NAM

Development of GMPEs for Response Spectral Accelerations and for Strong-Motion Durations (Version 1)

Julian J Bommer, Peter J Stafford, Benjamin Edwards, Bernard Dost and Michail Ntinalexis

Datum Juni 2015

Editors Jan van Elk & Dirk Doornhof

General Introduction

The hazard due to induced earthquakes is presented by the ground motion buildings and people are subjected to. The prediction of ground motion resulting from the earthquakes in the Groningen area induced by the production of gas, is therefore critical.

NAM has assembled a team of experts in the field of GMP (Ground Motion prediction) to prepare a methodology for assessing ground motions, due to the induced earthquakes in the Groningen area. This team is led by Julian Bommer and consists of academics from various universities and knowledge institutes.

Main members of this team are:

External Expert	Affiliation	Role	Main Expertise Area
Julian Bommer	Independent Consultant, London	Collaborator	Ground Motion Prediction and Site Response
Ben Edwards	University Liverpool	Collaborator	Ground Motion Prediction
Michail Ntinalexis	Independent	Collaborator	Ground Motion Prediction
Barbara Polidoro	Independent Consultant, London	Collaborator	Ground Motion Prediction
Adrian Rodriguez - Marek	Virginia Tech, USA	Collaborator	Site Response Assessment
Peter Stafford	Imperial College London	Collaborator	Ground Motion Prediction
Sinan Akkar	Bogazici, University Istanbul	Collaborator	Ground Motion Prediction

The research was done in close cooperation with experts from KNMI.

In the Technical Addendum to Winningsplan 2013, a Ground Motion Prediction methodology based on a catalogue of tectonic earthquakes in southern Europe, was presented. This methodology was therefore conservative in the sense that it predicted ground motions which in future are, in general, more likely to be adjusted downwards than upwards. This report described the first update of this Ground Motion Prediction methodology tailored to the Groningen situation. This update has led to a downward adjustment of assessed ground motions for larger earthquakes, resulting in a reduction of the assessed hazard.

A further update is in preparation. Building on the model for shallow geology of Groningen prepared by Deltares, this further update will also include the effect of local shallow ground and soil conditions,.

Assurance for this study is primarily based on cross-validation (peer review) between parties involved. The studies into the fragility of buildings will be reviewed by a panel of independent experts. The following experts have been invited.

External Expert	Affiliation	Role	Main Expertise Area
Gail Atkinson	Western University, Ontario, Canada	Independent Reviewer	Ground Motion Prediction
Hilmar Bungum	NORSAR, Norway	Independent Reviewer	Ground Motion Prediction
Fabrice Cotton	GFZ Potsdam, Germany	Independent Reviewer	Ground Motion Prediction
John Douglas	University of Strathclyde, UK	Independent Reviewer	Ground Motion Prediction
Jonathan Stewart	UCLA, California, USA	Independent Reviewer	Ground Motion Prediction
Ivan Wong	AECOM, Oakland, USA	Independent Reviewer	Ground Motion Prediction
Bob Youngs	AMEC, Oakland, USA	Independent Reviewer	Ground Motion Prediction

NAM

Title	**Development of GMPEs for Response Spectral Accelerations and for Strong-Motion Durations (V1)**		**Date**	June 2015
			Initiator	NAM
Autor(s)	Julian J Bommer, Peter J Stafford, Benjamin Edwards, Bernard Dost and Michail Ntinalexis	**Editors**	Jan van Elk Dirk Doornhof	
Organisation		**Organisation**	NAM	
Place in the Staudy and Data Acquisition Plan	Research Theme: Ground Motion Prediction The prediction of Ground Motion is central to the hazard assessment. This report describes an update of the Ground Motion Prediction methodology based on the Groningen situation.			
Directly linked research	(1) Hazard Assessment. (2) Fragility assessment of buildings in the Groningen region.			
Used data	Accelerograms from the accelerometers placed in the Groningen field.			
Associated organisations	KNMI			
Assurance	Assurance team has been arranged. Peer review is in progress.			

Development of Version 1 GMPEs for Response Spectral Accelerations and for Strong-Motion Durations

**Julian J Bommer, Peter J Stafford, Benjamin Edwards,
Bernard Dost & Michail Ntinalexis**

Version 2

21 June 2015

Version Control and Revision Record

Version / Date	Comments and Changes
Version 0 2nd March 2015	First draft. Chapters 1-7 only, no Appendices, issued for review and checking by authors
Version 1 16th March 2015	Complete draft including Chapter 8 and Appendices; several minor corrections from authors' review; issued as pre-read for workshops with ExxonMobil and Scientific Advisory Committee held in Assen on 8th/9th and 22nd/23rd April respectively
Version 2 21 June 2015	— Additional acknowledgements added — Executive Summary updated with the corrected sigma model and additional notes on interpolation and extrapolation of sigma values, and also to incorporate the R_{JB} model — Chapter 1 extended to include full overview of the final report including the inclusion of duration models — Section 2.2 updated with discussion on M_L-**M** relationship and ongoing work on this topic — Section 2.4 updated to reflect choice of two distance metrics (R_{epi} & R_{JB}) — Section 2.5 updated to reflect the decision not to use V_{S30} — Introduction to Chapter 4 modified to reflect decision to use simulations rather than European strong-motion data for models at larger magnitudes — Section 4.3 updated to reflect final purpose of European database — Introduction to Chapter 5 modified to reflect final GMPE development — Figure 5.23 replaced with clearer illustration of stress drop distribution — Section 5.4 extended to include discussion of low stress drop from induced and shallow earthquakes, including additional references from 2015 SSA meeting — Introduction to Chapter 6 updated to reflect final scope, including R_{JB} model — Section 6.2 updated to reflect final sigma model formulation and also the decision not to invoke a sigma reduction for larger magnitudes — Note in to Section 6.3 cross-referencing stress drop discussion in Section 5.4 — All plots in Section 6.4 replaced to reflect correction to sigma model — Additional plots in Section 6.5 to compare V0 and V1 spectral shapes — Figure 6.74 (previously mislabelled as 6.73) modified for consistent colours on individual curves — New Section added (6.6) to present alternative GMPEs using distance metric based on extended fault rupture — Minor correction to discussion of Kempton & Stewart (2006) residuals in Section 7.2 with reference to Figure 7.17 — Section 8.1 updated to discuss only risk sensitivity studies using R_{epi}- and R_{JB}-based GMPEs and not derivation of the latter (now in Section 6.6) — List of accelerograph stations identified for V_S measurements in Section 8.4 updated to reflect issues with access permits — References to Deltares reports on geological model and site response analyses added to Section 8.4 — Section 8.4 expanded to included overview of envisaged procedure for the development of the V2 GMPEs for response spectral ordinates — Additions, updates and minor corrections to Reference list — Correction of small typo in Appendix I — Appendix II extensively updated to reflect final sigma correction model

Table of Contents

Executive Summary

Ground-motion prediction equations have been derived for the estimation of peak ground acceleration (PGA) and response spectral accelerations, Sa(T), at oscillator periods, T, of 0.2, 0.5, 1.0 and 2.0 seconds, for application to the Version 1 seismic hazard and risk model for the Groningen field. In all cases, the equations predict the geometric mean of the two horizontal components of motion and yield values of acceleration in units of cm/s^2. The predictions are a function of only two independent variables: the moment magnitude, **M**, and the epicentral distance, R_{epi}, in km. The predicted accelerations correspond to the motions at the ground surface across the entire field, irrespective of local variations in soil conditions.

The equations have the following functional form, with variation of the quadratic scaling with magnitude for earthquakes above and below **M** 4.5:

$$\ln(Y) = c_1 + c_2 M + c_3 (M - 4.5)^2 + c_4 \ln\sqrt{R_{epi}^2 + [\exp(c_5 M + c_6)]^2} \qquad M \le 4.5 \qquad \text{(ES.1a)}$$

$$\ln(Y) = c_1 + c_2 M + c_{3a} (M - 4.5)^2 + c_4 \ln\sqrt{R_{epi}^2 + [\exp(c_5 M + c_6)]^2} \qquad M > 4.5 \qquad \text{(ES.1b)}$$

where *ln*(Y) is the acceleration in cm/s^2. In order to capture the large epistemic uncertainty in the ground-motion predictions, three models have been derived and assigned relative weights that would be applied to the models in a logic-tree formulation. The models have different coefficients for the median predicted accelerations and different values of the between-earthquake (inter-event) variability, τ, but with the same model for within-earthquake (intra-event) variability in all three cases. The three models and their weights are summarised in Table ES.1, with the coefficients of the equation for median values and the inter-event variability values presented separately in Tables ES.2-ES.4. The model for intra-event variability is explained subsequently.

Table ES.1. Summary of the 3 R_{epi}-based GMPEs in the ground-motion logic-tree

Model	Weight	Median Coefficients	Inter-event τ	Within-event ϕ
Upper	0.3	Table ES.3	Table ES.3	Table ES.5
Central	0.5	Table ES.2	Table ES.2	Table ES.5
Lower	0.2	Table ES.4	Table ES.4	Table ES.5

For applications requiring only a single ‘best estimate’ model, the central model may be deployed with a weight of unity, but the user making such a choice needs to be aware that there is very considerable epistemic uncertainty associated with the predictions for magnitudes above **M** 4 and that this uncertainty increases with the earthquake magnitude.

Table ES.2. Coefficients of Eq.(ES.1) and the inter-event variability of the CENTRAL GMPE model

	PGA	Sa(0.2s)	Sa(0.5s)	Sa(1.0s)	Sa(2.0s)
c_1	1.1563	2.4972	-0.0684	-4.3882	-7.8093
c_2	1.2732	1.1216	1.5742	2.2288	2.6929
c_3	-0.3394	-0.4314	-0.5416	-0.3549	-0.1520
c_{3a}	-0.1342	-0.0747	-0.2397	-0.4202	-0.4370
c_4	-1.5048	-1.4806	-1.2266	-1.1640	-1.1526
c_5	0.4233	0.4233	0.4233	0.4233	0.4233
c_6	-0.6083	-0.6083	-0.6083	-0.6083	-0.6083
τ	0.2810	0.3337	0.3216	0.3789	0.3547

Table ES.3. Coefficients of Eq.(ES.1) and the inter-event variability of the HIGHER GMPE model

	PGA	Sa(0.2s)	Sa(0.5s)	Sa(1.0s)	Sa(2.0s)
c_1	0.1638	1.5092	-1.7676	-5.9331	-8.5757
c_2	1.6566	1.4980	2.0695	2.6584	2.9277
c_3	-0.3236	-0.4312	-0.4308	-0.2273	-0.0983
c_{3a}	-0.2643	-0.2125	-0.4043	-0.5076	-0.4068
c_4	-1.5391	-1.4926	-1.2282	-1.1729	-1.1680
c_5	0.4233	0.4233	0.4233	0.4233	0.4233
c_6	-0.6083	-0.6083	-0.6083	-0.6083	-0.6083
τ	0.3581	0.416	0.3965	0.3965	0.3734

Table ES.4. Coefficients of Eq.(ES.1) and the inter-event variability of the LOWER GMPE model

	PGA	Sa(0.2s)	Sa(0.5s)	Sa(1.0s)	Sa(2.0s)
c_1	1.0490	2.1812	0.6494	-3.2480	-7.1140
c_2	1.1122	1.0202	1.2775	1.8682	2.4569
c_3	-0.3132	-0.3408	-0.5417	-0.4377	-0.2117
c_{3a}	-0.0942	-0.0544	-0.1430	-0.3306	-0.4442
c_4	-1.4529	-1.4670	-1.2223	-1.1500	-1.1324
c_5	0.4233	0.4233	0.4233	0.4233	0.4233
c_6	-0.6083	-0.6083	-0.6083	-0.6083	-0.6083
τ	0.2039	0.2514	0.2467	0.3612	0.3359

The logarithmic standard deviation, σ, is defined in terms of an inter-event component, τ, and an intra-event component, ϕ. The intra-event component is composed of a small-magnitude value, ϕ_{SM}, and a magnitude- and distance-dependent adjustment, $\delta\phi$ to correct for the use of a point-source distance metric:

$$\sigma = \sqrt{\tau^2 + \phi^2} = \sqrt{\tau^2 + (\phi_{SM}^2 + \delta\phi^2)} \qquad \text{(ES.2)}$$

The magnitude- and distance-dependent adjustment to the intra-event variability is defined as follows:

$$\delta\phi = SF.\frac{\varphi(z)}{\sigma_Z} \qquad M \geq 4 \quad \text{and} \quad R_{epi} > 0 \qquad \text{(ES.3a)}$$

$$\delta\phi = 0 \qquad M < 4 \quad \text{or} \quad R_{epi} = 0 \qquad \text{(ES.3b)}$$

where SF is the magnitude-dependent scaling factor, expressed as follows:

$$SF = \beta_1(M-4) + \beta_2(M-4)^2 \qquad \text{(ES.4)}$$

and $\varphi[\ ]$ is the normal probability density function, which is given by the following expression:

$$\varphi(z) = \frac{1}{\sqrt{2\pi}}\exp\left(\frac{-z^2}{2}\right) \qquad \text{(ES.5)}$$

The argument of this expression is given by:

$$z = \frac{\ln(R_{epi}) - \mu_Z}{\sigma_Z} \qquad \text{(ES.6)}$$

and the parameters of this expression are given by:

$$\mu_Z = \beta_3 + \beta_4(M-6.75) + \beta_5(M-6.75)^2 \qquad \text{(ES.7)}$$

$$\sigma_Z = \beta_6 \qquad \text{(ES.8)}$$

The coefficients of Eqs.(ES.2-8) are presented in Table ES.5.

Table ES.5. Coefficients of Eq.(ES.2-8) for the standard deviations of the predictions

	PGA	**Sa(0.2s)**	**Sa(0.5s)**	**Sa(1.0s)**	**Sa(2.0s)**
ϕ_{SM}	0.4918	0.4454	0.5146	0.4081	0.4133
β_1	0.20380	0.20284	0.20761	0.21116	0.21290
β_2	0.073419	0.080624	0.044808	0.018152	0.005130
β_3	3.39511	3.39511	3.39511	3.39511	3.39511
β_4	0.70978	0.70978	0.70978	0.70978	0.70978
β_5	0.0090045	0.0090045	0.0090045	0.0090045	0.0090045
β_6	1.03275	1.03275	1.03275	1.03275	1.03275

If sigma values are required at other periods, these may be obtained using the following:

- For periods up to 0.04 s, use the value at 0.01 s (*i.e*., for PGA)
- For longer periods, use linear interpolation against the logarithm of period
- For periods longer than 2 seconds, the sigma value is kept constant at the 2-second value

In order to enable more accurate estimation of the response spectral shapes, GMPEs only for median values of spectral accelerations at a larger number of oscillator periods have been derived using the same functional form of Eq.(ES.1). The coefficients are reported in Table ES.6; the coefficients c_5 and c_6 take the same values at all periods as those reported in Table ES.1. Inspection of normalised response spectra indicates that the spectral shapes corresponding to the two alternative (*i.e*., higher and lower) models are very similar.

Table ES.6. Coefficients of Eq.(ES.1) for the median accelerations of the central model at additional response periods for defining spectral shapes

Period (s)	c_1	c_2	c_3	c_{3a}	c_4
0.02	1.0630	1.2860	-0.3389	-0.1381	-1.4948
0.05	1.2086	1.2968	-0.2759	-0.1421	-1.5577
0.075	1.7287	1.2519	-0.2330	-0.1242	-1.6343
0.1	2.1158	1.1958	-0.2560	-0.1024	-1.6335
0.15	2.6975	1.1084	-0.3466	-0.0696	-1.5798
0.24	2.3014	1.1538	-0.4814	-0.0875	-1.4208
0.3	1.8760	1.2336	-0.5290	-0.1183	-1.3503
0.34	1.5623	1.2961	-0.5461	-0.1417	-1.3147
0.4	0.9941	1.3995	-0.5547	-0.1793	-1.2729
0.44	0.5759	1.4703	-0.5525	-0.2043	-1.2514
0.55	-0.6055	1.6561	-0.5276	-0.2665	-1.2116
0.6	-1.1450	1.7321	-0.5100	-0.2898	-1.2015
0.65	-1.6748	1.8049	-0.4895	-0.3109	-1.1946
0.7	-2.1667	1.8768	-0.4684	-0.3312	-1.1886
0.75	-2.6101	1.9457	-0.4481	-0.3504	-1.1829
0.8	-3.0183	2.0102	-0.4283	-0.3678	-1.1780
0.85	-3.3974	2.0707	-0.4089	-0.3833	-1.1737
0.9	-3.7500	2.1273	-0.3903	-0.3973	-1.1699
1.2	-5.4457	2.3918	-0.2925	-0.4494	-1.1569
1.4	-6.2732	2.5119	-0.2418	-0.4615	-1.1532
1.5	-6.6178	2.5591	-0.2206	-0.4628	-1.1521

If estimates of the spectral ordinates beyond 2 seconds are required, the ordinates at the longer periods should be obtained as follows, in all cases the starting point to transform the acceleration response spectrum to displacements via the pseudo-spectral relations:

- For **M** < 5.35, maintain the spectral displacement at T = 2 seconds constant
- For **M** ≥ 5.35, extrapolate linearly the displacement spectrum between periods of 1.5 and 2.0 seconds up to a period T_M, after which the displacement should be maintained constant; the period T_M is given by the following equation:

$$\log_{10}(T_M) = 0.5M - 2.3764 \quad \text{(ES.9)}$$

For the prediction of 5-75% Arias intensity significant durations, D_S, it is recommended to use the following implementation of the equation of Kempton & Stewart (2006):

$$\ln(D_S) = \ln\left[\frac{\left(\frac{\exp(6.02)}{10^{1.5M+16.05}}\right)^{\frac{-1}{3}}}{1.568x10^{7}} + 0.07\sqrt{R_{epi}^2 + 9} + 2.02\right] \quad \text{(ES.10)}$$

The total standard deviation is 0.52, with an inter-event component of 0.32 and an intra-event standard deviation of 0.42. Correlation coefficients for the predictions of these durations and of the spectral accelerations are as summarised in Table ES.7; if values at other response periods are required, it is recommended to use those presented by Bradley (2011).

Table ES.7. Correlation coefficients of the predictions of Sa and D_S

	PGA	**Sa(0.2s)**	**Sa(0.5s)**	**Sa(1.0s)**	**Sa(2.0s)**
ρ	-0.4690	-0.3159	-0.1183	-0.1023	-0.1550

Finally, a second version of the central GMPE has been generated using R_{JB} rather than R_{epi} as the distance metric. These alternative equations have been derived solely for the purposes of conducting sensitivity analyses to explore the impact on calculated hazard and risk of modelling earthquake sources as points (hypocentres) or as extended fault ruptures. The functional form of the equation is identical to that in Eq.(ES.1) and the coefficients for the median values are as listed in Table ES.8. There is no within-event sigma correction for this model; the within- and between-event components of the variability are also listed in Table ES.8.

Table ES.8. Coefficients of Eq.(ES.1) and the inter-event variability of the central R_{JB} GMPE

	PGA	**Sa(0.2s)**	**Sa(0.5s)**	**Sa(1.0s)**	**Sa(2.0s)**
c_1	1.1563	2.4972	-0.0684	-4.3882	-7.8093
c_2	1.2732	1.1216	1.5742	2.2288	2.6929
c_3	-0.3394	-0.4314	-0.5416	-0.3549	-0.1520
c_{3a}	-0.1342	-0.0747	-0.2397	-0.4202	-0.4370
c_4	-1.5048	-1.4806	-1.2266	-1.1640	-1.1526
c_5	0.4233	0.4233	0.4233	0.4233	0.4233
c_6	-0.6083	-0.6083	-0.6083	-0.6083	-0.6083
τ	0.2810	0.3337	0.3216	0.3789	0.3547
ϕ_{SM}	0.4918	0.4454	0.5146	0.4081	0.4133
$\delta\phi$	0.0	0.0	0.0	0.0	0.0

The stochastic simulations for the R_{JB} model used a different, and much denser, configuration of sources and receivers than thiose used for the derivation of the main R_{epi} model. In order to ensure that the comparisons are conducted using genuinely compatible and consistent equations based on point- and extended-source models, new R_{epi} equations were derived based on the same simulations. The coefficients for these equations—which do not override or replace the V1 models for the main hazard and risk calculations—are summarised in Table ES.9; in all other respects, these equations are the same as the central model presented earlier.

Table ES.9. Coefficients of Eq.(ES.1) and the inter-event variability of alternative R_{epi} GMPE

	PGA	**Sa(0.2s)**	**Sa(0.5s)**	**Sa(1.0s)**	**Sa(2.0s)**
c_1	1.3001	2.9388	0.4122	-4.1684	-7.8723
c_2	1.2988	1.1158	1.5157	2.2027	2.7270
c_3	-0.3790	-0.4200	-0.5752	-0.3933	-0.1531
c_{3a}	-0.1289	-0.0630	-0.1997	-0.3910	-0.4412
c_4	-1.5651	-1.6064	-1.2912	-1.1886	-1.1730
c_5	0.4233	0.4233	0.4233	0.4233	0.4233
c_6	-0.6083	-0.6083	-0.6083	-0.6083	-0.6083
τ	0.2810	0.3337	0.3216	0.3789	0.3547
ϕ_{SM}	0.4918	0.4454	0.5146	0.4081	0.4133
$\delta\phi$	Eq.(ES.3)	Eq.(ES.3)	Eq.(ES.3)	Eq.(ES.3)	Eq.(ES.3)

All of the coefficients for the models to predict response spectral accelerations are provided in the Excel file **Version 1 GMPE coefficients_21 June 2015.xlsx**, in which the sheets have the following correspondence to tables in this Executive Summary:

- Central: Table ES.2
- Higher: Table ES.3
- Lower: Table ES.4
- Delta_Phi: Table ES.5
- Shape: Table ES.6
- Central_RJB: Table ES.8
- Central_Repi: Table ES.9

Acknowledgements

Thanks to Jan van Elk who, with support from Dirk Doornhof, has provided the leadership and guidance for the work presented herein. Jan provided the conditions that enabled this study to be carried out in a timely fashion.

We are grateful to Professor Sinan Akkar and Dr Abdullah Sandıkkaya for provision of the updated European strong-motion database and the additional datasets of recordings from small-magnitude tectonic earthquakes in France, Italy and Switzerland. As well as compiling these valuable databases, Professor Akkar and Dr Sandıkkaya kindly responded to a number of questions regarding the derivation of the metadata. In addition to these important contributions, Professor Akkar participated in discussions held in London regarding the application of these data to the derivation of the Groningen ground-motion model and issues related to the functional form and independent variables in these equations, and made many contributions to these exchanges.

Drs Helen Crowley and Damian Grant, who led the development of the V1 fragility functions for buildings in the Groningen field, were particularly helpful in specifying the requirements of the ground-motion models in relation to the definition of the fragility functions and their application in the risk analyses. Dr Crowley in particular has contributed greatly through discussions—which are ongoing as we move towards V2—regarding the implementation of the GMPEs.

The authors are also indebted to numerous other individuals for constructive comments and probing questions at different stages of the GMPE development, including in particular Dr Dirk Kraaijpoel from KNMI, who offered many pertinent observations and suggestions at workshops and meetings related to the GMPE project, and to Drs Steve Oates and Stephen Bourne at Shell and also Dr Rui Pinho. Comments on the Version 0 GMPEs—in the context of a summary paper on the hazard model presented in the 2014 Winningsplan—by Dr John Douglas were also useful in bringing to our attention issues and ideas that have proved helpful in the development of this report.

This work, both for the development of the V1 equations and ideas regarding the ongoing development, has been enriched and enhanced by discussions with Dr Pauline Kruiver from Deltares, Professor Adrian Rodriguez-Marek of Virginia Tech and many of those listed above. The work has also benefited from feedback and discussions at Workshops held with ExxonMobil, the Scientific Advisory Committee (SAC) for Groningen, and the Dutch State Supervision of Mines (SodM).

Thanks are also due to Rein de Vries of ARUP and to Dr Barbara Polidoro for spotting typographical errors in an earlier draft.

1. Introduction

For the 2013 Winningsplan, ground-motion prediction equations (GMPEs) were developed for the prediction of peak ground acceleration (PGA) and peak ground velocity (PGV) as a result of induced earthquakes in the Groningen field. The equations were modified versions of the GMPEs derived using strong-motion data from Europe, the Mediterranean and the Middle East by Akkar *et al*. (2014a). The equations using hypocentral distance, R_{hyp}, were selected, and applied with an assumed field-wide 30-metre shear-wave velocity, V_{S30}, of 200 m/s and the assumption of normal faulting. The coefficients of the equations were modified below a certain magnitude—**M**4.2 for PGA, **M**3.8 for PGV—to fit the peak motions from 40 accelerograms obtained from 8 earthquakes by the KNMI network. The aleatory variability for the small-magnitude extension was assumed to be the same as that associated with the original equations. These preliminary GMPEs are described as part of the 2013 hazard model in Bourne *et al*. (2015).

For the Version 0 hazard and risk model derived in mid-2014, an additional 14 records were available from the **M**3.0 Leermens earthquake of February 2014. A very simple residual analysis suggested that the additional data did not warrant a modification of the 2013 GMPE, for which reason it was decided to retain those PGA and PGV equations for the Version 0 hazard and risk models (Bommer & Dost, 2014). The residual analyses did show, however, that the models did not fit the data well at short epicentral distances, which was concluded to be a consequence of the functional form of the Akkar *et al*. (2014a) equation and specifically the use of a fixed value for the near-source saturation term at all magnitudes. The addition of the Leermens records expanded the available dataset but not sufficiently to allow direct calculation of the aleatory variability.

This document describes the derivation of the GMPEs used in the Version 1 hazard and risk models. There are several fundamental differences with the Version 0 model, the first of these being the derivation of equations for response spectral ordinates at a number of oscillator periods in addition to PGA. The second major difference is that a continuous functional form is adopted across the full magnitude range rather than having an abrupt transition between the range of the Groningen database and the larger events considered in the hazard and risk calculations. Thirdly, to the extent that the data allow, the Version 1 models also include direct estimations of the sigma values (logarithmic standard deviations representing the aleatory variability in the predictions). Finally, rather than producing a single GMPE for each selected ground-motion parameter, a suite of alternative models is derived that capture both the current best estimate and the associated epistemic uncertainty. Additionally, a preliminary model for the prediction of strong-motion durations in the Groningen field has been included, together with a correlation model for the joint prediction of spectral accelerations and durations.

Following this brief introduction, Chapter 2 discussed the dependent and independent variables included in the V1 GMPEs. The strong-motion database for the Groningen field is presented in Chapter 3, together with an overview of existing and planned recording networks that are expected to provide a growing body of data. For exploratory purposes,

the existing database is used to calculate residuals with respect to some current GMPEs. Chapter 4 provides an overview of the database of recordings from tectonic earthquakes in Europe compiled for the purposes of this study, and the selection of a subset of these data used to check and evaluate the extrapolation of the GMPEs beyond the magnitude range covered by the current Groningen database. Chapter 5 describes two parallel studies conducted to develop the basic building blocks for the GMPEs: exploration of a suitable functional form and stochastic simulations using source, path and site parameters estimated from inversion of the Fourier spectra of the Groningen recordings; the inversions are fully documented in Appendix I. Chapter 6 then documents the derivation of the parametric equations including the sigma model, which for the GMPEs based on epicentral distance includes an adjustment for the use of the point-source approximation at larger magnitudes; the derivation of the adjustment is described in detail in Appendix II. In Chapter 7, the derivation of a very preliminary model for the estimation of duration of shaking for the Groningen earthquakes is presented, together with a model for the correlation of residual of spectral acceleration and duration to enable the joint prediction of these parameters.

An important point to stress here is that the V1 model is only a snapshot of the ongoing development of the hazard and risk model being developed by NAM for 2016 Winningsplan. Since periodic reporting is required, not only of ongoing progress in the development of the various elements of the hazard and risk models but also of the actual risk and hazard estimates, the V1 GMPEs have been used in the current hazard and risk estimates. As is pointed out in Chapter 6, this application needs to viewed bearing in mind that while there are several important advances embodied within the V1 equations, by virtue of being simply the current status of an ongoing process of building a sophisticated model—with field-specific data that are being gathered as the work progresses—some elements of the equation are not yet at a stage to be considered reliable. There is a serious shortcoming in the V1 GMPE in terms of the modelling of site response effects: the equations only represent the generic response of some average site conditions over the field (more precisely, over the recording network sites from which the accelerograms have been obtained) and, more importantly, the site response is modelled linearly. While this is an acceptable assumption for the small-magnitude earthquakes in the Groningen database, it will inevitably mean that the current predictions are likely to overestimate the ground motions from larger earthquakes at short source-to-site distances.

With this final point in mind, possibly the most important part of this report is Chapter 8, in which an overview is presented of the ongoing work to continue the development of the GMPEs towards the final hazard and risk models for the Winningsplan. Critical amongst these ongoing phases of development is inclusion of actual site conditions at each location across the field and the modelling of the non-linear response of these profiles to earthquake motions.

2. Dependent and Independent Variables

This Chapter discusses the parameters that will appear in the Version 1 equations, including the specific definition adopted for each variable.

2.1. Horizontal spectral accelerations

The GMPEs will predict values of the horizontal pseudo-spectral response acceleration for oscillators with an assumed equivalent viscous damping of 5% of critical, in addition to PGA (which is equivalent to the spectral acceleration at zero period). An equation for PGV may be derived but this would only serve for comparison with earlier hazard maps since it is not envisaged that the fragility functions will be derived in terms of this parameter.

GMPEs will ultimately be required at a large number of oscillator periods, primarily to enable the definition of complete response spectra for defining input to structural analyses. However, for the actual risk calculations, equations are only required for spectral accelerations at those response periods selected as being representative of the building types classified in the exposure database. The initial oscillator periods selected by those developing the fragility functions are 0.0 (PGA), 0.2, 0.5, 1.0 and 2.0 seconds.

Since each accelerogram has two horizontal components, a decision also needs to be made with regards to how to treat the spectral values from the two orthogonal components. In recent years it is has become standard practice to use the geometric mean of the two horizontal components, or one of the subtle variations of this definition proposed by Boore *et al*. (2006), as used in the NGA-West project to derive GMPEs for application to California and other active regions of shallow crustal seismicity (Abrahamson *et al*., 2008). The NGA-West2 equations (Gregor *et al.,* 2014) have used a slightly different definition of the horizontal component, but the resulting values have been shown to be almost identical to geometric mean values at short periods and only very slightly larger at longer periods (Boore, 2010). For this project, the geometric mean definition is adopted, not least because of the fact that for median motions it gives identical results to using both horizontal components as independent data points (Beyer & Bommer, 2006). This feature is attractive since GMPEs are also required for the duration of the ground motion, for which the use of both components may be more appropriate, as argued by Bommer *et al*. (2009). Moreover, joint (vector) predictions of accelerations and durations are required, and the Groningen data shows a very strong inverse relationship between PGA and duration when the data are examined as individual components.

Although the geometric mean component definition is adopted for the V1 GMPEs, part of the ongoing development work will be to produce a Groningen-specific model for the distribution of component-to-component ratios of the spectral ordinates, as explained in Section 8.7.

2.2. Earthquake magnitude

All GMPEs include magnitude as the basic measure of earthquake size and nowadays the most widely used scale is moment magnitude, **M**. This is the magnitude scale used, for example, in the NGA-West and NGA-West2 equations and all of the most recent European models (Douglas *et al*., 2014), amongst many others. Moment magnitudes are not, however, routinely calculated for small earthquakes and most national seismograph networks report some variation of the local magnitude, M_L, for smaller events. KNMI is no exception to this practice, although moment magnitudes are calculated for some events. Figure 2.1 shows the magnitude pairs for the 8 earthquakes used in the derivation of the 2013 models, which support the claim by KNMI that for magnitudes above a certain level (~3), the two scales may be considered equivalent.

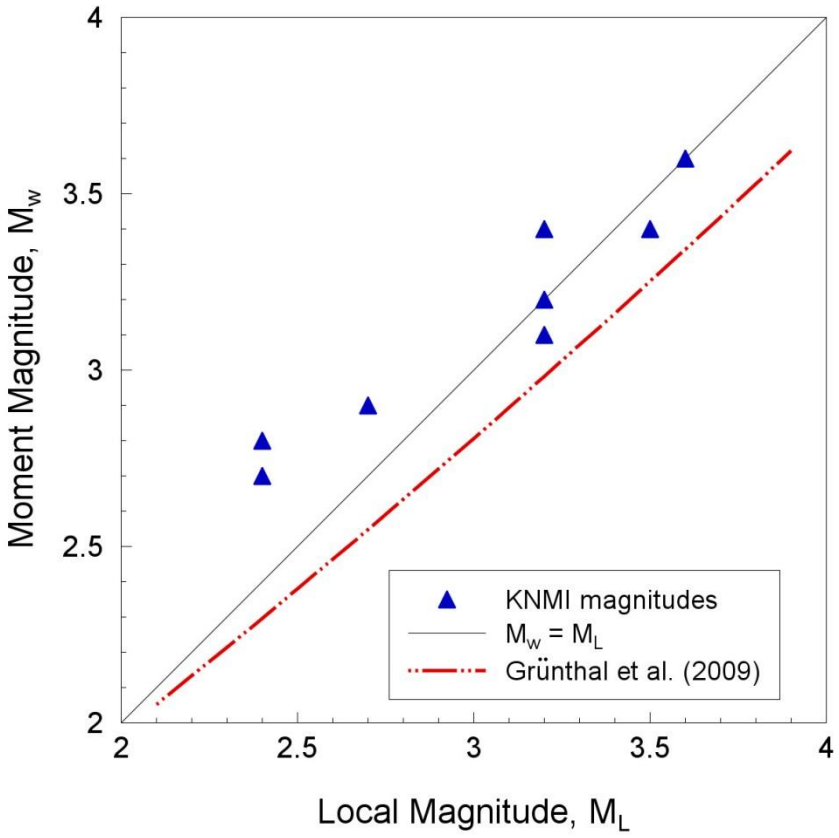


Figure 2.1. Comparison of local and moment magnitudes from KNMI for earthquakes in the Groningen field. Also shown is the widely-used European relationship between these two scales by Grünthal *et al*. (2009), which does not fit the data but does show the same general trend.

For the derivation of the Version 1 GMPEs, the assumption of equivalence in the two scales is invoked. In view of how the equations are developed, this is not a particularly critical issue since the assumed magnitudes are internally consistent and also consistent with those used in the earthquake catalogue, which makes the same assumption of equivalence. The nature of the relationship between moment and local magnitude in Groningen does have implications for the assumption of linear Gutenberg-Richter (G-R)

recurrence relationships, but in the derivation of the GMPEs it is only of significance when making comparisons to other GMPEs and in using the European dataset to explore the behaviour of the equations at larger magnitudes.

However, the issue of the moment magnitudes is noted does require greater attention and definitive documentation as part of the ongoing work. KNMI is continuing to work on this issue and a separate document on the both global and Groningen-specific relationships between these two magnitude scales is being produced.

2.3. Style-of-faulting

For at least a decade now, empirical GMPEs for active crustal regions have generally included a term for the influence of the style-of-faulting, except for the few cases where a single rupture mechanism dominates in the region covered by the equation. The pattern often observed is that reverse-faulting earthquakes produce the strongest motions, with normal-faulting earthquakes producing motions that may be slightly lower, on average, from those produced by strike-slip ruptures (*e.g.*, Bommer *et al*., 2003). This pattern has not been fully explained physically, and it may be that these apparent trends are the result of other influences not included in simple models. In more complex models, such as the NGA-West equations, the inclusion of the depth-to-top-of-rupture parameter, for example, modifies the apparent dependence on style-of-faulting (Abrahamson *et al*., 2008). Figure 2.2 shows the ratios of predicted ordinates from reverse and strike-slip faulting events to those from normal-faulting ruptures.

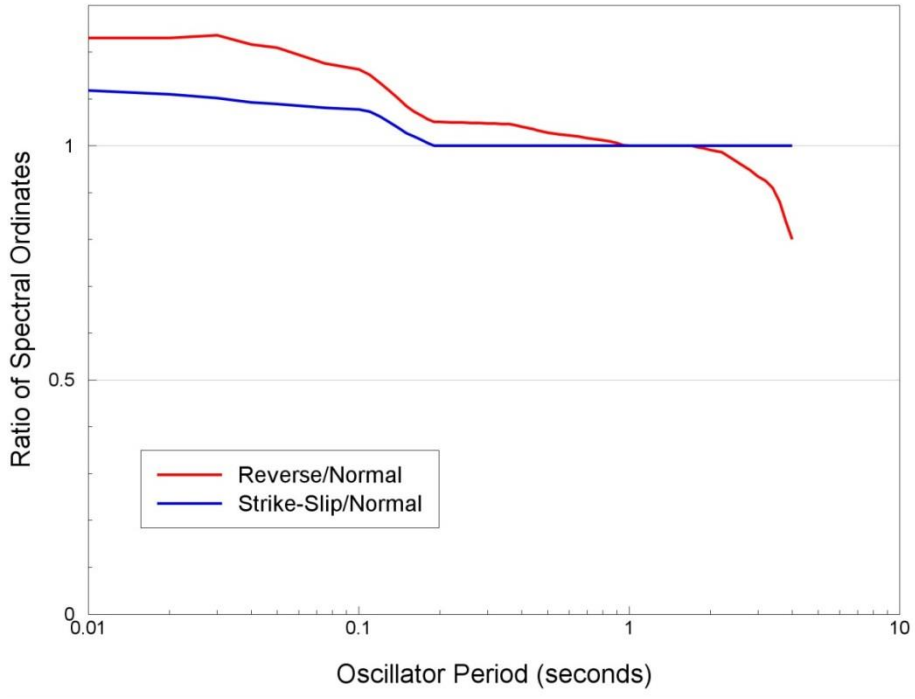


Figure 2.2. Ratios of spectral ordinates for different styles-of-faulting from the GMPEs of Akkar *et al*. (2014a), relative to the normal-faulting case

The earthquakes in the Groningen field have been found to have normal or strike-slip rupture mechanisms, or to be oblique combinations of these two faulting styles. However, the mechanisms are not known for most of the earthquakes in the ground-motion database for the field (Chapter 3) and there is no basis to assign, with any confidence, proportions of future events to the two styles-of-faulting. In addition to these considerations, it is not clear how much influence the style-of-faulting has for small-magnitude earthquakes in which even at relatively short distances the point-source approximation is reasonable.

Since it is not possible to model the specific influence of the style-of-faulting for the recorded motions from the Groningen field, and since the predicted amplitudes from the two mechanisms expected in the field are similar (Figure 2.2), it is decided to produce an equation without an explicit term for the style-of-faulting. This equation would be assumed to represent an average behaviour between normal and strike-slip faulting. Records from the European database (Chapter 4) of earthquakes with normal and strike-slip ruptures will be used without adjustment, whereas those from reverse-faulting earthquakes will be adjusted to an average of normal and strike-slip amplitudes obtained setting the factor F_N in Akkar *et al*. (2014a) to a value of 0.5; this is not the exact solution, but it is a very close approximation. For the selected periods for which the preliminary GMPEs are being developed (Section 2.1), the factors that will be applied to ordinates from reverse-faulting earthquakes are given in Table 2.1.

Table 2.1. Adjustment factors to transform accelerations from reverse-faulting earthquakes to the average of those from normal and strike-slip events, using the coefficients of Akkar *et al*. (2014a)

Period (s)	Adjustment Factor
0.0 (PGA)	0.862
0.01	0.860
0.2	0.952
0.5	0.973
1.0	1.000
2.0	1.009

2.4. Source-to-site distance

For the 2013 GMPEs, hypocentral distance (R_{hyp}) was selected as the distance metric in order to capture, to some extent, the influence of the shallow depth of the Groningen earthquakes, which are assumed to be occurring within the gas reservoir at a depth of about 3 km. This option can only be considered an approximate representation of the shallow depth since only a small proportion of the dataset used by Akkar *et al*. (2014a) is from very shallow earthquakes. In view of this, the use of R_{hyp}–based GMPEs derived from recordings of tectonic earthquakes may to some degree be a conservative option since it does not account for the possibly lower stress drops from very shallow earthquakes.

For the Version 1 GMPEs, the choice is made to move to the use of epicentral distance, R_{epi}. There are two factors motivating this choice. The first is that the Groningen events

may be assumed to all occur at a single focal depth, since the thickness of the reservoir is small compared to its depth below the surface. Therefore, the distinction between R_{epi} and R_{hyp} is unimportant because the depth can be accounted for with either metric. The second motivation is that part of the ongoing scope of work in this area is to develop GMPEs using a distance metric defined relative to extended fault ruptures, in order to perform sensitivity analyses to explore the impact on the calculated risk of locating larger (**M** > 5) events on faults, mapped or otherwise. For extended fault ruptures, the most commonly used metrics are the distance to the closest point on the fault rupture, R_{rup}, and the horizontal distance to the project of the fault rupture on the Earth's surface, R_{JB}. The latter is referred to as the Joyner-Boore distance, having first been proposed by Joyner & Boore (1981). Given the feature of approximately constant focal depths and the advantages of using R_{epi} noted above, the logical choice for an extended distance metric here is R_{JB} (given that the only objective is to build a model for Groningen and the transportability to other applications is not relevant). The metric R_{JB} will therefore be adopted for the extended fault version of the GMPEs—with an appropriate adjustment for the constant, shallow depth of the earthquakes—and thus adjusting the point-source model to the extended-source version will be more straightforward given the use of R_{epi} for the former. For the larger magnitude events, the use of R_{epi} does not really imply the loss of the capacity to account for the shallow focal depths since these are only being accounted for in a very crude way with the R_{hyp} metric. Additional constraints, combined with alternative models to represent the unavoidable epistemic uncertainty, will be invoked to model the influence of the shallow depth of these potential earthquakes.

A secondary advantage of using R_{JB} rather than R_{rup} relates to the use of the European strong-motion database for evaluating the extrapolations of the GMPEs to larger magnitudes. For the records in the European database, R_{JB} is known for nearly most records but R_{rup} only for 64% of the same records (Akkar *et al*., 2014b) and therefore the options for using the data in this way are improved if R_{JB} is the selected distance metric.

2.5. Site classification

The Akkar *et al*. (2014a) GMPEs represent site effects using the time-averaged shear-wave velocity over the uppermost 30 m at the site, V_{S30}. The site response is modelled to account for non-linear soil response, using the empirical functions derived by Sandıkkaya *et al.* (2013), as a function of V_{S30} and the predicted median PGA in reference rock, characterised by a V_{S30} of 750 m/s. The site amplification functions, in terms of amplitude, dependency on frequency, and the degree of non-linearity, represent any average model for all of the records and sites in the database used in the Sandıkkaya *et al.* (2013) study.

A very clear and important objective of the work plan to develop GMPEs for the Groningen field is to incorporate the effect of the dynamic response of the near-surface layers in the field on the motions at ground level. At the time of developing the Version 1 GMPEs, only very preliminary V_{S30} data were available for the field, and a great deal of inference and extrapolation have been invoked to obtain maps such as that presented in Figure 2.3. At

the same time, Deltares are beginning an ambitious programme of work to obtain V_S data for the Groningen field, including site-specific measurements at all of the accelerograph stations. Combined with the information obtained by the re-analysis of NAM's seismic reflection data from the reservoir imaging, all this information will allow not only to construct a detailed model of the velocity profiles across the field, but also to perform site response analyses to quantify the amplification factors for upcoming seismic waves (see Section 8.4).

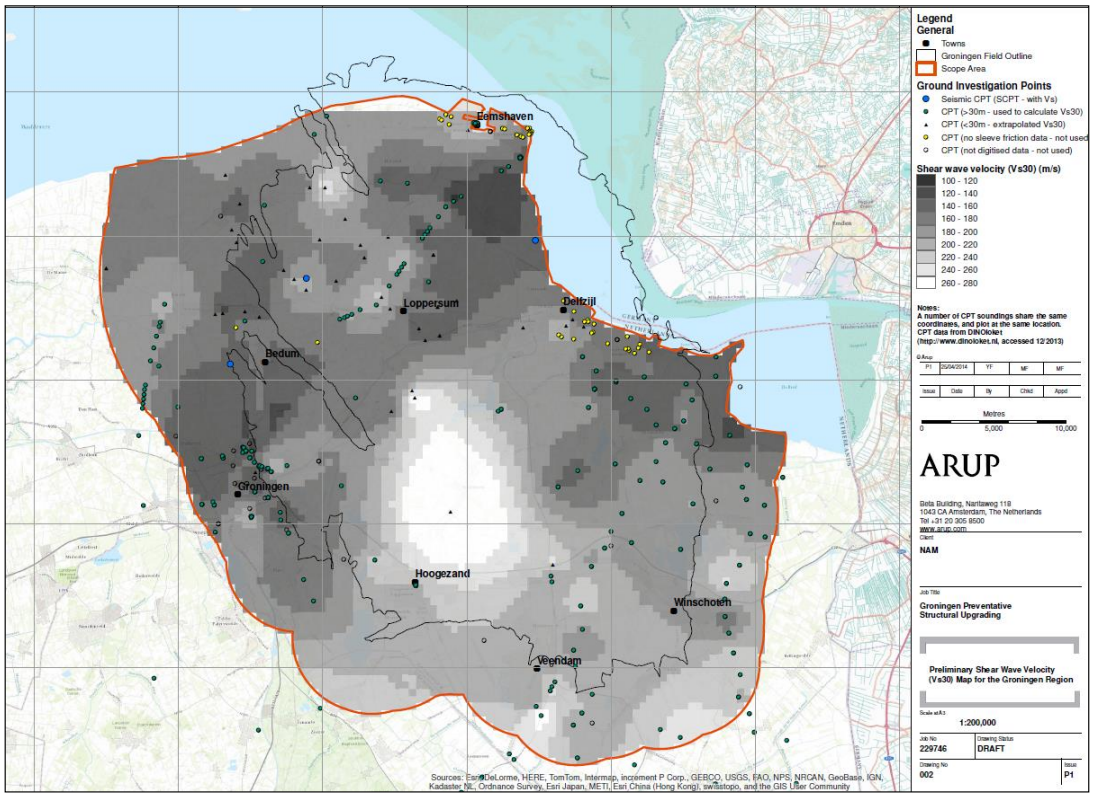


Figure 2.3. Preliminary map of V_{S30} values across the field compiled by ARUP, showing locations from which data was obtained.

Since two sites with different layering in the top 30 metres can have the same V_{S30}, their dynamic response can also be very different, and this is even more the case if there are also appreciable differences in the velocities profiles at depth below 30 metres (Figure 2.4). Incorporating V_{S30} into the GMPE for the Groningen field would bring the maximum benefit when combined with site-specific amplification functions, rather than generic factors such as those of Sandıkkaya *et al.* (2013).

For these reasons, it is not considered worthwhile incorporating V_{S30} into the GMPEs until such time that there are more reliable data on this parameter for both the recording stations and the field in general, and until field-specific site amplification factors have been derived. One could argue that even with a preliminary mapping of V_{S30} across the field, the explicit inclusion of site response effects would be useful to distinguish the ground shaking at softer locations. However, the range of V_{S30} across the field does not seem to vary

greatly, and even taking the most extreme values indicated in Figure 2.3, the differences in the predicted response spectra relative to the ordinates at a site with V_{S30} of 200 m/s are modest (Figure 2.5). The differences between the various curves plotted in Figure 2.5 may well be smaller than the differences between the generic amplification functions shown here and those that correspond to the specific conditions of the Groningen field.

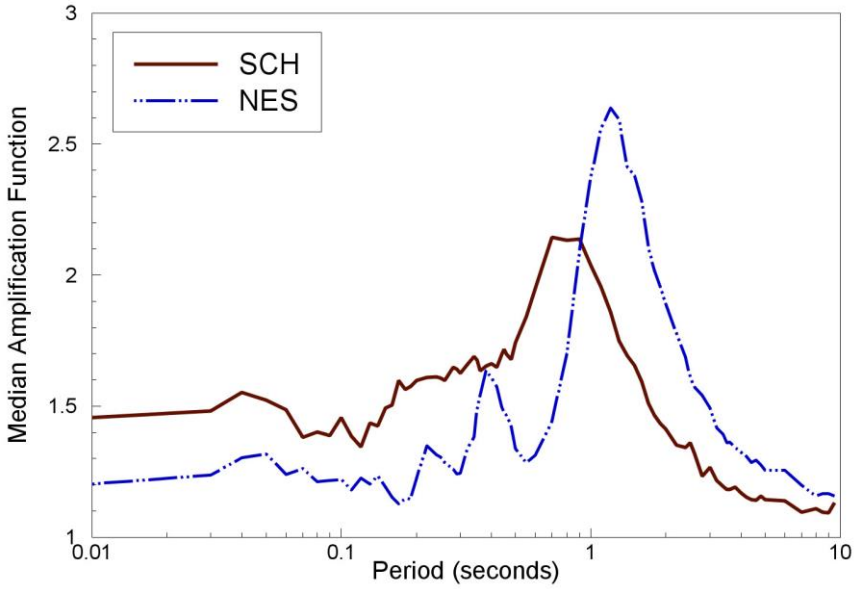


Figure 2.4. Median site amplification functions for the SCH and NES sites, which have V_{S30} values of 280 and 284 m/s, respectively (Papaspiliou *et al*., 2012).

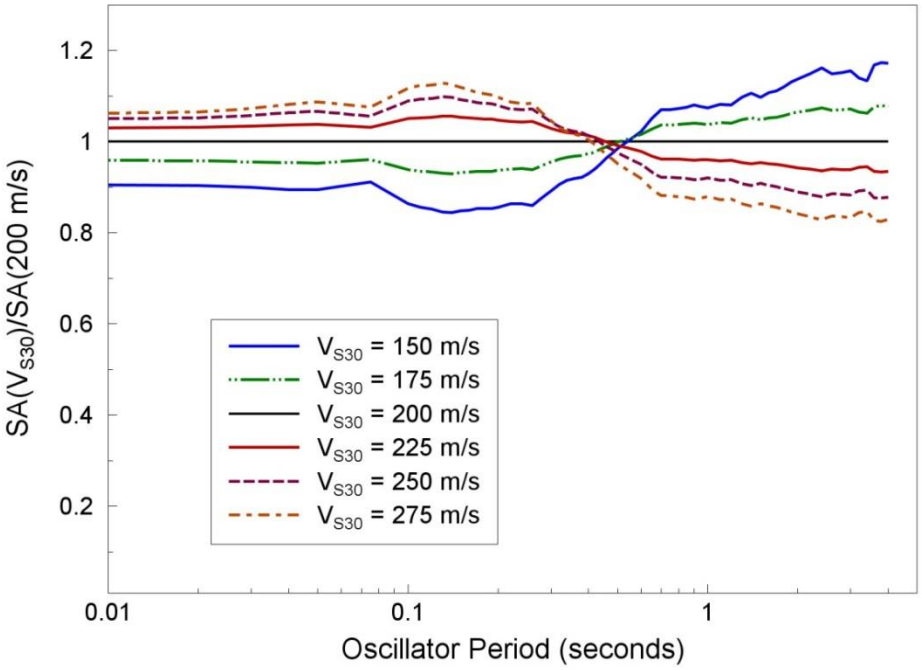


Figure 2.5. Ratios of median spectral ordinates for an **M** 4.5 normal-faulting earthquake at a hypocentral distance of R_{hyp} 5 km for sites with different V_{S30} values to those at a site with V_{S30} 200 m/s, obtained with the Akkar *et al*. (2014a) GMPEs

For the Version 1 GMPEs, therefore, it is decided to develop the equations for a single constant value of V_{S30}, as was done for the Version 0 model. The Version 1 GMPEs will still embody a number of important advances with regard to the earlier models, but the data available at the present time for incorporating site response is insufficient to warrant its inclusion. This decision does simplify the development of the Version 1 GMPEs, allowing other features to be incorporated more easily. The key focus in the development of the Version 2 equations will be the inclusion of Groningen-specific ground conditions and site response characteristics.

The final question to be addressed at this stage is the choice of the representative value of V_{S30} for the Version 1 GMPEs. The limited information on measured V_S values from the field indicates that 200 m/s is probably still a suitable representative V_{S30}. Selecting a lower value could be more safely conservative (although this would result in stronger non-linearity, which would reduce the higher values of acceleration) but this would also mean adopting model that was perhaps less well constrained by virtue of the distribution of the Sandıkkaya *et al.* (2013) dataset (Figure 2.6).

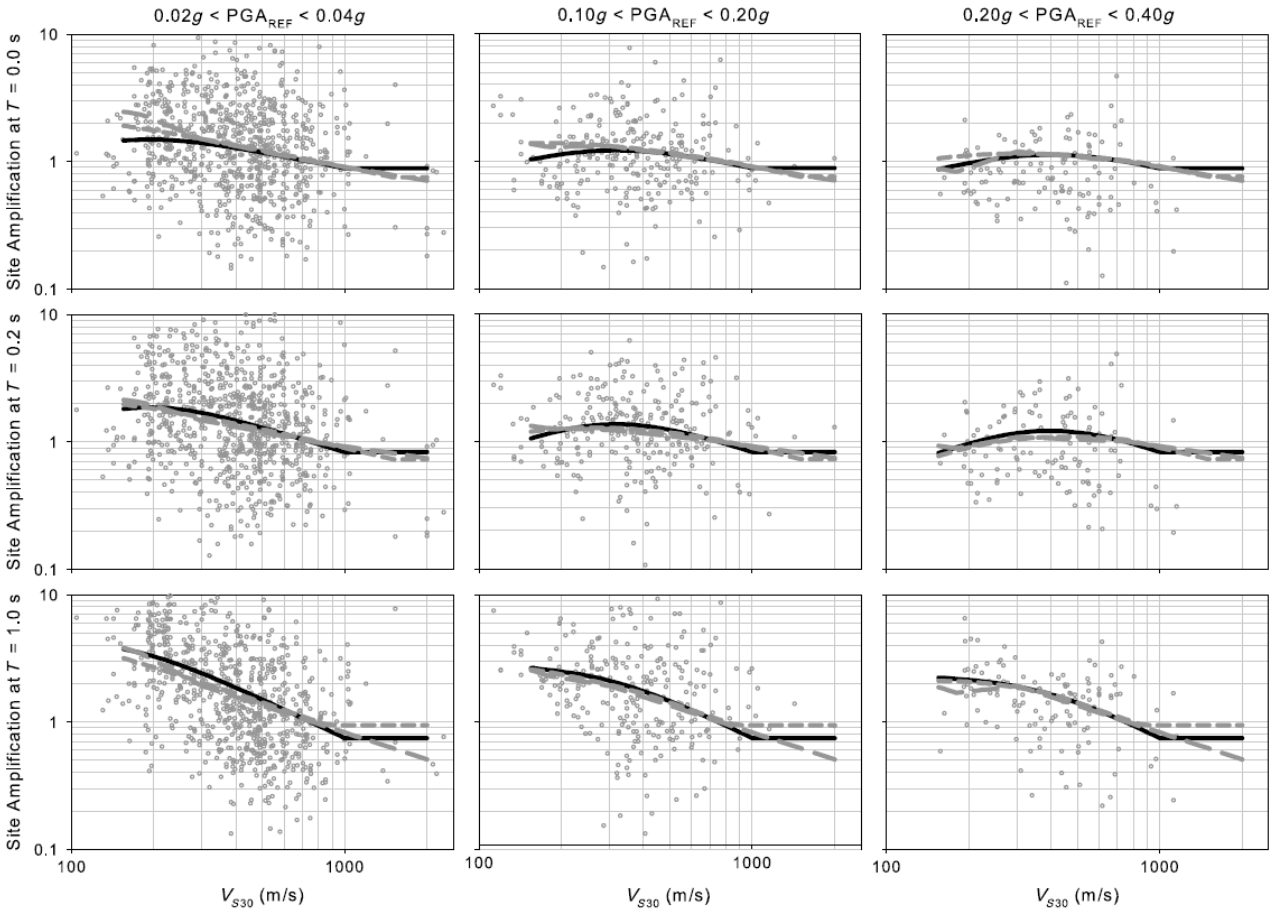


Figure 2.6. Site amplification functions (*black lines*) as a function of V_{S30} for different response periods and ranges of reference rock PGA from Sandıkkaya *et al*. (2013), together with the empirical data points. Note how sparse are the data for V_{S30} values <200 m/s.

In light of these considerations, the decision has been taken to assume a field-wide value of 200 m/s for V_{S30} and derive the Version 1 GMPEs conditioned to this value. For the purpose of making comparisons between the model predictions and the European strong-motion records, the procedure will be then to use the site amplification functions of Sandıkkaya *et al.* (2013)—but with the equation of Akkar *et al.* (2014a) for the prediction of PGA_{REF}, the value of peak ground acceleration in the reference rock condition—to transform each spectral ordinate to this assumed site condition. This will be done in two steps, effectively, first transforming the record from the V_{S30} at its recording station to the reference rock condition of 750 m/s, and then transforming the rock motion to the Groningen target of 200 m/s. This is discussed in more detailed in Section 4.3.

A final and important point to note is that the only objective of this work is to produce the best possible ground-motion prediction model for the Groningen field. There is no motivation to generate equations that could be transportable to any other location or region, even within the Netherlands. Therefore, the use of convenient but crude surrogate parameters such as V_{S30} is probably unnecessary, even if it is useful to have V_{S30} maps and classifications for comparative purposes. As explained in Section 8.4, the final GMPEs will probably be based on a zonation of the field with site amplification functions specified for each zone, obviating the need to employ simplified parameters such average shear-wave velocities or depths to horizons of a given shear-wave velocity as a measure of the sediment thickness.

3. Groningen Ground-Motion Database

The key to developing the GMPEs for the Groningen field are the recordings of ground shaking from induced earthquakes. This Chapter begins by providing an overview of the recording networks that are either in operation or being installed in the field, which will clearly provide a great deal of data over the coming months and years (leading, inevitably, to frequent updates and refinements of the GMPEs). The current database of recordings from the KNMI accelerograph network, which will be used as the basis for the derivation of the Version 1 GMPEs, is then briefly described.

3.1. Strong-motion networks in the Groningen field

KNMI has operated an accelerograph network in the Groningen field for several years. During period from late 2013 into late 2014, the network was expanded and upgraded. These station locations within the Groningen field are shown in Figure 3.1.

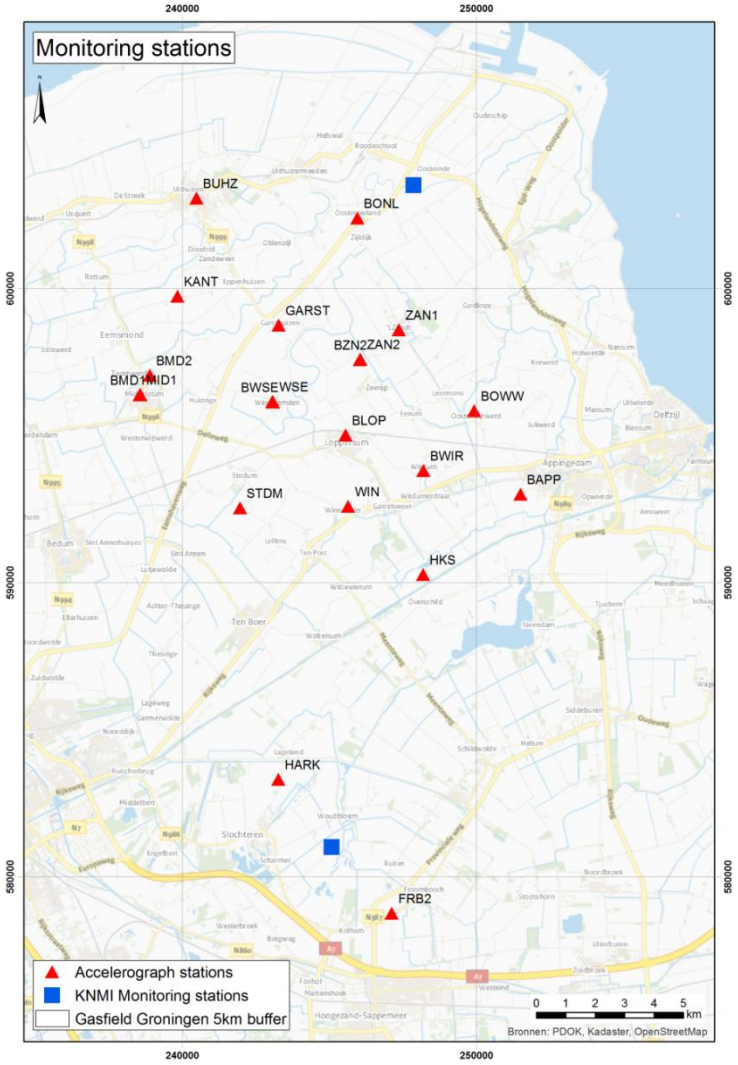


Figure 3.1. Locations of KNMI accelerographs (*red triangles*) in the Groningen field. The blue squares are two (SPY in the north, WDB in the south) of the five 200-m boreholes installed with geophones operated by KNMI within and around the field.

In addition to the accelerograph stations, KNMI has also operated since 1995 five boreholes of 200 m depth with geophones installed at 50 m intervals. Some of these also included a geophone at the ground surface but the surface instruments have not operated continuously for various reasons, and consequently there are relatively few events recorded both at the surface and within the boreholes. The geophone recordings from boreholes are considered to be of greater use for the study of site response characteristics than incorporated directly to the derivation of the GMPEs.

As part of the response to the Groningen earthquakes, NAM is installing two new deep boreholes with geophones inside the gas reservoir (at ~ 3 km depth) and 59 new 200-metre boreholes instrumented with geophones. The 59 geophone-instrumented boreholes, some of which have already been installed, will all be accompanied by an accelerograph at the surface, operated by KNMI. Once fully operational, these instruments (Figure 3.2) will lead to several tens of accelerograms being produced by all significant earthquakes in the field.

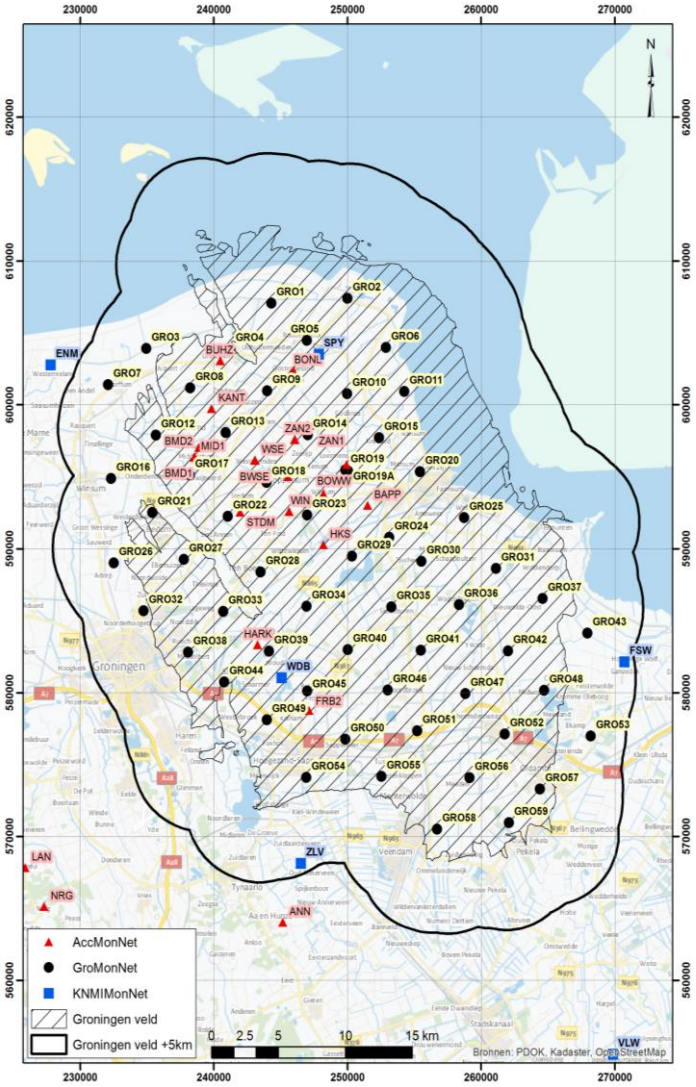


Figure 3.2. Locations of 59 instrumented boreholes and co-located accelerographs (*black circles*) being installed by NAM. Also shown are the locations of KNMI accelerographs (*red triangles*) and the five 200-m boreholes installed with geophones (*blue squares*).

The work on detailed site characterisation—in terms of V_S profiles—being conducted by Deltares is focusing primarily on the KNMI and NAM stations discussed above. These are expected to provide the main data that will be used in the ongoing development of the GMPEs used in the Groningen seismic hazard and risk modelling.

An additional 60 accelerographs have been installed (in clusters of three instruments at each of 20 locations) on the key facilities of the NAM gas production network in the field (Figure 3.3 and Table 3.1).

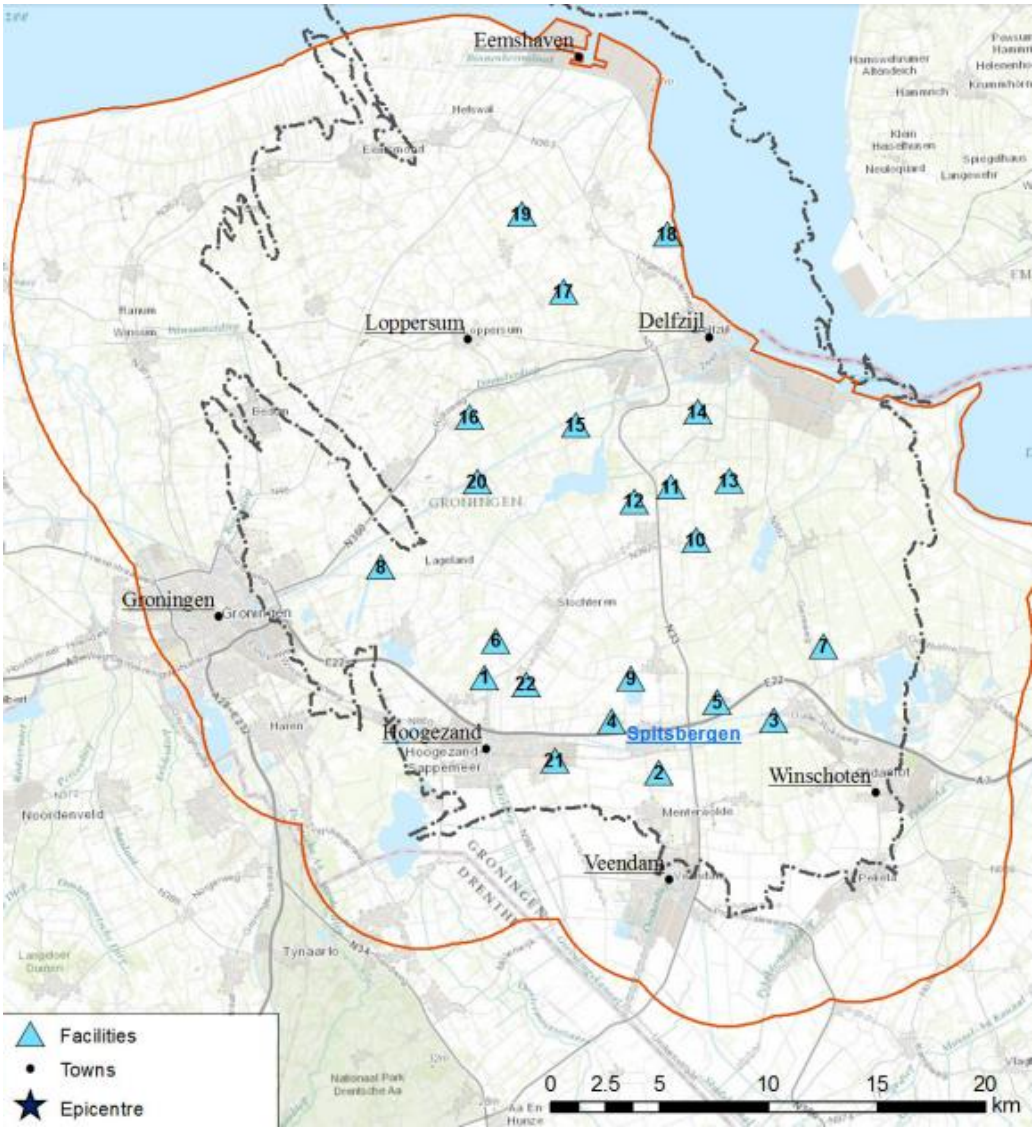


Figure 3.3. Locations of NAM facilities in the Groningen field (*blue triangles*) at which accelerographs are operating. The location names are listed in Table 3.1.

These instruments recorded two recent earthquakes from which recordings have not yet been added to the database from any of the networks. The first of these was the magnitude 2.8 Woudbloem event on 30 December 2014, which had epicentral coordinates of 53.208°N, 6.728°E, the second the magnitude 2.7 Wirdum earthquake of 6 January

2015, located at 53.324°N, 6.768°E. The purpose of the NAM accelerograph is to allow safe shut-down of the facilities if accelerations in excess of specific thresholds are exceeded, but the records obtained by these instruments will also be made available and are likely to be added to database used for derivation of the GMPEs, provided that reliable estimates of the site characteristics can also be obtained. The site characterisation programme being undertaken by Deltares may be extended to include these sites.

Table 3.1. Locations of NAM facilities in the Groningen field at which accelerographs are operating; see Figure 3.1.

No.	Location		No.	Location
1	Slochteren		12	Siddeburen
2	Tusschenklappen		13	Schaapbulten
3	De Eeker		14	Amsweer
4	Spitsbergen		15	Overschild
5	Scheemderzwaag		16	Ten Post
6	Kooipolder		17	Leermens
7	Zuiderpolder		18	Bierum
8	Eemskanaal		19	't Zandt
9	Zuiderveen		20	De Paauwen
10	Oudeweg		21	Sappemeer
11	Tjuchem		22	Froombosch

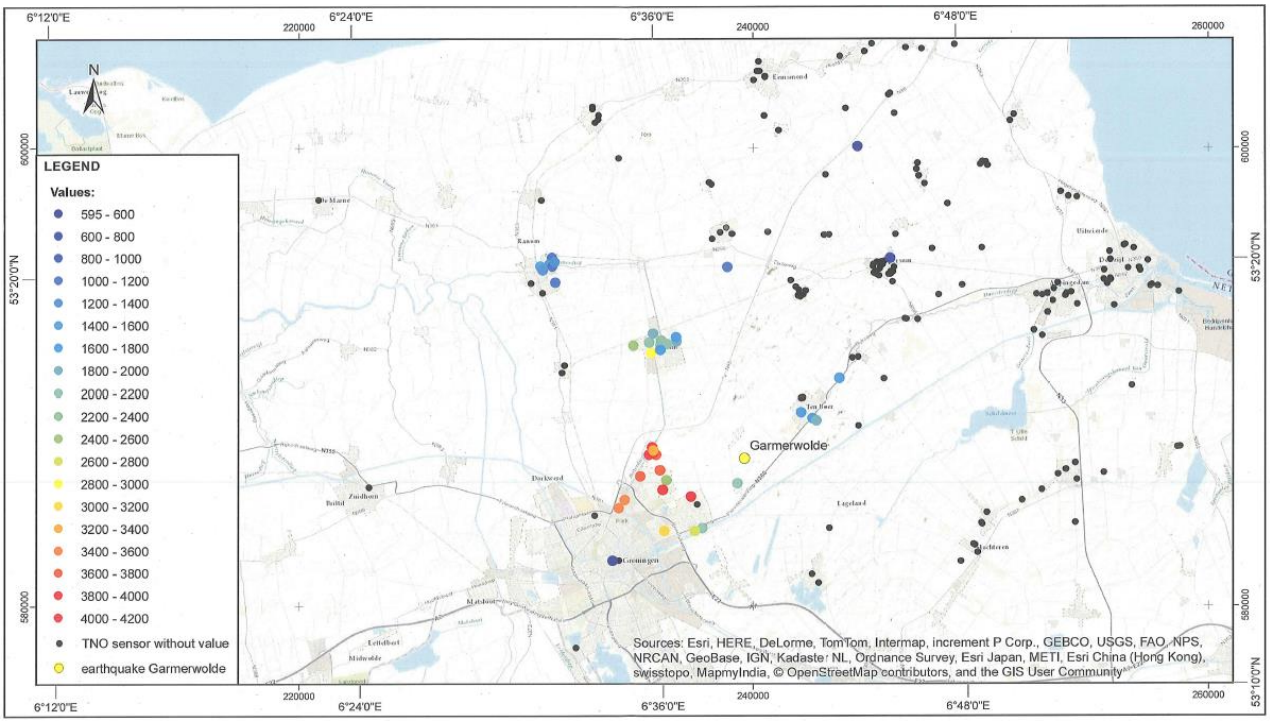


Figure 3.4. Locations of the TNO-installed accelerographs with indications of the level of ground shaking recorded in the **M**2.8 Garmerwolde earthquake of 30th September 2014

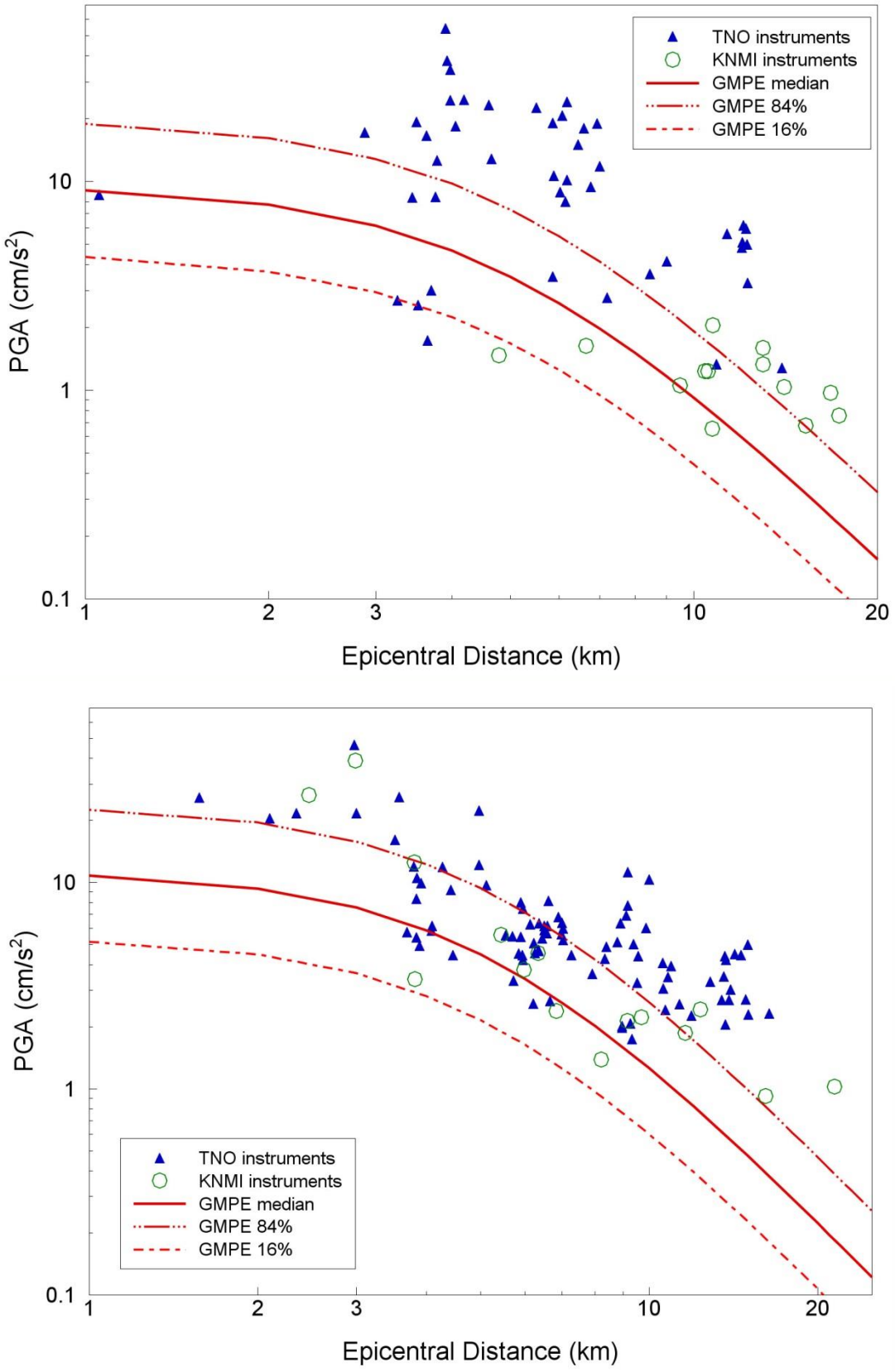


Figure 3.5. Comparison of the geometric mean PGA values recorded by the TNO-installed network (*blue triangles*) and the KNMI accelerographs (green open circles) from the **M**2.8 Garmerwolde earthquake of 30th September (*upper*) and the **M**2.9 Zandeweer earthquake of 5th November 2014 (*lower*) as a function of epicentral distance. Also shown, as red lines, are the median, 16-percentile and 84-percentile predictions from the Version 0 GMPEs of Bommer & Dost (2014)

A fourth network consists of some 200 accelerographs that have been installed by TNO, under contract to NAM, in some 20 public buildings and another 180 private homes. The latter instruments, the location of which cannot be made public because of Dutch legislation covering privacy, were installed in homes selected by requests made in response to an open invitation by NAM. Two earthquakes (30th September and 5th November 2014) have been recorded by these instruments, the locations of which are indicated in Figure 3.4. The PGA values from these instruments are compared with those from the KNMI instruments, and also with the predictions from the Version 0 GMPEs, in Figure 3.5, from which it can be appreciated that the values seem to be on the high side. This may be genuine, reflecting both the spatial variability of the ground motion as well as the influence of local ground conditions, but it is also likely to reflect the unusual practices in the installation of the instruments, which are mounted on angle brackets and often some distances from the floor (Figure 3.6).

The TNO-installed instruments are high-quality accelerographs (AS-73 accelerometers with GMS-plus recorders, from GeoSig) recording at a high sampling rate (250 per second, or a time interval of 0.004 s), but until the installation details of each instrument are known, it is judged preferable not to make incorporate these data into the derivation of the GMPEs.

Figure 3.6. Examples of the TNO-installed accelerographs in public buildings and private houses in the Groningen area. An instrument supported on an angle bracket attached to the side of a floor beam (*left*) and another on a bracket attached to the wall above the skirting board (*right*).

3.2. Recordings from the KNMI accelerograph network

A large number of recordings are now available from the KNMI instruments, although many of these are very rather small events. KNMI has made available 88 recordings from 12 earthquakes with magnitudes of between 2.6 and 3.6, obtained between August 2006 and November 2014. Some 60% of the records were obtained in the last two years, reflecting both an escalation of the seismic activity in the field and the expansion of the

recording network; more than half of the records were obtained in 2014 alone. Figure 3.7 shows the epicentres and recording stations for the 4 events producing most records.

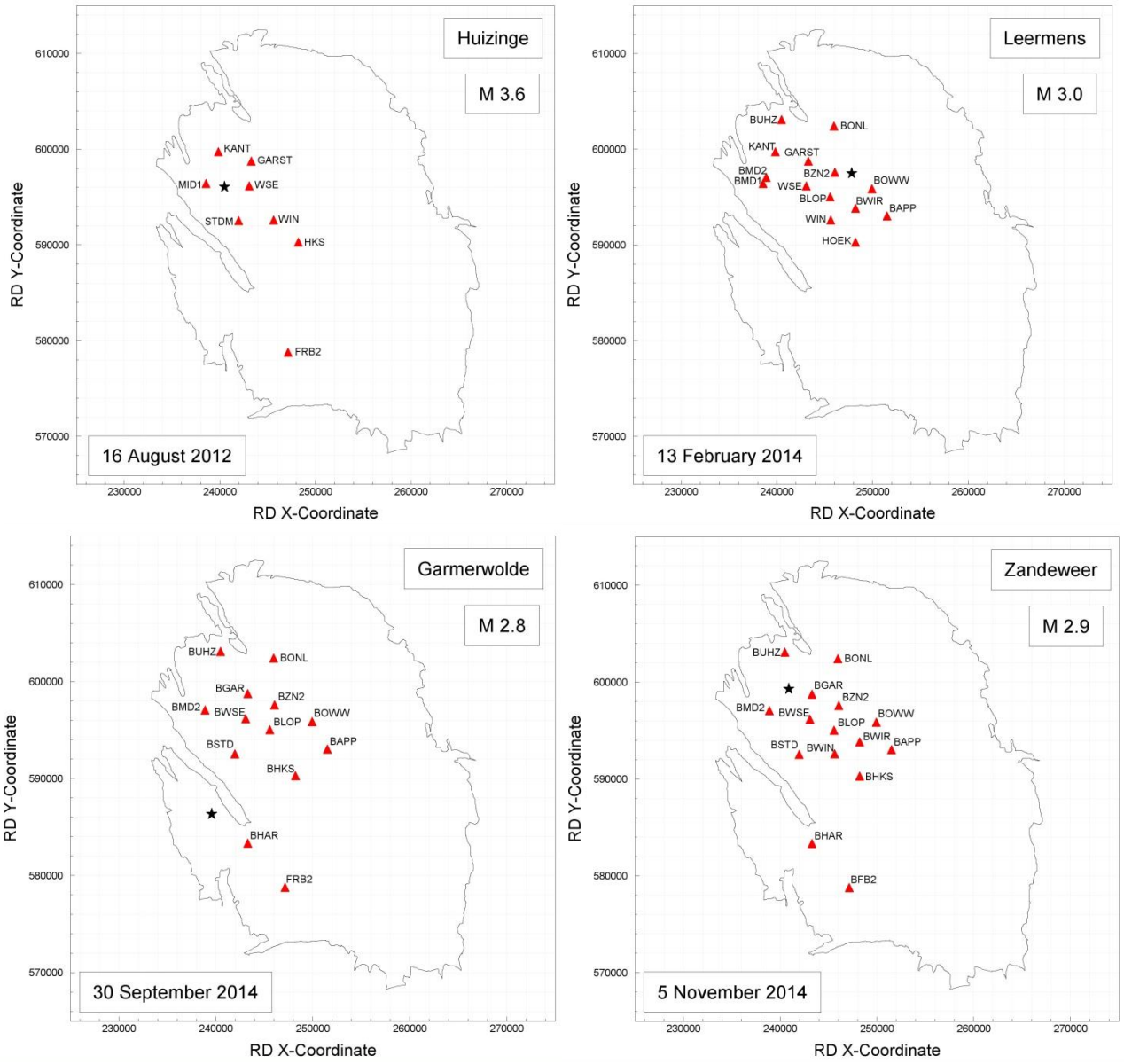


Figure 3.7. Locations of the epicentres (*black stars*) and recording stations (*red triangles*) of the four most productive earthquakes to date in terms of ground-motion data; these four earthquakes generated 50 of the 88 records in the current database

KNMI has advised not to use the three recordings from the FRB2/BFB2 station, located in the southern part of the field because of unusual, and currently unexplained, features of the recordings. These manifest as monochromatic high-frequency noise in the records, most clearly visible in the gradient of the Husid plots prior to the arrival of the strong shaking (Figure 3.8), which may be due to either a very localised site effect or due to some feature of the instrument installation.

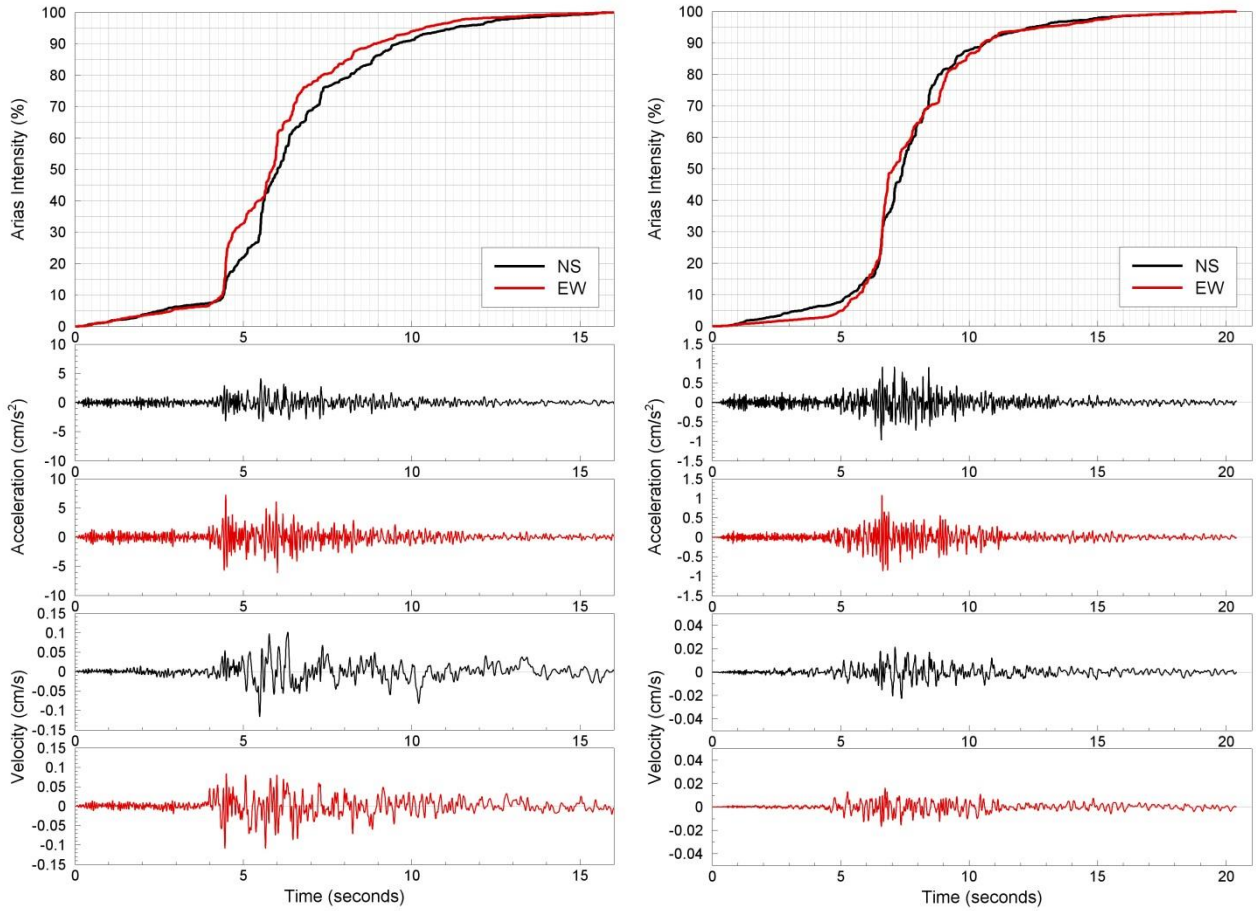


Figure 3.8. Recordings of the 2012 Huizinge earthquake at FRB2 (*left*) and of the 2014 Zanderweer earthquake at the upgraded BFB2 recording station (*right*) showing the two horizontal components in different colours. The top frames show the Husid plots depicting the accumulation of Arias intensity over time

The distribution of the data in magnitude-distance space is shown in Figure 3.9, including the subsequently excluded records from the FRB2/BFB2 station. The data are reasonably well distributed over distances up to 15 km from the epicentre and also within the magnitude range covered, but it should be noted that this corresponds to only 1.1 units of magnitude, making inferences about magnitude scaling rather uncertain. Figure 3.10 shows the geometric means of the PGA values against epicentral distances, grouped by magnitude ranges, and compared with the median predictions from the Version 0 GMPE. The comparison of curves and data for the larger records suggest that the rate of attenuation over short epicentral distances may be under-estimated by the Version 0 GMPEs, a point that is discussed in Chapter 5. A clear observation is that the levels of motion recorded to date are rather low: even looking at individual components, the maximum PGA in the database is 0.084*g* and the maximum PGV is 3.51 cm/s. Only 12 records have a PGA on at least component that is larger than 0.03*g*.

