \tag{5.23}$$

where, ΔAMP is the difference between the intra-event residuals at the surface and borehole. Combining Equations 5.19 and 5.23, ΔAMP can be rewritten as,

$$\Delta AMP = \delta W^G - \delta B^B = \delta S2S^{AMP} + \delta AMP \tag{5.24}$$

where, δAMP is a zero mean, normally distributed random variable with standard deviation ϕ_{AMP}, and $\delta S2S^{AMP}$ is also a zero mean, normally distributed random variable with standard deviation $\phi_{S2S^{AMP}}$. Table 5.11 shows the correlation coefficients between the residuals in Equation 5.20.

Table 5.11. Single-Station Residuals correlation Coefficients

	PGA	0.3 (sec)	1.0 (sec)
δW^B & $\delta S2S^{AMP}$	-0.1178	-0.1572	-0.1591
δW^B & δAMP	-0.2875	-0.1500	-0.2222
δW^B & δB^G	-0.0911	-0.0072	0.0777
$\delta S2S^{AMP}$ & δAMP	0.0000	0.0000	0.0000
$\delta S2S^{AMP}$ & δB^G	0.0236	0.0645	-0.0331
δAMP & δB^G	0.0459	0.0547	-0.0789

Note the negative correlation between the intra-event residual and mean residual amplification ($\delta S2S^{AMP}$), this dependency is what one would expect if nonlinear effects were taking place. The residual amplification, corrected by $\delta S2S^{AMP}$, δAMP is again negatively correlated with the intra-event residual, a further indication on of nonlinear effects. Notably, the correlation between $\delta S2S^{AMP}$ and δAMP extremely low. The remaining pairs are poorly correlated.

Figure 5.61 shows the correlation coefficients between all the components of the total residual at the ground surface. Since the random variables represented by their statistics in Equation 5.20 are not uncorrelated, Equation 5.20 should include these correlations as,

$$\begin{aligned}\sigma^G = &\sqrt{(\phi^B)^2 + (\phi_{S2S^{AMP}})^2 + \phi^2_{AMP} + \tau^2 + \cdots} \\ &\cdots\overline{2 * \rho_{\delta W^B - \delta S2S^{AMP}} * \phi^B * \phi_{S2S^{AMP}} + 2 * \rho_{\delta W^B - \delta AMP} * \phi^B * \phi_{AMP} + \cdots} \\ &\cdots\overline{2 * \rho_{\delta W^B - \delta B^G} * \phi^B * \tau + 2 * \rho_{\delta S2S^{AMP} - \delta AMP} * \phi_{S2S^{AMP}} * \phi_{AMP} + \cdots} \\ &\cdots\overline{2 * \rho_{\delta S2S^{AMP} - \delta B^G} * \phi_{S2S^{AMP}} * \tau + 2 * \rho_{\delta AMP - \delta B^G} * \phi_{AMP} * \tau}\end{aligned} \tag{5.25}$$

where, all the correlation coefficients needed are given in Table 5.11. The standard deviations of each of the involved random variables are given in Table 5.12, to be consistent with the estimates of the single station components of the Equation 5.25, the inter-event residual standard deviation (τ) is updated to represent only those events sampled in the

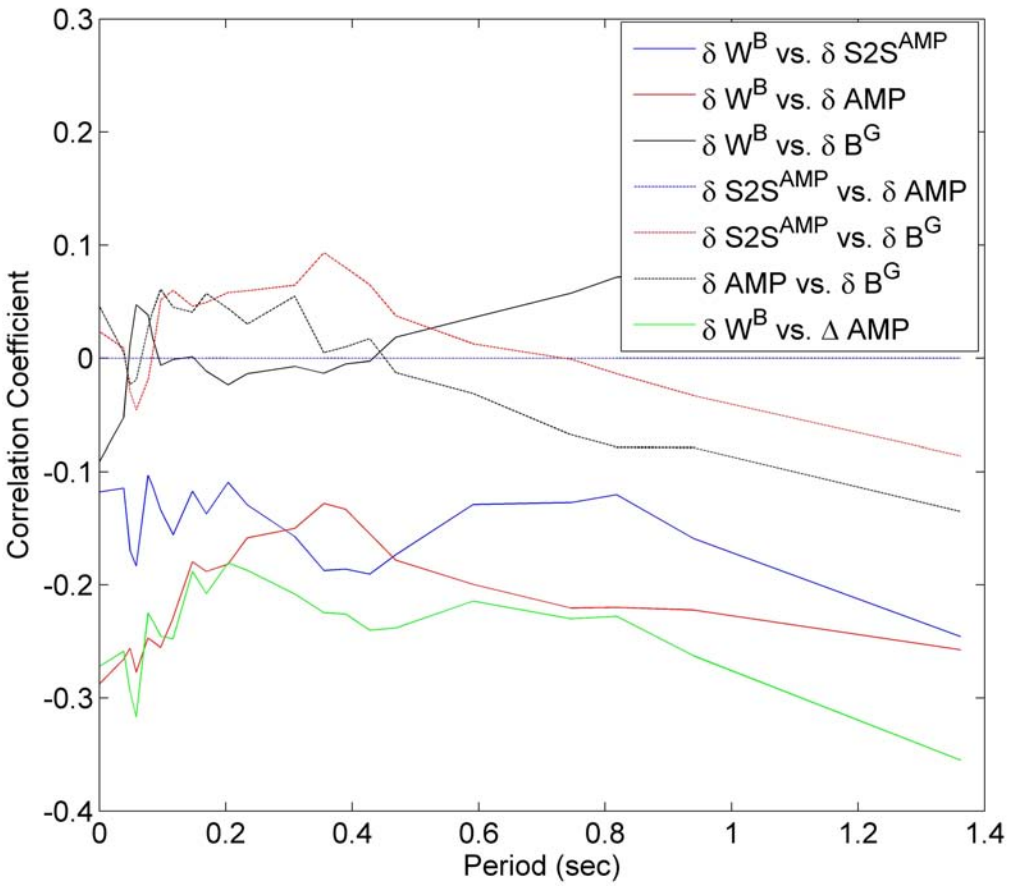


Figure 5.61. Correlation coefficient between the different components of the total ground surface residual. Correlation Coefficients are plotted against spectral period.

single-station subset 10 or more times. This change is also necessary to add the magnitude independent single-station residual variances, with inter-event variances that otherwise would be magnitude dependent. The difference between this value of τ and the ergodic estimate (i.e. the standard deviation of the mean of δB_e).

Table 5.13 shows the comparison between the empirical total standard deviation at the ground surface σ^G, with the sum of the disaggregated sources of variability (i.e. Equation 5.25). Note that for most spectral periods there is very good agreement between the two.

Table 5.12. Standard Deviation of Random Variables that compose the total Surface Standard Deviation (ergodic)

Period	ϕ^B	$\phi_{S2S^{AMP}}$	ϕ_{AMP}	τ
PGA	0.6603	0.3805	0.3067	0.5212
0.0384	0.6380	0.3681	0.3100	0.5092
0.0484	0.6190	0.3787	0.3153	0.4965
0.0582	0.6133	0.3854	0.3129	0.4872
0.0769	0.5939	0.4370	0.3012	0.4877
0.0844	0.6085	0.4527	0.2979	0.4821
0.0970	0.6120	0.4922	0.2910	0.4808
0.1167	0.6078	0.5019	0.2811	0.4849
0.1472	0.6010	0.4839	0.2592	0.4587
0.1691	0.6054	0.4926	0.2510	0.4593
0.2036	0.6067	0.4687	0.2425	0.4662
0.2340	0.5921	0.4730	0.2412	0.4616
0.3090	0.5773	0.4549	0.2301	0.4540
0.3551	0.5696	0.4555	0.2266	0.4579
0.3896	0.5517	0.4543	0.2240	0.4674
0.4274	0.5505	0.4483	0.2290	0.4666
0.4690	0.5418	0.4244	0.2359	0.4614
0.5913	0.5207	0.3876	0.2384	0.4717
0.7456	0.4964	0.3770	0.2607	0.4753
0.8180	0.4924	0.3610	0.2696	0.4781
0.9401	0.4935	0.3564	0.2828	0.4760
1.3622	0.5038	0.3577	0.3312	0.5062

The negative correlation between the borehole variability (δ^B) and the components of the variability in the amplification implies that if this correlation is taken into account in the Bazurro and Cornell methodology, the resulting standard deviation would be reduced. Equation 3.3 has to be modified to account for the correlation. Hence, the equation becomes

$$\sigma_{lnS_a^s(f)} \approx \sqrt{(c_1+1)^2 * \sigma^2_{lnS_a^r(f)} + \sigma^2_{lnAF(f)} + 2(c_1+1)\rho(f)\sigma_{lnS_a^r(f)}\sigma_{lnAF(f)}} \qquad (5.26)$$

Inclusion of the negative correlation would decrease standard deviations at the surface and hence result in a reduction in hazard curves.

Table 5.13. Empirical check for Equation 5.25

Period	Empirical Surface σ	Surface σ from eq. 5.25
PGA	0.8181	0.8567
0.0384	0.8179	0.8428
0.0484	0.8192	0.8256
0.0582	0.8248	0.8199
0.0769	0.8684	0.8675
0.0844	0.8722	0.8767
0.0970	0.8939	0.8944
0.1167	0.8948	0.8965
0.1472	0.8838	0.8837
0.1691	0.8758	0.8798
0.2036	0.8736	0.8782
0.2340	0.8678	0.8689
0.3090	0.8537	0.8459
0.3551	0.8455	0.8366
0.3896	0.8413	0.8310
0.4274	0.8311	0.8212
0.4690	0.8121	0.8006
0.5913	0.7963	0.7853
0.7456	0.7773	0.7677
0.8180	0.7690	0.7635
0.9401	0.7558	0.7508
1.3622	0.7472	0.7448

Figures 5.62 to 5.64 show $\delta S2S^{AMP}$ as a function of Vs_{30} and predominant period To. Note how the standard deviations at all periods reduce with increasing shear wave velocity. This would indicate lower variability in site response for stiffer soils.

5.3.4 Study of Single-Station Standard Deviations: path, magnitude, and distance effects

Single station standard deviation was shown to be a lower bound to the standard deviation that can be used in a PSHA analysis when the ergodic assumption on site response is removed. There is, however, large variability in single-station standard deviation from one station to another, as evidenced, for example, in Figure 5.50. One possibility for

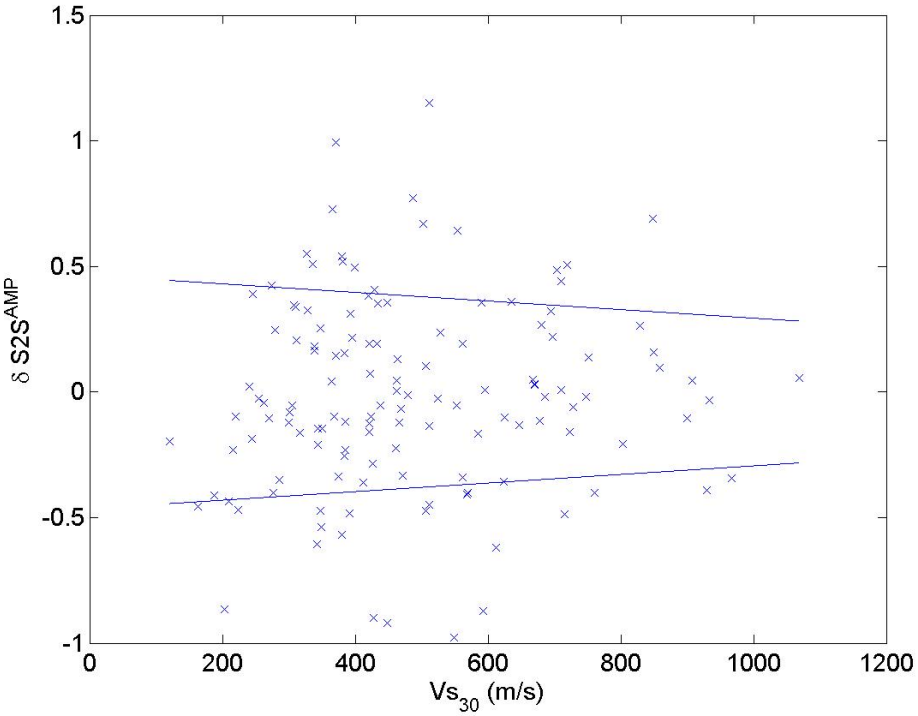


Figure 5.62. Mean residual amplification ($\delta S2S^{AMP}$) versus Vs_{30} for peak ground acceleration. Plus and minus 1 standard deviations are added to illustrate dependency

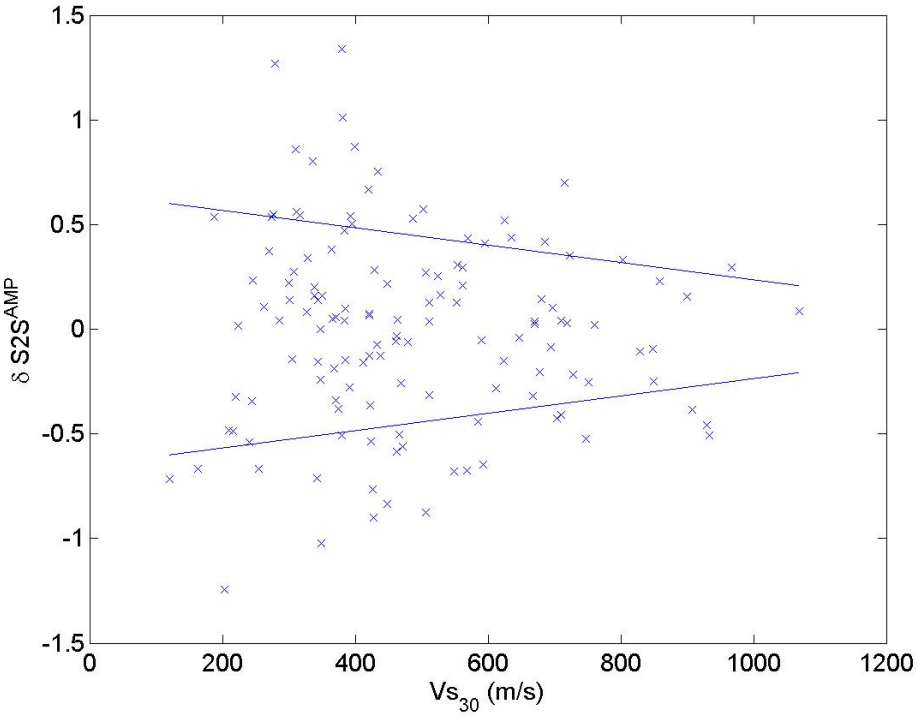


Figure 5.63. Mean residual amplification ($\delta S2S^{AMP}$) versus Vs_{30} for spectral period of 0.3 seconds. Plus and minus 1 standard deviations are added to illustrate dependency

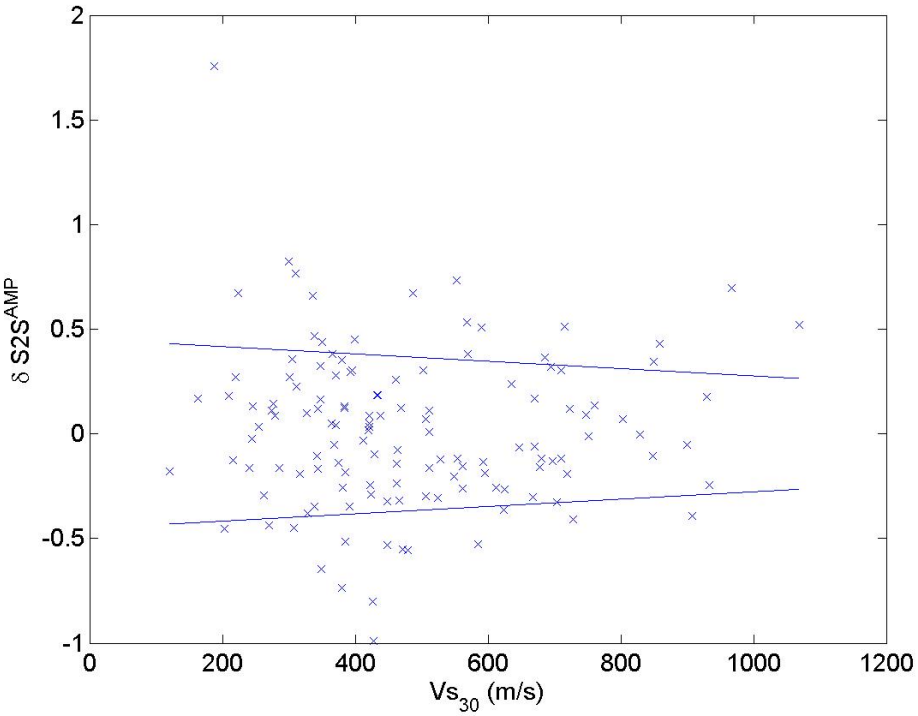


Figure 5.64. Mean residual amplification ($\delta S2S^{AMP}$) versus Vs_{30} for spectral period of 1.0 seconds. Plus and minus 1 standard deviations are added to illustrate dependency

this large variation in single station standard deviation is that a single station samples events from various sources and various azimuths. Figure 5.65 shows the single station standard deviation for $T = 0.3$ as a function of the standard deviation of the station-to-event azimuth. Observe a clear correlation of increasing standard deviation with increasing standard deviation azimuth. A similar trend can be seen in Figure 5.66, where single station standard deviations are shown only for records within a given bracket of station-to-event azimuth. When the azimuth bracket is reduced to 8 degrees, the single station standard deviation (intra-event) for $T = 0.3$ reduces to 0.424 from a value of 0.481 for all events. Table 5.14 shows a comparison of ϕ_{ss} and the bracketed ϕ_{ss}.