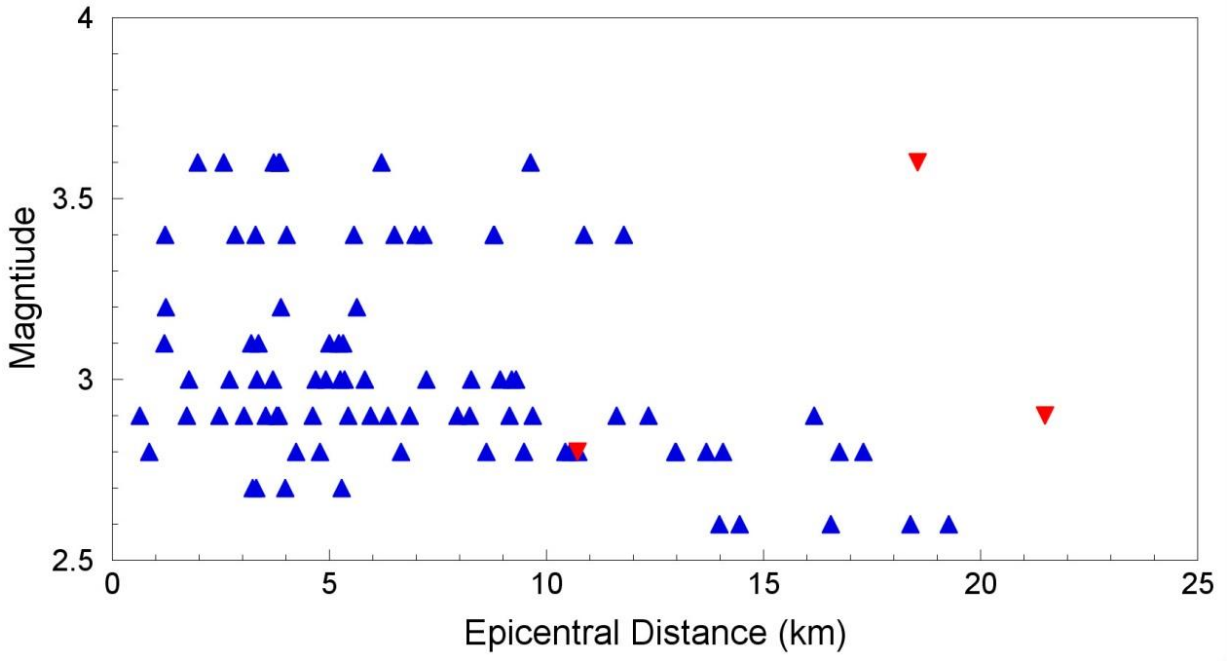


Figure 3.9. Distance-magnitude distribution of the Groningen ground-motion database; the red triangles are the recordings from the FRB2/BFB2 station, excluded from the current analyses

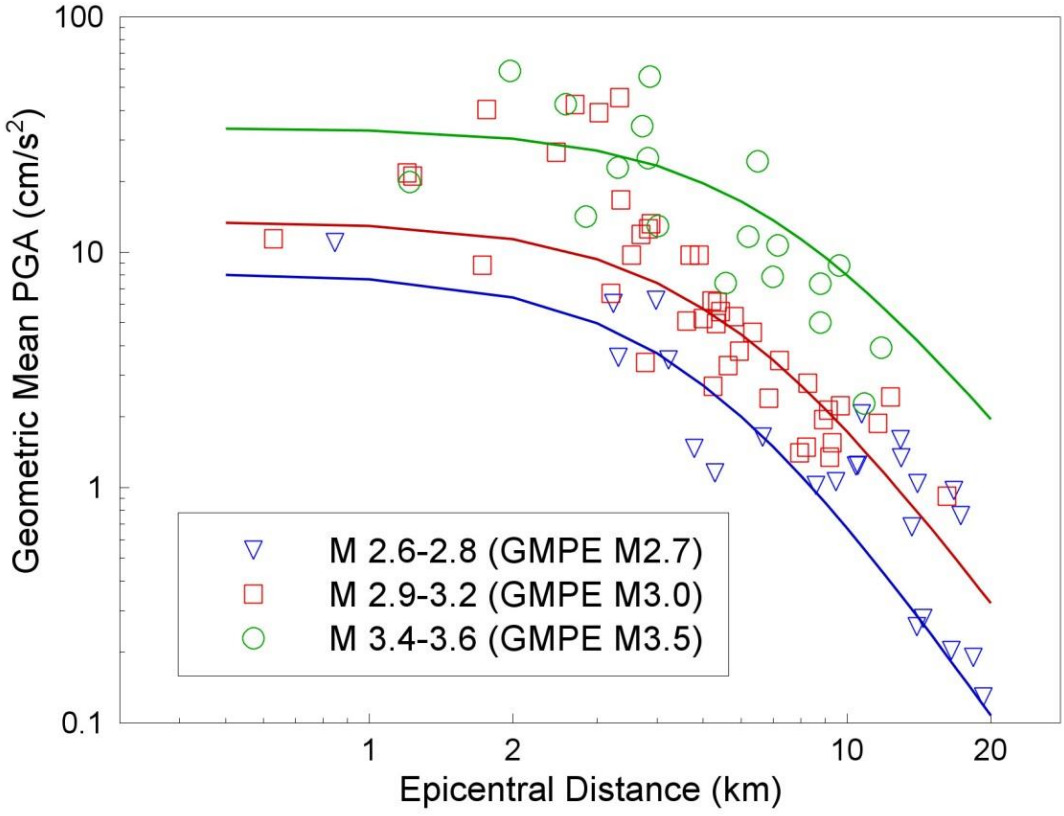


Figure 3.10. Geometric mean PGA values from the Groningen ground-motion database grouped by magnitude ranges and plotted against epicentral distances. Also shown for comparison are the median predicted values from the Version 0 GMPE

For the preliminary analyses presented herein, the records have only been processed by the application of simple baseline (linear or cubic) adjustments, where deemed necessary from inspection of the velocity trace obtained by integration of the acceleration time-series. In all cases, the pre-event memory—generally of a few seconds duration—was first removed, in order to make the record more manageable. In some cases, no baseline adjustment was required. Inspection of the Husid plots suggested that the Arias intensity was almost completely insensitive to the specific baseline applied, if any (Bommer *et al*., 2014). The response spectral ordinates were found to vary slightly for oscillator periods beyond about 2 seconds (Bommer *et al*., 2014), but for these very small earthquakes the long-period ordinates are in any case very low. One shortcoming of this approach is we do not have a clear guide on the usable period range, particularly at the long-period end (*e.g.,* Boore & Bommer, 2005; Akkar & Bommer, 2006); a number of studies have shown the high-frequency response ordinates are very insensitive to processing (*e.g*., Douglas & Boore, 2010; Akkar *et al*., 2011) and with the high sampling rate of these records (200 Hz), there are no concerns in this regard.

There are a few reasons for proceeding now with the simple baseline corrections rather than the application of appropriate filters to the records, including all of the following:

1. The processing of selecting appropriate filter cut-offs and applying the filters is time-consuming, whereas there is an imperative to generate the Version 1 GMPEs on a short time scale; for subsequent revisions of the GMPEs, more sophisticated processing techniques can be applied, but these are unlikely to have a significant impact on the values of the ground-motion parameters being predicted.
2. The analyses of the recordings by Dr Ben Edwards, as part of the stochastic inversions, will identify signal-to-noise ratios that will inform the selection of appropriate filter parameters.
3. These accelerograms are high-quality digital recordings, with relatively low noise; at the same time, in some of the records there is evidence for baseline shifts—which are not uncommon in digital records—which should preferably be removed through the application of piecewise-linear baseline adjustments on the velocity (Boore, 2001; Boore & Bommer, 2005). The application of filters would tend to effectively conceal these features—examples of which are shown in Figure 3.11—but it is preferable to remove them explicitly if this is possible. In such cases, there is a fit to the velocity trace over the interval judged to be linear, and the gradient (derivative) of this line is then subtracted from the corresponding portion of the acceleration time-series. Such piece-wise linear fits are applied sequentially, if needed, along the length of the record.

In conclusion, the Groningen database consists of 85 accelerograms—each with two horizontal components, which are aligned approximately NS and EW—from 12 earthquakes with magnitudes in the range from 2.6 to 3.6, and recorded at epicentral distances of up to 20 km. Considering the individual components, the recorded PGA values range from 0.11 cm/s^2 to 82.08 cm/s^2 (*i.e*., 0.0001*g* to 0.0837*g*).

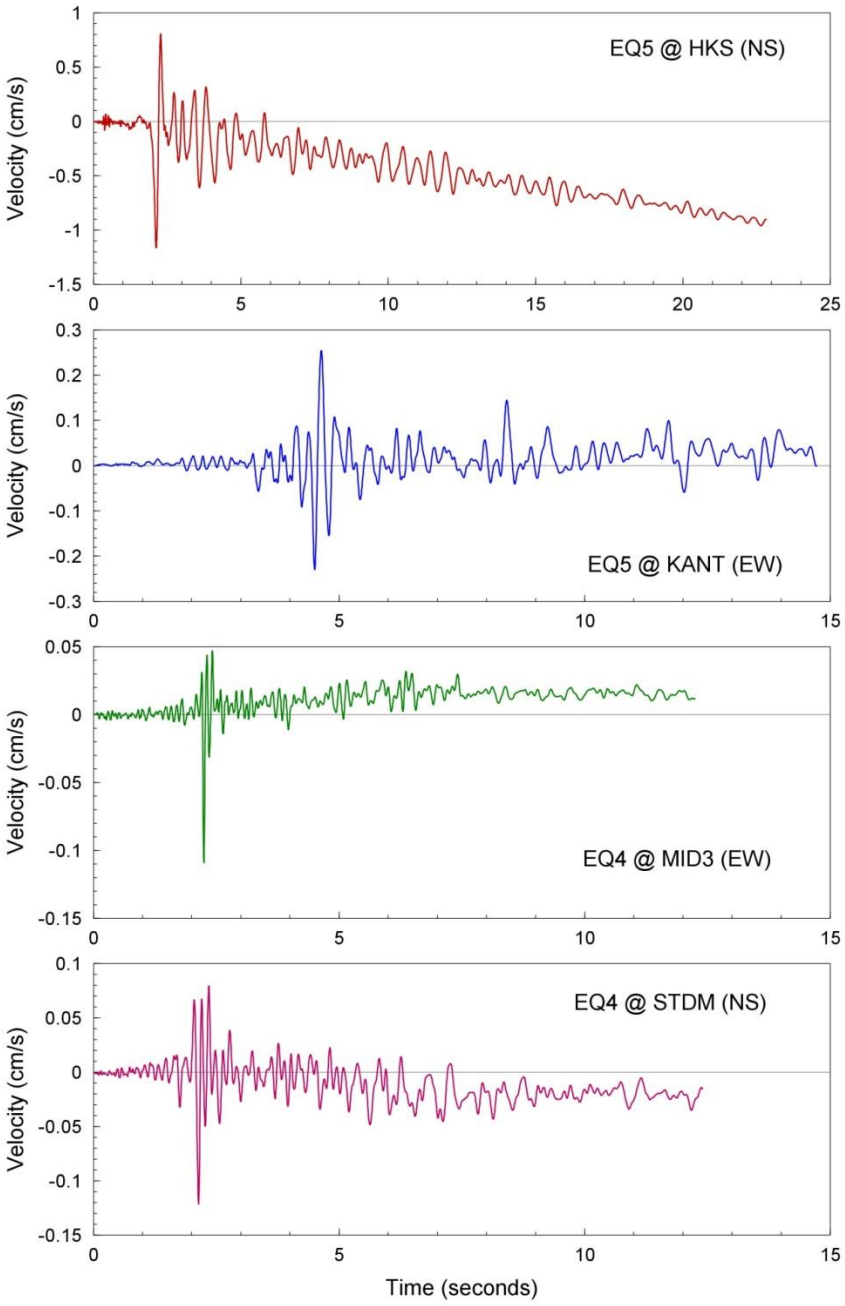


Figure 3.11. Examples of velocity traces of raw records (with pre-event memory removed) displaying discontinuous linear offsets in the baseline

3.3. Residuals relative to existing GMPEs

In order to obtain some preliminary insights into the characteristics of the Groningen dataset, some simple residual analyses are presented. Firstly, the PGA residuals of the complete dataset with respect to the Version 0 GMPE are calculated; this is a worthwhile exercise, given that only 40 records were used to derive the Version 0 model and there are now 85 records available. The additional records do alter the distribution of the data, in particular in terms of extending the distance range; the Version 0 dataset was limited to around ~10 km, whereas the distribution now gives reasonable coverage up to 20 km (Figure 3.12).

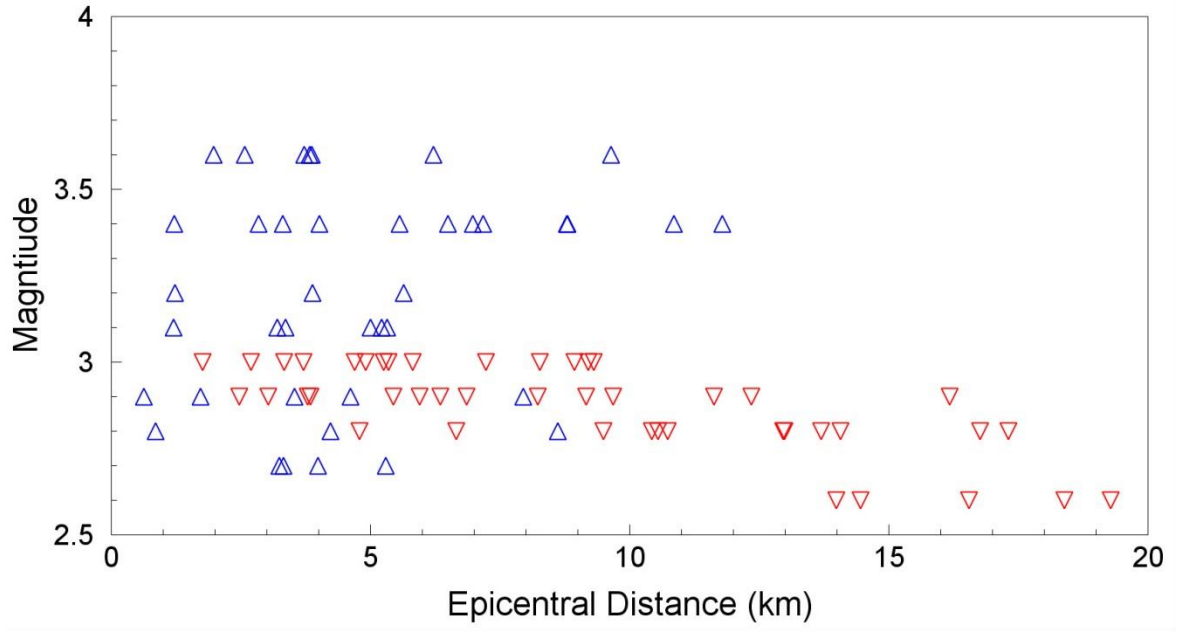


Figure 3.12. Magnitude-distance distribution of the dataset distinguishing between the records used to derive the Version 0 GMPEs (*blue triangles*) and the 45 additional recordings from the four most recent earthquakes (*red inverted triangles*).

Before entering into the details of these simple analyses of the recordings, it is worthwhile summarising the final characteristics of dataset whose distribution is illustrated in Figure 3.12. Table 3.2 lists the basic characteristics of the earthquakes and the usable recordings from each event. Figure 3.13 shows the location of the epicentres of the earthquakes within the Groningen field, with each event identified by its ID number from Table 3.2 and with its magnitude indicated in parentheses. The distribution is consistent with the historical patterns in the field, with the greatest concentration of events where the field compaction is greatest, although two of the more recent events are towards the south. The 6 events of **M** 3 and greater all cluster is fairly small area coincident with the centre of the compaction bowl.

Table 3.2. Earthquakes producing records from the Groningen accelerograph network

EQ	Date			Time		M	RD Coordinates		Name	Records	
ID	Y	M	D	H	M		X (m)	Y (m)		Event	Total
01	2006	VIII	8	05	04	3.4	242,159	596,659	Westeremden	4	4
02	2008	X	30	05	54	3.1	243,740	595,168	Westeremden	6	10
03	2009	V	8	05	23	2.9	246,479	597,129	Zeerjip	5	15
04	2011	I	19	19	39	2.7	238,780	593,075	Westerwijtwerd	4	19
05	2011	VI	27	15	48	3.4	248,253	591,487	Garrelsweer	8	27
06	2012	VIII	16	20	30	3.6	240,504	596,073	Huizinge	7	34
07	2013	I	19	20	10	2.8	248,515	589,488	Overschild	3	37
08	2013	II	7	23	19	3.2	240,085	600,945	Zandeweer	3	40
09	2014	II	13	02	13	3.0	247,804	597,489	Leermens	14	54
10	2014	IX	1	07	17	2.6	248,489	579,359	Froombosch	5	59
11	2014	IX	30	11	42	2.8	239,565	586,336	Garmerwolde	12	71
12	2014	XI	5	1	12	2.9	240,890	599,307	Zandeweer	14	85

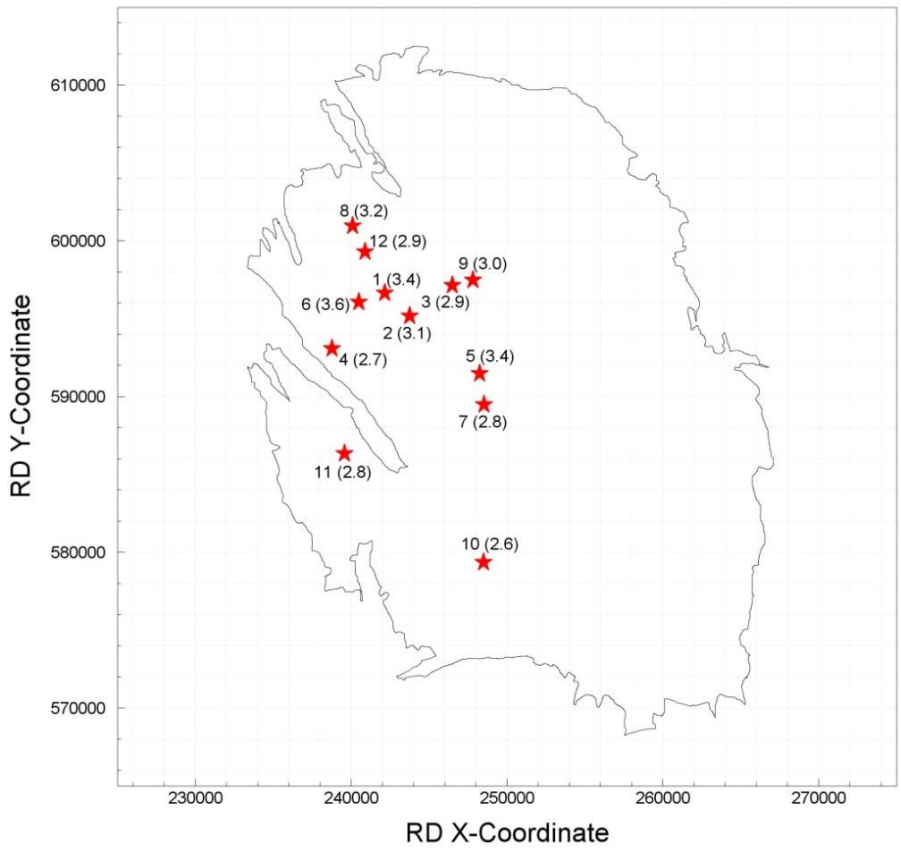


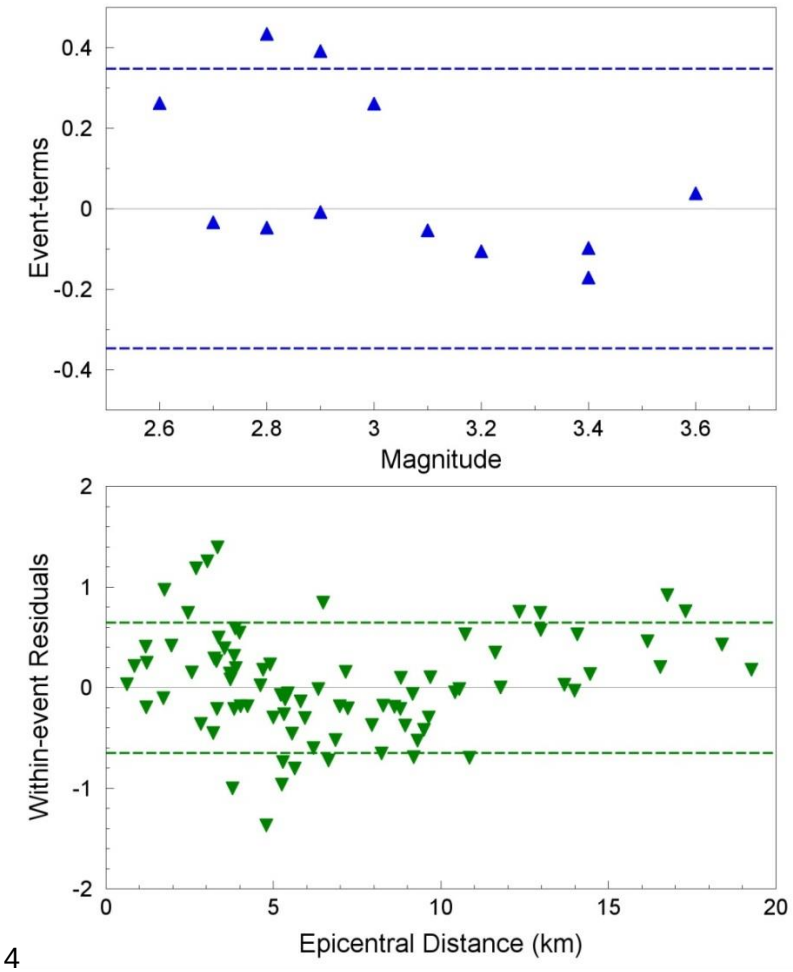
Figure 3.13. Locations and magnitudes of the 12 earthquakes represented in the database

Even though the database is still small, since it now includes a few well-recorded earthquakes, it is considered appropriate—and certainly more informative—to examine the residuals after their separation into between-event and within-event components. For this, the formulation of Abrahamson & Youngs (1992) is used to calculate the event-term, η_i for each earthquake as a function of the number of records, n_i, it contributes and the relative proportions of between-event (τ) and within-event (ϕ) sigma:

$$\eta_i = \frac{\tau^2 \sum_{j=1}^{n_i} (y_{i,j} - \mu_{i,j})}{n_i \tau^2 + \phi^2} \qquad (3.1)$$

where y is the logarithm of the ground-motion parameter and μ its predicted (logarithmic) median value. The between-event residuals are then obtained by subtracting this event term from the total residuals. The between- and within-event residuals of PGA are calculated in this way using the Version 0 GMPE and plotted against magnitude and distance, respectively, in Figure 3.14. The between-event residuals (or event terms) are generally rather small, except for 4 large positive values (implying under-prediction) that all correspond to the earthquakes that were not part of the Version 0 database. The event terms of the original database are on average slightly negative, suggesting that the original fit was imperfect—which is not surprising—and tending towards slight over-prediction in

terms of magnitude scaling. Even with these poorly fitting points, the calculated value of τ is just 0.21, which is considerably smaller than the assumed value for the Version 0 GMPE of 0.35, taken from Akkar *et al*. (2014a); however, any standard deviation calculated from such a small dataset should be interpreted with caution since it may easily be an underestimate as a result of the sample size.



4

Figure 3.14. Between- and within-event residuals of PGA with respect to the Version 0 GMPE plotted against magnitude and distance respectively. The dashed lines show the between-event (*upper*) and within-event (*lower*) sigma values.

The within-event residuals do not display any clear trends, apart from being apparently more dispersed at short distance (R_{epi} < 5 km). Considering only the very close-in recordings (R_{epi} ≤ 3 km), one can discern a tendency towards positive residuals, indicating under-estimation of the recorded values. The calculated value of ϕ from these residuals is 0.52, which again is considerably smaller than the assumed Version 0 value of 0.65; once again, however, the influence of the small sample size needs to be borne in mind. If we consider only the recordings from epicentral distances of less than 6 km, where the data are a little more abundant and appear to show a greater dispersion, the within-event sigma is calculated as 0.56, which is still smaller than the assumed value assigned to the Version 0 GMPE.

In order to obtain additional insights into the nature of the complete Groningen dataset, residuals for PGA and spectral accelerations at oscillator periods of 0.2, 0.5, 1.0 and 2.0 s are calculated using both the original Akkar *et al*. (2014a) GMPE and the Bindi *et al*. (2014) European GMPE that was also derived used using the RESORCE database (Akkar *et al*., 2014b). For the Akkar *et al*. (2014a) equation, consistent with parameter choices discussed in Chapter 2, the R_{epi}-based model is used, with predictions for the average ordinates from normal and strike-slip faulting earthquakes, for V_{S30} = 200 m/s. For the Bindi *et al*. (2014) GMPEs, the R_{hyp}-based model (rather than R_{JB}), for the same style-of-faulting assumption and site condition (the version of the equations based on site classes was not used). The between- and within-event residuals with respect to the two GMPEs are shown for the five ground-motion parameters indicated above in Figures 3.15 to 3.19.

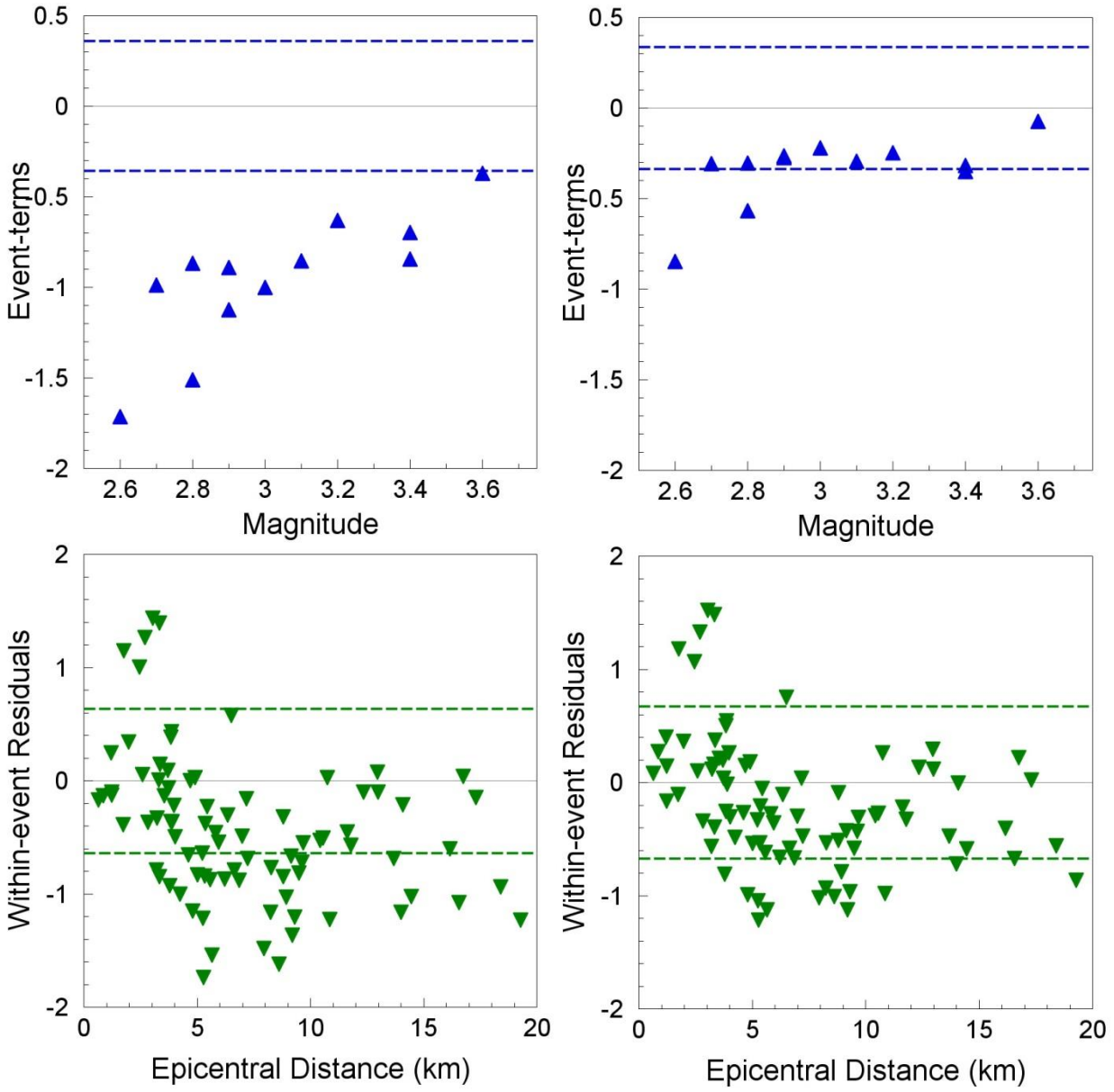


Figure 3.15. Between- and within-event residuals of PGA with respect to the GMPEs of Akkar *et al*. (2014a) (*left*) and Bindi *et al.* (2014) (*right*) plotted against magnitude and distance respectively. The dashed lines show the between-event (*upper*) and within-event (*lower*) sigma values.

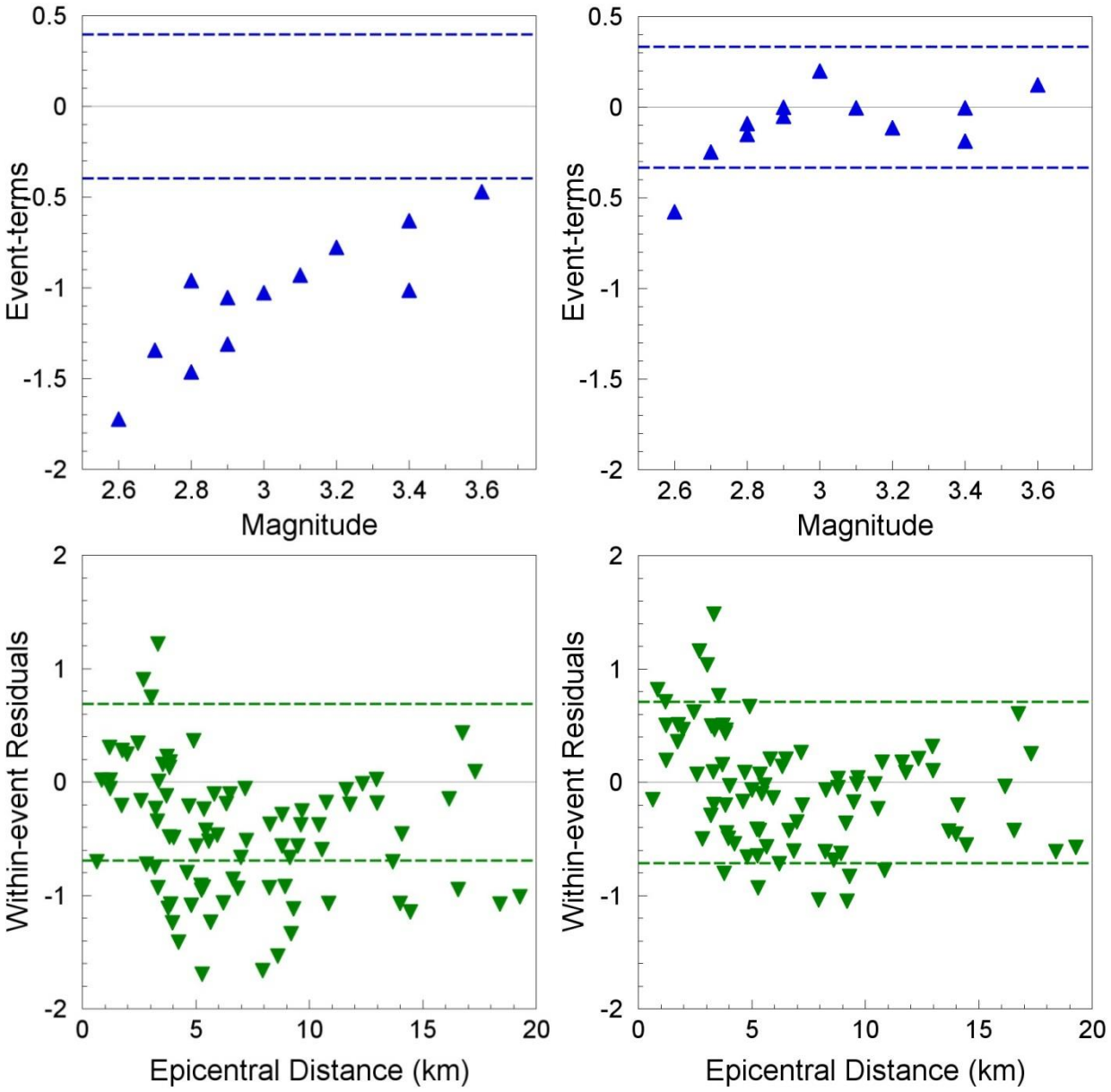


Figure 3.16. Between- and within-event residuals of SA(0.2s) with respect to the GMPEs of Akkar *et al*. (2014a) (*left*) and Bindi *et al*. (2014) (*right*) plotted against magnitude and distance respectively. The dashed lines show the between-event (*upper*) and within-event (*lower*) sigma values.

The residual patterns seen in Figures 3.15 to 3.19 are very interesting. For both GMPEs, as the response period increases from 0.01 second (*i.e*., PGA) to 2.0 seconds, the residuals move from a pattern of greater to lesser over-estimation. For the between-event residuals, this pattern culminates in a good fit (slight under-estimation) at 2.0 seconds for the Akkar *et al*. (2014a) model, whereas for the Bindi *et al*. (2014) equations the fit is almost perfect at 0.2 seconds, with severe under-estimation at long periods. In the Akkar *et al.* (2014a) between-event residuals, there are strong positive trends against magnitude at short periods, becoming weaker with increasing period until practically vanishing at 2.0 seconds. By contrast, the Bindi *et al*. (2014) between-event residuals show only a very

weak positive trend against magnitude at short periods, becoming mildly negative at longer periods.

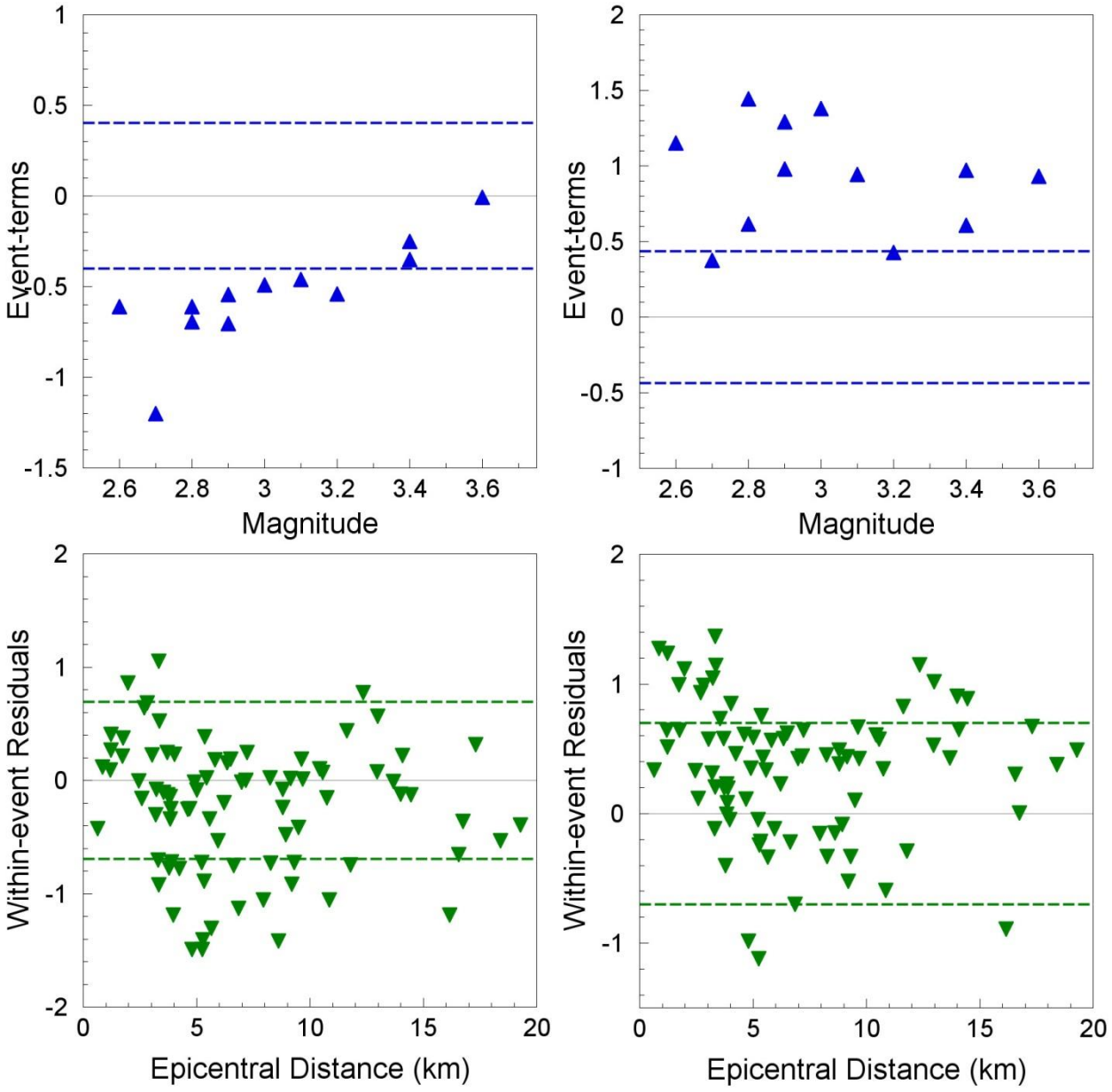


Figure 3.17. Between- and within-event residuals of SA(0.5s) with respect to the GMPEs of Akkar *et al*. (2014a) (*left*) and Bindi *et al.* (2014) (*right*) plotted against magnitude and distance respectively. The dashed lines show the between-event (*upper*) and within-event (*lower*) sigma values.

In terms of the within-event residuals, the fit is generally reasonably good in both cases at most periods, the exception being a trend to over-prediction at shorter response periods (0.01 and 0.2 s) for the Akkar *et al.* (2014a) equations. For both equations, a notable feature of the within-event residuals is an apparent trend of relative under-estimation at very short epicentral distances, with this trend decaying very rapidly over the first few kilometres. This may be due to the near-source saturation terms used in the equations,

which are independent of magnitude and take values that are probably too large for such small earthquakes. In the Akkar *et al.* (2014a) equation, the value is 7.5 km, for all response periods, whereas in the Bindi *et al.* (2014) equations it takes values of around 4.5-5 km for the periods considered, with the exception of 1.0 second where it is just 3.3 km. This interpretation of the effect is supported by the observation that at 1.0 second, this trend is less pronounced (Figure 3.18).

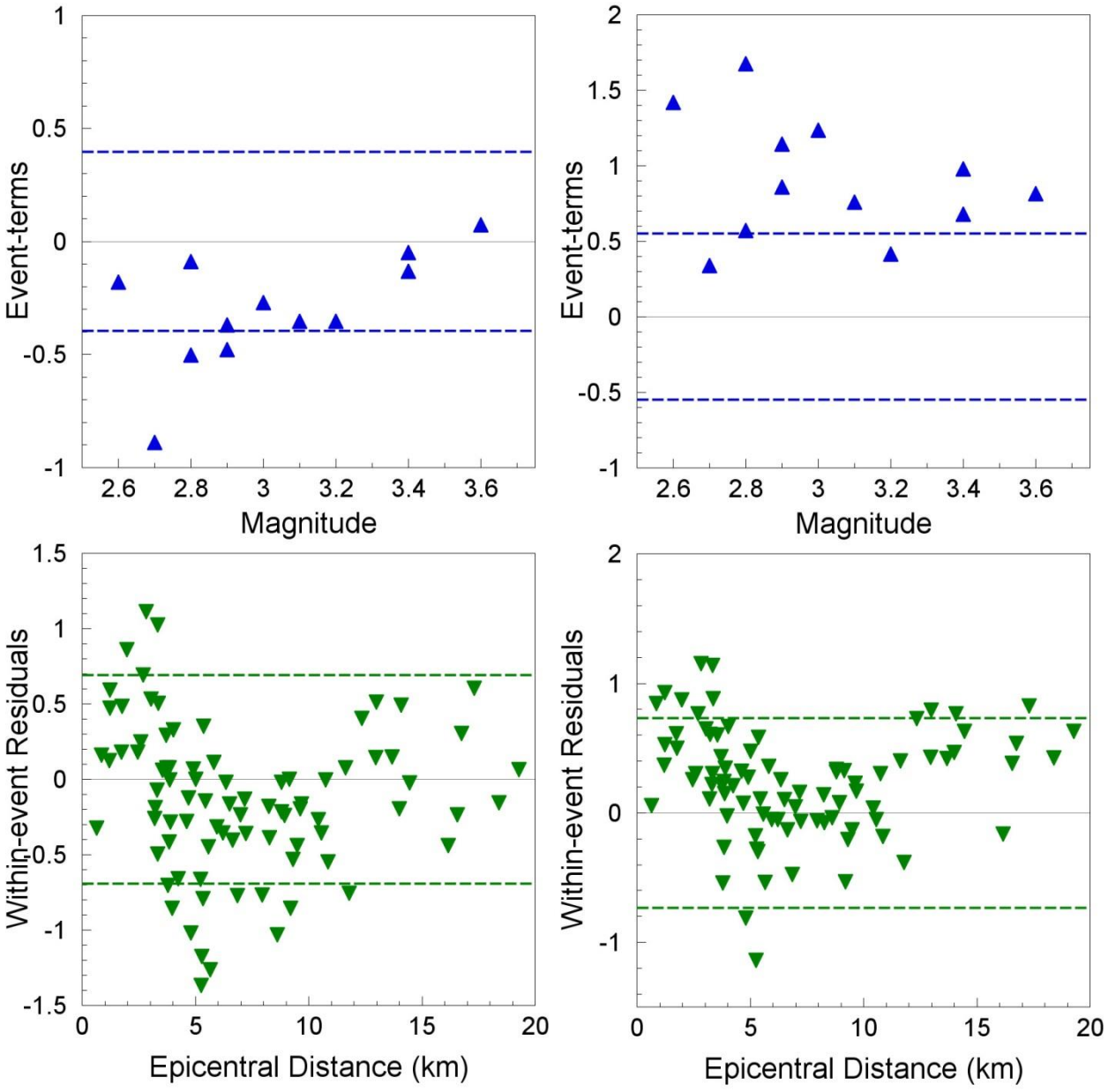


Figure 3.18. Between- and within-event residuals of SA(1.0s) with respect to the GMPEs of Akkar *et al*. (2014a) (*left*) and Bindi *et al*. (2014) (*right*) plotted against magnitude and distance respectively. The dashed lines show the between-event (*upper*) and within-event (*lower*) sigma values.

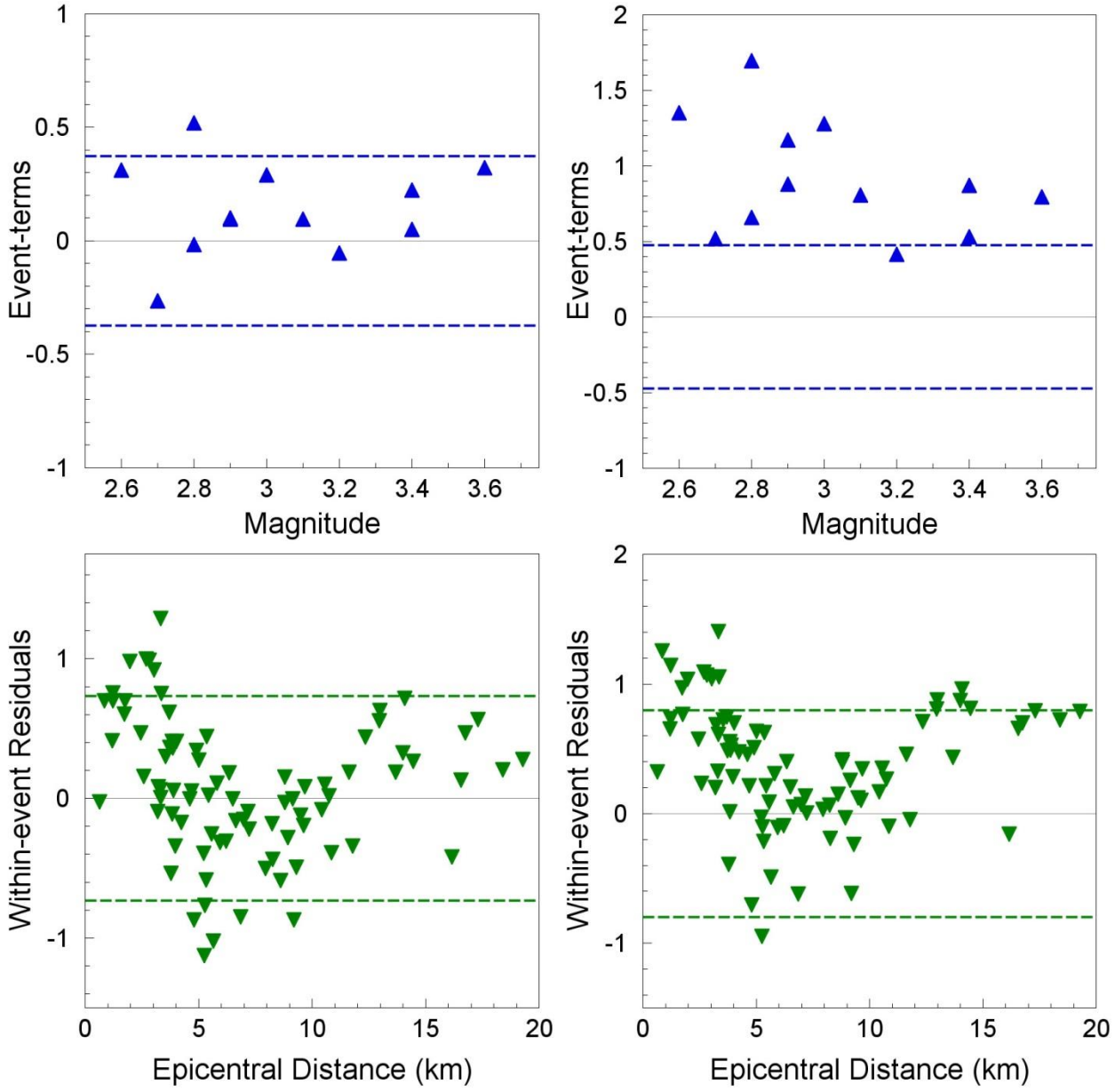


Figure 3.19. Between- and within-event residuals of SA(2.0s) with respect to the GMPEs of Akkar *et al*. (2014a) (*left*) and Bindi *et al.* (2014) (*right*) plotted against magnitude and distance respectively. The dashed lines show the between-event (*upper*) and within-event (*lower*) sigma values.

The differences in the residual patterns for the two equations are quite pronounced, which in itself is an interesting observation given that the models have very similar functional forms, the biggest difference being the inclusion of non-linear response effects in the site term of the Akkar *et al.* (2014a) equations. Both equations are derived from sub-sets of the RESORCE database (see Section 4.1), selecting only records from sites with known V_{S30} values. The distributions of the two datasets in magnitude-distance space are very similar, as can be appreciated from comparing Figure 3.20 with Figure 4.4. Both datasets cover the same magnitude range, with a lower limit of **M** 4. The dataset used by Bindi *et al.* (2014) is about 18% larger than that used by Akkar *et al*. (2014a) although it included data from only 6% more earthquakes; the key difference is that the Bindi *et al*. (2014) model

used data recorded up to distances of 300 km, as compared with 200 km for the other model.

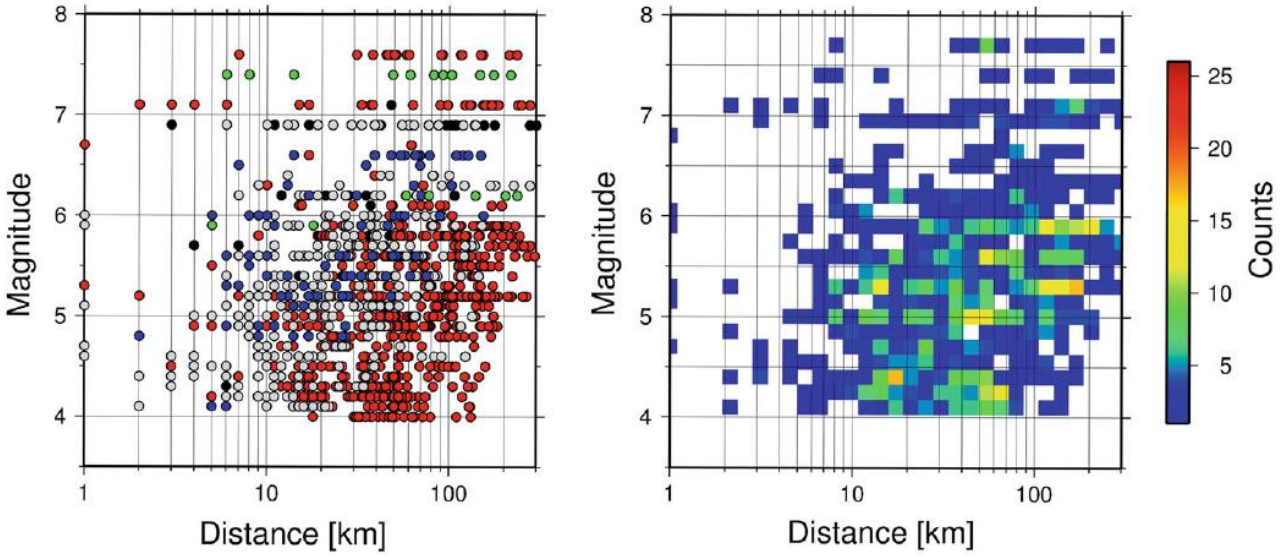


Figure 3.20. Magnitude-distance distribution of the data used in the V_{S30}-dependent GMPE for SA(20.1s) of Bindi *et al.* (2014). The colours in the left-hand plot indicate the country of origin of the data: Turkey (*red*), Italy (*grey*), Greece (*blue*), Iran (*green*), Iceland (*yellow*) and others (*black*).

Although the exact reasons for the differences in the two GMPEs may not be immediately apparent, it has been observed that the Bindi *et al*. (2014) equations have much stronger magnitude scaling at lower magnitudes than other recent GMPEs, including the Akkar *et al.* (2014a) model (Figure 3.21). For the current discussion, however, the divergence between these models is not the key issue. The most important observation is that whilst both equations seem to fit the Groningen data reasonably well at a particular oscillator period, neither equation could be simply adopted for this application across the range of periods at which predictions of spectral accelerations are required. A similar exercise could be performed using alternatives such as the NGA-West2 GMPEs (Gregor *et al*., 2014), although as noted previously these are all defined in terms of extended-rupture distance metrics. There is clearly a need, therefore, to generate new GMPEs for the Groningen field.

In closing this section, the variability of the Groningen residuals with respect to these two equations are estimated for the period at which each equation provides the best fit. The results are presented in Table 3.3. Clearly account must be taken of the small size of the Groningen dataset and the fact that this could easily lead to under-estimation of the variability. However, even with the imperfect fit in both cases—particularly at very short distances—the variability appears to be very much smaller than that associated with the original GMPEs. Given the fact that all of the Groningen events are occurring in the same source and with the exception of uppermost tens of metres the records correspond to very similar travel paths, this result is perhaps not entirely surprising. The potential implications for the hazard and risk assessments of the apparently small variability are significant.

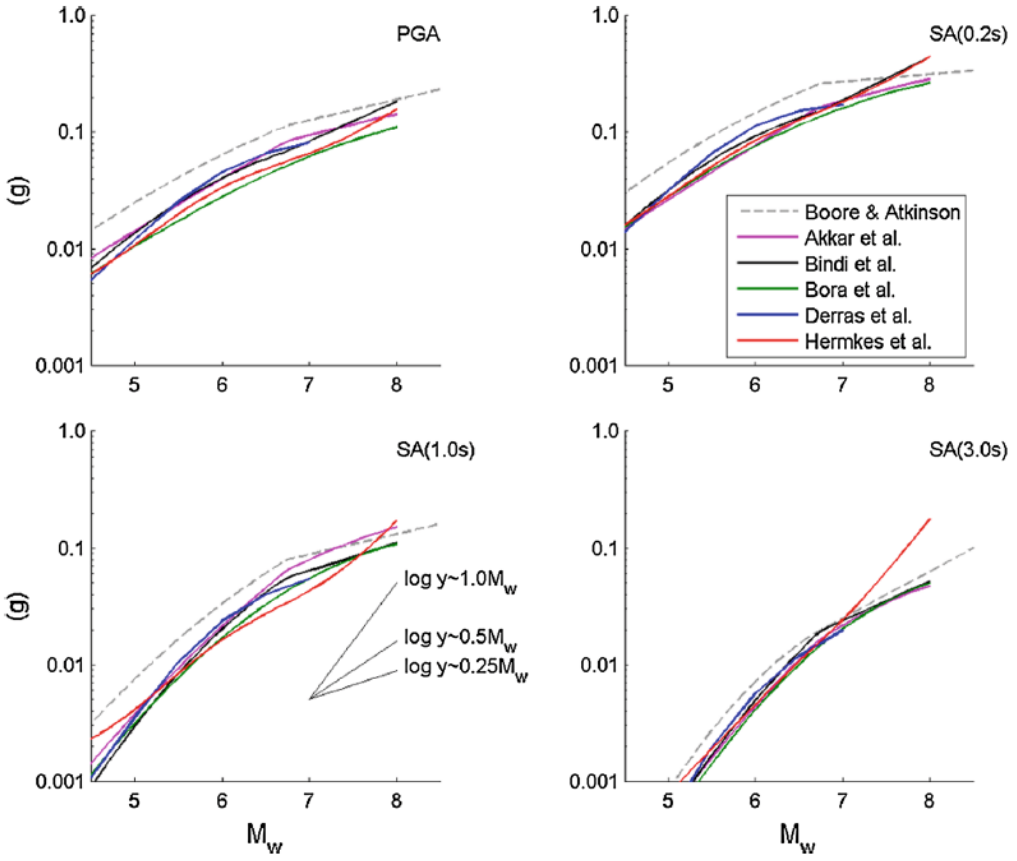


Figure 3.21. Comparison of the magnitude scaling of several GMPEs, including Akkar *et al*. (2014a) and Bindi *et al*. (2014), for PGA and spectral accelerations at three response periods (Douglas *et al*., 2014). Median predictions for R_{JB} = 30 km, V_{S30} = 760 m/s.

Table 3.3. Standard deviations of residuals computed for Groningen database with respect to the GMPEs of Bindi *et al.* (2014) and Akkar *et al*. (2014a).

Standard Deviation	**T =0.2 s, Bindi *et al.* (2014)**		**T = 2.0 s, Akkar *et al.* (2014a)**	
	Original GMPE	**Groningen Data**	**Original GMPE**	**Groningen Data**
ϕ	0.7111	0.5110	0.7333	0.4946
τ	0.3332	0.2099	0.3734	0.2433
σ	0.7853	0.5525	0.8229	0.5512

4. European Ground-Motion Database

The Groningen database presented in the previous Chapter provides invaluable insight into the nature and characteristics of the ground motions generated by induced earthquakes in the field to date. These local data will be used as the primary tool to generate the GMPEs and seismological theory will be used to extrapolate the models to the larger magnitudes to be considered in the hazard and risk calculations. Originally, an approach considered was to create a combined database of the Groningen data with European strong-motion data from tectonic earthquakes and then fit an appropriate functional form to the combined dataset. This approach would have been predicated on the assumption of similarity in ground motions from induced and tectonic earthquakes, as concluded, for example, by Douglas *et al*. (2013). We do not believe that there is an adequate basis for this assumption and consequently adopted the approach of using simulations, but nonetheless concluded that it would be very valuable to use other data to explore and evaluate the extrapolations to larger magnitudes, which will be subject to significant uncertainties. For this reason, the database of strong-motion recordings from tectonic earthquakes in Europe and the Middle East is used to guide the extension of new GMPEs to larger magnitudes.

This Chapter begins with an overview of the characteristics of the European database, slightly modified from that used to derive the GMPEs of Akkar *et al*. (2014a), considering in particular the subsets of this database that would be most suitable for this application. This is followed by an overview of new additions to the database, compiled specifically to enrich the database in the lower magnitude range of greatest relevance to the Groningen seismic hazard and risk study. The Chapter closes by identifying the combined database extracted from these sources that is considered appropriate for the analyses.

4.1. Updated Akkar *et al*. (2014) database

The compilation of a database of European strong-motion data was begun at Imperial College London by Professor N.N. Ambraseys and colleagues back in the 1970s. The database was expanded and enhanced through numerous projects, including EU-funded collaborations involving several partners throughout Europe and the Middle East. The most recent efforts to expand and improve the European database were undertaken as part of the EU-funded SHARE project, and the project SIGMA funded by EDF, which has created the RESORCE database (Akkar *et al*., 2014b). The RESORCE database effort has included the development of a database of uniformly estimated metadata and a uniformly processed databank of accelerograms coming primarily from the more seismically active countries of southern Europe, the Mediterranean and the Middle East, but a small number from less active areas in northwest Europe (Figure 4.1).

The metadata included in the RESORCE database includes moment magnitudes (as well as magnitude reported on other scales), focal depth, style-of-faulting, and the classification of the recording stations both in terms of generic site classes and, where available,

measured values of V_{S30}. Source-to-site distances were also calculated using four different distance metrics, two based on point representations of the seismic source (R_{epi}, R_{hyp}) and two relative to extended fault ruptures (R_{JB}, R_{rup}). The calculation of hypocentral distance requires a reliable estimate of focal depth, and the estimation of the rupture distance requires the geometry and location of the fault rupture plane; the Joyner-Boore distance can be estimated without necessarily having the rupture plane completely defined. Consequently, the numbers of records for which each distance metric is available vary (Figure 4.2): for all 5,882 records in the database R_{epi} is known but for 131 of these the absence of focal depths impedes the calculation of hypocentral distance. A total of 3,906 records are associated with an R_{JB} distance, but R_{rup} is available for only 2,490 records.

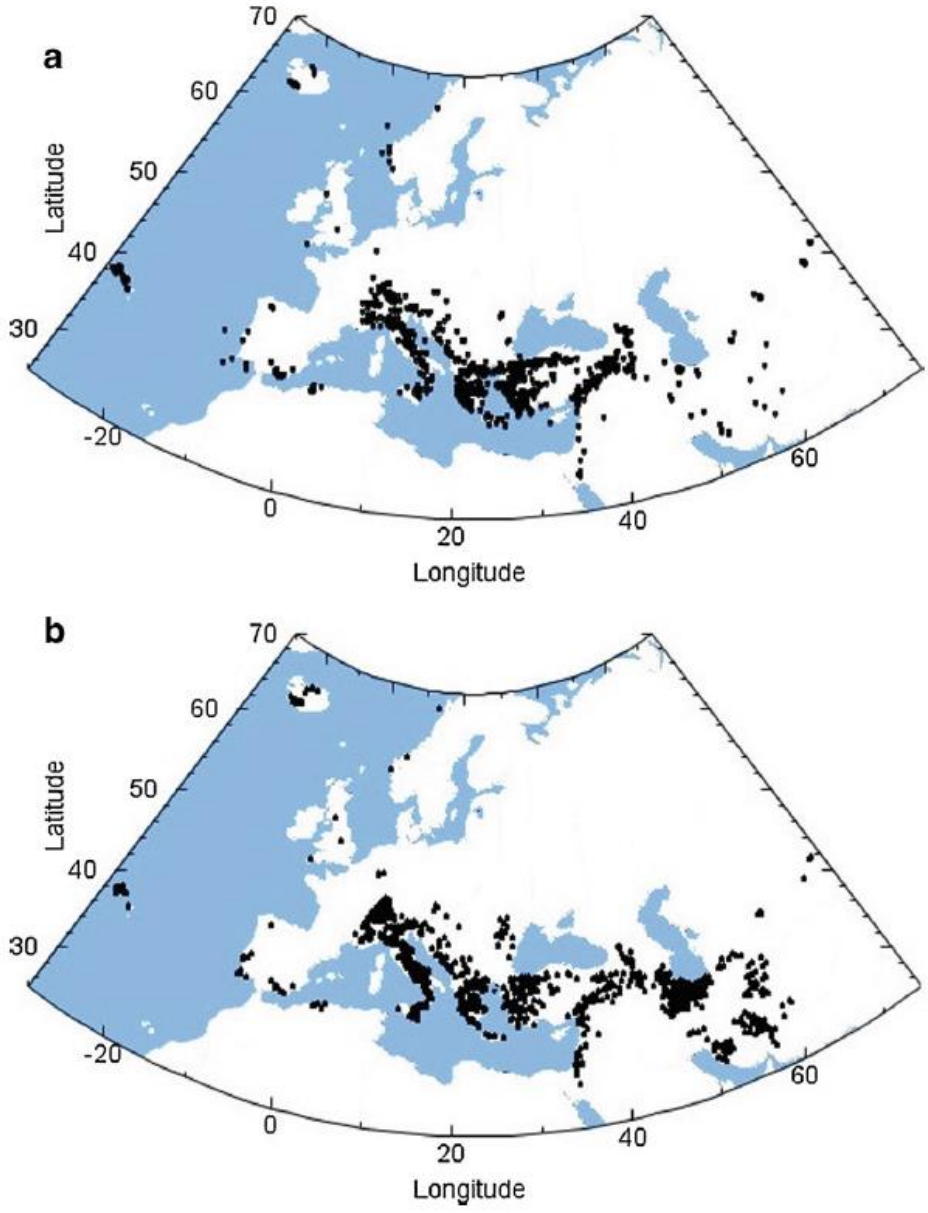


Figure 4.1. Geographical provenance of the records in the RESORCE strong-motion database by (a) earthquakes and (b) recording stations (Akkar *et al*., 2014b).

The fact that the rupture distance is known for less than two-thirds of the records having R_{JB} lends support to the decision to use the latter distance for the version of the GMPEs that will use an extended-rupture distance metric (Section 2.4) since the choice of R_{rup} would severely limit the available data. At this point, one might argue that it would therefore have been more appropriate to consider an alternative strong-motion database, the obvious candidate being the NGA-West2 database for which R_{rup} is calculated for a large proportion of the records (Figure 4.3). There are pragmatic considerations here, one of these being that the NGA-West2 acceleration time-histories are not currently available to all, and since we also need to develop GMPEs for durations this is an important limitation even if the flat-file of spectral accelerations were accessible. Another practical issue is that we have direct access to and communication with the developers of the European database, and can therefore obtain more detailed insights than might be the case with the NGA-West2 data.

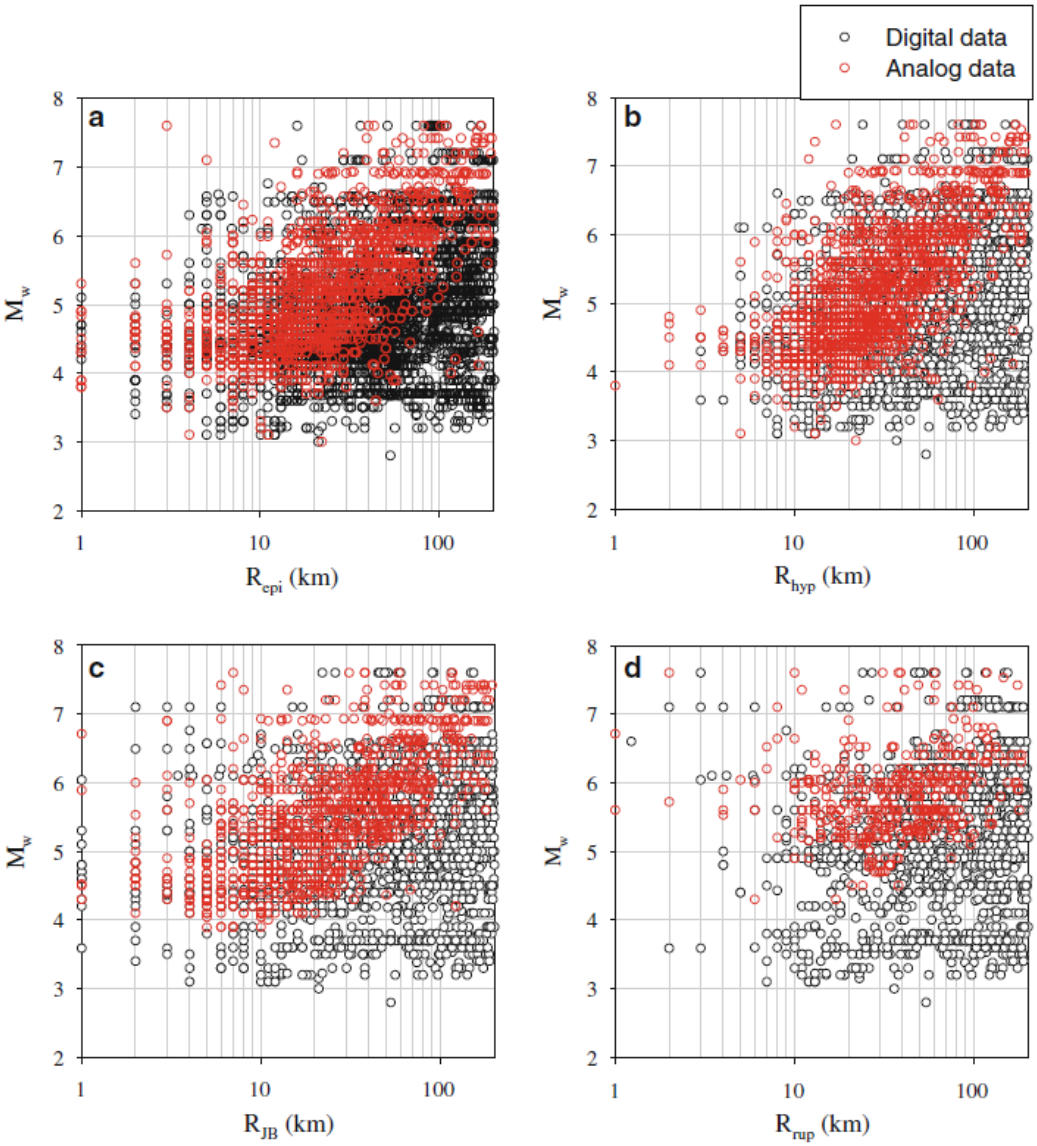


Figure 4.2. Magnitude-distance distributions of the RESORCE database in terms of (a) epicentral, (b) hypocentral, (c) Joyner-Boore and (d) rupture distance; the colour of the symbols indicates the type of recording instrument (Akkar *et al*., 2014b).

Even leaving aside such practical considerations, the true benefit of using R_{rup} rather than R_{JB}, given that there are very few records in either the European or NGA-West 2 databases from earthquakes as shallow as the Groningen events, is not clear. In effect, the benefit would ride on the assumption that the effects on ground motions of focal depth and horizontal distance are equivalent and interchangeable. This assumption may well not be valid, especially if stress drop varies with focal depth, which has been suggested to be the case for both natural (*e.g.*, Allen, 2012) and induced (*e.g.*, Hough, 2014) earthquakes.

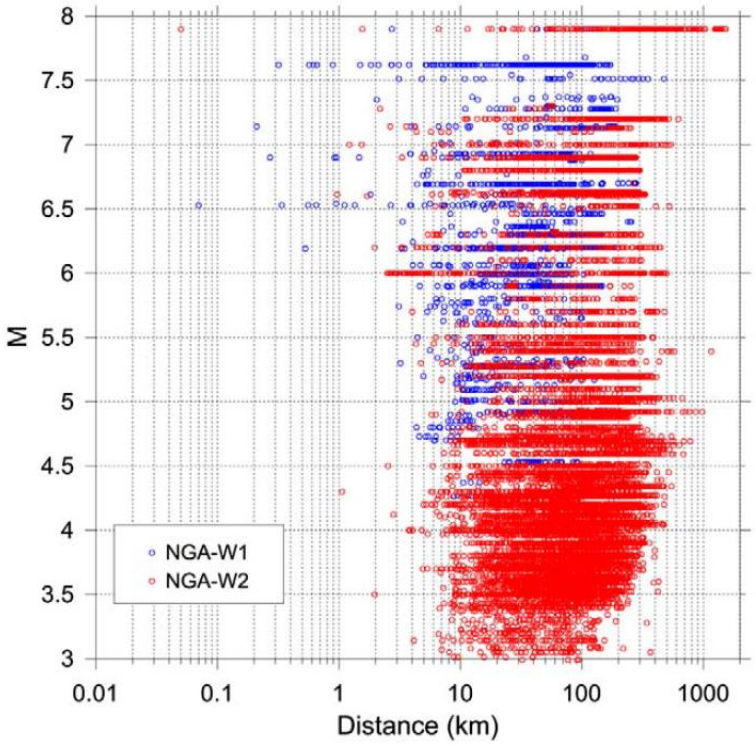


Figure 4.3. Magnitude-distance distribution of the NGA-West2 database (Ancheta *et al*., 2014)

In order to derive the Akkar *et al*. (2014a) GMPEs, a dataset (1,041 records from 221 earthquakes) was selected from the RESORCE database using the following criteria:

- Moment magnitudes, calculated directly rather than by conversion from another scale using empirical relationships, **M** ≥ 4
- Known style-of-faulting (normal, reverse or strike-slip)
- Focal depth < 30 km
- Calculated value of R_{epi}, R_{hyp} and R_{JB}
- Measured V_{S30} at recording site
- Only earthquakes with at least two recordings

The last criterion applies at PGA and very short response periods. Each record is used only up to a maximum period related on the applied filter parameters, and therefore at longer periods it may sometimes be the case that a single record remains from some

earthquakes. The distribution of the database in terms **M**-R_{JB}, grouped by different ranges of V_{S30} and with symbols reflecting the style-of-faulting, is displayed in Figure 4.4.

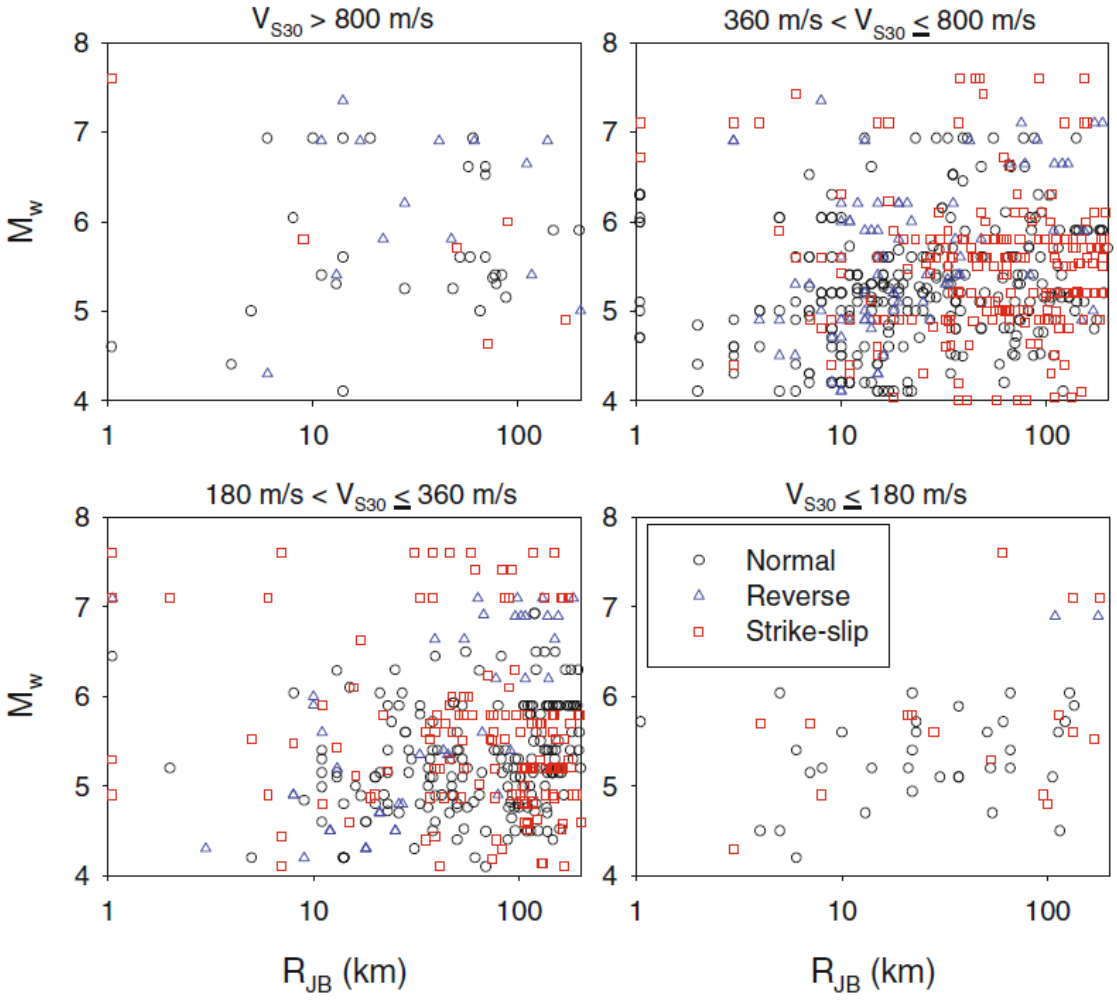


Figure 4.4. Magnitude-distance distribution of the Akkar *et al*. (2014a) database by V_{S30} range; the symbols reflect the style-of-faulting.

The database assembled for the Groningen project is essentially the same as that used by Akkar *et al*. (2014a), with some modifications to some metadata that resulted from ongoing work in the SIGMA project to refine and improve the RESORCE database. Additionally, this ongoing refinement work led to the removal of a total of 39 recordings due to the following reasons (some records were affected by more than one of these factors but for the primary motivation for removal is noted here):

- 13 records removed from RESORCE database because of concerns regarding the quality of the waveforms
- 14 records were judged to be lacking a reliable estimate of R_{JB}
- 5 records were removed because 2 re-evaluated focal depths ≥ 30 km

- 2 records removed because the **M** value for an earthquake was identified to be converted instead of calculated directly
- 5 records removed as a consequence of the previous removals leaving 5 singly-recorded earthquakes

The revised database therefore contains 1,002 records from 209 earthquakes. From this database, the next task is to select those records that are most likely to be useful to the derivation of Version 1 GMPEs for the Groningen field. The first step is to decide the criteria on which records will be selected or excluded, in terms of how different factors may influence the applicability of the data as potential analogues for motions from future earthquakes of **M** ≥ 4 in the field. This leads to the following considerations:

- **Magnitude**. The upper limit of the hazard integrations is **M**6.5 (Bourne *et al*., 2014), which corresponds to the estimated maximum magnitude, inferred from the extremely unlikely scenario of all of the field compaction at the end of production being released in a single seismic event. Since the hazard and risk calculations will include integrations up to this limiting value, the equations should be valid to **M**6.5, but there is no motivation to include larger earthquakes.
- **Distance**. The maximum separation of points on the boundary of the gas field is slightly less than 50 km. If it is assumed that earthquakes will only occur within the limits of the gas field, then taking account of the 5 km buffer around the field boundary to which the exposure database extends, the maximum distance considered in the hazard and risk calculations is on the order of 55 km. There is therefore no reason to extend the applicability of the GMPEs to greater distances; for the initial consideration of the European database, a maximum epicentral distance of 60 km is considered.
- **V_{S30}**. As was noted in Section 2.5, the Version 1 GMPEs will be developed for an assumed constant field-wide V_{S30} value of 200 m/s. For the Groningen recordings, it is assumed that the actual V_{S30} value at the accelerograph stations is sufficiently close to this to be a reasonable approximation. For the additional datasets at larger magnitudes, the spectral accelerations will be transformed using the site amplification factor, S, in the Akkar *et al*. (2014a) GMPE to the value that would be predicted for V_{S30} = 200 m/s. This transformation is made by first calculating the S factor to transform the motion for the station-specific V_{S30} and dividing the acceleration by this value to obtain the reference rock (V_{S30} = 750 m/s) motion. This is then multiplied by the S factor for the target V_{S30} of 200 m/s. The use of the Akkar *et al*. (2014a) model for making these transformations clearly introduces some uncertainty, firstly because the S factors are unlikely to be an accurate representation of the non-linear response of near-surface soils in the Groningen field. The second source of uncertainty is because the proposed procedure for making the transformation assumes applicability of the unmodified Akkar *et al*. (2014a) GMPE for calculating PGA_{REF}, the PGA value in the reference rock condition (*i.e*., V_{S30} = 750 m/s). As can be appreciated from Figure 4.4, a significant proportion of the database is from sites with V_{S30} values considerably larger than 200 m/s; there are a few records from sites with V_{S30} above 1,000 m/s. and the

hardest contributing site has an average shear-wave velocity of 2,000 m/s. The uncertainties associated with the transformation to a site velocity of 200 m/s can be expected to increase with the difference between the site and target V_{S30} values. For this reason, although all sites will be retained in the initial screening of the database, the V_{S30} values associated with the records will be tracked in order to allow some subsequent filtering on the basis of this parameter.

- **Style-of-faulting**. For the Groningen field, it is assumed that all earthquakes have either normal or strike-slip rupture mechanisms, as was noted in Section 2.3. The intention is to build a GMPE that does not include style-of-fault as an explicit parameter, but which rather is conditioned to be applicable to some 'average' of these two mechanisms (noting that many GMPEs predict comparable amplitudes of ground motion from these two styles-of-faulting). Records from earthquakes in the European database that are classified as having normal or strike-slip ruptures will be used without adjustment; those from reverse-faulting earthquakes will be adjusted to the average of strike-slip and normal using the factors previously discussed and presented in Table 2.1. Reverse-faulting earthquakes are the least well-represented of the three styles-of-faulting in the European database (Figure 4.5), so this adjustment is only applied to a relatively modest proportion of the records. However, it is considered worthwhile tracking the records from reverse events, in case it is subsequently decided that it would be desirable to reduce the influence of such adjustments. In passing it may be noted that normal-faulting earthquakes, which may be the dominant mechanism in the Groningen field, are poorly represented in the NGA-West2 database (Figure 4.6) that was discussed above as a potential alternative to the European strong-motion database; using that database would have required far more adjustments for style-of-faulting to be made.
- **Focal depth**. The current understanding is that the Groningen earthquakes are occurring—or at least originating—inside the gas reservoir, which is encountered at a depth of about 3 km in the Groningen field and is about 300 m in thickness. Imposing such a restriction on the focal depths of the European strong-motion records would effectively wipe out most of the available data, as can be appreciated from Figure 4.5. However, clearly focal depth is not an issue that can be simply neglected in selecting suitable data to constrain the larger magnitude events, especially in view of the decision to base the GMPEs on horizontal distance metrics (R_{epi} and R_{JB}). Consideration should also be given to the fact that the determination of focal depth is the most uncertain of all source parameters and significant errors may be associated with the some of the values reported in the database (including, for example, those shown as being equal to zero in Figure 4.5). From Figure 4.5 it is also easily discernible that several focal depths have been assigned fixed values (5 or 10 km). There is evidence for focal depths from shallow earthquakes being, on average, lower than those from deeper crustal events (*e.g.*, EPRI, 2006; Allen, 2012), even if this is not a universally accepted premise. Comparing intensities observed during induced and tectonic earthquakes in the Central and Eastern United States, Hough (2014) concluded that the induced events, by virtue of their shallower foci, are associated with lower stress drops. However, she also concluded that in the epicentral region—which is the key concern here—the

motions from the two types of events are comparable since the opposite effects of shorter travel paths and lower stress drops lead to comparable levels of motion. On this basis, a reasonable approach for this application is to remove the records from earthquakes in the lower crust, but not to aim to capture only the very shallowest earthquakes (other than through scaling to a focal depth of 3 km).

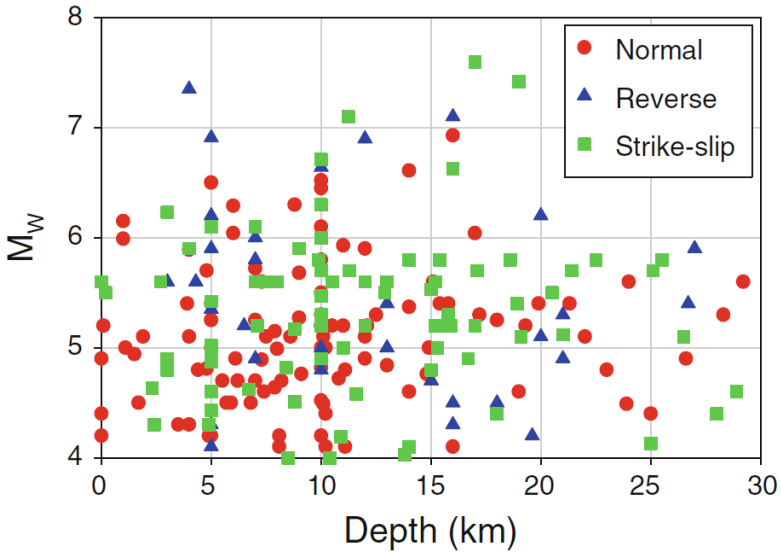


Figure 4.5. Magnitude-focal depth distribution of the earthquakes in the Akkar *et al*. (2014a) database, with symbols reflecting the style-of-faulting.

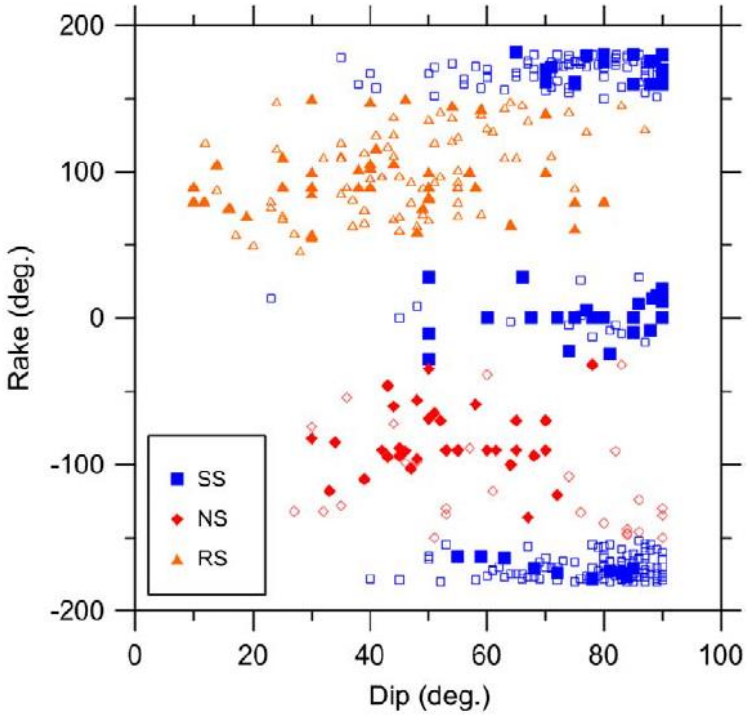


Figure 4.6. Rake and dip distribution of earthquakes in the NGA-West2 database, indicating the style-of-faulting (SS: strike-slip, NS: normal, RS: reverse); solid symbols are those in the NGA-West1 database, open symbols those added for NGA-West2 (Ancheta *et al*., 2014).

On the basis of the above considerations, a subset of the European database is extracted and the accelerations adjusted to be consistent with the assumed parameters of the Groningen data (*i.e*., normal or strike-slip ruptures, focal depth 3 km, V_{S30} 200 m/s). The characteristics of this reduced dataset and the procedures to be applied for their transformation to Groningen conditions are summarised in Section 4.4.

4.2. Additional small-magnitude recordings

In addition to the database of European strong-motion recordings used in the derivation of GMPEs such as those of Akkar *et al*. (2014a) and Bindi *et al*. (2014), additional databases were obtained from small-to-moderate magnitude earthquakes in Europe. To be considered useful for this application, only earthquakes with known style-of-faulting and moment magnitude estimates were considered; moreover, only recordings from stations with measured V_{S30} were included. The datasets come from France, Switzerland and Italy, and their distributions with respect to the basic parameters of interest are shown in Figures 4.7, 4.8 and 4.9 respectively. In view of the application, the earthquakes are distinguished in two groups, reverse-faulting events and others, and the recording sites are grouped into V_{S30} ranges that indicate the extent of the transformation that would be required to estimate the equivalent spectral accelerations on a site with V_{S30} of 200 m/s. The datasets contain the following numbers of recordings: 8 records from France, 47 from Switzerland and 116 from Italy.

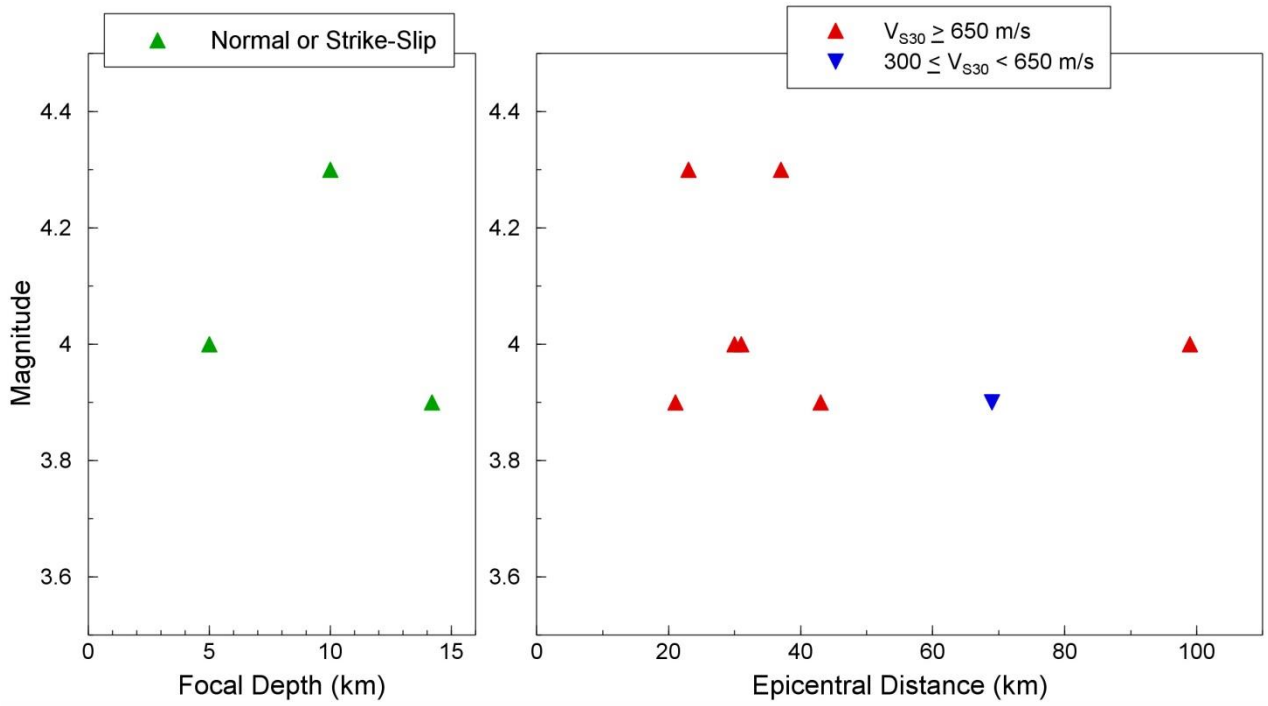


Figure 4.7. Distribution of the additional recordings from France, in terms of magnitude, depth and style-of-faulting (*left*) and magnitude, epicentral distance and V_{S30} (*right*)

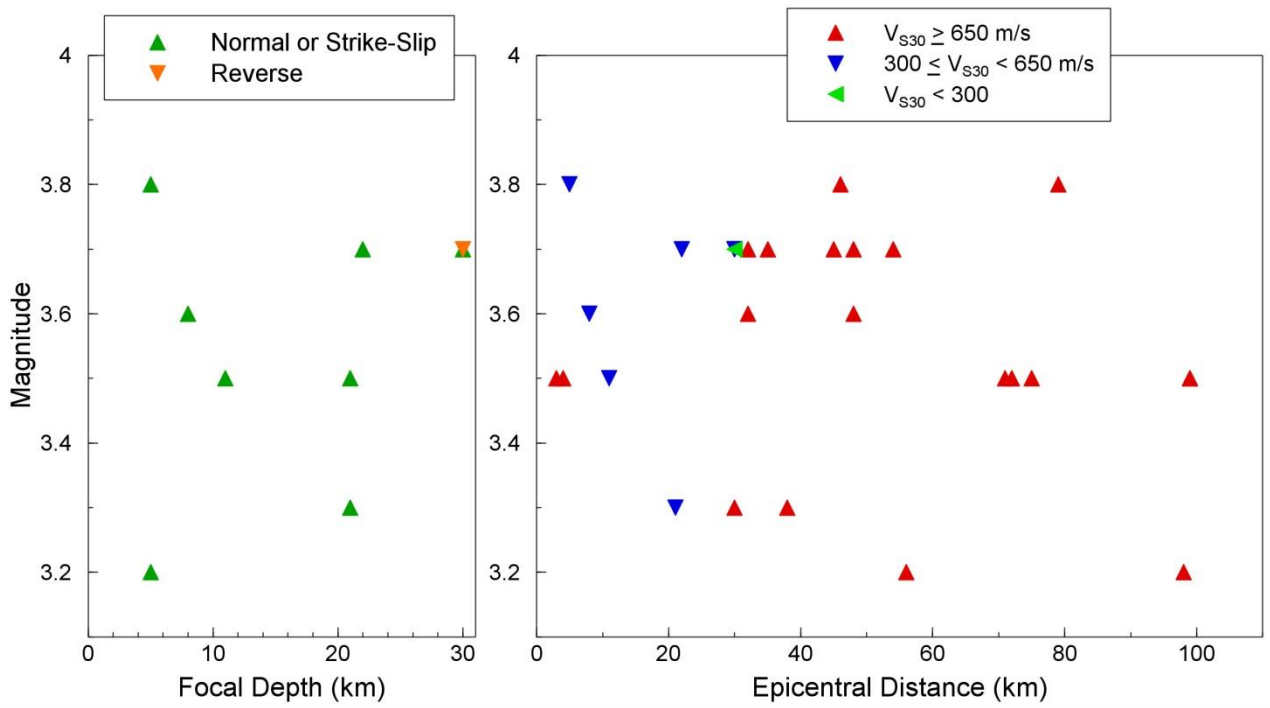


Figure 4.8. Distribution of the additional recordings from Switzerland, in terms of magnitude, depth and style-of-faulting (*left*) and magnitude, epicentral distance and V_{S30} (*right*)

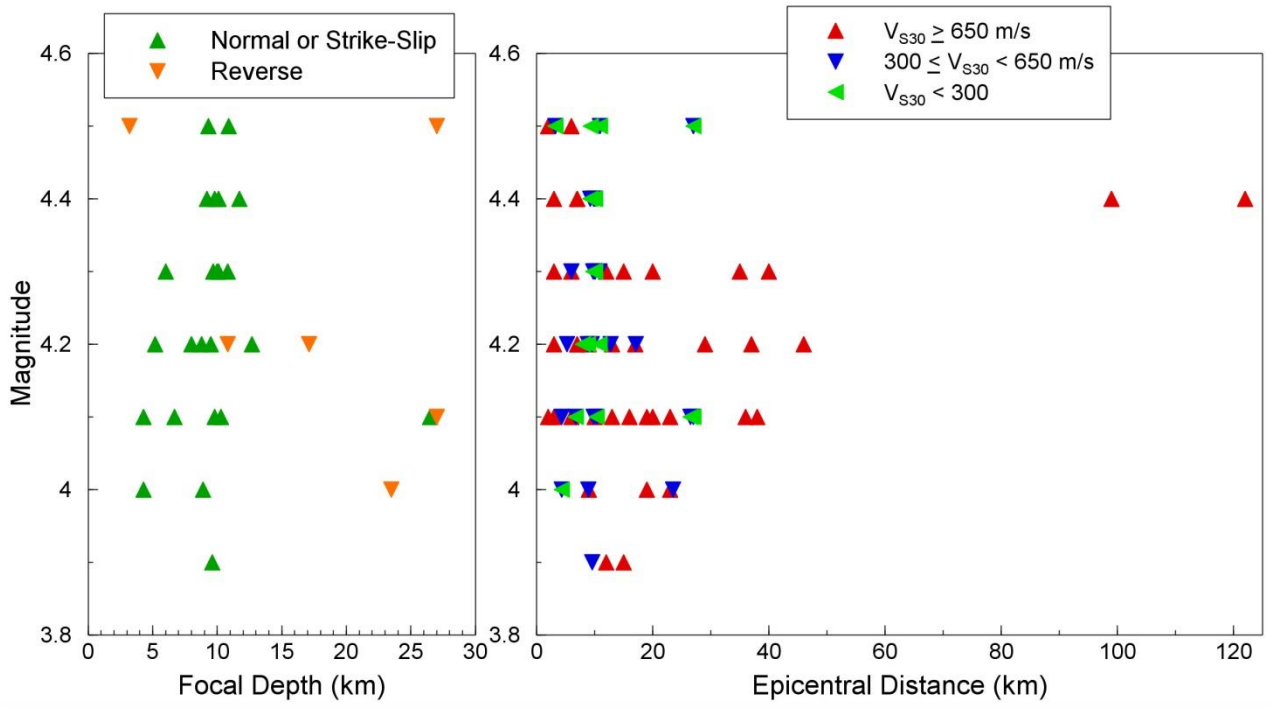


Figure 4.9. Distribution of the additional recordings from Italy, in terms of magnitude, depth and style-of-faulting (*left*) and magnitude, epicentral distance and V_{S30} (*right*)

The distributions of these datasets are worthy of brief consideration in terms of how well they correspond to the conditions in the Groningen field. The records from France (Figure 4.7) are all from very hard sites (V_{S30} > 1,000 m/s) with the exception of a single recording

obtained beyond the maximum distance of interest (60 km). None of the recordings has been obtained at a distance of less than 20 km, which for the small-magnitude range is where appreciable contributions to the computed hazard and risk may be expected.