Table 5.14. Single-Station Standard Deviations for Surface and Bracketed Azimuths

Period	Surf ϕ	Surf ϕ for 8° bracket
PGA	0.4967	0.5216
0.0384	0.4857	0.4220
0.0484	0.4725	0.4038
0.0582	0.4660	0.4179
0.0769	0.4774	0.4166
0.0844	0.4837	0.4192
0.0970	0.4912	0.4305
0.1167	0.5028	0.4346
0.1472	0.5149	0.4774
0.1691	0.5124	0.4909
0.2036	0.5067	0.4834
0.2340	0.5010	0.4779
0.3090	0.4812	0.4239
0.3551	0.4777	0.4285
0.3896	0.4653	0.4150
0.4274	0.4583	0.3989
0.4690	0.4531	0.3963
0.5913	0.4442	0.3915
0.7456	0.4258	0.3874
0.8180	0.4264	0.3882
0.9401	0.4258	0.4073
1.3622	0.4178	0.3897

This study concentrated on removing the ergodicity assumption on site response from site

response analyses. Figures 5.65 and 5.66 are an indication that if the ergodicity assumption on path is also removed, the single station standard deviation can be reduced much further. Studies in Taiwan (Lin et al., 2010; Chen and Tsai, 2002), Japan (Morikawa et al., 2008), and Southern California (Atkinson, 2006) showed single-station, single-path standard deviations for PGA as low as 0.365, 0.36, and 0.414, respectively. In this study, the value for a spectral period 0.03 seconds computed for stations with an azimuth range of 8 degrees is 0.422, for PGA the bracketed value is the only exception where higher than that obtained for all azimuths were obtained. Similar reduction is seen at other spectral periods. Table 5.15 shows a comparison with other studies, note the remarkable agreement in σ_{ss} from studies across very different regions.

Table 5.15. Single-station standard deviations for PGA

	This study	Atkinson (2006)	Lin et al. (2010)	Chen & Tsai (2002)
σ	0.799	0.711	0.680	0.731
τ	0.493		0.405	0.490
ϕ	0.629		0.546	0.543
σ_{ss}	0.672	0.617	0.619	0.631
ϕ_{sp}	0.387	0.414	0.365	

It is also important to consider the variability in single station standard deviation with magnitude and distance to rupture. Figure 5.67 shows the single station standard deviations for two ranges of magnitudes. The differences in standard deviation for the different magnitude ranges are significant, although it is important to note that the standard deviations obtained are sample size sensitive. This behavior is consistent for all periods, and as more data is gathered further validation would be desirable. When different ranges of distance are used (Figure 5.68), it is seen that the standard deviation decreases with increasing distance to the fault. This trend is also seen for intra-event residuals (e.g. Figure 5.22).

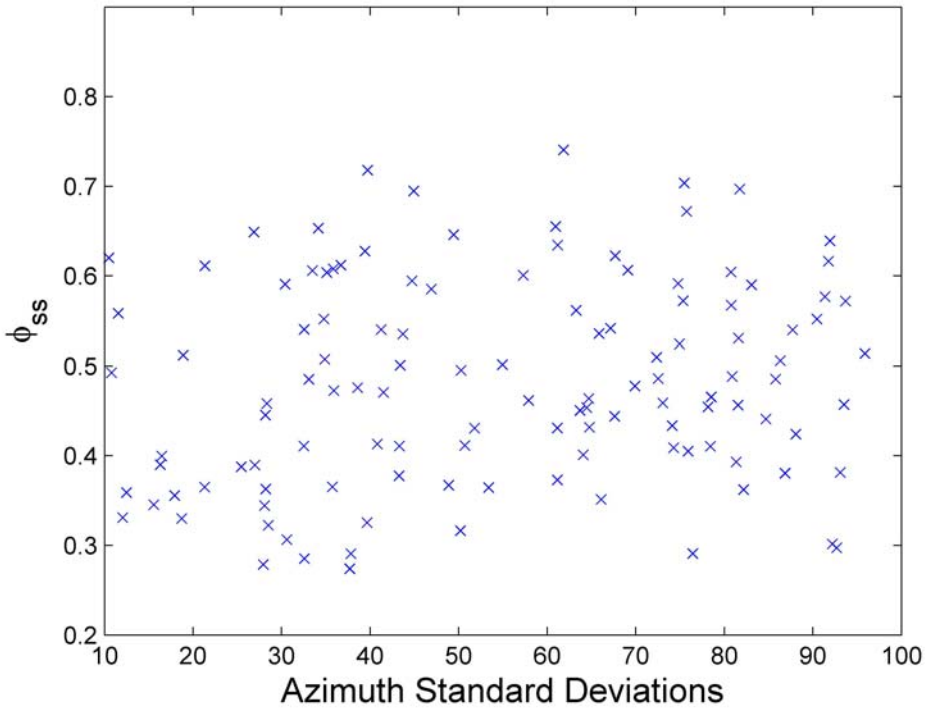


Figure 5.65. Single station standard deviations for $T = 0.3$ versus standard deviation of station-to-event azimuths. These two quantities are positively correlated ($\rho = 34\%$). Only stations with more than 15 records are considered for this plot.

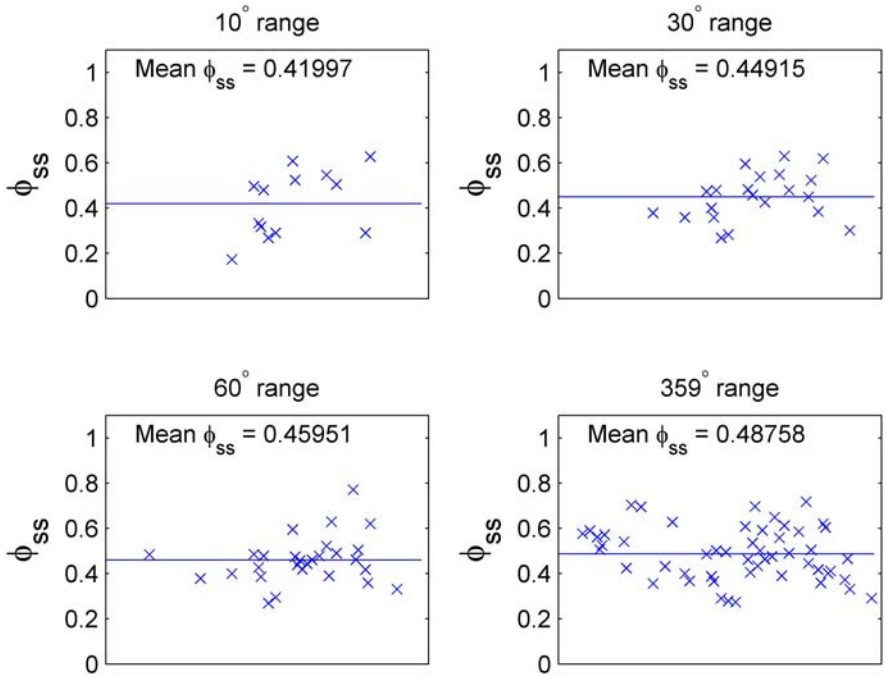


Figure 5.66. Single station standard deviations for $T = 0.3$ of records sampled from a varying range of station-to-event azimuth. Only station with more than 15 records are considered for this plot.

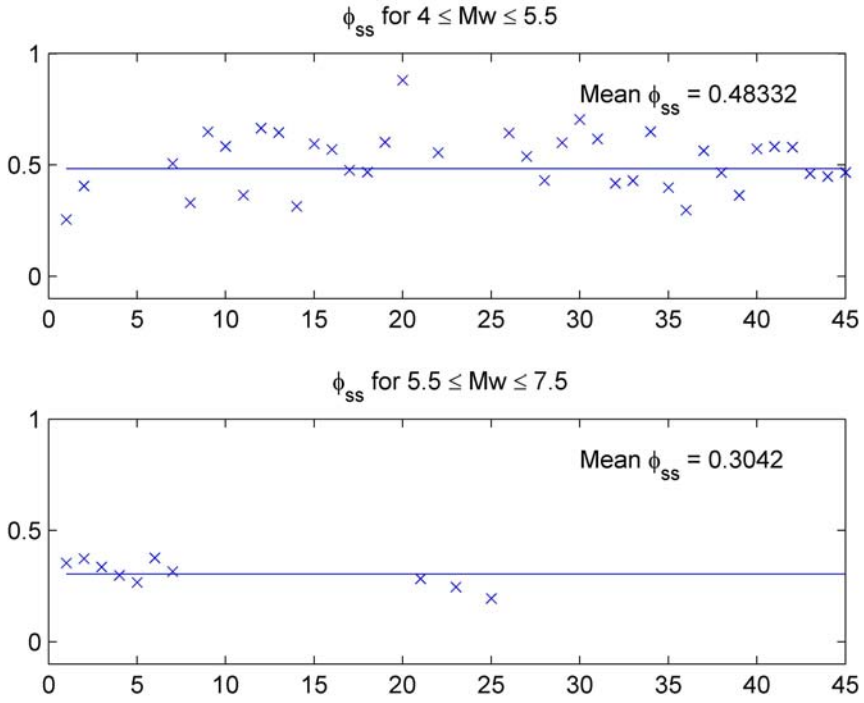


Figure 5.67. Single station standard deviations for different magnitude ranges for $T = 0.3$. Only station with more than 10 records are considered for this plot, and only those with 6 or more records are within each magnitude bracket are plotted.

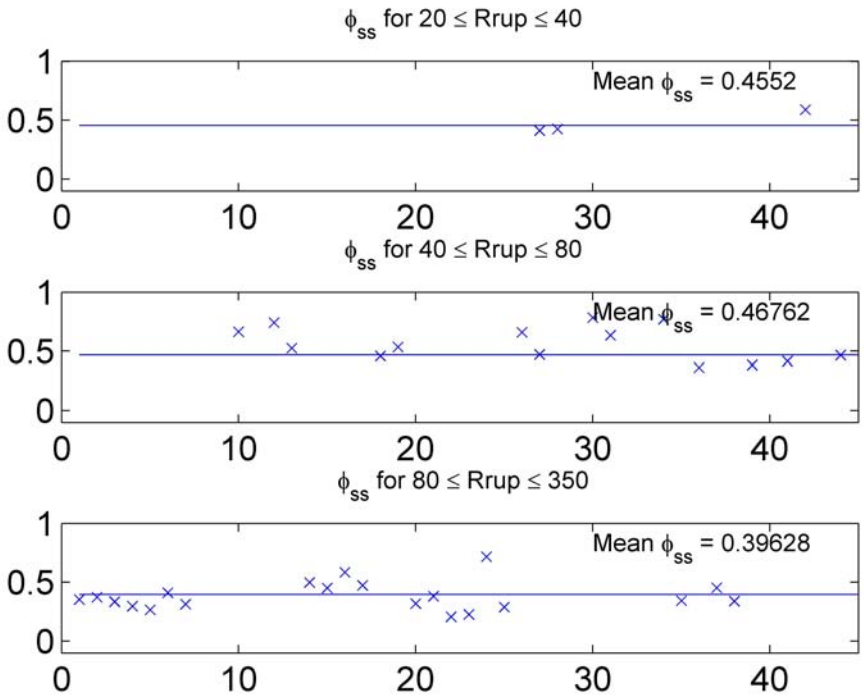


Figure 5.68. Single station standard deviations for $T = 0.3$ and for different ranges of distance to the rupture plane. Only station with more than 10 records are considered for this plot. Mean single station standard deviation of these records is 0.48.

5.4 PSHA Example

This section presents an example application of the proposed methodology along with current *ergodic* estimates, and estimates using single-site ground motion records. All within the framework of a probabilistic seismic hazard analysis (PSHA).

The seismicity affecting the sample site will be defined as follows,

- Areal Source of seismicity of 100 by 100 kilometers.
- Activity rate of 0.38, for Mw between 4 and 7
- Truncated exponential Mw distribution ($\lambda = 0.8$).

The selected site is is located at coordinates 34.4856^o N and 130.9069 W and, corresponds to station KMMH09 in the KiK-net database. The Vs_{30} of the site is 399.79 (m/s), and the shear-wave velocity profile is given in Figure 5.69.

In this example we will look at the ground motion predictions and hazard curve for a spectral period of 0.1 seconds. Figure 5.70 shows the residuals of the records at the site, with a mean deviation from the predicted $Sa(T = 0.1)$ of $\delta S2S$= 0.28, this corresponds to the site-to-site variability. The standard deviation of these residuals is σ_{ss_s}=0.36.

Figure 5.71 shows the hazard curve for the ergodic condition, that is using the median GMPE prediction for the ground surface, and the corresponding standard deviation. Figure 5.72 shows the hazard curve for the site-specific case, that is when $\delta S2S$ is known and σ_{ss_s} can be used.

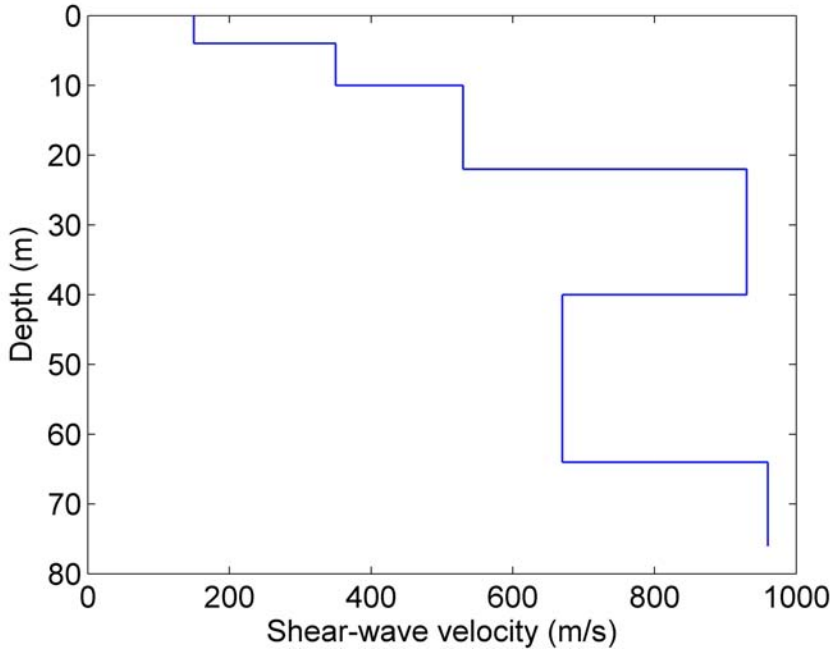


Figure 5.69. Shear-wave velocity profile for site KMMH09

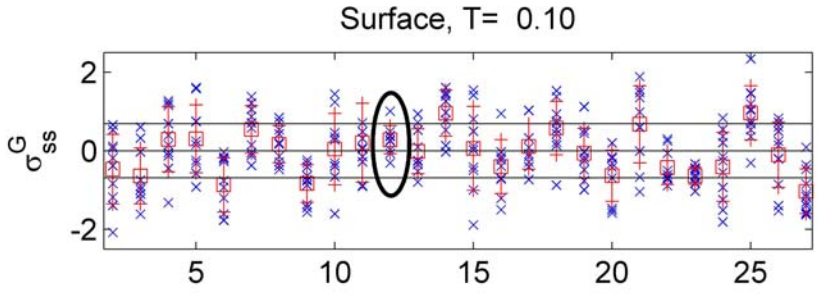


Figure 5.70. Total residuals at site KMMH09, for spectral period of 0.1 seconds.

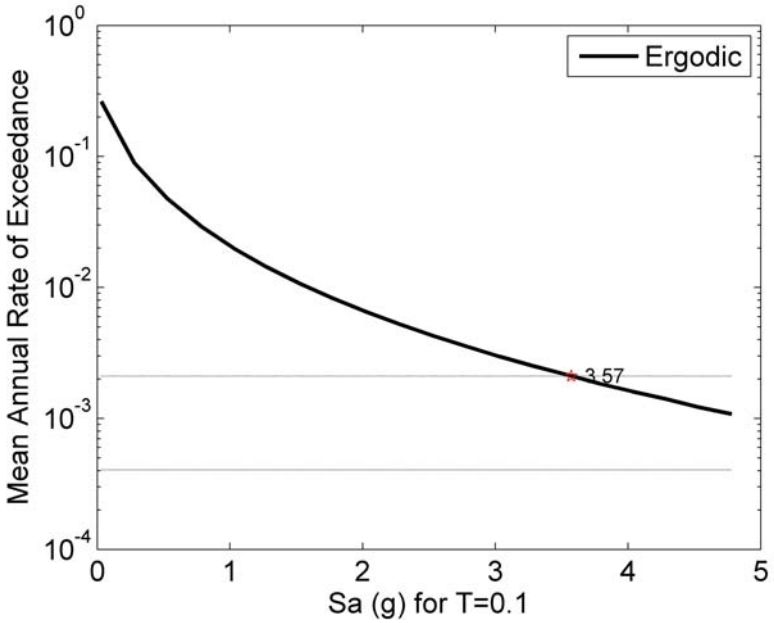


Figure 5.71. Ergodic Hazard Curve

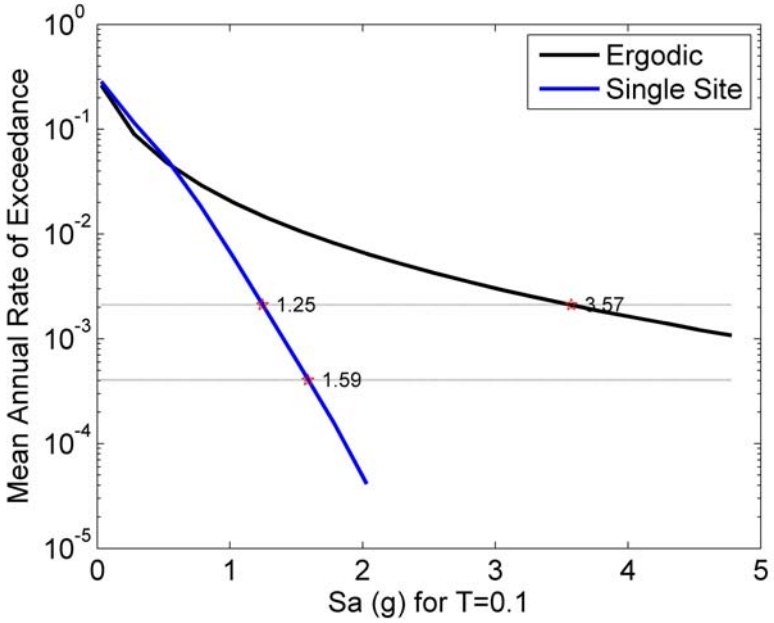


Figure 5.72. Ergodic and Single-Site Hazard Curves

Note the that the reduction in design values for long return periods is significant. For short return periods, the single-site hazard curve is increases because of the positive value of

$\delta S2S$, this shows the bias of the ergodic prediction. Figure 5.73 shows the prediction of the single-site hazard curve using the median borehole prediction, site response, and the standard deviation known for the site (σ_{ss_s}). This last curve matches reasonably well the single-site curve, but it is important to know that in a general case we do not have the necessary information to use σ_{ss_s}.

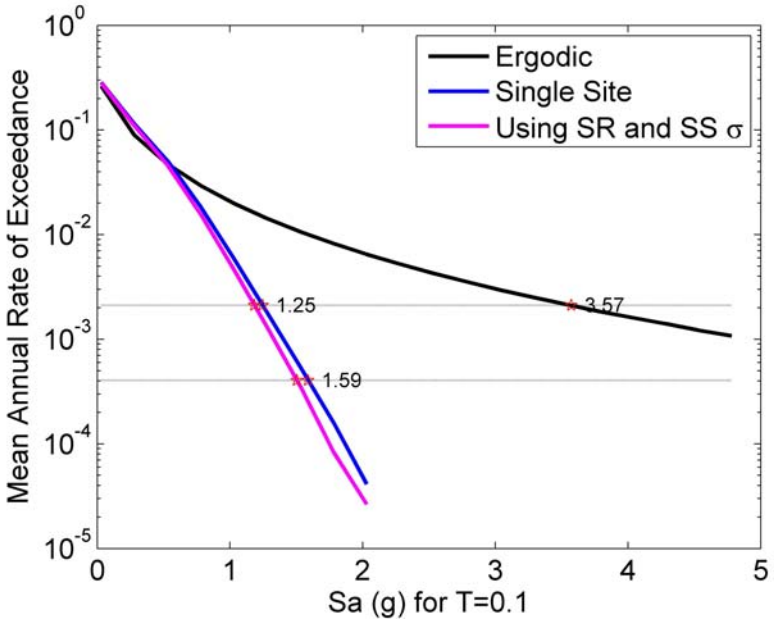


Figure 5.73. Site-specific hazard curve using borehole ground motion estimate and site response.

Using the second method proposed in Chapter 3, a fourth hazard curve is calculated with the median estimate for borehole and the ergodic standard deviation but without the site-to-site amplification variability (i.e. $\phi_{S2S^{AMP}}$). This hazard curve is included in Figure 5.74, note the reduction on design values for long return periods is significant.

Figure 5.75 shows another important advantage of the proposed approach, the reduction of bias in ground motion estimates. Because the selected site is positively bias with respect to the median prediction (i.e. $\delta S2S > 0$), the predicted ground motion using site response is higher than the median prediction, which impacts the seismic hazard and can be observed in Figure 5.75.

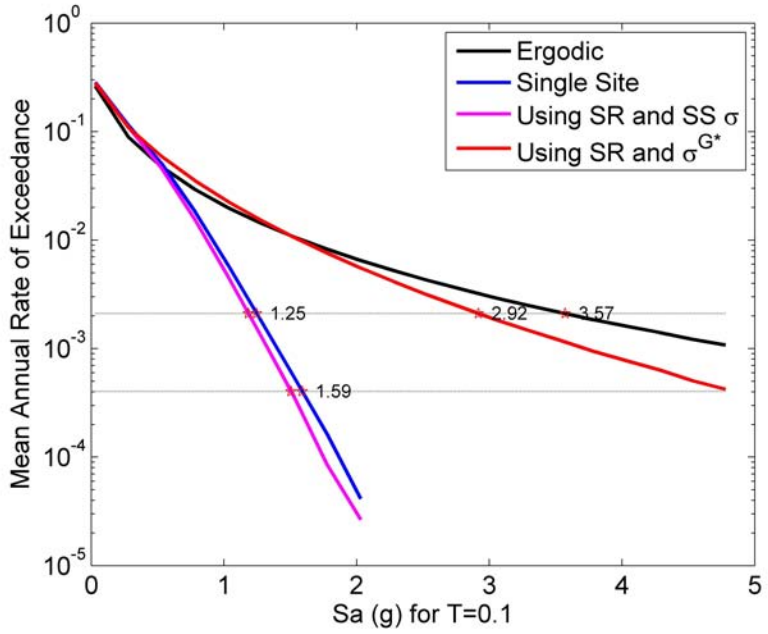


Figure 5.74. Hazard curve using site response and partially ergodic standard deviation

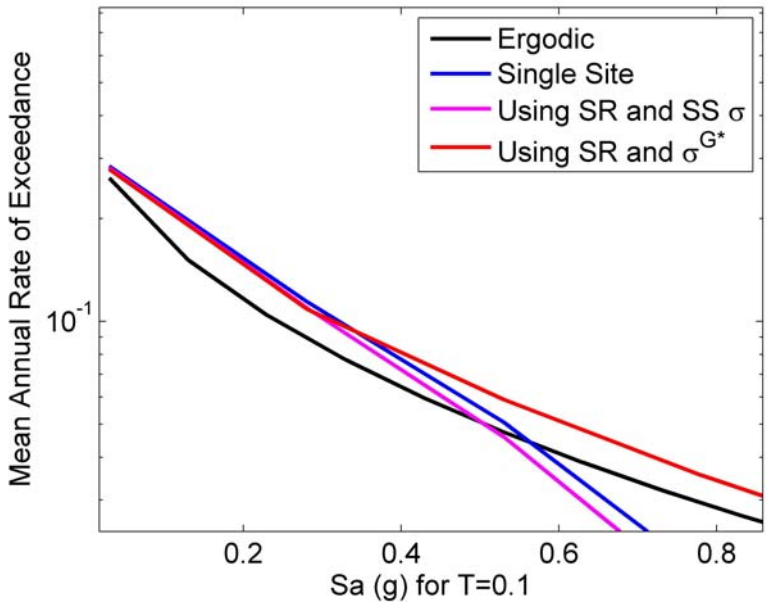


Figure 5.75. Short return period portion of the hazard curves. Note the misfit of the site response generated curves with respect to the ergodic curve, this reflects the reduction of the bias in the ground motion predictions.

All the hazard curves shown so far consider linear site response. The GMPE's consider only linear site terms because the KiK-net database is governed by low magnitude events that in most cases did not produce nonlinear soil behavior. However, in our PSHA example long return periods are considered, which imply higher intensities and therefore would induce nonlinear soil behavior. To consider nonlinearity in site response, the Bazzurro and Cornell approach is used.

Site response analyses were performed using the linear-equivalent approach. Modulus degradation curve used for all layers was the Dobry and Vucetic (1987) curve for clays with plasticity index between 10 and 20, and the density was taken as 1.8 g/cm^3. The ground motions used as input correspond to those used in Section 4.2.

Figure 5.76 shows the results of this exercise. Observe the significant reduction in the design values for long return periods.

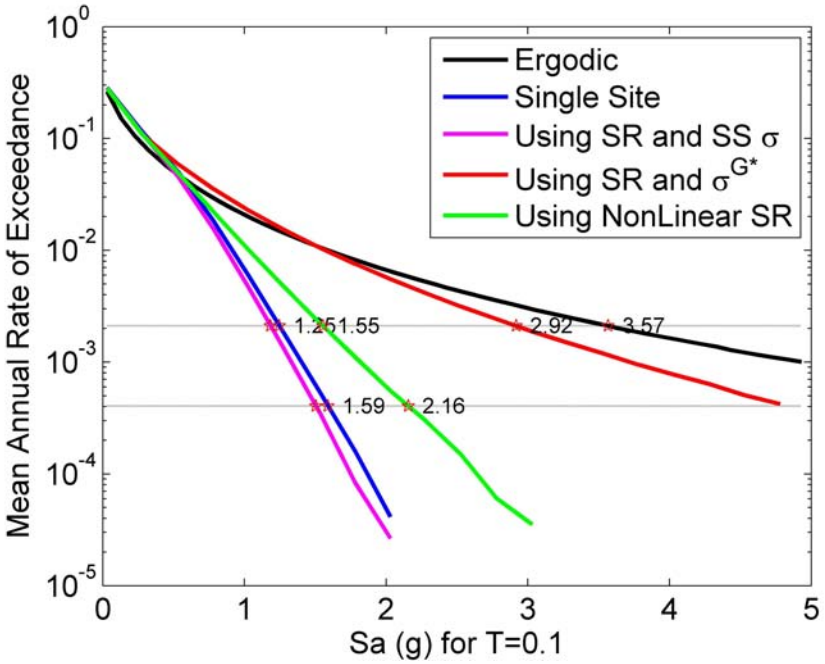


Figure 5.76. Hazard curves including nonlinear site response.

Chapter 6

Conclusions

6.1 Summary

The results of seismic hazard analyses are strongly influenced by the uncertainty in the input parameters. The motivation for this dissertation is to better account for the components of these uncertainties. One important step towards reducing uncertainty is to identify the parts of the phenomena that are epistemic, and therefore reducible, and the parts that constitute aleatoric uncertainty.

Throughout this dissertation the total uncertainty in the predicted ground motion was divided into different components. The components were explained and where possible bounds to these uncertainties were set. This is significant in the context of Probabilistic Seismic Hazard Analysis (PSHA), as if a part of the uncertainty is bounded it means that when designs for long return periods (i.e. important structures) are needed the magnitudes of the parameters involved (e.g. magnitude, site amplification) cannot increase indefinitely. Although there is no evidence in the literature to cap the total uncertainty in ground motion predictions, the ability to limit a part of it could be significant.

In order to avoid counting the site variability twice in a site-specific hazard analysis, one needs to use an input motion that does not include that uncertainty. This is not easily done with current methodologies, as most GMPE are developed using records from multiple earthquakes and multiple locations, hence the ground motion estimates includes site-to-site variability. This dissertation approaches this problem by studying the site response part of the total uncertainty, this is done using surface and borehole data available from the KiK-net database.

Three different problems were studied . First, a random field model for shear wave velocity profiles was developed. This model can be used to obtain realizations of random fields for Monte Carlo type analyses. Second, a ground motion prediction equation for the Kik-net database was developed. This database was selected because it has sufficient records to compute stable estimates for the statistics of ground motion residuals at a single station. Finally, the statistics of residuals at a single station were studied and their components were analyzed.

6.2 Significant Findings

This dissertation provides means to perform site-specific seismic hazard analysis, relaxing in part the ergodic assumption, and hence reducing uncertainty in the ground motion estimates.

Site response can be a significant contributor to the seismic hazard. Moreover, this is a component of ground motion uncertainty that could be reduced. For this reason random fields were studied as an option to manage the variability of shear-wave velocity profiles that in turn can result in site response variability. Random fields can be used to generate profiles that match the observed profiles to the degree of uncertainty reached with the site exploration, and then used to propagate bedrock ground motions. By using random fields

the engineer is allowed to use all available information on the site of interest to constrain the wave propagation characteristics of the site. Constraining the variability of the shear-wave velocity profiles, produces three important benefits,

- Better reflecting the variability in all inputs to a PSHA analysis.
- Reducing bias in the ground motion estimate.
- Possible reduction of the total uncertainty.

One significant contribution of this dissertation to the use random fields in generating vertical shear-wave velocity profiles is the use of stationary gaussian models. Typical random profile generation approaches use non-stationary gaussian models, that require the fitting of many of their parameters before the user is able to generate profiles appropriate for his/her study. The proposed stationary gaussian model uses the strong correlation that exists between adjacent layers rather than distance between points. This allows for very simple and fast calibration of the model to any data set, in particular if at a site of interest site exploration is performed the random field models can be used to account for the varying degrees of knowledge on the site profile.

Several random field models were developed and compared in terms of reproducing the statistics of the underlying database, its correlation structure, and their ability to reproduce site response. The two models that compared the best with the database (i.e. the real profiles), were the one layer lag stationary gaussian model and the non-stationary gaussian model proposed by Toro (1995). The conclusion is that the one layer lag stationary model is the simplest of all the tested models and the one with best behavior by the measures that were tested.

The typical assumption in PSHA is that the uncertainty of the entire sample of sites and earthquakes is equal to the uncertainty of an individual site. This so called ergodic

assumption leads to high uncertainties because many different seismotectonic regions, wave paths, and sites are lumped together. In this dissertation single-site data were examined. A noteworthy observation is that the reduction of uncertainty is significant when moving from using a general GMPE (i.e. one that uses the ergodic assumption), to that observed uncertainty at single sites that have recorded more that 10 earthquakes.

In order to study the uncertainty at these sites, three ground motion prediction equation were developed. One for the ground surface, one for the borehole, and one that used the combined data set, for surface and borehole together. This last approach is novel in that it uses borehole and surface data to constrain the terms that are not surface or borehole specific. Conclusions from the analysis of the residuals from these GMPEs include that the Combined model produces similar results to those of the models for surface or borehole. The mean residuals for each earthquake are remarkably similar between the three equations, which validates the use of a single measure for borehole and ground surface event residual.

Single station residuals are a sample of what would be the variability at a specific site (i.e. without using ergodic assumption). This variability can be reduced even more if a limited range of earthquake azimuths is considered for each station, forcing a unique path and possibly source. Path and source effects are a significant contribution to the total uncertainty, allowing to reduce the intra-event standard deviation by about 0.2 (log units of pseudo-spectral acceleration). A striking observation is that while total uncertainty in different regions of the world are significantly different, estimates of single-station standard deviation and single-station, single-path standard deviation are very similar.

Finally, a methodology for site-specific PSHA studies for sites where site response is needed was proposed. The methodology proposed by Bazzurro and Cornell (2004b) was used. However, the methodology was modified in two important ways. First, it was demonstrated that the standard deviation of the input ground motion should be that of the borehole records, which is significantly lower than the standard deviation at the surface. Second, it

was demonstrated that borehole residuals are negatively correlated with the amplification factor and this correlation can result in a reduction of the uncertainty at the surface.

6.3 Recommendations for Further Study

Several issues worth of additional study were identified throughout the conduct of the research presented in this dissertation. The most important issues are highlighted below

- The ground motion prediction equations developed in this study assume linear soil behavior. However, several indicators point to non-linear soil behavior for some stations. It is recommended that the GMPE developed herein be modified to account for soil nonlinearity. The effect of nonlinearity on the distribution of the components of single-station residuals and their interdependence must also be studied.

- Study of the KiK-net data clearly indicated that consideration of single-path would reduce the single-station standard deviation. More systematic methods should be used to compute single-path standard deviation. Moreover, methodologies to compute path terms that would be used in PSHA must be developed.

- The proposed methodology takes advantage of the existence of borehole ground motion data at the KiK-net array. In a typical PSHA study, these data is unavailable. This implies that the methodology should be generalized to cases where site response is computed from surface records. This implies taking into consideration the strong, but not perfect, correlation between rock outcrop and borehole ground motions.

- Residuals with respect to hypocenter depth were observed in the GMPE. This implies that the GMPE can be improved by introducing this parameter in the equation. Moreover, strong dependence of standard deviation with distance to the fault was observed. Standard deviation can be parameterized to capture this dependency, with

possible reduction in the overall uncertainty.

Bibliography

Abrahamson, N., Atkinson, G., Boore, D., Bozorgnia, Y., Campbell, K., Chiou, B., Idriss, I. M., Silva, W., Youngs, R., 2008. Comparisons of the nga ground-motion relations. Earthquake Spectra 24 (1), 45–66.

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

Al-Atik, L., Abrahamson, N., Cotton, F., Scherbaum, F., Bommer, J., Kuehn, N., 2010. The variability of ground-motion prediction models and its components. To be submitted to Seismological Research Letters.

Anderson, J. G., 2010. personal communication.

Anderson, J. G., Brune, J., 1999. Probabilistic seismic hazard analysis without the ergodic assumption. Seism. Res. Lett. 70, 19–28.