The recordings from Switzerland are also predominantly from hard sites, with just over half the records in the uppermost V_{S30} category: 4 of the records in this site class correspond to about 700 m/s V_{S30}, the rest from sites with 30-m shear-wave velocities over 1,000 m/s, even reaching values in excess of 3,000 m/s. There is a single record from a soft site (V_{S30} 275 m/s) but this is from an earthquake with focal depth of 30 km. If earthquakes deeper than 15 km and recordings obtained either beyond 60 km or on hard sites are excluded, just 10 of the 47 Swiss records remain (from one normal and one strike-slip earthquake, with magnitudes of 3.5 and 3.6, and depths of 11 and 8 km, respectively). However, since these events are in the same magnitude range as that covered by the Groningen data, these data are not useful helping to bridge the extension to larger magnitudes.

Applying the same three exclusionary criteria to the more abundant Italian dataset results in a reduction from 116 to 54 records from 24 earthquakes, two of which have reverse mechanisms. This reduced dataset includes 14 records from events with focal depths less than 7 km and 22 records obtained at epicentral distances of less than 20 km. This would therefore appear to be the most promising of the three additional datasets.

The final decision regarding whether, and if so how, to make use of these small-magnitude datasets from France and Italy is conditioned on what they could add to the existing European database in the range of explanatory variables of relevance to the Groningen hazard and risk models.

4.3. Final database for GMPE development

As noted in the introduction to this chapter, one option for developing the new GMPEs would have been to perform regressions through a dataset that combines the Groningen data at small magnitudes and the European database at larger magnitudes, with a functional form that applies across the whole magnitude range. A potential shortcoming of such an approach is that it would lead to a relatively poor fit to the Groningen data, in order to also obtain a reasonable fit to the larger-magnitude recordings, although to some extent this could be addressed by assigning proportionally higher weighting to the local data. Another issue to consider is whether the combination of the datasets would lead to any discontinuity in the magnitude range defining the boundary between the two datasets. For these reasons, the choice is made to use stochastic simulations to extend the GMPEs to magnitudes above the upper limit of the current Groningen database, as explained in Chapters 5 and 6. However, the European data provide a useful tool for evaluating these extrapolations of models based on the Groningen data and specifically for informing the logic-tree weights to be assigned to the alternative models (Section 6.3). Therefore, the

first step is to extract from the existing European database records that are consistent with the target ranges of the Groningen models, or at least amenable to suitable adjustments:

1. Earthquakes with moment magnitude **M** ≤ 6.5
2. Earthquake with focal depth ≤ 15 km
3. If the style-of-faulting is Reverse, then the spectral accelerations are transformed to the expected equivalent corresponding to an average of strike-slip and normal earthquakes, using the factors listed in Table 2.1
4. If the focal depth is greater than 3 km (which is generally the case), the R_{hyp}-based model of Akkar *et al.* (2014a) will be used to calculate the ratio between the median motion expected on reference rock for the magnitude and distance combination of the record and the accelerations that would be expected if the depth were instead 3 km. This ratio can then be used to transform the observed acceleration. This is a slightly conservative approach since it effectively assumes that there is no variation of stress drop with depth but given the constant near-source saturation terms of 7.5 in the GMPE, the impact will be modest for recordings at short epicentral distances and almost negligible at more distant sites.
5. Finally, using the R_{hyp}-based model of Akkar *et al.* (2014a) once again, but with the distance calculated using the actual epicentral distance and a depth of 3 km, ratios are calculated of the expected motions on the V_{S30} of the site and the target value of 200 m/s, which effectively requires division by one S-factor to transform the motions to reference rock and then a second S-factor to transform the rock motion to V_{S30} of 200 m/s.

Figures 4.10 to 4.12 show the distributions of the European database and the additional small-magnitude recordings after application of the magnitude and distance limits (steps 1 and 2 in the list above), indicating the distributions with respect to the parameters considered in steps 3-5. The lower magnitude limit of the European database is **M** 4, as has been noted previously; for the additional datasets, recordings of events of **M** 3.6 and smaller are removed since the focus is the extension of the model beyond the current limit of the Groningen data. The Swiss recording stations on sites with V_{S30} in excess of 3,000 m/s are also removed since these are so far beyond the limit of applicability of the Akkar *et al*. (2014a) site adjustment factors. As it turns out, this eliminates all of the Swiss data, so the additional European records are predominantly those from Italy with a small number from France. Inspection of these plots shows that the additional data clearly improve the **M**-R_{epi} distribution at the low magnitude range and extend the lower limit from 4.0 to 3.9. The additional data are predominantly not from reverse-faulting earthquakes, which are known to be rather poorly represented in the European database. Although Figure 4.10 may suggest that data from reverse-faulting earthquakes could be discarded to avoid the uncertainty associated with the adjustments, the factors to be applied (Table 2.1) are not sufficiently different from unity for this to be a major concern. Similarly Figure 4.11 might lead one to consider applying a more severe focal depth cut-off but there would only be a strong motivation for doing this at short distances, where the data is sufficiently sparse to justify retaining the full dataset displayed.

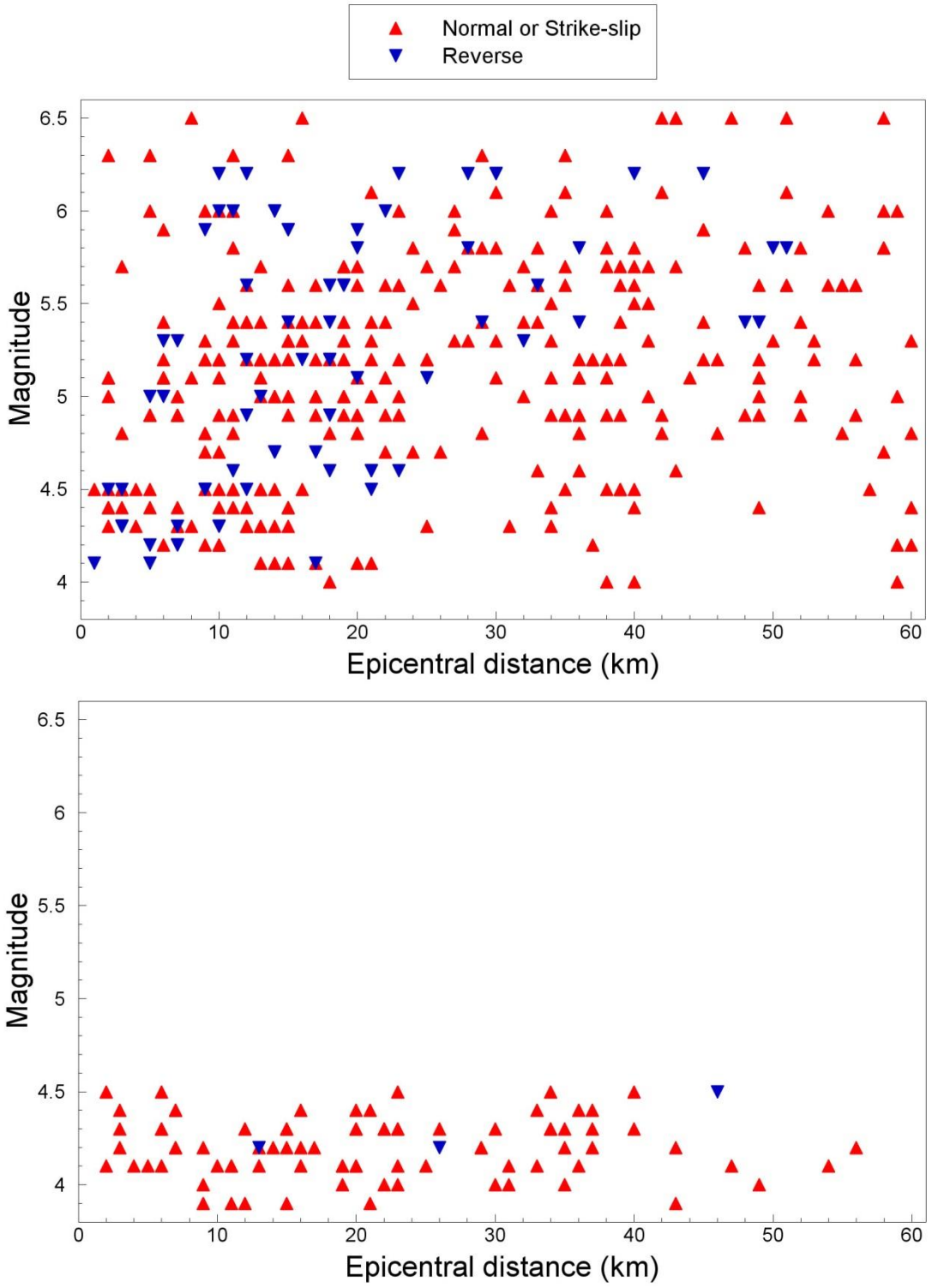


Figure 4.10. Magnitude-distance distribution of the European strong-motion database indicating the style-of-faulting of the contributing earthquakes, for the updated Akkar *et al.* (2014a) database (*upper*) and the additional small-magnitude data from Italy and France (*lower*)

On the basis of the distribution shown in Figure 4.12, one might also consider it beneficial to remove records from sites with V_{S30} very different from the value of 200 m/s currently assumed to be applicable across the Groningen field. The basis for such a measure would be to remove records that appear to require the largest transformation. However, given

that a generic pan-European site amplification function will be applied for the adjustment—and to date no investigation has been made to ascertain how well this might represent the site amplification factors in the Groningen field—the records from softer sites (which are perhaps more likely to have pronounced amplification effects) are potentially more uncertain.

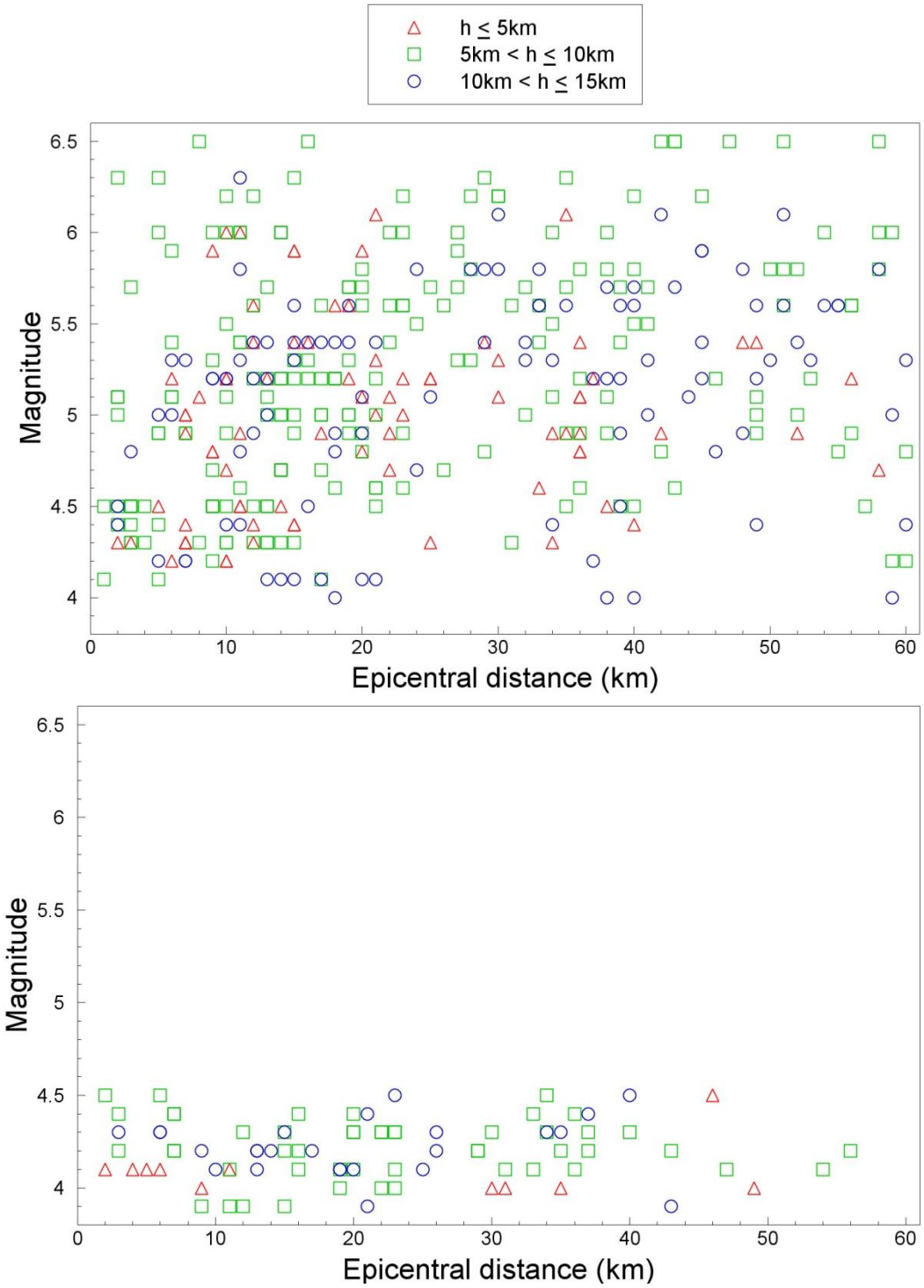


Figure 4.11. Magnitude-distance distribution of the European strong-motion database indicating the focal depths of the contributing earthquakes, for the updated Akkar *et al.* (2014a) database (*upper*) and the additional small-magnitude data from Italy and France (*lower*)

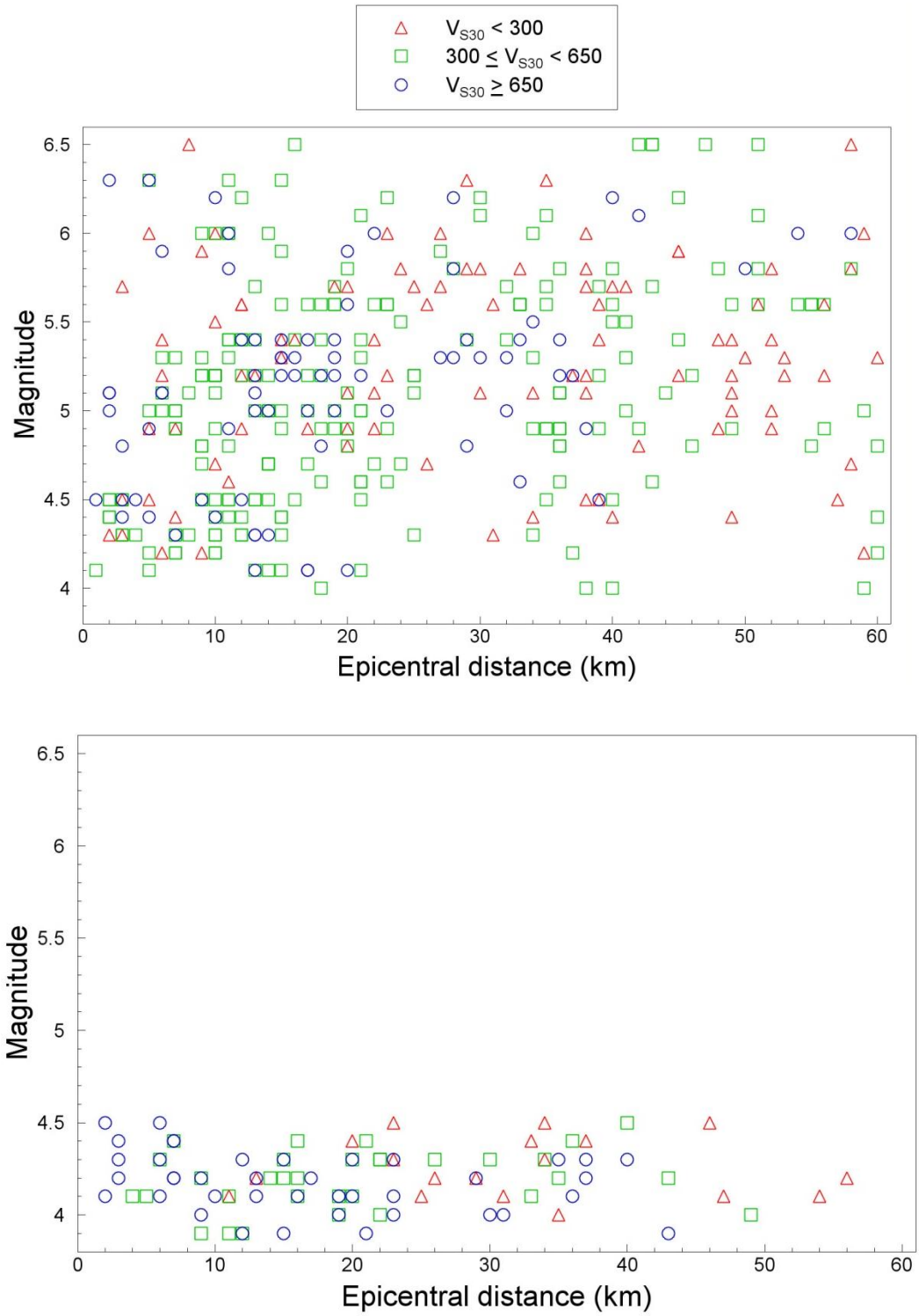


Figure 4.12. Magnitude-distance distribution of the European strong-motion database indicating the V_{S30} values of the recording sites, for the updated Akkar *et al.* (2014a) database (*upper*) and the additional small-magnitude data from Italy and France (*lower*)

Therefore, in conclusion, at this stage no other filters are applied to the dataset and the records that will be used in exploring the extrapolation of the equations to larger magnitudes will be those shown in both plots of Figures 4.10 to 4.12.

5. Exploratory Analyses and Functional Form

For the 2013 Winningsplan, Bommer & Dost (2014) derived GMPEs for PGA and PGV by adjusting selected GMPEs derived for tectonic earthquakes to be consistent with the limited database of recordings from the Groningen field in the lower magnitude range. For the Version 1 GMPEs a fundamentally different approach is adopted whereby a model is developed that in the first instance is designed to provide a good fit to the now expanded Groningen database (Chapter 3) and this model will then be extrapolated, with adjustments as necessary, to larger magnitudes. This extrapolation will be guided by existing GMPEs, seismological theory, stochastic simulations and also the European database discussed in Chapter 4. Given the inevitably very large epistemic uncertainty in the ground-motion predictions for magnitudes of **M** 4 and greater, it is likely that multiple options for the extrapolations will be considered (Figure 5.1).

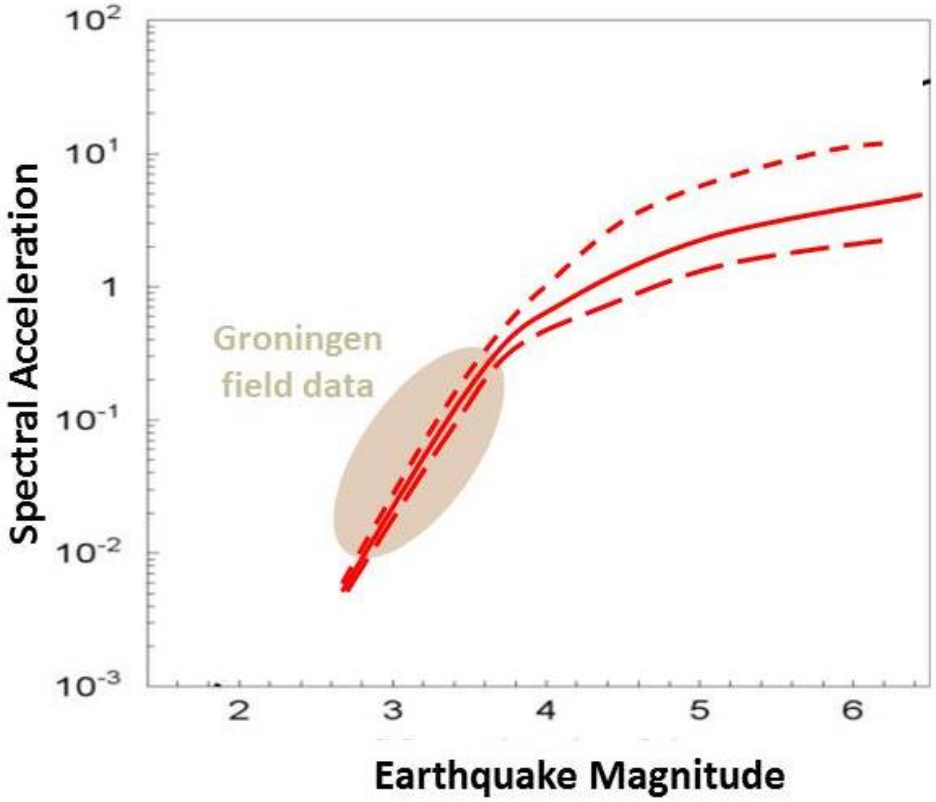


Figure 5.1. Schematic illustration of the process of fitting equations to the small-magnitude Groningen data and generating extrapolations of the predictions to larger magnitudes. The multiple extrapolations are intended to capture the epistemic uncertainty that inevitably with increase with the earthquake magnitude.

In this Chapter, a suitable functional form for the Groningen data is explored, bearing in mind that the records cover a very narrow range of magnitudes. Theoretical and empirical considerations for the extrapolation of this functional form to larger magnitudes are then discussed to guide the final functional form.

5.1. Functional form to fit the Groningen recordings

Examination of the plots of within-event residuals against distance for the Groningen data with respect to European GMPEs derived from recordings of tectonic earthquakes reveal a strong negative slope in the residuals over the first few kilometres from the epicentre (Figures 3.14-19). This suggests that the amplitudes of the ground motions decay very rapidly over short distances, which is also consistent with the very strong inverse relationship observed between amplitudes and durations of the motion (Figure 5.2.). The pattern observed in this plot is consistent with the hypothesis of a strong influence of refraction and reflection by the high-velocity layers immediately above the gas reservoir. Travel paths that are close to vertical will lead to signals at the surface with relatively high amplitudes, and short durations consistent with the magnitude of these events; ray paths exiting the source at larger take-off angles will undergo repeated reflections and refractions, leading to elongated signals of very low amplitude.

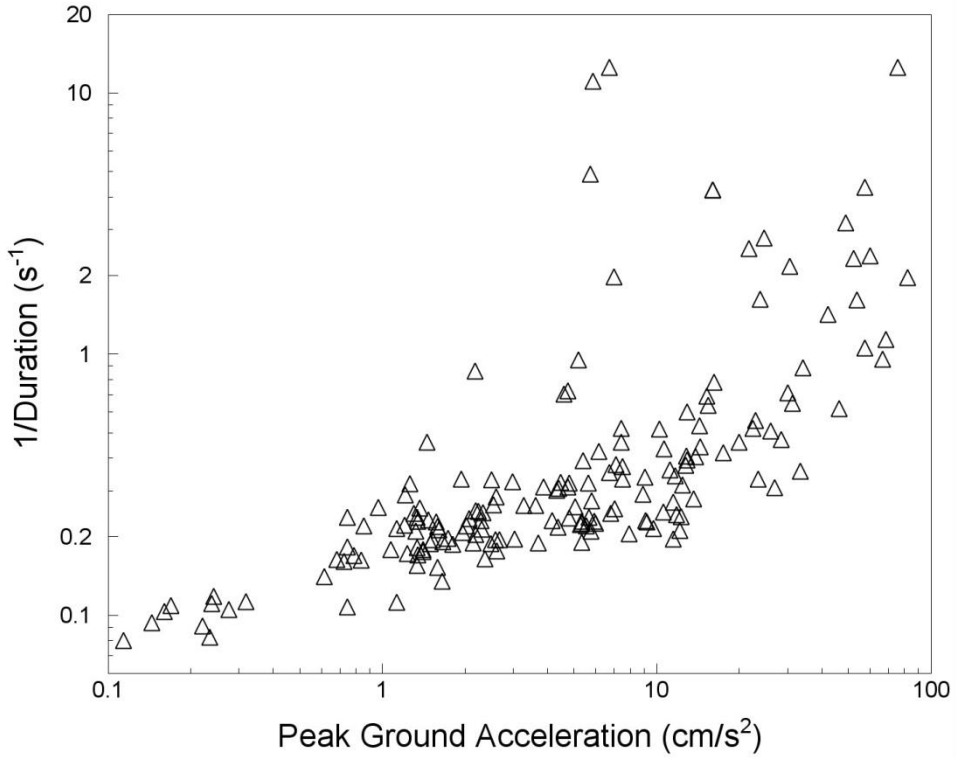


Figure 5.2. Plot of PGA values and reciprocal values of significant durations (5-75% of the total Arias intensity) for the Groningen database

The misfit of the Groningen spectral accelerations at very short distances, as seen in the residual plots, was interpreted in Chapter 3 as being primarily the consequence of the large constant values—at any given response period—of the distance included to represent near-source saturation. Even with the magnitude dependence introduced into this term in the Version 0 GMPE, the pattern persisted; however, as noted by Bommer & Dost (2014), the constraint that the distance had to be equal to 7.5 km at **M** 4.2 to avoid a discontinuity in the equation meant that the values were always larger than desired. With

this is mind, a starting point for considering a suitable adjustment could be the model of Yenier & Atkinson (2014) for the near-source saturation distance, *h*, which was proposed for more realistic models from point-source simulations:

$$\log_{10}(h) = 0.43M - 1.72 \qquad (5.1)$$

The equation, which has an associated standard deviation of 0.19, was derived from empirical data and is applicable to earthquakes of **M** 6 and larger. It is also noteworthy that the value of *h* was proposed for use in conjunction with R_{rup} by Yenier & Atkinson (2014). However, it may be considered as a starting point for the exploratory analysis. Figure 5.3 shows the median values predicted by Eq.(5.1), which for magnitudes of less than 6 is being extrapolated below its strict lower limit of applicability. Also in the Figure is the bi-linear relationship included in the Version 0 GMPE, which was constrained to be equal to 0.5 km at **M** 1.5 and was obliged to converge to 7.5 km—the value in the Akkar *et al*. (2014a) GMPE—at the hinge magnitude of **M** 4.2, above which the European GMPE was used without modification. The large values of *h* that resulted in the magnitude range 2.6-3.6 contribute to the poor behaviour of the residuals at short distances (Figure 3.14).

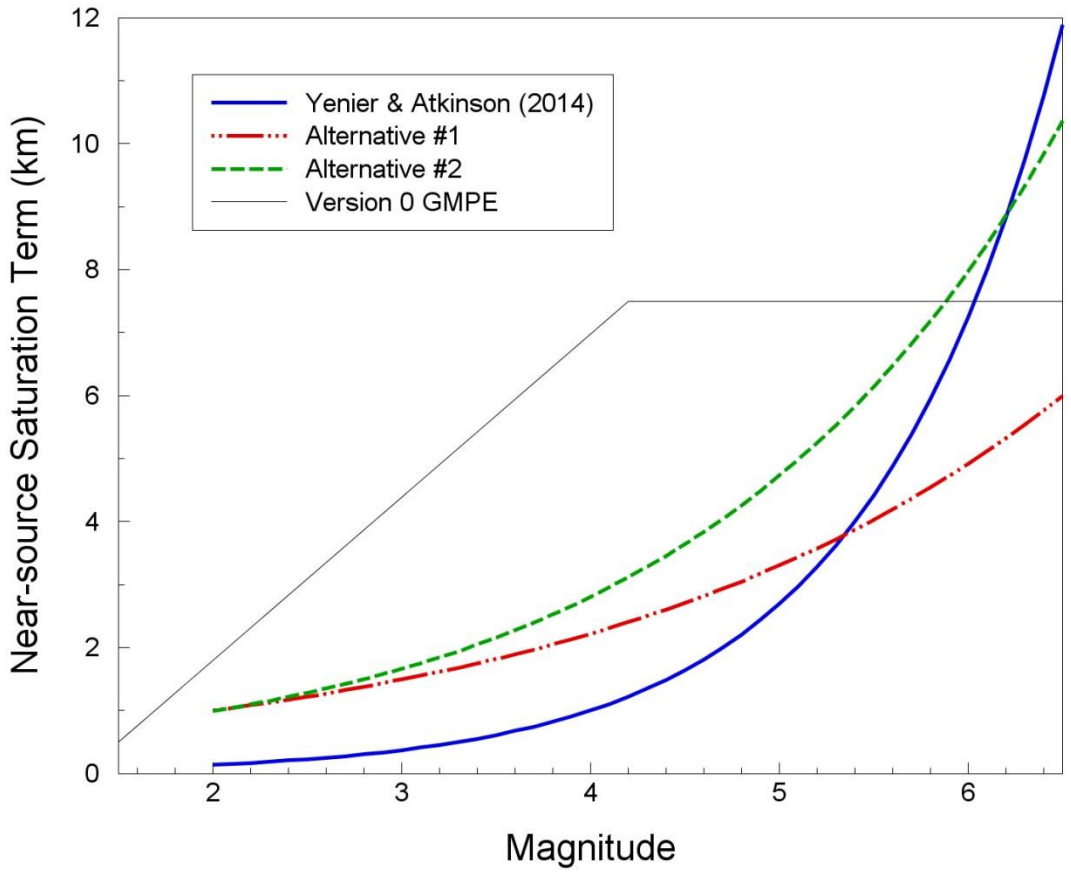


Figure 5.3. Comparison of the near-source saturation term proposed by Yenier & Atkinson (2014) with that used in the Version 0 GMPE and two alternative models proposed for the purposes of exploratory analyses. Note that below **M** 6, the Yenier & Atkinson (2014) model is being extrapolated beyond its intended range of applicability.

To explore the sensitivity to using alternative near-source saturation terms, two alternative relationships were derived, both arbitrarily constrained to be equal to 1 km at **M** 2, the first constrained to yield a value of 6 km at **M** 6.5 (since the shallow focal depths may result in less near-source saturation than for tectonic events):

$$\log_{10}(h) = 0.1729M - 0.3458 \tag{5.2}$$

The second was constrained to give better approximation to the Yenier & Atkinson (2014) relationship in the magnitude range 6-6.5 (taking into account that the standard deviation in that relationship is equal to a factor of 1.55 on *h*):

$$\log_{10}(h) = 0.2271M - 0.4607 \tag{5.3}$$

With this magnitude-dependence in the distance term, and taking account of the fact that we are only interested in rather short distances (max. 60 km), the multiplier on the geometric spreading term in the equation can now be a simple scalar. Moreover, the magnitude interval is too small to constrain anything other than linear scaling, so the functional form is:

$$\ln(Y) = c_1 + c_2 M + c_4 \ln\sqrt{R_{epi}^2 + h^2} \tag{5.4}$$

Expressing the distance saturation term algebraically, this becomes:

$$\ln(Y) = c_1 + c_2 M + c_4 \ln\sqrt{R_{epi}^2 + [\exp(c_5 M + c_6)]^2} \tag{5.5}$$

For fixed values of c_4, c_5 and c_6, the problem is reduced to a linear least squares regression. Using each of the three near-source saturation terms—Eq.(5.1) to Eq.(5.3)—good fit to data, in terms of total residuals not showing any trend or offset with respect to either magnitude or distance, were found with c_4 values of -1.4, -1.5 and -1.6 respectively, yielding the following three alternative equations:

$$\ln(PGA) = -3.0429 + 2.3048M - 1.4\ln\sqrt{R_{epi}^2 + [\exp(0.99M - 3.96)]^2} \tag{5.6}$$

$$\ln(PGA) = -2.6752 + 2.2651M - 1.5\ln\sqrt{R_{epi}^2 + [\exp(0.398M - 0.796)]^2} \tag{5.7}$$

$$\ln(PGA) = -2.7617 + 2.3844M - 1.6\ln\sqrt{R_{epi}^2 + [\exp(0.523M - 1.061)]^2} \tag{5.8}$$

Figure 5.4 compares the total residuals for the current Groningen database (85 records from 12 earthquakes) plotted against magnitude and distance. The residuals display successively improved performance of these three equations, with narrower distributions in each case. The residuals associated with three records at distances ~3 km with large

positive residuals—therefore implying under-prediction by the models—do not reduce appreciably, however, as the overall residual distributions get better. These records may simply represent exceptionally high-amplitude recordings (*i.e*., large positive epsilon values); this is supported by the fact that in all three cases the E-W components of these three accelerograms are associated with large PGA values (> 50 cm/s^2) and very short durations (< 1 second). Since the final goal is vector predictions of amplitudes and durations, these excursions can be handled without having a disproportionate impact on the estimated risk.

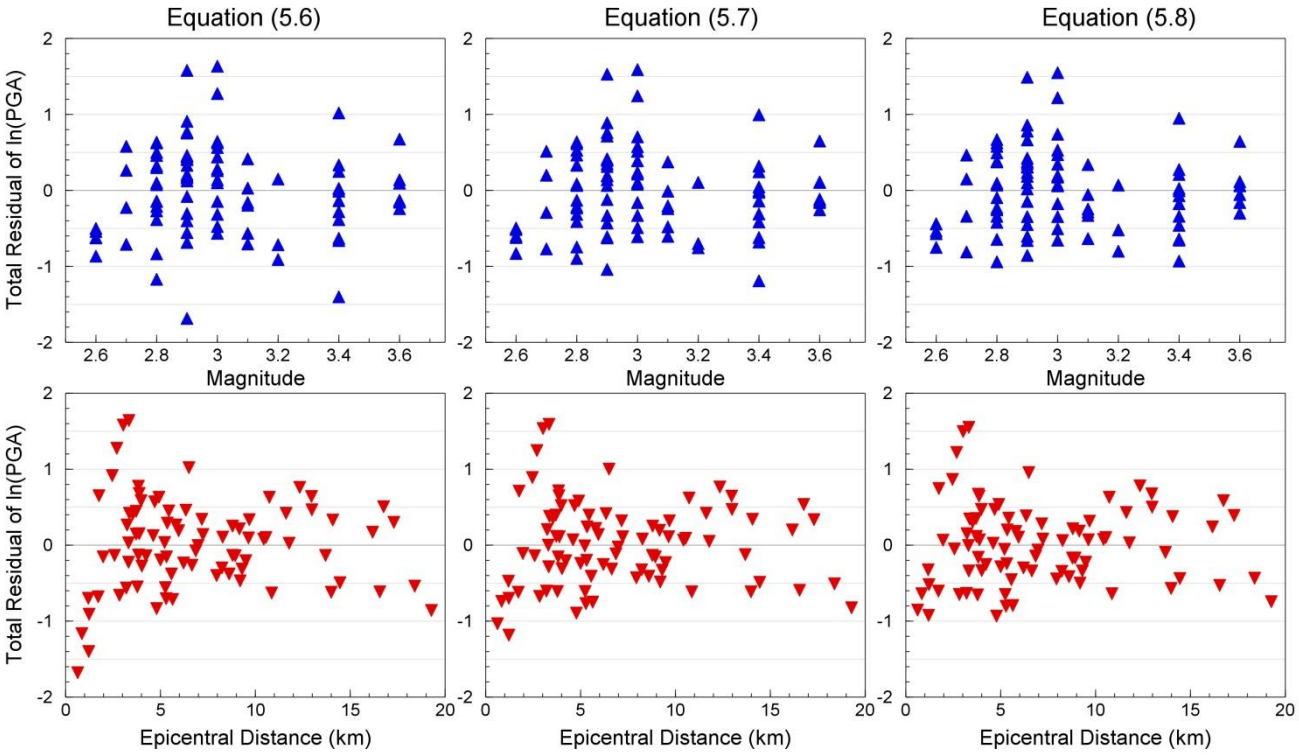


Figure 5.4. Total residuals of the natural logarithms of the 85 geometric mean PGA values from the Groningen dataset obtained with three different simple fits to the data, plotted against magnitude (*upper row*) and epicentral distance (*lower row*)

Figure 5.5 compares the predictions of median PGA values from the three equations with those obtained from the Akkar *et al*. (2014a) equation, setting V_{S30} to 200 m/s and using the average of the coefficients for normal and strike-slip earthquakes for the latter. The predicted values are plotted against magnitude for various distances. Several observations can be made, including the obvious and expected feature that for larger magnitudes, the simple equations presented above predict absurdly high values of PGA due to the linear scaling of amplitudes with magnitude. Clearly, the extension of the model to magnitudes above 3.6 would need to introduce a non-linear scaling term to avoid this effect.

The next interesting observation is that the differences among the predicted values from three different equations from the Groningen data are only large at very short distances. In the epicentral region, the predicted accelerations are strongly dependent on the near-

source saturation value, although the impact of increasing the *h* value diminishes as this parameter grows. At larger distances (just a few kilometres from the epicentre), the differences are confined to the predictions for larger magnitudes and reflect the differences in the *h* values: for Eq.(5.8), in which the *h* value at **M** 6 was fixed to be of the same order as that in the Yenier & Atkinson (2014) model embedded in Eq.(5.6), the predictions are rather similar.

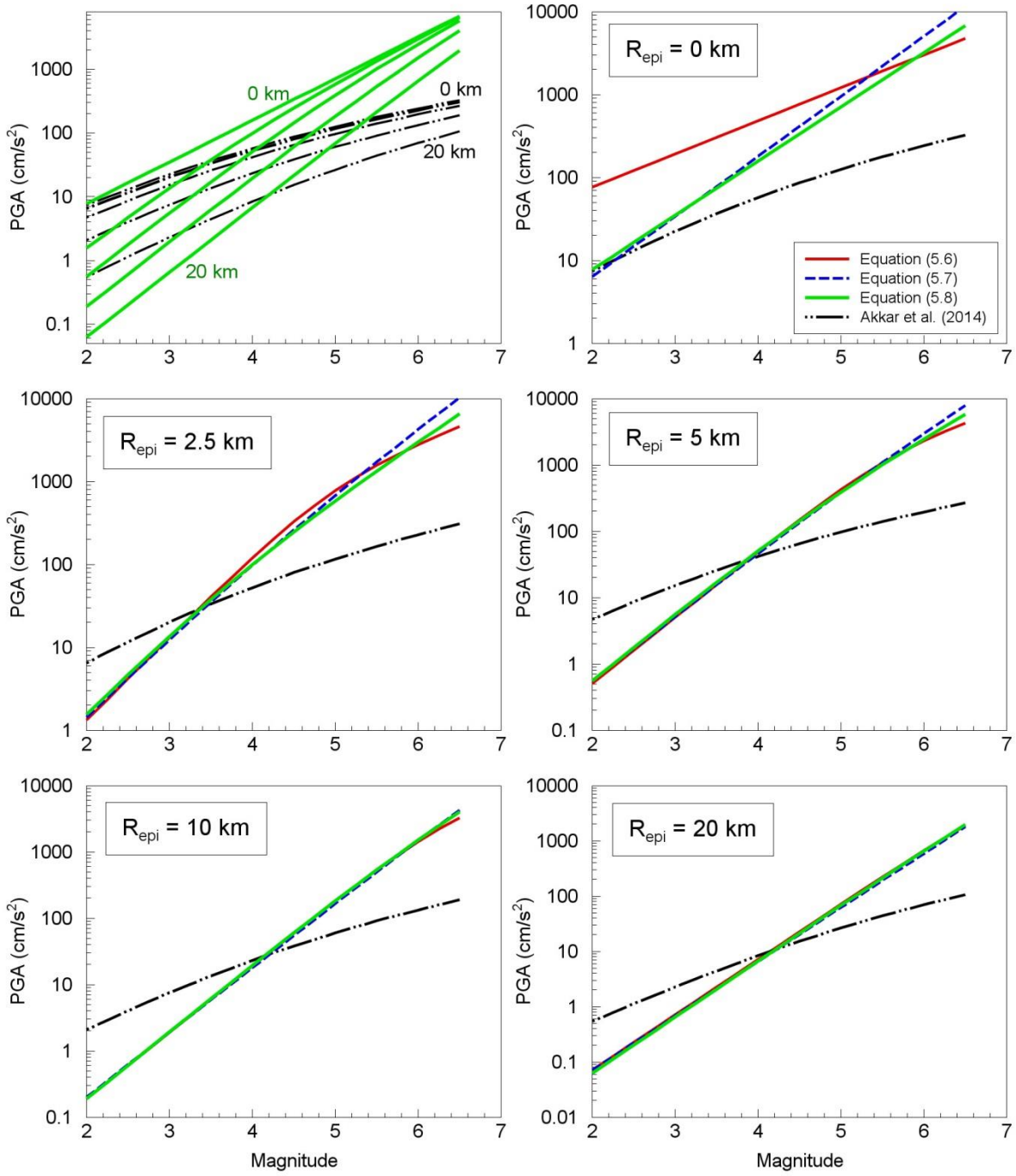


Figure 5.5. Predicted median PGA values for different magnitude and distance combinations from Eqs.(5.6) to (5.8) and from Akkar *et al*. (2014a), using V_{S30} = 200 m/s and the average of normal and strike-slip faulting for the latter

The other feature of Figure 5.5 worthy of brief comment is the relative amplitudes of predicted PGA values from the Groningen equations and from the European model of Akkar *et al.* (2014a). This comparison is best made considering only one of the Groningen equations, for which Eq.(5.8) seems the logical choice in view of the residual plots. The top left-hand frame of Figure 5.5 compares the median predictions from this equation with those from Akkar *et al*. (2014a). A number of observations can be made from this comparison, including the fact that the Groningen data indicate more rapid attenuation over short distances from the epicentre. General tendency of the Akkar *et al*. (2014a) model to predict larger values at small magnitudes—except at very short distances where the near-source saturation term overrides this feature—may be a result of the general tendency of empirical GMPEs to lead to over-prediction when extrapolated beyond their lower magnitude limit (Bommer *et al*., 2007; Atkinson & Morrison, 2009; Chiou *et al*., 2010), although this effect, as has been seen previously, would be somewhat less pronounced—for PGA—were the comparison made with Bindi *et al*. (2014). The steep slope of the magnitude scaling inferred from the Groningen data is also likely to be strongly influenced by the kappa effect of the soft soils encountered in the field, as discussed in Section 5.4.

5.2. Empirical GMPEs derived from Groningen data

On the basis of the considerations in the previous section, the chosen functional form for the Groningen data is that presented as Eq.(5.5), repeated here for ease of reference:

$$\ln(Y) = c_1 + c_2 M + c_4 \ln\sqrt{R_{epi}^2 + [\exp(c_5 M + c_6)]^2} \qquad (5.9)$$

Rather than using the model of Yenier & Atkinson (2014)—which after all was not intended for application to small-magnitude earthquakes—or either of the arbitrarily-chosen alternatives depicted in Figure 5.3, it was decided to determine the magnitude-dependence of the near-source saturation term as part of the fitting. Random effects maximum likelihood regressions (Pinheiro & Bates, 2004) were applied to estimate the coefficient c_1, c_2, c_4, c_5 and c_6 for each of the 5 ground-motion parameters. The results obtained indicated markedly different behaviour from that modelled by Eq.(5.3), which was judged to the best estimate of coefficients c_1 and c_2 in the exploratory analyses. Moreover, for the three longest response periods, the behaviour modelled by the values obtained for c_5 and c_6 is actually unphysical and contrary to all expectations (Figure 5.6).

In response to this finding, it was decided to perform a more complex regression across the five periods to obtain the single pair of values for c_5 and c_6 that would best satisfy the data. This yielded values of 0.4233 for c_5 and -0.6083 for c_6. Figure 5.7 compares the resulting magnitude-dependent scaling of the near-source term with that of Yenier & Atkinson (2014) and that represented by Eq.(5.3). By coincidence, the final model is remarkably close to the one found, by trial and error, to provide a good fit to the PGA data.

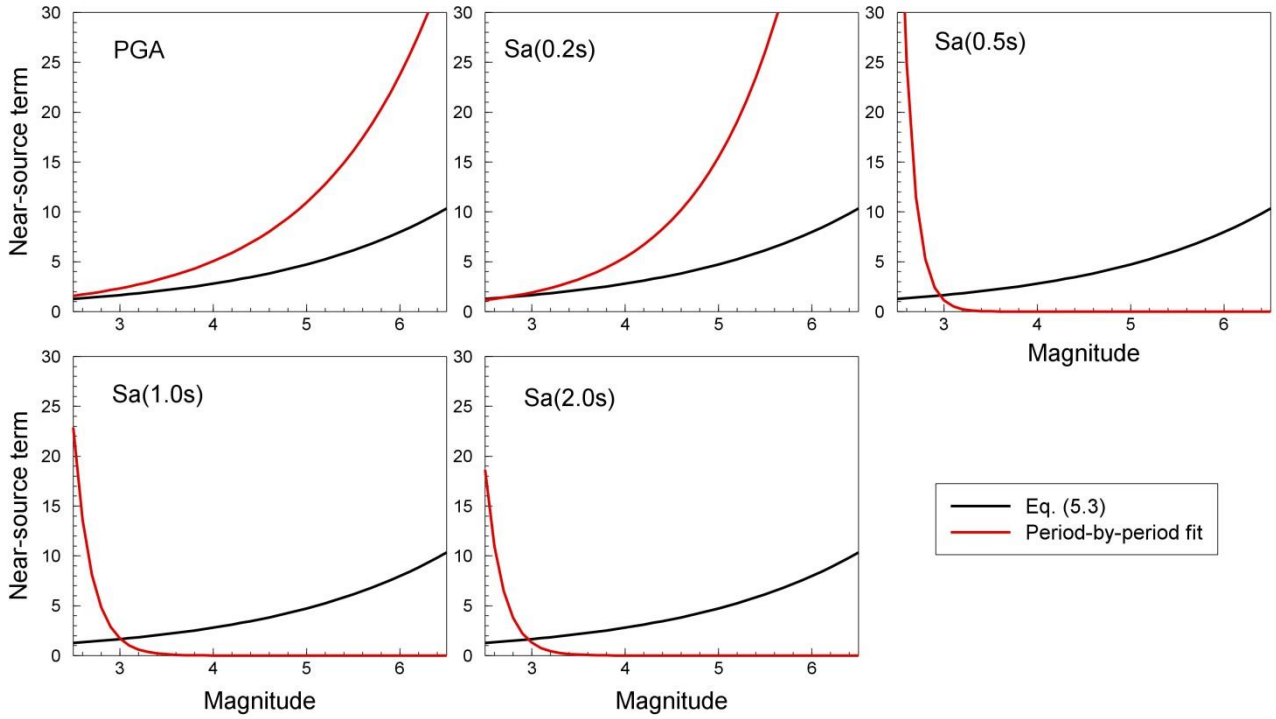


Figure 5.6. Magnitude-dependent depth terms obtained from regression of Eq.(5.9) at each period (*red*) compared with the assumed model of Eq.(5.3) (*black*)

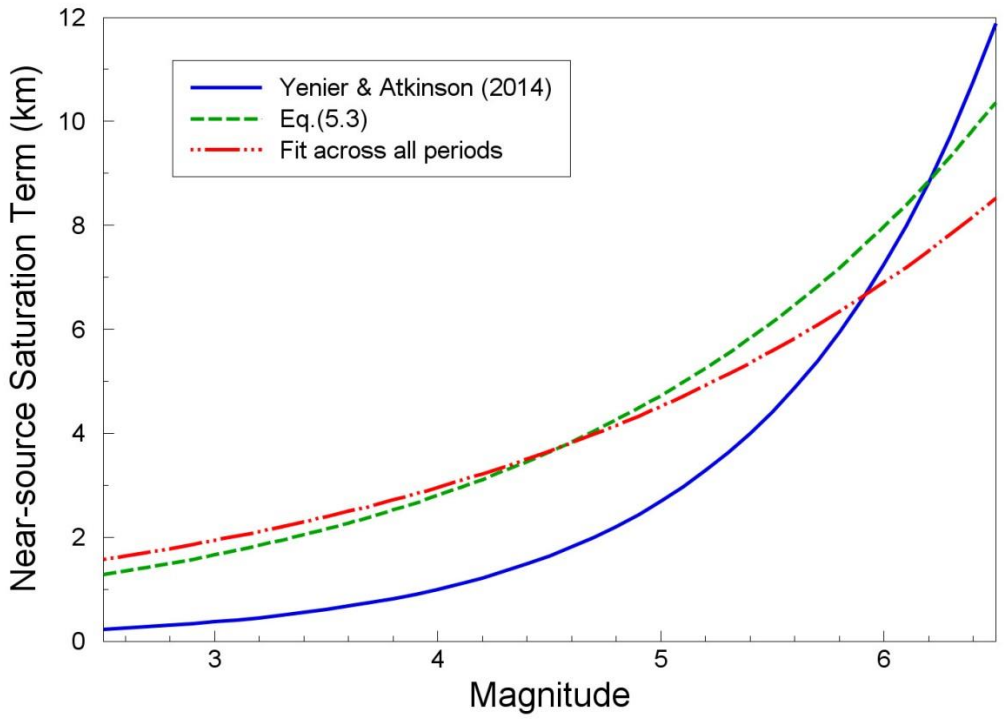


Figure 5.7. Comparison of the near-source saturation term proposed by Yenier & Atkinson (2014) with that in Eq.(5.3) and the final values obtained by regressions across all five response periods simultaneously. models proposed for the purposes of exploratory analyses. As in Fig.5.3 it should be noted that below **M** 6, the Yenier & Atkinson (2014) model is being extrapolated beyond its intended range of applicability.

These values of c_5 and c_6 were then held constant and regressions repeated using Eq.(5.9) to determine c_1, c_2 and c_4, period by period, yielding the coefficients and standard deviations recorded in Table 5.1. The first and most striking observation that can be made is that the aleatory variability appears to be rather low when compared with GMPEs such as Akkar *et al*. (2014a) and Bindi *et al*. (2014). The final sigma model is discussed in Section 6.2, but it may be noted in this stage that although these rather small standard deviations are encouraging, they have been obtained from a rather small dataset and may therefore underestimate the true variability of the Groningen ground motions. At the same time, given that the earthquakes are from a common source and the waves are travelling through very similar sub-surface materials from the reservoir to within a few tens of metres of the surface, one would expect less variability than is found for equations derived from datasets covering several regions.

Table 5.1. Coefficients and standard deviations of Eq.(5.9) obtained from regressions

	PGA	**Sa(0.2s)**	**Sa(0.5s)**	**Sa(1.0s)**	**Sa(2.0s)**
c_1	-2.8209	-3.5853	-6.2052	-7.9647	-8.8322
c_2	2.4215	2.6958	3.0228	3.0875	2.9086
c_4	-1.6621	-1.3405	-0.9669	-1.0105	-1.0770
c_5	0.4233	0.4233	0.4233	0.4233	0.4233
c_6	-0.6083	-0.6083	-0.6083	-0.6083	-0.6083
τ	0.2039	0.2514	0.2467	0.3612	0.3359
ϕ	0.4831	0.4404	0.5208	0.4121	0.4170
σ	0.5243	0.5071	0.5762	0.5479	0.5355

Figures 5.8 to 5.10 show the residuals from the regressions, plotting inter-event residuals against magnitude and intra-event residuals against epicentral distance. The plots all indicate that the regressions produce a good fit to the data, albeit that there are some apparent fluctuating trends over distance discernible in the intra-event residuals. In terms of bias, straight-line fits to the residuals in all cases indicate that there is effectively no bias at all, with the exception of the intra-event residuals at longer periods. For spectral accelerations at both 1 and 2 seconds, there is a small trend (almost identical for both parameters) in the intra-event residuals with a positive gradient. These trends indicate a bias to over-estimation of the median by a little less than 4% at 0 km distance, varying towards an under-estimation of just under 7% at 20 km from the epicentre. Although this means that some refinement of the model for the accelerations at long periods may be possible, the trends are not sufficiently strong for the models not to be considered suitable for the current stage of development.

The models represented by Eq.(5.9) and the coefficients in Table 5.1 provide a good explanation for the field data currently available. However, the dataset is small and it needs to be borne in mind that it represents only 12 earthquakes, which may or may not be representative of the actual distribution of source characteristics (such as stress drop) for Groningen earthquakes in this magnitude range. The consequence of this is that as the database of field recordings expands, it may be found that the model for the median

motions at small magnitudes shifts, up or down, and this epistemic uncertainty needs to be reflected in the ground-motion logic tree even at small magnitudes.

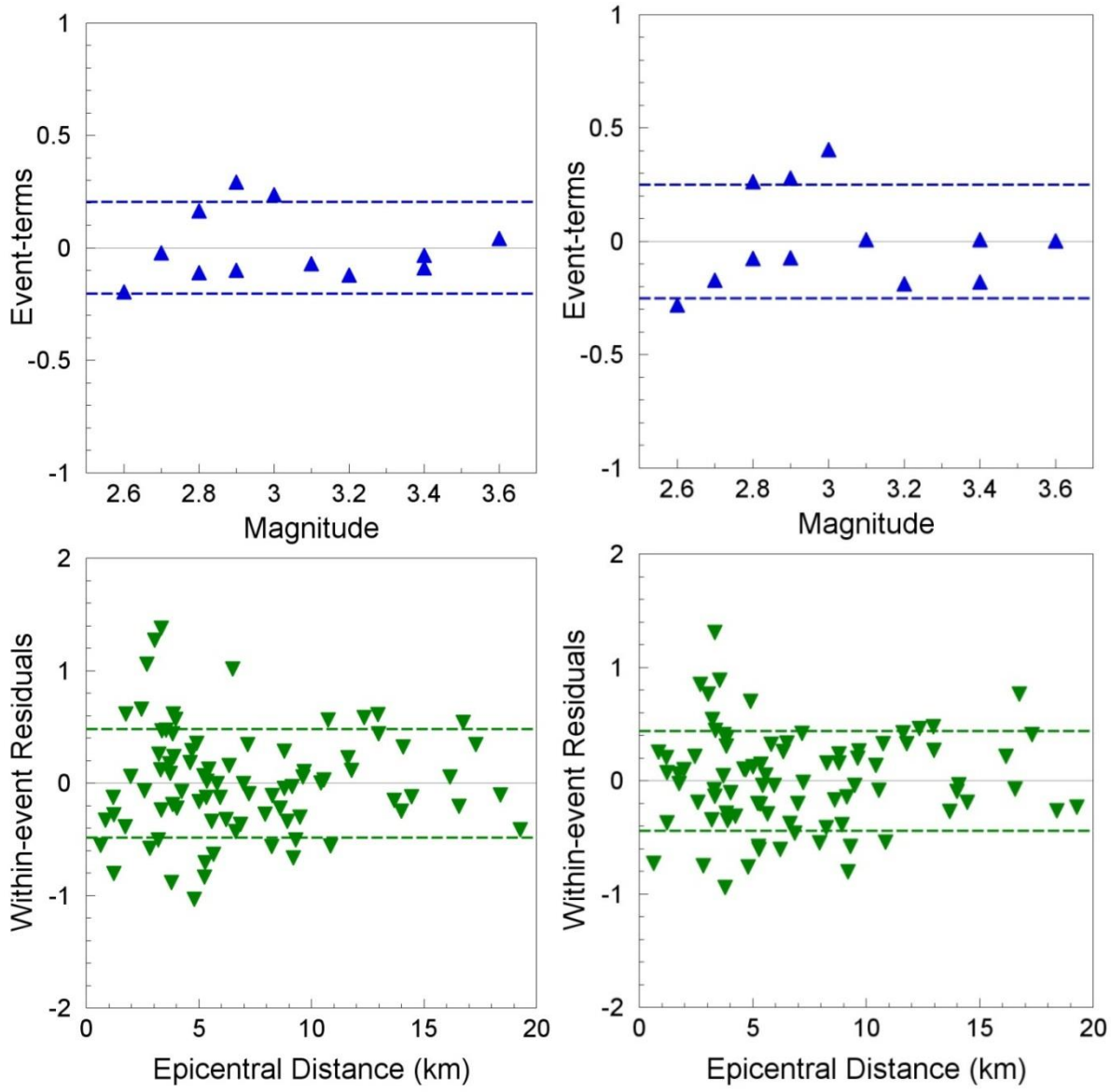


Figure 5.8. Residuals of the Groningen data with respect to Eq.(5.9) and the coefficients in Table 5.1 for PGA (*left*) and the spectral acceleration at 0.2 s (*right*); the dashed lines indicate the inter- and intra-event standard deviations, respectively, in the upper and lower plots.

The epistemic uncertainty in the median ground-motion predictions is therefore not zero in the narrow range of small magnitudes represented by the current database (**M** 2.6 to 3.6). As the predictions are extended to larger magnitudes—and ultimately all the way to **M** 6.5—the epistemic uncertainty associated with the median accelerations must inevitably increase with the degree of extrapolation. The current Groningen database is insufficient to constrain anything other than linear magnitude scaling, and as is discussed in the next section, the combination of small magnitudes and soft site conditions mean that this linear scaling is likely to be an approximation to non-linear scaling. As was already seen in

Section 5.1 and Figure 5.5, extrapolation of the small-magnitude linear trends to larger magnitudes results in very unrealistic predictions of the ground-motion amplitudes. The single most important challenge in the development of GMPEs for the Groningen hazard and risk assessments is the extension of the models to larger magnitudes.

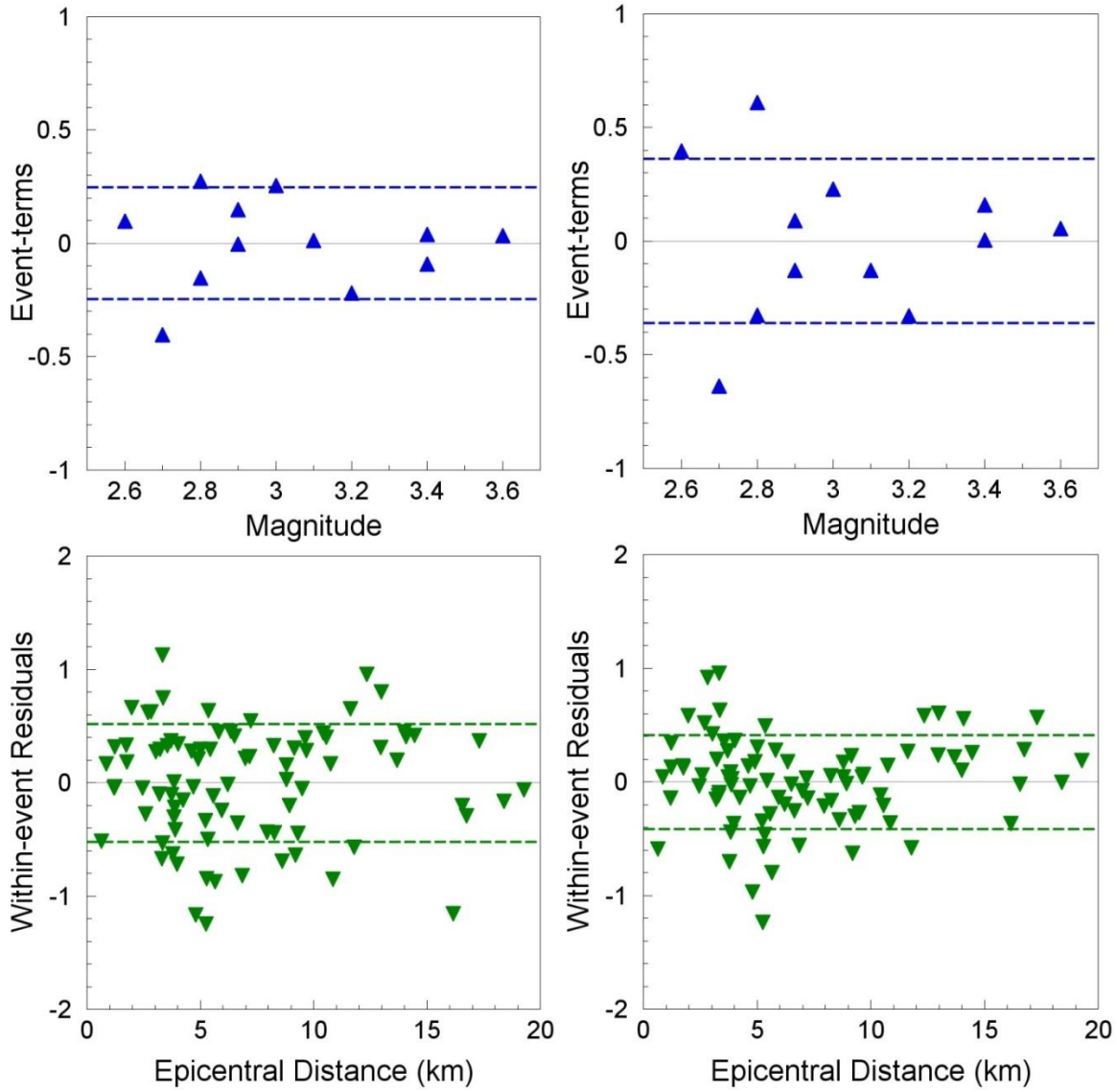


Figure 5.9. Residuals of the Groningen data with respect to Eq.(5.9) and the coefficients in Table 5.1 for the spectral accelerations at 0.5 s (*left*) and at 1.0 s (*right*); the dashed lines indicate the inter- and intra-event standard deviations, respectively, in the upper and lower plots.

Although the upper limit currently considered in these calculations is **M** 6.5, which represents a very appreciable extrapolation from the current upper limit of **M** 3.6 in the database, the hazard and risk estimates are likely to be dominated by somewhat smaller events, as has been indicated by early disaggregation studies based on the Version 0 model. This is to be expected and is a common feature observed in probabilistic seismic hazard analysis, which results from the balance between two features: the recurrence relationship indicates decreasing frequency of earthquakes with increasing magnitude (for

a *b*-value of 1, the decrease is an order of magnitude for each additional unit of magnitude, with a more rapid drop off in the vicinity of the upper limit on magnitude), whereas the accelerations increase exponentially with increasing magnitude, although the non-linearity in the magnitude scaling means that towards larger magnitudes the effect diminishes. Together these patterns render contributions from earthquakes of magnitude close to the upper limit of **M** 6.5 practically irrelevant, but earthquakes of **M** 4-5, and even a little larger, are likely to have a major impact on the estimated risk. Extrapolating the predictions into this range is therefore vitally important. Two related resources are available to guide this extrapolation, seismological theory and stochastic simulations of ground motions, which are discussed in the next two sections.

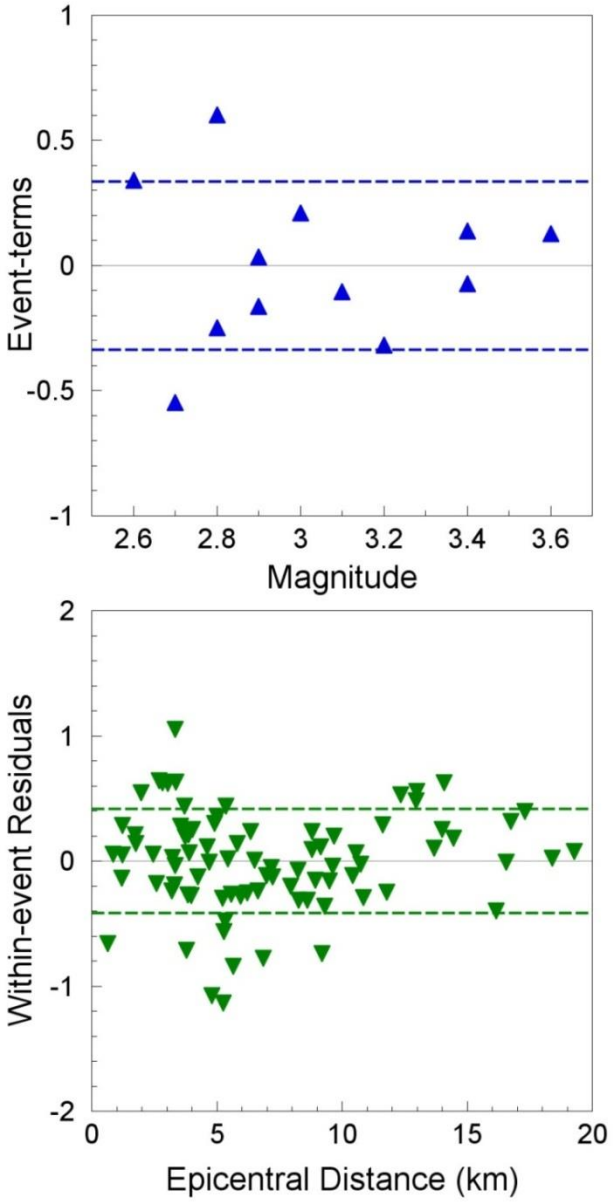


Figure 5.10. Residuals of the Groningen data with respect to Eq.(5.9) and the coefficients in Table 5.1 for the spectral accelerations at 2.0 s; the dashed lines indicate the inter- and intra-event standard deviations, respectively, in the upper and lower plots.

5.3. Theoretical considerations for the functional form

Seismological theory defines expected relationships between the amplitude of ground motions and parameters characterising the earthquake source, the travel path to the site, and the velocity profile at the site. The first point that must be highlighted in this regard is that seismological theory expresses the ground motion in terms of the Fourier amplitude spectrum, whereas our interest is in predicting the ordinates of the response spectrum. The response at most oscillator frequencies is actually closely related to the Fourier amplitude at the same frequency, except for the highest frequencies (> 10-20 Hz), as illustrated in Figure 5.11. Therefore, the distinction does need to be borne in mind but at least for the longer response periods of interest (0.5, 1.0 and 2.0 seconds), a degree of equivalence may be assumed for discussion purposes.

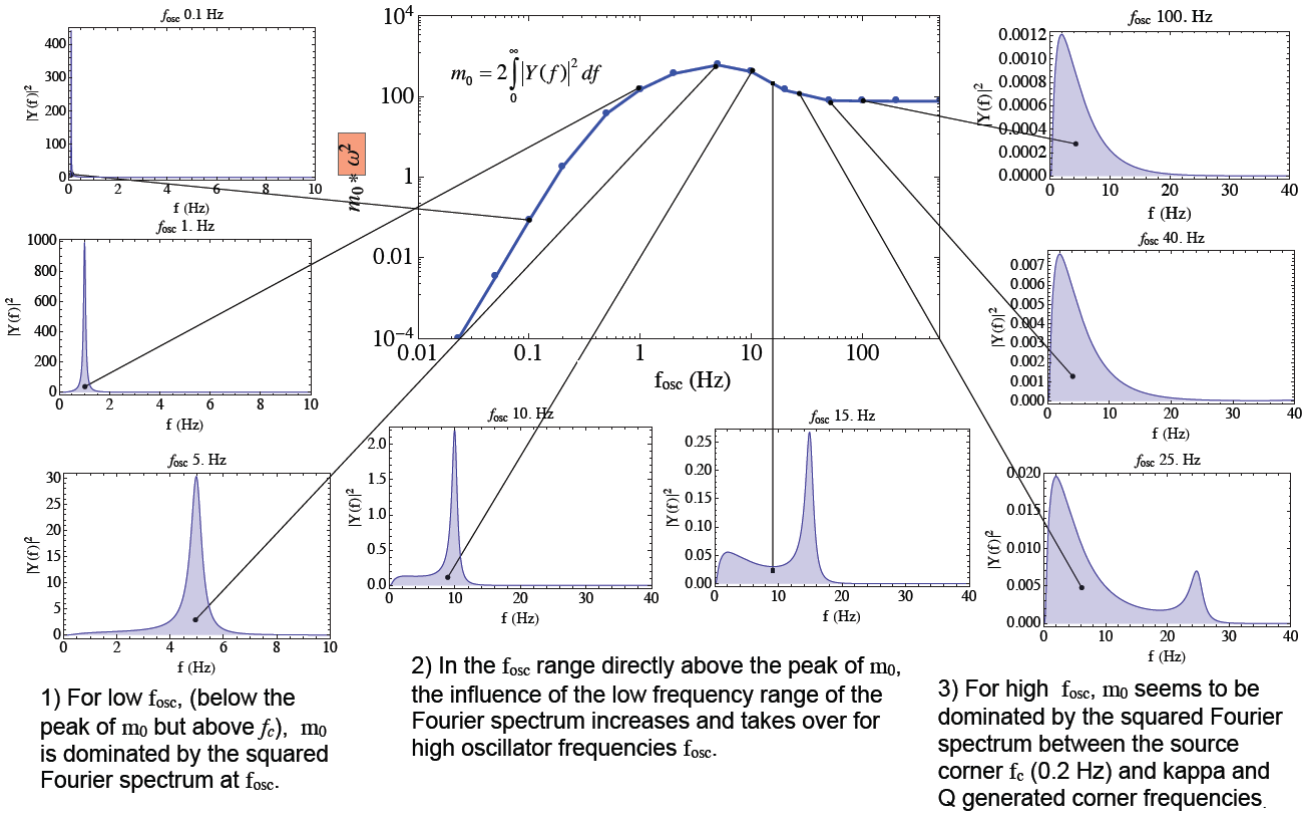


Figure 5.11. Relationship between the Fourier amplitude spectra and the acceleration response spectrum shown in the large central plot, indicating which parts of the Fourier spectrum contribute to the response at different oscillator frequencies (*Courtesy of Professor Frank Scherbaum, University of Potsdam, Germany*)

The theoretical shape of the Fourier spectrum of acceleration is illustrated in Figure 5.12, which is based on a model first proposed by Aki (1967) and the formulation for the source spectrum later developed by Brune (1970). The spectrum is a function of the seismic moment, M_0, and a parameter generally referred to as the stress drop, $\Delta\sigma$, even though it is more correctly referred to as simply the stress parameter (Atkinson & Beresnev, 1997). The shape of the spectrum is defined by the corner frequency, f_0, which is a function of M_0

and Δσ. The corner frequency is also proportional to the reciprocal of the rise time, which may be considered as the source duration or the time taken for the fault to propagate along the length of the rupture. Since for moderate to large earthquakes the velocity of rupture is generally on the order of 3 km/s, the source duration is proportional to the rupture length, which grows exponentially with the magnitude. Hence as magnitude increases, the corner frequency becomes shorter, as illustrated in the plot.

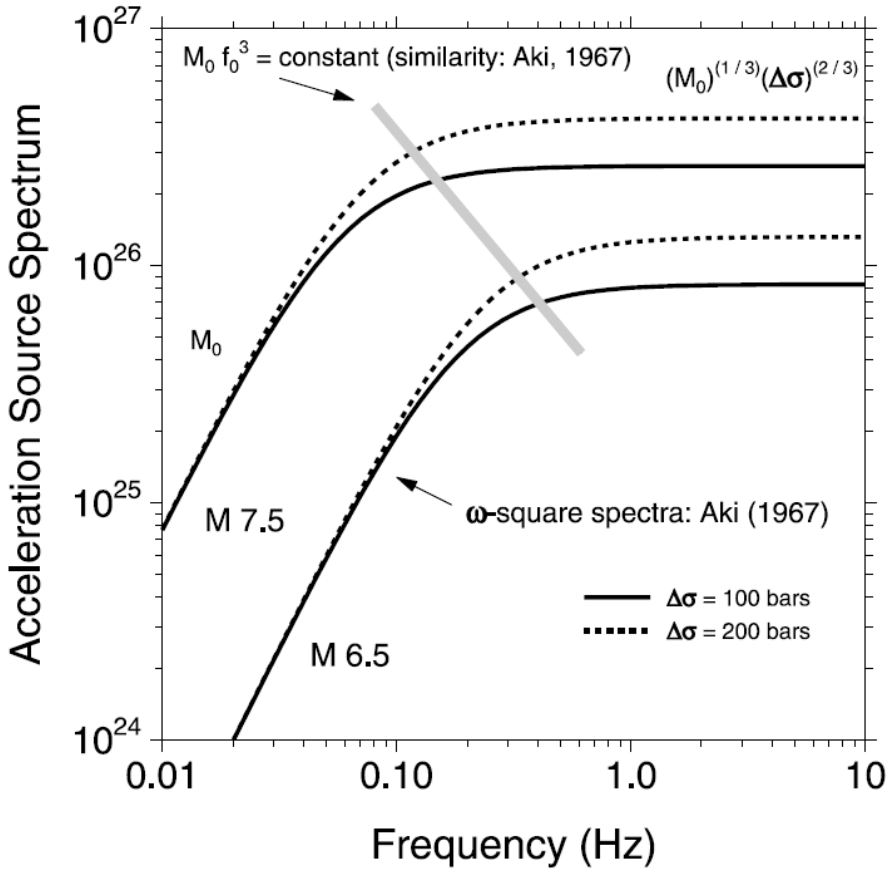


Figure 5.12. Fourier amplitude spectra of acceleration for earthquakes of **M** 6.5 and 7.5 and two values of stress drop (Boore, 2003)

Immediately apparent from this figure is the fact that the scaling of the two spectra is not constant across the frequency range, with the increase from **M** 6.5 to **M** 7.5 resulting in greater source strength at low frequencies than those frequencies above the corner frequency. Consequently, for any given frequency there will be a non-linearity in the amplitude scaling as the magnitude increases and the corner frequency moves from being above to below the frequency in question. This non-linear magnitude scaling of spectral ordinates was discussed in detail by Fukushima (1996) and is illustrated schematically in Figure 5.13. Using the formulation of Fukushima (1996), Stafford & Bommer (2012) formulated a relationship for the scaling of response spectral ordinates as function of magnitude and stress drop, considering both constant stress drop and magnitude-dependent stress drop, illustrated in Figures 5.14 and 5.15 respectively.

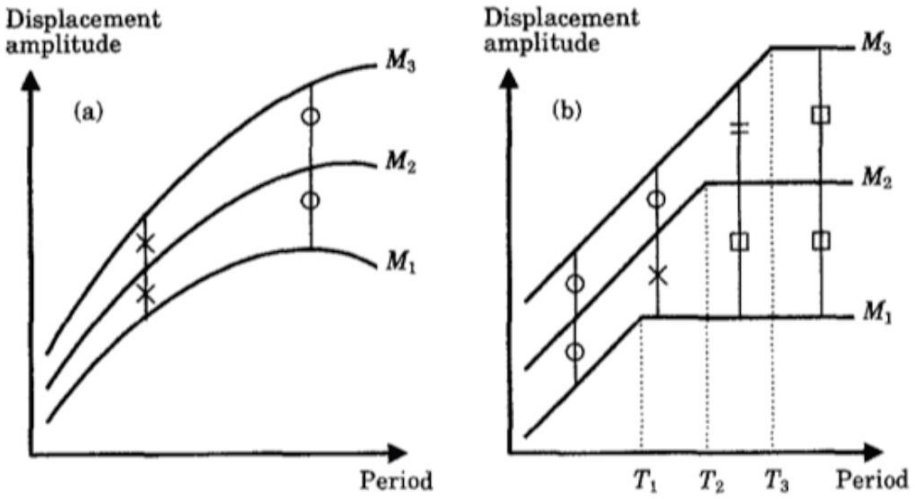


Figure 5.13. Scaling of the Fourier amplitude spectra of displacement with constant scaling (*left*) and with the non-linear magnitude scaling (*right*) that results from the magnitude dependence of the corner frequency (Fukushima, 1996)

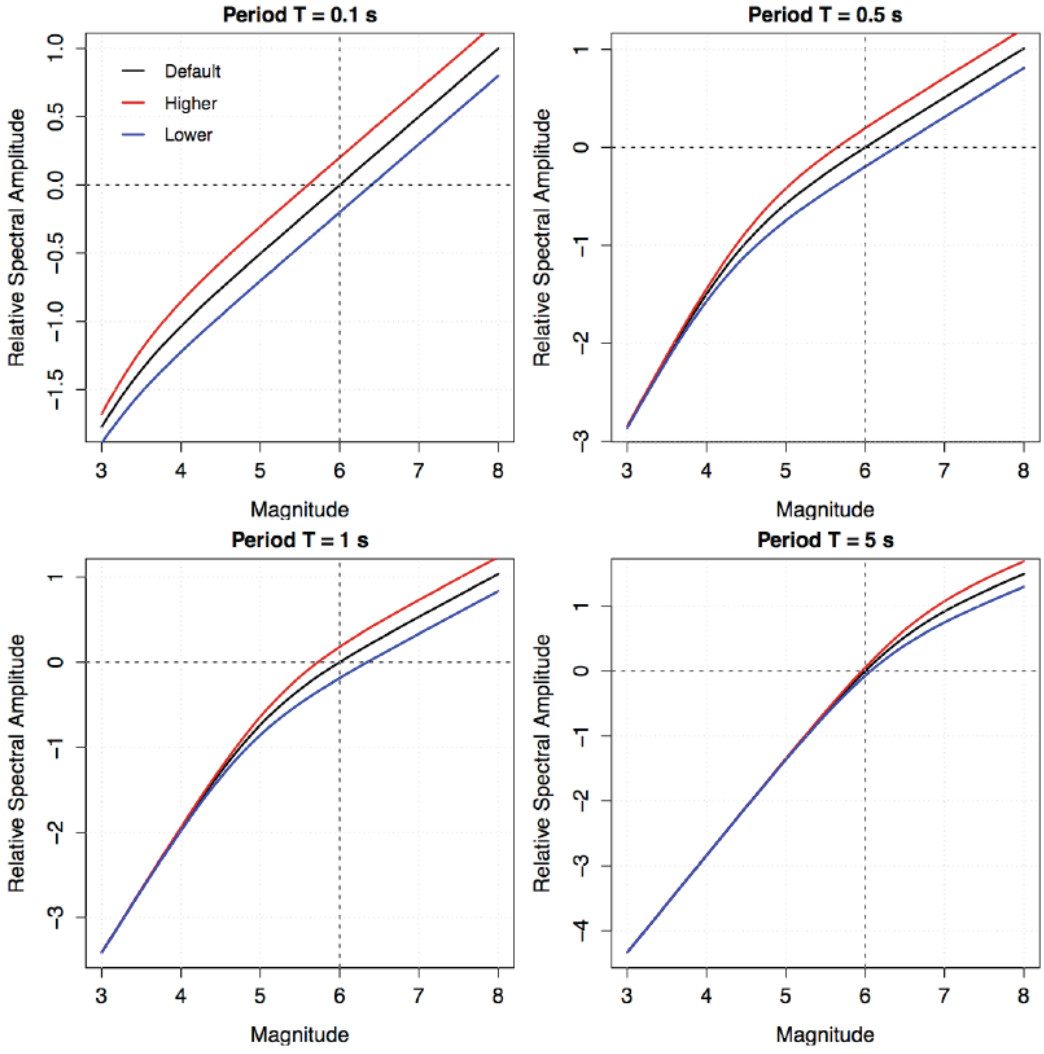


Figure 5.14. Scaling of response spectral ordinates at four oscillator periods with magnitude for different stress drops: the black line is for 30 bars and the other two lines for increasing and decreasing this value by a factor of 2 (Stafford & Bommer, 2012)

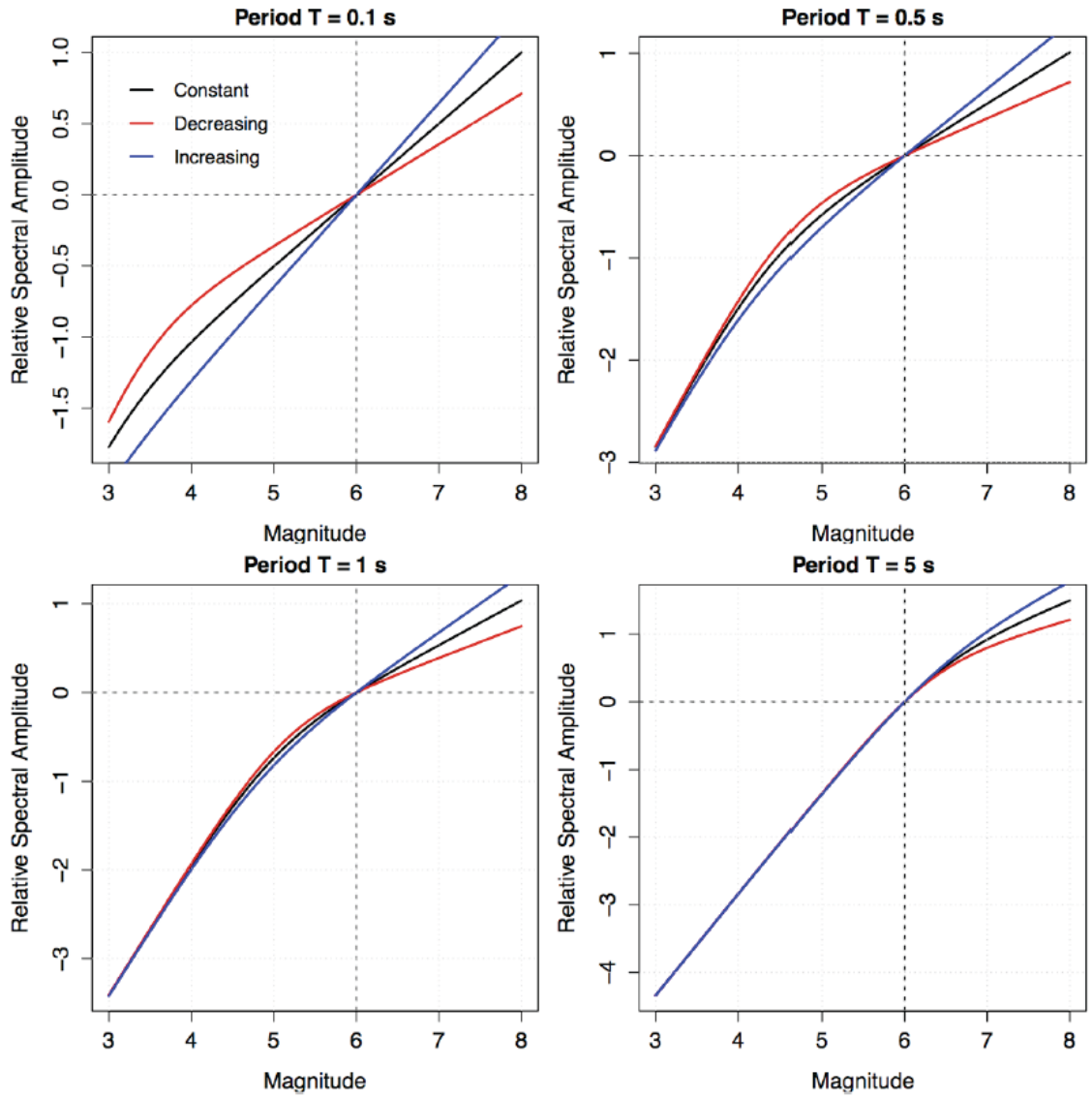


Figure 5.15. Scaling of response spectral ordinates at four oscillator periods with magnitude for different stress drops: the black line is for constant 30 bars and the other two lines for stress drops increasing or decreasing with magnitude and also have a value of 30 bars at magnitude **M** 6 (Stafford & Bommer, 2012)

Even greater non-linearity in the magnitude scaling occurs at even larger magnitudes as a result of the geometric effects of the fault rupturing across the full width of the seismogenic crust, after which the rupture will grow only in length and consequently more and more of the source of radiated energy will be remote from any point of observation close to the fault and hence attenuate sufficiently not to contribute to the motions at a given site. Figure 5.16 illustrates how this effect can be observed in ground-motion recordings; the effect, even for PGA, occurs beyond **M** 6.5 and hence this is not of direct concern to the development of the Groningen model.

All of the discussion so far relates to factors that cause the scaling of ground motions to become non-linear at higher magnitudes. There is also a reason that the scaling becomes non-linear in range of very small magnitudes, and here again it relates to the corner frequency of the Fourier amplitude spectrum. The damping by the upper part of the crust,

which is generally represented by the parameter kappa (κ; Anderson & Hough, 1984), leads to decay of the high-frequency waves (Figure 5.17).