Andrade, J. E., Borja, R. I., 2006. Quantifying sensitivity of local site response models to statistical variations in soil properties. Acta Geotechnica 1 (1), 3–14.

Asten, M. W., Boore, D. M., 2005. Comparison of shear-velocity profiles of unconsolidated

sediments near the coyote borehole (ccoc) measured with fourteen invasive and non-invasive methods. U.S. Geol. Surv. Open-File Rept. 2005-1169.

Atkinson, G. M., 2006. Single-Station Sigma. BULLETIN OF THE SEISMOLOGICAL SOCIETY OF AMERICA 96 (2), 446–455.

Baker, J. W., Cornell, C. A., 2006. Spectral shape, epsilon and record selection. Earthquake Engineering & Structural Dynamics 35 (9), 1077–1095.

Baturay, M. B., Stewart, J. P., 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, C. A., 2004a. Ground-Motion Amplification in Nonlinear Soil Sites with Uncertain Properties. Bulletin of the Seismological Society of America 94 (6), 2090–2109.

Bazzurro, P., Cornell, C. A., 2004b. Nonlinear Soil-Site Effects in Probabilistic Seismic-Hazard Analysis. Bulletin of the Seismological Society of America 94 (6), 2110–2123.

Boore, D. M., 1983. Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra. Bulletin of the Seismological Society of America 73 (6), 1865–1894.

Boore, D. M., 2004. Can site response be predicted? Journal of Earthquake Engineering 8 (1), 1–41.

Boore, D. M., Atkinson, G. M., 2008. Ground-motion prediction equations for the average horizontal component of pga, pgv, and 5%-damped psa at spectral periods between 0.01 s and 10.0 s. Earthquake Spectra 24 (1), 99–138.

Boore, D. M., Thompson, E. M., 2007. On using surface-source downhole-receiver logging to determine seismic slownesses. Soil Dynamics and Earthquake Engineering 27 (11), 971 – 985.

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

Borcherdt, R. D., 2002. Empirical Evidence for Acceleration-Dependent Amplification Factors. Bulletin of the Seismological Society of America 92 (2), 761–782.

Brown, L. T., Boore, D. M., Stokoe, Kenneth H., I., 2002. Comparison of shear-wave slowness profiles at 10 strong-motion sites from noninvasive sasw measurements and measurements made in boreholes. Bulletin of the Seismological Society of America 92 (8), 3116–3133.

Castellaro, S., Mulargia, F., Rossi, P. L., 2008. Vs30: Proxy for Seismic Amplification? Seismological Research Letters 79 (4), 540–543.

CEN, 2003. Eurocode 8: Design of structures for earthquake resistance, part 1: general rules. Seismic actions and rules for buildings, stage 49 draft, Comite´ Europe´en de Normalisation. Brussels.

Chen, Y.-H., Tsai, C.-C. P., 2002. A stable algorithm for regression analyses using the random effects model. Bulletin of the Seismological Society of America 92 (1), 1984–1991.

Conover, W., 1980. Practical Nonparametric Statistics. Wiley, New York.

Cotton, F., Pousse, G., Bonilla, F., Scherbaum, F., 2008. On the Discrepancy of Recent European Ground-Motion Observations and Predictions from Empirical Models: Analysis of KiK-net Accelerometric Data and Point-Sources Stochastic Simulations. Bulletin of the Seismological Society of America 98 (5), 2244–2261.

Cramer, C. H., 2003. Site-Specific Seismic-Hazard Analysis that is Completely Probabilistic. Bulletin of the Seismological Society of America 93 (4), 1841–1846.

Darendeli, M., 2001. Development of a new family of normalized modulus reduction and material damping curves, Ph.D. thesis. University of Texas, Austin, Texas.

Dobry, R., Borcherdt, R. D., Crouse, C. B., Idriss, I. M., Joyner, W. B., Martin, G. R., Power, M. S., Rinne, E. E., Seed, R. B., 2000. New site coefficients and site classification system used in recent building seismic code provisions. Earthquake Spectra 16 (1), 41–67.

Dobry, R., Oweis, I., Urzua, A., 1976. Simplified procedures for estimating the fundamental period of a soil profile. Bulletin of the Seismological Society of America 66 (4), 1293–1321.

Dobry, R., Vucetic, M. (Eds.), 1987. Dynamic properties and seismic response of clay deposits. Proceedings of the International Symposium on Geotechnical Engineering of Soft Soils, Mexico City, Mexico.

Douglas, J., Gehl, P., Bonilla, L. F., Scotti, O., Regnier, J., Duval, A.-M., Bertrand, E., 2009. Making the most of available site information for empirical ground-motion prediction. Bulletin of the Seismological Society of America 99 (3), 1502–1520.

EPRI, E. P. R. I., 1993. Guidelines for Determining Design Basis Ground Motions. Report, TR-102293, Palo Alto: California.

Fenton, G., Griffiths, D., 2005. Three-dimensional probabilistic foundation settlement. Journal of Geotechnical and Geoenvironmental Engineering 131 (2), 232–239.

Fenton, G., Griffiths, D., Williams, M., 2005. Reliability of traditional retaining wall design. Géotechnique 55 (1), 55 – 62.

Fenton, G. A., 1994. Error evaluation of three random-field generators. Journal of Engineering Mechanics 120 (12), 2478–2497.

Finn, W. L., 1991. Geotechnical engineering aspects of microzonation. Proceedings of the 4^{th} International Conference on Seismic Zonation Vol. 1, 199–259.

Fujiwara, H., Aoi, S., Kunugi, T., Adachi, S., 2004. Strong-motion observation networks of nied : K-net and kik-net. abstract.
URL www.cosmos-eq.org/Projects/Fujiwara_paper.pdf

Fukushima, Y., 1996. Scaling relations for strong ground-motion prediction models with m2 terms. Bull. Seismol. Soc. Am. 86, 329–336.

Goldstein, P., Dodge, D., Firpo, M., , Minner, L., 2003. SAC2000: signal processing and analysis tools for seismologists and engineers, in The IASPEI International Handbook of Earthquake and Engineering Seismology, W. H. K. Lee, H. Kanamori, P. C. Jennings and C. Kisslinger (Editors). Academic Press, London.

Goulet, C. A., Stewart, J. P., 2009. Pitfalls of deterministic application of nonlinear site factors in probabilistic assessment of ground motions. Earthquake Spectra 25 (3), 541–555.

IBC, I. C. C., 2006. 2006 International Building Code. Washington: DC.

Idriss, I., Seed, H., 1970. Soil moduli and damping factors for dynamic response analysis. Report, EERC 70-10. Berkeley: University of California.

Kiban-Kyoshin, N., 2010. Database. available from http://www.kik.bosai.go.jp/, last accessed on May 2010.

Kim, H., Santamarina, J., 2008. Spatial variability: drained and undrained deviatoric load response. Géotechnique 58 (10), 805 – 814.

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.

Kottke, A., Rathje, E. M., 2008. A semi-automated procedure for selecting and scaling recorded earthquake motions for dynamic analysis. Earthquake Spectra 24 (4), 911–932.

Kramer, S. L., 1996. Geotechnical Earthquake Engineering. Pearson Education Inc.

Kramer, S. L., Stewart, J. P., 2004. Earthquake Engineering, from Engineering Seismology to Performance Based Engineering. CRC Press, Ch. 4.

Lin, P., Abrahamson, N., Walling, M., Lee, C.-T., Chiou, B., Cheng, C., 2010. Repeatable path effects on the standard deviation for empirical ground motion models. To be submitted.

Manolis, G. D., 2002. Stochastic soil dynamics. Soil Dynamics and Earthquake Engineering 22 (1), 3 – 15.

Marosi, K. T., Hiltunen, D. R., 2004. Characterization of spectral analysis of surface waves shear wave velocity measurement uncertainty. Journal of Geotechnical and Geoenvironmental Engineering 130 (10), 1034–1041.

Martin, A. J., Diehl, J. G., 2004. Practical experience using a simplified procedure to measure average shear- wave velocity to a depth of 30 meters (vs30). 13 World Conference on Earthquake Engineering, Vancouver, B.C., Canada.

Morikawa, N., Kanno, T., Narita, A., Fujiwara, H., Okumura, T., Fukushima, Y., Guerpinar, A., 2008. Strong motion uncertainty determined from observed records by dense network in Japan. Journal of Seismology 12 (4), 529–546.

Moss, R. E. S., 2008. Quantifying Measurement Uncertainty of Thirty-Meter Shear-Wave Velocity. Bulletin of the Seismological Society of America 98 (3), 1399–1411.

Park, D., Hashash, Y. M., 2004. Soil damping formulation in nonlinear time domain site response analysis. Journal of Earthquake Engineering 8 (2), 249 – 274.

Pawlowsky-Glahn, V., Olea, R. A., 2004. Geostatistical analysis of compositional data. Studies in mathematical geology. Oxford University Press, New York.

PEER, 2010. Next generation attenuation ground motion database. http://peer.berkeley.edu/nga/index.html (last accessed on May 2010).

Phoon, K., Kulhawy, F., 1999. Characterization of geotechnical variability. Can. Geotech. J. 36 (4), 612 – 624.

Pousse, G., Berge-Thierry, C., Bonilla, F., Bard, P. Y., 2005. Eurocode 8 design response spectra evaluation using the k-net japanese database. Journal of Earthquake Engineering 9 (4), 547–574.

Rahman, M. S., Yeh, C. H., 1999. Variability of seismic response of soils using stochastic finite element method. Soil Dynamics and Earthquake Engineering 18 (3), 229 – 245.

Rathje, E. M., Kottke, A. R., Trent, W. L., 2010. Influence of input motion and site property variabilities on seismic site response analysis. Journal of Geotechnical and Geoenvironmental Engineering 136 (4), 607–619.

Rathje, E. M., Ozbey, M. C., 2006. Site-specific validation of random vibration theory-based seismic site response analysis. Journal of Geotechnical and Geoenvironmental Engineering 132 (7), 911–922.

Renault, P., Heuberger, S., Abrahamson, N. A., 2010. Pegasos refinement project: An improved psha for swiss nuclear power plants. Proceedings of the 14^{th} ECEE European Conference on Earthquake Engineering.

Reyes, D., Rodriguez-Marek, A., Lizcano, A., 2009a. A hypoplastic model for site response analysis. Soil Dynamics and Earthquake Engineering 29 (1), 173 – 184.

Reyes, D. K., Montalva, G. A., Rodriguez-Marek, A., Lizcano, A., 2009b. Propagation of uncertainty in non-linear site amplification. Submitted to Engineering Geology.

Rodriguez-Marek, A., Bay, J. A., Park, K., Montalva, G. A., Cortez-Flores, A., Wartman, J., Boroschek, R., 2010. Engineering analysis of ground motion records from the 2001 m[sub w] 8.4 southern peru earthquake. Earthquake Spectra 26 (2), 499–524.

Rodriguez-Marek, A., Bray, J. D., Abrahamson, N. A., 2001. An empirical geotechnical seismic site response procedure. Earthquake Spectra 17 (1), 65–87.

Rodriguez-Marek, A., Montalva, G. A., 2010. Uniform hazard spectra for site-specific applications including uncertainties in site-response. Final Report for Award 08HQGR0086. USGS.

Schnabel, P., Lysmer, J., Seed, H., 1972. Shake - a computer program for equation response analysis of horizontally layered sites. Report, EERC, 72–12.

Searle, S. R., 1971. Linear Models. Wiley, New York.

Stewart, J. P., Liu, A. H., Choi, Y., 2003. Amplification factors for spectral acceleration in tectonically active regions. Bulletin of the Seismological Society of America 93 (1), 332–352.

Thompson, E. M., Baise, L. G., Kayen, R. E., Guzina, B. B., 2009. Impediments to Predicting Site Response: Seismic Property Estimation and Modeling Simplifications. Bulletin of the Seismological Society of America 99 (5), 2927–2949.

Toro, G., 1995. Probablistic models of site velocity profiles for generic and site-specific ground-motion amplification studies. Department of Nuclear Energy Brookhaven National Laboratory, Upton, New York.

Toro, G., 2006. Site-wide Probabilistic model of shear-wave velocity profiles at the Savannah River Site. Aiken, South Carolina. Bechtel Savannah River, Inc., Aiken, South Carolina.

Wills, C. J., Silva, W., 1998. Shear-wave velocity characteristics of geologic units in california. Earthquake Spectra 14 (3), 533–556.

Appendix A

Additional Plots of Ground Motion Residuals

This appendix presents supplementary figures for the intra- and inter-event residuals. Residuals are plotted against time-averaged upper 30 meter shear-wave velocity (Vs_{30}), predominant period (To), depth to shear-wave velocity of 800 (m/s), and Magnitude considering only events that where recorded within 20 kilometers the fault.

A.1 Intra-Event Residuals

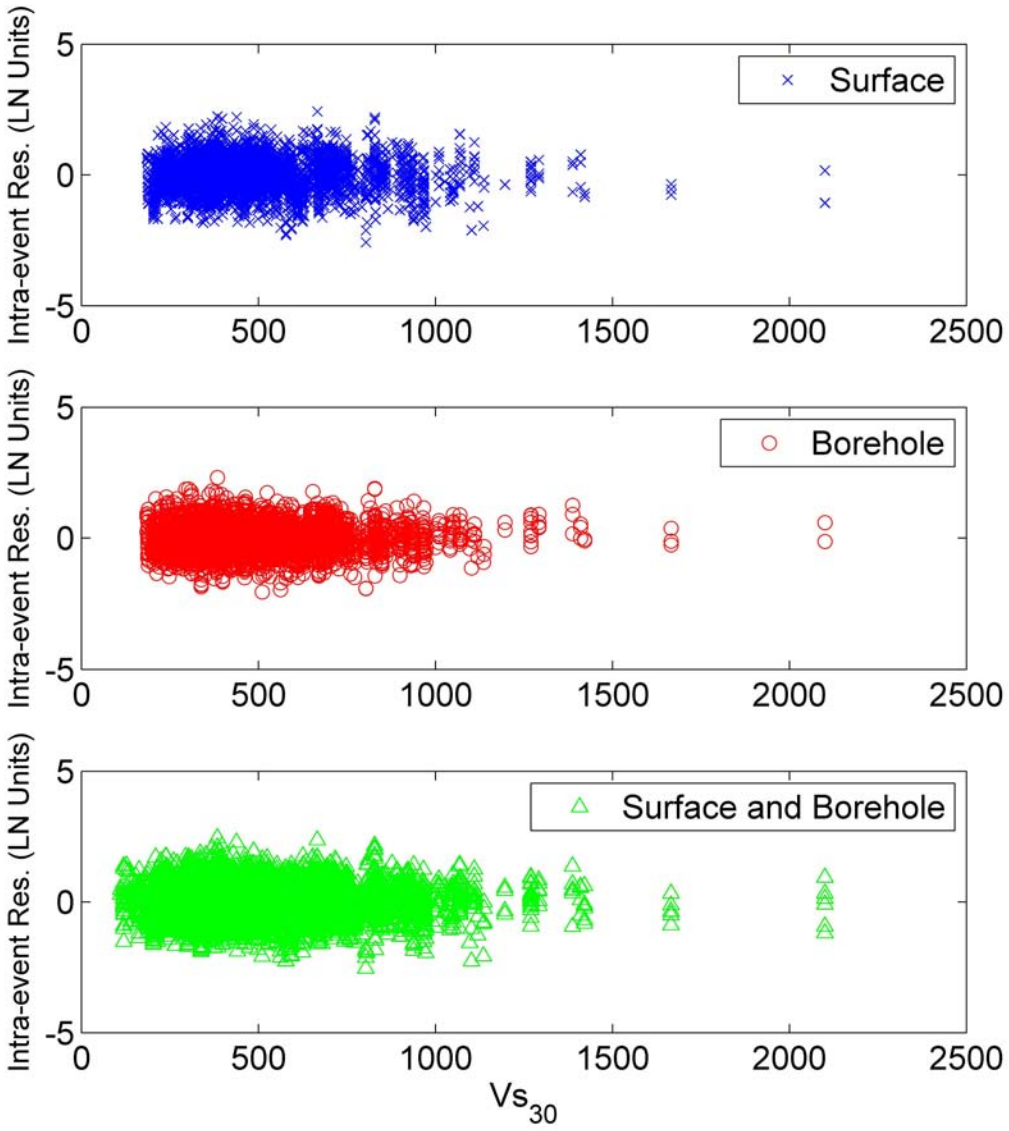


Figure A.1. Intra-Event Residuals for spectral period of 0.03 seconds and the models for Surface, Borehole, and Combined versus Vs_{30}.

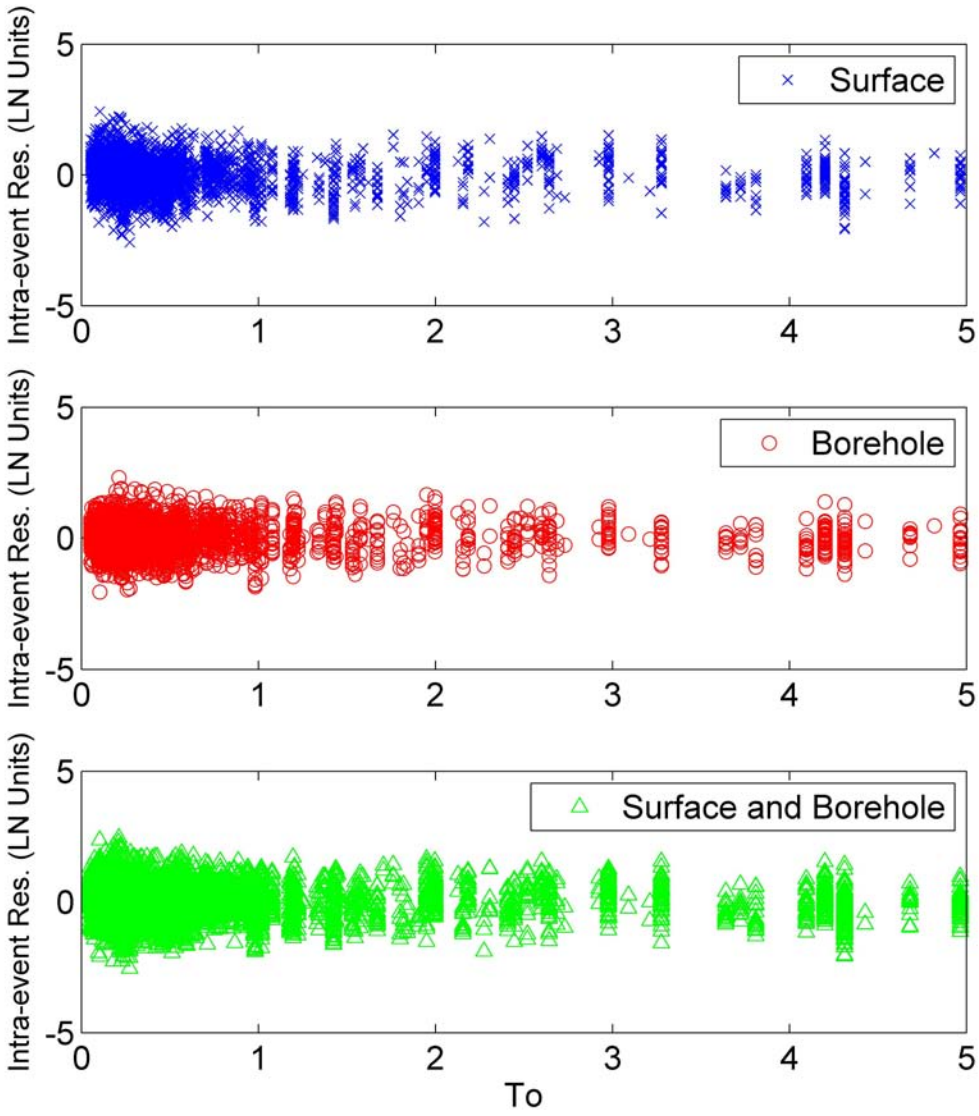


Figure A.2. Intra-Event Residuals for spectral period of 0.03 seconds and the models for Surface, Borehole, and Combined versus predominant period (To).

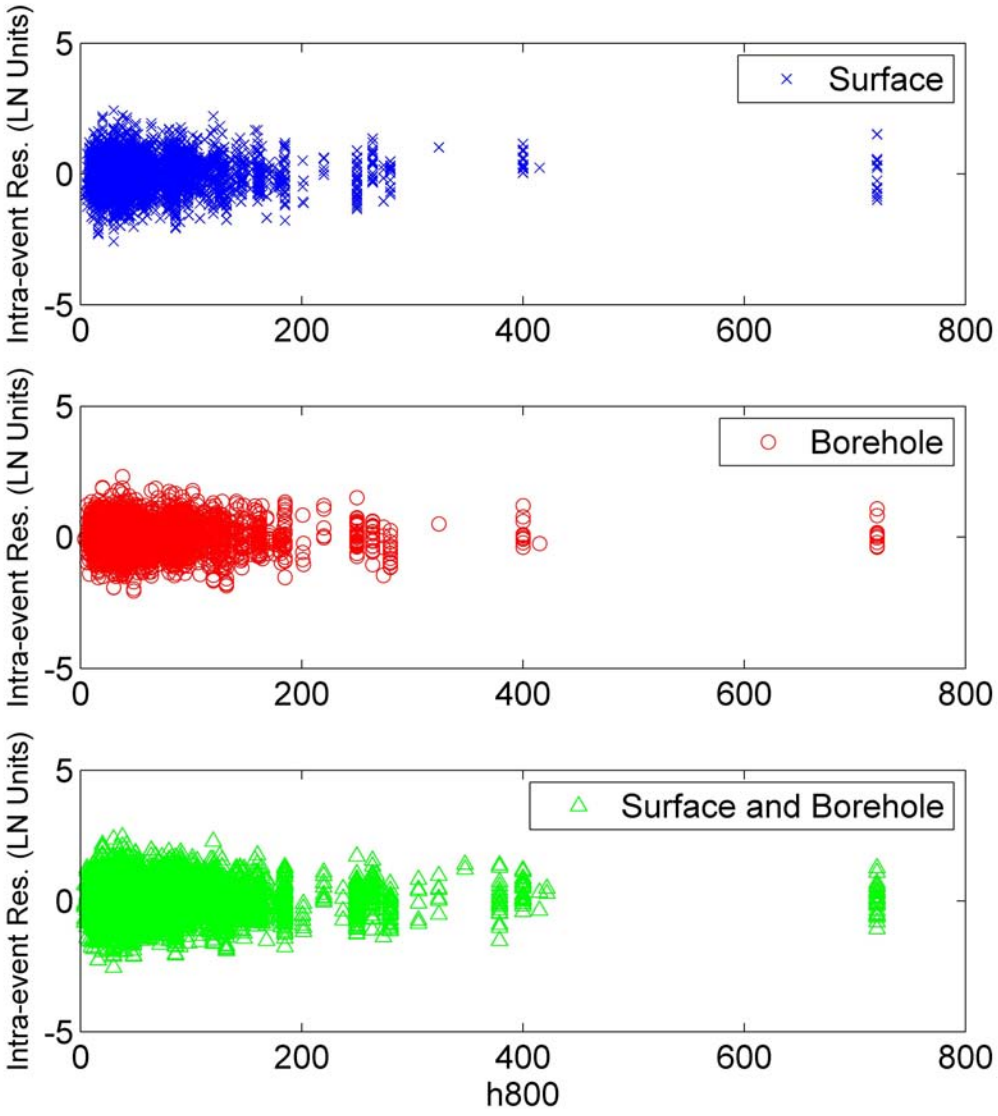


Figure A.3. Intra-Event Residuals for spectral period of 0.03 seconds and the models for Surface, Borehole, and Combined versus depth to shear-wave velocity of 800 (m/s).

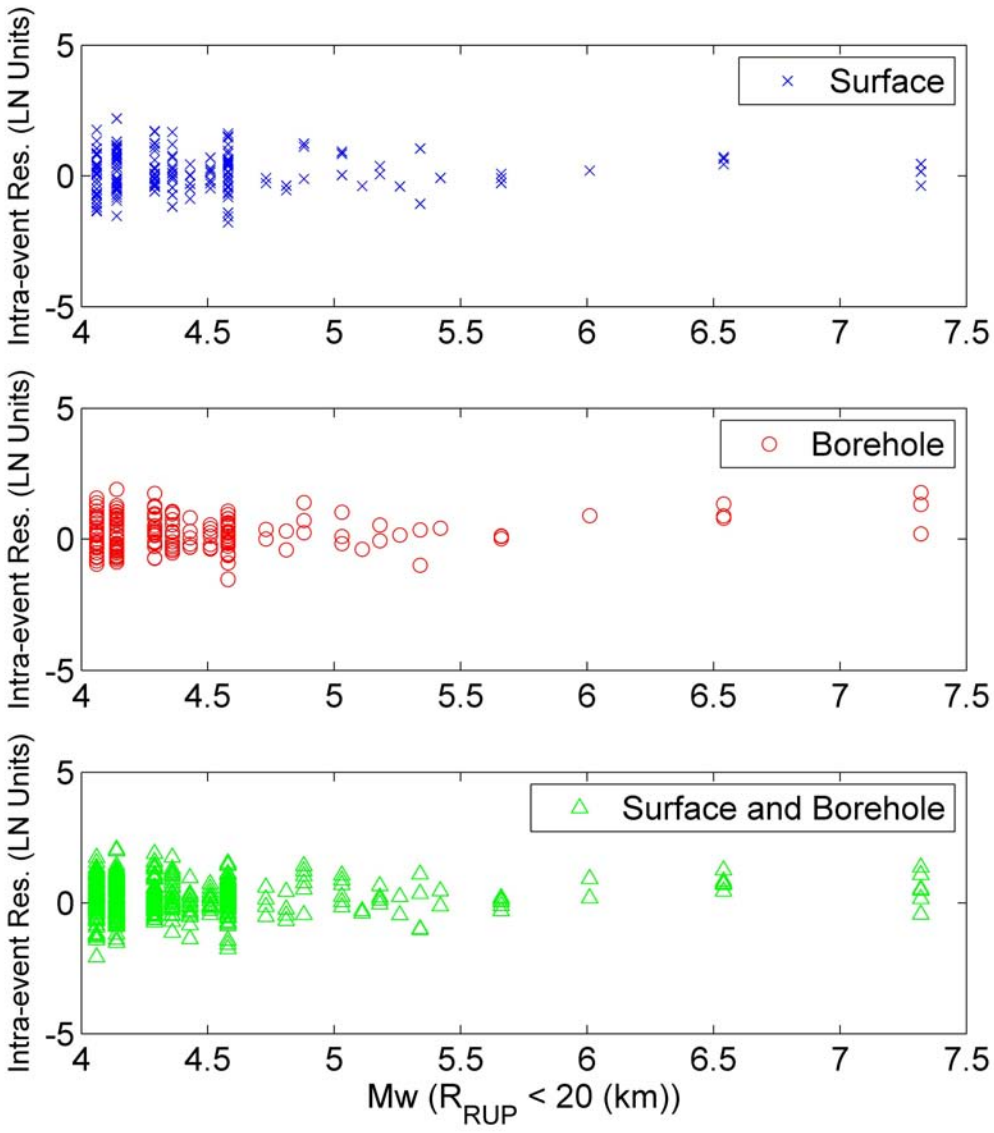


Figure A.4. Intra-Event Residuals for spectral period of 0.03 seconds and the models for Surface, Borehole, and Combined versus Magnitude, considering only events closer than 20 (km).

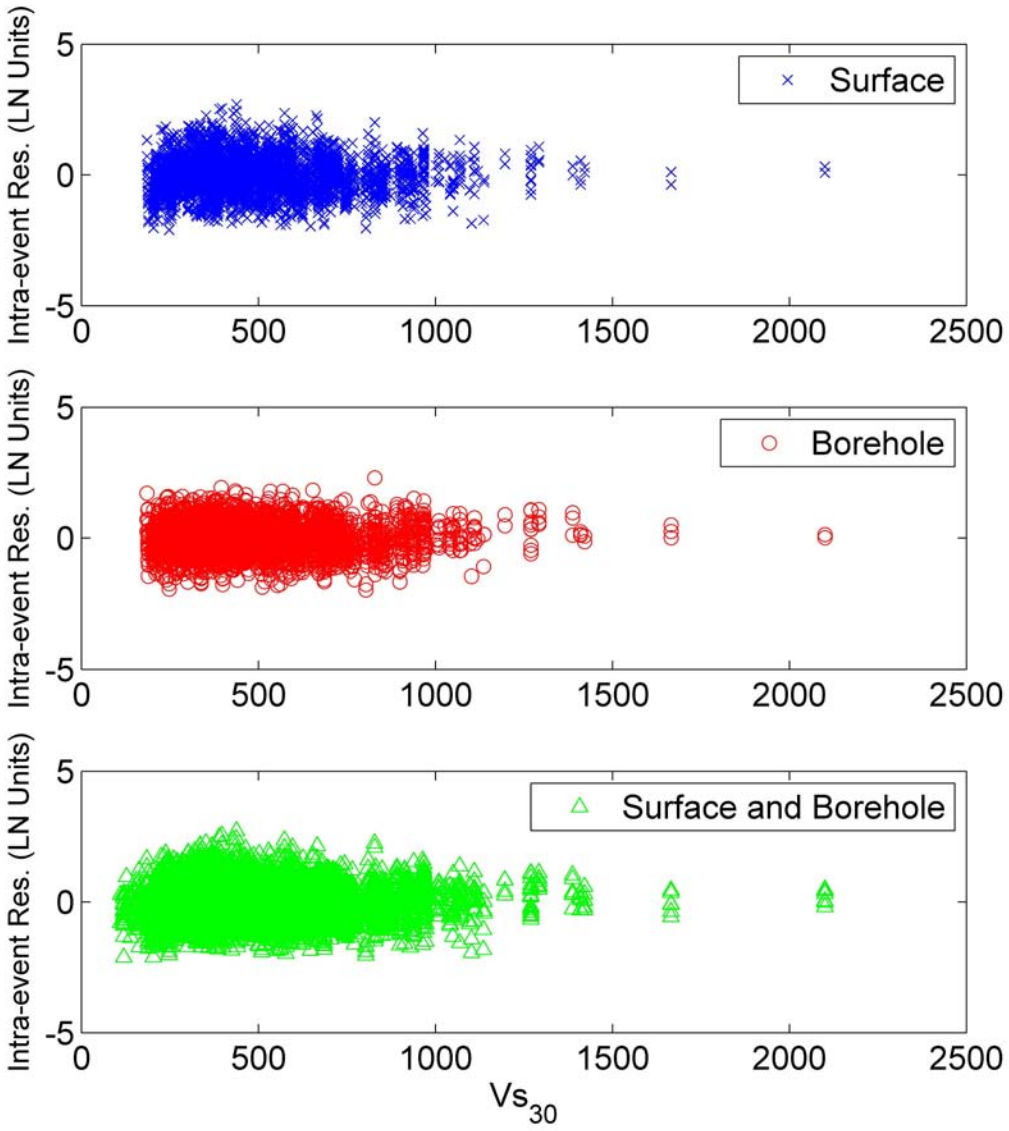


Figure A.5. Intra-Event Residuals for spectral period of 0.2 seconds and the models for Surface, Borehole, and Combined versus Vs_{30}.

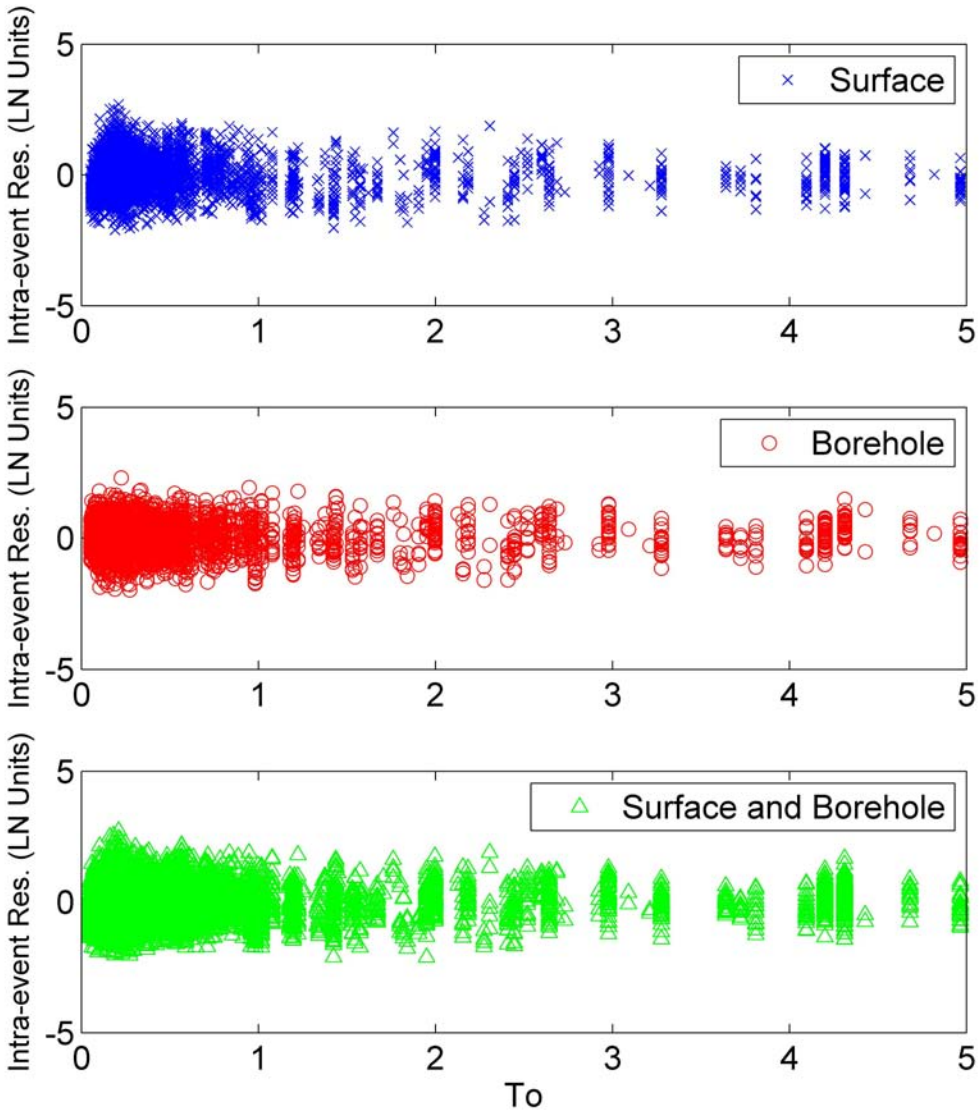


Figure A.6. Intra-Event Residuals for spectral period of 0.2 seconds and the models for Surface, Borehole, and Combined versus predominant period (To).