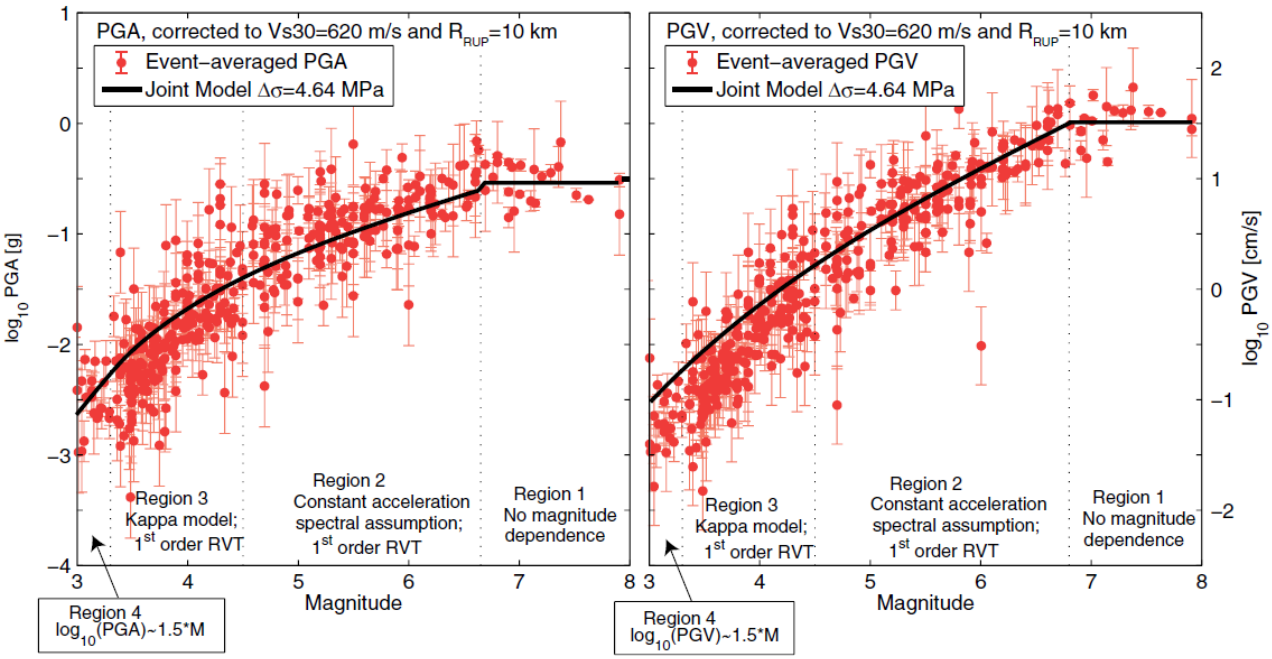


Figure 5.16. Event-averaged values of PGA (*left*) and PGV (*right*) from earthquakes in the NGA-West2 database, adjusted to 10 km distance and V_{S30} 620 m/s (Baltay & Hanks, 2014)

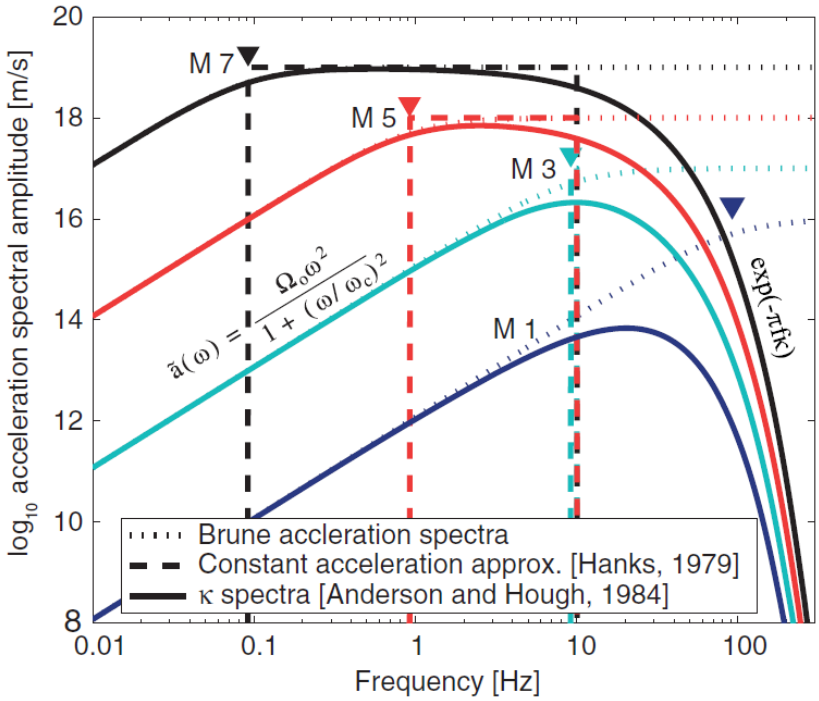


Figure 5.17. Brune Fourier spectrum for different magnitudes and a stress drop of 50 bars (*dotted lines*) and their modification by kappa (*solid lines*); the dashed lines represent another model but it is not pertinent to these discussions (Baltay & Hanks, 2014)

Therefore, where kappa is high (which will be the case for soft ground as encountered in the Groningen field), for small earthquakes the corner frequency may be located in the frequency range where the motion is filtered, hence the amplitudes will be lower than expected from linear scaling of magnitude. This results in an increased gradient of the magnitude scaling at low magnitudes, as a function of the kappa value, as shown by Douglas & Jousset (2011) and Baltay & Hanks (2014) and illustrated in Figures 5.18 and 5.19 respectively.

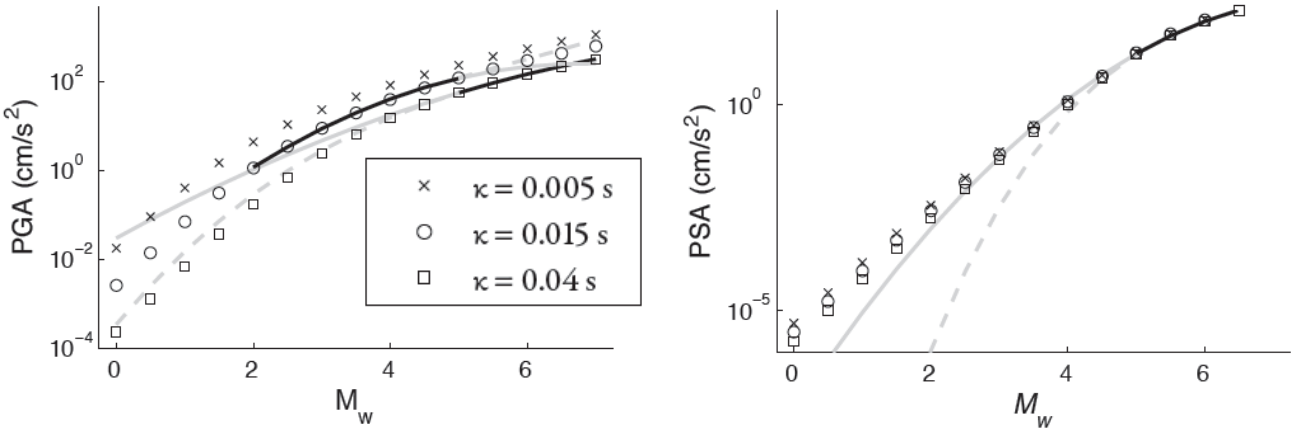


Figure 5.18. Magnitude scaling of PGA (*left*) and Sa(1.0s) (*right*) from stochastic simulations for different kappa values and a stress drop of 100 bars (Douglas & Jousset, 2011)

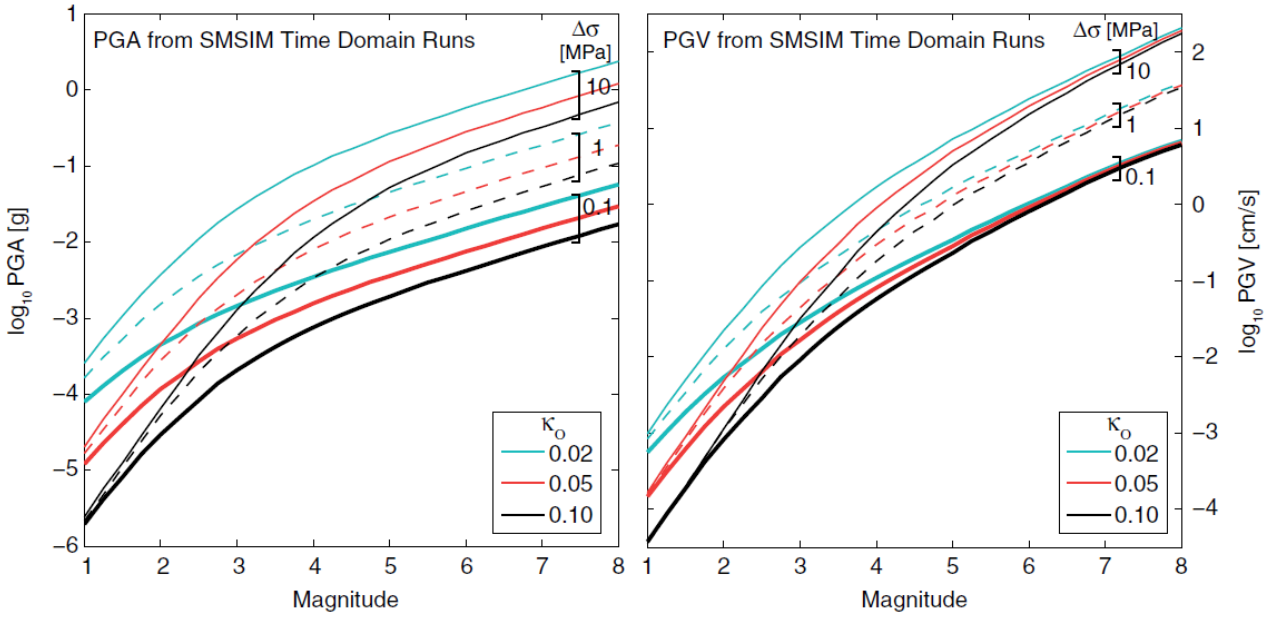


Figure 5.19. Magnitude scaling of PGA (*left*) and PGV (*right*) from stochastic simulations for different combinations of stress drop and kappa (Baltay & Hanks, 2014)

Consequently, GMPEs for spectral ordinates, particularly at short oscillator periods, covering the magnitude range from **M** 2.5 to **M** 6.5 will inevitably require non-linear scaling with magnitude. Over a wider range of magnitude one might expect even cubic

dependence on magnitude (Douglas & Jousset, 2011) or multiple segments of different quadratic scaling (Figure 5.20), but for the range of magnitudes in question two different ranges of quadratic scaling should suffice; insistence on extending the model down to **M** 1.5 would challenge this simplification, but such small earthquakes do not contribute appreciable to the hazard or the risk. However, as noted in Section 5.2, the empirical data from the Groningen field can only constrain a model for linear scaling of magnitude. Therefore, the challenge remains how to extrapolate the model to larger magnitudes.

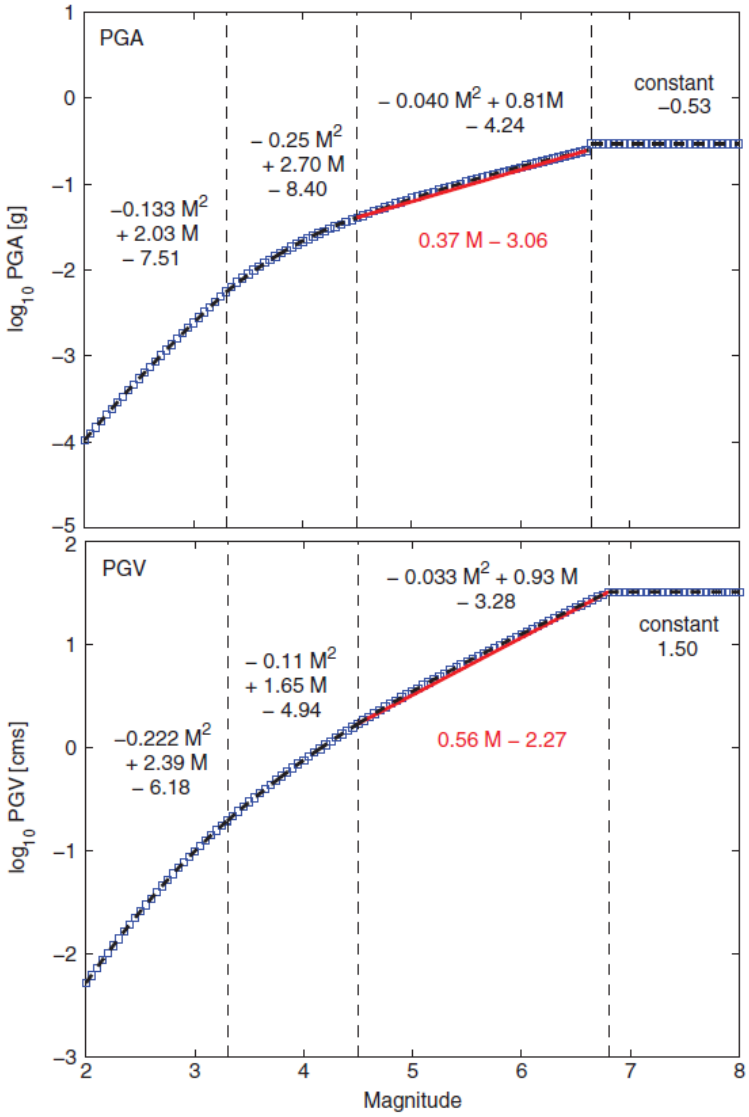


Figure 5.20. Model for magnitude scaling of PGA (*upper*) and PGV (*lower*) with three ranges of quadratic scaling and constant amplitudes at very larger magnitudes (Baltay & Hanks, 2014)

Although one option is to combine the Groningen dataset with the European dataset presented in Chapter 4, this is likely to prove complicated and therefore a preferable approach for the preliminary equations for response spectral ordinates is to use stochastic simulations, as discussed below in Section 5.4. The preceding discussions provide a basis for the magnitude scaling to be expected.

Before looking at stochastic simulations for the Groningen field, it is also worth briefly considering how seismological theory can inform us about the expected decay of ground-motion amplitudes with distance. A great deal could be written on this subject but here a very brief discussion of three key points is sufficient, particularly given that we are interested in predictions over relatively short distances. Earthquake ground-motion amplitudes decay over distance due to two distinct process: anelastic attenuation, represented by the quality factor, Q, and geometric spreading. Theoretically, the geometric spreading for far-field S-waves over the first tens of kilometres (before effects of reflection from the Moho and other path-related factors come into play) is spherical, which for amplitudes of motion implies a decay that is proportional to 1/R, where R is the distance from the source (Figure 5.21). Interestingly, the coefficients c_3 on the distance term in Eq.(5.9) were found to be close to -1 for longer periods (Table 5.1). In fact, however, the 1/R scaling is actually only theoretically valid for the far field and close to the source the rate of spreading may be steeper; note that In Figure 5.21 the shortest distance is 10 km. Another point worthy of note here is that for a point source, the R in question is the hypocentral distance, R_{hyp}, whereas we are developing a model in terms of R_{epi}. It should also be noted that empirical ground-motion data from short distance is generally unable to discriminate the anelastic attenuation effect and no attempt was made to include such a term in the empirical equations derived in Section 5.2; however, the anelastic attenuation of high-frequency motions may partly explain the small values of c_3 found for PGA and Sa(0.2s).

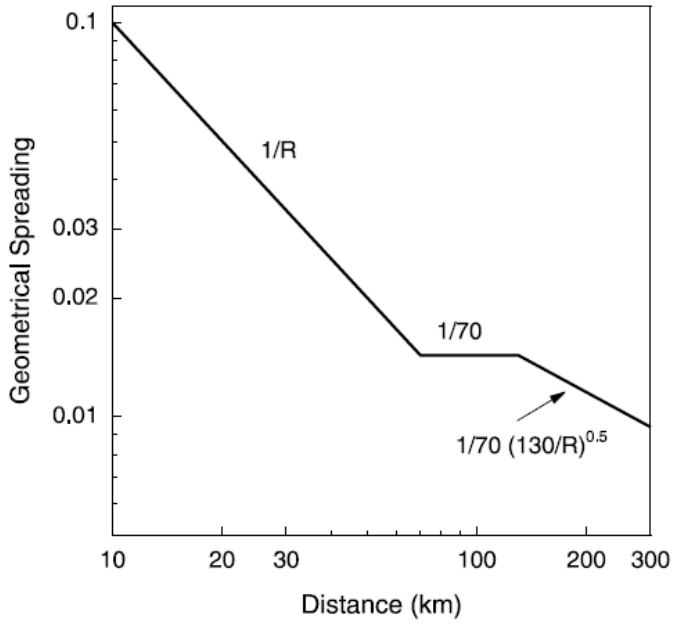


Figure 5.21. Model for geometrical spreading of ground motions (Boore, 2003)

The next issue is the saturation of motions at short distances. This saturation manifests as a flattening of the attenuation curve at short distances and is observed in most empirical GMPEs. The saturation has been related to several related causes, including the effect of the source being at some depth, even in the case of surface rupture (Baltay & Hanks, 2014), and to the effect of the extension of the rupture leading to greater attenuation from

more distant portions of the rupture at a given observation location (Yenier & Atkinson, 2014). The extent of the region of flattening is clearly dependent on the magnitude of the earthquake (Anderson, 2000), which is reflected in the functional form of the Equation (5.9). If the depth of the seismic source is partly responsible for the phenomena modelled by the near-source saturation term, then for the shallow Groningen earthquakes, it is to be expected that the value of the 'effective depth' should be rather small, perhaps even at larger magnitudes; this is actually consistent with the model that has been chosen (Figure 5.7).

In closing this brief discussion on magnitude scaling and attenuation with distance, we can address the interaction of the two in terms of the magnitude dependence of the spreading. Cotton *et al.* (2008) have demonstrated that while the spreading of Fourier amplitude spectra is independent of magnitude, this does not hold for response spectral ordinates (Figure 5.22).

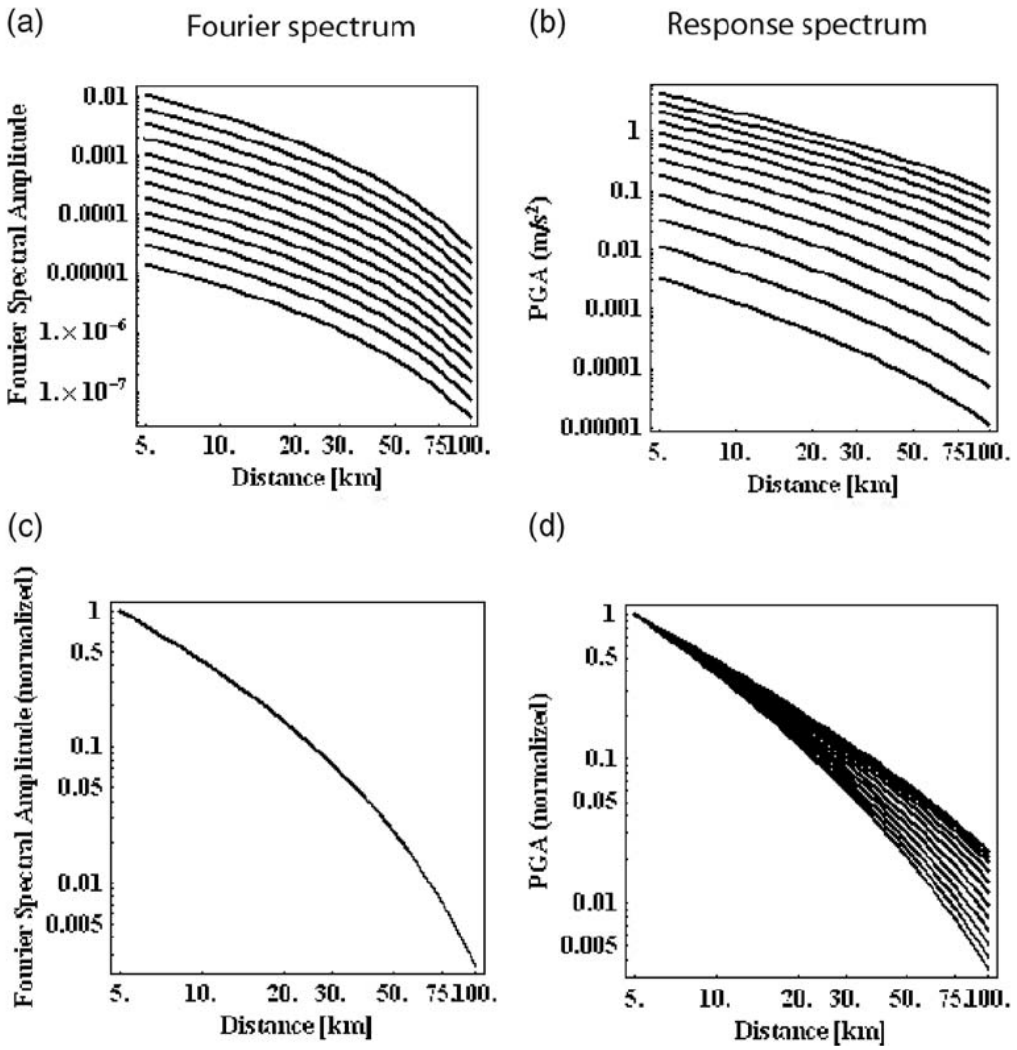


Figure 5.22. Fourier (a) and response (b) spectral accelerations at 50 Hz from stochastic simulations for magnitudes from 2.0 to 7.5 in 0.5 intervals; the lower curves in (c) and (d) are normalised to 1 at a distance of 5 km (Cotton *et al*., 2008)

The magnitude dependence of geometric spreading is reflected in the fact that most modern GMPEs including a magnitude-dependent term as a multiplier on the term defining attenuation with distance. Such a magnitude-dependent multiplier, however, is not required for our model since the magnitude dependence of the attenuation is accounted for by the magnitude-dependent near-source saturation term in the equation. This is similar to the way that the magnitude-dependence of the attenuation has been captured in several earlier GMPEs (*e.g*., Campbell, 1997).

5.4. Stochastic simulations for Groningen earthquakes

On the basis of the preceding discussions in Section 5.3, one could conclude that such theoretical considerations could provide a reliable basis for extrapolating the empirical equations derived in Section 5.2 to cover the full range of magnitudes considered in the hazard and risk calculations for Groningen. The most convenient way to apply such a theoretically-informed extrapolation is through the use of point-source stochastic simulations, as originally proposed by Hanks & McGuire (1981) and subsequently expanded upon by Boore (1983, 2003). Indeed, much of the work presented in Section 5.3 was actually derived using such point-source simulations, generally using the SMSIM software of Boore (2005a).

In order to generate stochastic simulations of ground motions in the Groningen field, the first stage is to obtain estimates of the parameters ($\Delta\sigma$, Q, κ, etc) that characterise the source, path and site. The same suite of accelerograms from the Groningen field that were used for the derivation of the empirical models presented in Section 5.2 was deployed for this purpose.

In order to guide the scaling of ground-motion predictions to larger magnitudes stochastic simulation approaches can be used (Boore, 2003, 2009). Stochastic simulation models produce synthetic acceleration time-series, from which we can obtain corresponding damped response spectra. They are based on two fundamental inputs for a given earthquake-site scenario (*e.g*., defined by magnitude, distance, source depth, *etc*.): (i) a model of the Fourier spectra and (ii) a shaking duration model. To define the inputs (i) and (ii) we typically take advantage of spectral analysis of small earthquakes in the area of study (Edwards *et al*., 2008). In the case of the Groningen Field, 12 events have so far been analysed in detail (Appendix I). From these analyses we define several parameters (earthquake stress-drop, geometrical spreading, Q, *etc*.) which are used to build the required input models for the stochastic simulations.

Using regional seismicity that has been recorded over a number of years we are able to define probability distributions of the stochastic model input parameters (Rietbrock *et al*., 2013; Atkinson & Boore, 2006). However, due to the limited number of recordings that have so far been subject to detailed analysis in the Groningen Field, we adopt an alternative approach. We use the inputs already derived from the 12 recorded events as a basis for the range of possible inputs. A coarse grid-search around those parameters is

then performed, finding the combination that best fits the observed response spectra database (Bommer *et al*., 2014). In this way the predictions should, within the permitted constraints on source shape, match the empirical GMPE predictions from Section 5.2 in the range of magnitudes already recorded.

In more detail, the inputs required for the stochastic simulations are: a geometrical spreading model; a Q and κ_0 model, a stress-drop and a duration model. For the duration model (T_d) we assume that there are two contributions, source (T_s) and path (T_p). The source model is simply given by $T_s=1/f_0$, with f_0 the source corner frequency defined by the Brune (1970) ω-squared source spectrum. The path duration is based on a numerical analysis by Herrmann (1985), who found that $T_p = 0.05R_{hyp}$. For the geometrical decay model we currently assume that a model with $1/R_{hyp}^{x}$ is appropriate; this may be revised at a later stage based on ongoing numerical waveform modelling. The parameter *x* is the exponent of decay (typically *x* is in the range of 1 to 1.3). The equivalent geometrical spreading exponent (c_4) for the GMPE was close to 1 at long periods (*i.e*., 1 to 2 s, where Q should not have a big impact on observed attenuation). For the Q-κ_0 model, it was observed that Q is rather poorly defined due to the short distances in the recording database (see Appendix I). Nevertheless, based on broadband analysis an average value of 260 was consistent with the data; based on high-frequency spectral fitting the value was 150. Within a 1-sigma confidence, the two estimates overlap. KNMI also note strong Q in the region (*e.g*., using borehole spectral ratios). For the site specific attenuation (κ_0) the average (across all sites) was 0.06 s (broadband fits) or 0.05 s (high frequency fits). For the stress drop the event-average was 7 bars and the record average 9 bars, although it should be noted that the range of values from event to event was large (of the order 1 to 50 bar, Figure 5.23).

The input amplification, in this case the average over the field, was based on spectral analysis of recorded waveforms and is consistent with an impedance contrast between the source V_S = 2600 m/s and the surface V_S = 200 m/s. In order to account for near-source saturation of ground motion for larger events the simulation is performed at a modified distance (related to a pseudo-depth) that accounts for near-field ground-motion saturation for larger events; for the sake of consistency we take the same form as used in the GMPE (Section 5.2).

Based on these initial observations and sensitivity tests we define 36 possible models based on the combination of: *x* = 1.0 and 1.1; Q = 150, 250; κ_0 = 0.05, 0.06 and 0.07 s; stress drop 10, 30 and 90 bar. The simulations were compared to the recorded response spectra at PGA, 0.2, 0.5, 1 and 2 s. The total misfit and variance across all periods was averaged leading to a single measure of bias and spread (sigma) for each combination of simulation parameters. The best fitting model is found to have the following parameter combination: Q = 150, κ_0 = 0.06 s, and stress drop = 30 bars; geometrical spreading exponent *x* = 1.0. This model was used for the central estimate with the stress drop valid across the magnitude range.

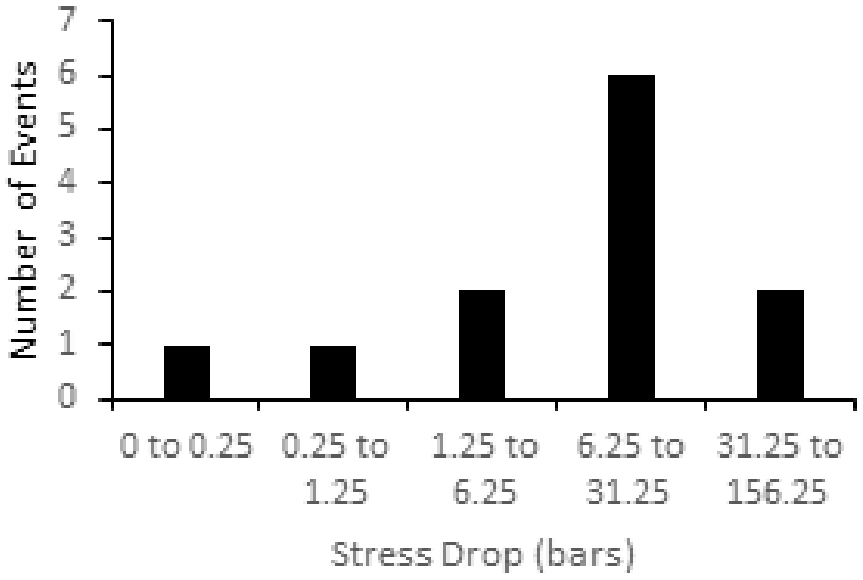


Figure 5.23. Historgram of stress drops (in bins) determined for the 12 events analysed.

Figures 5.24 to 5.27 compare the spectral accelerations obtained from the simulations with the median predicted values obtained from the empirical model presented in Section 5.2, in both the magnitude range where the Groningen data are concentrated and also for the full range of relevant magnitudes, at four different epicentral distances. The agreement in the small magnitude range is generally good, which is a reassuring outcome given that the two sets of estimated accelerations have been obtained using very different approaches, albeit that the stochastic parameters were tuned to provide a reasonable agreement with the recorded motions. A perfect match would not be expected given that the empirical model is constrained to follow the selected functional form of the regression equation Eq. (5.9). The fact that the agreement is best at short distances is also a welcome result, given that the dominant contributions likely to come from such distances.

In the lower plots of each figure that cover a wider magnitude range, the empirical equations are clearly being extrapolated beyond their range of applicability and the generation of these plots is purely for comparative purposes, as well as for illustrating the benefit of using the simulations to control the extrapolation of the predictions to larger magnitudes. The most striking feature of these plots, which is fully expected, is the very strong divergence of the two sets of curves seen at larger magnitudes. The divergence begins at a magnitude that increases with increasing response period. At the upper limit of **M** 6.5, the empirical model predicts short-period accelerations more than two orders of magnitude greater than those obtained from the stochastic simulations. There is also a notable divergence at smaller magnitudes, where the stochastic simulations indicate a steeper scaling with magnitude than is modelled by the empirical equations; this feature is consistent with the high κ values for the field and the influence this exerts on the apparent magnitude scaling, as discussed in Section 5.3. This small-magnitude divergence is evident below **M** 2.5 and is therefore of little concern since such small earthquakes are not expected to contribute significantly to the hazard or the risk.

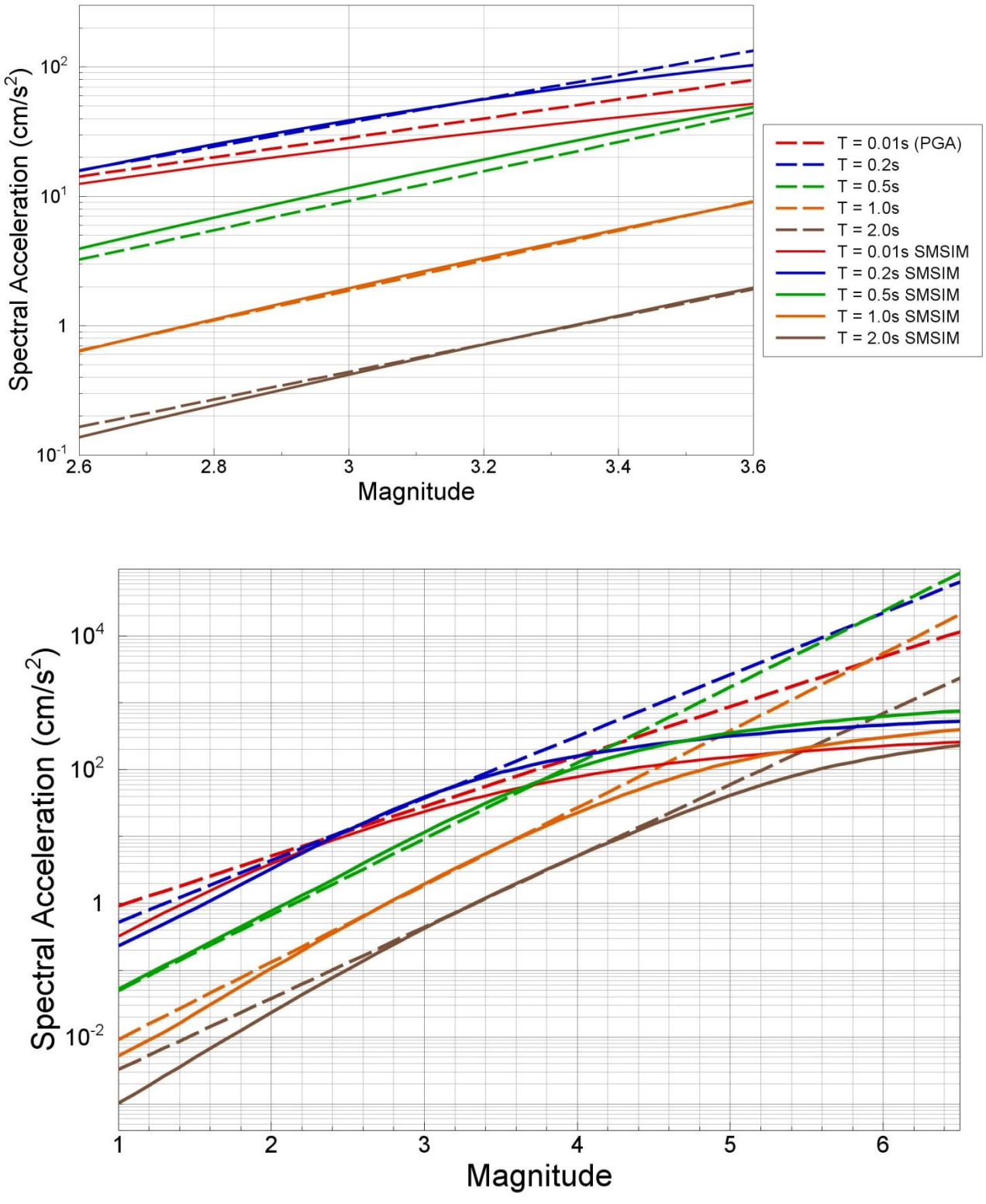


Figure 5.24. Comparison of median predictions from the empirical GMPEs derived in Section 5.2 with the stochastic ground-motion simulations at R_{epi} = 0 km for the magnitude range of the data (*upper*) and the full magnitude range (*lower*)

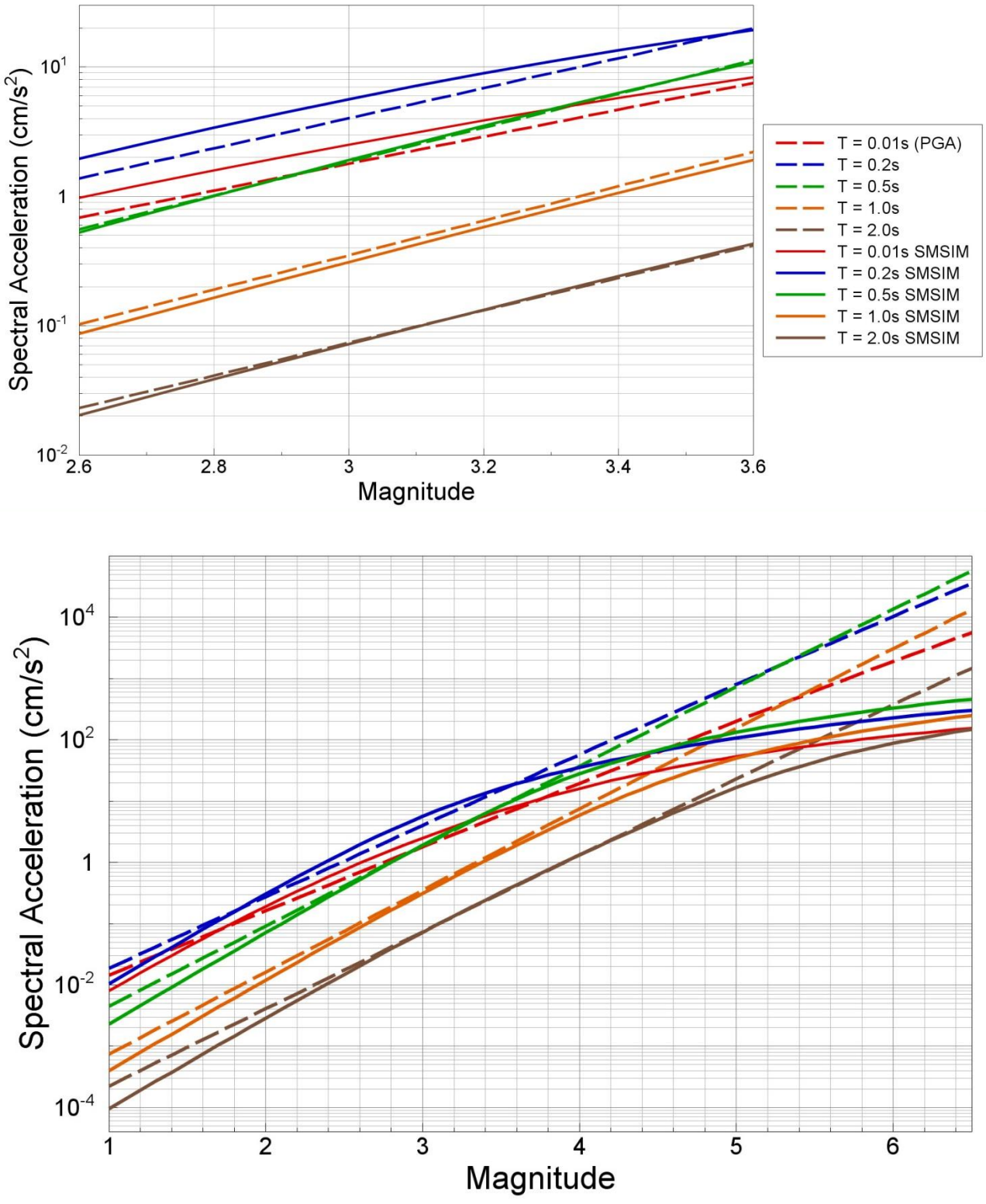


Figure 5.25. Comparison of median predictions from the empirical GMPEs derived in Section 5.2 with the stochastic ground-motion simulations at R_{epi} = 10 km for the magnitude range of the data (*upper*) and the full magnitude range (*lower*)

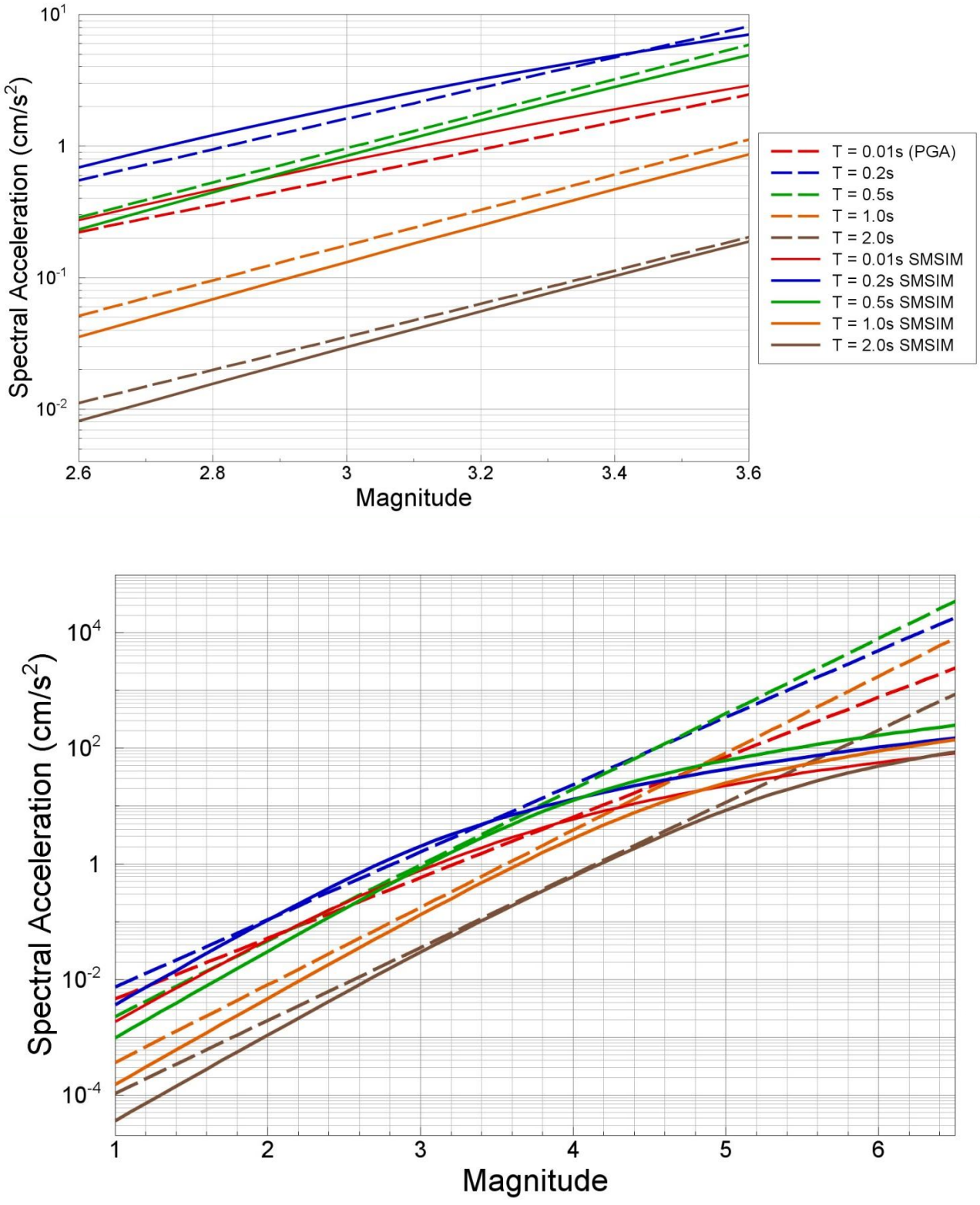


Figure 5.26. Comparison of median predictions from the empirical GMPEs derived in Section 5.2 with the stochastic ground-motion simulations at R_{epi} = 20 km for the magnitude range of the data (*upper*) and the full magnitude range (*lower*)

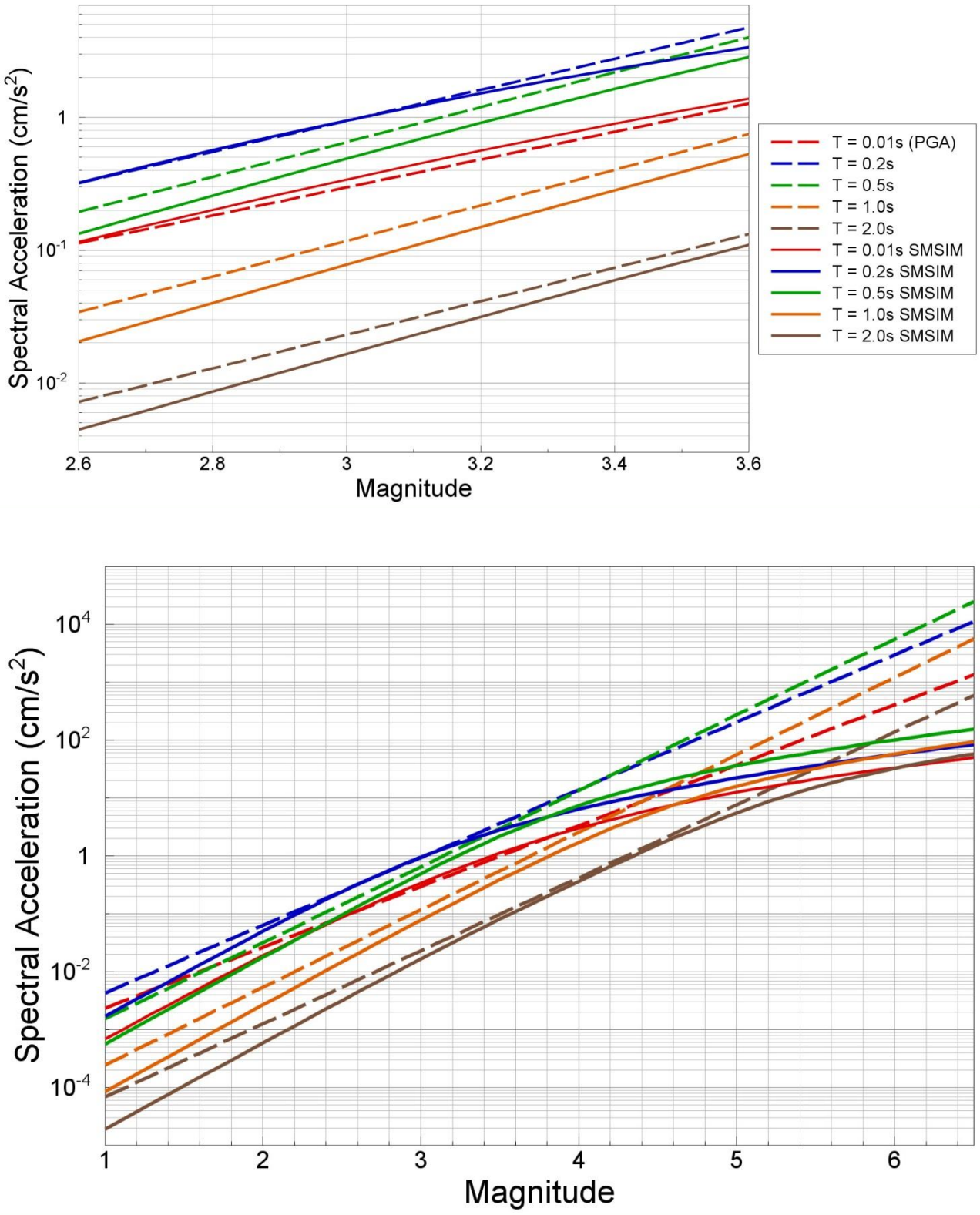


Figure 5.27. Comparison of median predictions from the empirical GMPEs derived in Section 5.2 with the stochastic ground-motion simulations at R_{epi} = 30 km for the magnitude range of the data (*upper*) and the full magnitude range (*lower*)

In order to cover the epistemic uncertainty we defined lower and upper estimate models. The lower model was based on the average stress-drop determined for the 12 events: 10 bar, constant over the entire magnitude range. For the recorded response spectral database this was found to lead to somewhat smaller ground-motions than the best-fitting 30 bar model. It will also lead to ground motions at large magnitude that are lower than typically observed in strong ground-motion datasets from tectonic, yet covers the possibility of low stress drop events occurring at shallow depths. The upper model has the same form as the central model at **M** 2.5 and scales the stress drop linearly in the log-space so that it reaches 100 bar at **M** 4.5, after which it remains constant. This will lead to ground motions that are consistent with those observed for larger tectonic events, which tend to have stress drops in this order of magnitude (Edwards & Fäh, 2013a). This model for magnitude-dependent Δσ is supported by a number of studies (Figure 5.28).

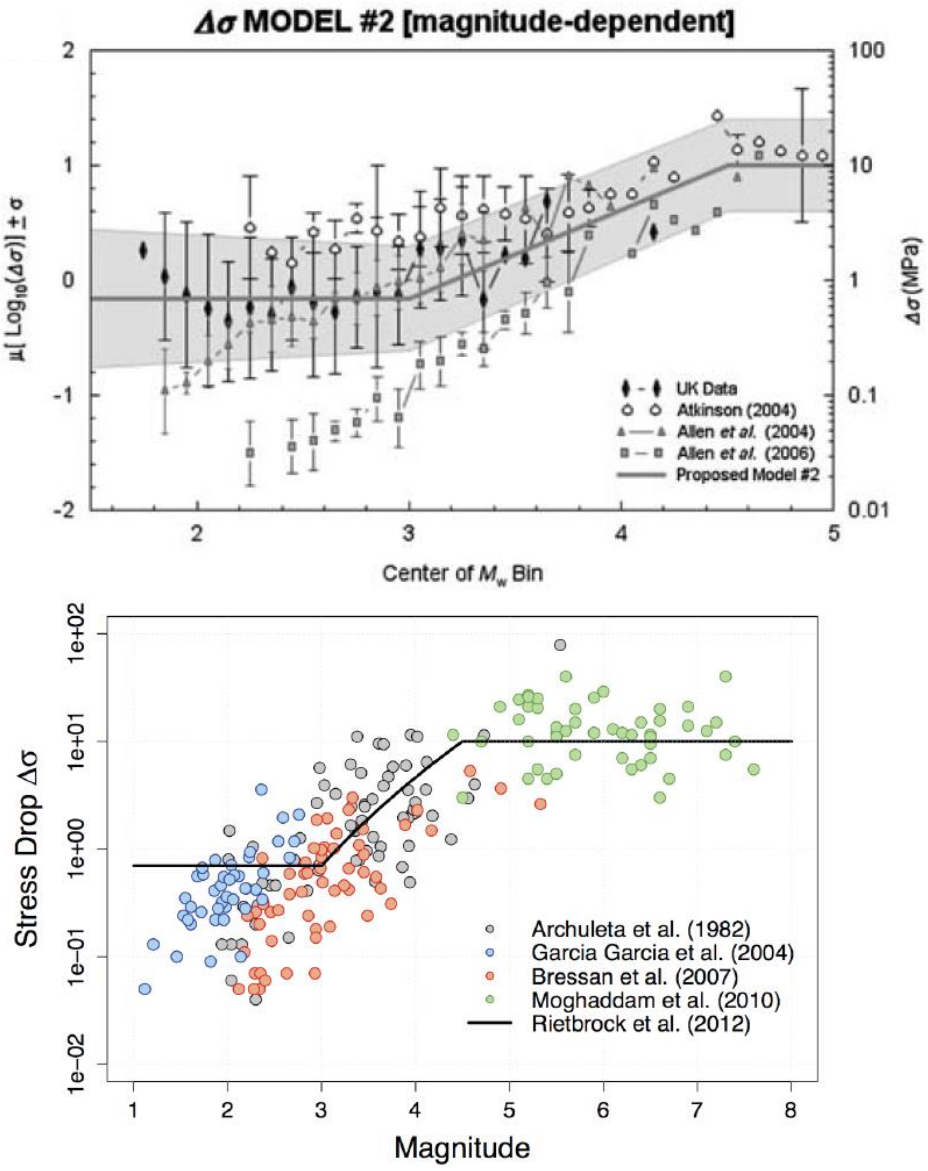


Figure 5.28. The magnitude-dependent stress drop model adopted by Rietbrock *et al*. (2013) and comparison with datasets from different studies that lend support to the model. Upper plot from Rietbrock *et al.* (2013), the lower from Stafford & Bommer (2012)

The approach of using stochastic simulations with different ranges of input parameters, in particular the stress drop, in order to generate alternative models covering the range of epistemic uncertainty has been used for both tectonic (Toro *et al*., 1997) and induced (Douglas *et al*., 2013) earthquakes, and is consistent with the approach proposed by Atkinson *et al.* (2014) to capture epistemic uncertainty in ground-motion logic trees. Figures 5.29 to 5.33 show the accelerations obtained from the simulations using the three alternative sets of inputs (*i.e*., stress drop models) as a function of magnitude and distance.

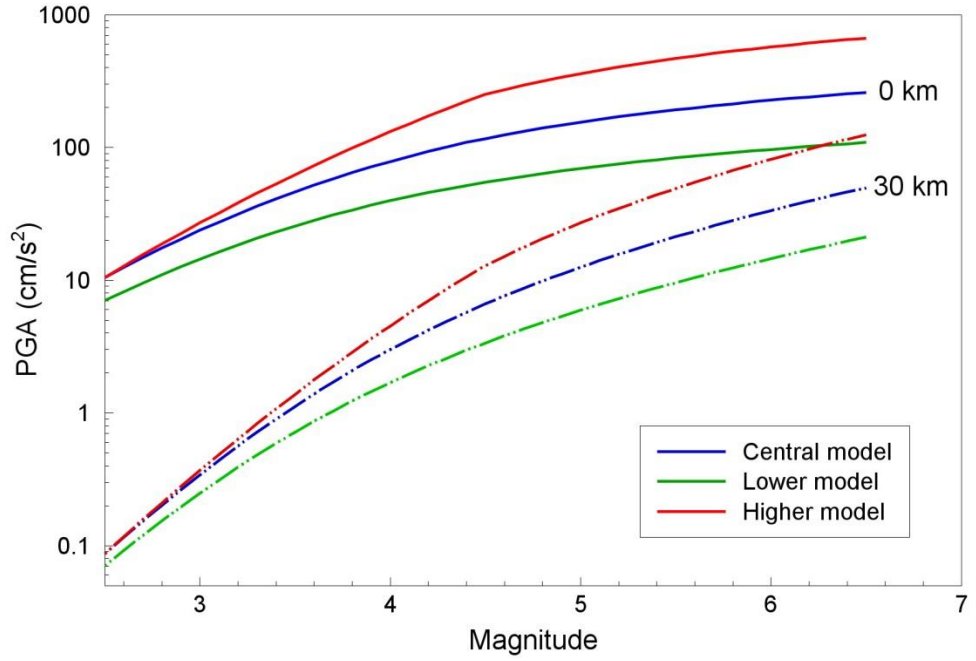


Figure 5.29. PGA values obtained from the three alternative stochastic models

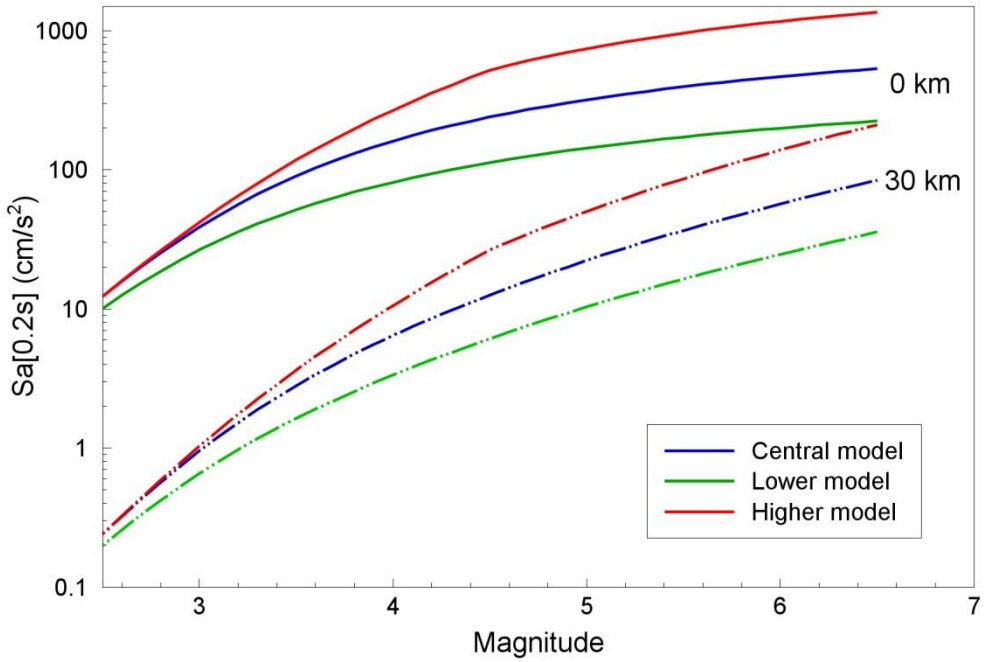


Figure 5.30. Sa(0.2s) values obtained from the three alternative stochastic models

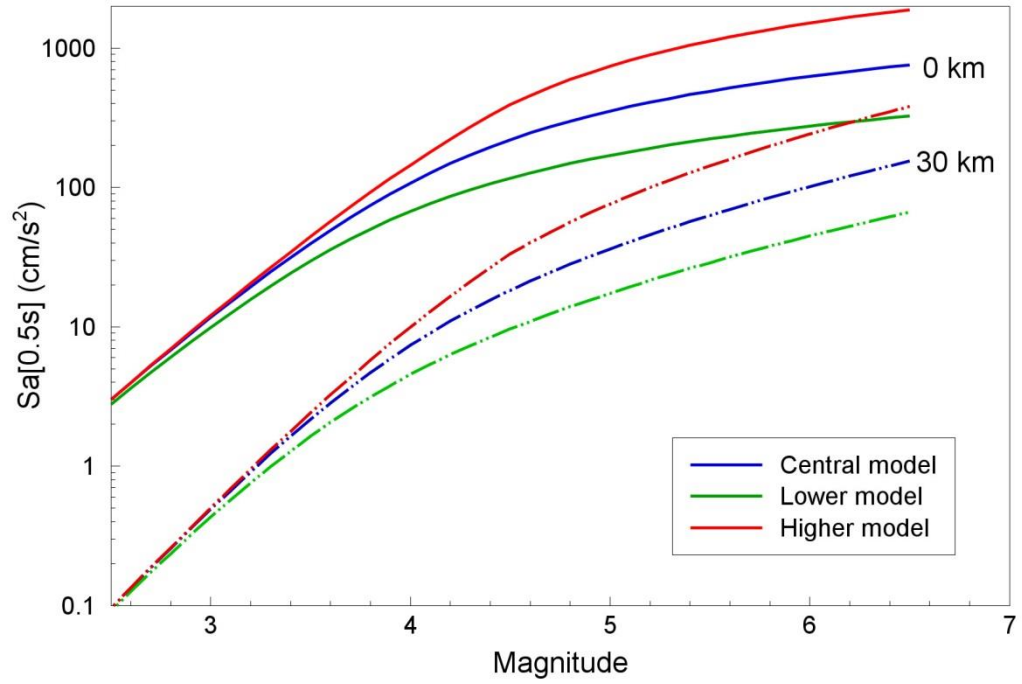


Figure 5.31. Sa(0.5s) values obtained from the three alternative stochastic models

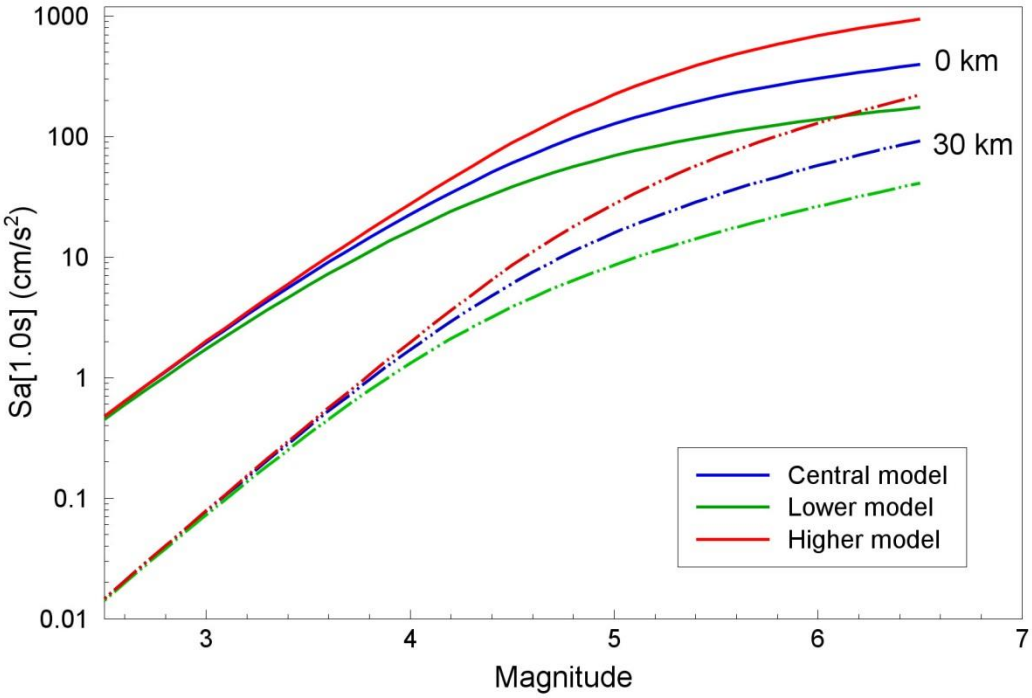


Figure 5.32. Sa(1.0s) values obtained from the three alternative stochastic models

In all cases the curves display the patterns that would be expected, with increasing divergence among the predicted accelerations with increasing magnitude, reflecting the increasing epistemic uncertainty associated with moving beyond the bounds of the data. For the shorter oscillator periods, there is also some separation of the predictions in the small-magnitude range, indicating that epistemic uncertainty is being considered even

where there are data. Although less apparent, this epistemic uncertainty in the small-magnitude range does persist at longer response periods. The degree of separation of the curves at greater magnitudes indicates that considerable epistemic uncertainty has been captured, which is appropriate in view of the current state of knowledge.

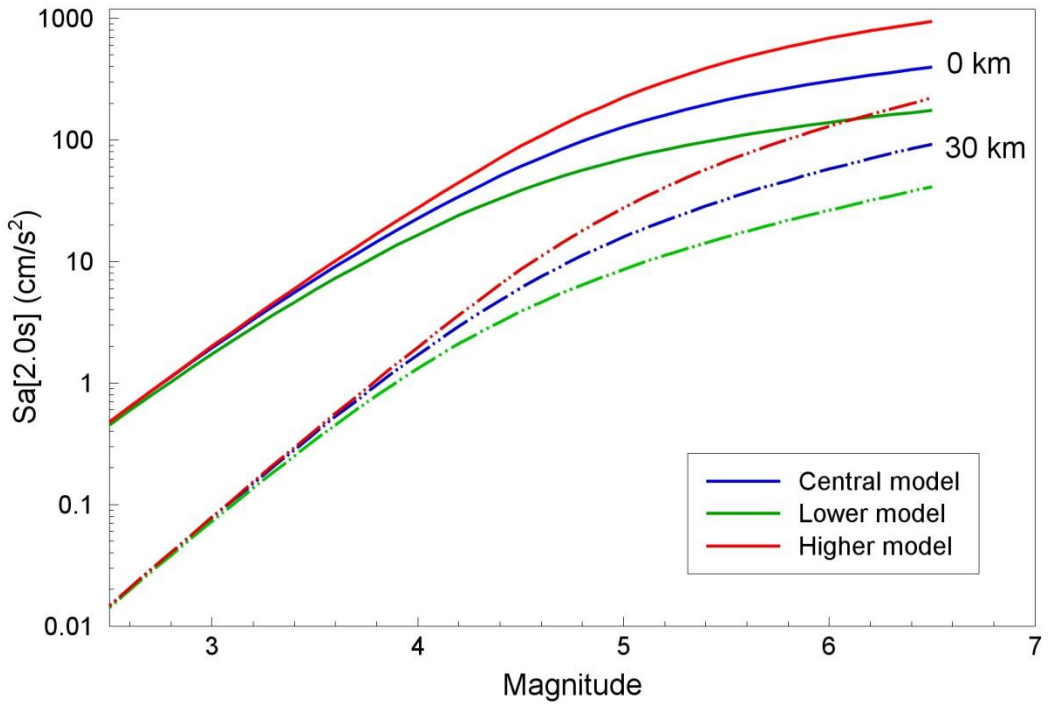


Figure 5.33. Sa(2.0s) values obtained from the three alternative stochastic models

Since the choice of the different values of the stress parameter selected for the simulations that underly the three models is of fundamental importance, it is worth closing this section with a little more discussion of the basis for this selection. The implicit assumption is that the key parameter in defining the extrapolations of the GMPEs to larger magnitudes is the values of the stress parameter, Δσ, and it should be carefully noted that the selected values represent the possible location of the median of the distribution of stress drops for earthquakes in the Groningen field. The variability of stress drops from one earthquake to another is reflected in the inter- or between-event standard deviation, τ, although the stress drop variability found from source studies is generally much greater than would be implied from the inter-event variability in GMPEs (Cotton *et al*., 2013). With this in mind, Figure 5.34 shows the stress drops determined for the 12 earthquakes by the inversions, and superimposed are the stress drop models underlying the three GMPEs. The error bars on the stress drop values are estimated following the same procedure used by Viegas *et al.* (2010) to estimate the confidence intervals on estimates of corner frequency. The plot suggests that three models encompass the range of possible median stress drops implied by this small dataset, which is clearly too limited to make any robust inferences about the actual distribution of stress drops in the Groningen field, even in the small magnitude range.

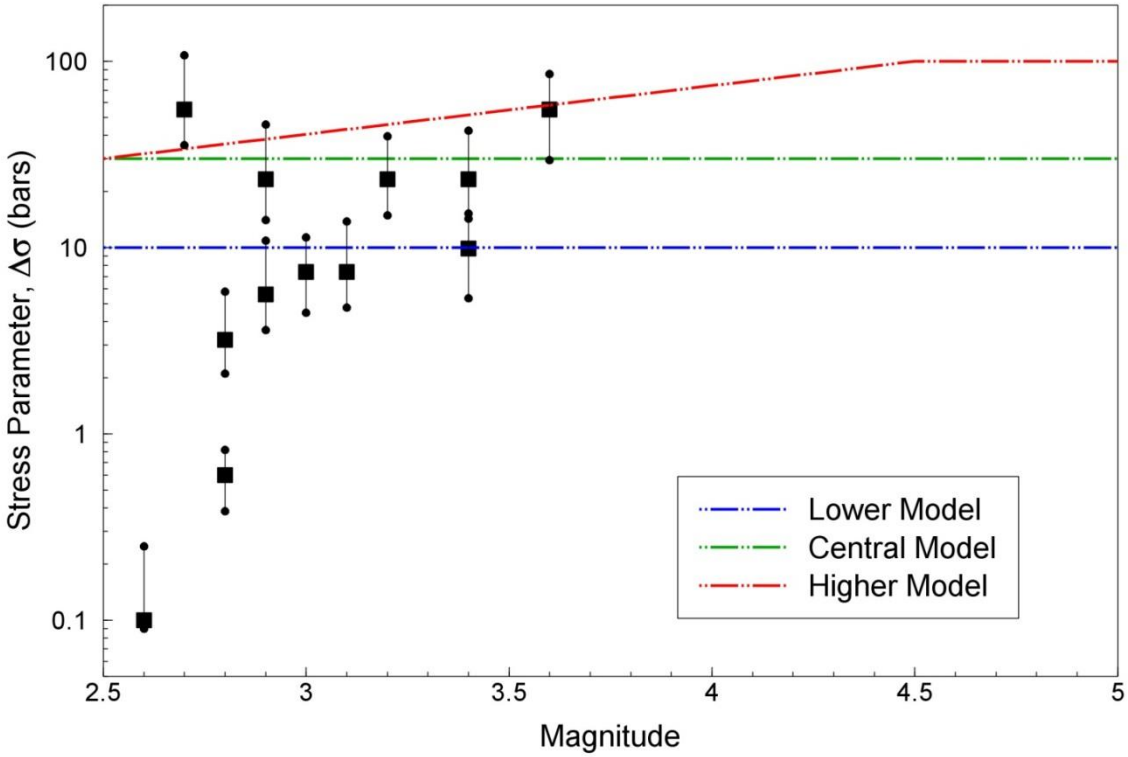


Figure 5.34. Individual stress parameter estimates, with vertical bars indicating confidence intervals estimated using the procedure of Viegas *et al*. (201), compared with the stress drop models underlying the three sets of stochastic simulations generated to capture the range of epistemic uncertainty in the predictions of spectral accelerations

The relative weights assigned to the three stress drop values—and hence to the branches of the ground-motion logic-tree—are presented and discussed in Section 6.3, but it is worthwhile devoting some space here to discussing the adopted stress drop values, and in particular the value of 30 bars assigned to the central (*i.e*., best estimate) model. As has been noted earlier in the report, Hough (2014) inferred from analysis of intensity data obtained for tectonic and induced earthquakes in the Central and Eastern United States, that stress drop from the latter are systematically lower. Hough (2014) attributed this observation to the shallower depths of induced events, with the consequence that in the epicentral region the reduction of motions due to lower stress drop may be balanced out by the shorter travel paths. The concept of lower stress drops for shallower crustal earthquakes has also been proposed for tectonic earthquakes: in developing new GMPEs for a region of Australia, Allen (2012) produced separate equations for shallow and deeper crustal earthquakes, the former yielding lower accelerations at all distances.

The issue of stress drop and focal depth received considerable attention at the 2015 meeting of the Seismological Society of Meeting, which included an all-day session on induced seismicity. For example, Viegas *et al.* (2015) concluded that "*on average, the reservoir events have low static and dynamic stress drops*" and Neighbors *et al*. (2015) report remarkably low stress drop (~3 bars) for induced earthquakes in Oklahoma, alluding to focal depth as a potential explanation. Boyd *et al.* (2015) also concluded "*that stress drops are considerably lower for potentially induced earthquakes, possibly due to their*

relatively shallow focal depth." Cramer (2015) concludes that the primary reason for the low stress drops observed for induced earthquakes is focal depth and that for similar focal depths, induced and tectonic events may not be distinct. Wong e*t al.* (2015) extend these arguments and note that "*the shallow nature of induced earthquakes in a low Q environment and the resulting potentially lower stress drops has also been recently suggested as a cause for lower ground motions as compared to tectonic earthquakes*." In summary, while we acknowledge that this is an area of ongoing research and that definitive conclusions have yet to be reached, there is a body of evidence and the judgement of several leading seismologists to support lower stress drops—and consequently lower high-frequency ground motions—from induced earthquakes, both as a result of their shallow depths and the fact that the source of energy release will generally be located in zones with high attenuation characteristics. In the case of the Groningen earthquakes, we would also note that the mechanism by which the earthquakes are being induced—namely, relaxation of the upper 3 km of the crust due to reduction of gas pressure in the reservoir—is also unlikely to generate events of very high stress drop, even if the participating faults have not experienced slip for a long time. At the same time, it is important to emphasise that the possibility of the induced earthquakes in the Groningen field being comparable to tectonic events is retained as a viable model through the inclusion of the upper stress drop value (reaching 100 bars for **M** 4.5 and above).

6. Ground-Motion Model for Groningen

On the basis of the analyses and discussions presented in Chapter 5, a clear approach to the development of the Version 1 GMPEs has been defined. In this Chapter, the final models are developed using both the empirical data from the field and the stochastic simulations (Section 6.1). Using the empirical data and the sigma penalty for using point-source distance metrics (Appendix II), sigma models are proposed for the equations (Section 6.2). Expert judgements regarding the source parameters adopted for the stochastic simulations and comparison of predicted accelerations with the values observed in the transformed European dataset are used to then assign weights to three alternative models (Section 6.3). In Section 6.4 the new equations for the prediction of PGA are compared with the V0 predictions. In Section 6.5 additional stochastic simulations are used to provide guidance on appropriate response spectral shapes to interpolate between the 5 oscillator periods at which the predictions are provided. Finally, alternative GMPEs are derived using R_{JB} rather than R_{epi} as the distance metric (Section 6.6).

6.1. GMPEs for median ground-motion amplitudes

In view of the good agreement between the empirical GMPEs derived in Section 5.2 and the stochastic simulations (Section 5.4) in the magnitude range covered by the Groningen database, the procedure adopted to develop the final GMPEs is split into two parts. For the coefficients that define the median predictions, an appropriate functional form is fitted to the stochastically simulated spectral accelerations. Then, the sigma values in the small-magnitude range are computed using the Groningen dataset presented in Section 3.2, and then extended to larger magnitudes; the sigma model derivation is described in Section 6.2.

In order to capture the non-linearity in the magnitude scaling, a quadratic magnitude term is added to the function form used in Section 5.2. Additionally, since the equations are required to model the ground motions over 4 units of magnitude (M 2.5-6.5), a hinge magnitude, $\overline{M}$, is defined and the quadratic scaling allowed to vary above and below this threshold (see Figure 5.20).

$$\ln(Y) = c_1 + c_2 M + c_3 (M - \overline{M})^2 + c_4 \ln\sqrt{R_{epi}^2 + [\exp(c_5 M + c_6)]^2} \qquad M \le \overline{M} \qquad (6.1a)$$

$$\ln(Y) = c_1 + c_2 M + c_{3a} (M - \overline{M})^2 + c_4 \ln\sqrt{R_{epi}^2 + [\exp(c_5 M + c_6)]^2} \qquad M > \overline{M} \qquad (6.1b)$$

As before, the coefficients c_5 and c_6 are held constant at 0.4233 and -0.6083 respectively, and the other coefficients determined from the regression analyses. In the first regressions, the value of $\overline{M}$ was included as one of the parameters to be optimised in the fitting, but it was found that the value oscillated in value as the period increased, which could lead to irregular spectral shapes. For example, for the central model (Δσ = 30 bars)

the values obtained for PGA and spectral accelerations at 0.2, 0.5, 1.0 and 2.0 seconds were 4.7, 4.1, 5.0, 3.4 and 4.6. Using instead a fixed value of 4.5 at all periods results in only a very slight increase in the standard deviations and hence this was chosen as the constant value to be constrained in all the regressions. With the value of the hinge magnitude fixed at **M** 4.5 for all cases, the regressions were performed on the three sets of simulated ground motions corresponding to the central model and lower and higher alternatives. The coefficients for the three median models are reported in Tables 6.1-6.3.

Table 6.1. Coefficients of Eq.(6.1) for the central (Δσ = 30 bars) model

	PGA	**Sa(0.2s)**	**Sa(0.5s)**	**Sa(1.0s)**	**Sa(2.0s)**
c_1	1.1563	2.4972	-0.0684	-4.3882	-7.8093
c_2	1.2732	1.1216	1.5742	2.2288	2.6929
c_3	-0.3394	-0.4314	-0.5416	-0.3549	-0.1520
c_{3a}	-0.1342	-0.0747	-0.2397	-0.4202	-0.4370
c_4	-1.5048	-1.4806	-1.2266	-1.1640	-1.1526
c_5	0.4233	0.4233	0.4233	0.4233	0.4233
c_6	-0.6083	-0.6083	-0.6083	-0.6083	-0.6083
$\overline{M}$	4.5	4.5	4.5	4.5	4.5

Table 6.2. Coefficients of Eq.(6.1) for the lower (Δσ = 10 bars) model

	PGA	**Sa(0.2s)**	**Sa(0.5s)**	**Sa(1.0s)**	**Sa(2.0s)**
c_1	1.0490	2.1812	0.6494	-3.2480	-7.1140
c_2	1.1122	1.0202	1.2775	1.8682	2.4569
c_3	-0.3132	-0.3408	-0.5417	-0.4377	-0.2117
c_{3a}	-0.0942	-0.0544	-0.1430	-0.3306	-0.4442
c_4	-1.4529	-1.4670	-1.2223	-1.1500	-1.1324
c_5	0.4233	0.4233	0.4233	0.4233	0.4233
c_6	-0.6083	-0.6083	-0.6083	-0.6083	-0.6083
$\overline{M}$	4.5	4.5	4.5	4.5	4.5

Table 6.3. Coefficients of Eq.(6.1) for the higher (Δσ = 100 bars) model

	PGA	**Sa(0.2s)**	**Sa(0.5s)**	**Sa(1.0s)**	**Sa(2.0s)**
c_1	0.1638	1.5092	-1.7676	-5.9331	-8.5757
c_2	1.6566	1.4980	2.0695	2.6584	2.9277
c_3	-0.3236	-0.4312	-0.4308	-0.2273	-0.0983
c_{3a}	-0.2643	-0.2125	-0.4043	-0.5076	-0.4068
c_4	-1.5391	-1.4926	-1.2282	-1.1729	-1.1680
c_5	0.4233	0.4233	0.4233	0.4233	0.4233
c_6	-0.6083	-0.6083	-0.6083	-0.6083	-0.6083
$\overline{M}$	4.5	4.5	4.5	4.5	4.5

In order to obtain some insight into how well these models fit the simulated motions—which are not referred to as data—Figures 6.1 to 6.15 show the 'residuals' against magnitude and distance for each model and ground-motion parameter. The residuals are simply the natural logarithms of the ratios of the stochastically-simulated motions to those predicted by these models, without distinction between inter- and intra-event components.

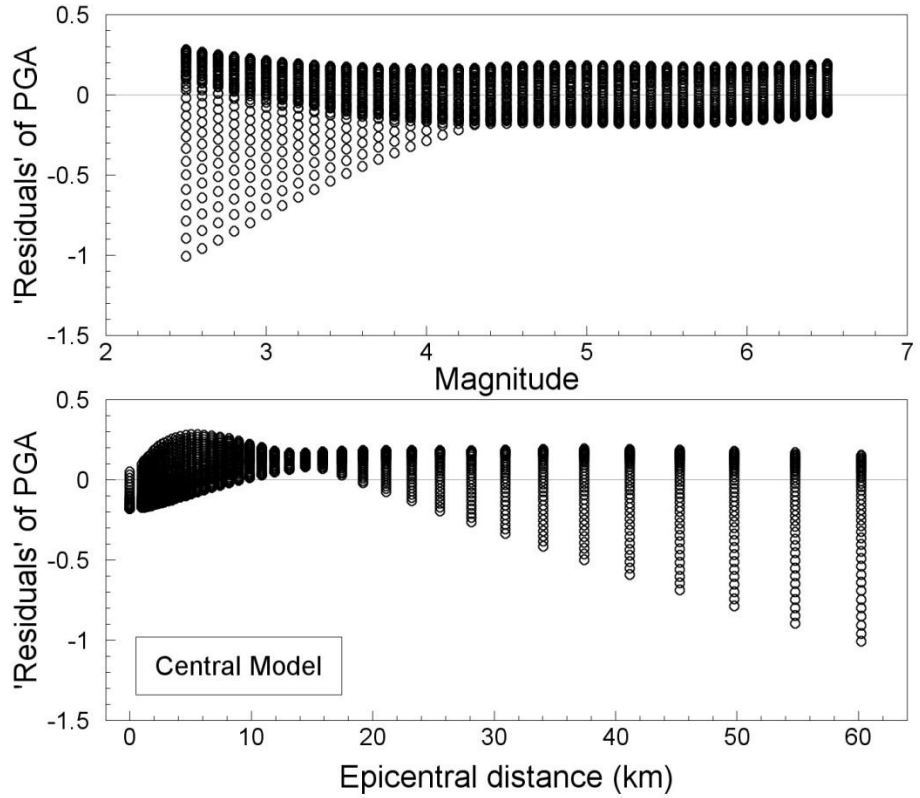


Figure 6.1. Computed 'residuals' of the stochastic simulations (Δσ = 30 bars) for PGA relative to the regression model in Eq.(6.1) and Table 6.1

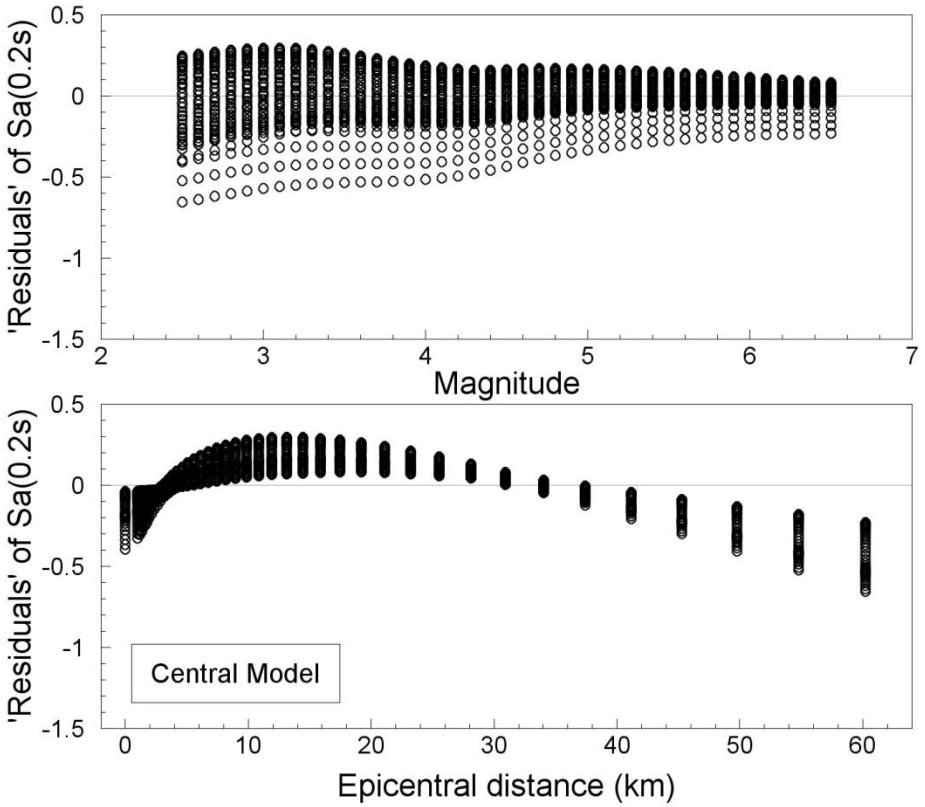


Figure 6.2. Computed 'residuals' of the stochastic simulations (Δσ = 30 bars) for Sa(0.2s) relative to the regression model in Eq.(6.1) and Table 6.1

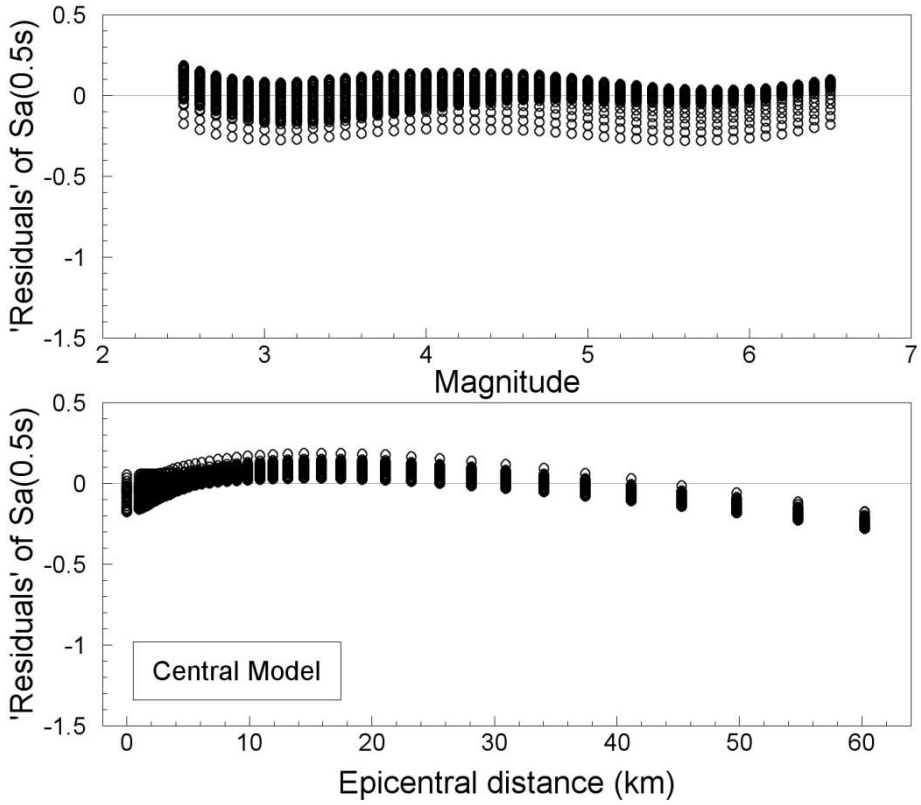


Figure 6.3. Computed 'residuals' of the stochastic simulations (Δσ = 30 bars) for Sa(0.5s) relative to the regression model in Eq.(6.1) and Table 6.1

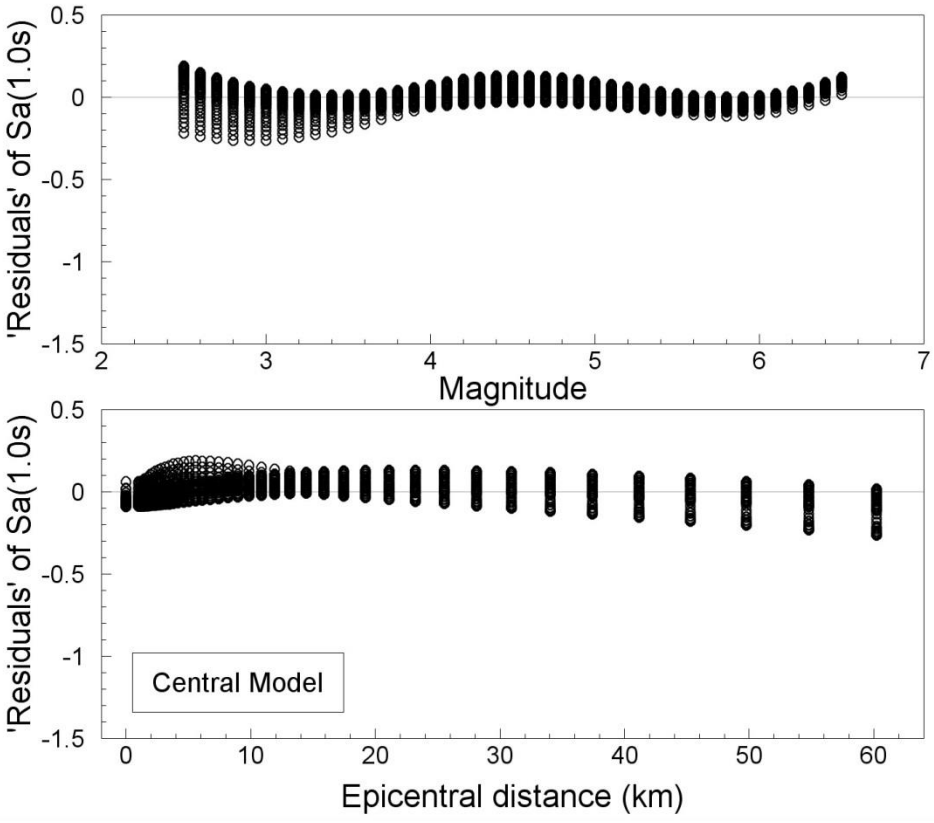


Figure 6.4. Computed 'residuals' of the stochastic simulations (Δσ = 30 bars) for Sa(1.0s) relative to the regression model in Eq.(6.1) and Table 6.1

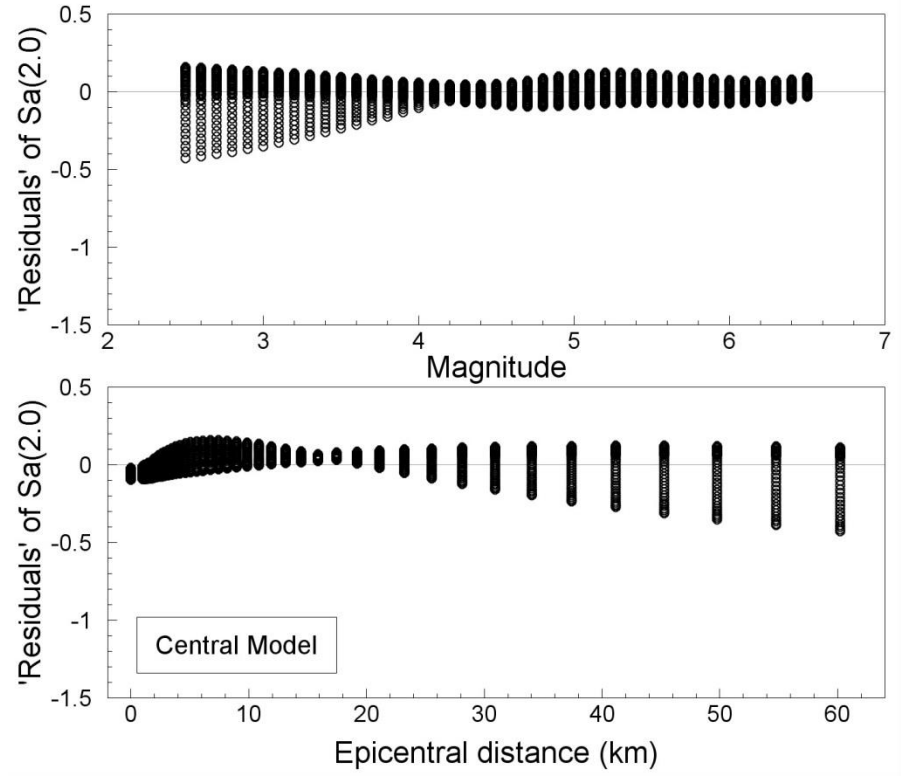


Figure 6.5. Computed 'residuals' of the stochastic simulations (Δσ = 30 bars) for Sa(2.0s) relative to the regression model in Eq.(6.1) and Table 6.1

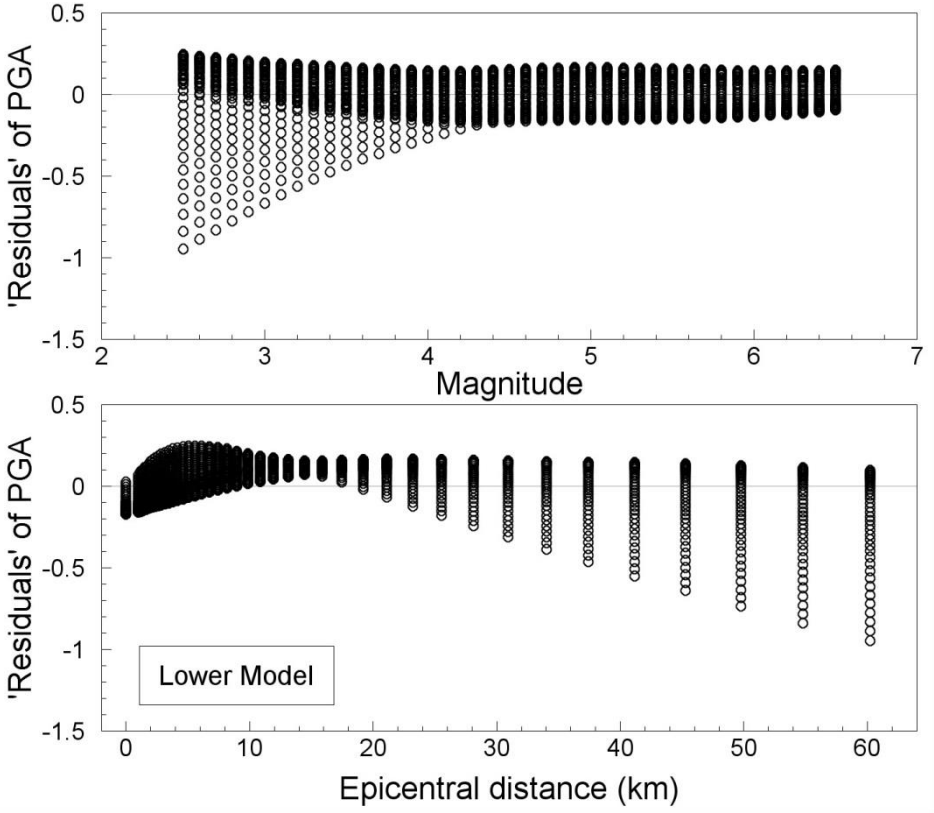


Figure 6.6. Computed 'residuals' of the stochastic simulations (Δσ = 10 bars) for PGA relative to the regression model in Eq.(6.1) and Table 6.2

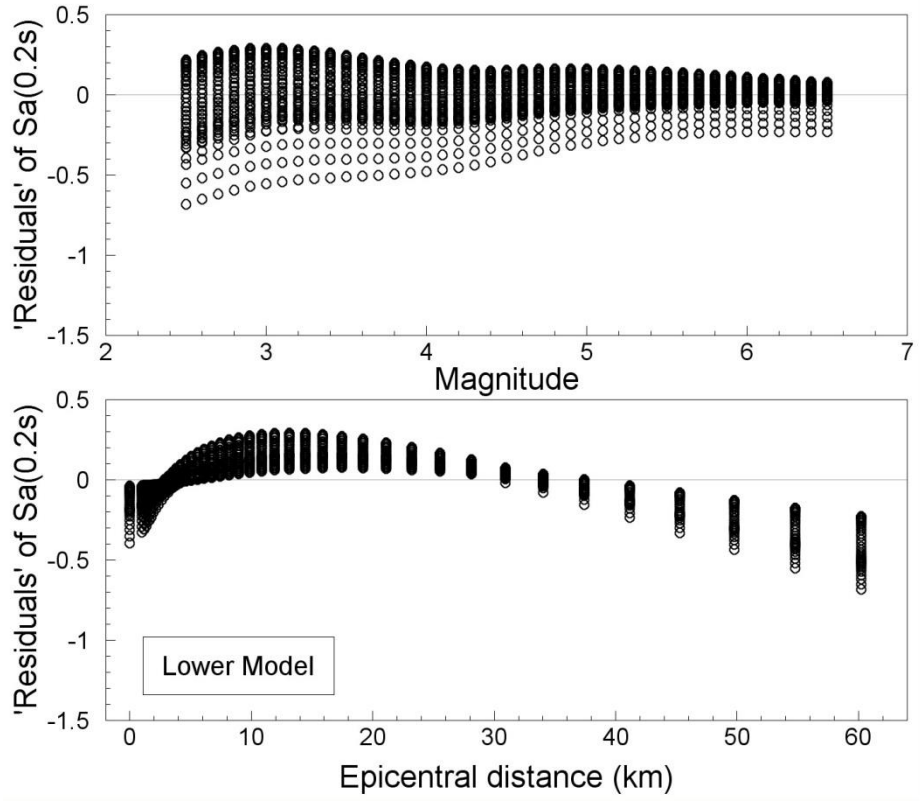


Figure 6.7. Computed 'residuals' of the stochastic simulations (Δσ = 10 bars) for Sa(0.2s) relative to the regression model in Eq.(6.1) and Table 6.2

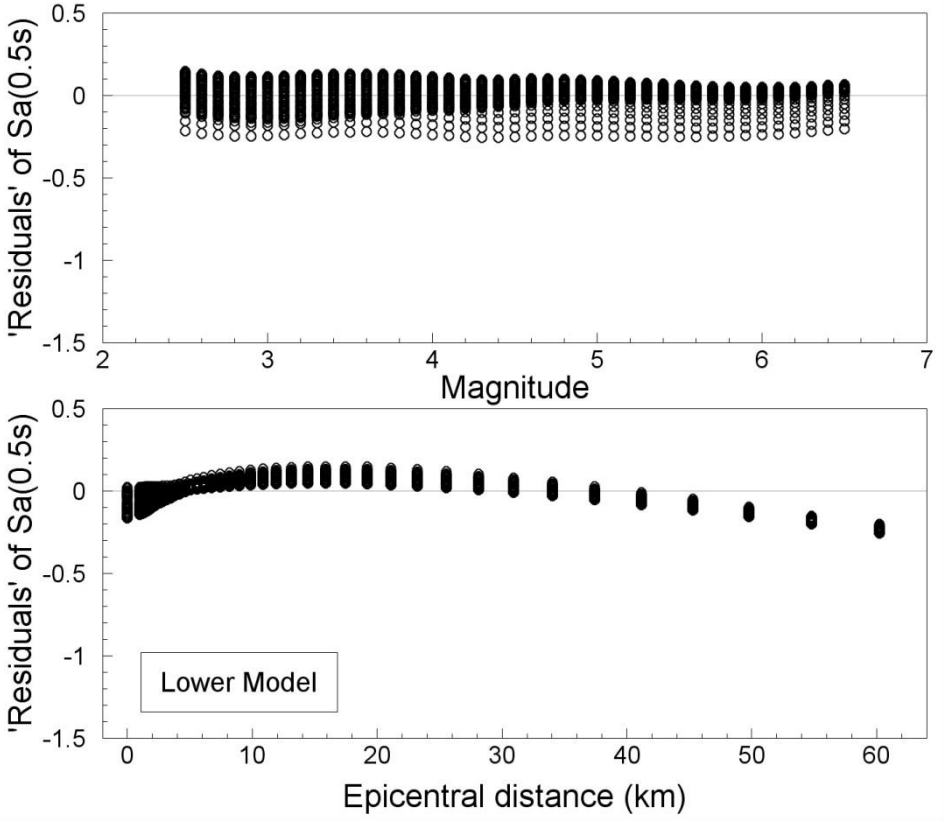


Figure 6.8. Computed 'residuals' of the stochastic simulations (Δσ = 10 bars) for Sa(0.5s) relative to the regression model in Eq.(6.1) and Table 6.2

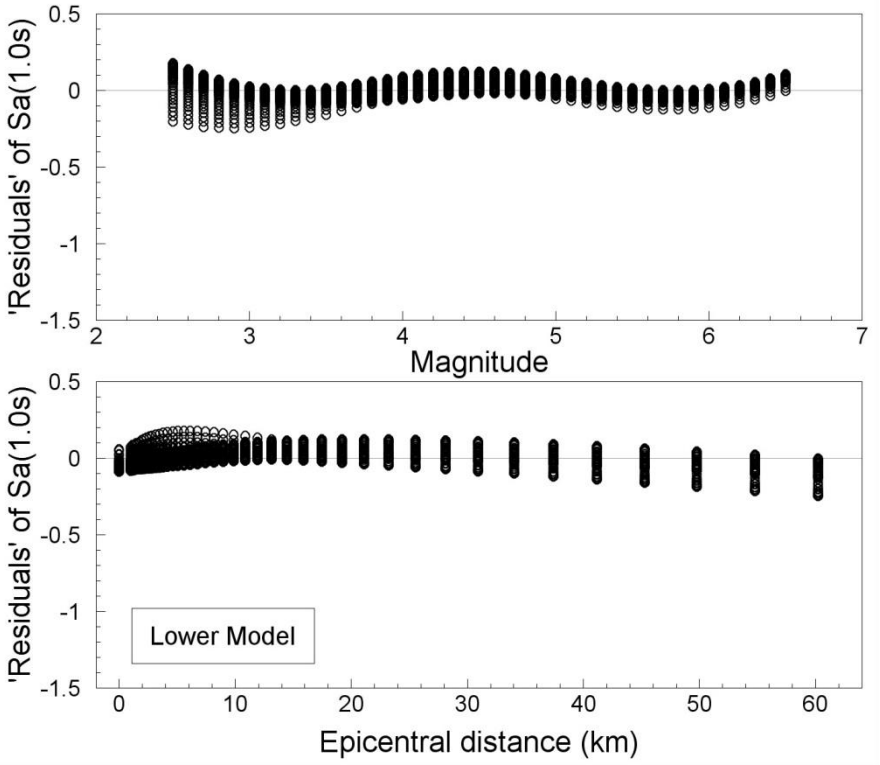


Figure 6.9. Computed 'residuals' of the stochastic simulations (Δσ = 10 bars) for Sa(1.0s) relative to the regression model in Eq.(6.1) and Table 6.2

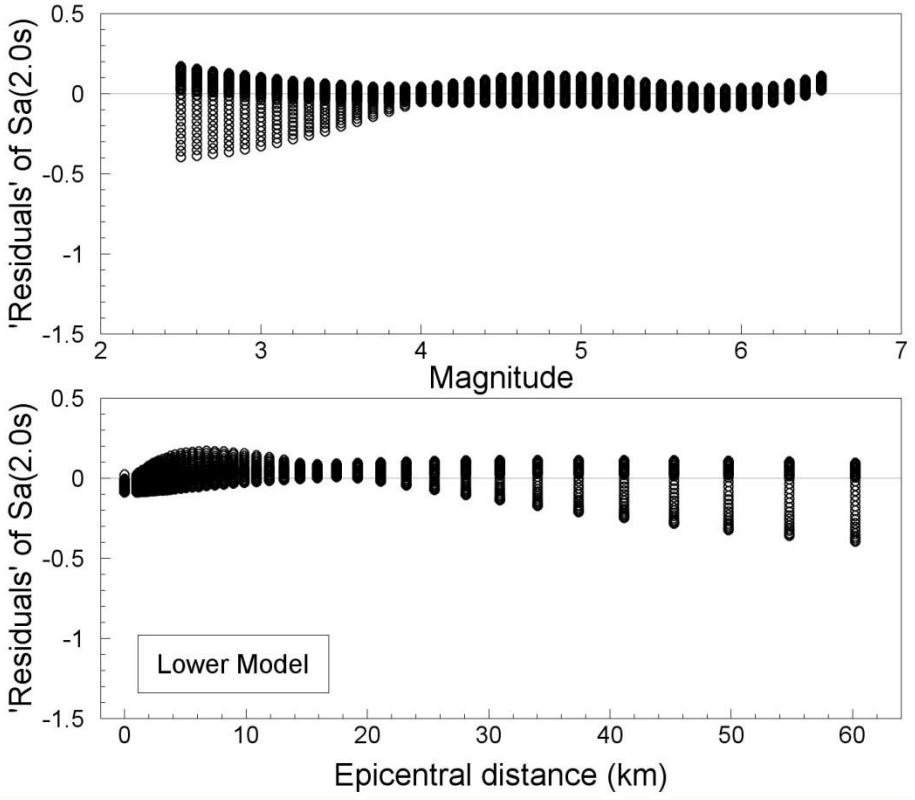


Figure 6.10. Computed 'residuals' of the stochastic simulations (Δσ = 10 bars) for Sa(2.0s) relative to the regression model in Eq.(6.1) and Table 6.2

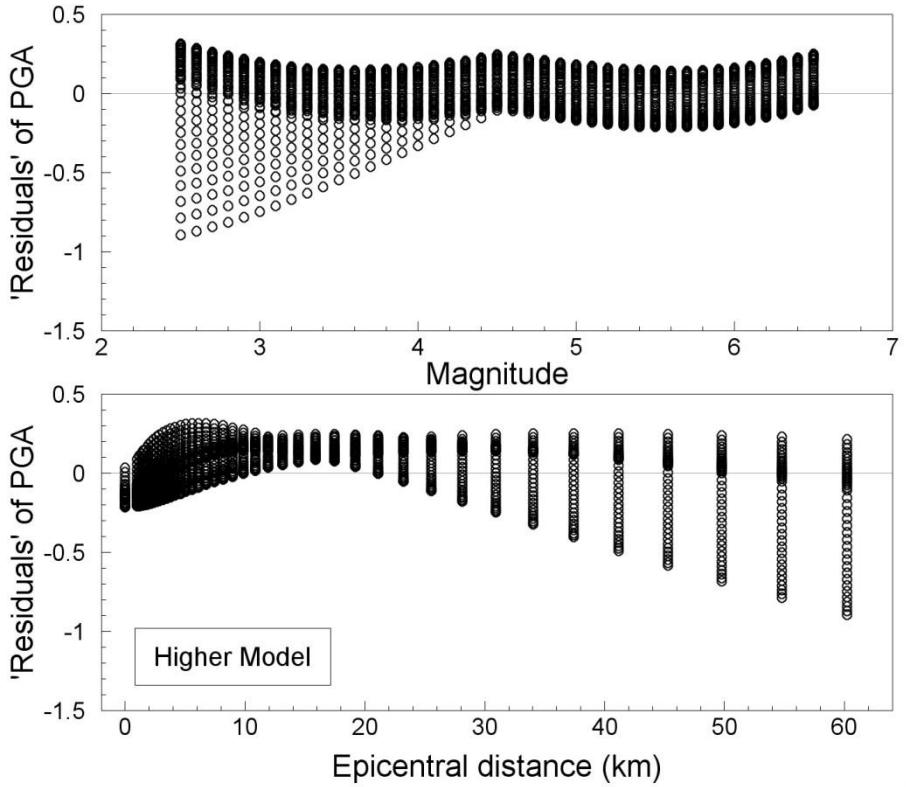


Figure 6.11. Computed 'residuals' of the stochastic simulations (Δσ = 100 bars) for PGA relative to the regression model in Eq.(6.1) and Table 6.3

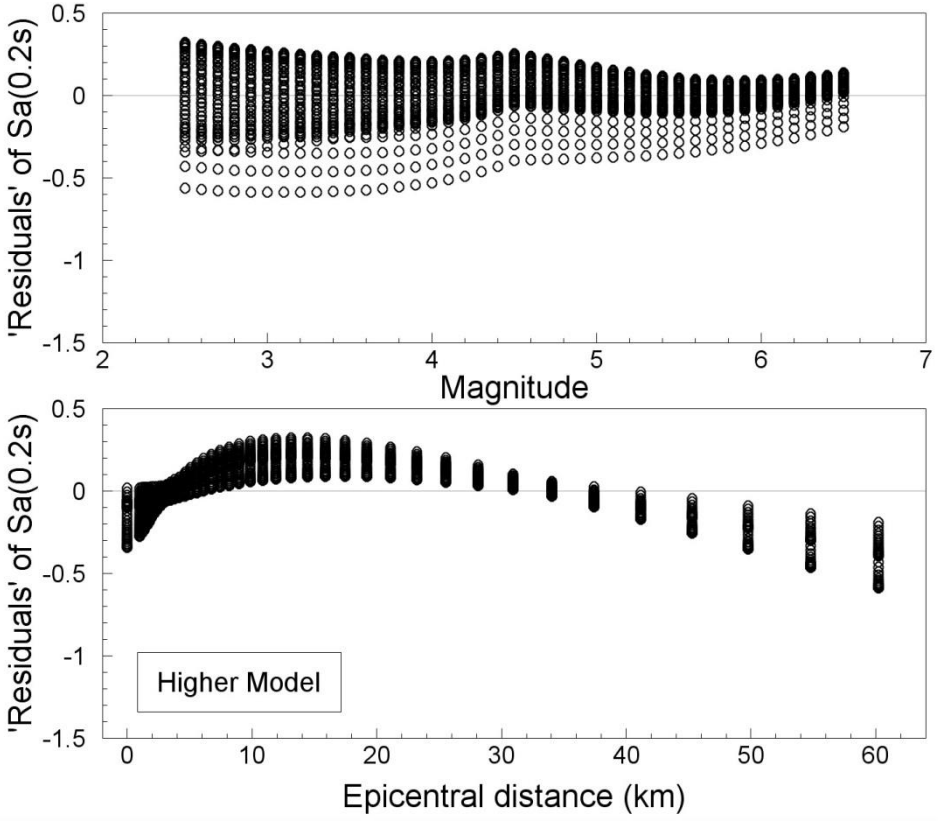


Figure 6.12. Computed 'residuals' of the stochastic simulations (Δσ = 100 bars) for Sa(0.2s) relative to the regression model in Eq.(6.1) and Table 6.3

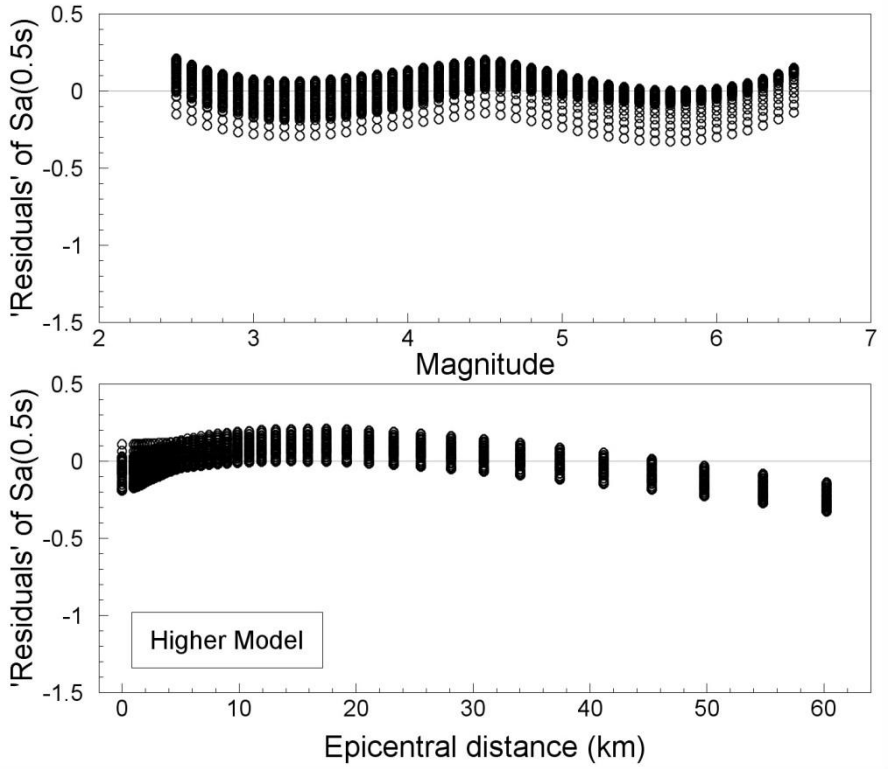


Figure 6.13. Computed 'residuals' of the stochastic simulations (Δσ = 100 bars) for Sa(0.5s) relative to the regression model in Eq.(6.1) and Table 6.3

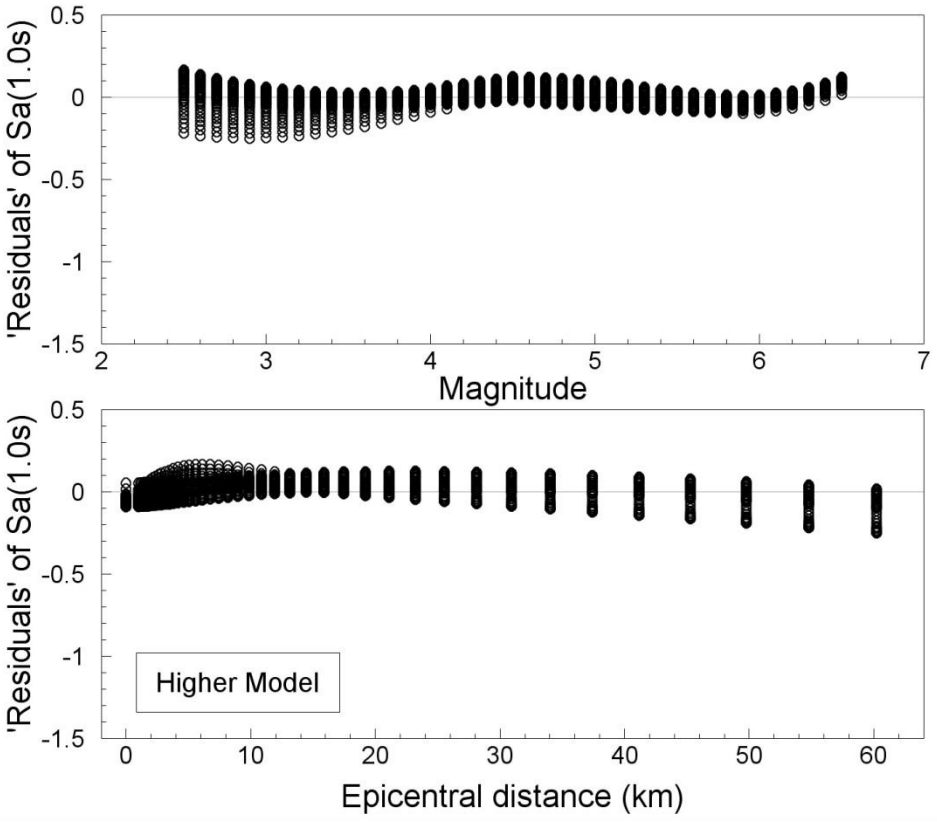


Figure 6.14. Computed 'residuals' of the stochastic simulations (Δσ = 100 bars) for Sa(1.0s) relative to the regression model in Eq.(6.1) and Table 6.3

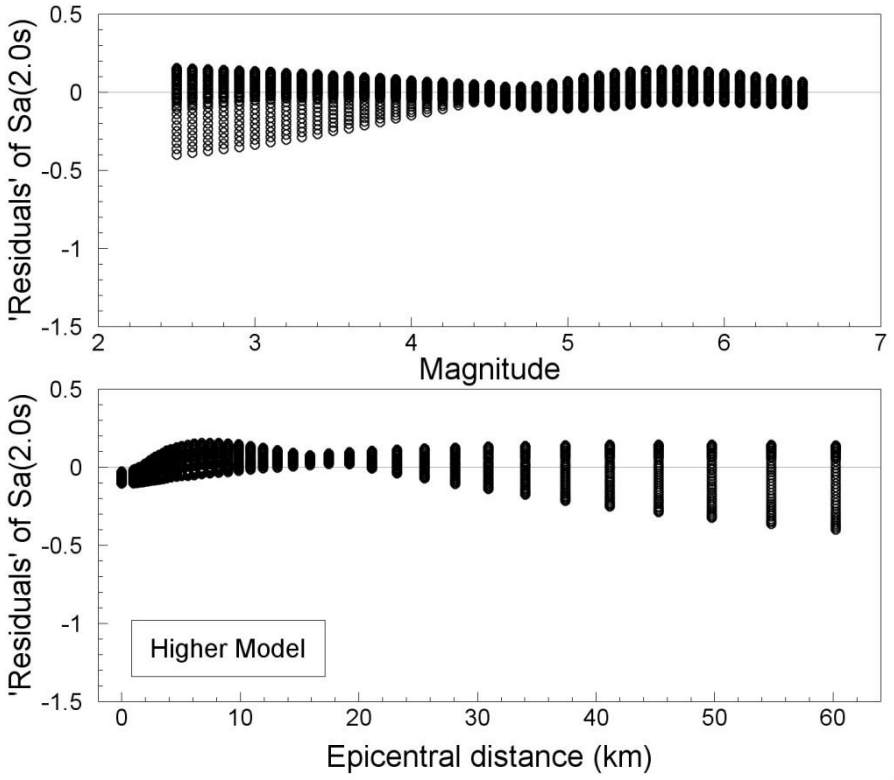


Figure 6.15. Computed 'residuals' of the stochastic simulations (Δσ = 100 bars) for Sa(2.0s) relative to the regression model in Eq.(6.1) and Table 6.3

The patterns observed in all of these plots suggest that a reasonable fit has been obtained, notwithstanding some fluctuations in the trends, which are the result of the simplified functional form of the regression model. The most striking feature, on first inspection, are the rather large negative 'residuals' observed at large distances (which also show up at the smaller magnitudes), particularly at the shorter oscillator periods. This is simply the result of the anelastic attenuation (Q) included in the stochastic simulations but not explicitly modelled in the empirical regression equation.

Figures 6.16 to 6.27 compare the median accelerations obtained from the regressions with the accelerations from the stochastic simulations, for various combinations of magnitude and distance. In all cases, the agreement is generally good for most of the ground-motion parameters across the ranges of magnitudes and distances considered. The greatest divergence between the stochastically-simulated motions and the median predictions from the regression models occur at shorter periods, especially at 0.2 seconds where the regression model is above the simulations at 0 km and then below them at 10 km, with the difference diminishing with increasing magnitude. There is also some divergence for the predictions of PGA, but in this case it sometimes persists at larger magnitudes. For the spectral accelerations at the longer periods (0.5, 1.0 and 2.0 seconds), the agreement is very good. These results are both interesting and encouraging, but they only reflect how well the regressions using the rather simple functional form of Eq.(6.1) are able to replicate the stochastic simulations rather than how well the final model represents the Groningen

recordings, given that the stochastic simulations for the central model were not a perfect match with the empirical model obtained directly from these data (Figures 5.24 to 5.27).

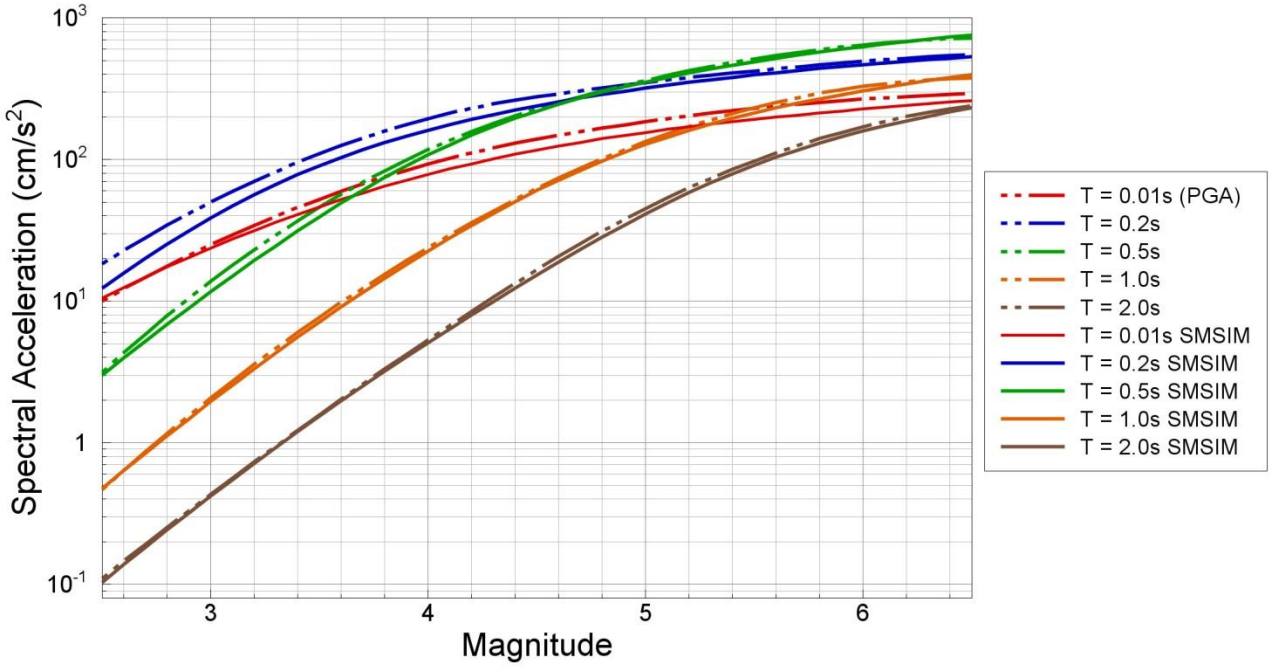


Figure 6.16. Comparison of the stochastic simulations for the central model with the median predictions from the regression of Eq.(6.1) and coefficients in Table 6.1 at 0 km

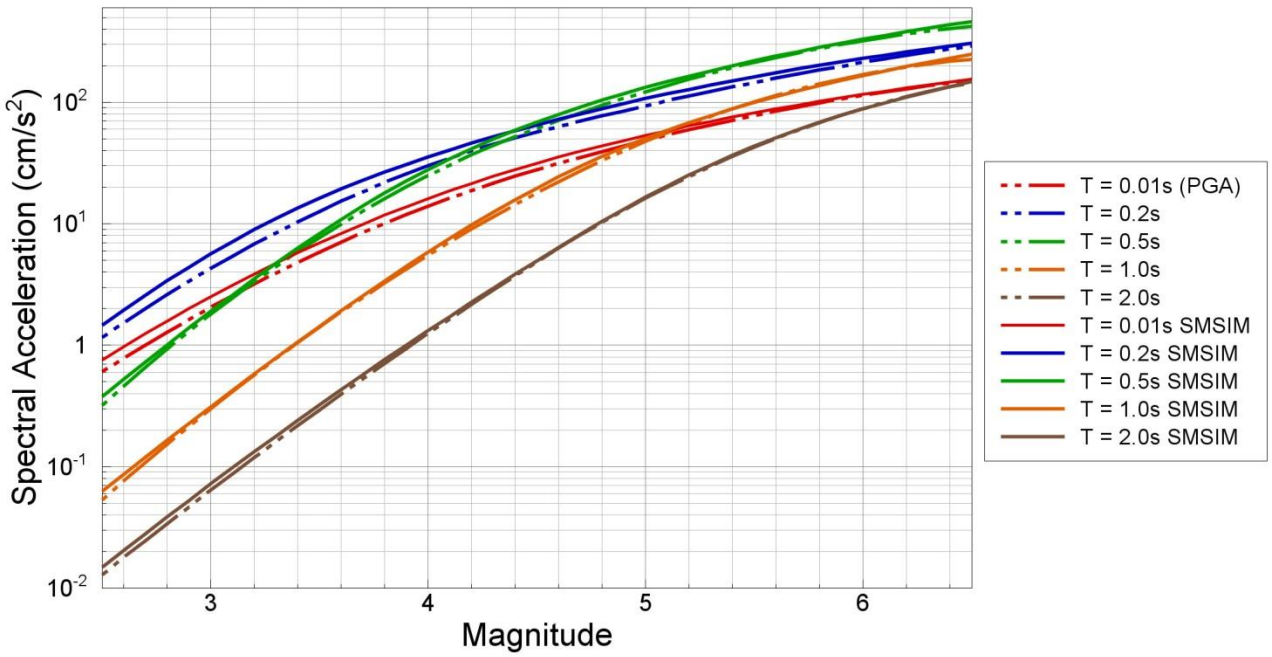


Figure 6.17. Comparison of the stochastic simulations for the central model with the median predictions from the regression of Eq.(6.1) and coefficients in Table 6.1 at 10 km

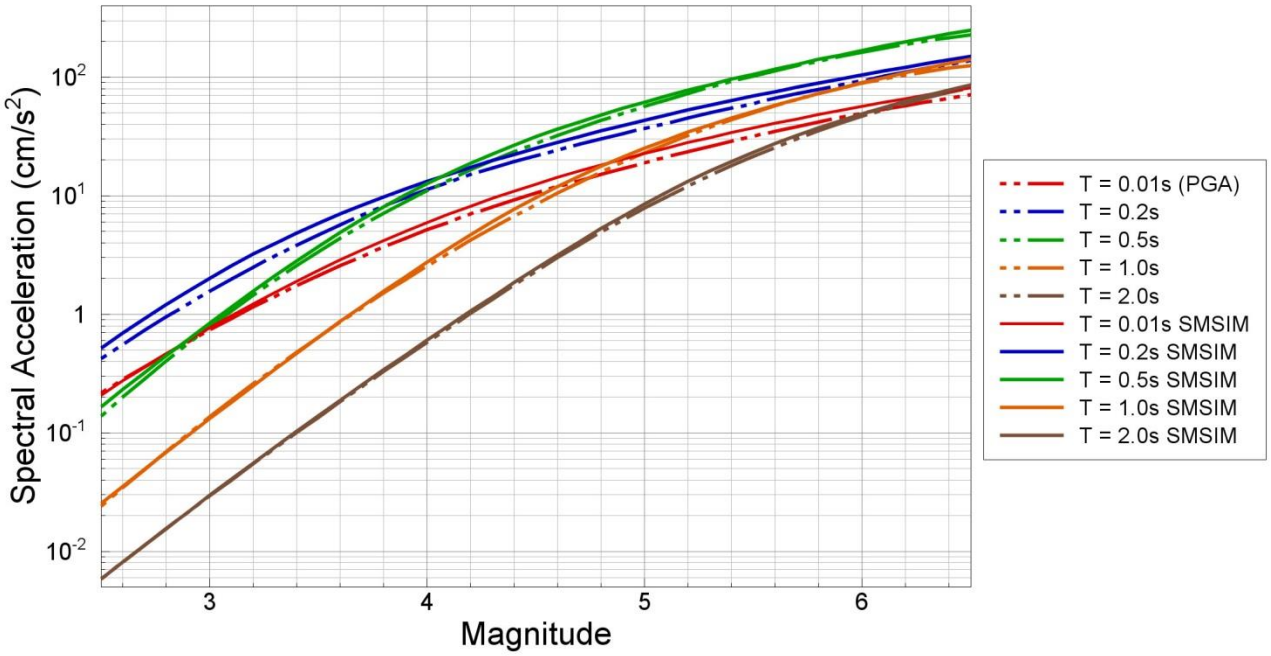


Figure 6.18. Comparison of the stochastic simulations for the central model with the median predictions from the regression of Eq.(6.1) and coefficients in Table 6.1 at 20 km

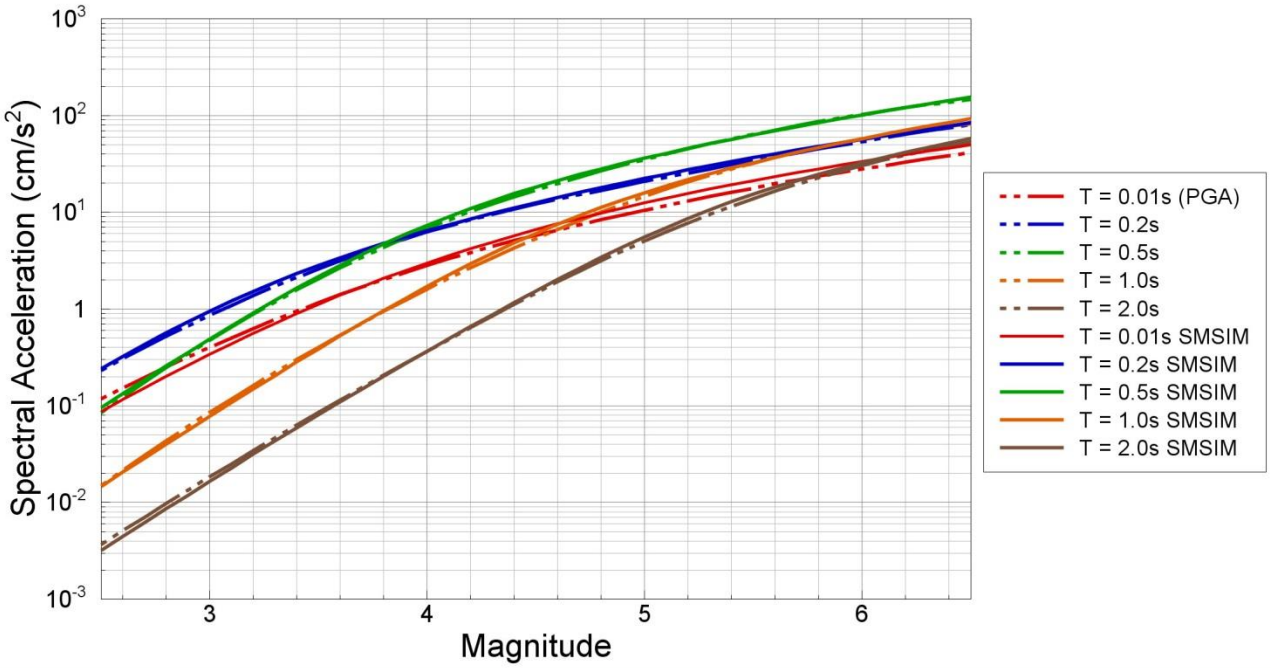


Figure 6.19. Comparison of the stochastic simulations for the central model with the median predictions from the regression of Eq.(6.1) and coefficients in Table 6.1 at 30 km

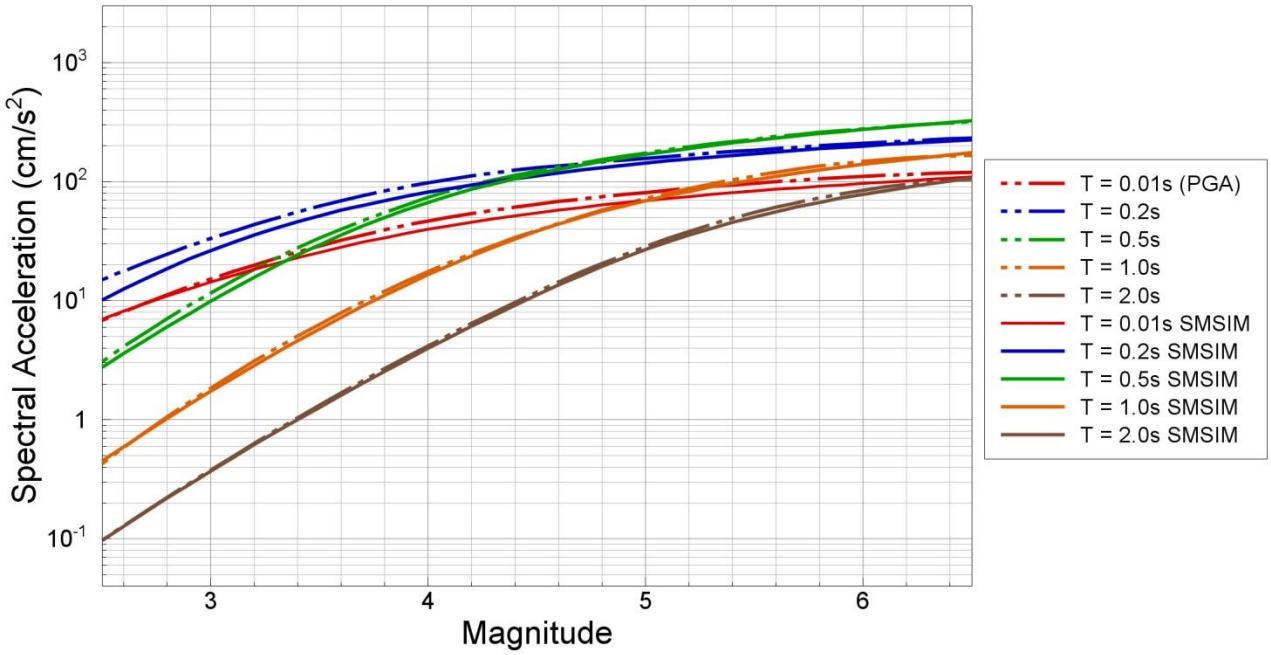


Figure 6.20. Comparison of the stochastic simulations for the lower model with the median predictions from the regression of Eq.(6.1) and coefficients in Table 6.2 at 0 km

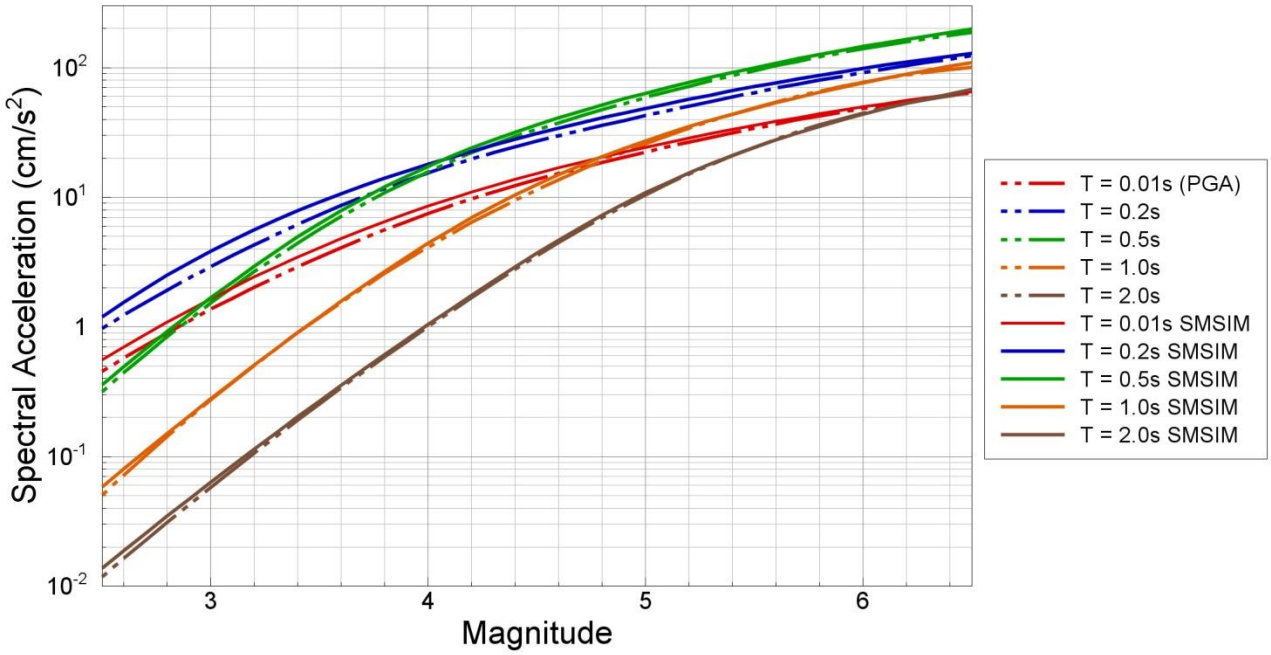


Figure 6.21. Comparison of the stochastic simulations for the lower model with the median predictions from the regression of Eq.(6.1) and coefficients in Table 6.2 at 10 km

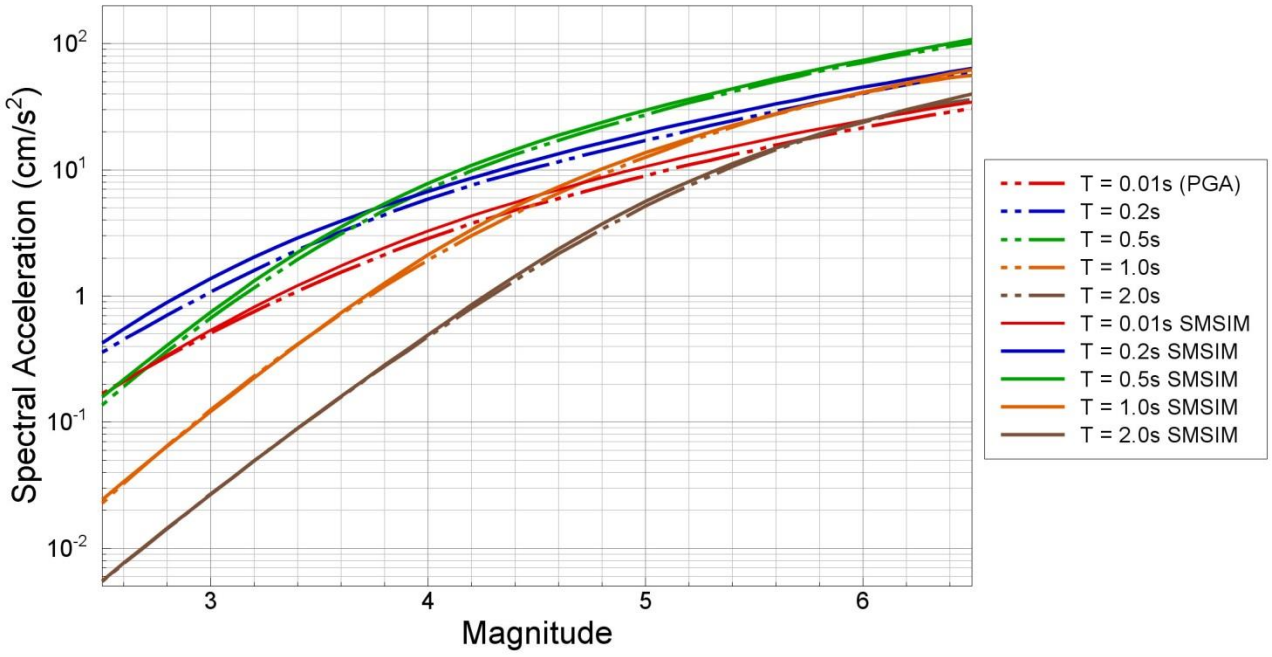


Figure 6.22. Comparison of the stochastic simulations for the lower model with the median predictions from the regression of Eq.(6.1) and coefficients in Table 6.2 at 20 km

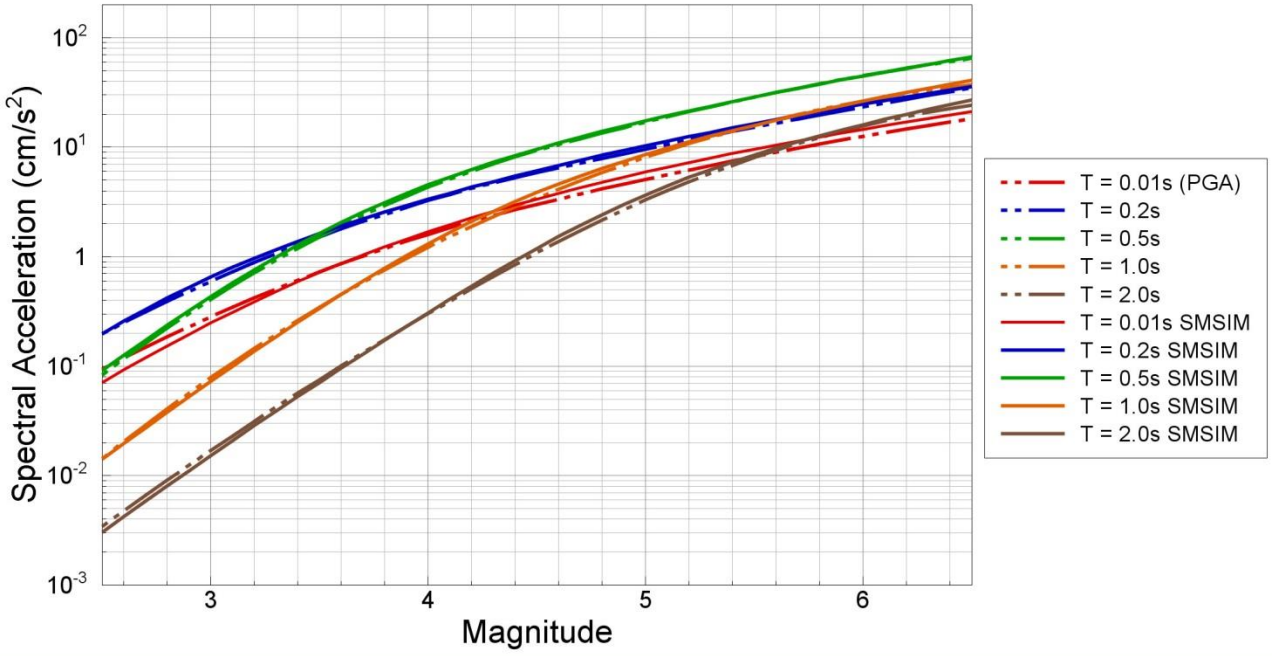


Figure 6.23. Comparison of the stochastic simulations for the lower model with the median predictions from the regression of Eq.(6.1) and coefficients in Table 6.2 at 30 km

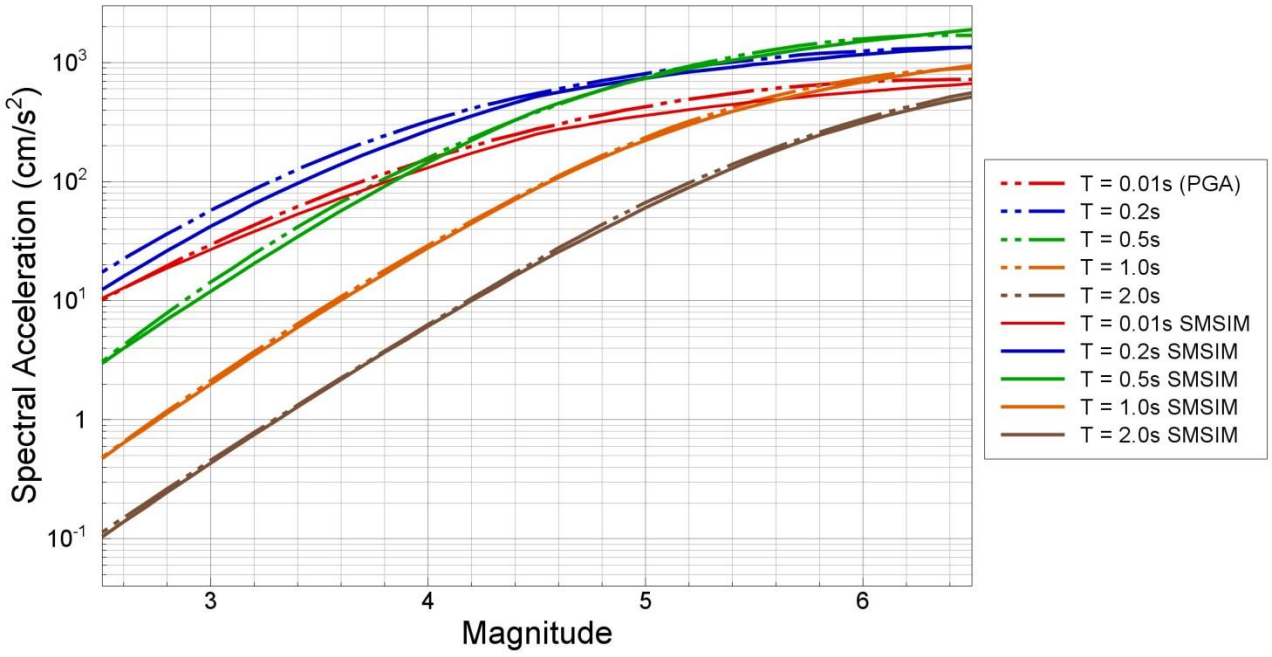


Figure 6.24. Comparison of the stochastic simulations for the higher model with the median predictions from the regression of Eq.(6.1) and coefficients in Table 6.3 at 0 km

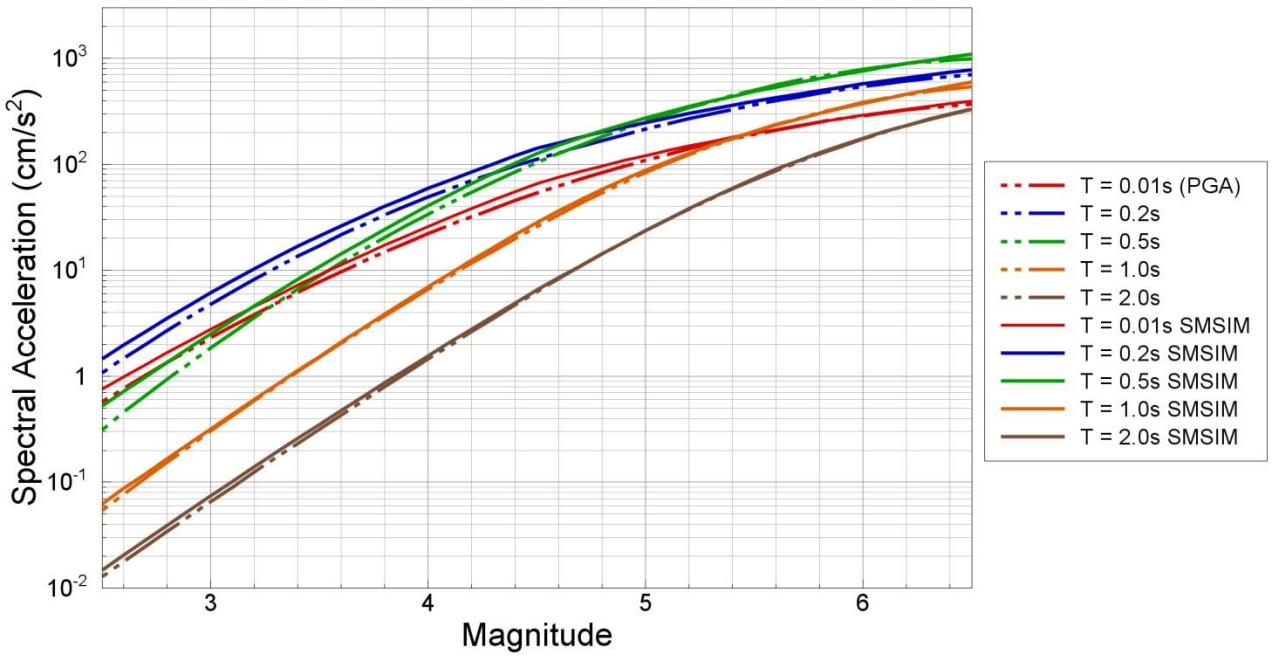


Figure 6.25. Comparison of the stochastic simulations for the higher model with the median predictions from the regression of Eq.(6.1) and coefficients in Table 6.3 at 10 km

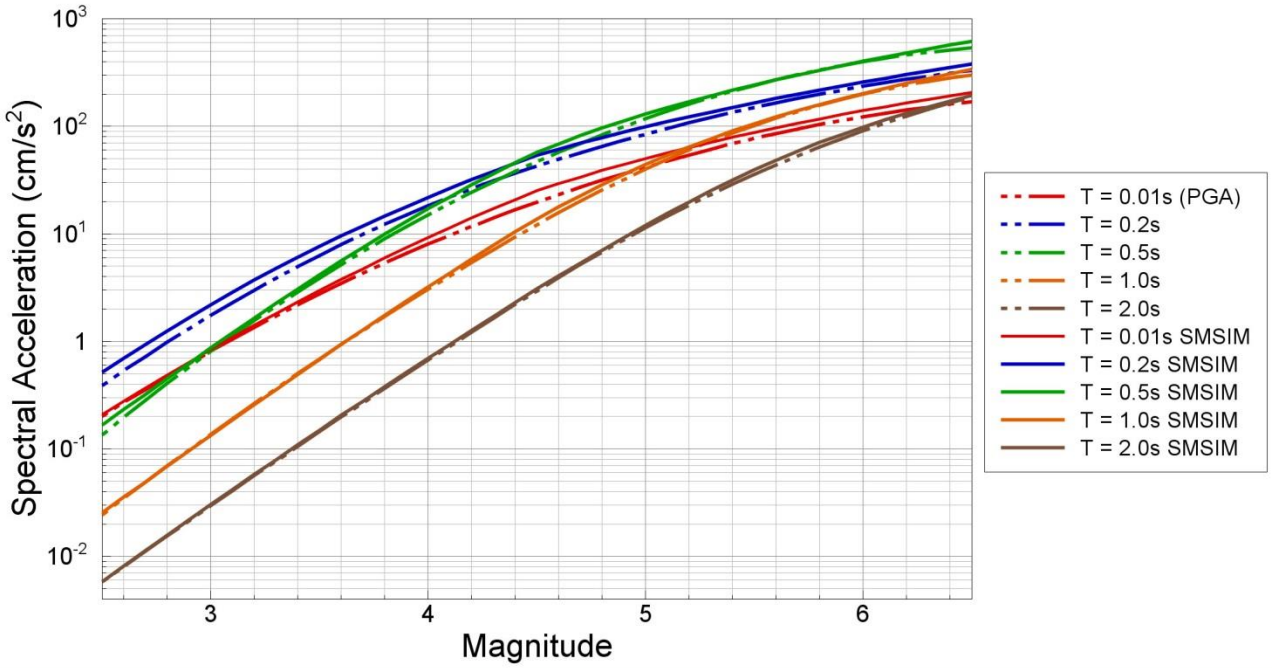


Figure 6.26. Comparison of the stochastic simulations for the higher model with the median predictions from the regression of Eq.(6.1) and coefficients in Table 6.3 at 20 km

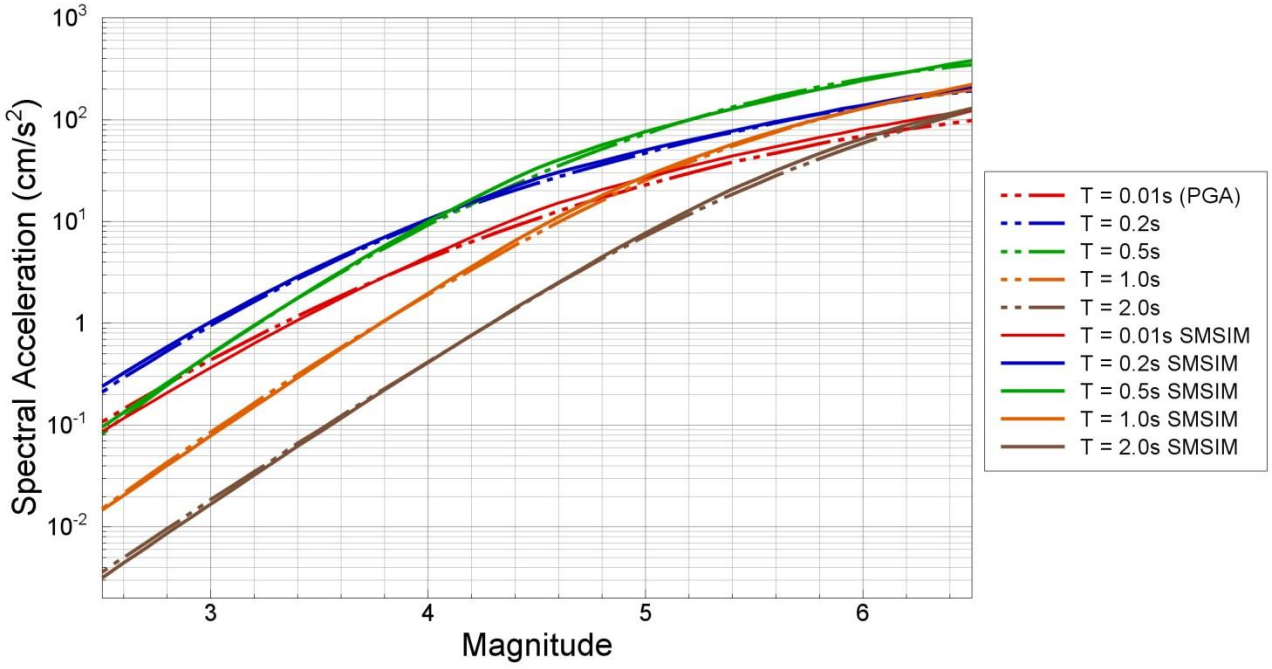


Figure 6.27. Comparison of the stochastic simulations for the higher model with the median predictions from the regression of Eq.(6.1) and coefficients in Table 6.3 at 30 km

For completeness, Figures 6.28 to 6.30 show the median predicted PGA values from the three different models plotted against epicentral distance for various magnitude values covering the range of applicability of the equations. Several features become clearly apparent from these plots, including the strongly non-linear scaling with magnitude and the magnitude dependence of the near-source saturation.

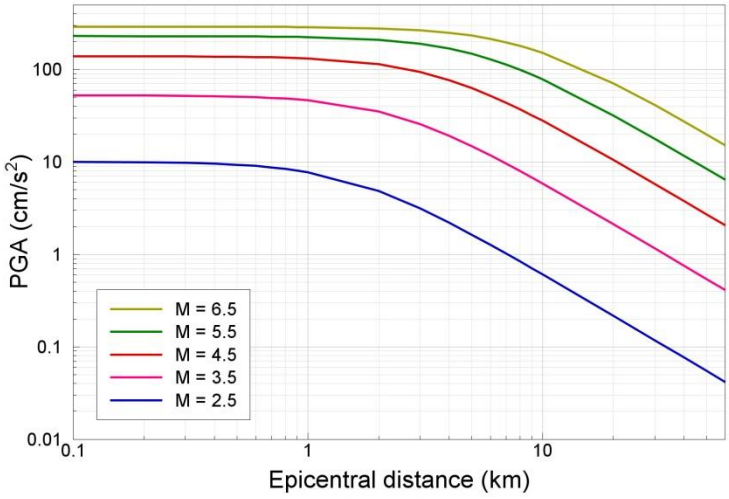


Figure 6.28. Predicted median values of PGA against distance, for different magnitudes, obtained from the central GMPE

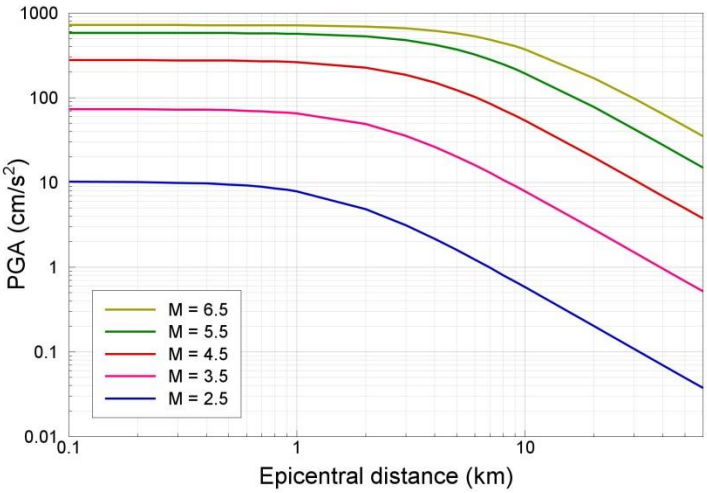


Figure 6.29. Predicted median values of PGA against distance, for different magnitudes, obtained from the higher GMPE

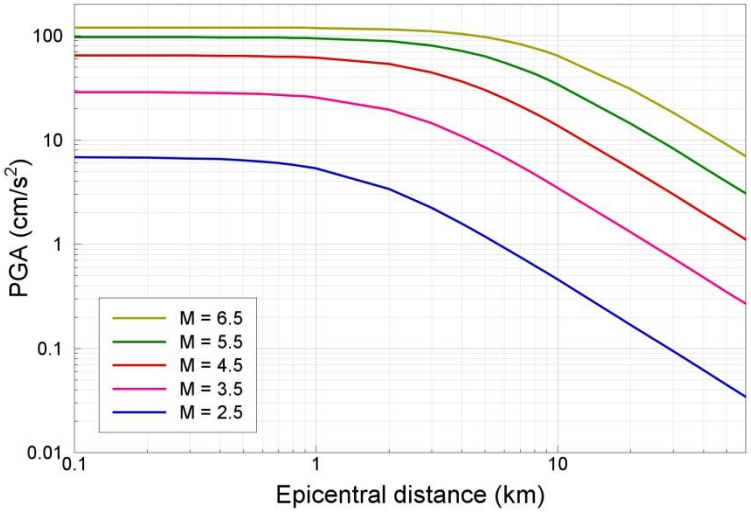


Figure 6.30. Predicted median values of PGA against distance, for different magnitudes, obtained from the lower GMPE

6.2. Sigma models for Version 1 GMPEs

The residuals of the Groningen recordings with respect to the new models were calculated, in the same way as the residuals obtained with respect to the European GMPEs that were presented in Section 3.3. Figures 6.31 to 6.35 show the inter- and inter-event residuals of the field data with respect to the central model presented in the previous section (Table 6.1) and compares these with the residuals obtained with respect to the empirical models derived directly from the data in Section 5.2; the latter residuals were previously shown in Figure 5.8 to 5.10, but are repeated here in order to facilitate the comparison.

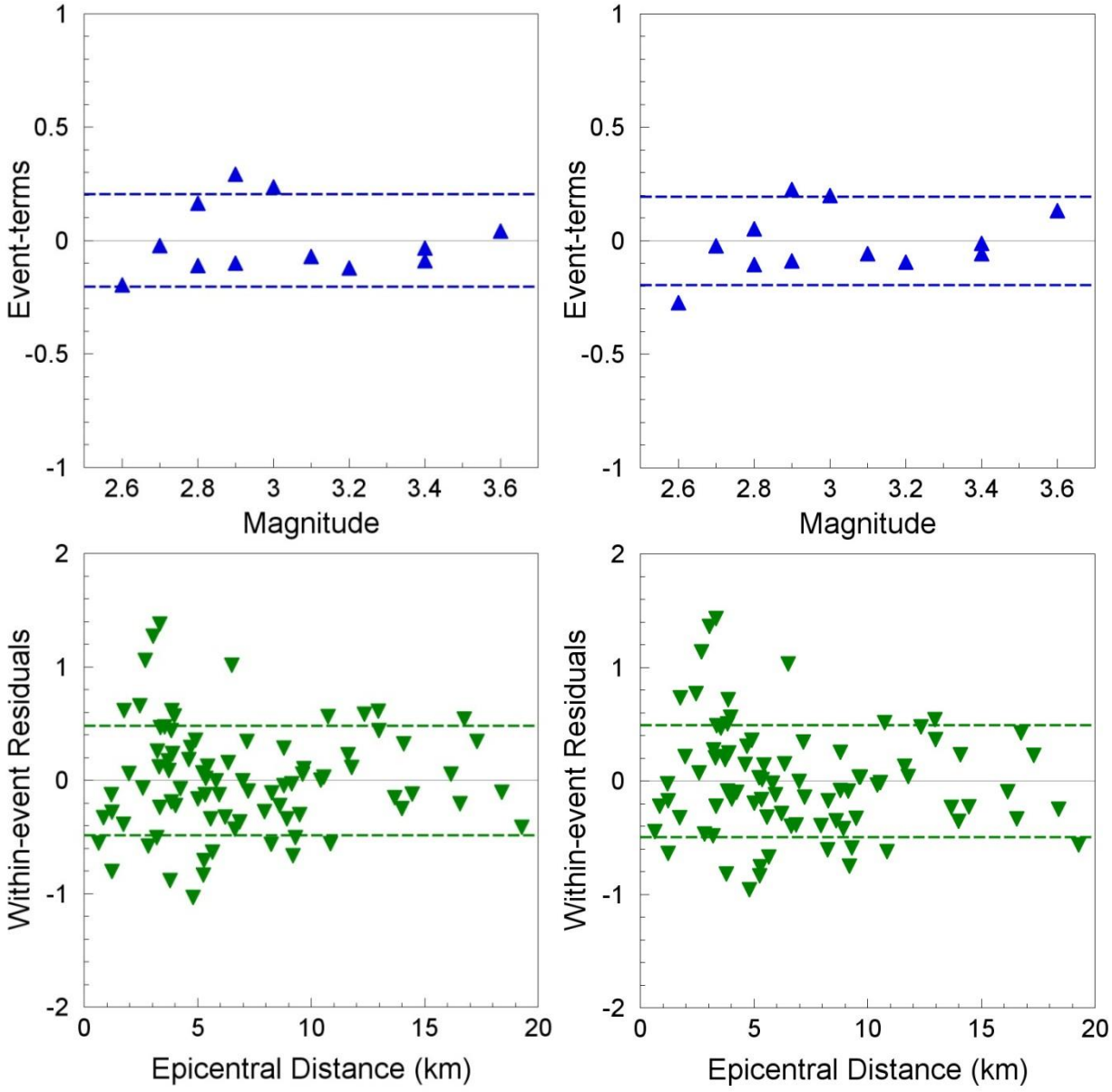


Figure 6.31. Residuals of the Groningen PGA values with respect to Eq.(5.9) and the coefficients in Table 5.1 (*left*) and with respect to Eq.(6.1) and the coefficients in Table 6.1 (*right*); the dashed lines indicate the inter- and intra-event standard deviations, respectively, in the upper and lower plots.

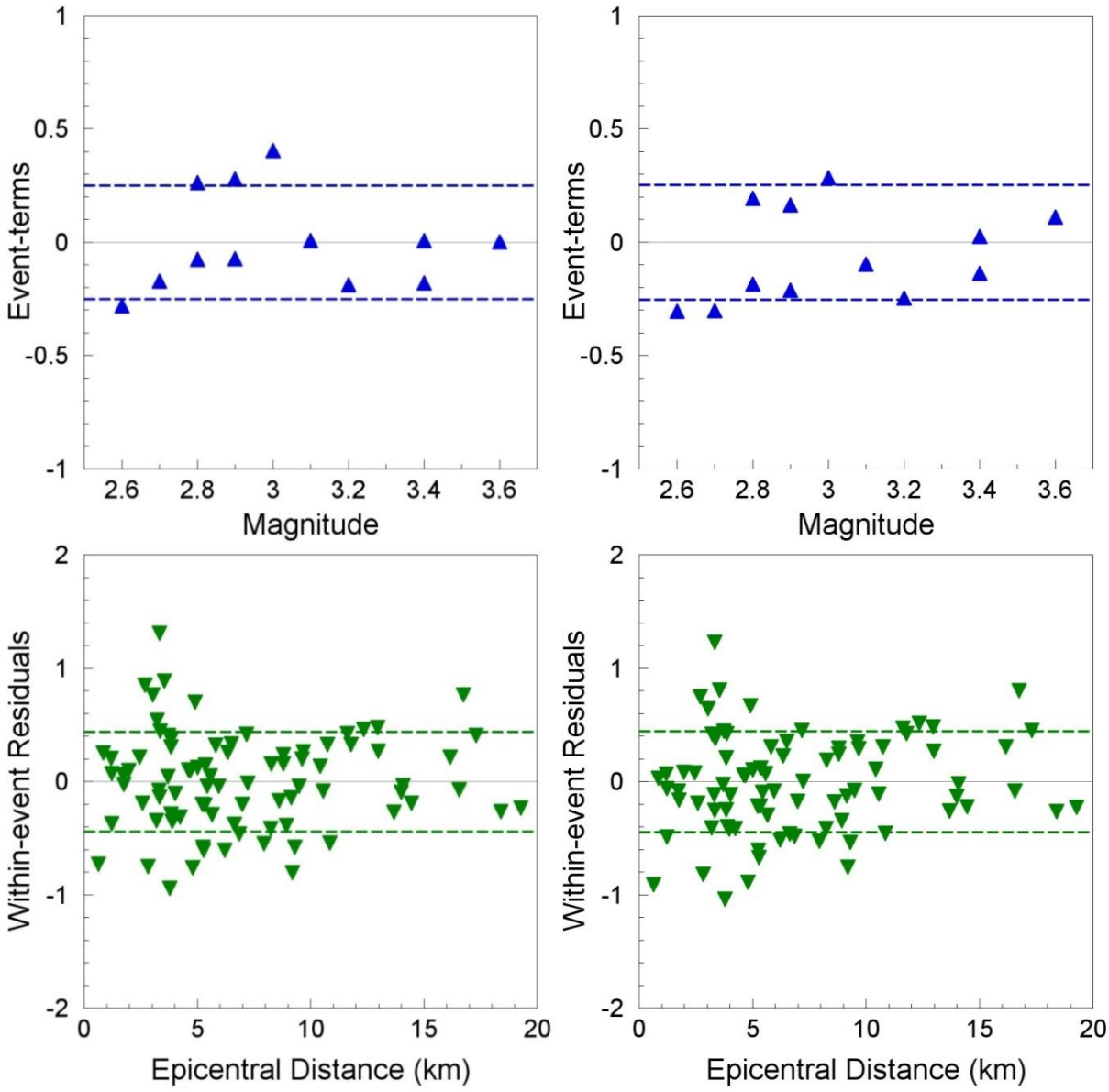


Figure 6.32. Residuals of the Groningen Sa(0.2s) values with respect to Eq.(5.9) and the coefficients in Table 5.1 (*left*) and with respect to Eq.(6.1) and the coefficients in Table 6.1 (*right*); the dashed lines indicate the inter- and intra-event standard deviations, respectively, in the upper and lower plots.

The intra-event residual plots for the Groningen data with respect to both the empirical models derived directly by regression analyses on the acceleration values and the final equations derived by regression on the stochastic simulations are practically indistinguishable on preliminary visual inspection, for all five ground-motion parameters. There are some differences, but these are rather small and the intra-event sigma (ϕ) terms are very similar for both models; this is also confirmed by inspection of the values listed in Table 6.4. The inter-event residuals, however, show significantly greater dispersion for the final model rather than for the preliminary empirical model and this generally results in larger values of the inter-event sigma, τ, especially at the longer periods of 0.5 and 1.0 seconds; again, this is confirmed by the listed values in Table 6.4.

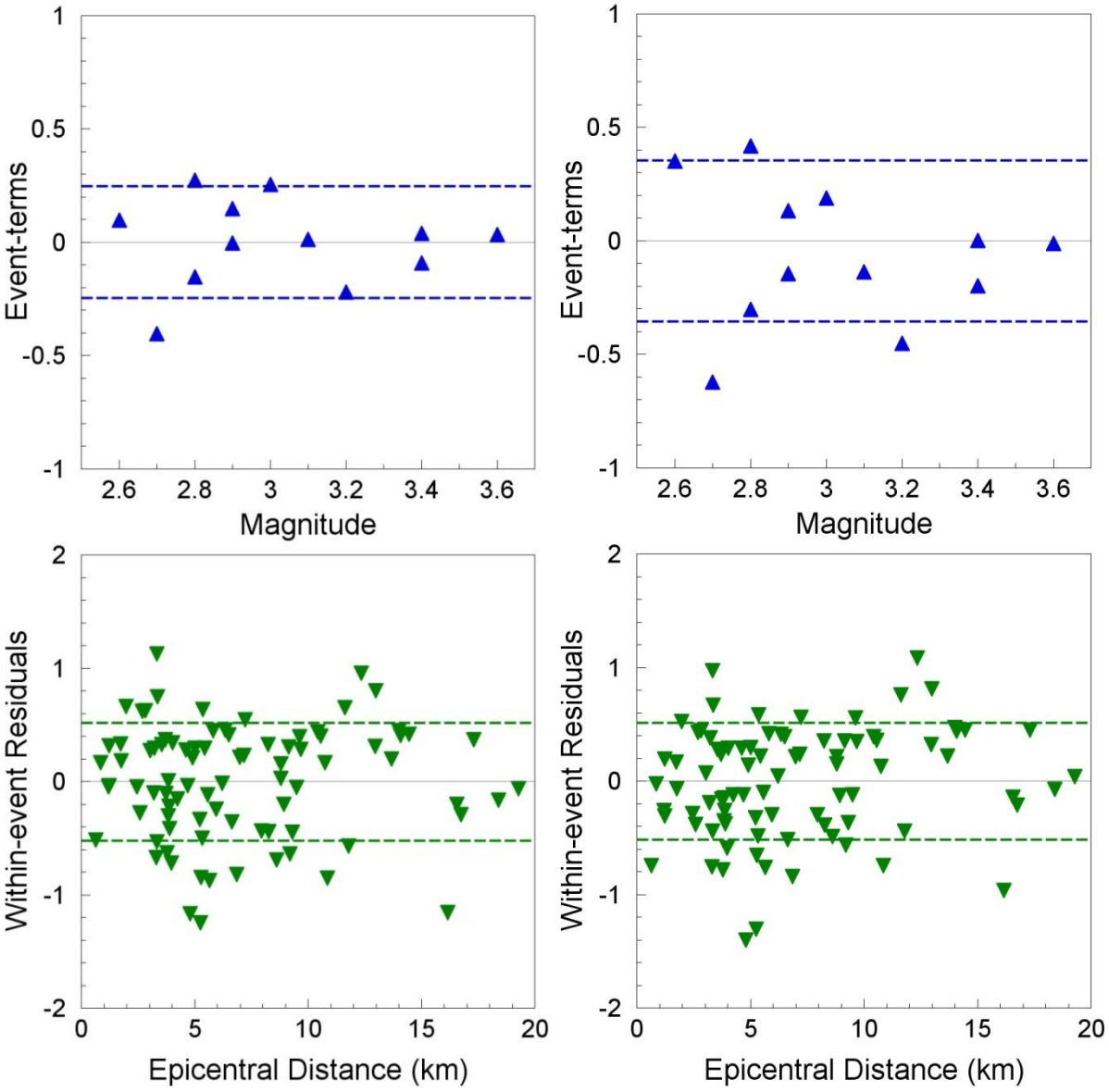


Figure 6.33. Residuals of the Groningen Sa(0.5s) values with respect to Eq.(5.9) and the coefficients in Table 5.1 (*left*) and with respect to Eq.(6.1) and the coefficients in Table 6.1 (*right*); the dashed lines indicate the inter- and intra-event standard deviations, respectively, in the upper and lower plots.

The larger inter-event variability could be partly explained by the fact that in the empirical regressions, the optimal event-term—which could be interpreted as representing how much more or less energetic each earthquake source is in comparison to other events of the same magnitude, which in essence is a reflection of the stress drop—is estimated for each earthquake whereas in the stochastic simulations a constant $\Delta\sigma$ is assumed for all the earthquakes. However, if this were the only factor influencing the differences, they should manifest primarily at shorter periods and not, as is the case, at 0.5 and 1.0 seconds. This indicates that there is clearly scope for refining these preliminary models, but this will be done in any case in the near future, using both an expanded Groningen database (see Section 8.3) with site characterisation of the recording stations (see Section

8.4), and the stochastic simulations will benefit from additional constraint provided by full waveform modelling (see Section 8.2). For the preliminary model, the patterns of the residuals do not suggest that there is any reason to reject the model—especially since the biases in all cases are found not to be statistically significant—and in view of the fact that the total sigma values are found to not be markedly larger for the final model than for the initial empirical model (Figure 6.36).

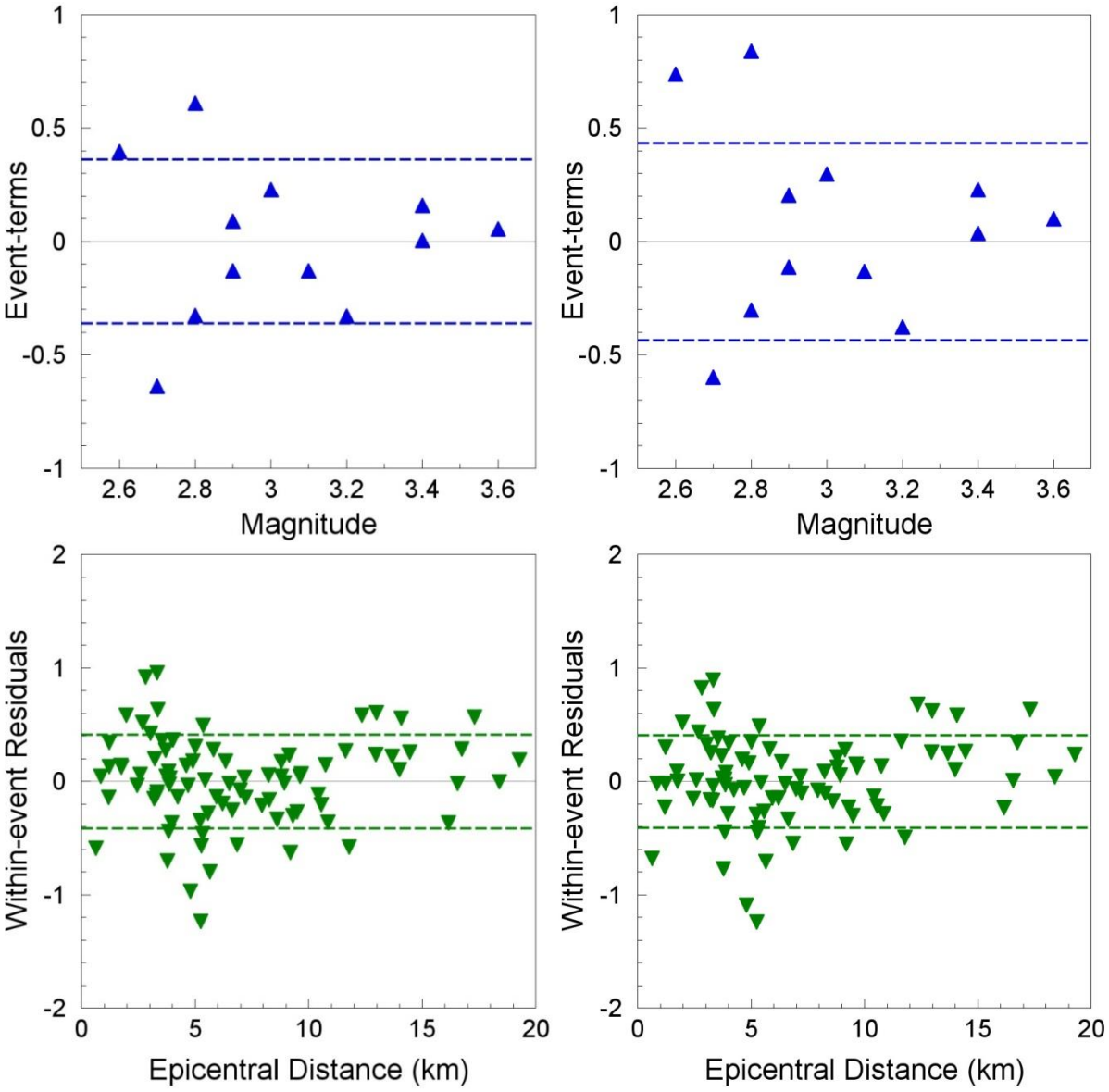


Figure 6.34. Residuals of the Groningen Sa(1.0s) values with respect to Eq.(5.9) and the coefficients in Table 5.1 (*left*) and with respect to Eq.(6.1) and the coefficients in Table 6.1 (*right*); the dashed lines indicate the inter- and intra-event standard deviations, respectively, in the upper and lower plots.

As noted above, the sigma values with respect to the final central model are listed in Table 6.4, together with those from the original empirical model (from Table 5.1), for comparison,

and those obtained from both of the alternative new models. The total sigma values from all four models are plotted in Figure 6.36. In all cases, unsurprisingly, the variability from direct regression on the data is the lowest value, but the differences for the models obtained from the stochastic simulations are only slightly larger. For PGA and Sa(0.2s), only the lower alternative model has sigma values that are not almost identical to those from the empirical model; at the longer periods, all the sigma values are very similar and somewhat larger than those from the empirical model. For additional comparison, Table 6.4 also includes the values from the R_{epi}–based model of Akkar *et al.* (2014a), which are appreciably larger at all periods; this is also confirmed visually in Figure 6.36.

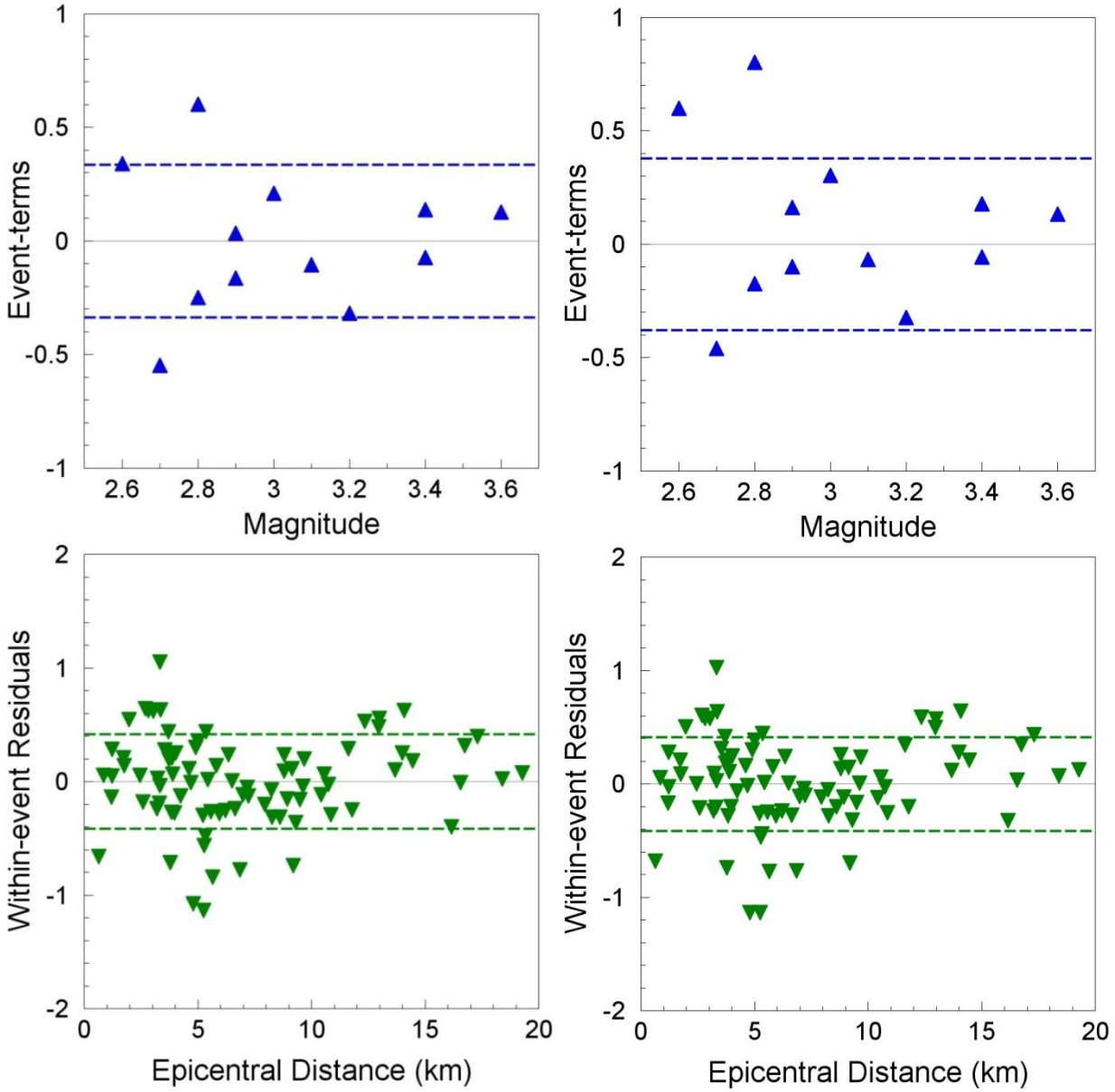


Figure 6.35. Residuals of the Groningen Sa(2.0s) values with respect to Eq.(5.9) and the coefficients in Table 5.1 (*left*) and with respect to Eq.(6.1) and the coefficients in Table 6.1 (*right*); the dashed lines indicate the inter- and intra-event standard deviations, respectively, in the upper and lower plots.

Table 6.4. Standard deviations of Eq.(5.9) obtained from regressions

Model	Sigma	PGA	Sa(0.2s)	Sa(0.5s)	Sa(1.0s)	Sa(2.0s)
Central	τ	0.1953	0.2539	0.3551	0.4349	0.3779
	ϕ	0.4918	0.4454	0.5146	0.4081	0.4133
	σ	0.5291	0.5127	0.6253	0.5964	0.5600
Lower	τ	0.2334	0.2908	0.3478	0.4261	0.3645
	ϕ	0.4975	0.4460	0.5152	0.4081	0.4141
	σ	0.5495	0.5324	0.6216	0.5900	0.5517
Higher	τ	0.1902	0.2509	0.3695	0.4400	0.3894
	ϕ	0.4874	0.4444	0.5142	0.4081	0.4129
	σ	0.5232	0.5103	0.6331	0.6001	0.5675
Empirical	τ	0.2039	0.2514	0.2467	0.3612	0.3359
	ϕ	0.4831	0.4404	0.5208	0.4121	0.4170
	σ	0.5243	0.5071	0.5762	0.5479	0.5355
Akkar *at al.* (2014a) R_{epi} model	τ	0.3581	0.4160	0.3965	0.3965	0.3734
	ϕ	0.6375	0.6851	0.6922	0.6922	0.7333
	σ	0.7312	0.8015	0.7977	0.7977	0.8229

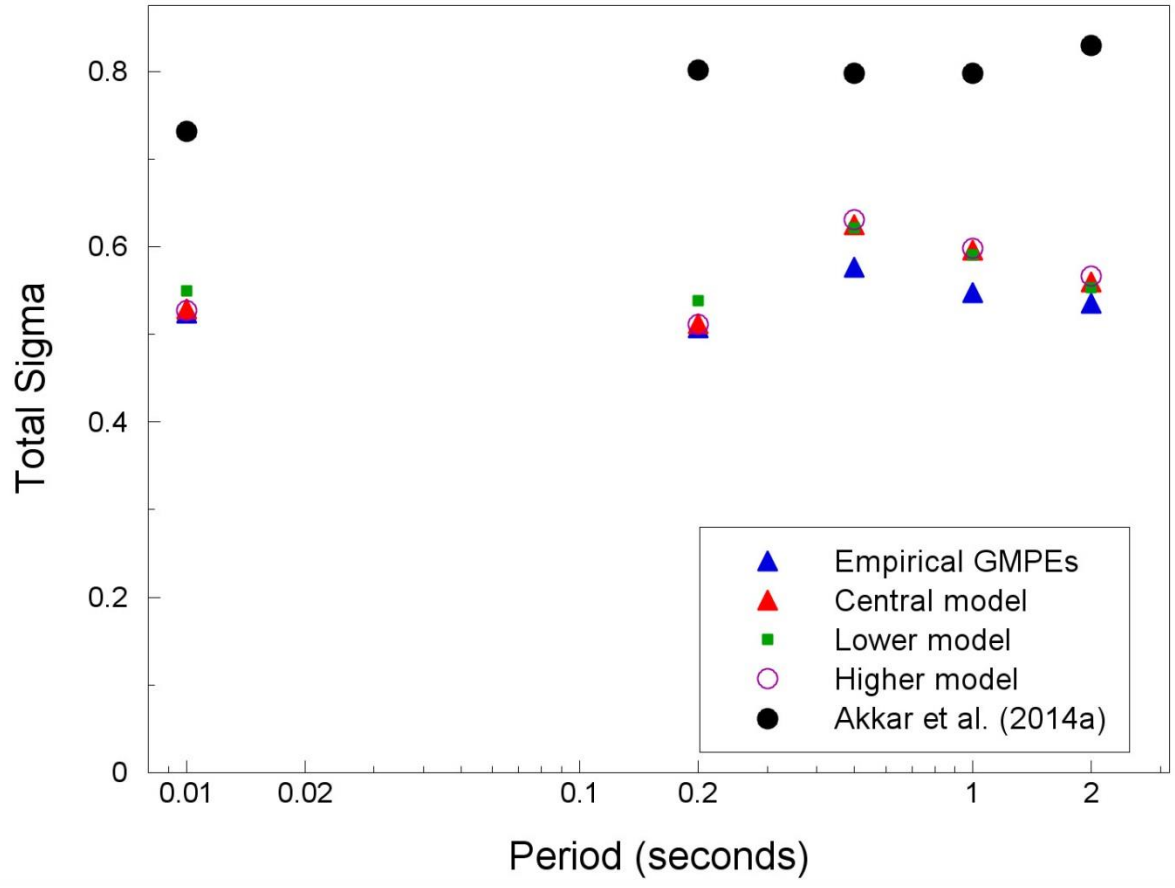


Figure 6.36. Total sigma values calculated from the residuals of the Groningen data with respect to the empirical GMPEs presented in Section 5.2 and the three final models based on the stochastic simulations (Section 5.4), and those from the R_{epi}-based model of Akkar *et al*. (2014a).