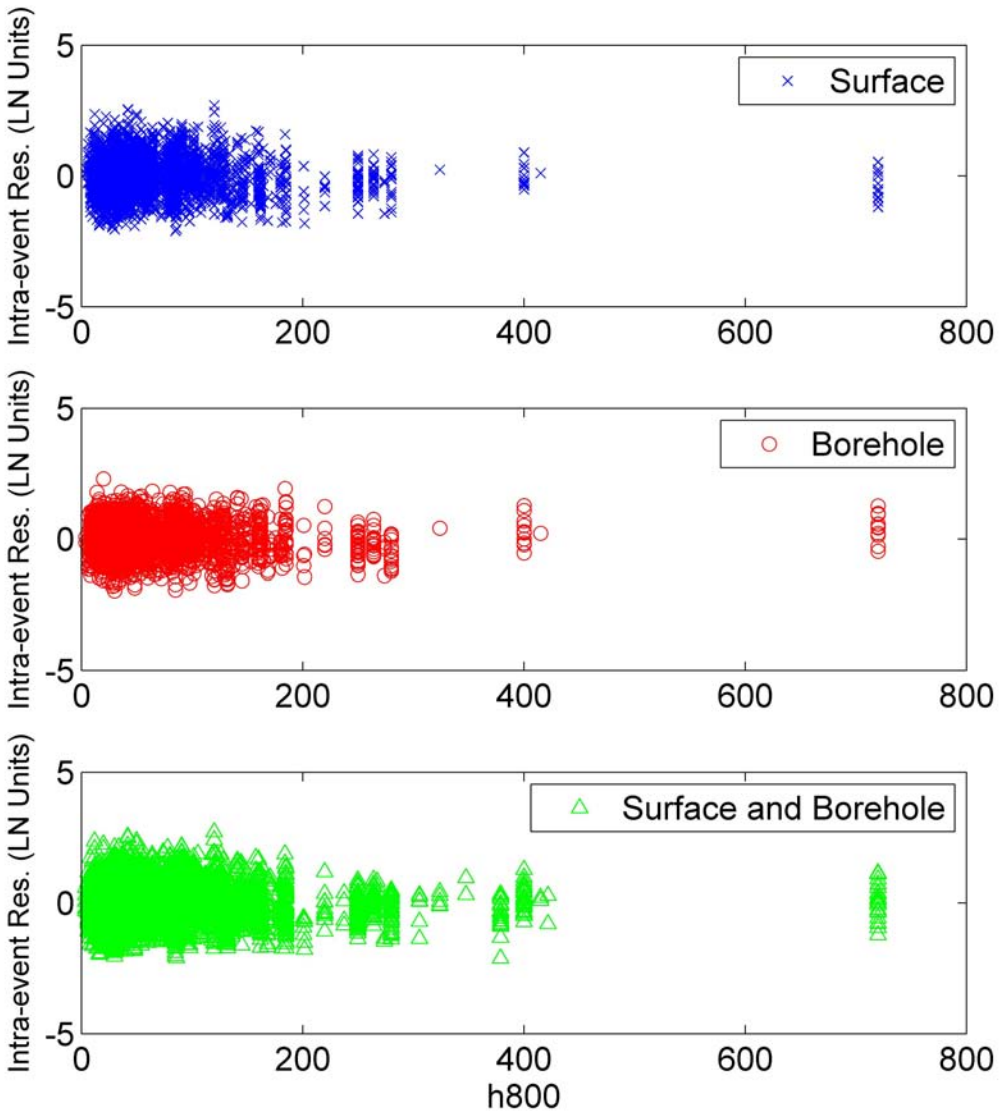


Figure A.7. Intra-Event Residuals for spectral period of 0.2 seconds and the models for Surface, Borehole, and Combined versus depth to shear-wave velocity of 800 (m/s).

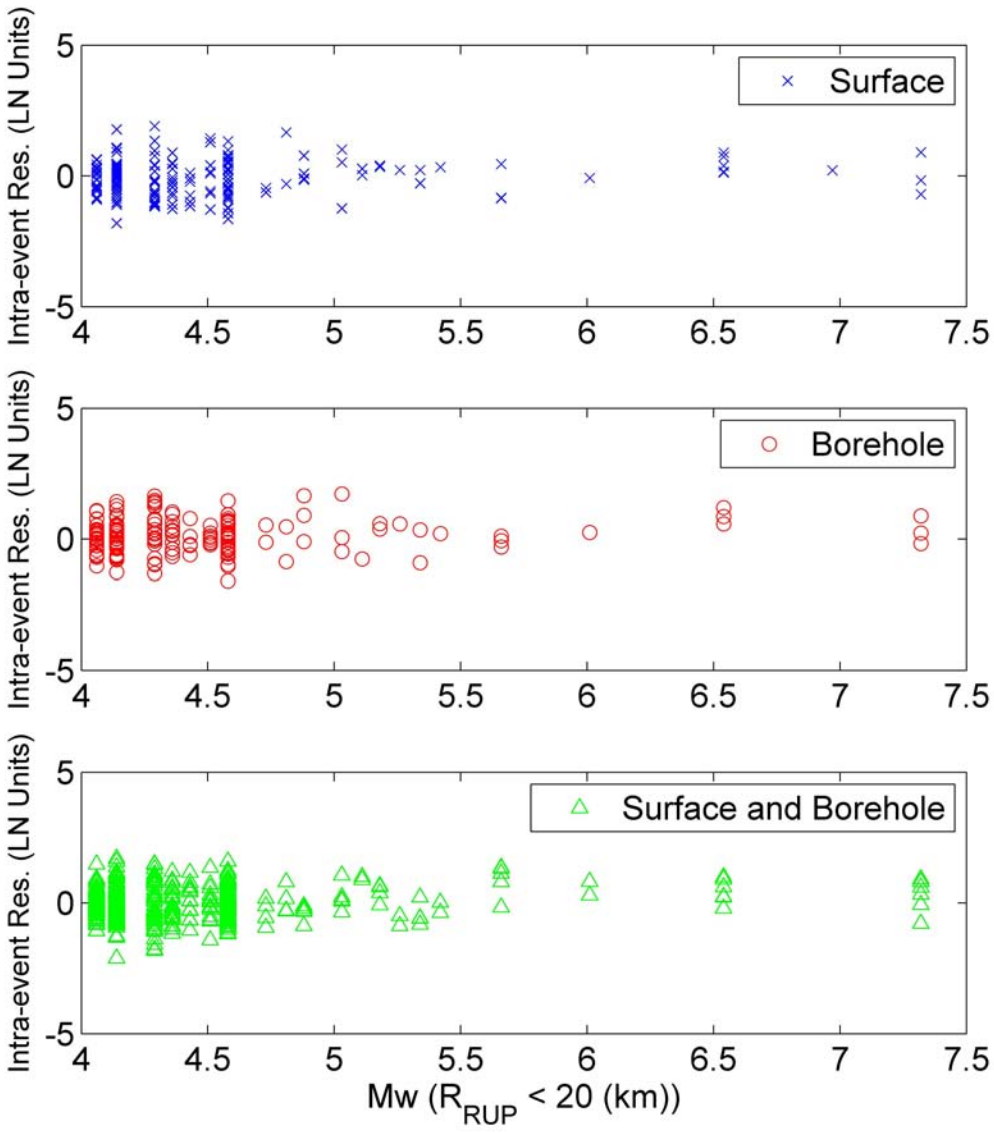


Figure A.8. Intra-Event Residuals for spectral period of 0.2 seconds and the models for Surface, Borehole, and Combined versus Magnitude, considering only events closer than 20 (km).

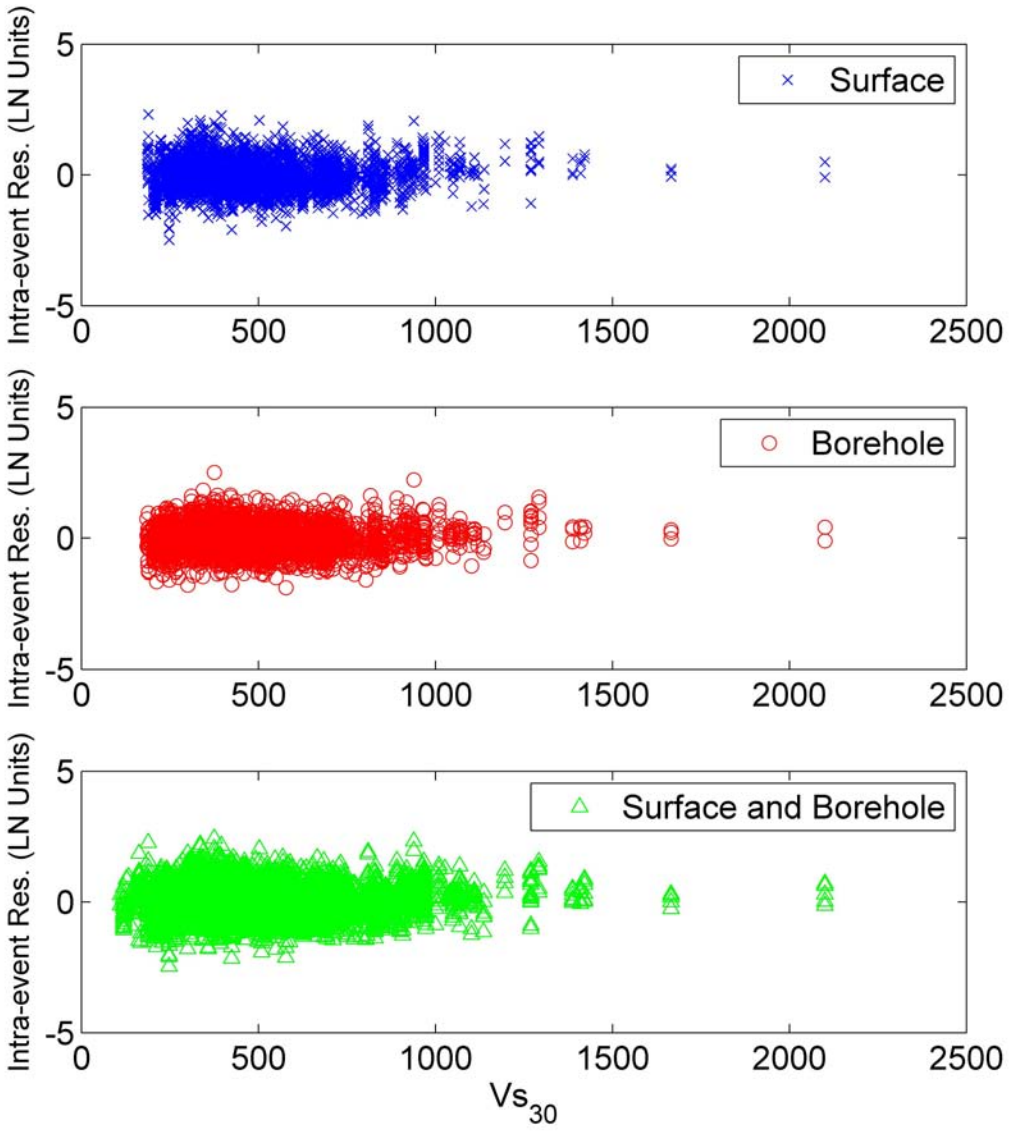


Figure A.9. Intra-Event Residuals for spectral period of 0.6 seconds and the models for Surface, Borehole, and Combined versus Vs_{30}.

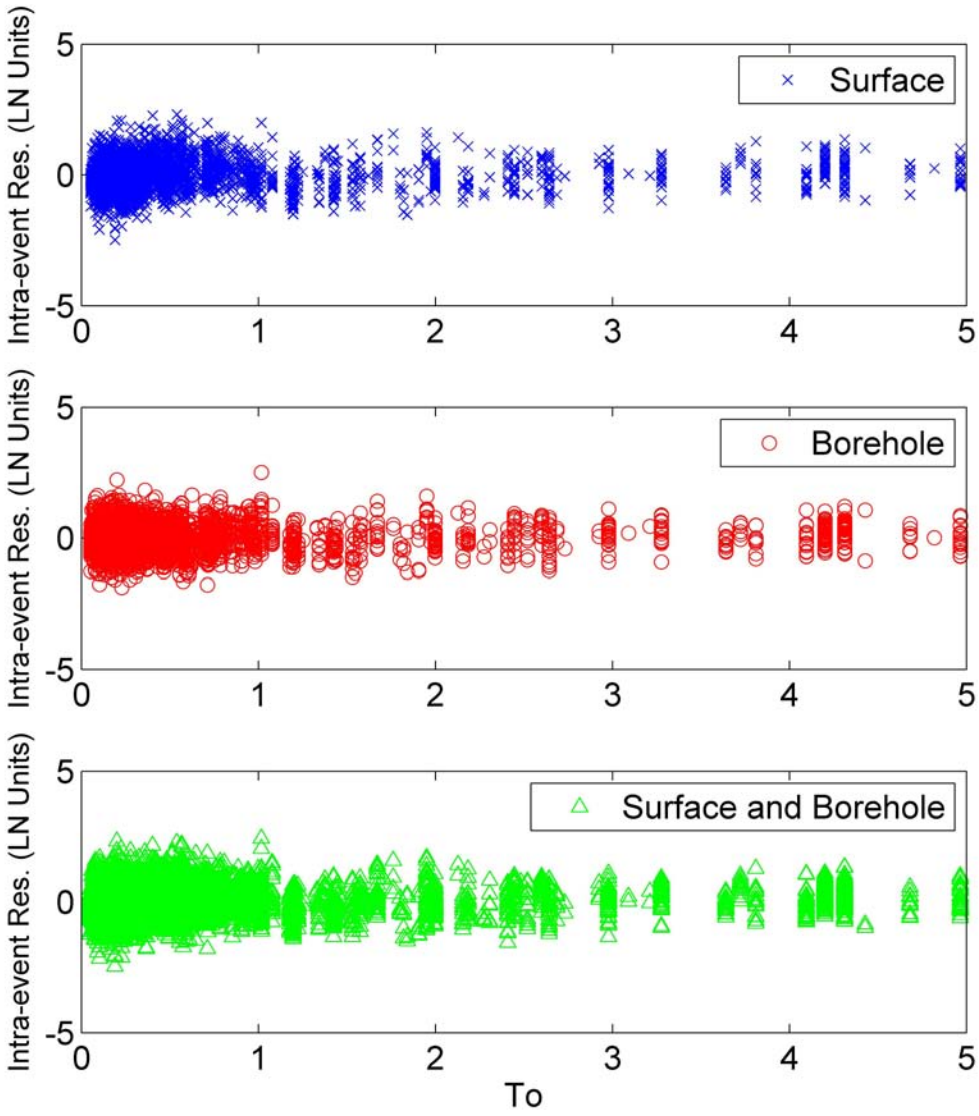


Figure A.10. Intra-Event Residuals for spectral period of 0.6 seconds and the models for Surface, Borehole, and Combined versus predominant period (To).

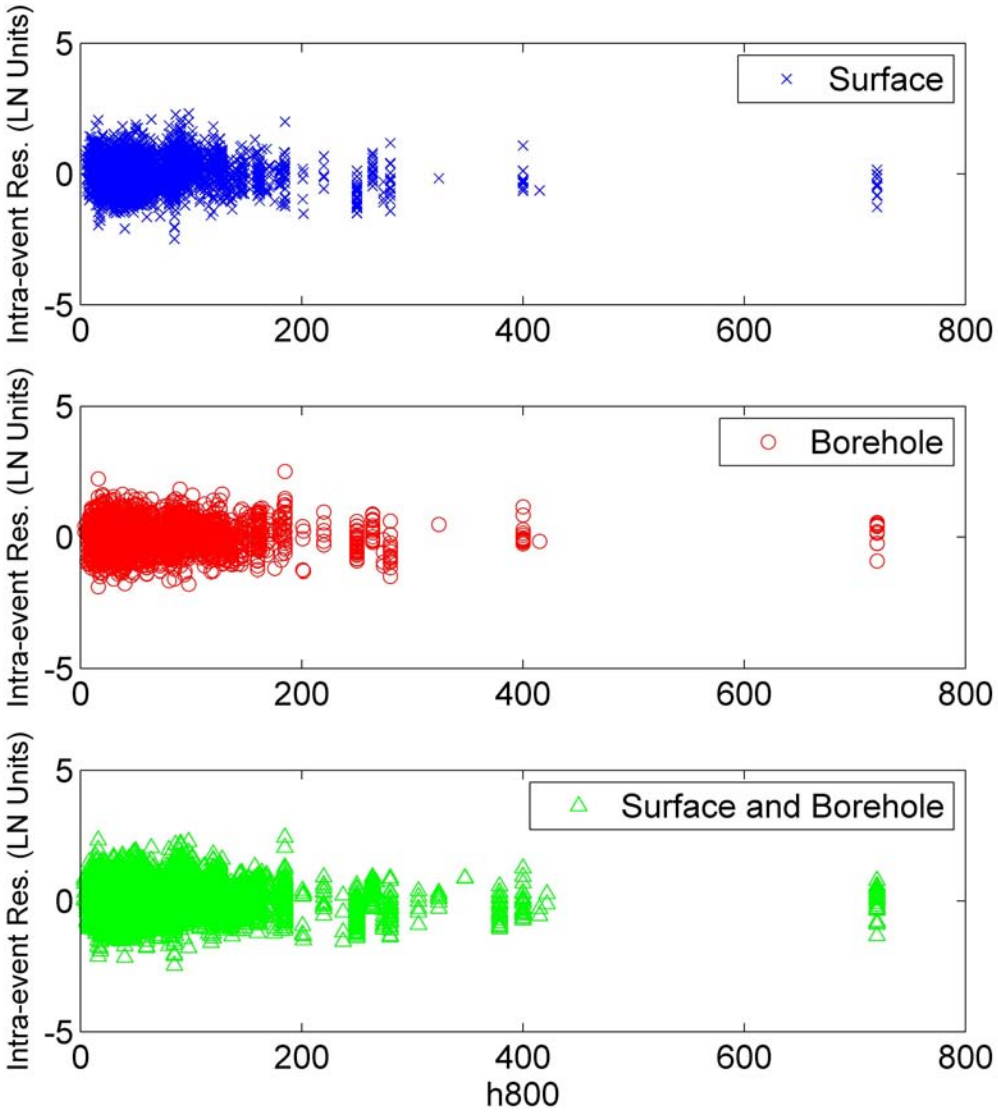


Figure A.11. Intra-Event Residuals for spectral period of 0.6 seconds and the models for Surface, Borehole, and Combined versus depth to shear-wave velocity of 800 (m/s).