Just as in the development of the median ground-motion prediction models described in Section 6.1, there must also be consideration for the epistemic uncertainty in the development of the sigma model, in particular with regards to the sigma values at larger magnitudes. The reasoning followed in the development of the sigma model is as follows, starting with a reminder of the basic breakdown of the total variability into between-earthquake (inter-event) and within-earthquake (intra-event) components, represented by τ and ϕ respectively:

$$\sigma = \sqrt{\tau^2 + \phi^2} \tag{6.2}$$

The first issue to be addressed is the fact that the GMPEs are being derived in terms of epicentral distance, which means that the earthquake source is being treated as a point. Whilst this is perfectly appropriate for the small-magnitude earthquakes in the current Groningen database, it is recognised that this is an unrealistic model for larger earthquakes having extended rupture sources. Although alternative GMPEs will be developed using a distance metric based on extended ruptures (see Section 8.1), for the Version 1 hazard and risk model, the R_{epi}-based equations are used. Therefore, it is appropriate to apply a penalty to the sigma values for larger earthquakes with extended fault ruptures that will be represented as point sources. Since this penalty is related to the geometry of the source and the spatial variation of ground motions, it is effectively a correction to the intra-event variability and is designated $\delta\phi$; the penalty is applied, by summing variances, to the intra-event variability inferred from the small-magnitude Groningen data, ϕ_{SM}, hence Eq.(6.2) becomes:

$$\sigma = \sqrt{\tau^2 + (\phi_{SM}^2 + \delta\phi^2)} \tag{6.3}$$

The derivation of the sigma penalty is fully explained in Appendix II of this report, and is summarised here for completeness. The magnitude- and distance-dependent adjustment to the intra-event variability is defined as follows:

$$\delta\phi = SF.\frac{\varphi(z)}{\sigma_Z} \qquad M \geq 4 \quad \text{and} \quad R_{epi} > 0 \tag{6.4a}$$

$$\delta\phi = 0 \qquad M < 4 \quad \text{or} \quad R_{epi} = 0 \tag{6.4b}$$

where SF is the magnitude-dependent scaling factor, expressed as follows:

$$SF = \beta_1(M-4) + \beta_2(M-4)^2 \tag{6.5}$$

and $\varphi[\ ]$ is the normal probability density function, which is given by the following expression:

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-z^2}{2}\right) \tag{6.6}$$

The argument of this expression is given by:

$$z = \frac{\ln(R_{epi}) - \mu_Z}{\sigma_Z} \tag{6.7}$$

and the parameters of this expression are given by:

$$\mu_Z = \beta_3 + \beta_4(M - 6.75) + \beta_5(M - 6.75)^2 \tag{6.8}$$

$$\sigma_Z = \beta_6 \tag{6.9}$$

The coefficients of Eqs.(6.3-6.9) are presented in Table 6.5. The condition specified in Eq.(6.4) that the correction is only applied for earthquakes of magnitude 4 or greater since for smaller events the rupture dimensions may be assumed to be sufficiently small for the two distance metrics (R_{epi} and R_{JB}) to be considered equivalent and thus for no adjustment to the variability to be needed. Figure 6.37 shows the adjustment to the within-event variability for PGA as a function of distance for a range of magnitudes; the plot confirms how as the magnitude approaches 4, the adjustments become vanishingly small.

Table 6.5. Coefficients of Eqs.(6.3-6.9) for the intra-event component of the variability

	PGA	**Sa(0.2s)**	**Sa(0.5s)**	**Sa(1.0s)**	**Sa(2.0s)**
ϕ_{SM}	0.4918	0.4454	0.5146	0.4081	0.4133
β_1	0.20380	0.20284	0.20761	0.21116	0.21290
β_2	0.073419	0.080624	0.044808	0.018152	0.005130
β_3	3.39511	3.39511	3.39511	3.39511	3.39511
β_4	0.70978	0.70978	0.70978	0.70978	0.70978
β_5	0.0090045	0.0090045	0.0090045	0.0090045	0.0090045
β_6	0.8970	0.9025	0.8688	0.8297	0.7994

To explain the form of the correction, it was found to have a shape—when plotted against the logarithm of distance—that is very well approximated by the shape of a normal distribution. Therefore, the equations presented above effectively define the values of *ln*(R_{epi}) that represent the mean and standard deviation of that distribution, with the mean increasing with the magnitude. Although the adjustments shown in Figure 6.37 reach rather large values, it should be borne in mind that the full set of curves is shown for illustrative purposes only and that the largest magnitude considered in the hazard and risk model for Groningen is **M** 6.5. Moreover, the adjustments are added to the small-magnitude estimate of within-event variability as variances hence the final impact on the

total standard deviation is modest even for this largest value of magnitude considered in the hazard and risk calculations. This is illustrated by Figure 6.38, which shows the effect of the adjustments on the total sigma for PGA as a function of distance and magnitude.

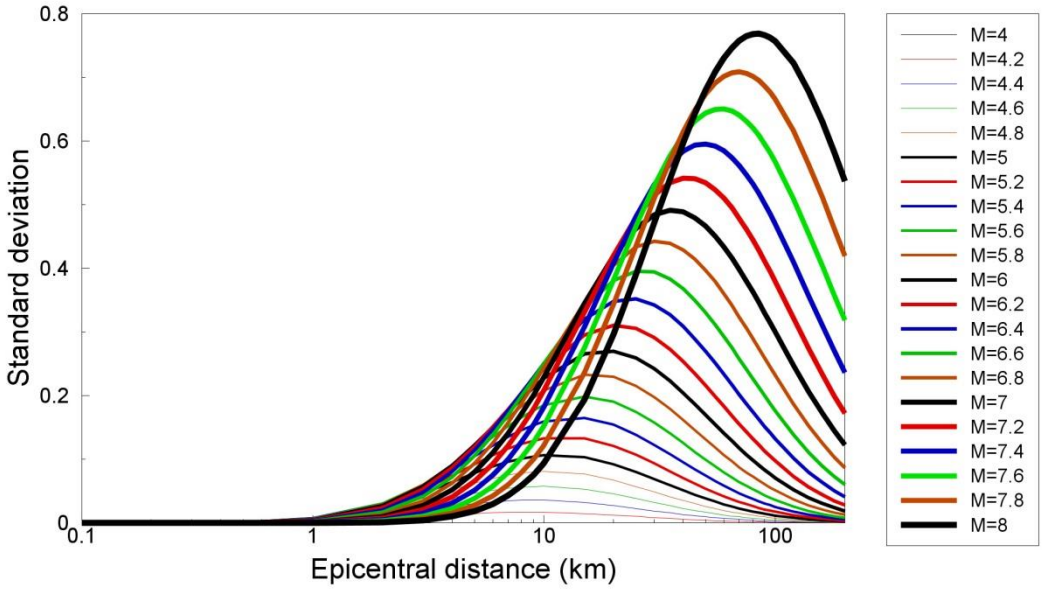


Figure 6.37. The values of the correction $\delta\phi$ for PGA predictions as a function of distance for a range of magnitude values

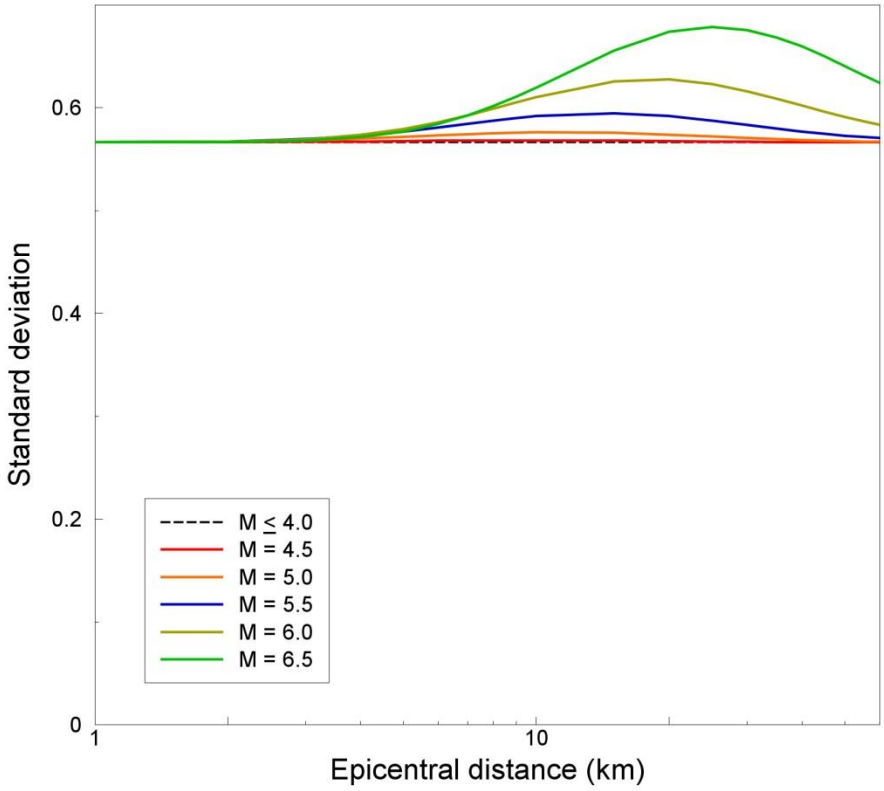


Figure 6.38. Total sigma on PGA as a function of magnitude and distance

To complete the intra-event standard deviation model, we also need to specify the value of the small-magnitude intra-event sigma, ϕ_{SM}, for which the following considerations are made. The values obtained from the empirical regressions on the Groningen data are potentially under-estimated because of the small sample size. On the other hand, 85 records is actually not a very sparse sample for the M-R space of 1.1 magnitude unit and 20 km covered by the data. The within-event variability may also be under-estimated since only a limited number of site profiles are currently sampled (*i.e*., the recording station sites) with respect to the full variation over the field. In view of these considerations, the final decision was to adopt the values of intra-event sigma obtained from the central model—fitted to the stochastic simulations—since it results in values that are a little larger than those obtained from the direct empirical regressions (Figure 6.36), particularly at longer periods. These chosen values of ϕ_{SM} are included in the first line of Table 6.5.

Several modern GMPEs invoke heteroskedastic models for sigma, in which the standard deviation is smaller for large-magnitude earthquakes than for smaller earthquakes (*e.g*., Abrahamson *et al*., 2008; Strasser *et al*., 2009). Confidence in our sigma model for the small magnitude range covered by the Groningen data could lead us to propose a decrease rather than increase in sigma for larger magnitudes by invoking the patterns seen in such heteroskedastic sigma models, or at least to off-set the increase being modelled as the within-event variability penalty for using a point-source distance metric. However, instead we choose not to make any such reduction of the sigma explicitly, noting instead that the variability at larger magnitudes may actually be slightly greater than implied by our geometric penalty as a result of variability in the rupture process along the fault. Any reduction of sigma at larger magnitudes is therefore assumed to cancel out any excessive simplification in our adjustment model.

For the inter-event sigma, the value of τ obtained from the empirical regressions may be considered a lower bound estimate, especially since the sample contains only 12 earthquakes and in view of the fact that the four most recent events show distinctly different event terms (which serves as a warning that surprises may be expected as additional data are acquired from the expanding networks of recording instruments in the field). On the other hand, the inter-event sigmas associated with the GMPE of Akkar *et al.* (2014a) for tectonic earthquakes—combining data from several regions and source types—may be considered a pessimistic upper bound for the inter-event sigma from larger earthquakes in the field. This leads to a selection of three alternative models, these two bounding values and a central model based on τ values that lie mid-way between these extremes (Table 6.6).

Table 6.6. Alternative models for the inter-event component of the variability

Model	PGA	Sa(0.2s)	Sa(0.5s)	Sa(1.0s)	Sa(2.0s)
Higher	0.3581	0.4160	0.3965	0.3965	0.3734
Middle	0.2810	0.3337	0.3216	0.3789	0.3547
Lower	0.2039	0.2514	0.2467	0.3612	0.3359

For some applications the complete response spectrum may need to be defined, which presents a challenge since at this early stage GMPEs have only been developed for five response periods. Guidance on the spectral shape to allow interpolation of the median predictions is provided in Section 6.5. To obtain standard deviations at the missing periods, it is recommended to apply the following steps, based on the shape of the predicted spectra and observed patterns from other GMPEs:

- For periods up to 0.04 s, use the value at 0.01 s (*i.e*., for PGA)
- For longer periods, use linear interpolation against the logarithm of period
- For periods beyond 2 second, the sigma value at this period are kept constant

Examples of the resulting standard deviations are shown in Figure 6.39.

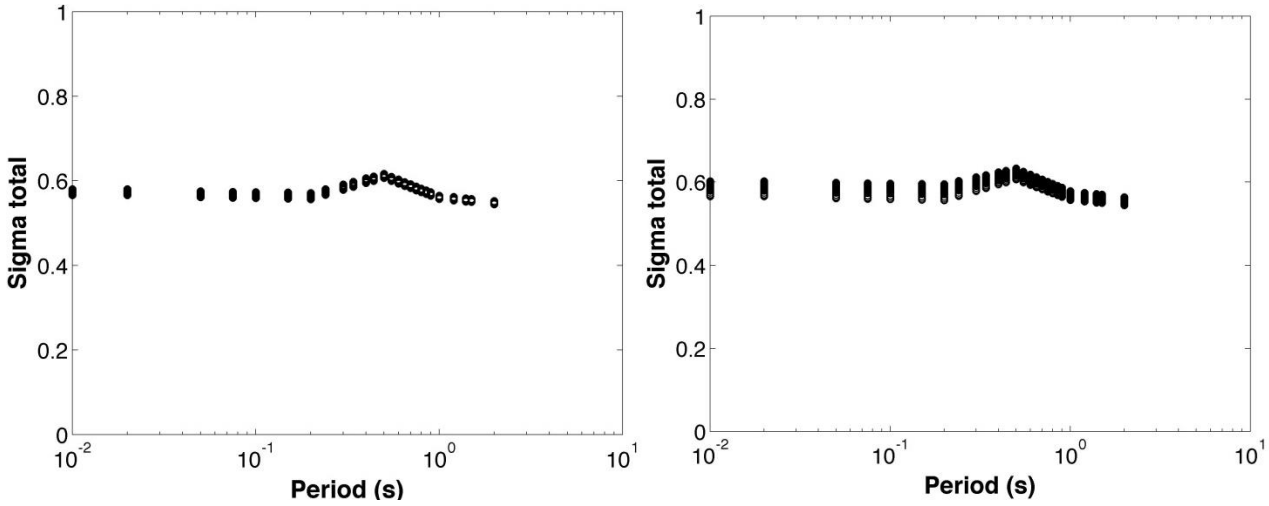


Figure 6.39. Interpolated values of total sigma for a range of distances and for magnitudes **M** 4.5 (*left*) and **M** 5.5 (*right*) (Courtesy of Helen Crowley)

6.3. Logic-tree weights for Version 1 ground-motion model

From Sections 6.1 and 6.2 we have two 3-branch models capturing both our best estimates and the range of epistemic uncertainty of the predicted median values of spectral acceleration and the associated sigma values. Following standard practice in site-specific PSHA studies (*e.g*., Bommer *et al*., 2015), the procedure would now be to assign weights to each of the set of branches—three weights summing to unity for the median branches and another three, also summing to one, for the sigma branches—and thus generate a logic-tree with 9 branches, the weights on which would be obtained from the product of the weights on the median and sigma branches represented in the individual combination. However, for the Version 1 risk model—in which the potential loss of life due to damage in approximately 250,000 buildings is to be calculated using Monte Carlo simulations—it is necessary to maintain the logic-tree as simple as possible. Therefore, the final logic-tree will consist of three branches which will be formed by pairing the models for the medians and for sigma values. In order to have three models that collectively span

a wide range of uncertainty and also are distinctly different—not overlapping, in accordance with current thinking on ground-motion logic-trees (Atkinson *et al*., 2014)—the decision is to couple the larger variability estimates with the models predicting higher median values, and vice versa. The formulation, and the associated weights, are summarised in Table 6.7.

Table 6.7. Branches for the Version 1 GMPE logic-tree

Branch	**Median**	τ	ϕ	**Weight**
1	Higher (Table 6.3)	Higher (Table 6.6)	Table 6.5	0.3
2	Central (Table 6.1)	Middle (Table 6.6)	Table 6.5	0.5
3	Lower (Table 6.2)	Lower (Table 6.6)	Table 6.5	0.2

The weight of 0.5 on branch #2 reflects the fact that this is the best estimate model, being based on a good fit to the Groningen data at small magnitudes and extrapolations using a slightly conservative estimate of the median stress drop at larger magnitudes. The sigma value is based on measured intra-event variability, slightly inflated and corrected for the point-source assumption at larger magnitudes, combined with an estimate of the inter-event variability that is safely conservative. To assign a weight greater than 0.5 to this branch would lead to an excessively peaked distribution of predicted accelerations, which would not be a true reflection of the current state of knowledge. Moreover, we believe that there is an appreciable and growing body of evidence to support the hypothesis of smaller stress drop values for induced earthquakes of shallow focal depth, as discussed in Section 5.4.

The question then is how to distribute the remaining 0.5 weight between the two remaining models. In the absence of any basis for believing one of these branches is more likely than the other, each branch would carry a weight of 0.25; the final assignments, as seen in Table 6.7, are not very far from this outcome, but we have chosen to give the branch #1 a slightly higher weight. This branch is actually based on the assumption of the induced earthquakes being essentially the same as tectonic earthquakes, in terms of medians and sigmas, which is conservative. However, the maintenance of the low stress drop of 10 bars, accompanied by the current low estimate of inter-event variability, corresponding to branch #3 is considered rather optimistic. The important conclusion is that the three very different models, with this rather even distribution of branch weights, define a probability distribution of the predicted motions that is consistent with very considerable epistemic uncertainty for magnitudes above **M** 4. This is entirely appropriate for the Version 1 GMPE and it is to be hoped that subsequent refinements—and additional data—will lead to some reduction of the overall uncertainty (see Chapter 8).

In order to obtain additional insight into the behaviour of the three median models, especially with respect to motions from tectonic earthquakes, the median predictions are compared with the subset of the European database selected in Section 4.3. The recorded accelerations were adjusted to 3 km focal depth and to V_{S30} 200 m/s, and in the case of

records from reverse-faulting earthquakes, to the average of normal and strike-slip events. For simple graphical inspection, the records were then all adjusted, using Eq.(6.1) and the central model whose coefficients are in Table 6.1, to a common epicentral distance of 10 km. The mean acceleration for each earthquake is then plotted together with an indication of the range of motions, against magnitude, and compared with the median predictions from the three models. These comparisons are shown in Figures 6.40 to 6.44 for the five ground-motion parameters in consideration.

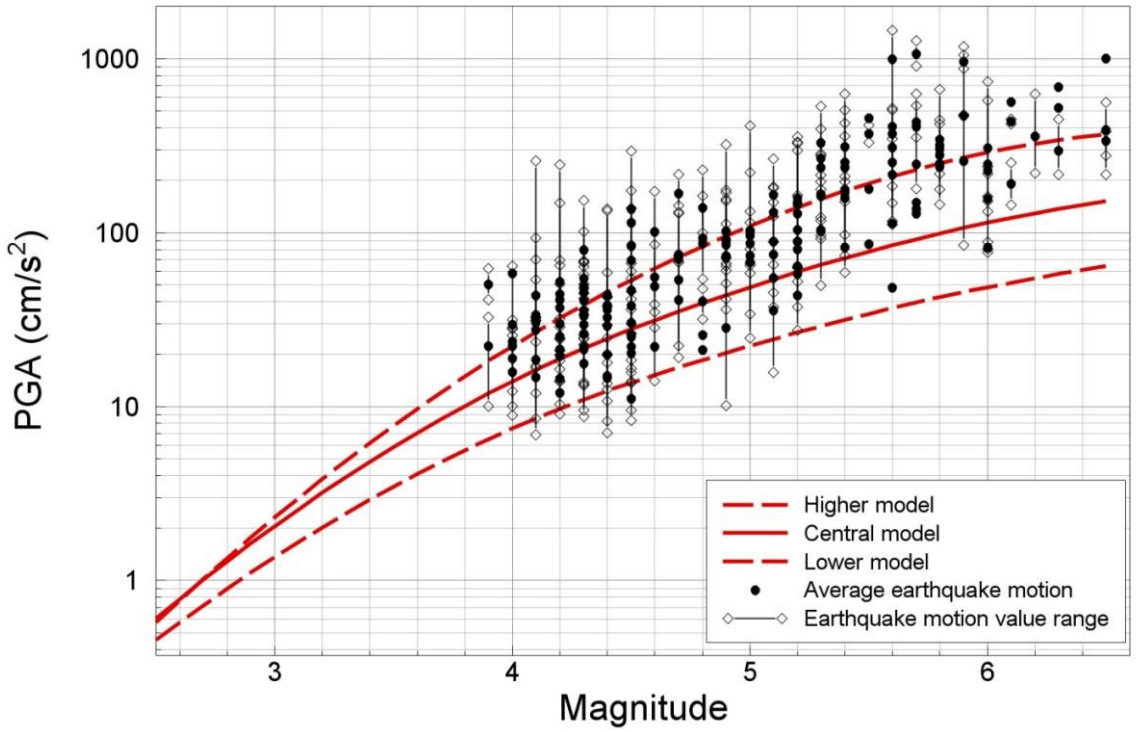


Figure 6.40. PGA values from European strong-motion data adjusted to 3 km focal depth, non-reverse faulting, V_{S30} 200 m/s and R_{epi} 10 km, compared with median predictions from three GMPEs. Black dots are average accelerations from each earthquake, after adjustments.

These comparison plots suggest that the upper model is actually a very good fit to the European tectonic data for PGA and Sa(0.2s). At longer periods, the upper model predicts values that are high with regards to the European data, which might lead one to consider assigning different weights to the logic-tree branches at different response periods. However, there are several reasons to avoid such action, starting with the fact that it is not actually consistent with the way the models have been derived. There is no compelling reason for there to be excellent agreement between the higher branch of the logic-tree and the European strong-motion data; the differences at longer periods may be related to the site effects in the field and how they are modelled, as discussed in Section 6.4. Another reason not to vary the weights from period-to-period is that it undermines the assertion of the central branch being the best estimate and could also lead to unusual spectral shapes if the logic-tree branches were fully implemented in a probabilistic framework.

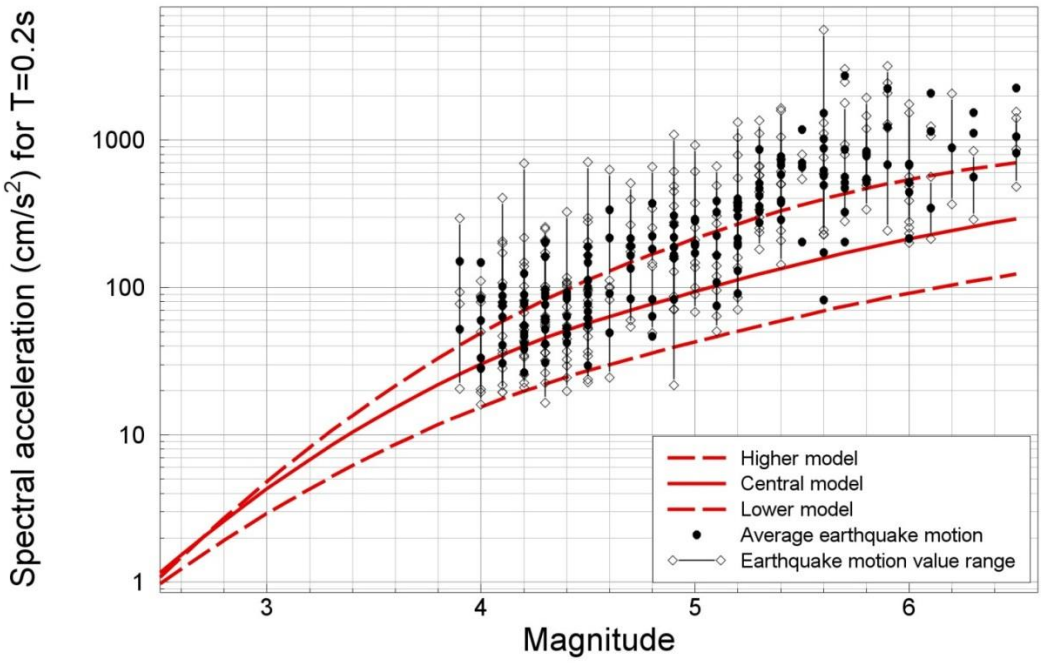


Figure 6.41. Sa(0.2s) values from European strong-motion data adjusted to 3 km focal depth, non-reverse faulting, V_{S30} 200 m/s and R_{epi} 10 km, compared with median predictions from three GMPEs. Black dots are average accelerations from each earthquake, after adjustments.

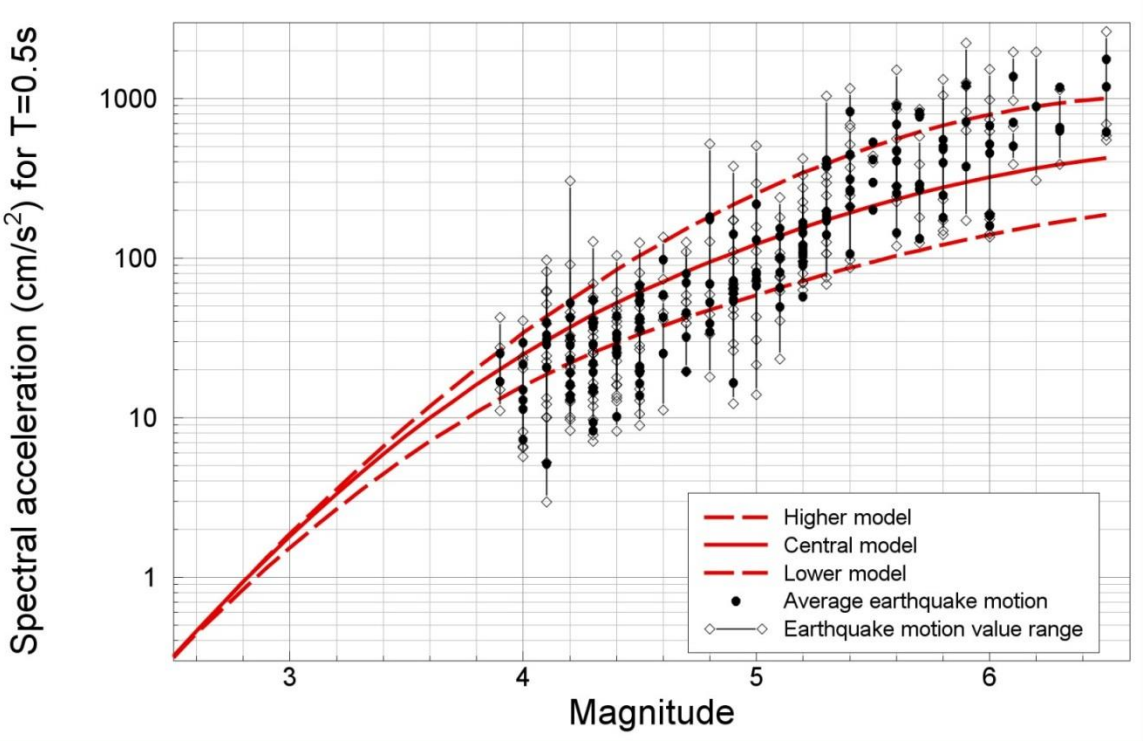


Figure 6.42. Sa(0.5s) values from European strong-motion data adjusted to 3 km focal depth, non-reverse faulting, V_{S30} 200 m/s and R_{epi} 10 km, compared with median predictions from three GMPEs. Black dots are average accelerations from each earthquake, after adjustments.

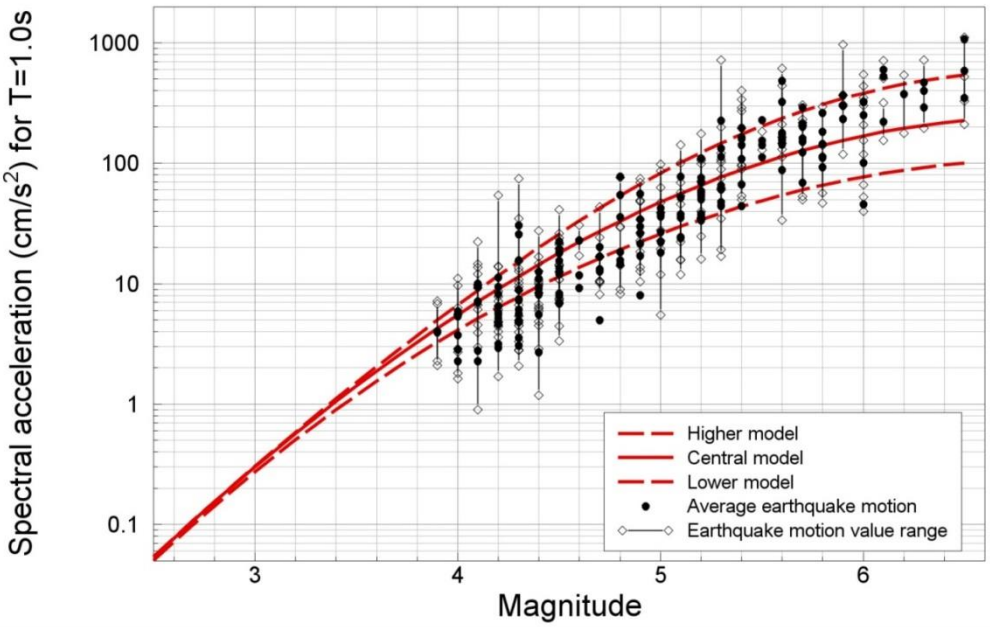


Figure 6.43. Sa(1.0s) values from European strong-motion data adjusted to 3 km focal depth, non-reverse faulting, V_{S30} 200 m/s and R_{epi} 10 km, compared with median predictions from three GMPEs. Black dots are average accelerations from each earthquake, after adjustments.

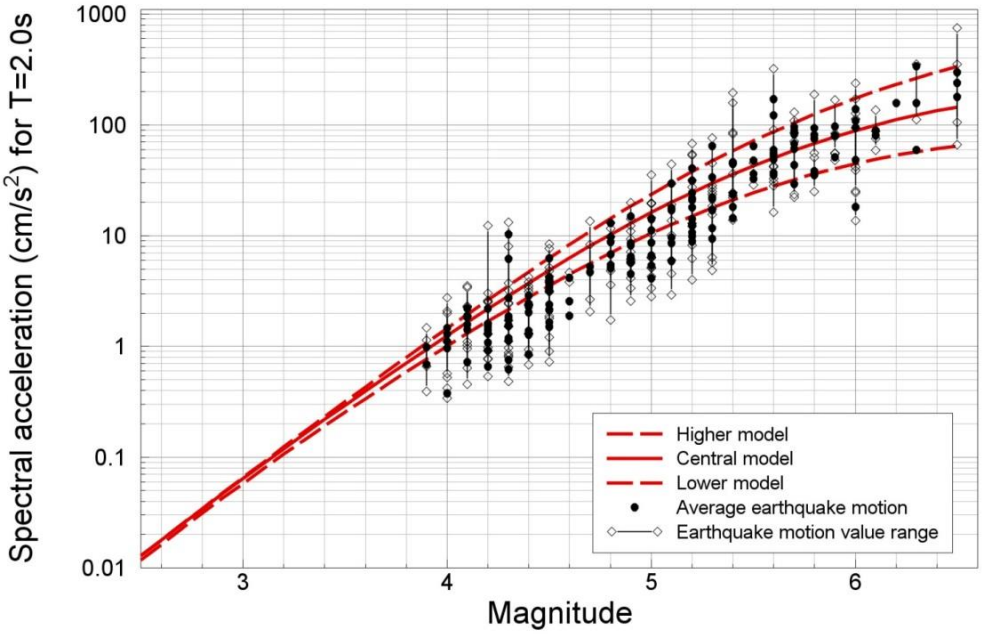


Figure 6.44. Sa(2.0s) values from European strong-motion data adjusted to 3 km focal depth, non-reverse faulting, V_{S30} 200 m/s and R_{epi} 10 km, compared with median predictions from three GMPEs. Black dots are average accelerations from each earthquake, after adjustments.

More detailed insight can be obtained from presenting plots of the total residuals plotted against magnitude and distance from each of the three models individually. Figures 6.45 to 6.49 show the total residuals of the European data calculated with respect to the three median prediction models, and plotted against both magnitude and distance. The plots confirm the same patterns that were observed in the figures discussed above.

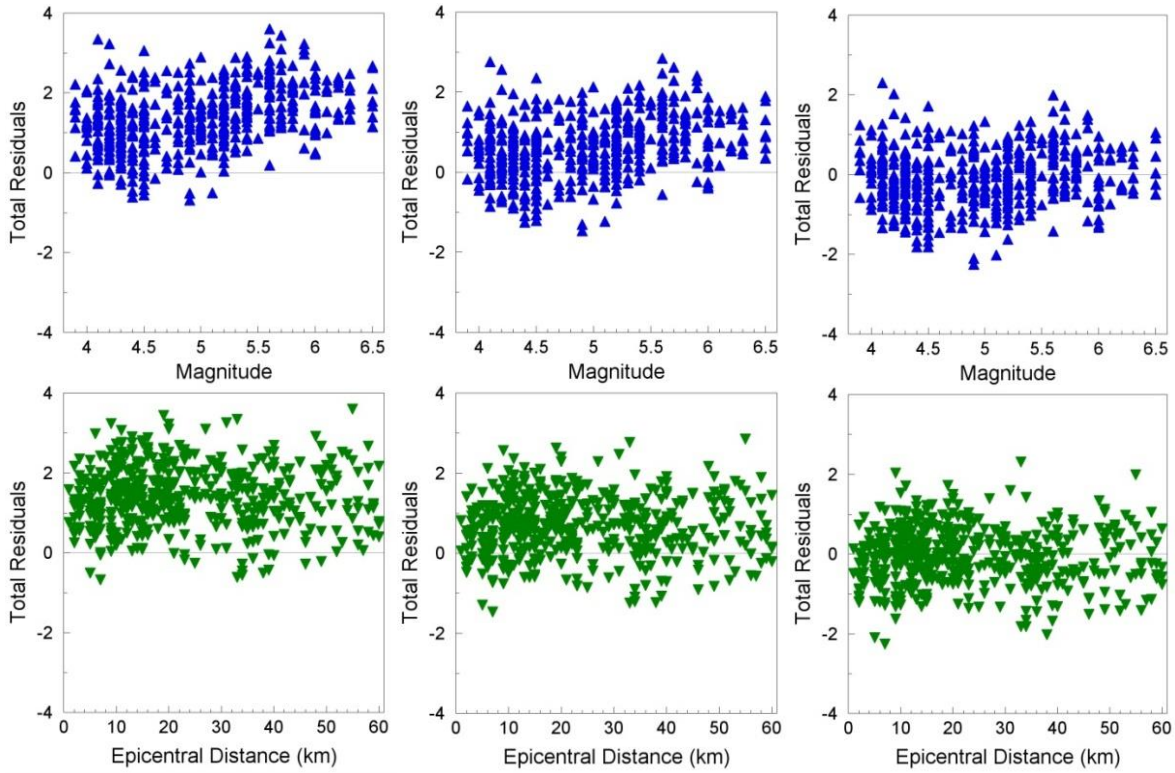


Figure 6.45. Total residuals of PGA from the transformed European dataset with respect to the lower (*left*), central (*middle*) and higher (*right*) models

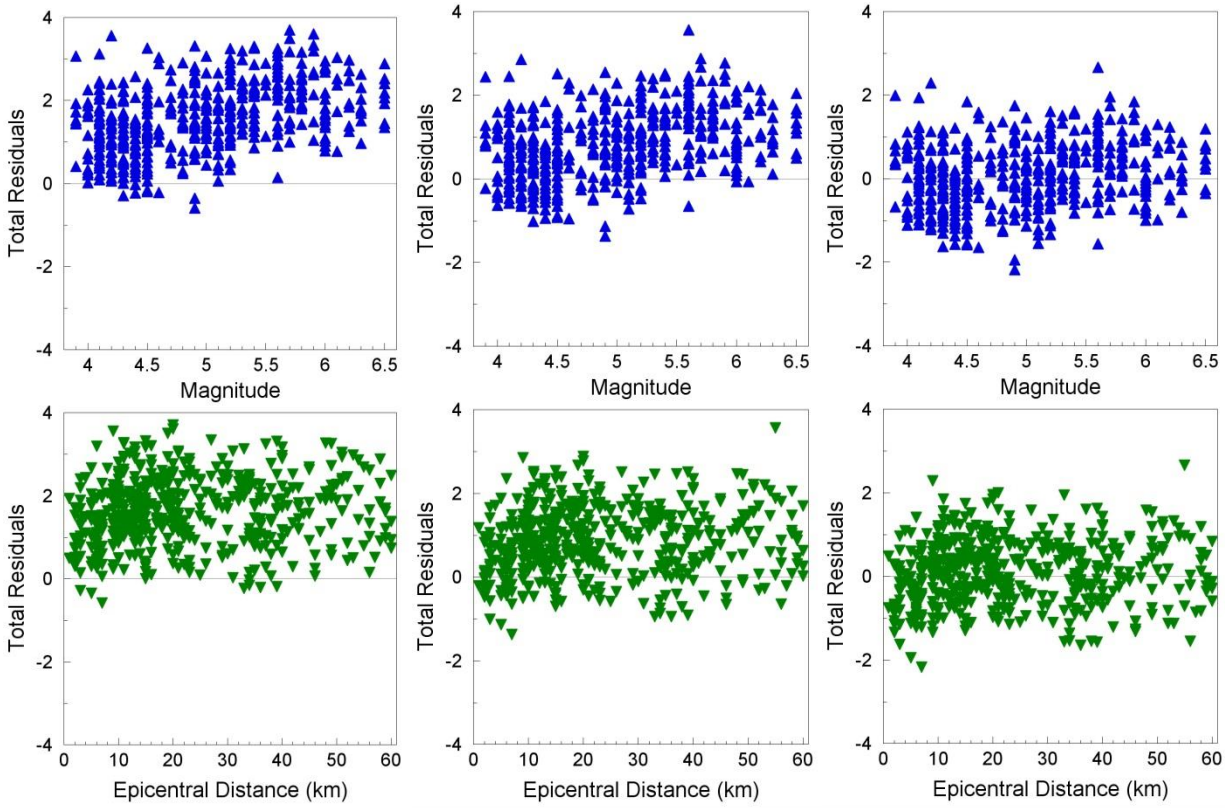


Figure 6.46. Total residuals of Sa(0.2s) from the transformed European dataset with respect to the lower (*left*), central (*middle*) and higher (*right*) models

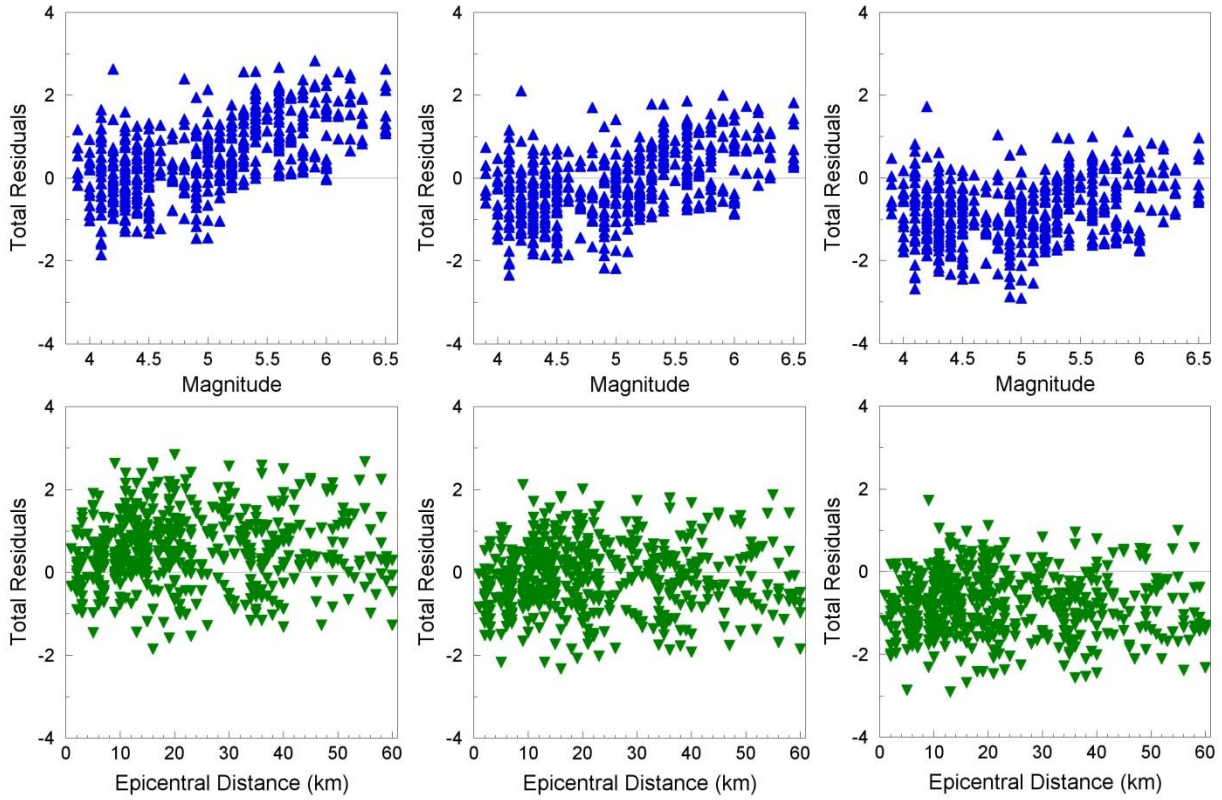


Figure 6.47. Total residuals of Sa(0.5s) from the transformed European dataset with respect to the lower (*left*), central (*middle*) and higher (*right*) models

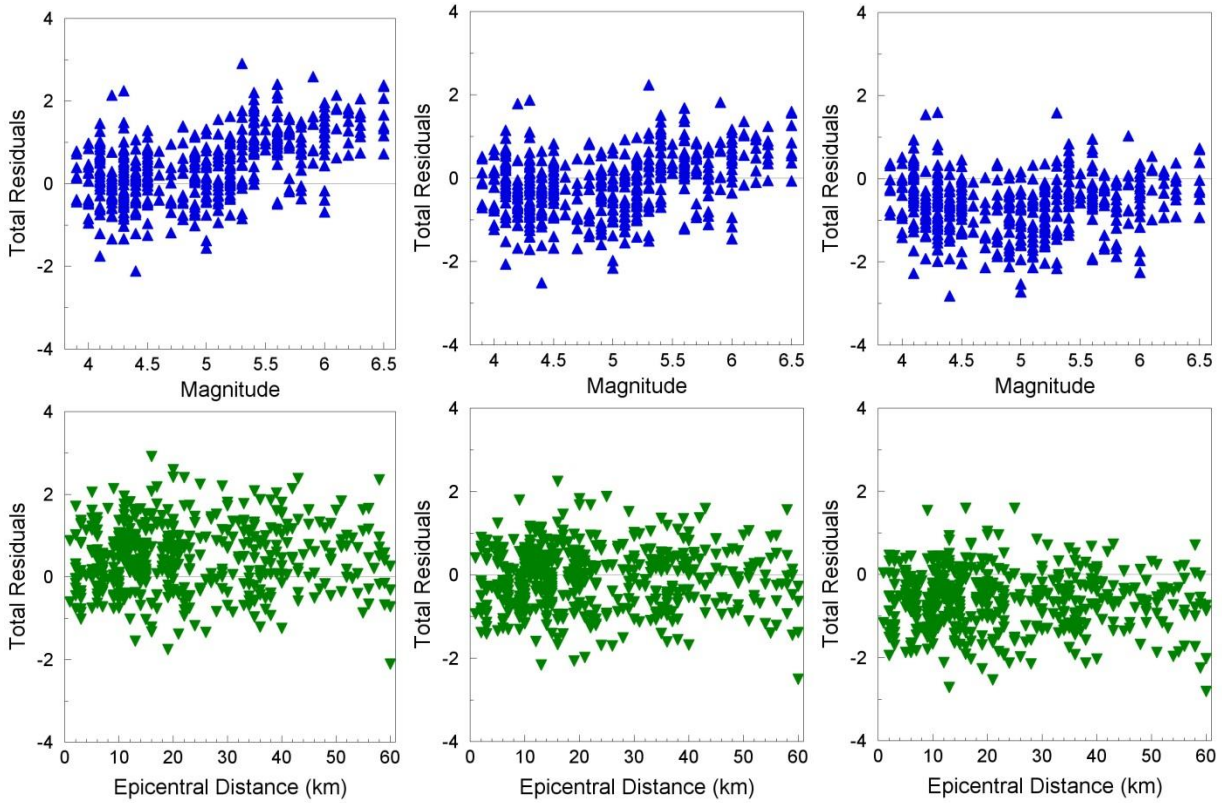


Figure 6.48. Total residuals of Sa(1.0s) from the transformed European dataset with respect to the lower (*left*), central (*middle*) and higher (*right*) models

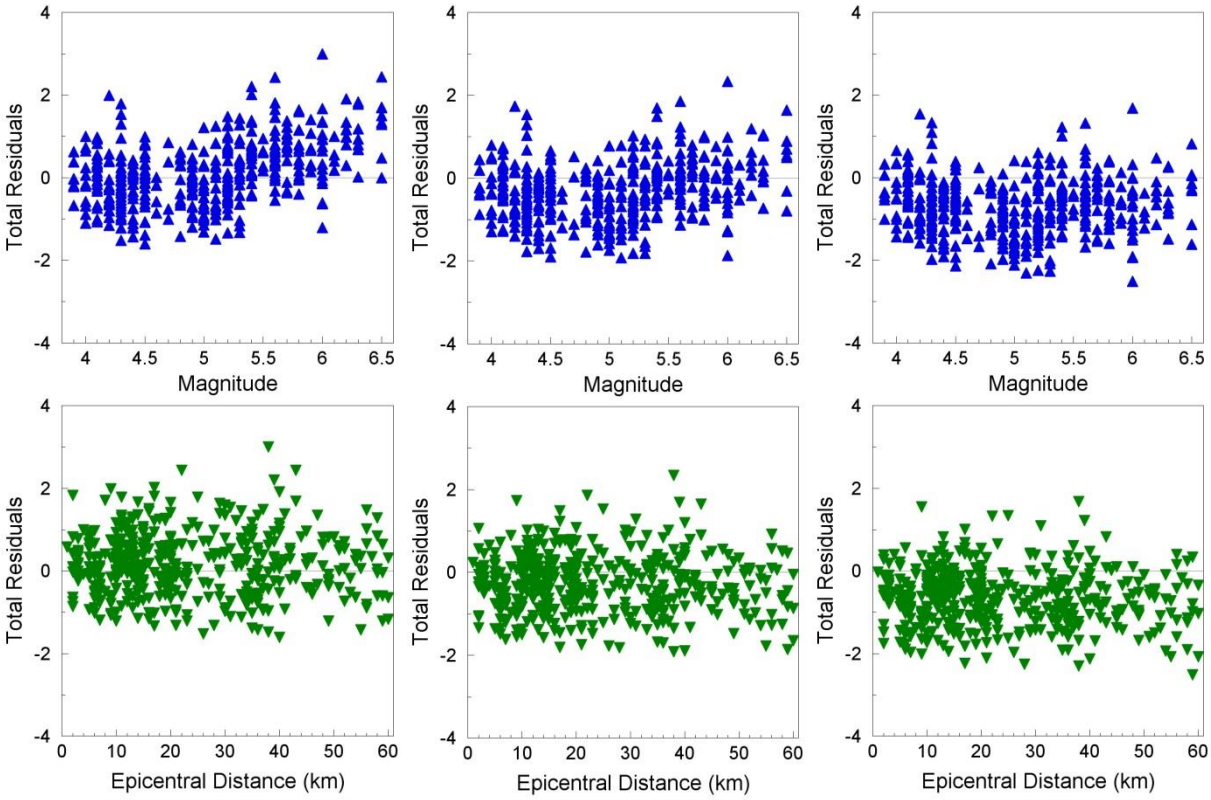


Figure 6.49. Total residuals of Sa(2.0s) from the transformed European dataset with respect to the lower (*left*), central (*middle*) and higher (*right*) models

6.4. Comparison of Version 0 and Version 1 PGA predictions

Since questions regarding the changes from the GMPE used in the 2013 Winningsplan are inevitable, this section compares the predictions of geometric mean horizontal PGA from the Version 0 and Version 1 GMPEs. In the first set of plots (Figures 6.50-5.56) the median and 97.7-percentile (two standard deviation) predictions are compared for a specified magnitude, and plotted against epicentral distance, using the central V1 equation. The observed patterns for the medians are that the central V1 equation predicts values that are generally lower than those from the V0 equation, with the difference between the two equations increasing with earthquake magnitude. The exception to this is at small magnitudes, since in the V1 equation the magnitude dependence of the near-source saturation term is modelled more realistically, leading to higher values close to the epicentre. The V1 model shows less rapid attenuation with distance at longer distances. The differences are unsurprising given that the V0 model was a single equation based on intentionally conservative assumptions, whereas the central V1 model is the best estimate within an appreciable range of epistemic uncertainty.

The ratio of 97.7-percentile to 50-percentile predictions is constant for the V0 model, but increases with increasing magnitude for the V1 model as a result of the point-source correction to the intra-event variability.

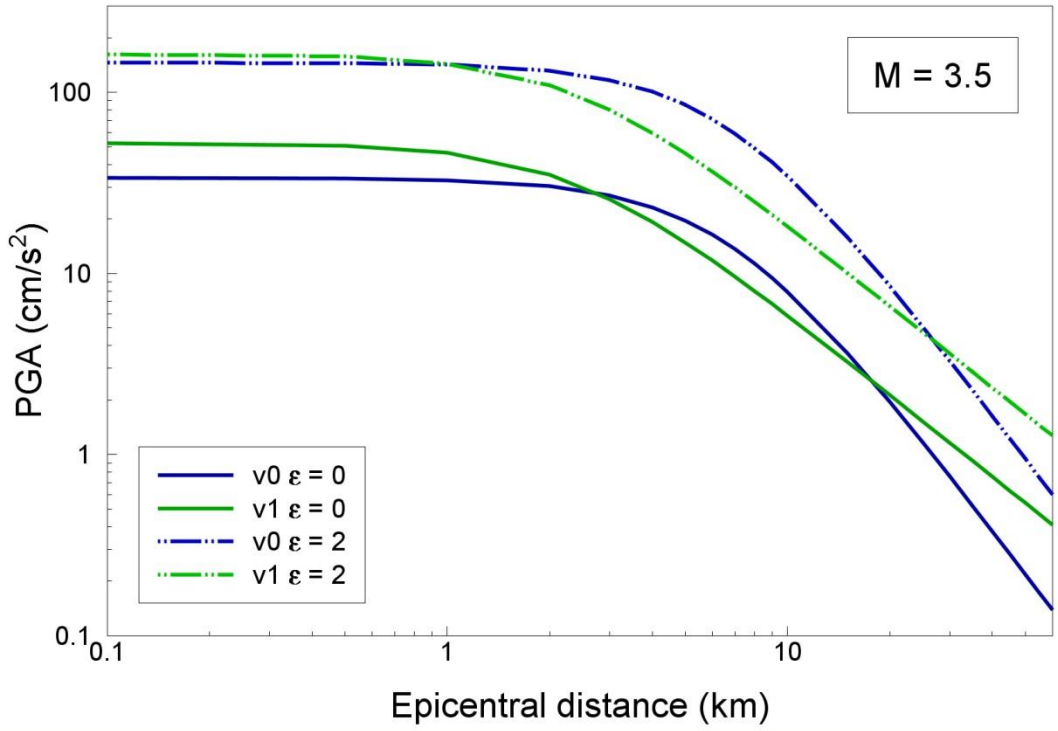


Figure 6.50. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and central Version 1 GMPEs for an earthquake of **M** 3.5

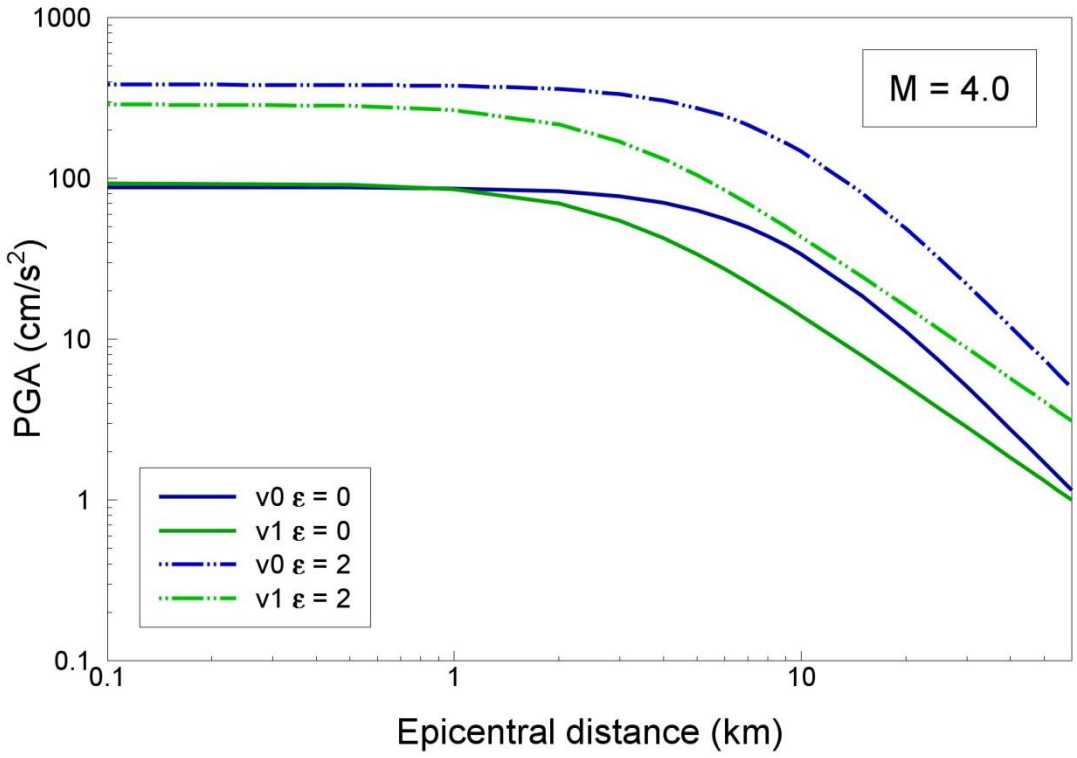


Figure 6.51. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and central Version 1 GMPEs for an earthquake of **M** 4.0

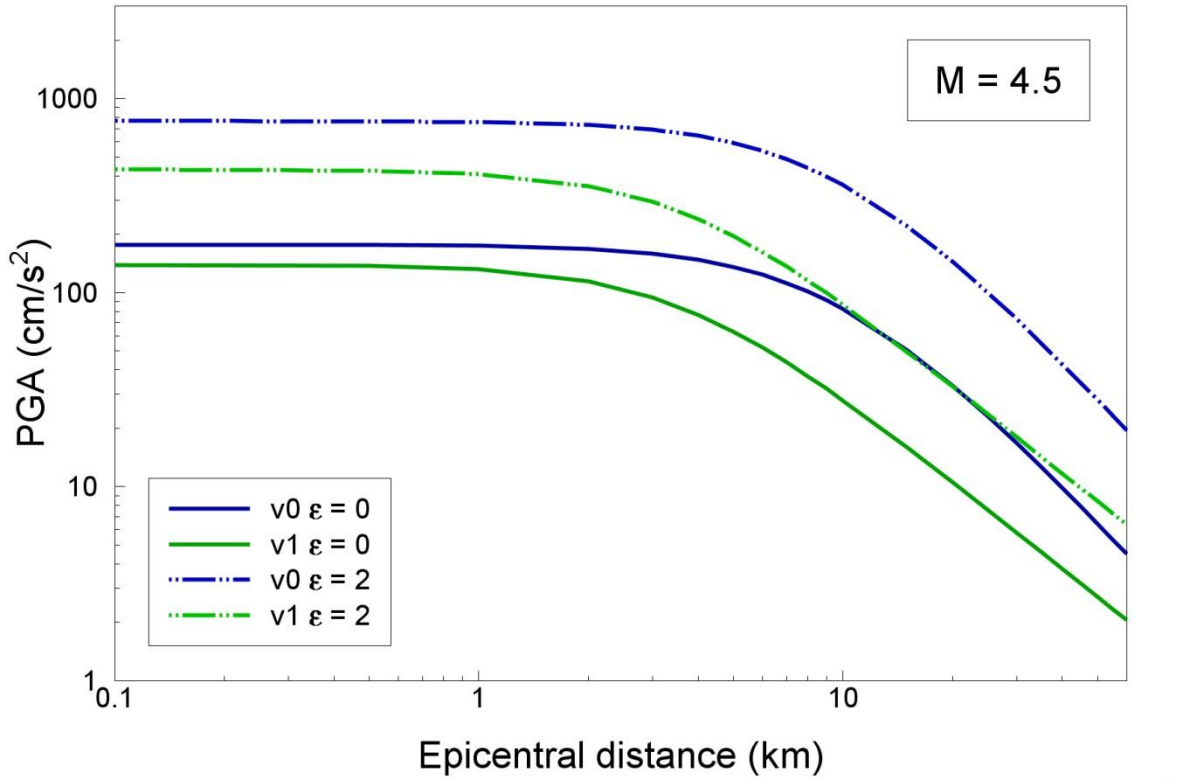


Figure 6.52. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and central Version 1 GMPEs for an earthquake of **M** 4.5

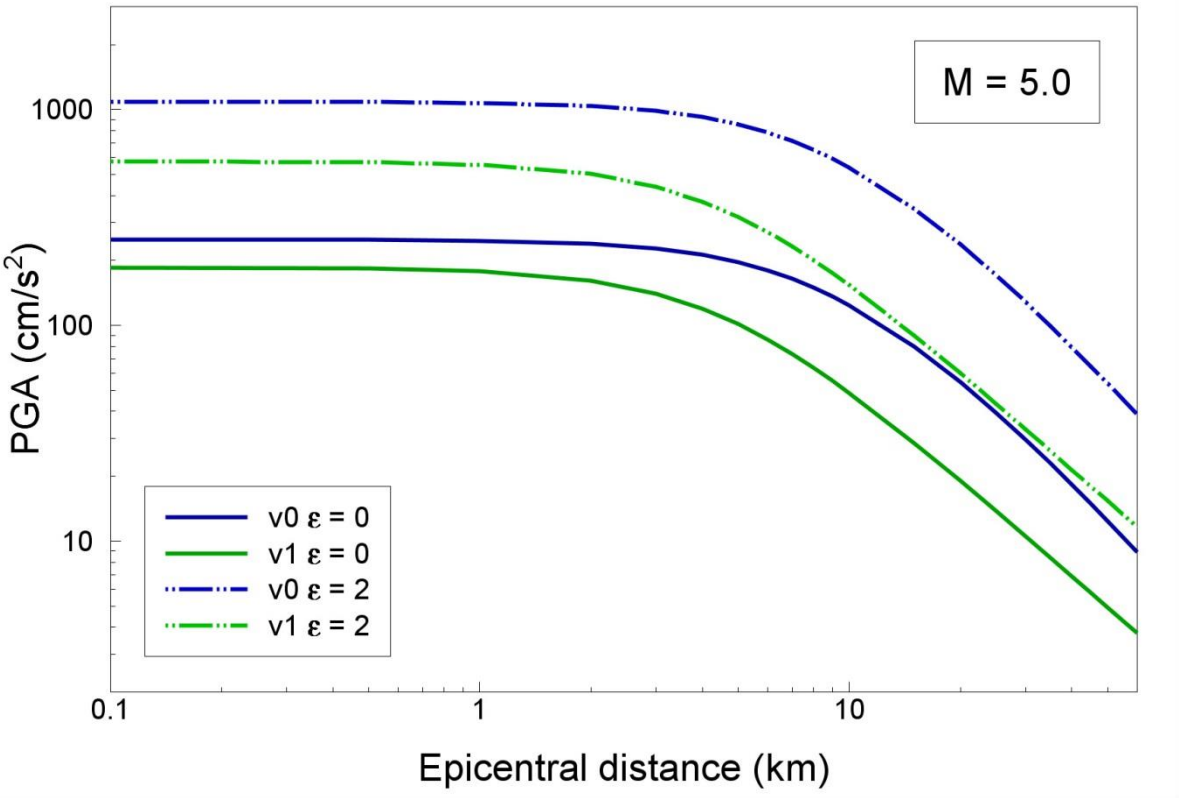


Figure 6.53. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and central Version 1 GMPEs for an earthquake of **M** 5.0

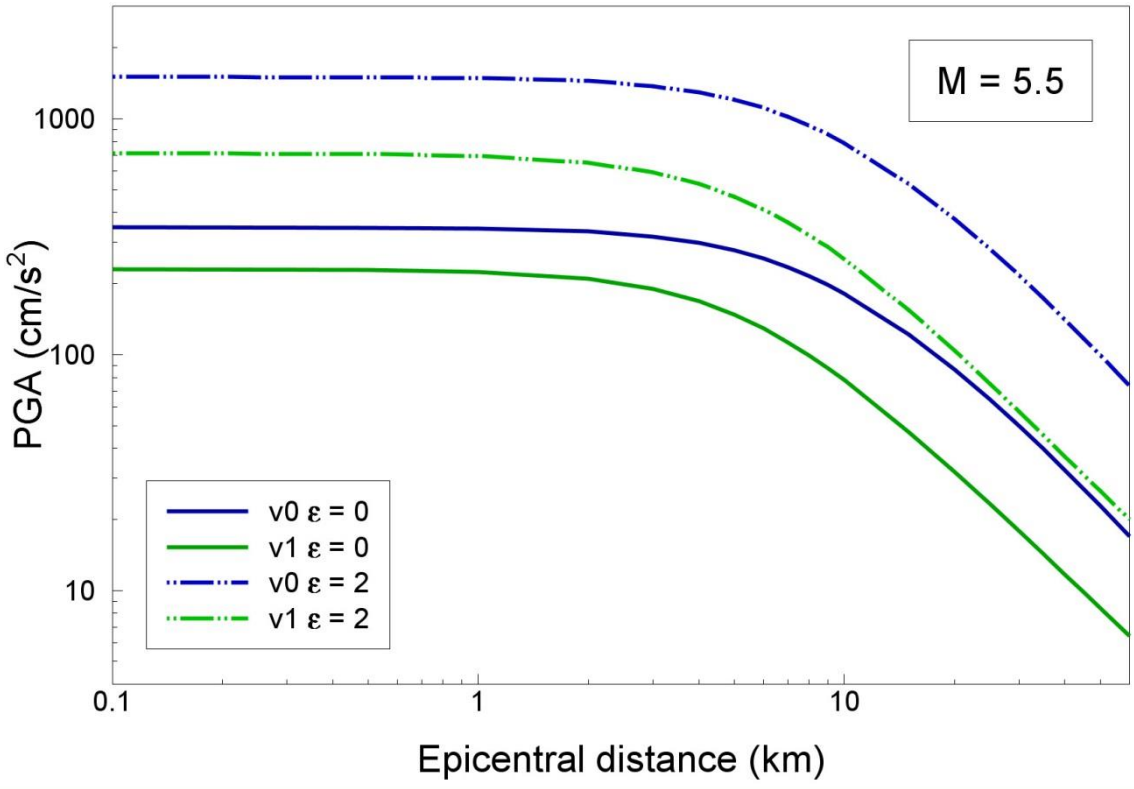


Figure 6.54. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and central Version 1 GMPEs for an earthquake of **M** 5.5

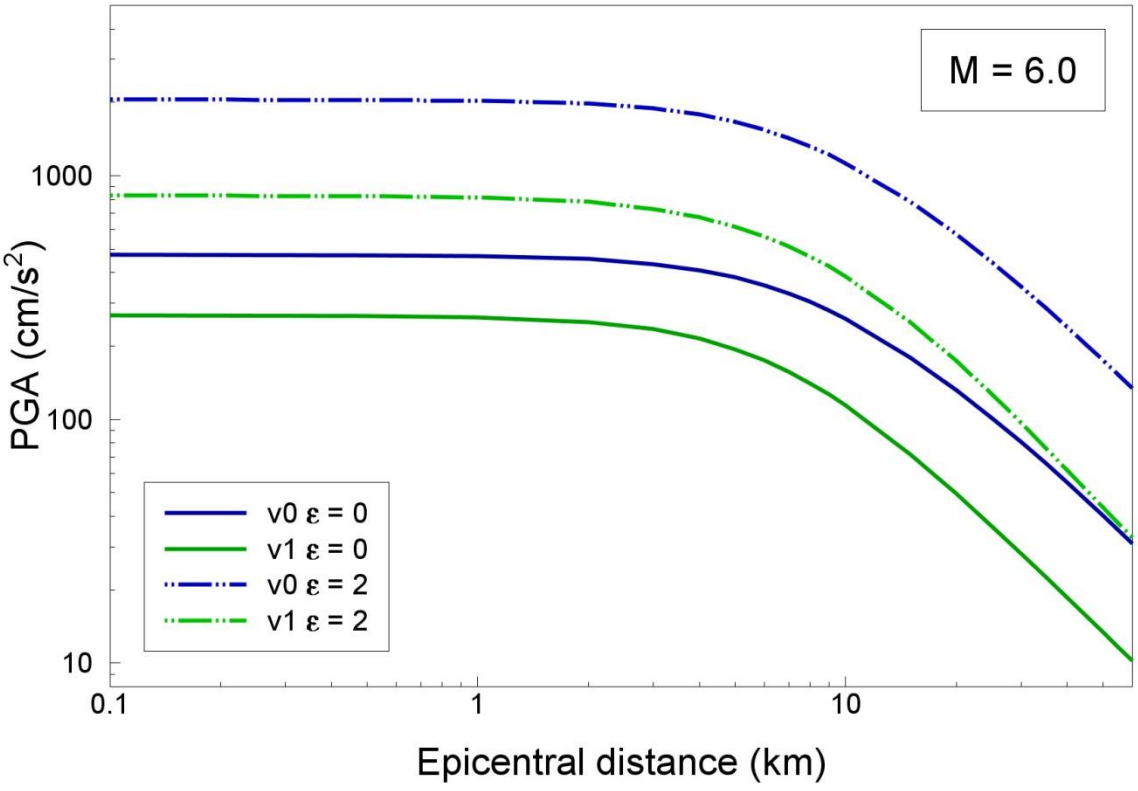


Figure 6.55. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and central Version 1 GMPEs for an earthquake of **M** 6.0

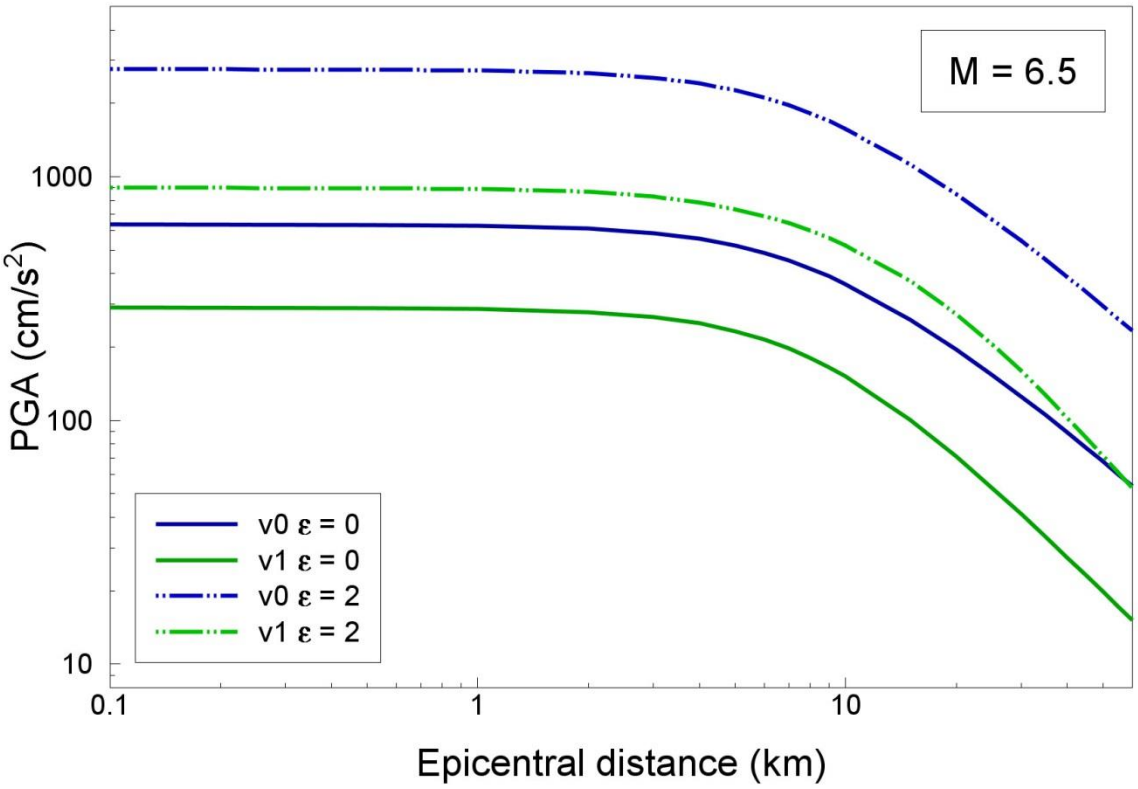


Figure 6.56. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and central Version 1 GMPEs for an earthquake of **M** 6.5

The second set of plots (Figures 6.57-6.63) reproduce those just shown, but now also include the median and 97.7-percentile predictions from both the higher and lower alternative GMPEs that are part of the Version 1 model, include to capture the unavoidable epistemic uncertainty associated with these predictions.

In these rather busy plots, the first thing to focus on is the comparison between the V0 model and the upper V1 model, which are represented by the solid blue and red lines respectively. It can be appreciated that at short distances, the higher V1 model is generally higher than the median V0 model, with the difference between the two models decreasing as the magnitude increases; the models are in rather close agreement for **M** 6.5. This is consistent with the higher V1 model attempting to be consistent with GMPEs for tectonic models, whereas the V0 model intended to replicate the Groningen motions at small magnitudes. At longer distances, the higher V1 model is generally higher. For the 97.7-percentiles, the patterns are similar except for the influence of the somewhat smaller sigma values in the V1 model.

The lower V1 model is generally below the other models, as would be expected, but interestingly for **M** 3.5 and very short epicentral distances, the lower V1 model is closer to the V0 equation than the central V1 model.

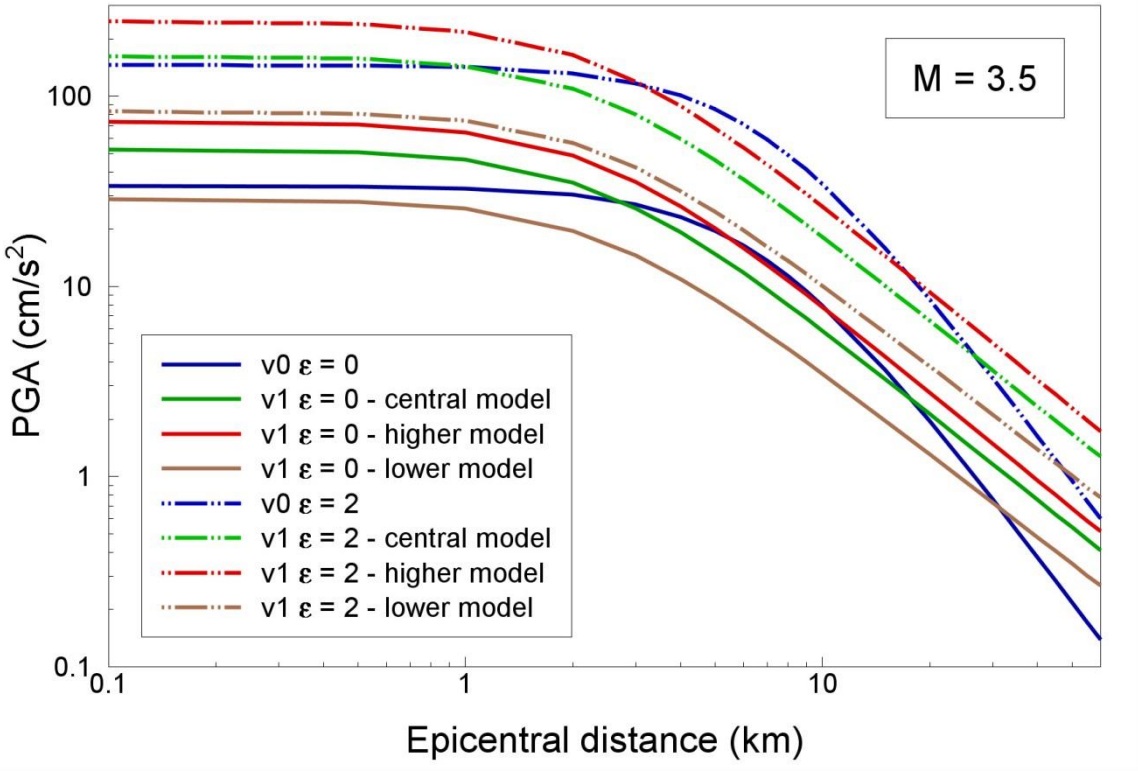


Figure 6.57. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and central, upper and lower Version 1 GMPEs for an earthquake of **M** 3.5

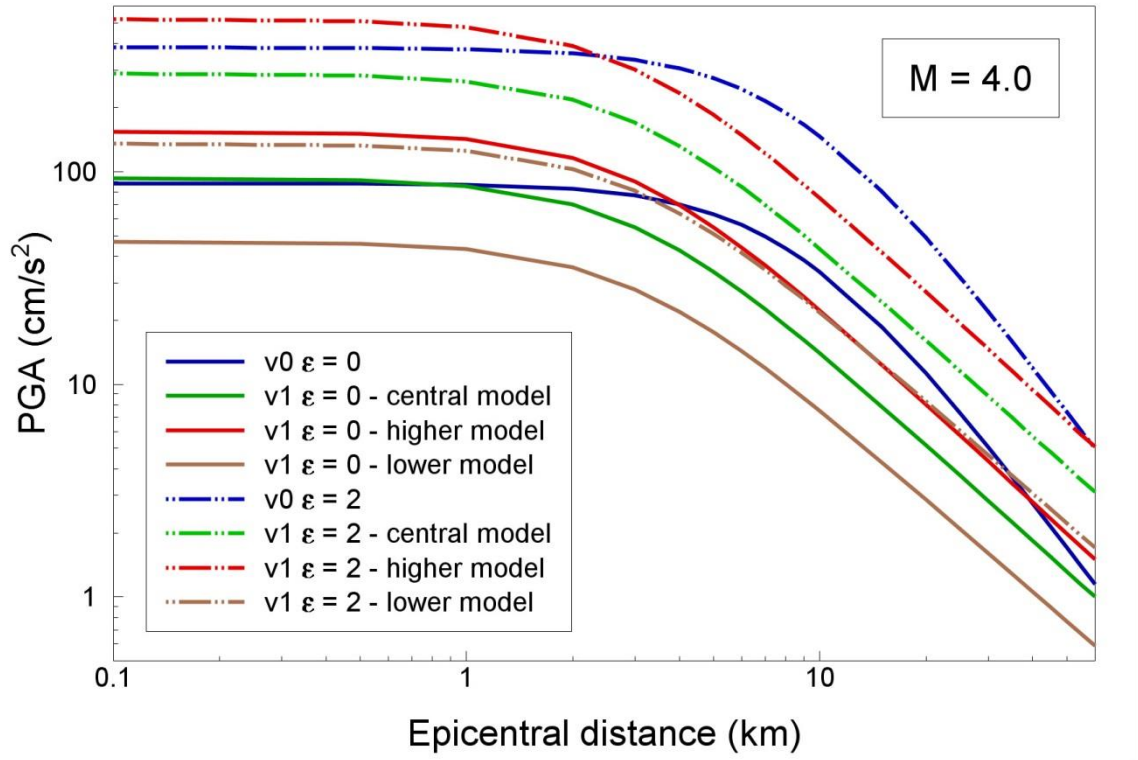


Figure 6.58. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and central, upper and lower Version 1 GMPEs for an earthquake of **M** 4.0

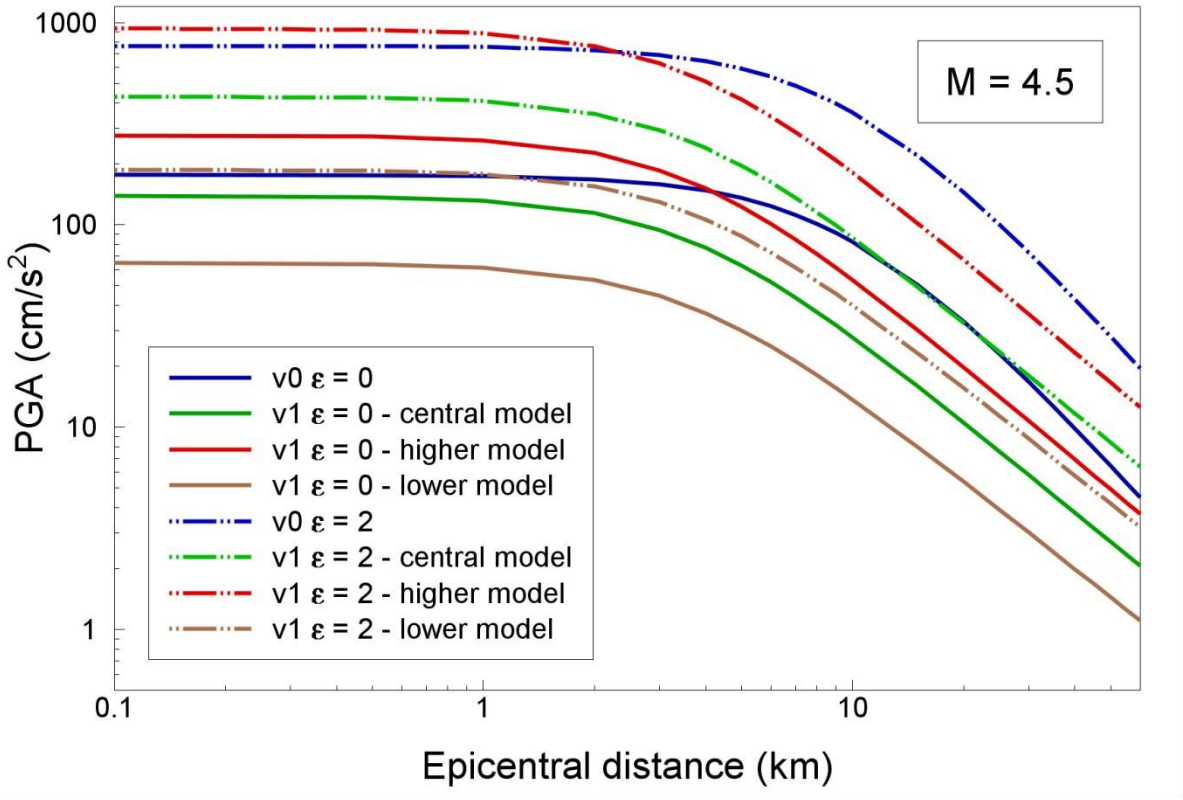


Figure 6.59. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and central, upper and lower Version 1 GMPEs for an earthquake of **M** 4.5

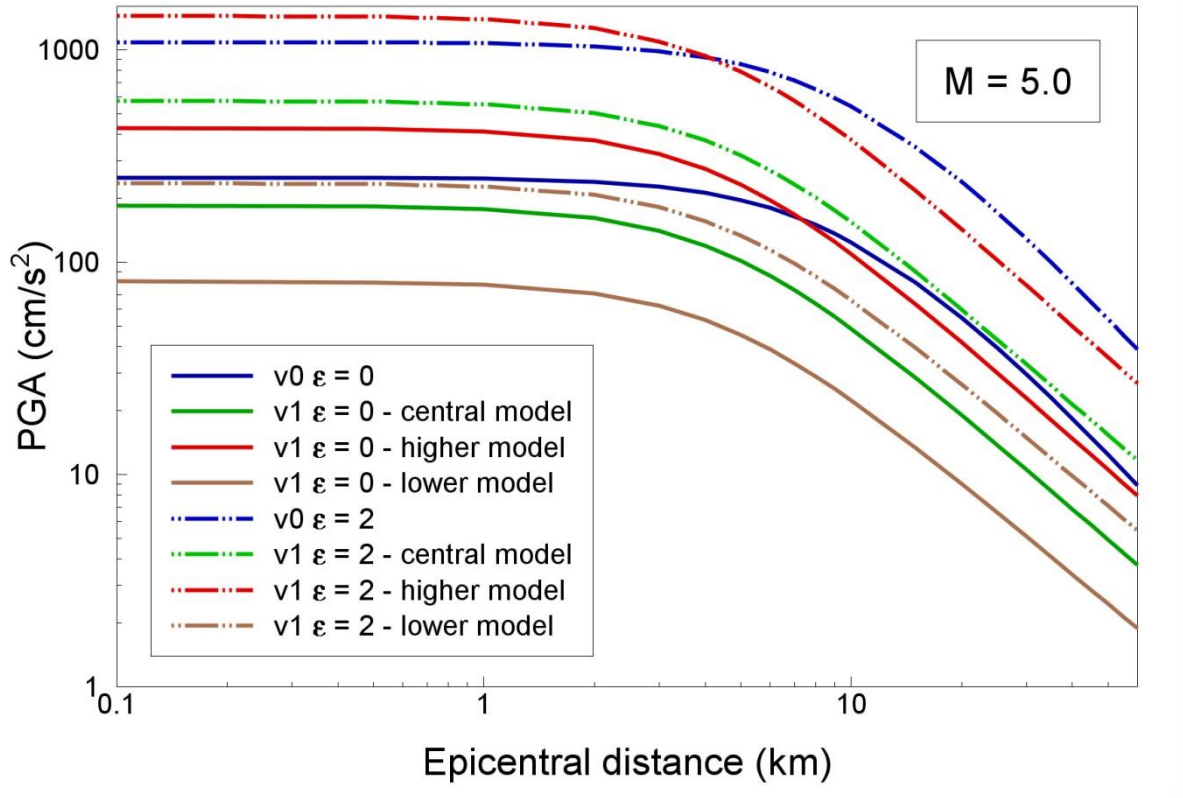


Figure 6.60. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and central, upper and lower Version 1 GMPEs for an earthquake of **M** 5.0

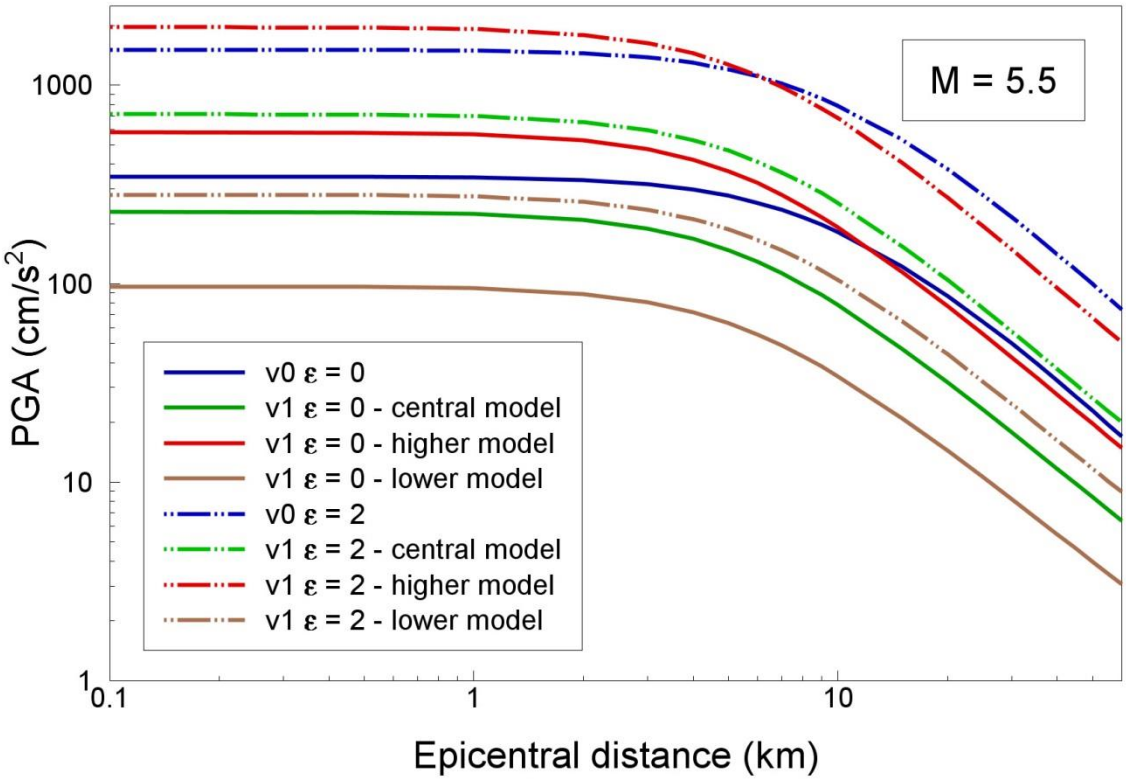


Figure 6.61. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and central, upper and lower Version 1 GMPEs for an earthquake of **M** 5.5

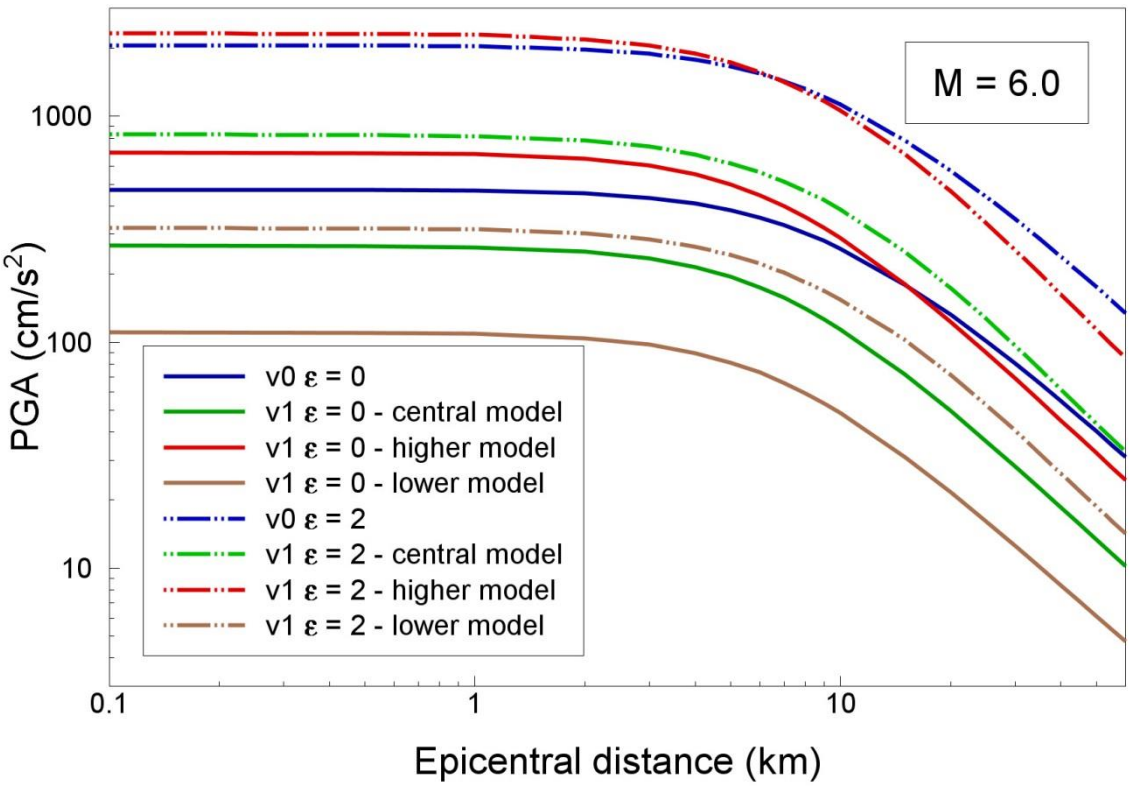


Figure 6.62. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and central, upper and lower Version 1 GMPEs for an earthquake of **M** 6.0

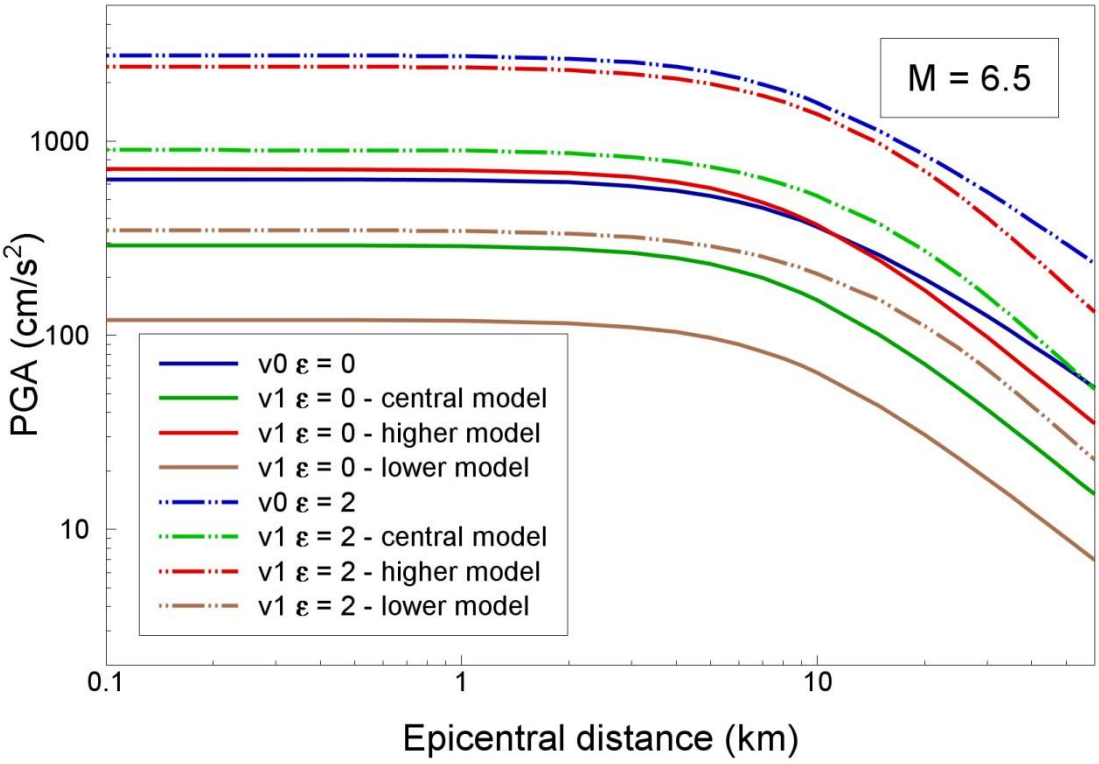


Figure 6.63. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and central, upper and lower Version 1 GMPEs for an earthquake of **M** 6.5

The final set of plots (Figure 6.64-6.70) compares the median and 97.7-percentile predictions from the V0 GMPE with the mean predictions from the V1 equations. The mean values at both exceedance levels are calculated using the logic-tree weights that have been assigned to the three models:

- Lower w = 0.2
- Central w = 0.5
- Higher w = 0.3

The weighted mean PGA for each **M**-R_{epi}-ε combination is an indication of the overall effect of the models, even if this is not a direct indication of the influence the logic-tree would have on the calculation of the mean hazard or risk (for which the weighted average of the resulting probabilities of exceedance are calculated). The observed patterns are rather similar to those seen in the first comparisons between the V0 and central V1 models, but with the weighted mean V1 predictions yielding values that are higher than the central model. Interestingly, given the disaggregation results from the Version 0 hazard and risk calculation that showed dominant contributions from the magnitude range **M** 4-5, the two sets of median predictions at short distances are very similar in this magnitude range. As would be expected, however, the mean V1 model decays more rapidly with distance (because of the magnitude-dependent near-source saturation term) and yields much lower 97.7-percentile values (because of the smaller sigma).

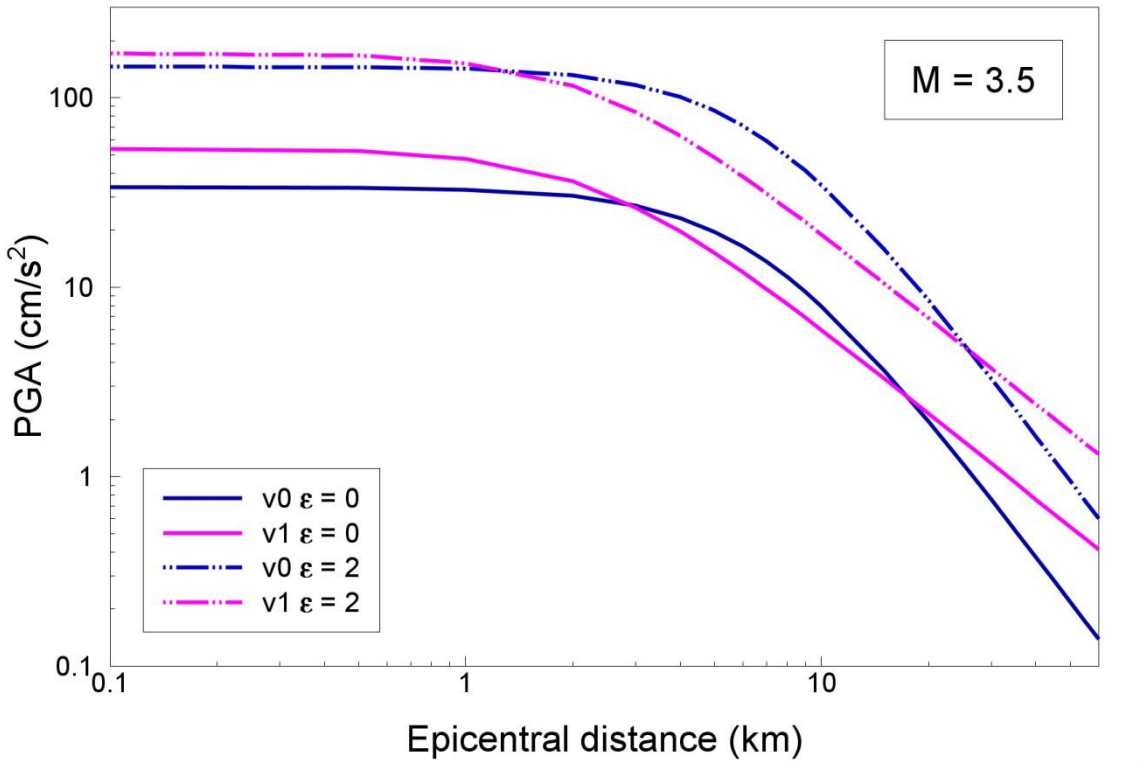


Figure 6.64. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and weighted mean of the Version 1 GMPEs for an earthquake of **M** 3.5

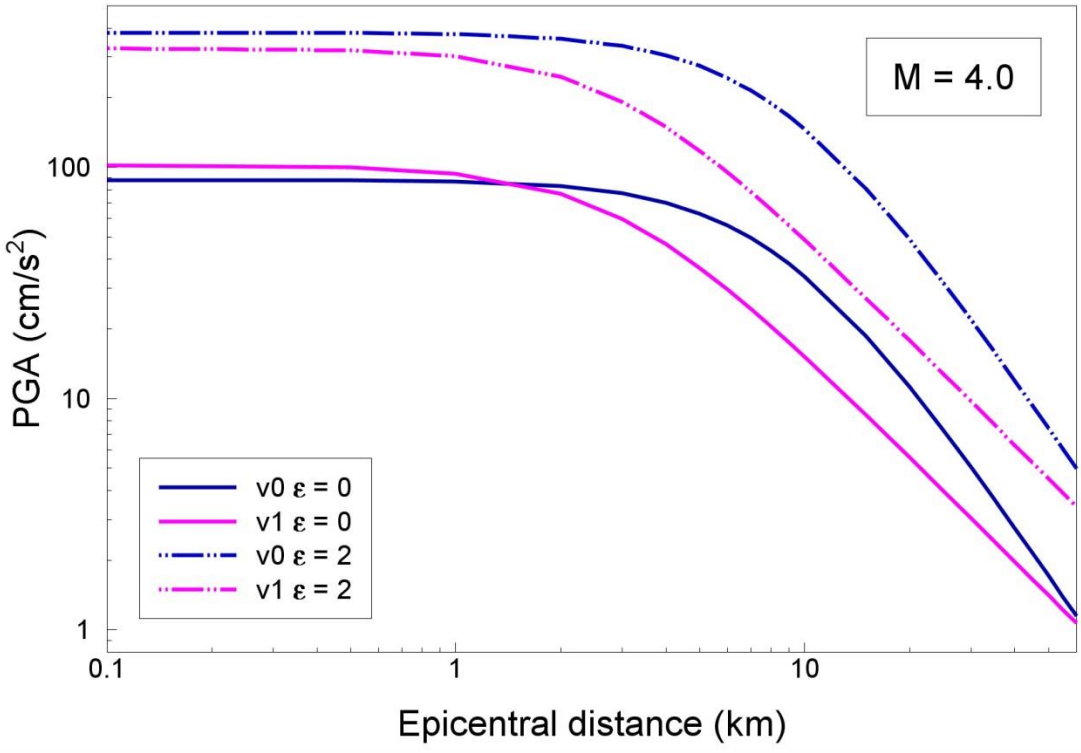


Figure 6.65. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and weighted mean of the Version 1 GMPEs for an earthquake of **M** 4.0

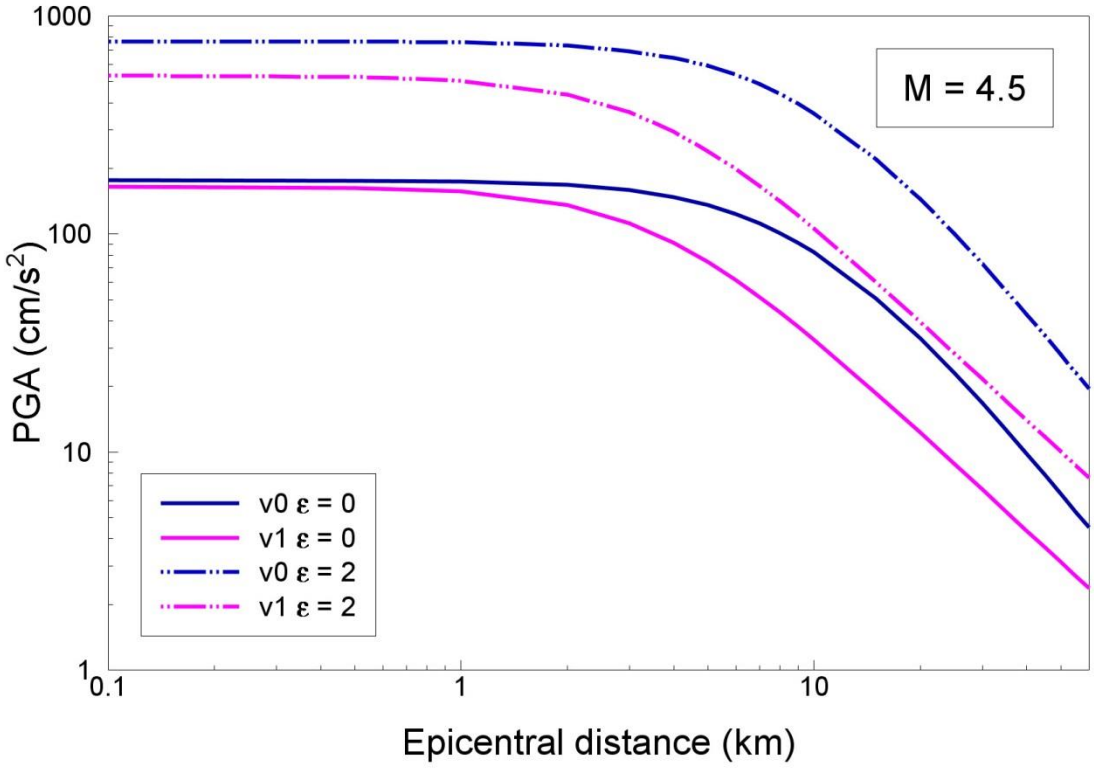


Figure 6.66. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and weighted mean of the Version 1 GMPEs for an earthquake of **M** 4.5

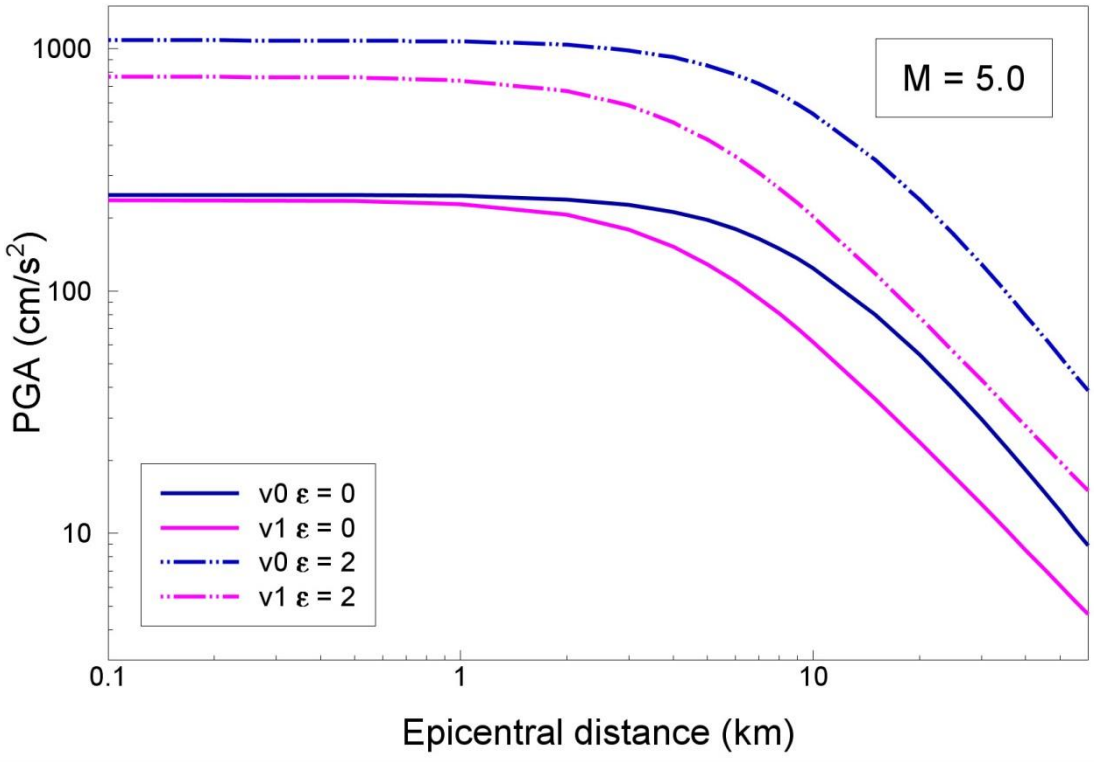


Figure 6.67. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and weighted mean of the Version 1 GMPEs for an earthquake of **M** 5.0

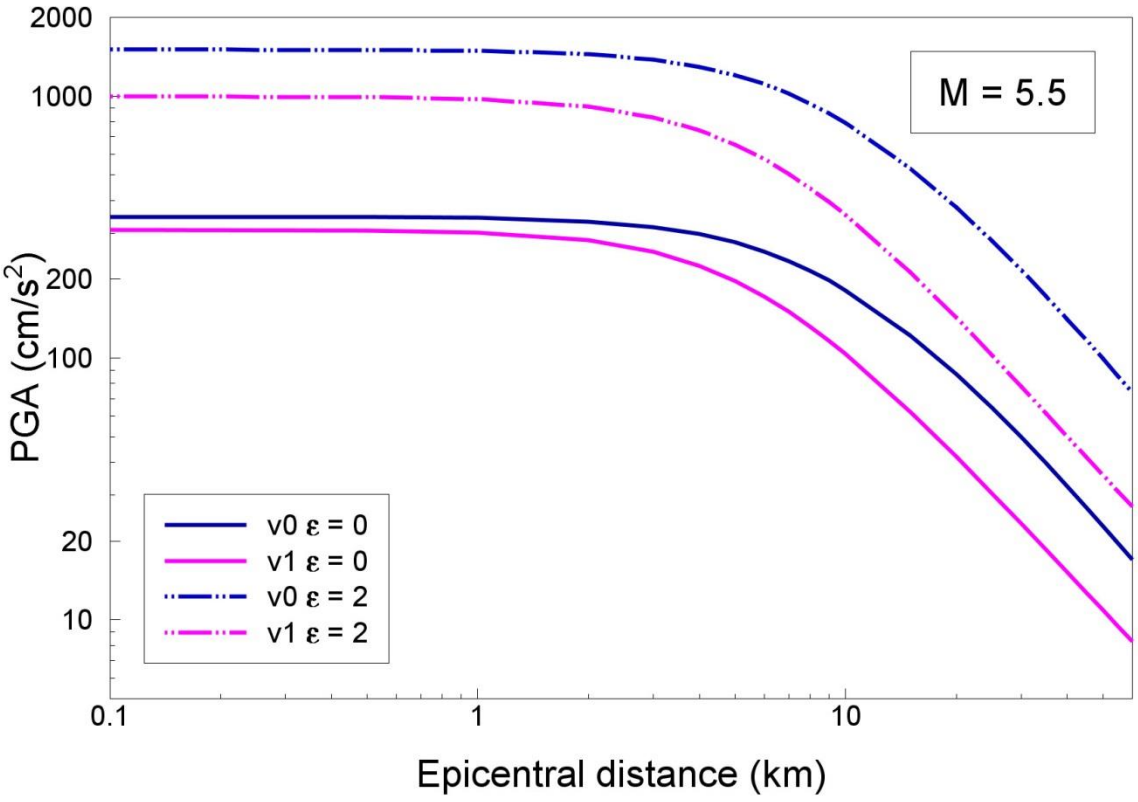


Figure 6.68. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and weighted mean of the Version 1 GMPEs for an earthquake of **M** 5.5

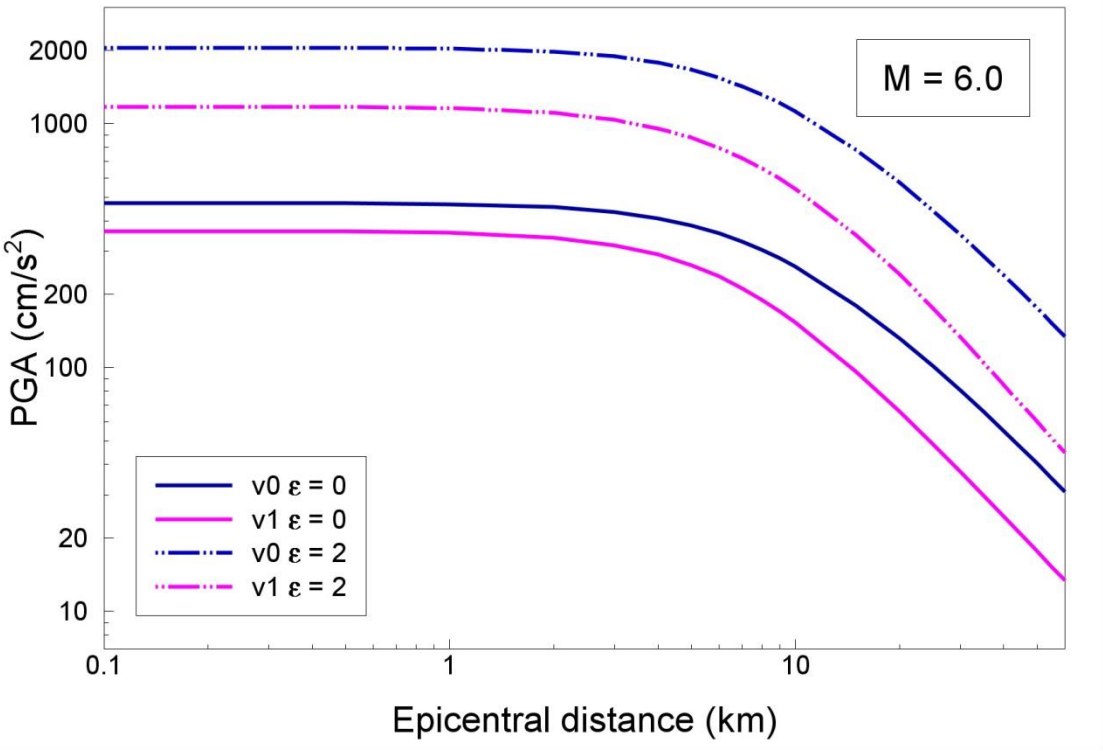


Figure 6.69. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and weighted mean of the Version 1 GMPEs for an earthquake of **M** 6.0

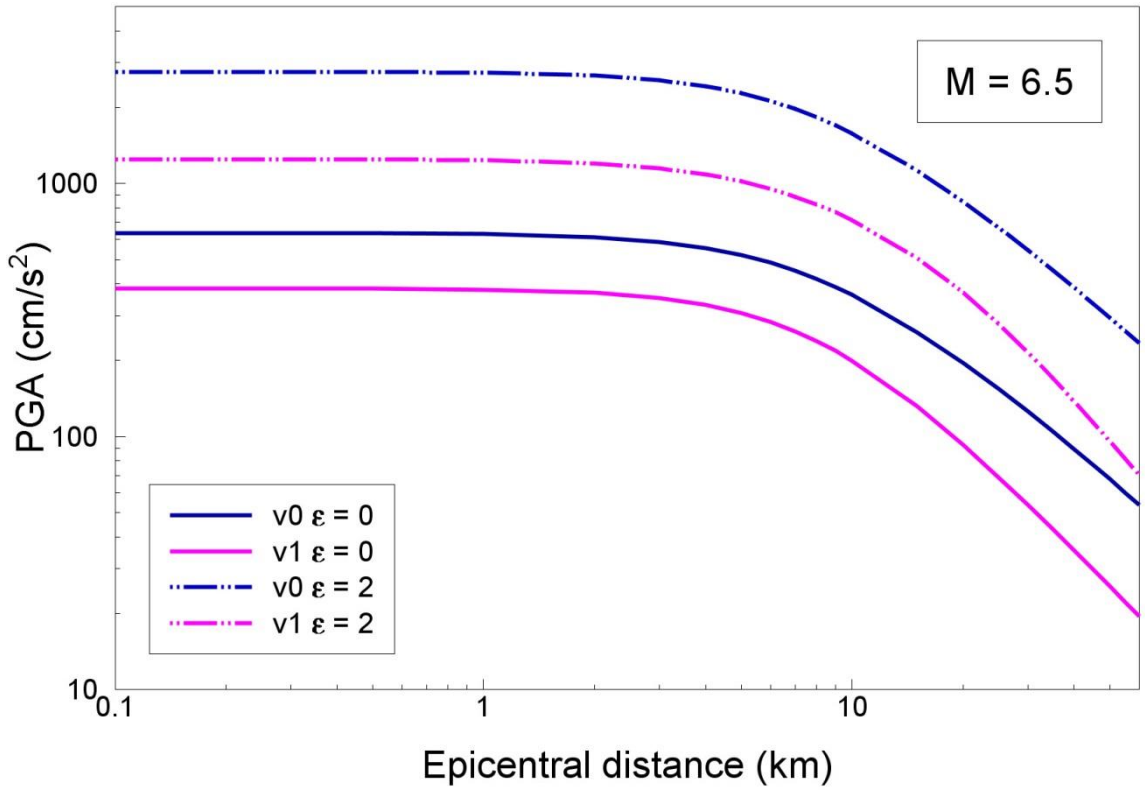


Figure 6.70. Comparison of 50- and 97.7-percentile PGA predictions from the Version 0 and weighted mean of the Version 1 GMPEs for an earthquake of **M** 6.5

6.5. Response spectral shapes

For the purposes of the Version 1 risk model, the fragility functions for all building types are being grouped so that they are defined by the spectral accelerations at one of the 5 selected oscillator periods (0.01, 0.2, 0.5, 1.0 and 2.0 seconds). However, the derivation of the fragility functions requires definition of complete response spectral shapes, and in order to avoid uncertain interpolation across the large intervals between the five periods, additional equations—for medians only—were derived. These equations were obtained by first repeating the stochastic simulations for the central model (*i.e.*, $\Delta\sigma$ = 30 bars) and then performing regressions to fit the functional form of Eq.(6.1)—with the coefficients c_5 and c_6 fixed to the same constant values—to the simulated motions. The resulting coefficients are in Table 6.8.

The resulting spectral forms can be appreciated from Figures 6.71 and 6.72 which show the predicted median spectral ordinates for different combinations of magnitude and distance. Figure 6.71 suggests that the spectral shapes are not very sensitive to distance at longer periods (from the spectral peak upwards), which is consistent with the very modest changes in the coefficient c_4 at periods beyond 0.5 seconds, and the fact that the coefficients c_5 and c_6 are held constant. The spectral shapes are, however, sensitive to magnitude, as confirmed by Figure 6.72, with the period at which the spectral acceleration reaches a peak increasing with increasing magnitude (as is expected). Noteworthy is just

how relatively high are the ordinates at longer periods, even at lower magnitudes, suggesting that the ground conditions are deep layers of soft soils (which we know to be the case from the field data available to date).

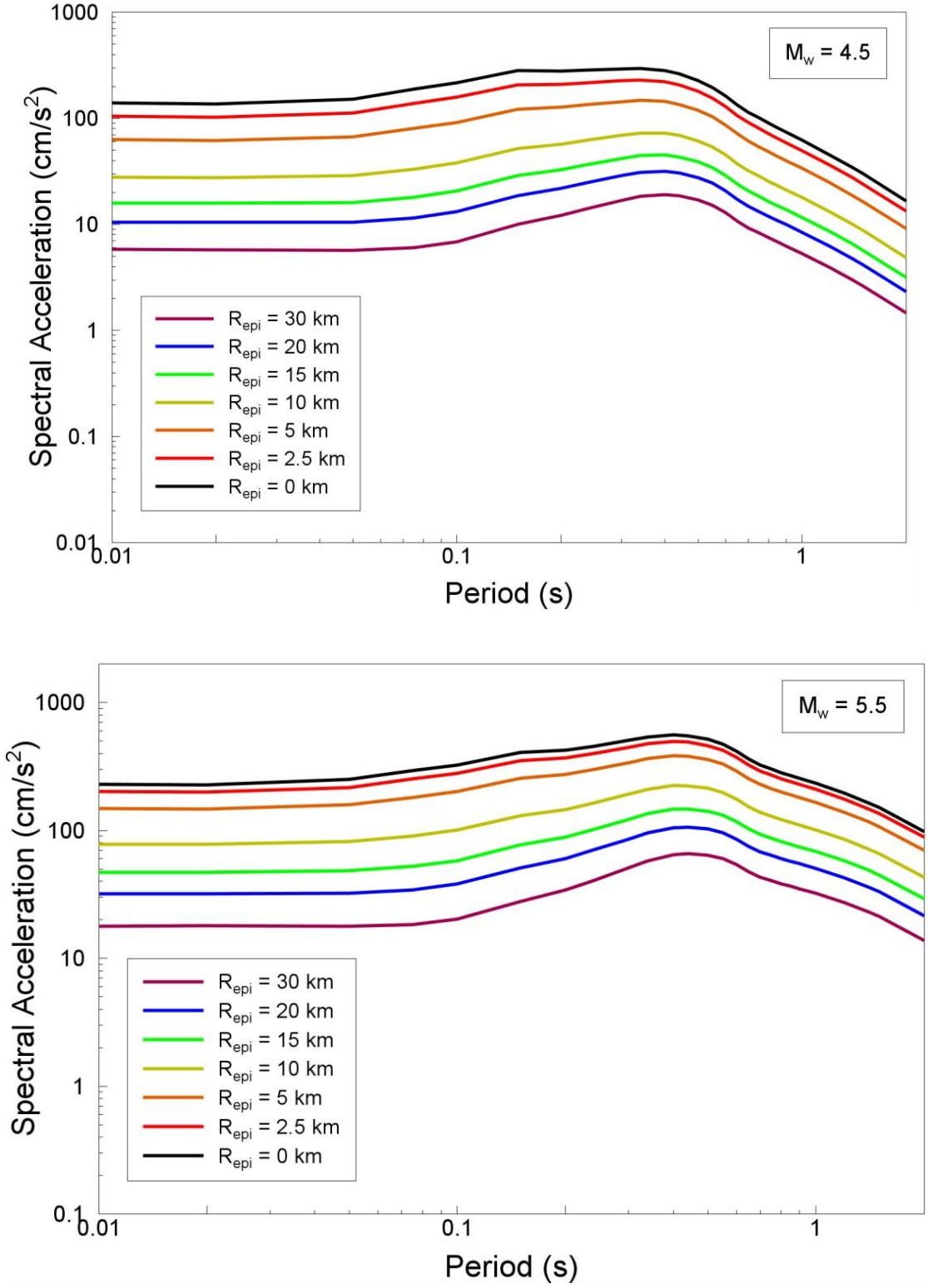


Figure 6.71. Predicted median response spectra for **M** 4.5 (*upper*) and **M** 5.5 (*lower*) at different epicentral distances.

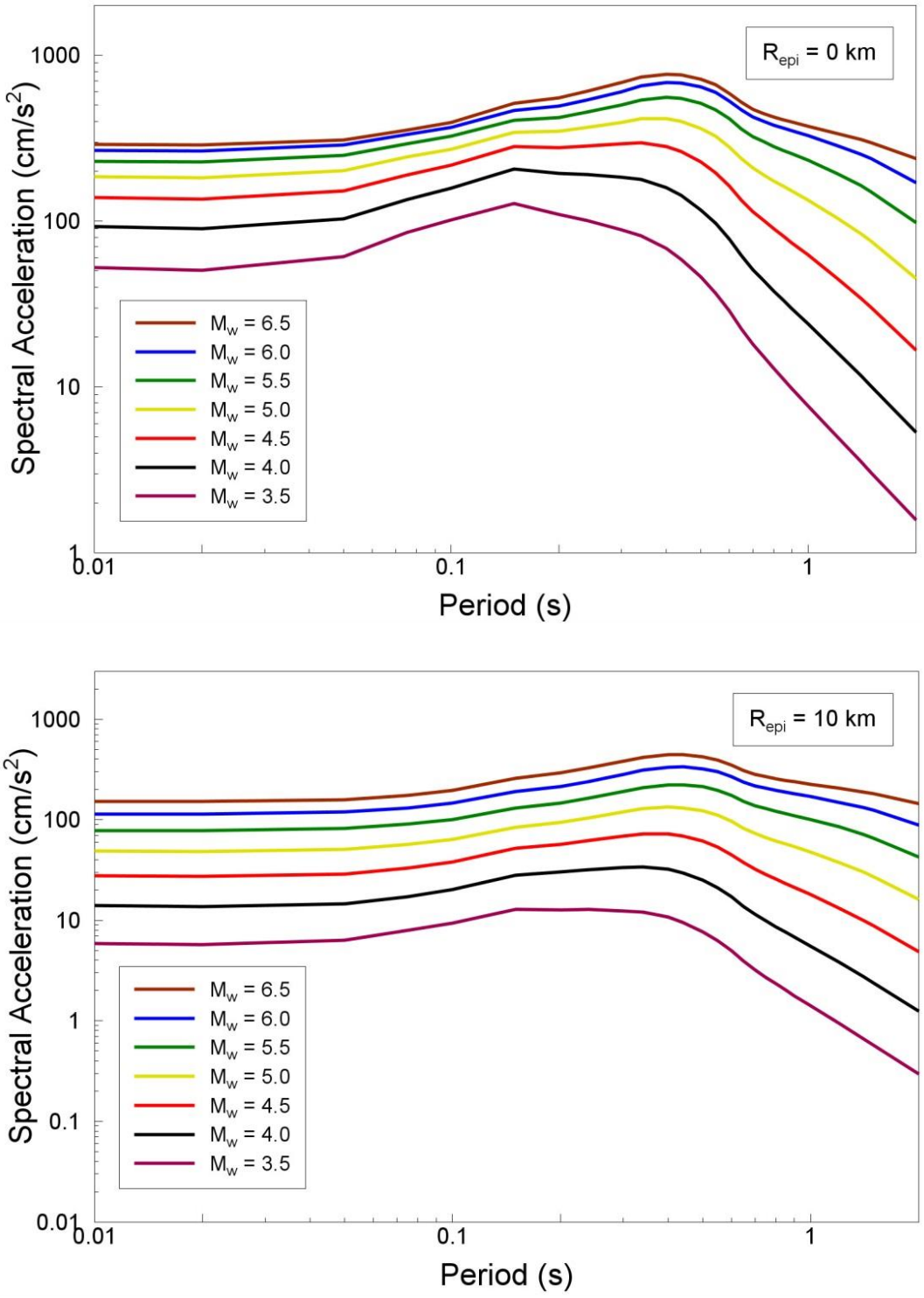


Figure 6.71. Predicted median response spectra at R_{epi} = 0 km (*upper*) and R_{epi} = 10 km (*lower*) for different earthquake magnitudes

This is also seen clearly in Figure 6.73, which shows normalised spectral shapes from the central model and from the original GMPEs of Akkar *et al*. (2014a) for three different magnitude-distance combinations. The differences at longer periods between the predictions from the Groningen model and those from the European model may reflect a

relatively small proportion of deep soil sites in the European data set, but equally it could result from the fact that the Groningen GMPEs are modelling only linear soil response, which may be leading to over-estimation of the spectral ordinates for larger magnitudes at short distances. The plots in Figure 6.73 show that the spectral shapes associated with the higher and lower alternative Groningen GMPEs are similar to those from the central model, whence the coefficients in Table 6.8 could be used to obtain spectral shapes to be applied to any of the three logic-tree branches.

Table 6.8. Coefficients of Eq.(6.1) for the median accelerations of the central model at additional response periods for defining spectral shapes

Period (s)	c_1	c_2	c_3	c_{3a}	c_4
0.02	1.0630	1.2860	-0.3389	-0.1381	-1.4948
0.05	1.2086	1.2968	-0.2759	-0.1421	-1.5577
0.075	1.7287	1.2519	-0.2330	-0.1242	-1.6343
0.1	2.1158	1.1958	-0.2560	-0.1024	-1.6335
0.15	2.6975	1.1084	-0.3466	-0.0696	-1.5798
0.24	2.3014	1.1538	-0.4814	-0.0875	-1.4208
0.3	1.8760	1.2336	-0.5290	-0.1183	-1.3503
0.34	1.5623	1.2961	-0.5461	-0.1417	-1.3147
0.4	0.9941	1.3995	-0.5547	-0.1793	-1.2729
0.44	0.5759	1.4703	-0.5525	-0.2043	-1.2514
0.55	-0.6055	1.6561	-0.5276	-0.2665	-1.2116
0.6	-1.1450	1.7321	-0.5100	-0.2898	-1.2015
0.65	-1.6748	1.8049	-0.4895	-0.3109	-1.1946
0.7	-2.1667	1.8768	-0.4684	-0.3312	-1.1886
0.75	-2.6101	1.9457	-0.4481	-0.3504	-1.1829
0.8	-3.0183	2.0102	-0.4283	-0.3678	-1.1780
0.85	-3.3974	2.0707	-0.4089	-0.3833	-1.1737
0.9	-3.7500	2.1273	-0.3903	-0.3973	-1.1699
1.2	-5.4457	2.3918	-0.2925	-0.4494	-1.1569
1.4	-6.2732	2.5119	-0.2418	-0.4615	-1.1532
1.5	-6.6178	2.5591	-0.2206	-0.4628	-1.1521

The differences between the spectral shapes shown in Figure 6.73 are so marked—and since the absence of non-linear soil response in the V1 GMPEs is clearly the most significant shortcoming—an experiment was conducted in making an approximate adjustment was made to the response spectral shapes as follows: for each scenario considered, the spectrum was generated using the V1 GMPEs and then equivalent rock spectrum generated by removing the linear part of the Akkar *et al*. (2014a) site response term for V_{S30} = 200 m/s. The non-linear spectrum was then obtained by applying the non-linear site amplification term from the same GMPE for the same V_{S30} value. Although this is a generic rather than region-specific adjustment for non-linear soil response, the impact for stronger scenarios is quite pronounced, as shown in Figure 6.74. This also confirms the great importance of incorporating non-linear site response into the GMPEs as one of the key elements in the next phase of development, as discussed in Section 8.4.

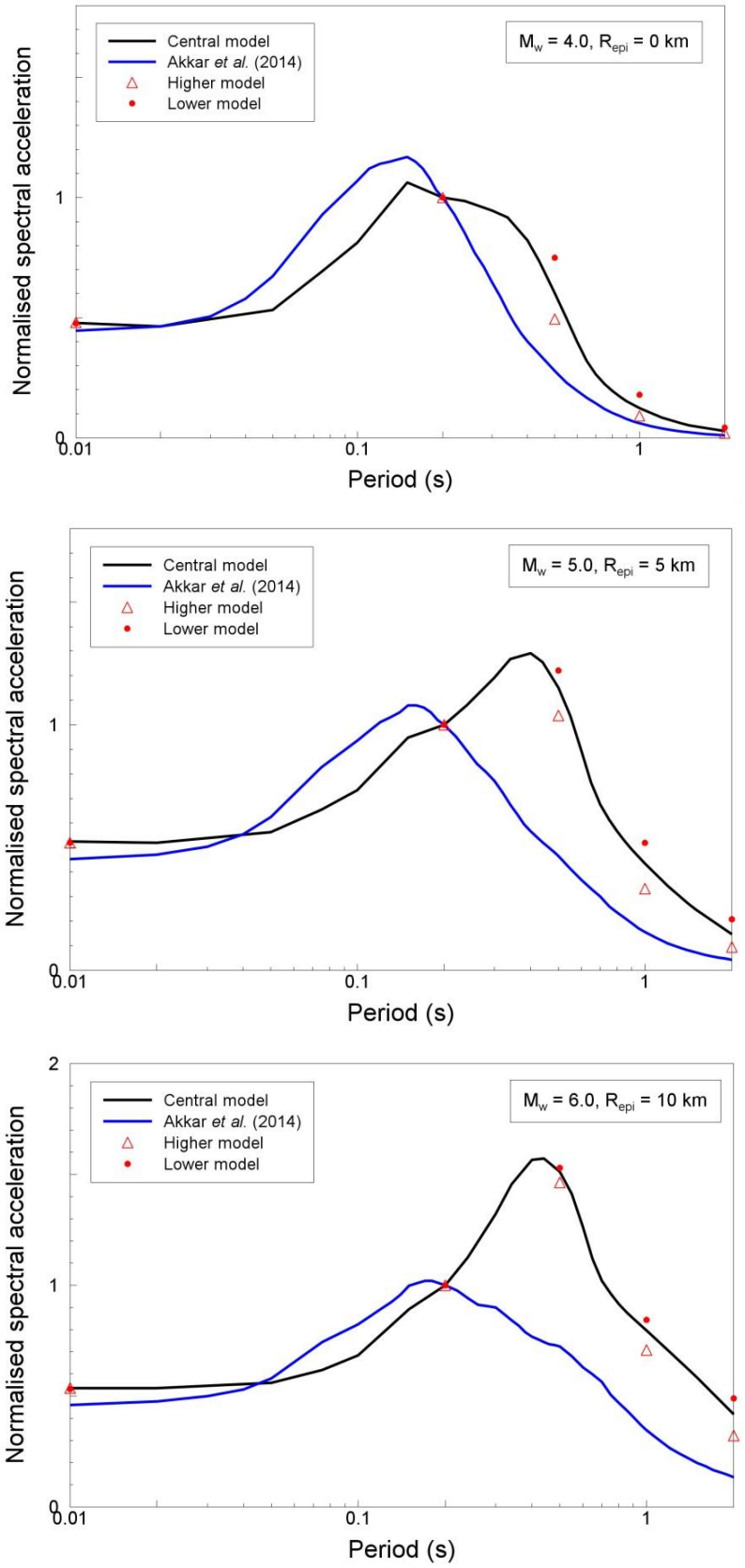


Figure 6.73. Predicted median response spectra, normalised to the ordinate at 0.2 second, from the central Groningen model and the original Akkar *et al*. (2014a) GMPEs, for different combinations of magnitude and distance. The symbols show the normalised ordinates from the higher and lower Groningen GMPEs.

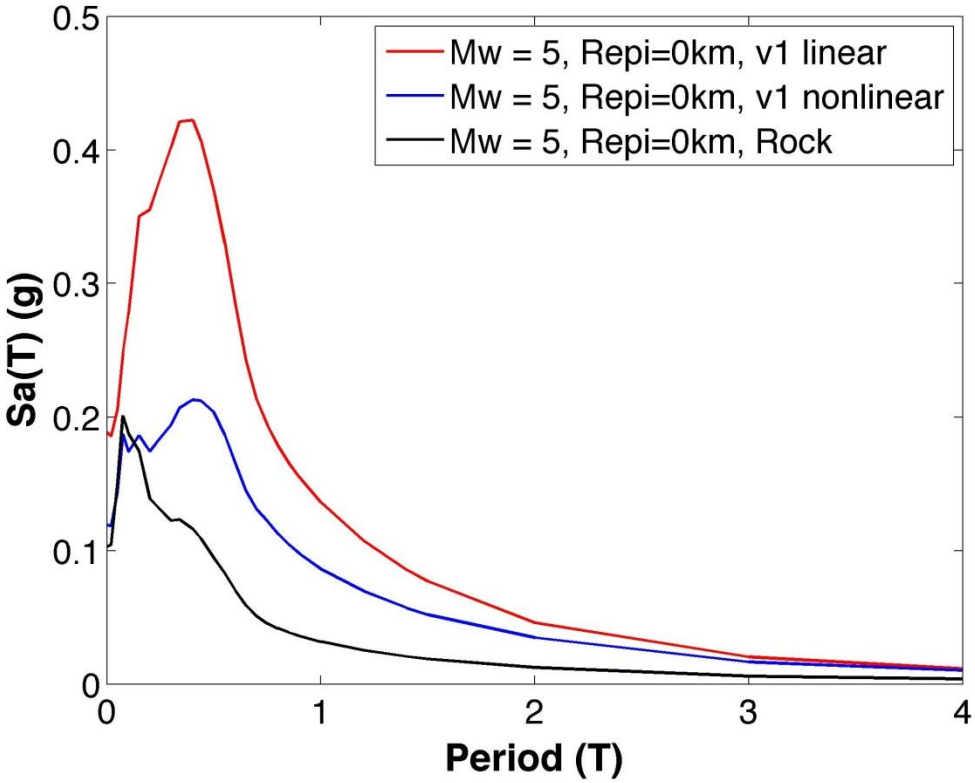


Figure 6.74. Predicted median response spectrum from the V1 model for an earthquake of **M** 5.0 at the epicentre and adjustments for non-linear response using the linear site response term from Akkar *et al*. (2014a) to transfer the spectrum to rock (V_{S30} 760 m/s) and then applying the non-linear site term for V_{S30} 200 m/s to this rock spectrum. Figure courtesy of Helen Crowley.

Since the derivation of the Version 1 fragility functions involves the use of equivalent elastic periods and equivalent damping to represent the deformed shape of structures responding non-linearly to seismic excitation, estimates of the response spectral displacements at periods greater than 2 seconds are sometimes needed. Figure 6.75 shows predicted median displacement spectra at the epicentre for a range of magnitudes (split into two groups for greater clarity), plus their extrapolation to longer periods obtained by linearly extending the interval between 1.5 and 2.0 seconds.

These extrapolations are sensitive to the shape of the spectrum in the interval from 1.5 to 2.0 seconds and to any difference in the ordinates at these two periods. For the smaller earthquakes, the small differences that lead to a decaying displacement spectrum at longer periods may be an artefact resulting from inaccuracies in the predictions. For an earthquake of magnitude of **M** 5, it is possible that intermediate ordinates at the periods between 1.5 and 2.0 seconds would show a much greater degree of flattening than is apparent in these plots. All of these observations lead us to recommend that for smaller earthquakes, the displacements should simply be held constant at the 2-second ordinate to obtain the spectrum at longer periods. By assuming equivalence between the corner frequency of the Fourier spectrum and the period, T_M, at which the spectral displacements reach a plateau—as imposed, for example, in the GMPEs of Abrahamson & Silva (2008)—then using the Brune (1970) formulation the relationship between this period and

magnitude, for the assumed central stress drop value of 30 bars, is given by the following equation:

$$\log_{10}(T_M) = 0.5M - 2.3764 \qquad (6.8)$$

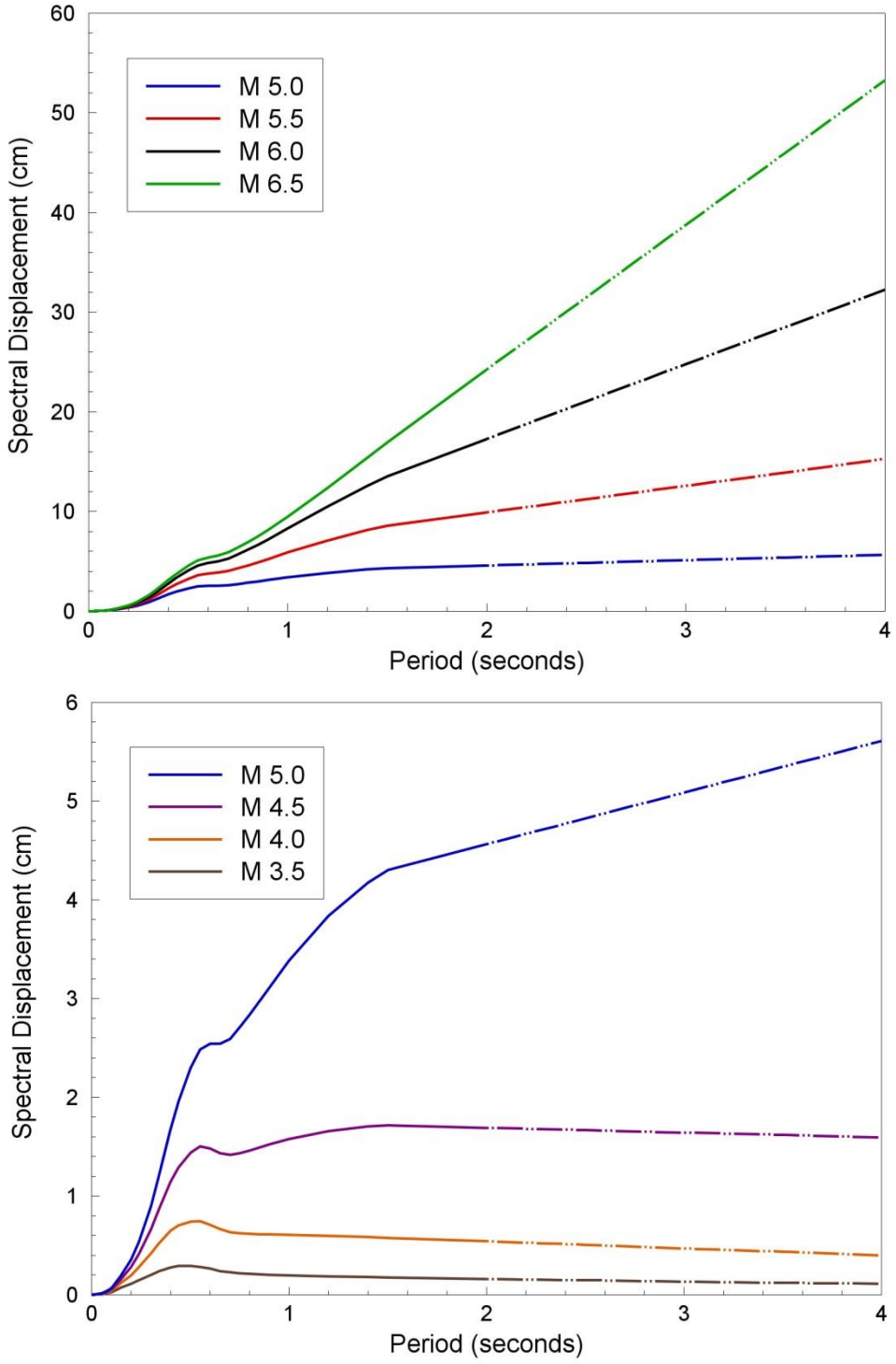


Figure 6.75. Predicted median displacement spectra at R_{epi} = 0 km for various magnitudes, and the linear extrapolation of the final part of the spectrum

This equation yields a value of 2 seconds for **M** 5.35, so if estimates of the spectral ordinates beyond 2 seconds are required, the ordinates at the longer periods should be obtained as follows, in all cases the starting point to transform the acceleration response spectrum to displacements via the pseudo-spectral relations:

- For **M** < 5.35, maintain the spectral displacement at T = 2 seconds constant
- For **M** ≥ 5.35, extrapolate linearly the displacement spectrum between periods of 1.5 and 2.0 seconds up to a period T_M, after which the displacement should be maintained constant; the period T_M is given by Eq.(6.8).

6.6. R_{JB}-based GMPEs

As explained in Section 2.4, for pragmatic reasons for computational efficiency, in the V1 hazard and risk models earthquake sources are modelled as points (hypocentres) for which the use of a GMPE based on epicentral distance is internally consistent. However, for larger earthquakes it is clearly physically more realistic to represent the extension of the fault rupture although there is a computational penalty to pay for this refinement. Therefore, in order to decide whether the V2 and subsequent GMPEs will continue to use R_{epi} or should adopt the R_{JB} distance metric, a sensitivity analysis will be performed as described in Section 8.1. In order to conduct this experiment, a GMPEs are required that are derived in an identical fashion to those presented above in Sections 6.1 and 6.2 but using R_{JB} instead of R_{epi} as the measure of source-to-site separation.

In the case of stochastic simulations for the V1 GMPEs implemented by the code SMSIM this is the hypocentral distance, R_{hyp}, the distance between the earthquake hypocentre and the site. The reason for this is that hypocentral distance is compatible with simple models of geometrical spreading. For example, in an elastic homogeneous full space the loss of amplitude is proportional to $1/R_{hyp}$. The assumption made by using point-source distance metrics is that all energy is released at a point in space. In practice, even for small events this is not the case. The disparity is accommodated in the v1 GMPE by using a pseudo-distance, which is formed by the epicentral distance along with a saturation term, dependent on magnitude, *h*(M):

$$R = \sqrt{\mathrm{R}_{\mathrm{epi}}^2 + h(M)^2} \tag{6.9}$$

R_{epi} is the distance between the earthquake epicentre and the site of interest, and *h*(M) is defined through direct regression, as explained in Section 6.1:

$$h(M) = \exp(0.4233M - 0.6083) \tag{6.10}$$

This functional form provides predictions that appropriately saturate in the near-field for larger earthquakes—with h(M) greater than the true source depth—but it still results in a

'bulls-eye' shape to the ground-motion field at the surface. In reality both observed and fully simulated ground-motion fields are commonly found to take the form of a race-track, centred around the surface projection of the fault. In this updated V1 GMPE the aim is to reflect this more realistic surface distribution of ground motion. GMPEs for large tectonic earthquakes achieve this by adopting a finite-distance metric, R_{JB}, which is the closest distance to the surface projection of the fault. In the case of stochastic simulations there are two options to include this effect: the first is to construct sources using a summation of distributed point sources. The simulated ground motion and corresponding measured finite-distance metrics can then be regressed using an appropriate functional form. However, such approaches require a significant amount of information that is typically not available: such as slip distribution. This limitation can be overcome by bootstrapping simulations with randomization of the unknown parameters, although it is rather inefficient. Alternative approaches include the so-called effective distance metric R_{eff} (Boore, 2009) and the method recently implemented by Yenier & Atkinson (2015). The concept of R_{eff} is that, for a given (finite) source and receiver location, R_{eff} represents the distance that provides compatible attenuation characteristics to those that would be observed by summing contributions from sub-faults across the fault surface. It has been shown that this approach provides similar results to the more involved process of constructing and randomizing a finite fault from a distribution of point-sources (Boore, 2009) in the case that fault rupture characteristics (*e.g.*, slip distribution, hypocentre location) are not known, as is the case for PSHA.

The additional input for the R_{eff} approach (with respect to the point-source approach) is:

a) the source dimensions (length and width);
b) the source orientation (strike) – not known in this case;
c) the frequency (f_Q) at which to match attenuation using R_{eff} vs the sum over sub-faults. The default is 10 Hz: Boore (2009) states that "*the results are not sensitive to the choice of* f_Q", so we do not consider further investigation of this term necessary.

A clear unknown in this case is the strike of the fault. However, since we aim to determine generic ground-motions at R_{JB} distances (rather than an event-specific scenario), it is sufficient to simply use a single strike (*e.g.*, 0º) 'instrumented' with a dense array of receivers. In SMSIM the receiver location can be defined using an arbitrary grid: the output calculates the corresponding R_{JB} distance (in addition to other commonly used metrics) with the simulated ground motion.

For consistency we adopted the same choices and models in the simulation of the extended rupture characteristics as those used in deriving the adjustments for the sigma when using point-source distances, as described in Appendix II. The equation for estimating fault rupture lengths, L (km), is from Wells & Coppersmith (1994):

$$\log_{10}(L) = 0.59M - 2.44 \qquad (6.11)$$

with a logarithmic standard-deviation of 0.16. In addition, the corresponding rupture width, W (km), is given by the following equation from Wells & Coppersmith (1994):

$$\log_{10}(W) = 0.32M - 1.01 \quad (6.12)$$

with a standard deviation of 0.15. In order to account for the uncertainty in L and W, the standard deviation is sampled at -2.8570, -1.3556, 0, 1.3556 and 2.8570 sigma, with appropriate weights (0.0113, 0.2221, 0.5333, 0.2221 and 0.0113 respectively) based upon equating moments and using Gauss-Hermite quadrature (Miller & Rice, 1983). The depth of the sources in Groningen is considered to be within the reservoir, at 3km. This is considered to be the hypocentre, such that the 'finite' faults will extend symmetrically, until they reach the surface, at which point the increasing width will be entirely accommodated by downward propagation. The dip of faults in the Groningen field are generally found to be 'near-vertical' (Figure 6.76) and hence, consistent with the simulations to estimate the impact of point-source metrics on aleatory variability (Appendix II), we assume vertical faults.

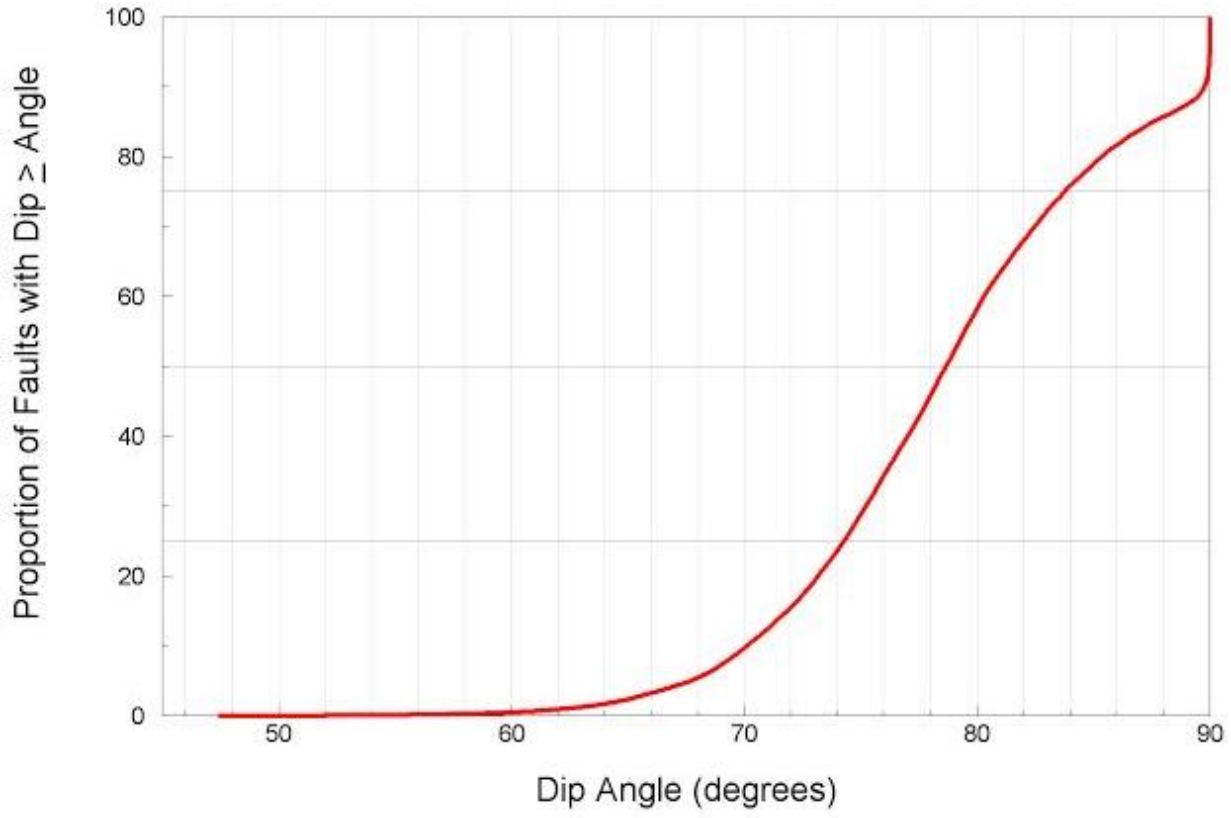


Figure 6.76. Cumulative distribution of dips of fault segments in the Groningen field

In order to maintain consistency with the previous iteration of the main V1 GMPEs, the saturation term, Eq.(6.9), is implemented in addition to the R_{eff} approach. Whilst this may seem to double-count saturation effects, we consider the former to account entirely for point-source based saturation, while the R_{eff} approach accounts for the effect of distributing those point sources in space. Similarly to the calculations performed for

assessing the sigma for the Re_{pi} model, the distribution of recording sites was defined by a grid, centred on the middle of the fault strike with inter-station spacing of 1 km in the first 10 km, 2 km spacing from 10 to 20 km, 4 km spacing from 20 to 32 km, then 10 km spacing from 40 km to the maximum station distance of 70km (Figure 6.77).

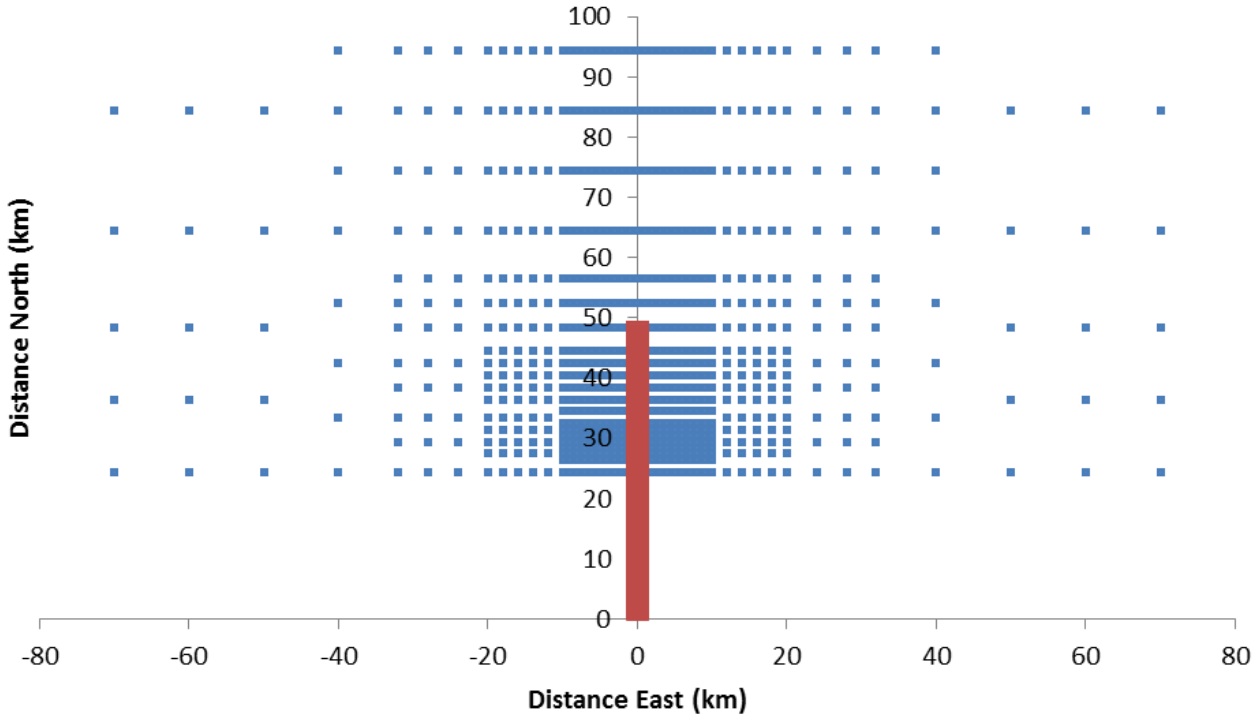


Figure 6.77. Station distribution (*blue*) around the strike of a large earthquake (*red*) earthquake fault rupture

The source-site receiver combinations depicted in Figure 6.77 led to a much larger number of simulated motions for each magnitude and distance combination than those generated for the original R_{epi}-based GMPE. In order to ensure that the two suites of equations—based on point-source and extend-source distance metrics—generated for the sensitivity study described in Section 8.1 were consistent in all respects other than the distance metrics, the R_{epi} values was also calculated for each source-receiver combination in the simulations. The median values of spectral acceleration for each **M**-R_{epi} combination obtained in this way are essentially identical to those obtained from the original simulations, but for the purposes of the exploratory GMPEs, regressions were performed on both the R_{epi} and R_{JB} datasets.

With the datasets generated, regressions were performed using the same functional form as in Eq.(6.1) and following exactly the same procedures. Although the simulations were generated for a large magnitude and distance range (**M** 1.0-6.6 and distances out to 100 km), the regressions are performed for a restricted range of these parameters so that we cover **M** 2.5-6.6 and distances out to 60km, consistent with the limits of application for the Groningen hazard and risk model.

The regressions were based upon a weighted non-linear least squares (no random effects because all the 'data' are simulated with perfectly uniform sampling of the independent variables). The only departure from just a standard non-linear least squares regression analysis is that weights are used to reflect the fact that the rupture lengths for any given magnitude are considered random. However, rather than simply randomly sampling value of the rupture length for each magnitude, as noted above the simulations consider values of epsilon (epsilon in terms of the rupture lengths) that are the same as those used for the variance correction model in Appendix II. So, the stochastic simulations generate one RVT-based motion for five values of rupture length for each magnitude (corresponding to the five epsilon values indicated above) and then these weights are used to reflect that fact that the epsilon 0 value is much more likely to occur than the epsilon 2.857 value, *etc*. The resulting coefficients for the two models are presented in Tables 6.9 and 6.10.

Table 6.9. Coefficients of Eq.(6.1) and the inter-event variability of the central R_{JB} GMPE

	PGA	**Sa(0.2s)**	**Sa(0.5s)**	**Sa(1.0s)**	**Sa(2.0s)**
c_1	1.1563	2.4972	-0.0684	-4.3882	-7.8093
c_2	1.2732	1.1216	1.5742	2.2288	2.6929
c_3	-0.3394	-0.4314	-0.5416	-0.3549	-0.1520
c_{3a}	-0.1342	-0.0747	-0.2397	-0.4202	-0.4370
c_4	-1.5048	-1.4806	-1.2266	-1.1640	-1.1526
c_5	0.4233	0.4233	0.4233	0.4233	0.4233
c_6	-0.6083	-0.6083	-0.6083	-0.6083	-0.6083
τ	0.2810	0.3337	0.3216	0.3789	0.3547
ϕ_{SM}	0.4918	0.4454	0.5146	0.4081	0.4133

Table 6.10. Coefficients of Eq.(6.1) and the inter-event variability of alternative R_{epi} GMPE

	PGA	**Sa(0.2s)**	**Sa(0.5s)**	**Sa(1.0s)**	**Sa(2.0s)**
c_1	1.3001	2.9388	0.4122	-4.1684	-7.8723
c_2	1.2988	1.1158	1.5157	2.2027	2.7270
c_3	-0.3790	-0.4200	-0.5752	-0.3933	-0.1531
c_{3a}	-0.1289	-0.0630	-0.1997	-0.3910	-0.4412
c_4	-1.5651	-1.6064	-1.2912	-1.1886	-1.1730
c_5	0.4233	0.4233	0.4233	0.4233	0.4233
c_6	-0.6083	-0.6083	-0.6083	-0.6083	-0.6083
τ	0.2810	0.3337	0.3216	0.3789	0.3547
ϕ_{SM}	0.4918	0.4454	0.5146	0.4081	0.4133

The values in Tables 6.9 and 6.10 also include the within-event and between-event standard deviations, which are taken to be the same as for the central V1 models, as presented in Section 6.2. The only difference is that for the R_{JB}-based model, the adjustment to the within-event variance for the use of a point-source approximation is clearly not needed, and therefore for this equation $\delta\phi = 0$.

At this point, a few comments are warranted regarding the decision to maintain the same functional form and the same near-source saturation term for both the R_{JB} model and the new R_{epi} model. In this regard, it is worth recalling that the near-source saturation term was derived using regressions on the empirical data from the field, for which R_{epi} and R_{JB} are not distinguishable metrics because of the magnitude range of the data. If we did have empirical data over a greater magnitude range then it would be reasonable to assume that we might find different h(M) terms for R_{epi} and R_{JB}. However, in practice if we independently derived separate h(M) terms for R_{epi} and R_{JB} we could end up with large differences purely due to the uncertainties in the independent R_{epi} and R_{JB} values that would most likely be large enough to cause small inconsistencies in the two h(M) functions. The other point to bear in mind is that the simulations use an effective distance that is calibrated to the h(M) function derived for the R_{epi} case, so using the same function for both R_{epi} and R_{JB} is consistent.

Another point worthy of note, although already discussed with regards to the original equations, is that the functional form does not include an explicit term to account for anelastic attenuation. That is, the functional form does not account in any way for Q effects that are generated from the stochastic simulations. We know from the residual plots that there are some trends with respect to distance that could potentially be removed, or at least reduced. However, we deliberately have not looked at this because of the desire to maintain consistency with the R_{epi}-based model that was already implemented for the hazard and risk calculations. One other point about these Q trends to emphasise, however, is that it is not simply a case of being able to say that our models are conservative by not including a term for Q. The reason for this is that the simulated values show some deviations from a linear scaling of log-amplitude with log-distance. What we do is effectively fit a straight line through these ‘data’. Hence for large distances the model may be conservative, but at intermediate distances it can be under-predicting the motions. Whether the models are actually under-estimating motions at some distances in reality depends upon how ‘true’ the Q estimates are and the judgement is that from the preliminary inversions of the field data, with a rather crude geometric spreading model and a simple network-averaged amplification function, there is probably very considerable uncertainty in this parameter.

A consequence of using the same functional form to perform regressions on a much larger dataset of R_{epi}-based values is that the slightly different equations are obtained, despite the median values of the simulations being essentially the same as for the main V1 models. The two sets of median predictions are compared in Figures 6.78 to 6.82. The differences manifest mainly at short distances and do not seem to show any very strong trend with magnitude, but to diminish with increasing response period. The clear observation at the shorter response periods is that the new R_{epi} models are predicting slightly higher median values than the original V1 models. This is an interesting observation but not one of major importance given that the differences between the two sets of curves are small compared to the range of predictions captured in the logic-tree. The sampling of simulations as used for the derivation of the V1 equations was entirely consistent with standard practice in the generation of stochastic GMPEs.

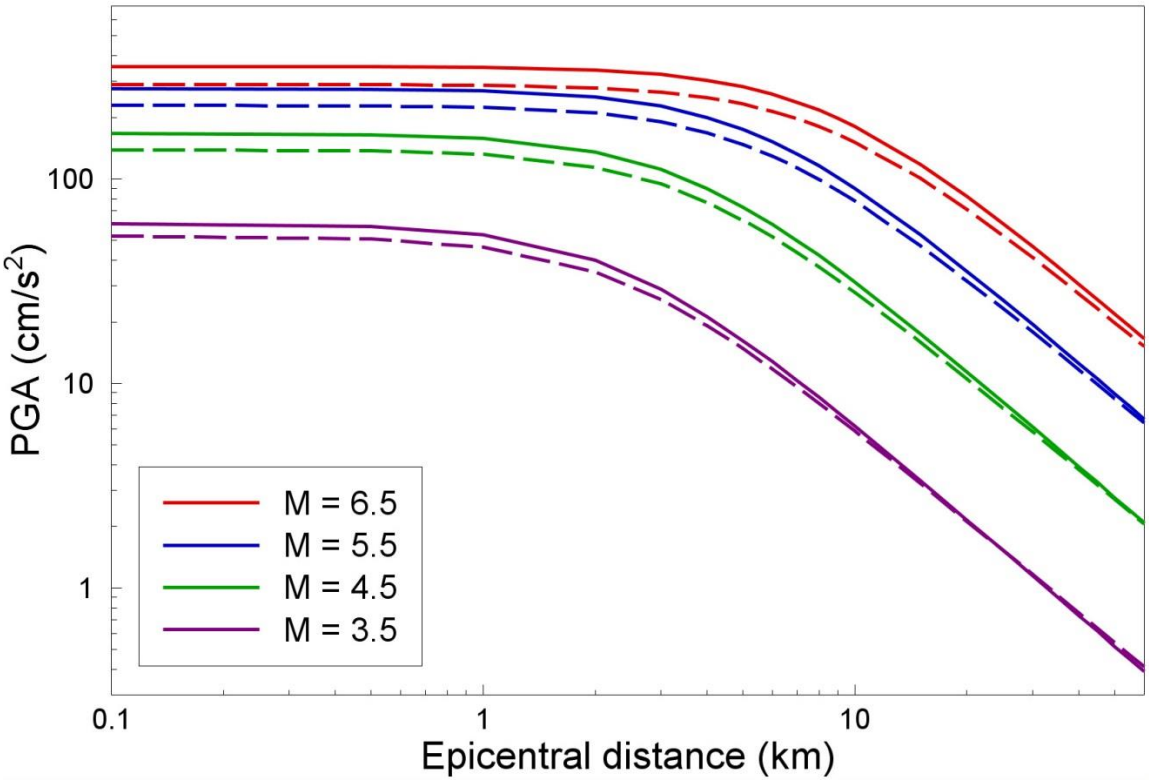


Figure 6.78. Comparisons of median predictions of PGA from the original (*dashed*) and alternative (*solid*) R_{epi}-based GMPEs as a function of magnitude and distance

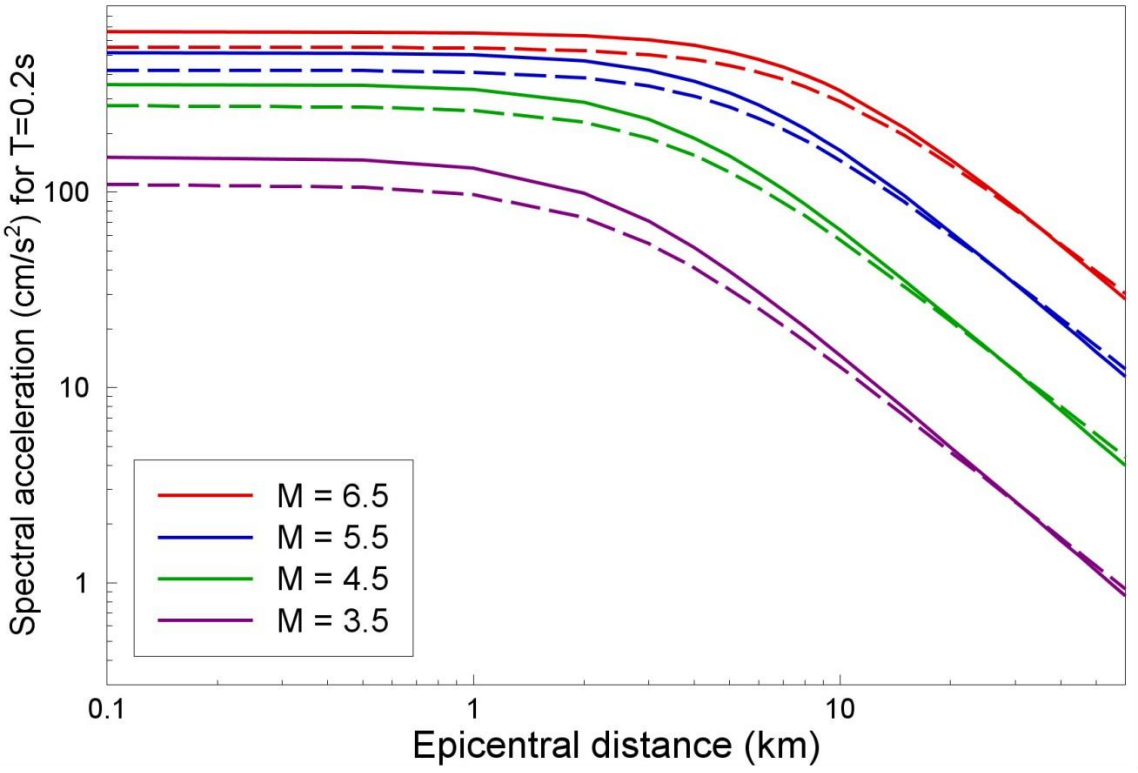


Figure 6.79. Comparisons of median predictions of Sa(0.2s) from the original (*dashed*) and alternative (*solid*) R_{epi}-based GMPEs as a function of magnitude and distance

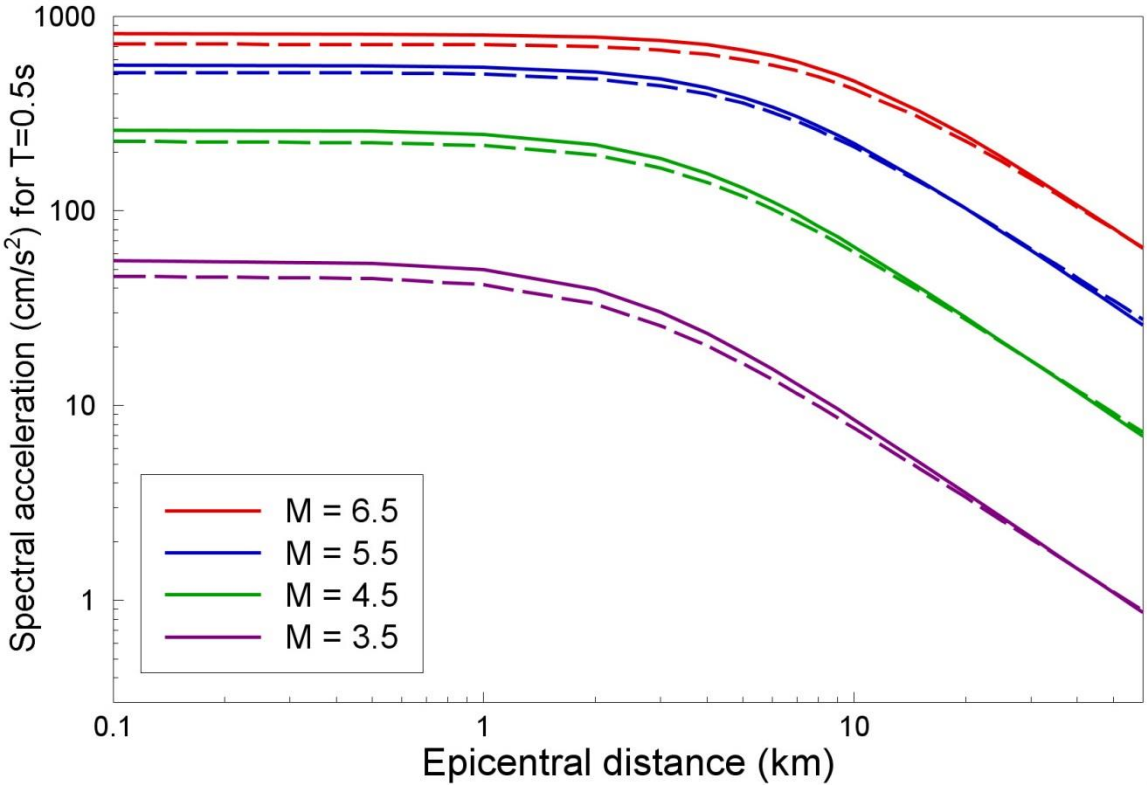


Figure 6.80. Comparisons of median predictions of Sa(0.5s) from the original (*dashed*) and alternative (*solid*) R_{epi}-based GMPEs as a function of magnitude and distance

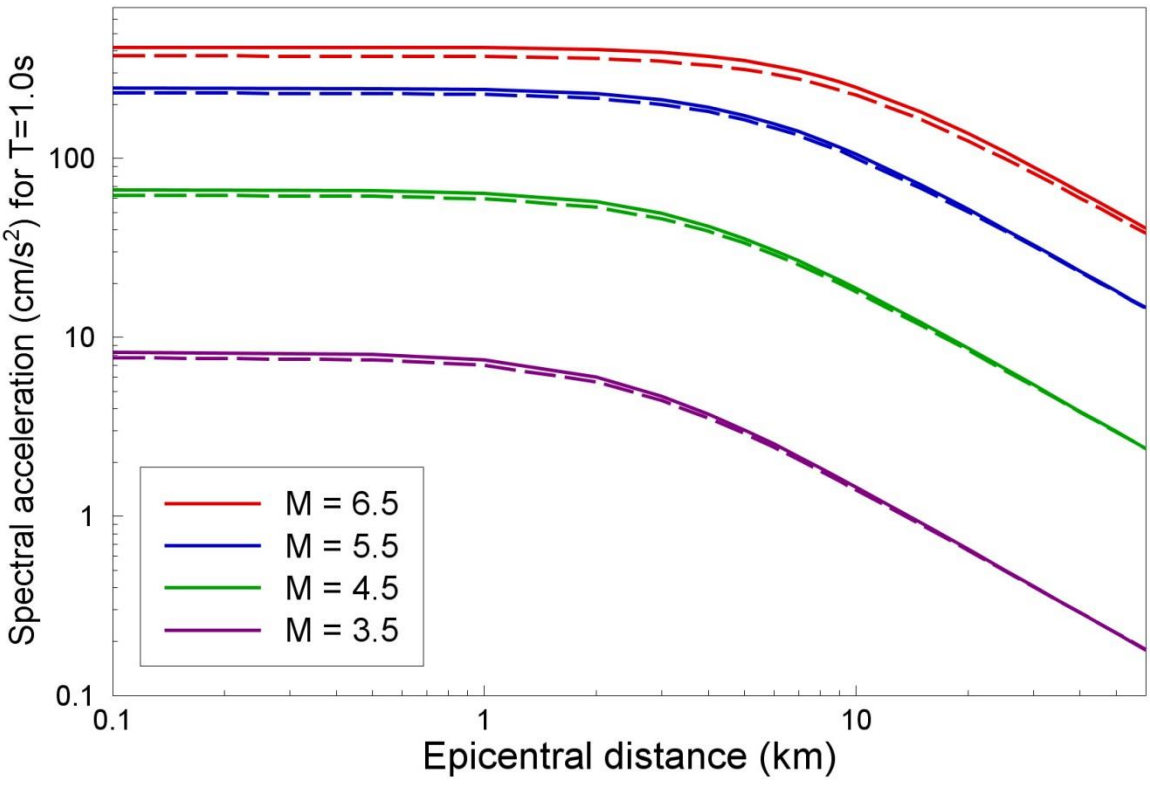


Figure 6.81. Comparisons of median predictions of Sa(1.0s) from the original (*dashed*) and alternative (*solid*) R_{epi}-based GMPEs as a function of magnitude and distance

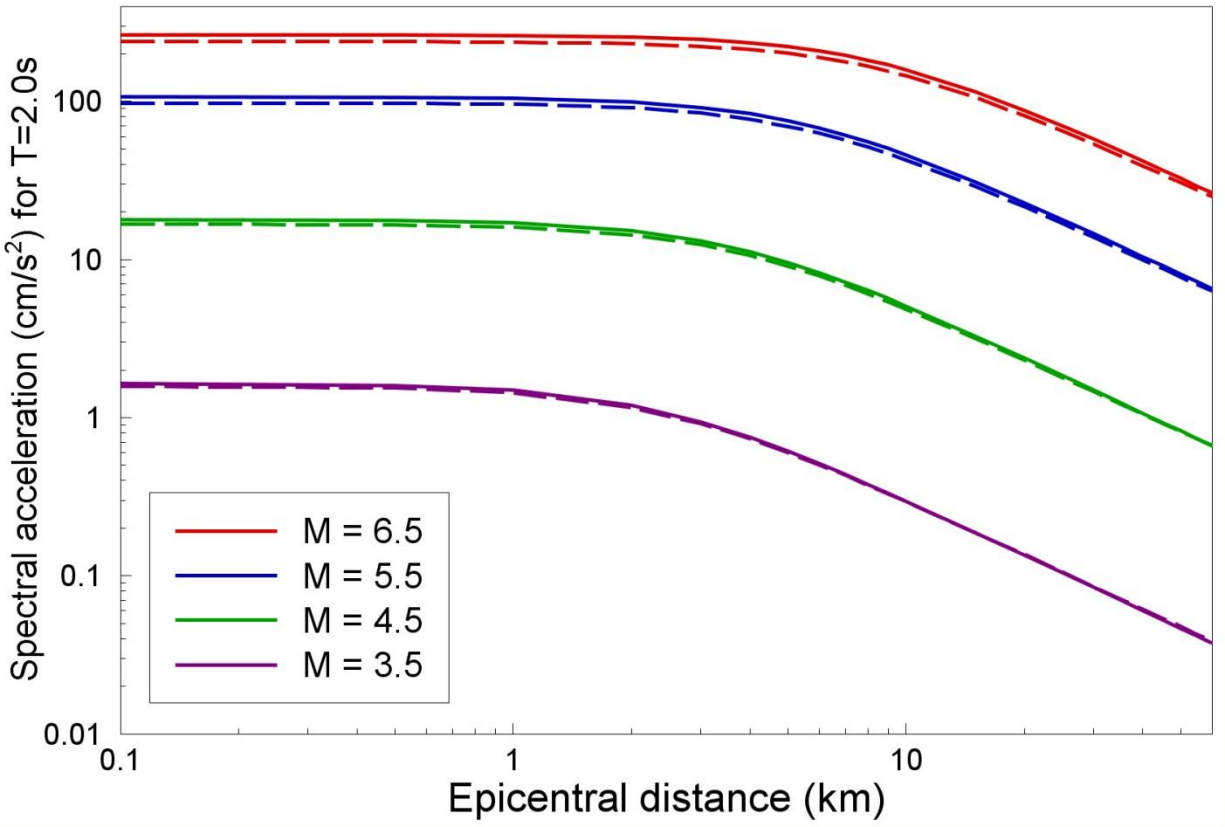


Figure 6.82. Comparisons of median predictions of Sa(2.0s) from the original (*dashed*) and alternative (*solid*) R_{epi}-based GMPEs as a function of magnitude and distance

To close this section and chapter, we present two sets of comparisons between the new R_{epi} and R_{JB} equations. Figures 6.83 to 6.87 compare the attenuation curves for different magnitudes plotted against 'distance', using the metric native to each equation. This is potentially somewhat misleading since the x-axis values do not correspond to the same quantity for both curves, but nonetheless the plots are useful for determining whether the behaviour predicted by the models is as expected. Indeed, it was in generating such plots comparing the R_{JB} equations with the original V1 equations that we realised that unexpected differences (higher values from the R_{JB} model at short distances) were arising as a result of the equations being derived from simulations over very different combinations of source and site locations. The patterns observed are very much as would be expected, with the R_{epi} model predicting larger values of acceleration than the R_{JB} model for a given distance, with the difference being strongly dependent upon magnitude (and increasing with the size of the earthquake) and diminishing—and effectively vanishing for the smaller magnitudes—with increasing distance. The differences seen even for the smallest magnitude (**M** 4) may be a little surprising but the differences are rather small and may, in part, be the result of using an identical near-source saturation term for both equations.

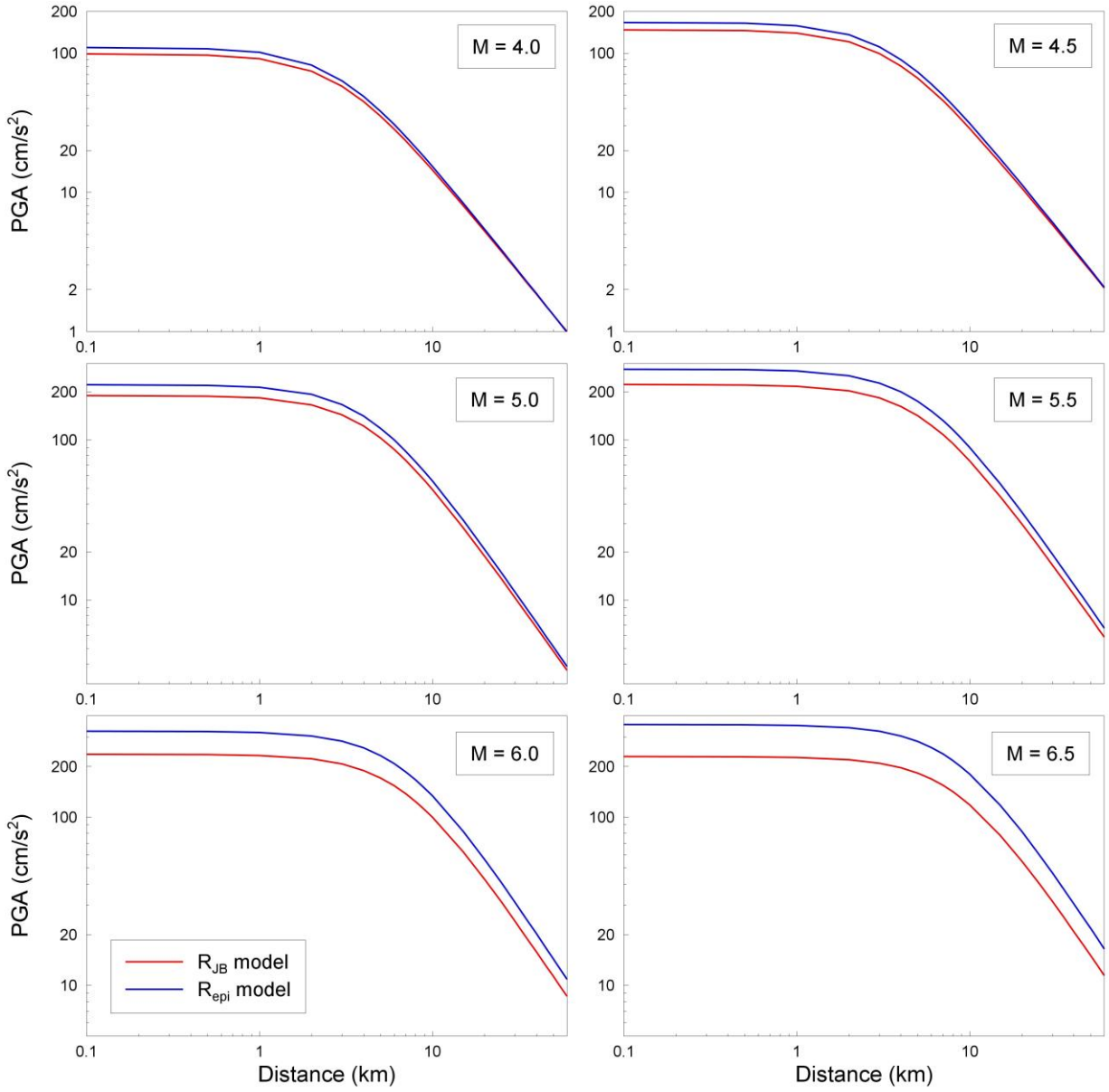


Figure 6.83. Comparisons of median predictions of Sa(0.01s) from the new R_{epi}- (*blue*) and R_{JB}- (*red*) based GMPEs as a function of magnitude and distance

In order to also make physically more meaningful comparisons, Figures 6.85 to 6.89 show the predictions from the new R_{epi} and R_{JB} models in a different way; since the observed patterns are consistent across the period range, the results are shown only for the limiting periods. Rather than plotting these together with each model using its native metric, the plots are shown against R_{JB} with the same scheme used in the derivation for simulations sampled to estimate the corresponding R_{JB} value for each R_{epi}, in a manner that in effect is not dissimilar to that proposed by Scherbaum *et al.* (2004) for converting one distance metric to another. However, in order to see the relationship more clearly, the R_{epi}-based predictions are plotted for intervals of azimuths from strike-normal (90º) to strike-parallel (0º). Similar observations can be made at all five response periods: firstly, for the smallest magnitude considered (**M** 4), the predictions from both models are essentially identical, except at very short distances (< 1 km) for which the R_{epi} models yield slightly higher values. This small difference is likely to be the result of imposing exactly the same near-

source saturation on both models. For greater magnitudes, larger differences emerge and the strong azimuthal dependence of the R_{epi} values becomes apparent. The 90º azimuth curve represents the strike-normal path, which for the assumed model of vertically-dipping strike-slip ruptures, yields identical values of the two distance metrics for a given location. For these comparisons, as would be expected, the R_{epi}-based prediction is consistently greater than that from the R_{JB}-based model, and the difference increases with magnitude. In practice, the differences in the effects would be further amplified by the larger sigma values associated with the R_{epi}-based model for larger magnitudes at certain distances.