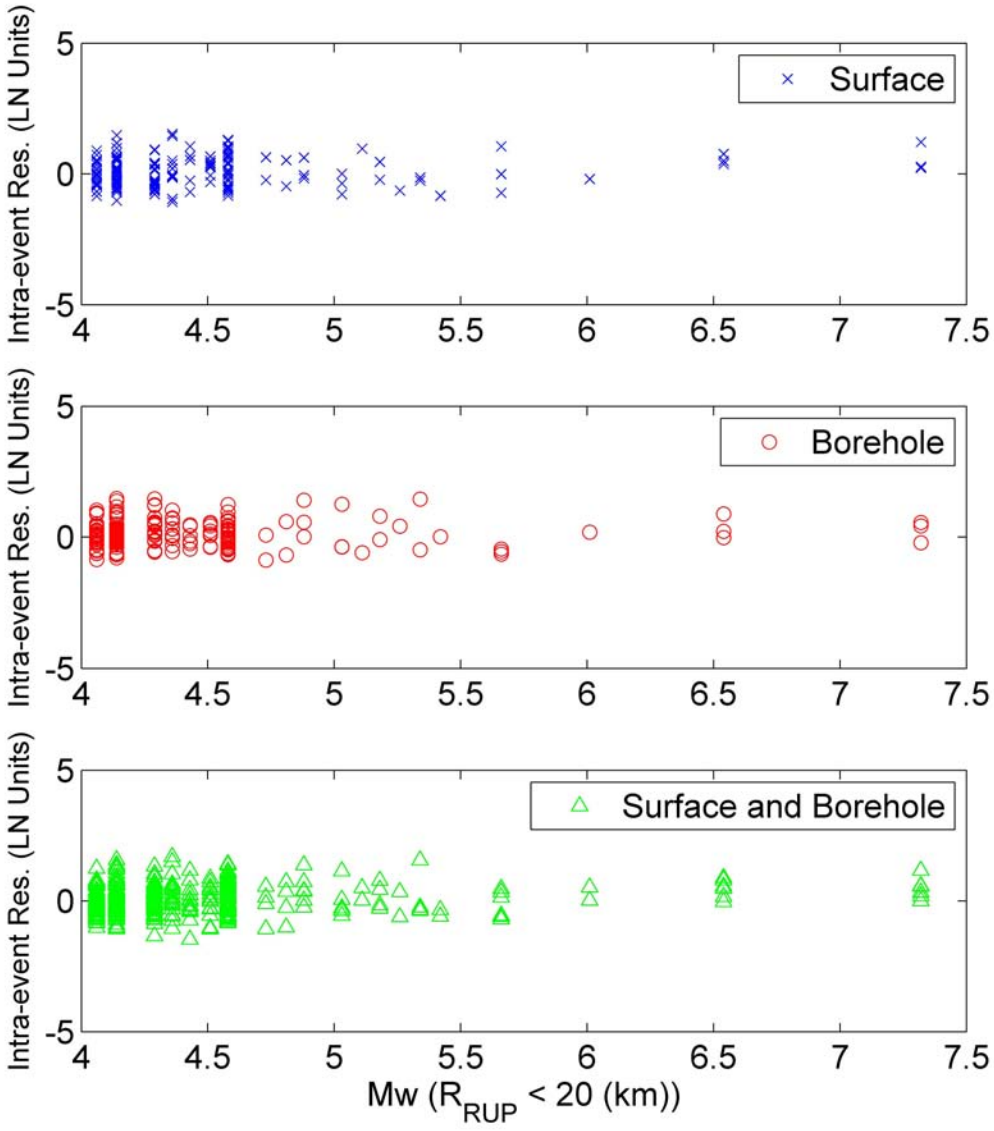


Figure A.12. Intra-Event Residuals for spectral period of 0.6 seconds and the models for Surface, Borehole, and Combined versus Magnitude, considering only events closer than 20 (km).

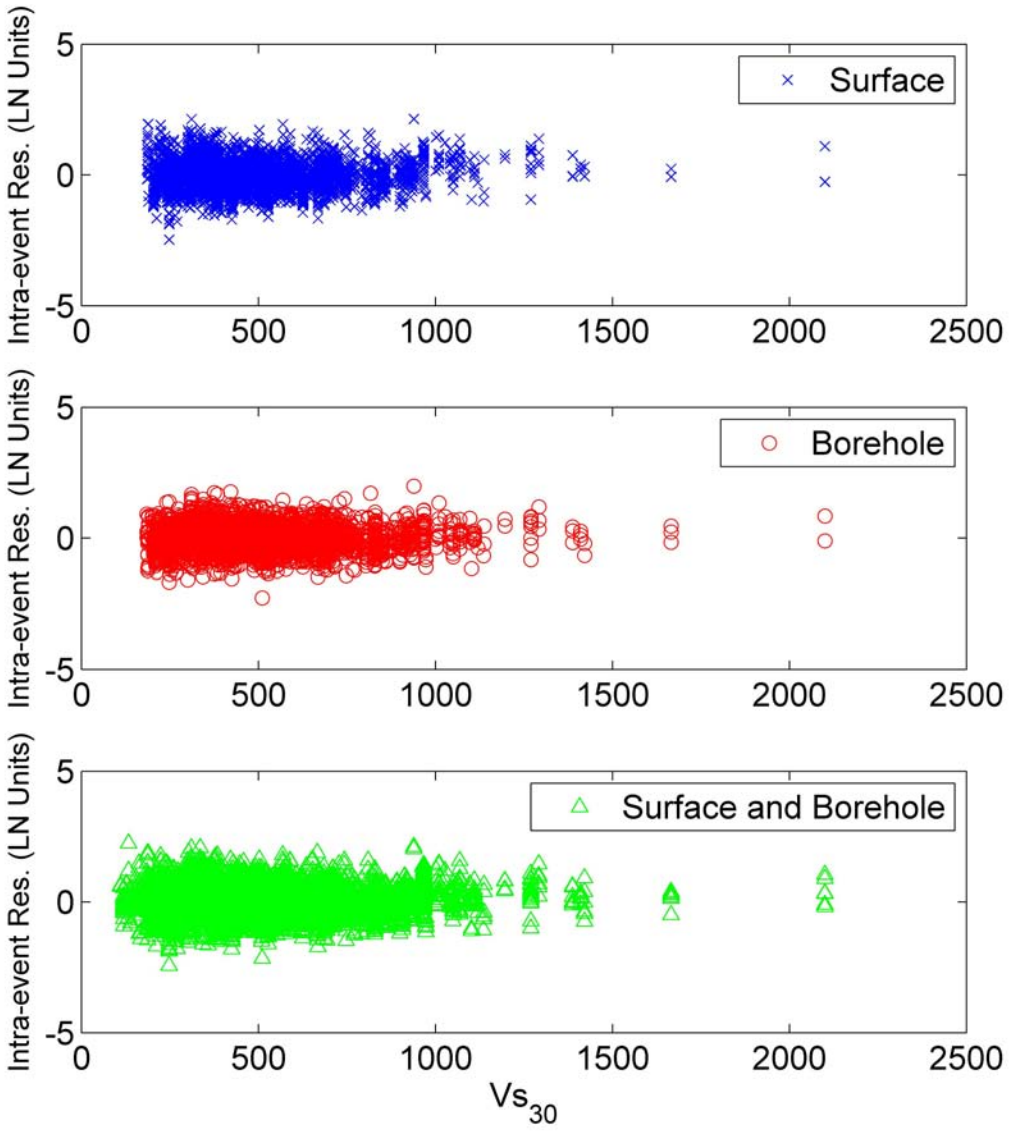


Figure A.13. Intra-Event Residuals for spectral period of 1.0 second and the models for Surface, Borehole, and Combined versus Vs_{30}.

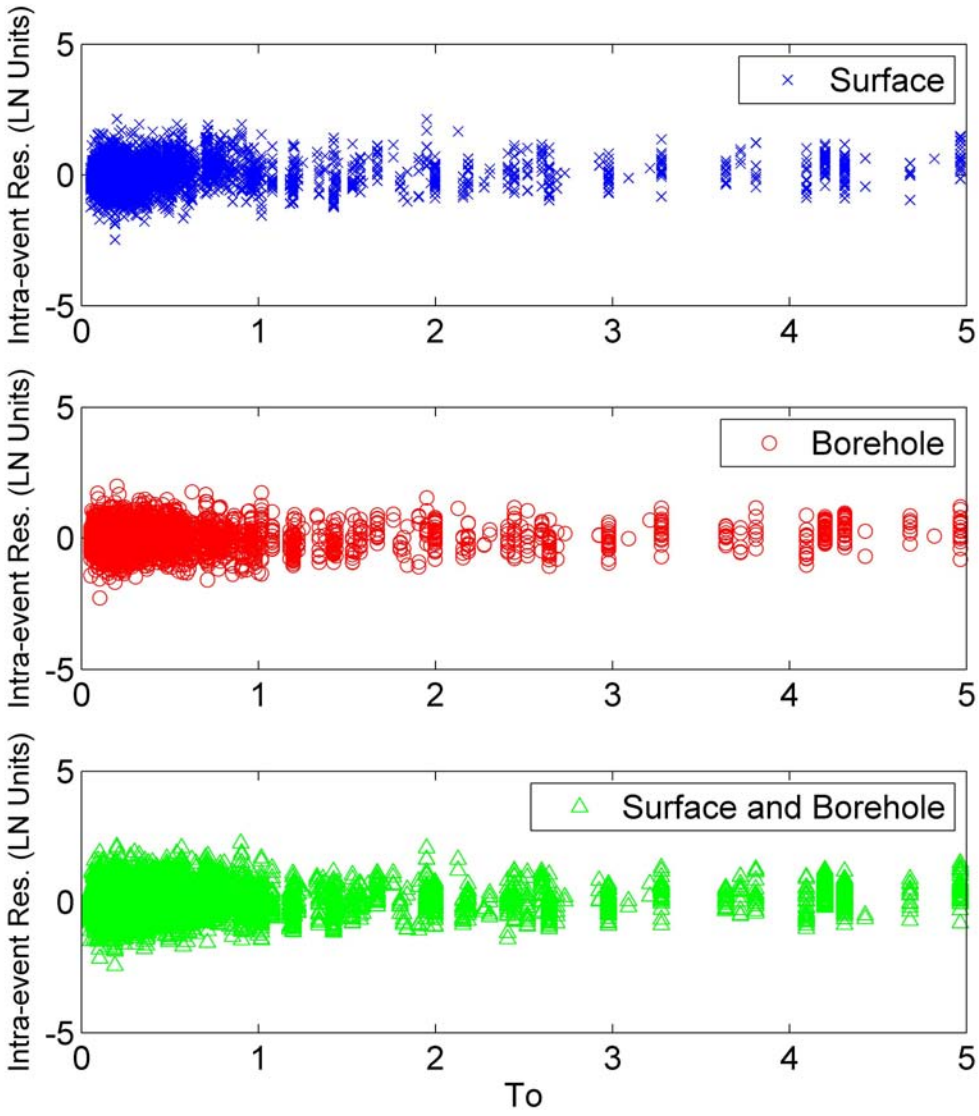


Figure A.14. Intra-Event Residuals for spectral period of 1.0 second and the models for Surface, Borehole, and Combined versus predominant period (To).

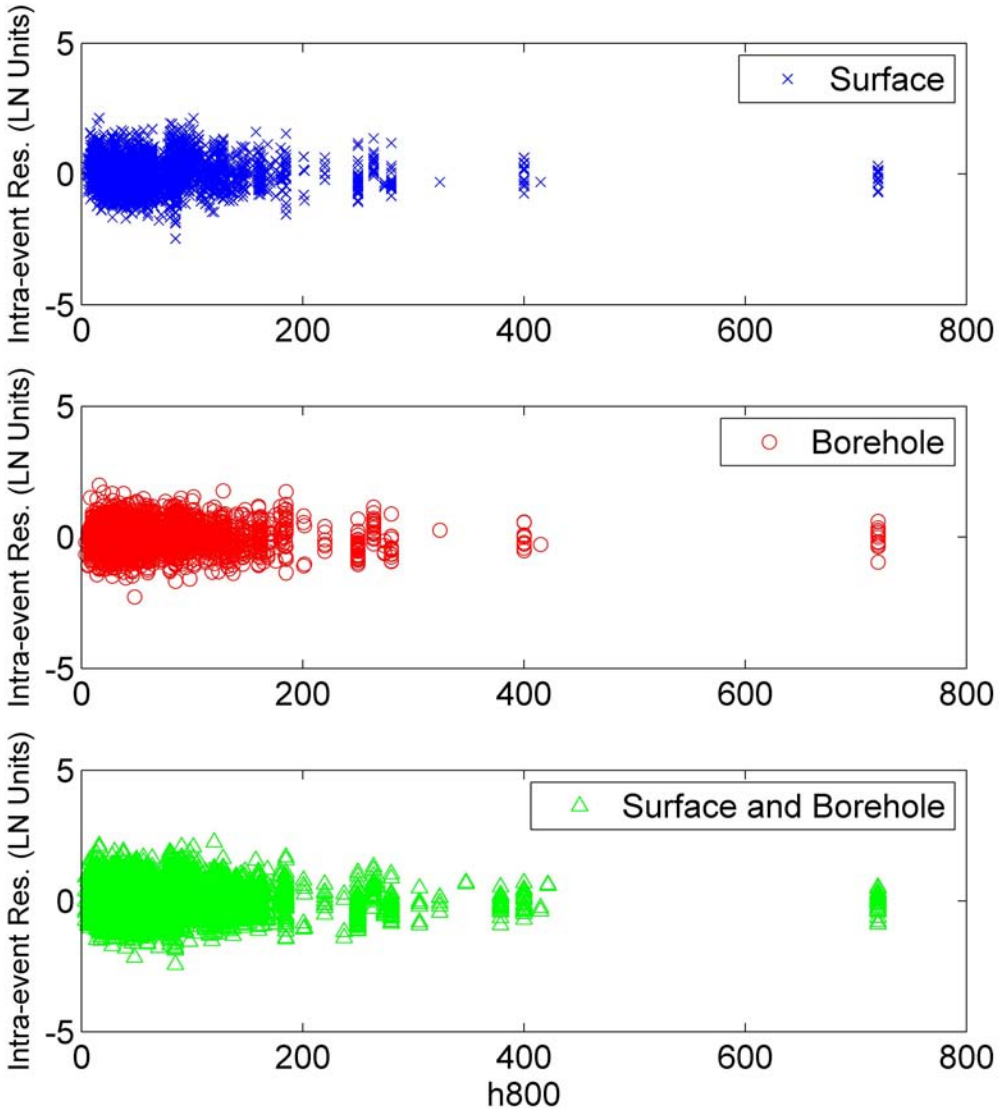


Figure A.15. Intra-Event Residuals for spectral period of 1.0 second and the models for Surface, Borehole, and Combined versus depth to shear-wave velocity of 800 (m/s).

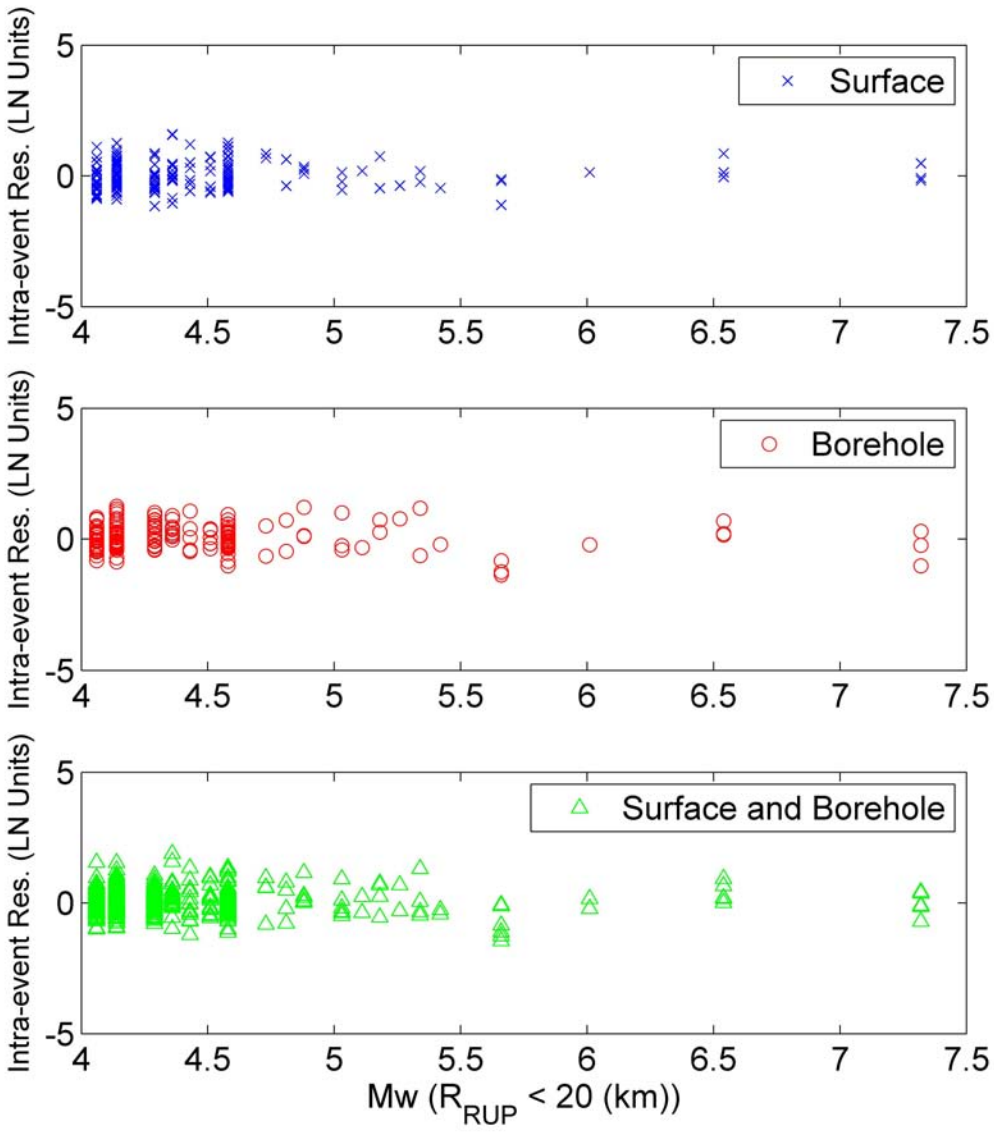


Figure A.16. Intra-Event Residuals for spectral period of 1.0 second and the models for Surface, Borehole, and Combined versus Magnitude, considering only events closer than 20 (km).

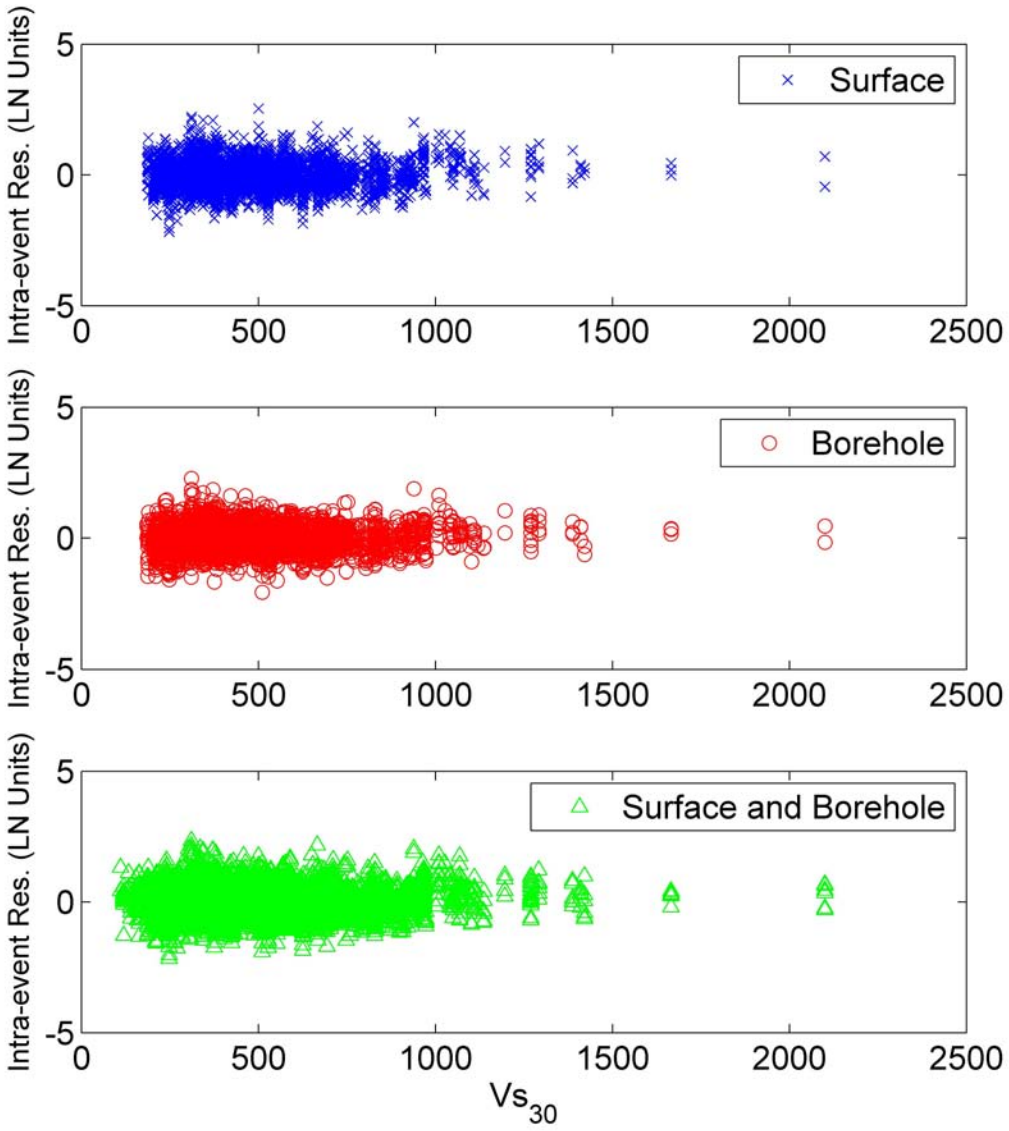


Figure A.17. Intra-Event Residuals for spectral period of 1.4 seconds and the models for Surface, Borehole, and Combined versus Vs_{30}.

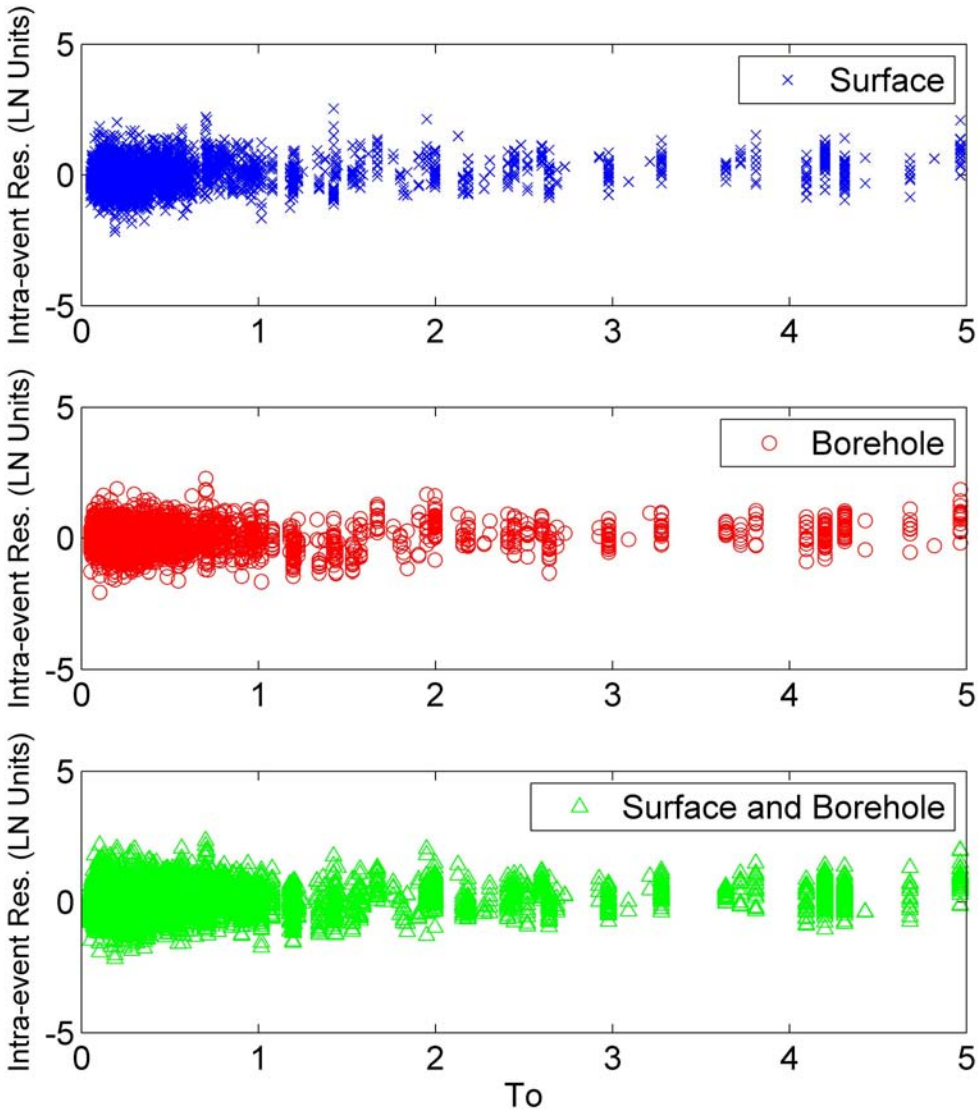


Figure A.18. Intra-Event Residuals for spectral period of 1.4 seconds and the models for Surface, Borehole, and Combined versus predominant period (To)

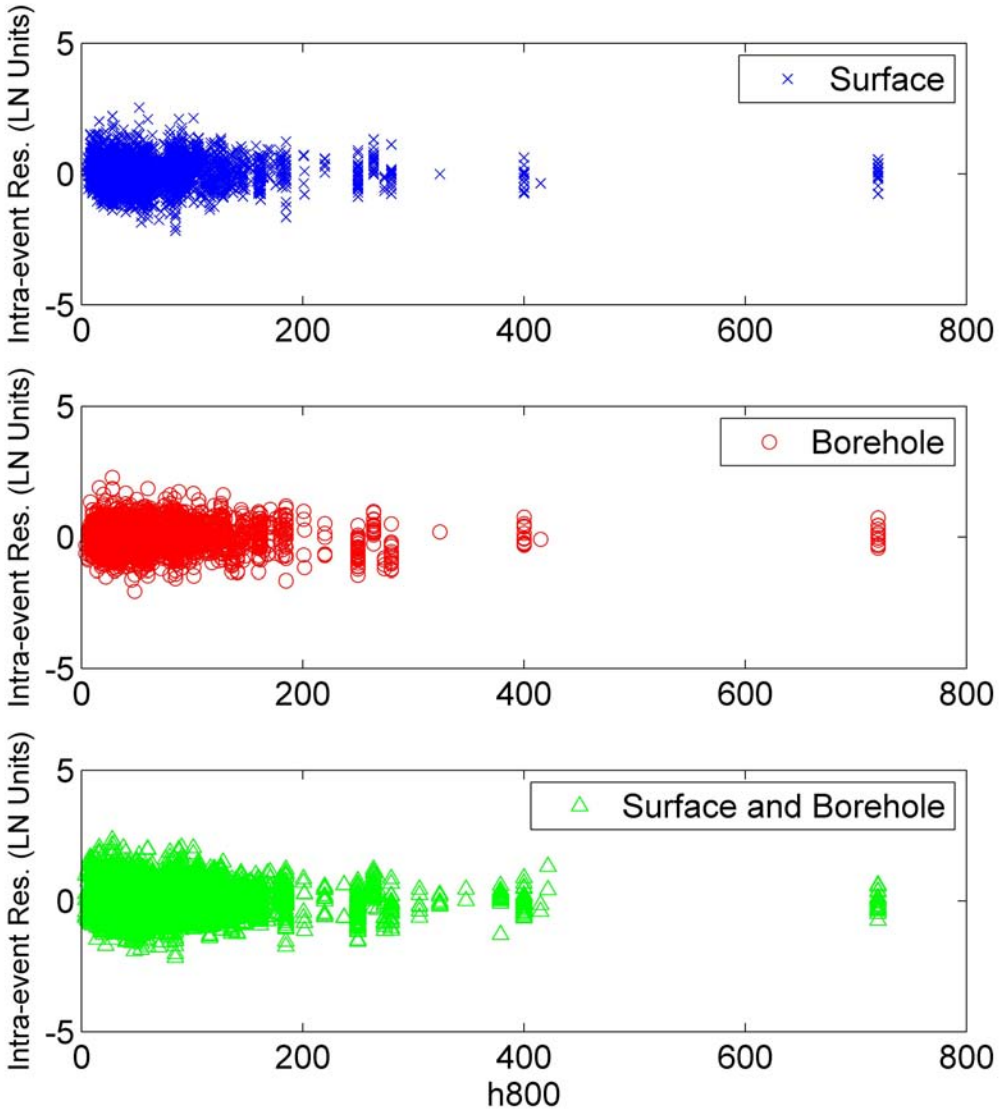


Figure A.19. Intra-Event Residuals for spectral period of 1.4 seconds and the models for Surface, Borehole, and Combined versus depth to shear-wave velocity of 800 (m/s).

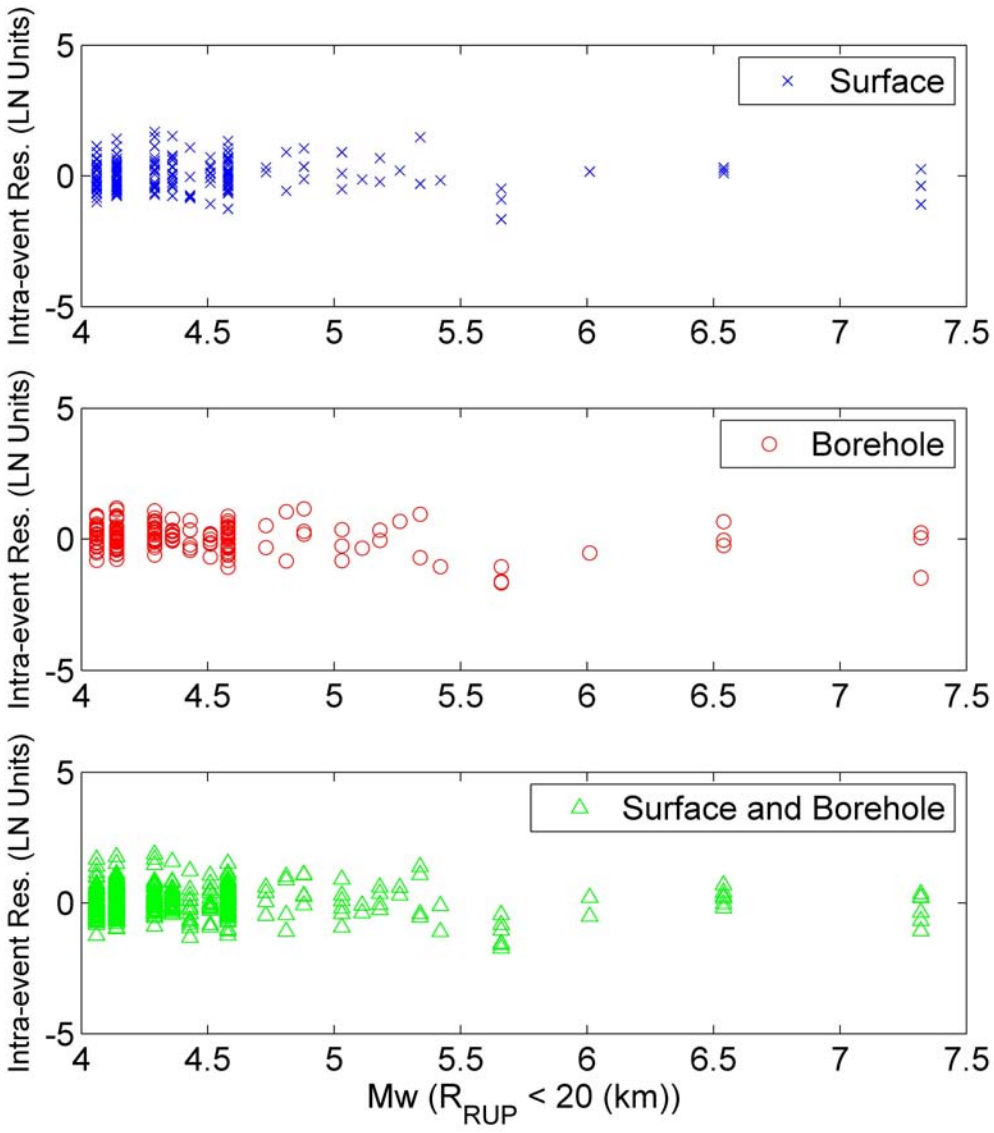


Figure A.20. Intra-Event Residuals for spectral period of 1.4 seconds and the models for Surface, Borehole, and Combined versus Magnitude, considering only events closer than 20 (km)