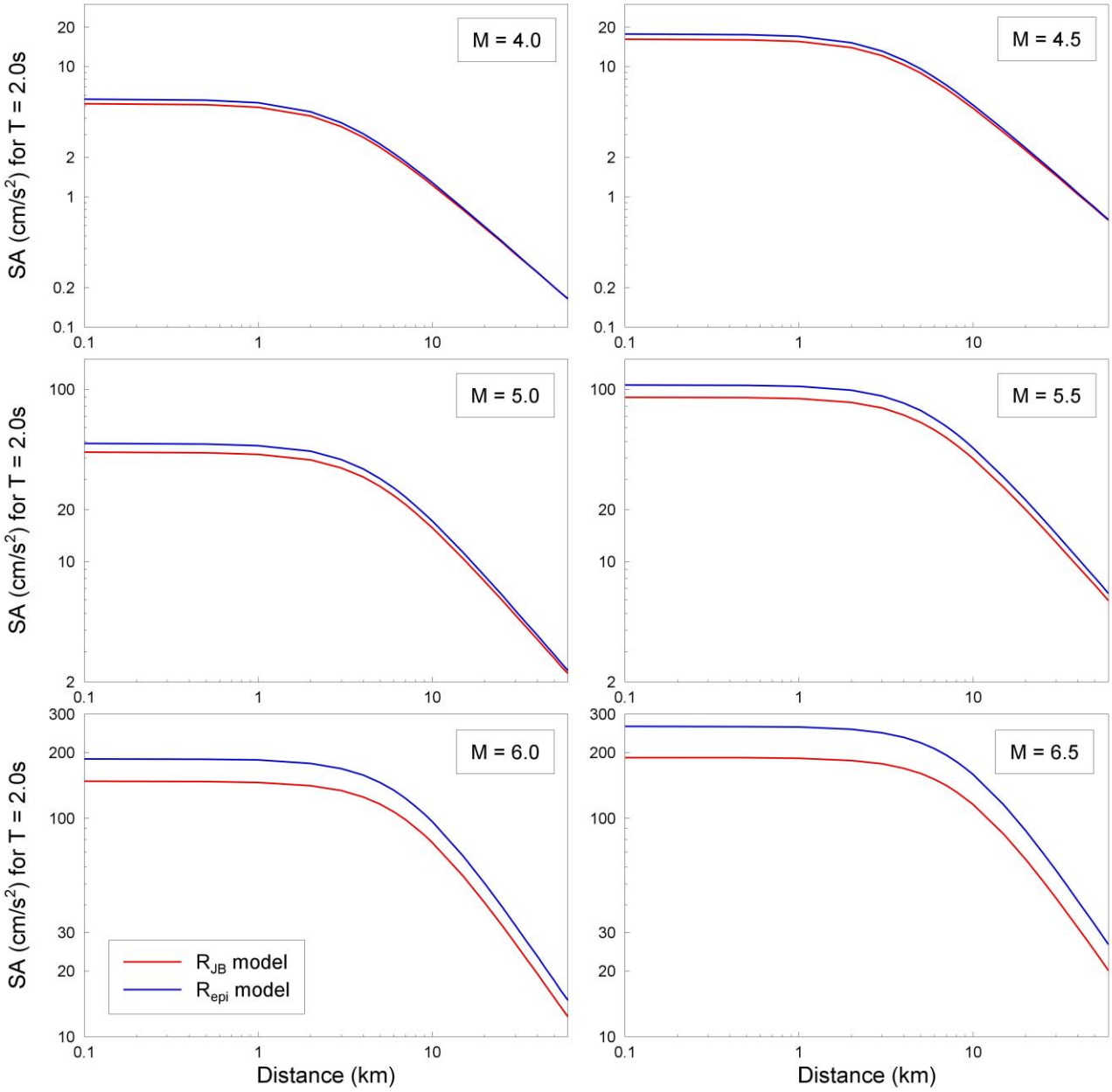


Figure 6.84. Comparisons of median predictions of Sa(2.0s) from the new R_{epi}- (*blue*) and R_{JB}- (*red*) based GMPEs as a function of magnitude and distance

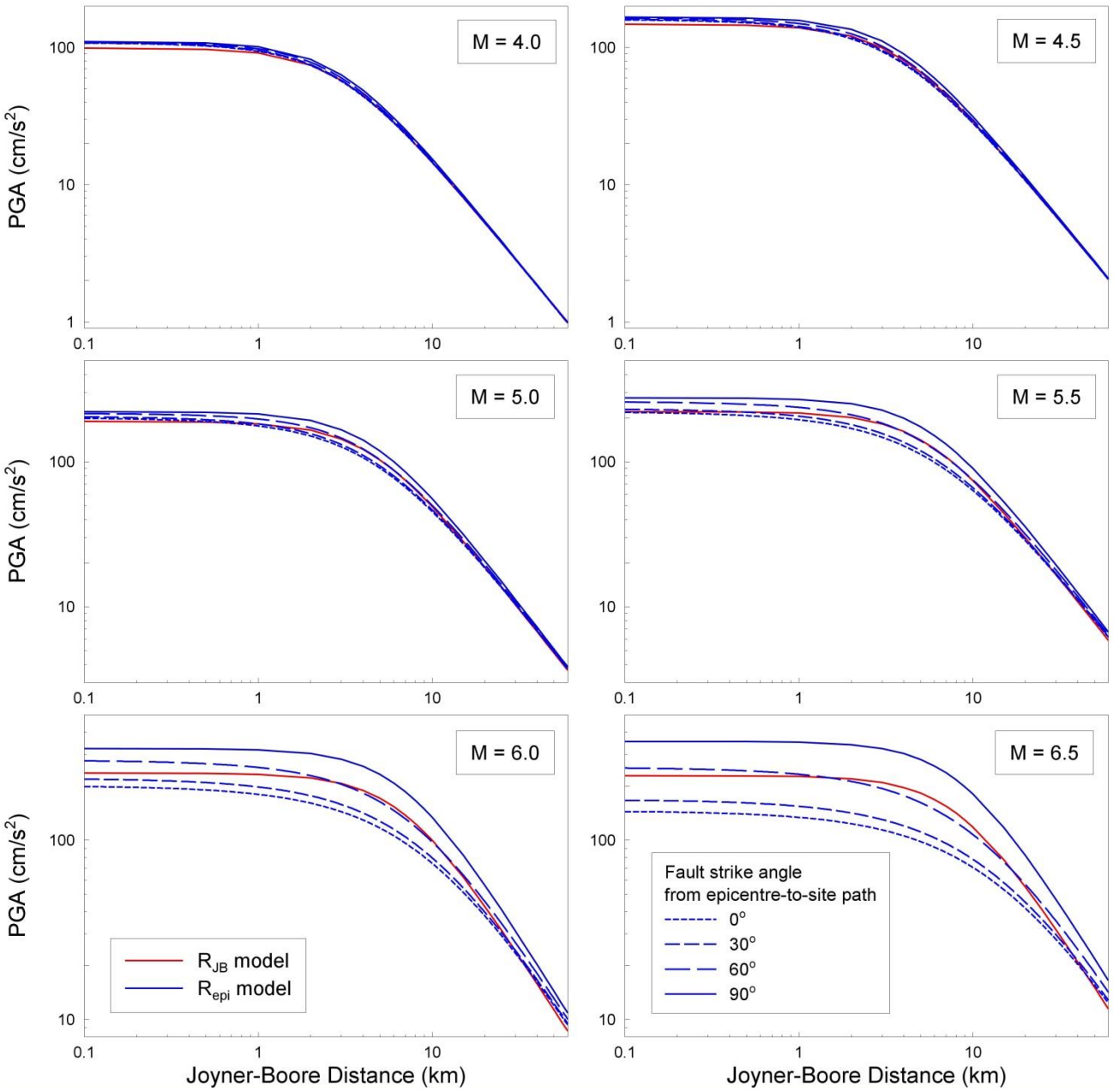


Figure 6.85. Comparisons of median predictions of PGA from the R_{JB} (*red*) and R_{epi} (*blue*) GMPEs as a function of magnitude and R_{JB} distance; the R_{epi}-based predictions are grouped by azimuth

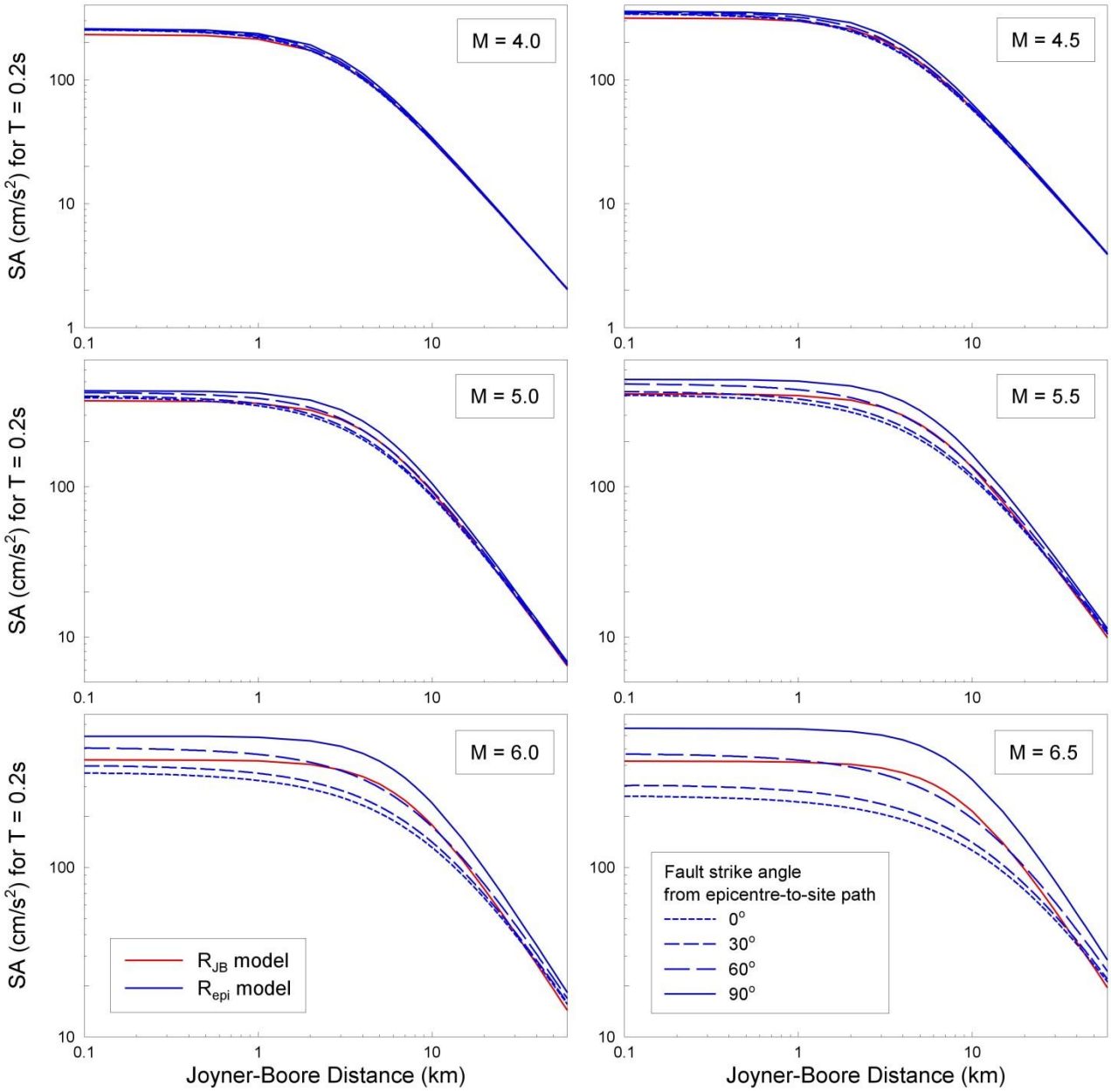


Figure 6.86. Comparisons of median predictions of Sa(0.2s) from the R_{JB} (*red*) and R_{epi} (*blue*) GMPEs as a function of magnitude and R_{JB} distance; the R_{epi}-based predictions are grouped by azimuth

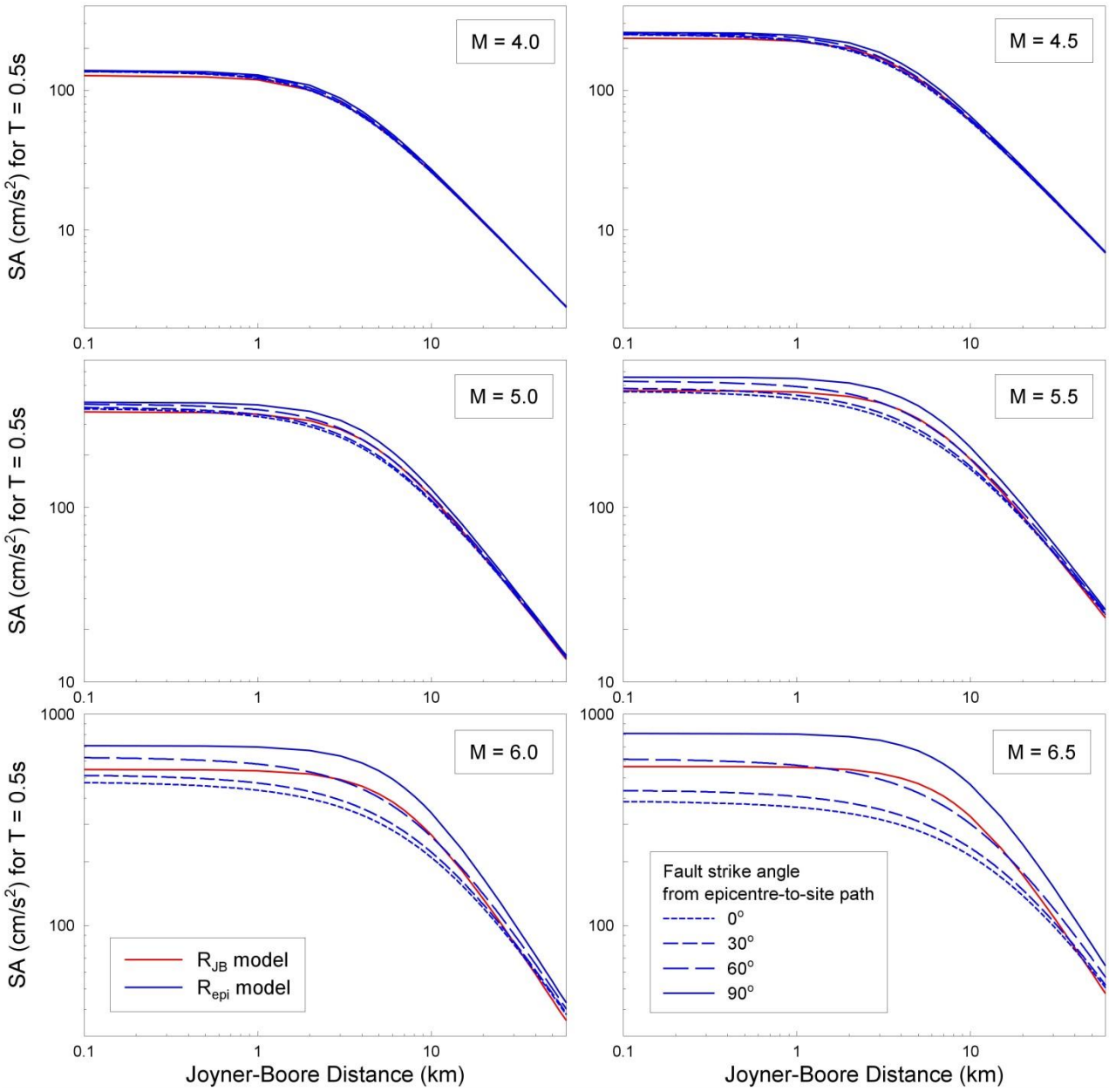


Figure 6.87. Comparisons of median predictions of Sa(0.5s) from the R_{JB} (*red*) and R_{epi} (*blue*) GMPEs as a function of magnitude and R_{JB} distance; the R_{epi}-based predictions are grouped by azimuth

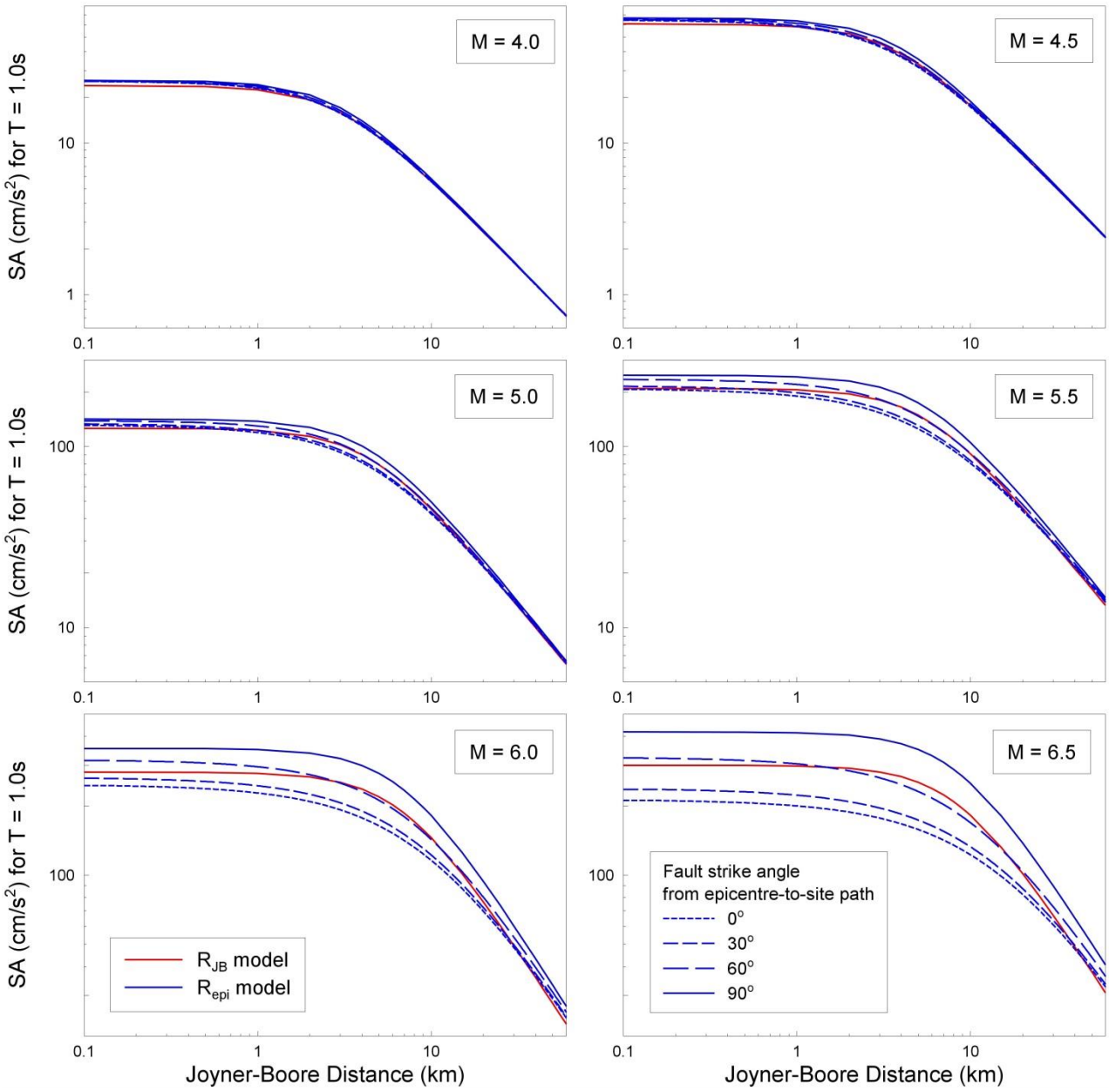


Figure 6.88. Comparisons of median predictions of Sa(1.0s) from the R_{JB} (*red*) and R_{epi} (*blue*) GMPEs as a function of magnitude and R_{JB} distance; the R_{epi}-based predictions are grouped by azimuth

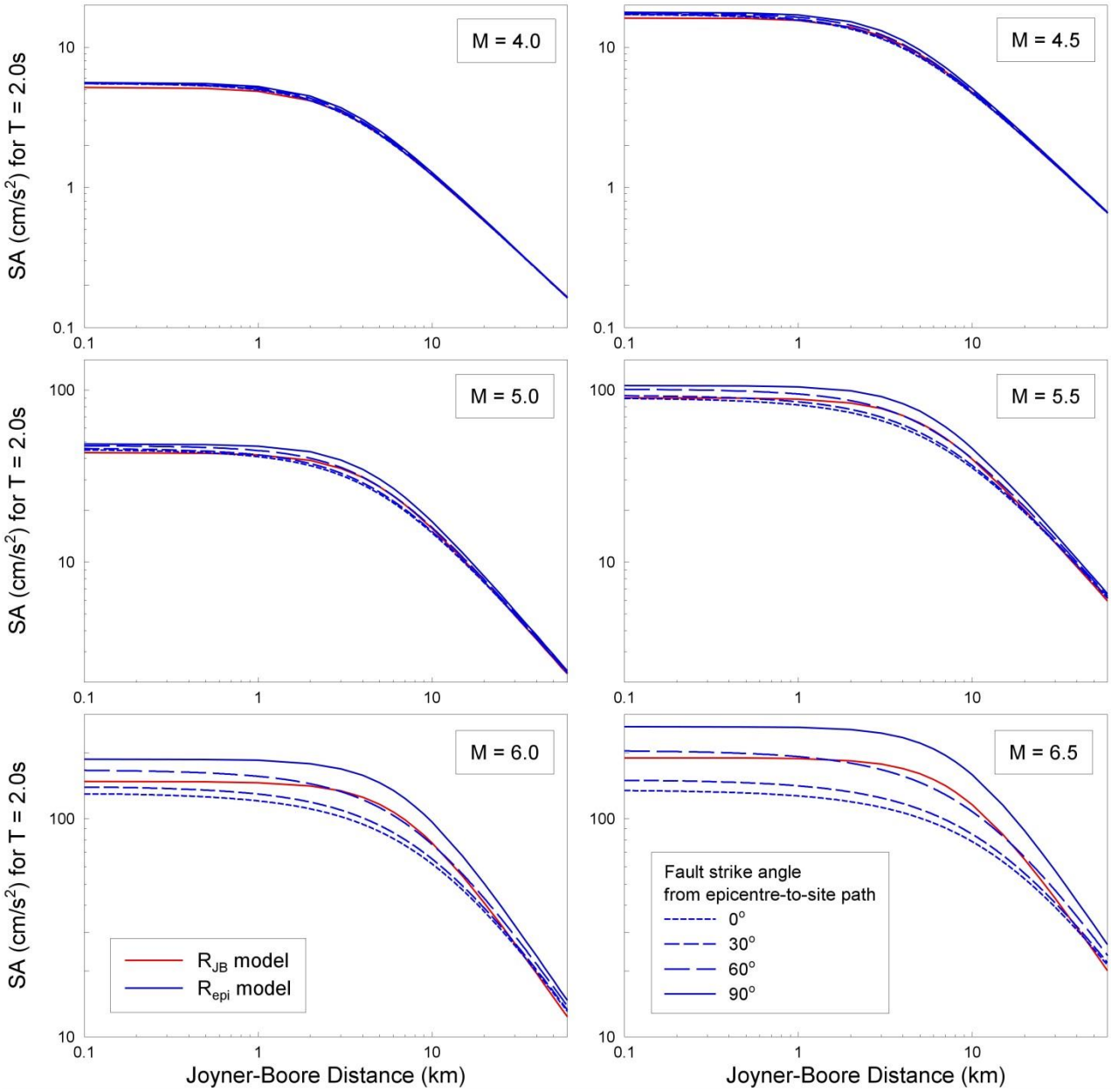


Figure 6.89. Comparisons of median predictions of Sa(2.0s) from the R_{JB} (*red*) and R_{epi} (*blue*) GMPEs as a function of magnitude and R_{JB} distance; the R_{epi}-based predictions are grouped by azimuth

7. PRELIMINARY GMPES for DURATION

The effect of the duration of ground shaking is generally not considered in structural design, other than with regard to its influence on scaling factors to adjust response spectral ordinates to alternative damping ratios (*e.g.*, Stafford *et al*., 2008). However, the influence of duration has been recognised as being important in the seismic assessment of buildings (*e.g.*, Iervolino *et al*., 2006; Chandramohan *et al*., 2015), particularly those constructed from materials such as masonry that is prone to experience degradation of both strength and stiffness under dynamic loading (*e.g.*, Bommer *et al*., 2004). For this reason, the development of the fragility functions for the buildings in the Groningen field aims to take account of the ground-motion duration, for which reason a predictive equation is required for the chosen definition of the ground-motion duration.

7.1. Selection of duration definition

A large number of definitions of the duration of earthquake have been put forward in the technical literature, and application of these definitions to a given accelerogram can produce very different estimates of the duration of strong motion. Bommer & Martínez-Pereira (1999) classified all of the published definitions into three categories:

- Bracketed duration, D_B. This is defined as the interval between the first and last excursions of a specified threshold of acceleration.
- Uniform duration, D_U. This is defined as the sum of the intervals during which the acceleration is above a specified threshold.
- Significant duration, D_S. This is the interval over which some specified portion of the total energy in the record—usually calculated as the integral of the squared acceleration over time—is accumulated.

Each of the definitions can be applied to the actual ground motion or to the response of an oscillator or other structure to the ground motion. The definitions can also be applied with absolute or relative thresholds: for example, the bracketed duration can be defined by excursions of an acceleration level of, say, 0.1*g*, or of a fraction of the PGA.

Since duration has little significance in isolation, its influence is usually coupled with a direct measure of the amplitude of the motion, such as response spectral acceleration. This being the case, it makes more sense to then use definitions defined using relative measures, which will generally result in durations that increase with distance as well as with magnitude. The significant duration is the most widely-used definition, the original interval being that related to the accumulation from 5% to 95% of the total Arias intensity of the record (Trifunac & Brady, 1975). In more recent years, there has been a tendency to move towards using the interval from 5% to 75% of the total energy in the record, which is more likely to isolate the strongest portion of the record that generally corresponds to the shear waves. However, other variations have been proposed, such as the proposal by Boore & Thompson (2014) to use a duration that is equal to twice the interval between

20% and 80% of the total energy being accumulated as a surrogate for the 5-95% definition with improved performance for smaller earthquakes. Work is underway to explore which definition will work best for identifying the strong-motion portion of the recorded motions from the Groningen field considering multiple variations of the bracketed and significant duration definitions with relative thresholds (Figure 7.1).

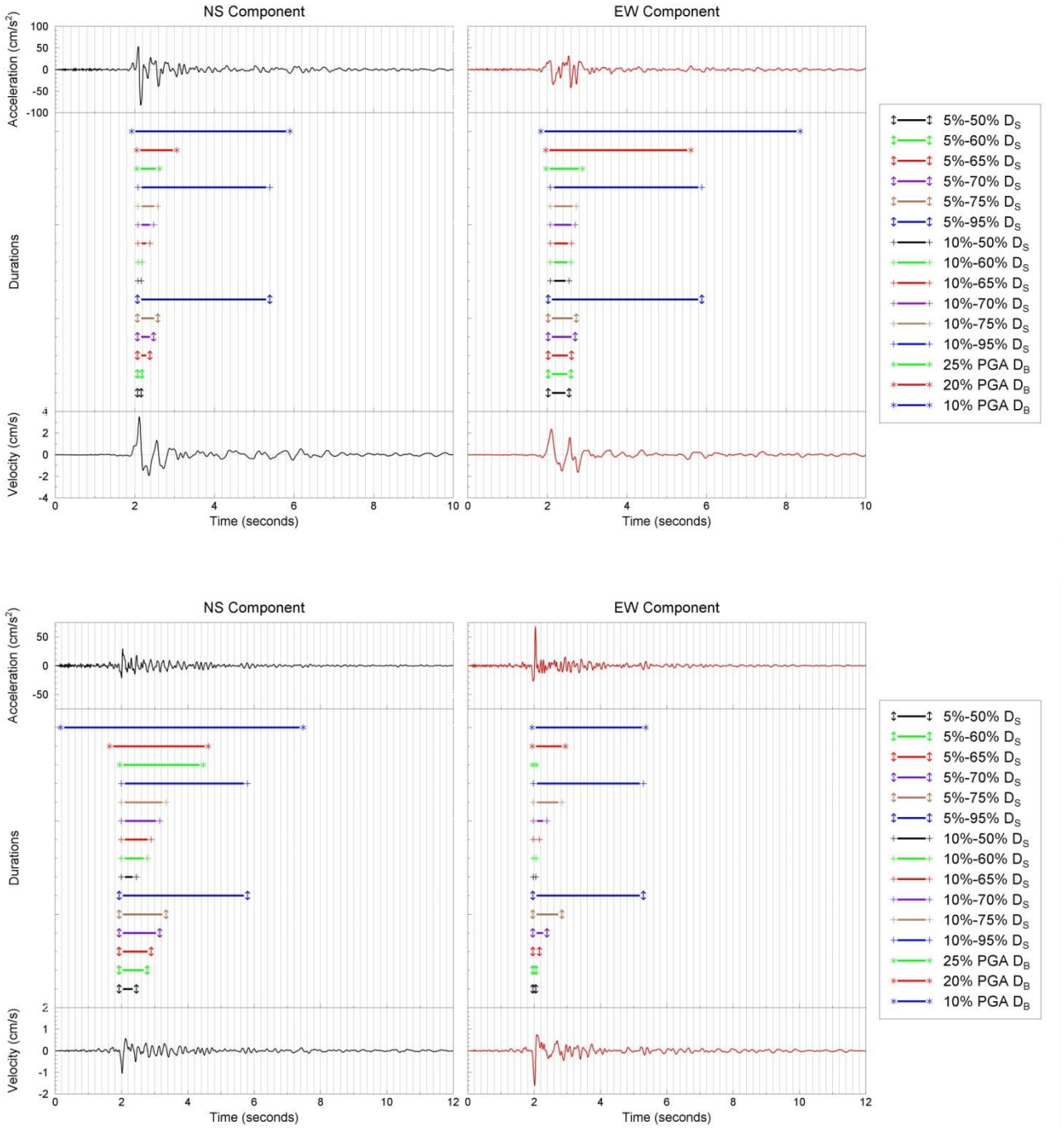


Figure 7.1. Time-series from the 2012 Huizinge earthquake at MID1 (*upper*) and from the 2014 Leermens earthquake at BLOP (*lower*) showing the duration intervals defined by D_B using different portions of the PGA and by D_S with different starting and finishing percentages of the total Arias intensity; for each record, the acceleration signal is shown at the top and the velocity at the bottom

The final choice of duration definition, however, will not only depend on the ability to yield consistent results for the Groningen motions and the generation of stable prediction equations, but also that it is useful—or efficient, in the terminology of earthquake engineering—as a parameter for quantifying structural response. The procedure will be to identify a reduced number of definitions that perform reasonably well from a seismological perspective and pass these to the fragility group to identify those that are the most efficient predictors of structural damage. At this stage, given the time constraints for developing the Version 1 risk model, it has been agreed to use the 5-75% of Arias intensity-based significant duration. This is also convenient since due to its widespread application, there are many tools and models referenced to this duration definition that are available to the team responsible for the derivation fragility functions. This definition is illustrated in Figure 7.2.

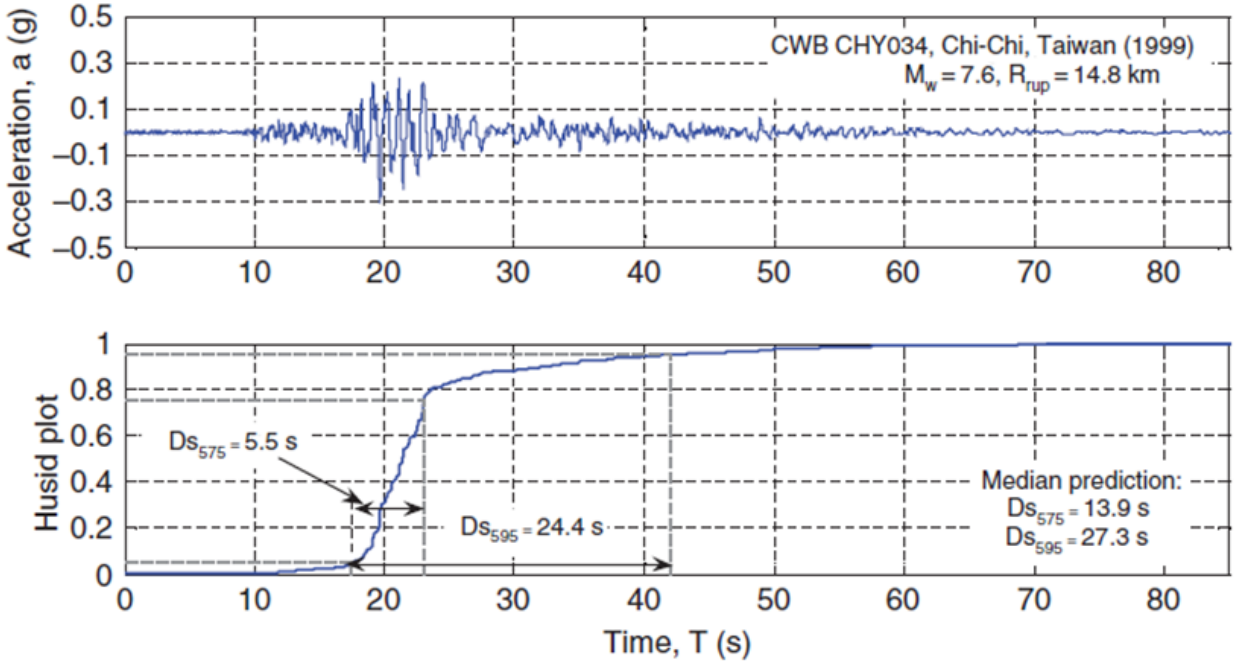


Figure 7.2. Acceleration time-history (*upper*) and Husid plot (*lower*) showing the accumulation of Arias intensity over time and the definition of the 5-75% and 5-95% significant durations (Bradley, 2011)

7.2. Development of GMPEs for duration

As for PGA and the spectral acceleration ordinates, the current Groningen database is clearly insufficient to enable the derivation of GMPEs for duration that could be expected to provide reliable predictions over the full range of magnitudes of interest, namely from **M**2.5 to **M**6.5. In this case, however, unlike for the prediction of spectral accelerations, stochastic simulations would not provide a useful tool for the extrapolation of predictions to larger magnitudes for the simple reason that the simulations actually require an estimate of the ground-motion duration as an input. Therefore, to produce a preliminary model, the most promising approach is to adjust an existing GMPE for $D_{S5\text{-}75}$, derived from recordings of tectonic earthquakes, in a similar fashion to what was done for the Version 0 GMPEs

(Bommer & Dost, 2014). For the prediction of the $D_{S5\text{-}75}$ duration, there are two GMPEs available that are widely applied in practice, those of Kempton & Stewart (2006) and Bommer *et al.* (2009). Both equations are derived from the NGA-West database of recordings from crustal earthquakes (Chiou *et al*., 2008) and give comparable predictions of $D_{S5\text{-}75}$, as shown in Figure 7.3.

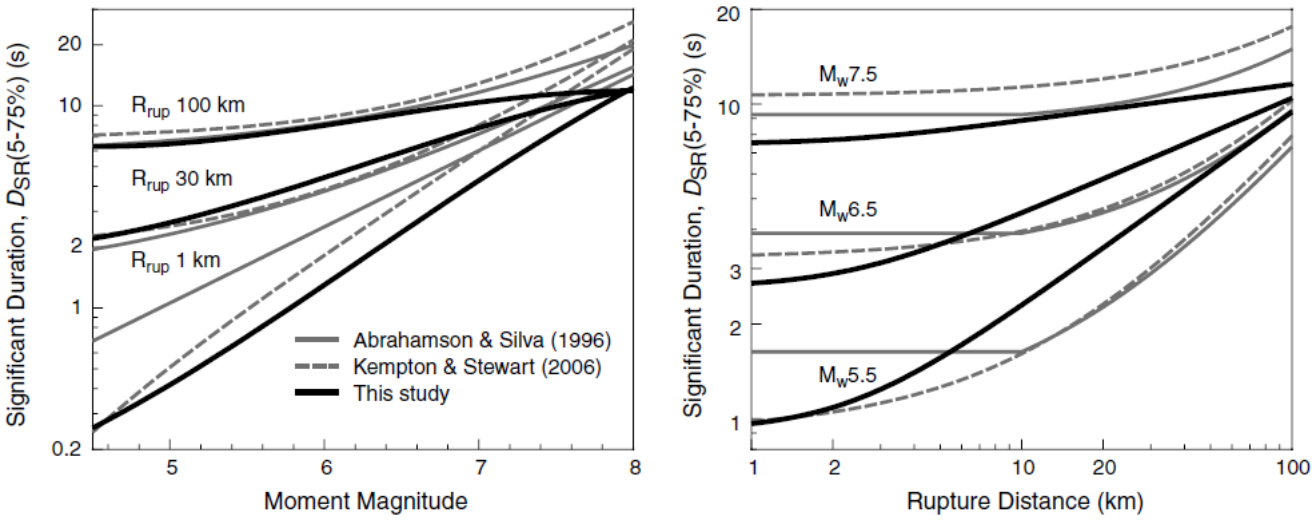


Figure 7.3. Comparison of median predictions of significant duration from the GMPEs of Kempton & Stewart (2006) and Bommer *et al*. (2009), together with another earlier model.

Given that both of these equations are being extrapolated appreciably beyond their ranges of applicability, the possible application—with adjustments as needed—of both equations is considered, not least because both models may provide additional, and different, insights into the nature of the Groningen ground motion durations. We start with the Bommer *et al*. (2009) equation since it has a simpler functional form and may therefore be more amenable to adjustments to fit the Groningen data in the small-magnitude range. The equation for the natural logarithm of $D_{S5\text{-}75}$ includes terms that are linear functions of the logarithm of V_{S30} and of the depth-to-top-of-rupture, Z_{TOR}; setting these parameters to the assumed values of 200 m/s and 3.0 km, respectively, the equation becomes:

$$\ln(D_{S5-75}) = -7.3229 + 1.2619M + (2.0063 - 0.252M)\ln\sqrt{R_{rup}^2 + 2.3316^2} \qquad (7.1)$$

For the small-magnitude earthquakes in the current database, it is reasonable to assume that R_{rup} and R_{hyp} are equivalent, hence the equation can be expressed in terms of R_{epi}:

$$\ln(D_{S5-75}) = -7.3229 + 1.2619M + (2.0063 - 0.252M)\ln\sqrt{R_{epi}^2 + 14.4364} \qquad (7.2)$$

Following the same procedures used in Section 3.3, the between-event and within-event residuals of the Groningen durations—using the geometric mean of the horizontal components were calculated with respect to Eq.(7.2) and are presented in Figure 7.4. The

equation of Bommer *et al*. (2009) was derived using both horizontal components from each record but a correction to sigma value for the component-to-component variability was also provided to enable the sigma for the geometric mean component to be calculated. The first, and most striking, observation that can be made is the very large event terms for all of the earthquakes, reflecting a strong and consistent under-prediction of the durations by the GMPE, by an average factor in excess of four. The residuals do a weak negative trend with magnitude, even though it is questionable if the data are sufficient to distinguish such a pattern. The within-event residuals also show a general tendency to be under-predicted, albeit by much small factors and with no discernable trend. The scatter of the residuals over short distances (< 7 km) is very appreciable.

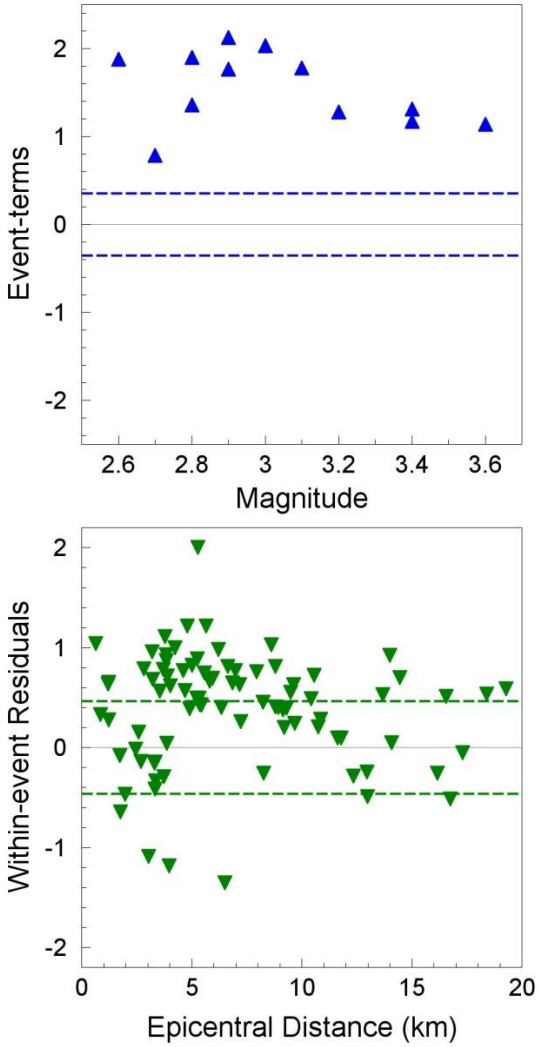


Figure 7.4. Between-event (*upper*) and within-event (*lower*) residuals of the Groningen durations with respect to the GMPE of Bommer *et al*. (2009),

Before making any adjustments to Eq.(7.2) to obtain a better fit to the Groningen data, it is worth briefly considering the distribution of the durations with respect to magnitude and distance. Figure 7.5 shows the durations in 3D space as a function of magnitude and distance, which suggests that the higher values of duration—perhaps not surprisingly—are

associated with records obtained at greater distances. Figure 7.6 shows the geometric mean durations of records obtained at epicentral distances of less than 10 km. There is a no perceptible trend in the data with respect to magnitude at all (although a straight line fit to these data has a negative gradient), which is somewhat surprising even taking into account the fact that the data only span 1.1 units of magnitude.

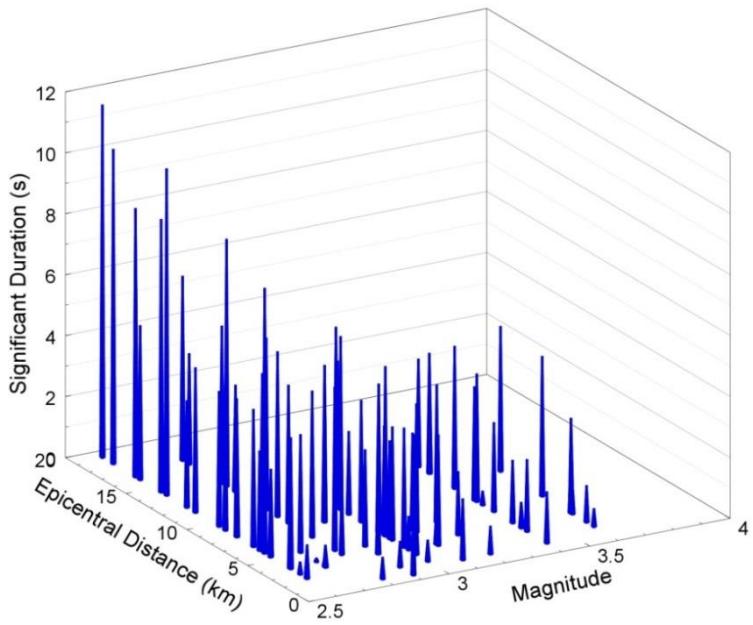


Figure 7.5. Distribution of $D_{S5\text{-}75}$ durations of the Groningen records (geometric means) with respect to magnitude and epicentral distance

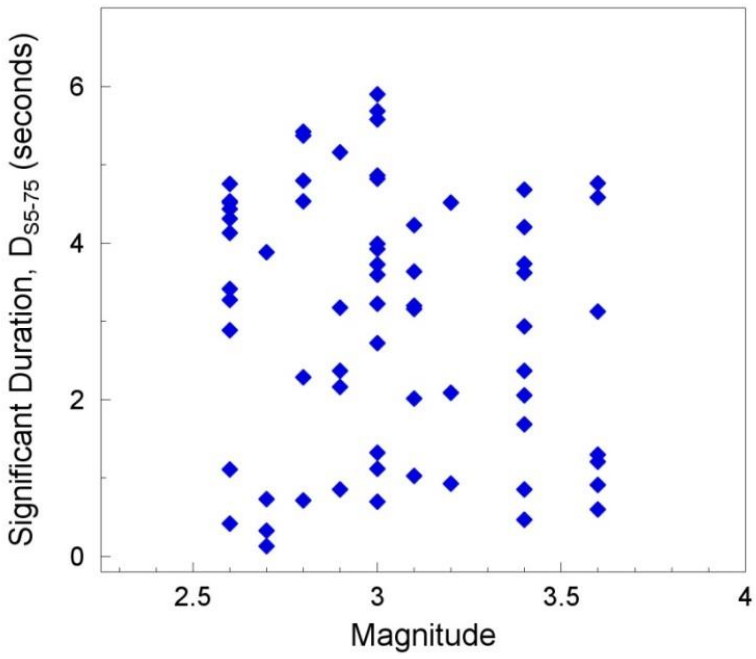


Figure 7.6. $D_{S5\text{-}75}$ durations of the Groningen records obtained within 10 km of the epicentre against earthquake magnitude

Nonetheless, given that the data show both a large positive offset and a mild negative gradient with magnitude with respect to predictions from Eq.(7.2), an adjustment is made to the constant term (+4.7) and the magnitude scaling coefficient (-0.8868) to provide an improved fit to the data; the adjusted equation then is as follows:

$$\ln(D_{S5-75}) = -2.6229 + 0.3751M + (2.0063 - 0.252M)\ln\sqrt{R_{epi}^2 + 14.4364} \qquad M \le 5.3 \qquad (7.3)$$

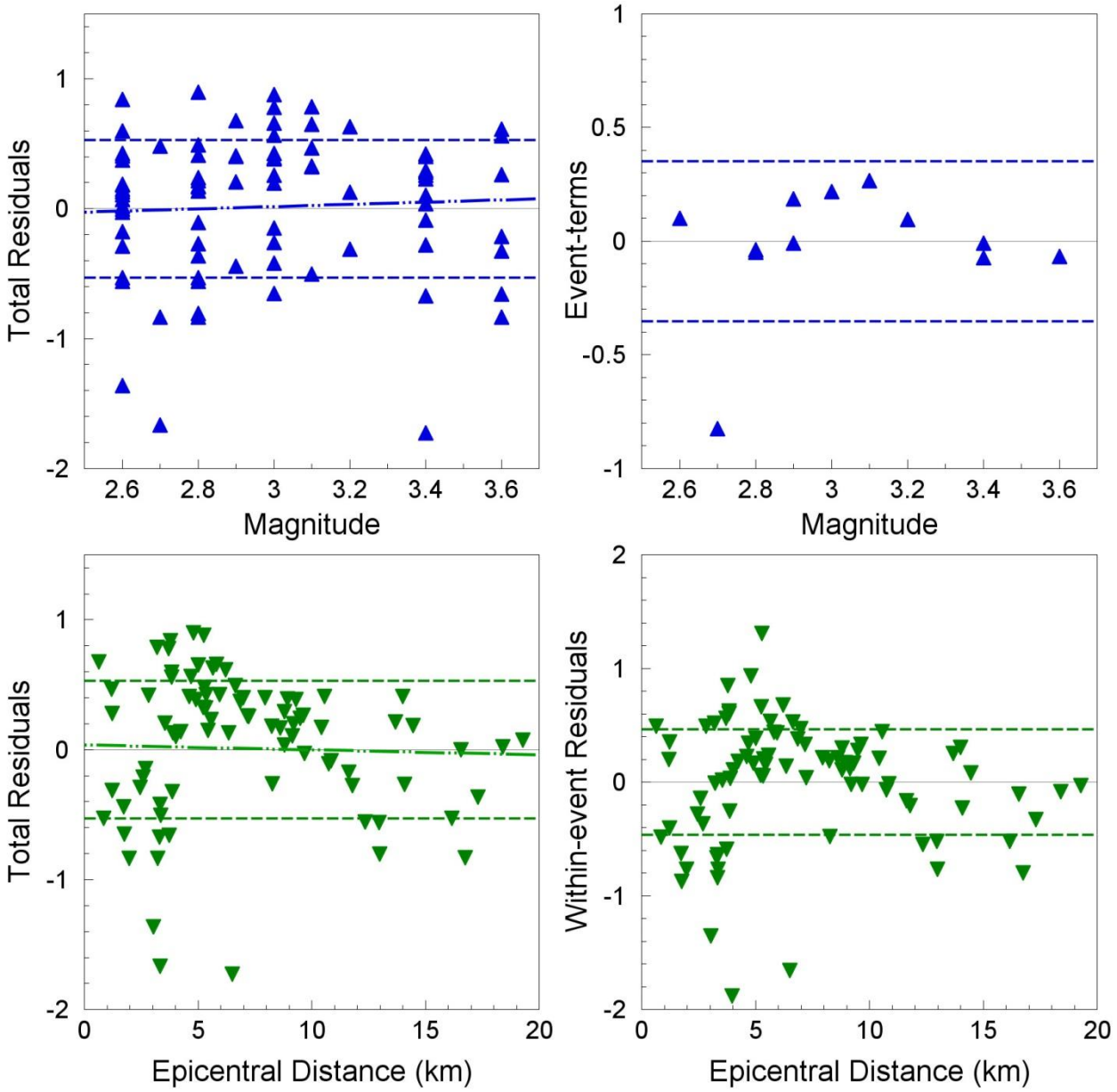


Figure 7.7. *Left*: Total residuals of D_{S5-75} durations of the Groningen records with respect to Eq.(7.3); the lines with short dashes correspond to the total sigma from Bommer *et al*. (2009) and the other dashed line the best fit to the residuals. *Right*: Between-event (*upper*) and within-event (*lower*) residuals with respect to the same GMPE; dashed lines are inter- and intra-event sigmas

For earthquakes of magnitude greater than **M** 5.3, the original equation of Bommer *et al*. (2009)—*i.e*., Eq.(7.2)—should be used. The total residuals of the Groningen data with

respect to Eq.(7.3) are shown in the left-hand plot of Figure 7.7, from which it can be seen that the model provides a reasonable fit in terms of capturing the general trend of the data.

The inter- and intra-event residuals, calculated using the sigma values of the original Bommer *et al*. (2009) GMPE, are shown in the right-hand panel of Figure 7.7. The inter-event residuals show a remarkably good distribution, with the single but very marked exception of the large negative event-term (implying severe over-prediction of the durations) for the **M** 2.7 earthquake of 19th January 2011. This earthquake was recorded by just 4 accelerographs (Figure 7.8) all of which were at relatively short distances from the epicentre, and yet the durations for individual components of the four records vary by a factor of almost 50 (Figure 7.9). This is then reflected in the intra-event residuals in Figure 7.7, although there is no clear linear trend in the data. The smallest and largest individual within-event residuals correspond to records obtained at distances of 4 and 5.3 km from this earthquake. This indicates that the development of a robust model for predicting the durations of ground motions in the Groningen field will be a challenging task; options for addressing this challenge are discussed in Section 8.9.

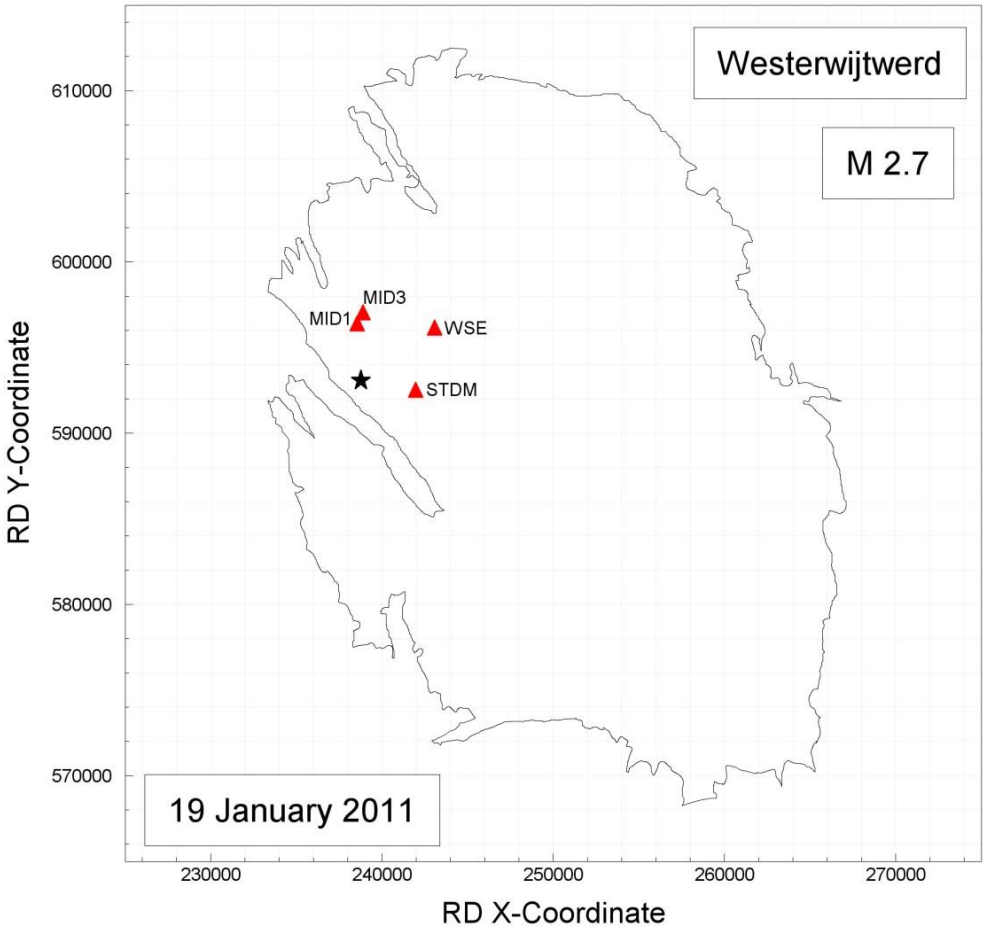


Figure 7.8. Epicentral location and accelerographs recording the 19th January 2011 earthquake.

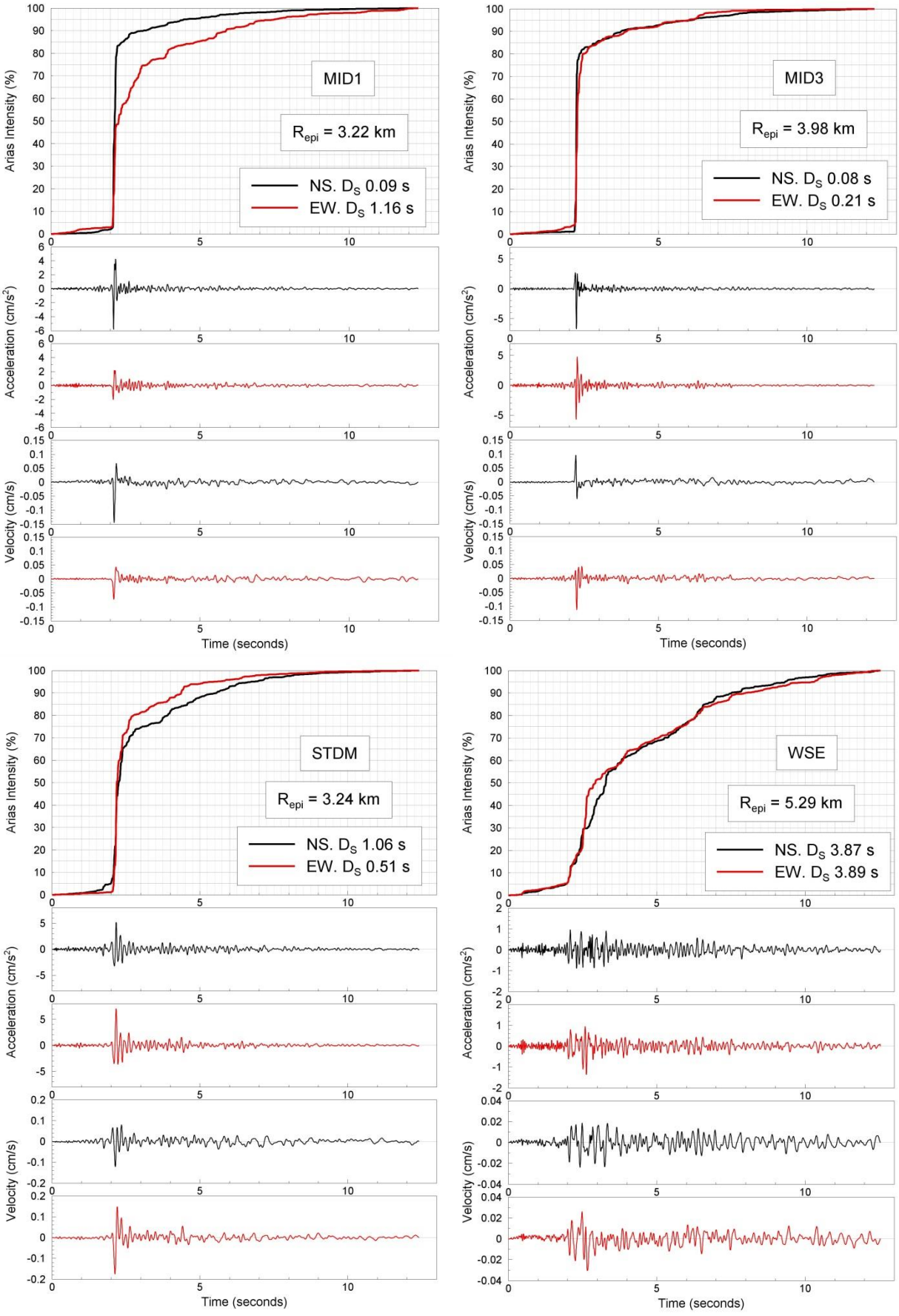


Figure 7.9. Time-series and Husid plot of recordings of the 19th January 2011 earthquake.

Notwithstanding the large unusual residual plots in Figure 7.7, the data do not suggest that the total sigma value associated with the Bommer *et al.* (2009) would not be an appropriate surrogate to adopt for use in conjunction with Eq.(7.3), even if this might be associated with a smaller inter-event component and an apparently larger intra-event variability. However, even though the model does seem to provide a reasonable fit to the Groningen data—notwithstanding the unexplained fluctuations over short distances—the combination with the Bommer *et al*. (2009) GMPE for estimating durations from events of magnitude larger than **M** 5.3 produces a rather strange relationship (Figure 7.10).

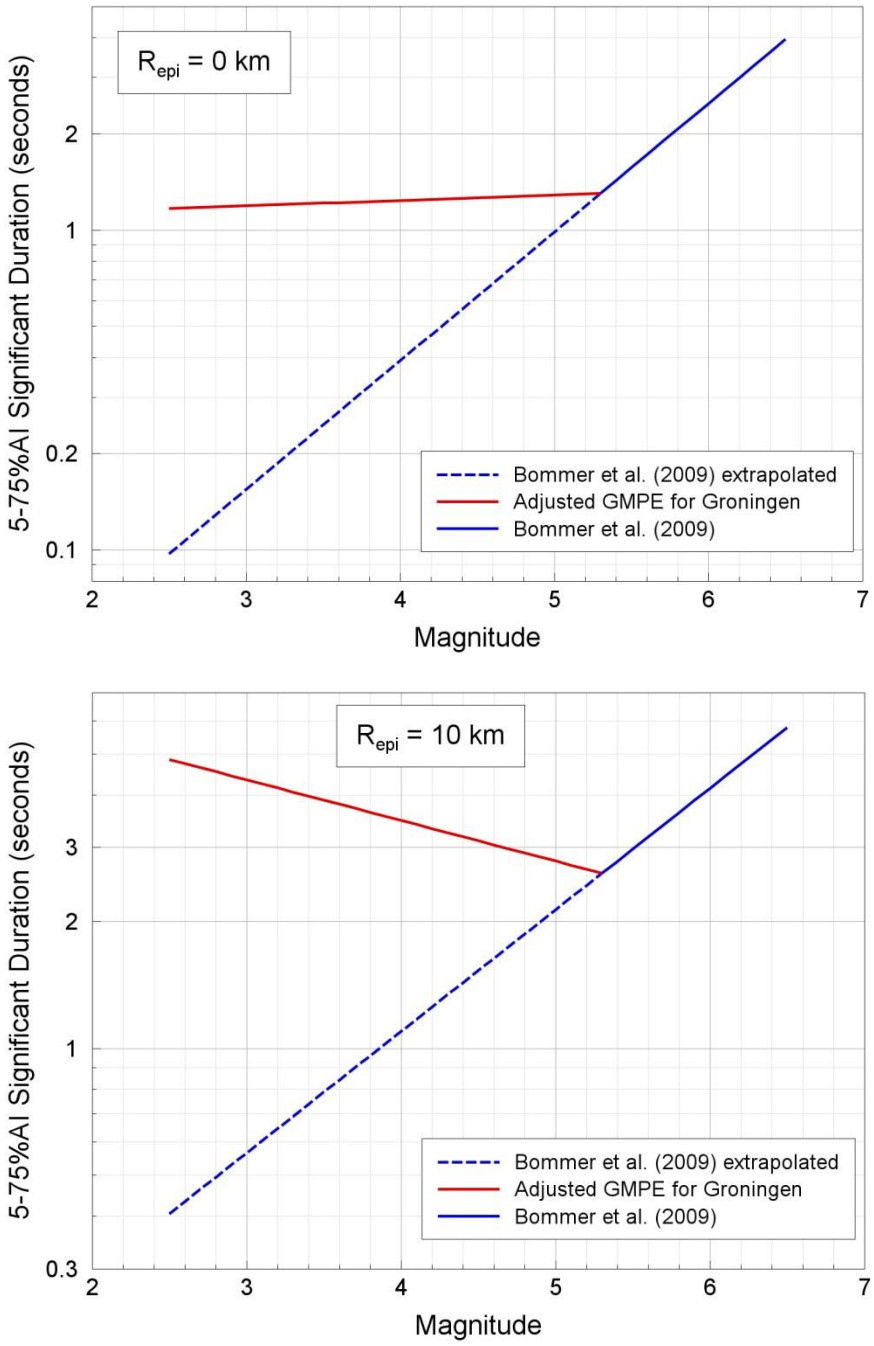


Figure 7.10. Predicted median accelerations from Eq.(7.3) for smaller earthquakes and Eq.(7.2) for events of magnitude **M** 5.3 or greater

The implied variation of durations with magnitude from **M** 4 and above—first decreasing and then increasing as the magnitude becomes larger—is clearly unphysical. Although the durations of the Groningen motions appear to exhibit a genuinely unusual behaviour, which may also raise questions regarding the applicability of standard duration definitions to these very low-amplitude motions, this is the result of trying to make simple adjustments to an existing model in order to match the local data. One needs to bear in mind, of course, the fact that we are applying GMPEs derived from recordings of larger magnitude earthquakes, which is already known to lead to erroneous predictions of accelerations when extrapolated to events smaller than their lower limit of applicability (*e.g.*, Bommer *et al.*, 2007). However, in order to render the adjusted model more physically acceptable, some adjustments are made to prevent the behaviour shown in the lower plot of Figure 7.10. Since the Groningen durations show a rather weak dependence on magnitude (Figure 7.6), we fix the durations for any given epicentral distance at the value corresponding to a magnitude of **M** 3.1, which is the mid-point of the range covered by the Groningen database. The durations are then held constant at this level for magnitudes below the level at which the original Bommer *et al*. (2009) formula, for the same distance, yields the same duration. Setting the magnitude to **M** 3.1 in Eq.(7.2), the corresponding magnitude in Eq.(7.2) can be obtained algebraically by setting the two predicted durations to be equal. The resulting values are shown in Figure 7.11 together with a best-fit linear trend, the coefficients of which have been very slightly adjusted to yield integer results at the limiting values of 0 and 15 km epicentral distances, namely **M** 5.1 and **M** 6.5. The equation of this fit is as follows, with R_{epi} expressed in kilometres:

$$M = 5.1 + 0.0933 R_{epi} \tag{7.4}$$

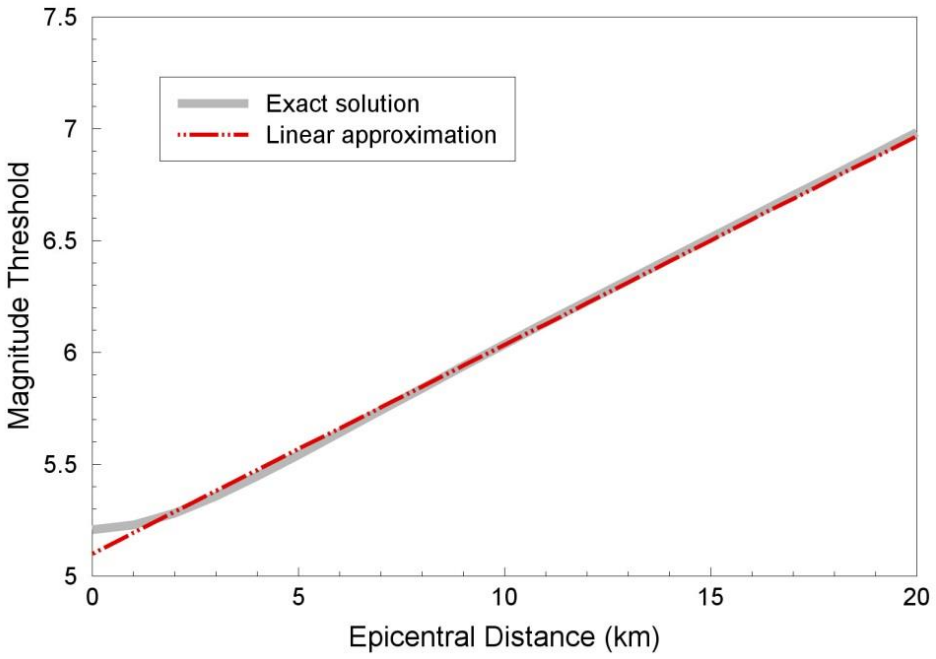


Figure 7.11. Magnitude thresholds, as a function of distance, below which the durations predicted by Eq.(7.2) is held constant, showing the exact solution and the linear approximation of Eq.(7.4)

Figure 7.12 compares the durations predicted by this model with the Groningen data, from which it can be seen that the fit is very reasonable. The model predictions are illustrated in Figure 7.13.

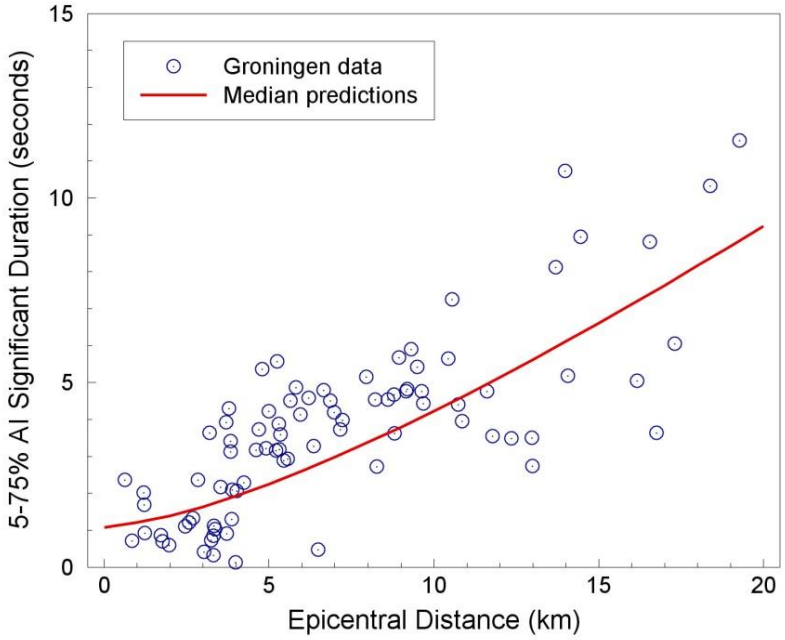


Figure 7.12. Comparison of observed durations of Groningen motions with median predictions from Bommer *et al*. (2009) using the threshold magnitude at each distance predicted by Eq.(7.4)

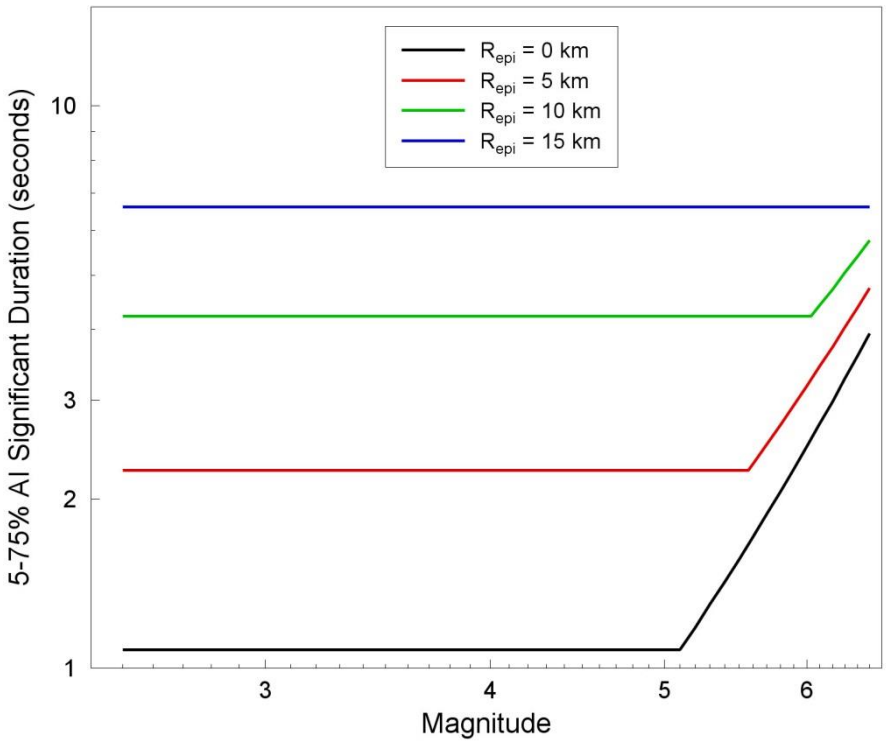


Figure 7.13. Median predictions of duration from the adjusted version of Bommer *et al*. (2009)

Although this provides a model that could be deployed in the hazard and risk calculations, the implied independence of duration on magnitude for earthquakes of **M** 5 and smaller is not easily defensible. Moreover, a decision would need to be made as to how to model the durations at epicentral distances beyond 15 km, since the model would imply that even for the small-to-moderate magnitudes currently observed in the field these durations would correspond to the expectations from tectonic earthquakes on the order of **M** 7 or even larger.

In view of the rather unsatisfactory results from modifying the Bommer *et al.* (2009) model to be consistent with the Groningen data at small magnitudes, the next stage was to explore the applicability of the equation of Kempton & Stewart (2006). This model is somewhat more complicated, as noted previously, but for the case of the 5-75%AI significant duration, the equation for median predictions can be reduced to:

$$\ln(D_S) = \ln\left[\frac{\left(\frac{\exp(6.02)}{10^{1.5M+16.05}}\right)^{\frac{-1}{3}}}{1.568x10^{7}} + 0.07\sqrt{R_{epi}^2+9} + 0.82 - 0.0013(V_{S30})\right] \quad (7.5)$$

The equation actually uses the R_{rup} distance metric, which herein is assumed equivalent to R_{hyp} for small-magnitude earthquakes; V_{S30} is expressed in m/s. A second version of the model is presented by Kempton & Stewart (2006) to include the effects of deep soil basins, which are characterised by the depth (in metres) to the horizon with a shear-wave velocity of 1.5 km/s, $Z_{1.5}$:

$$\ln(D_S) = \ln\left[\frac{\left(\frac{\exp(6.02)}{10^{1.5M+16.05}}\right)^{\frac{-1}{3}}}{1.568x10^{7}} + 0.07\sqrt{R_{epi}^2+9} + 2.73 - 0.0013(V_{S30}) - 0.00075(Z_{1.5})\right] \quad (7.6)$$

Kempton & Stewart (2006) provide two alternative basin-effects models, one for which the earthquake source is located within the basin and another for which it is remote from the basin where the site is located; the former has been adopted in this case. Kempton & Stewart (2006) also provide a third model that does not make this distinction, and they actually recommend that one for general use but in the case of Groningen, the source-site separation would suggest that the so-called CBL (coincident source and site basin locations) model is appropriate.

As before, the value of 200 m/s is assumed for V_{S30} and, based on information provided by Remco Romijn of NAM, an initial estimate of $Z_{1.5}$ is 1,200 m. Using these values, the residuals of the Groningen durations are calculated using both Eq.(7.5) and Eq.(7.6), and the results are shown in Figure 7.14. Although there is a large scatter of the within-event residuals, the event terms obtained with the model that does consider basin events

indicate a rather good fit to the data, and certainly much better than that obtained with the GMPE of Bommer *et al*. (2009), as can be appreciated by comparison with Figure 7.4. Although the Groningen field is overlain by thick layers of soft deposits, it is not necessarily only basin effects contributing to the extended durations, which are currently interpreted as being in large part due to the refractions and reflections of upcoming seismic waves in the high-velocity Zechstein formation above the gas reservoir. With this being the case, the basin-effect model is effectively being used as a surrogate in this case, in which case the optimal model may not be that having a basin depth of 1,200 m. In order to explore alternative options, the residuals were re-calculated with basin depths of 600 m and 1,800 m (Figure 7.15). From these plots, it would be appear that the value of 600 m produces the best fit to the data, although there is still a very appreciable scatter of the within-event residuals and a marked over-prediction of the durations at short distances (R_{epi} < 5 km).

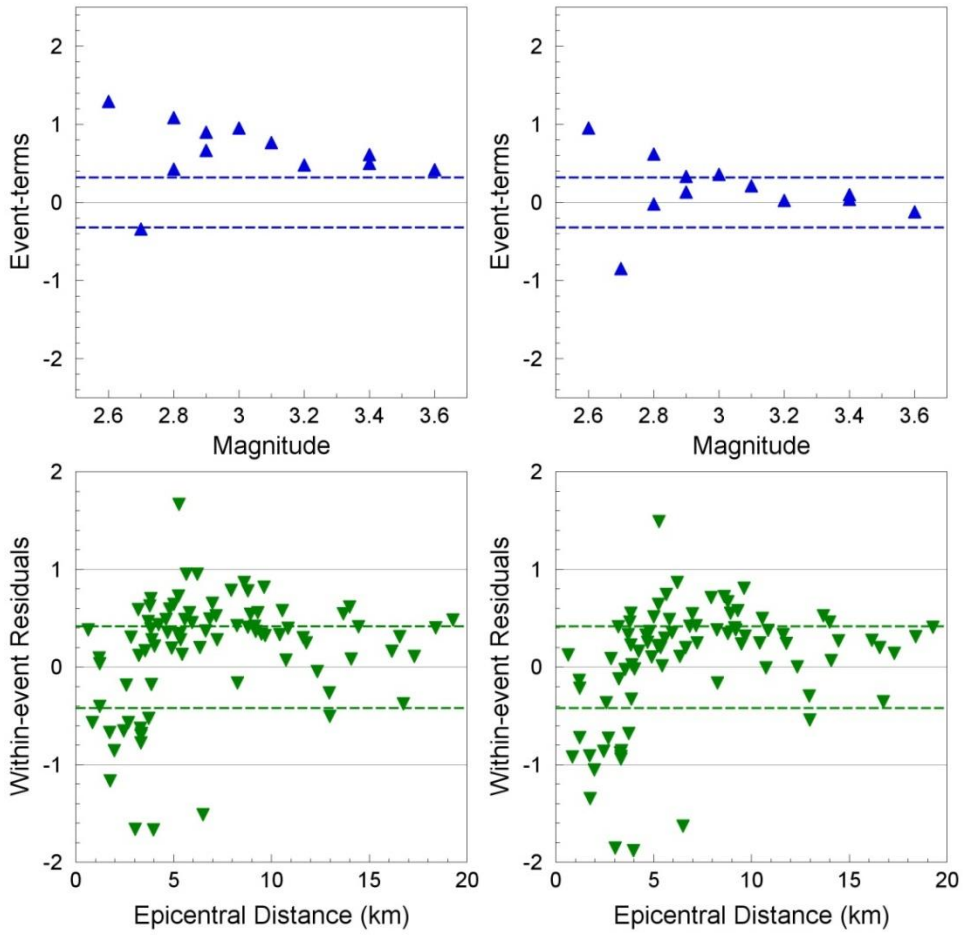


Figure 7.14. Between- and within-event residuals of the Groningen durations with respect to the GMPE of Kempton & Stewart (2006) without (*left*) and with (*right*) basin effects, assuming a basin depth of 1,200 m

Comparison of the left-hand frame in Figure 7.15 with the right-hand frame of Figure 7.7 suggests that between the adjusted Bommer *et al*. (2009) GMPE and the Kempton & Stewart (2006) GMPE applied with a basin depth of 600 m, there is relatively little to

choose between them. The Kempton & Stewart (2006) model, however, has the advantage of providing a more physically acceptable performance over the full magnitude range, as confirmed by Figure 7.16, which compares the predictions from this model with those shown previously in Figure 7.13 using the modified Bommer *et al*. (2009) equation. This plot shows that in fact the overall shapes of the curves are not so different and both indicate weak—in one case, null, by design—scaling with magnitude in the small-magnitude range. The more striking feature of the Kempton & Stewart (2006) model is the much weaker scaling with distance, which is also borne out by the comparison made in Figure 7.17. This figure repeats the comparison of median predictions of duration with the Groningen data shown in Figure 7.12, but now adding in the predictions from the Kempton & Stewart (2006) model as well. Since the latter model still has magnitude-dependence, rather than using the central value of **M** 3.1, the curves are plotted for the bounding magnitude values of 2.6 and 3.6. The impact of the unit change of magnitude is effectively negligible and the variation of duration with distance modelled by the Kempton & Stewart (2006) is rather weak, leading to consistent over-estimation of the duration at short distances (R_{epi} < 4 km) and consistent under-prediction for R_{epi} > 8 km.

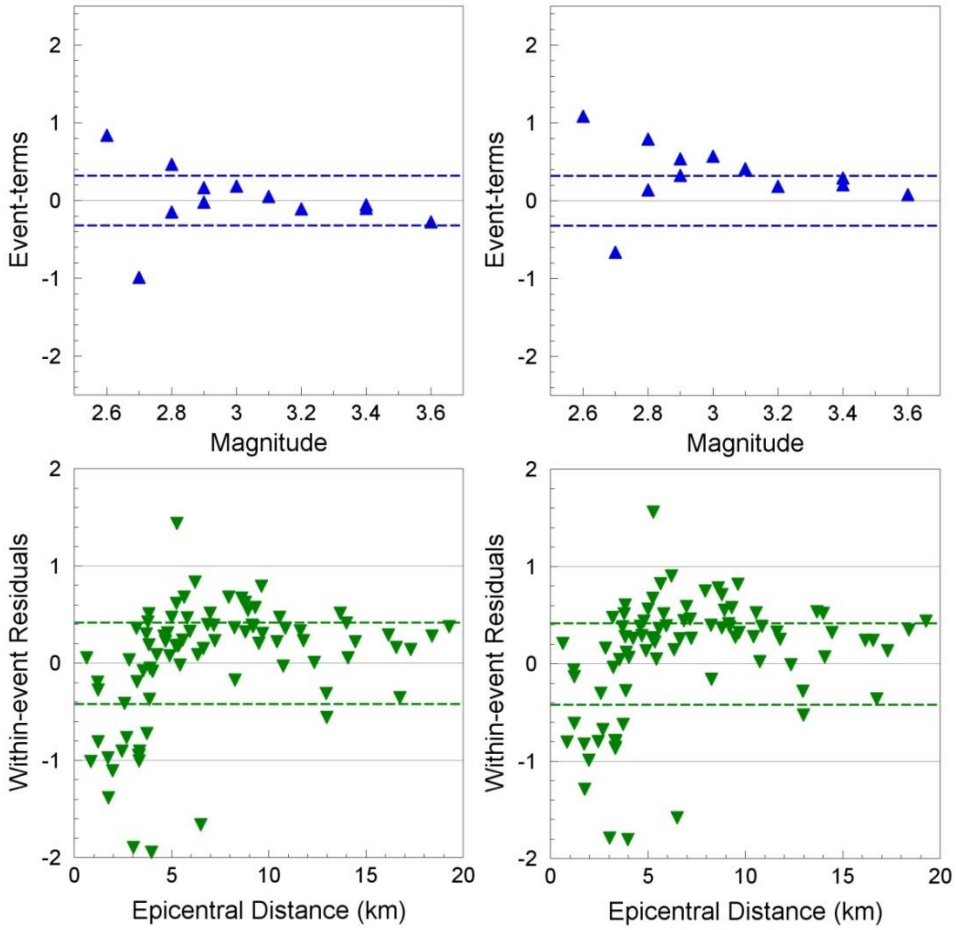


Figure 7.15. Between- and within-event residuals of the Groningen durations with respect to the GMPE of Kempton & Stewart (2006) assuming basin depths of 600 m (*left*) and 1,800 m (*right*)

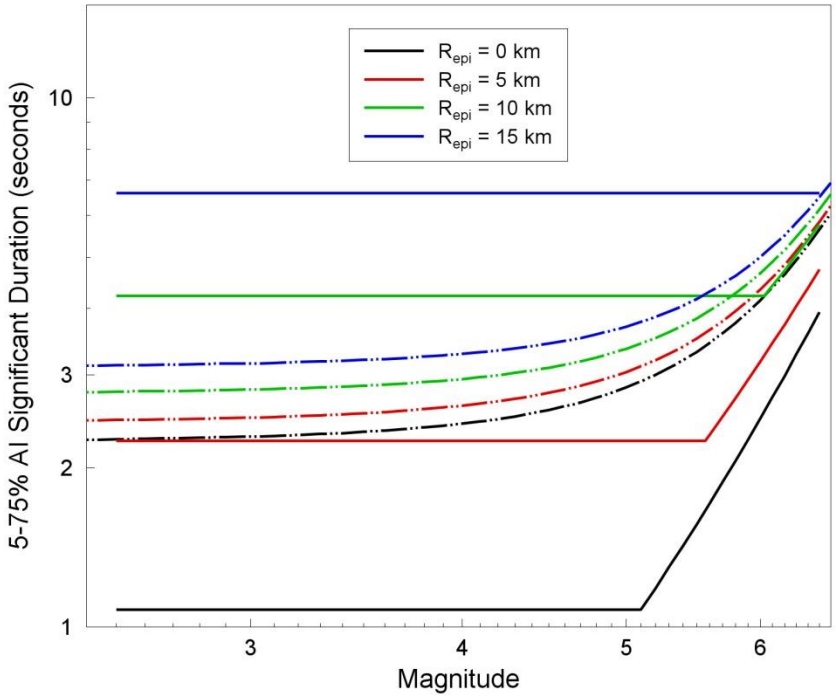


Figure 7.16. Median predictions of duration from the adjusted version of Bommer *et al*. (2009) and from the Kempton & Stewart (2006) model with a basin depth of 600 m

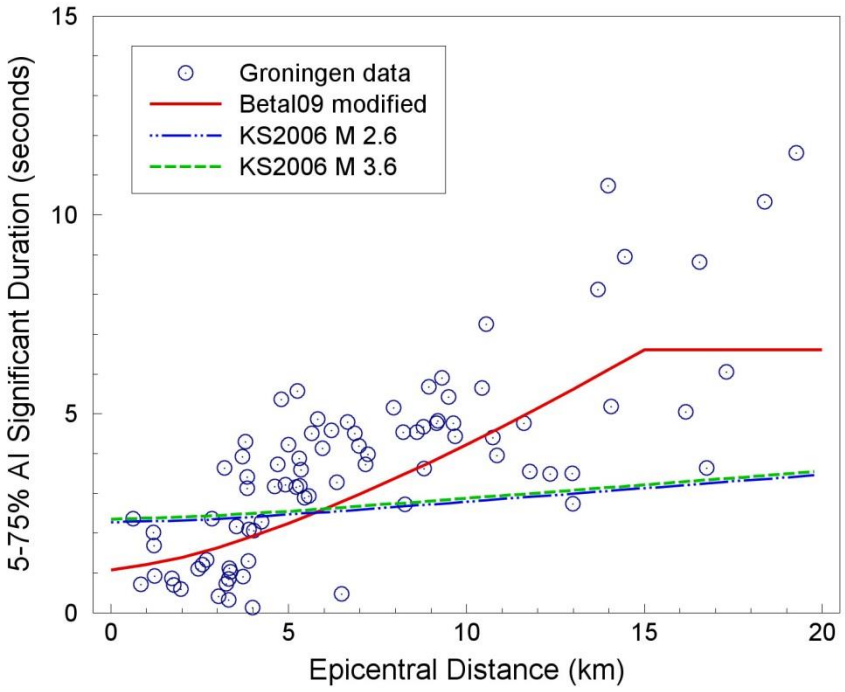


Figure 7.17. Comparison of observed durations of Groningen motions with median predictions from Bommer *et al*. (2009) using the threshold magnitude at each distance predicted by Eq.(7.4)—with the distance saturation effect illustrated in Figure 7.16— and the predictions from Kempton & Stewart (2006) for **M** 2.6 and **M** 3.6.

Finally, we have two potentially applicable models, each of which would appear to have relative strengths and weaknesses. The choice of which of the two is more appropriate to

use in conjunction with the Version 1 GMPEs for response spectral ordinates is not unambiguously clear and would depend to some extent on the specific requirements of the application of the equation in the fragility function derivation. For now, both models are kept in contention, with the final choice being discussed in Section 7.4.

For completeness, the final version of the Kempton & Stewart (2006) equation with the adopted parameter values can be expressed as follows:

$$\ln(D_S) = \ln\left[\frac{\left(\frac{\exp(6.02)}{10^{1.5M+16.05}}\right)^{\frac{-1}{3}}}{1.568x10^{7}} + 0.07\sqrt{R_{epi}^2 + 9} + 2.02\right] \qquad (7.7)$$

Once again, the assumption is made that the variability in the original equation can be adopted for its small-magnitude application as well. The sigma values for the two models are recorded in Table 7.1.

Table 7.1. Sigma values associated with duration prediction equations

Sigma	Eq.(7.3) Bommer *et al.*	Eq.(7.7) Kempton & Stewart
τ	0.3527	0.32
ϕ	0.4304	0.42
σ	0.5289	0.53

7.3. Correlation of residuals of duration and accelerations

An important aspect of including duration in the risk assessment is to account for the inverse correlation between the duration of the shaking and the amplitude of the acceleration. As Figure 7.18 shows very clearly, for the 85 records in the current Groningen database, there is a very clear pattern that components associated with higher peaks of acceleration (> 50 cm/s^2) are associated with durations of 1 second or shorter, whereas all the records with longer durations (longer than, say, 6 seconds) are associated with extremely low amplitudes (< 5 cm/s^2). The lower plot in Figure 7.18 shows that the former type of records (high PGA, short duration) are all obtained at epicentral distances of less than 4 km, whereas the long-duration, low-amplitude records are from recording stations at epicentral distances of at least 8 km. These observations are consistent with the current hypothesis that waves leaving the reservoir at take-pffangles that are not nearly vertical are subjected to multiple refractions and reflections by the high-velocity Zechstein salt formation overlying the reservoir—and the even higher velocity anhydrite layers within the Zechstein—leading to signals outside the epicentral area being dominated by multiple indirect arrivals.

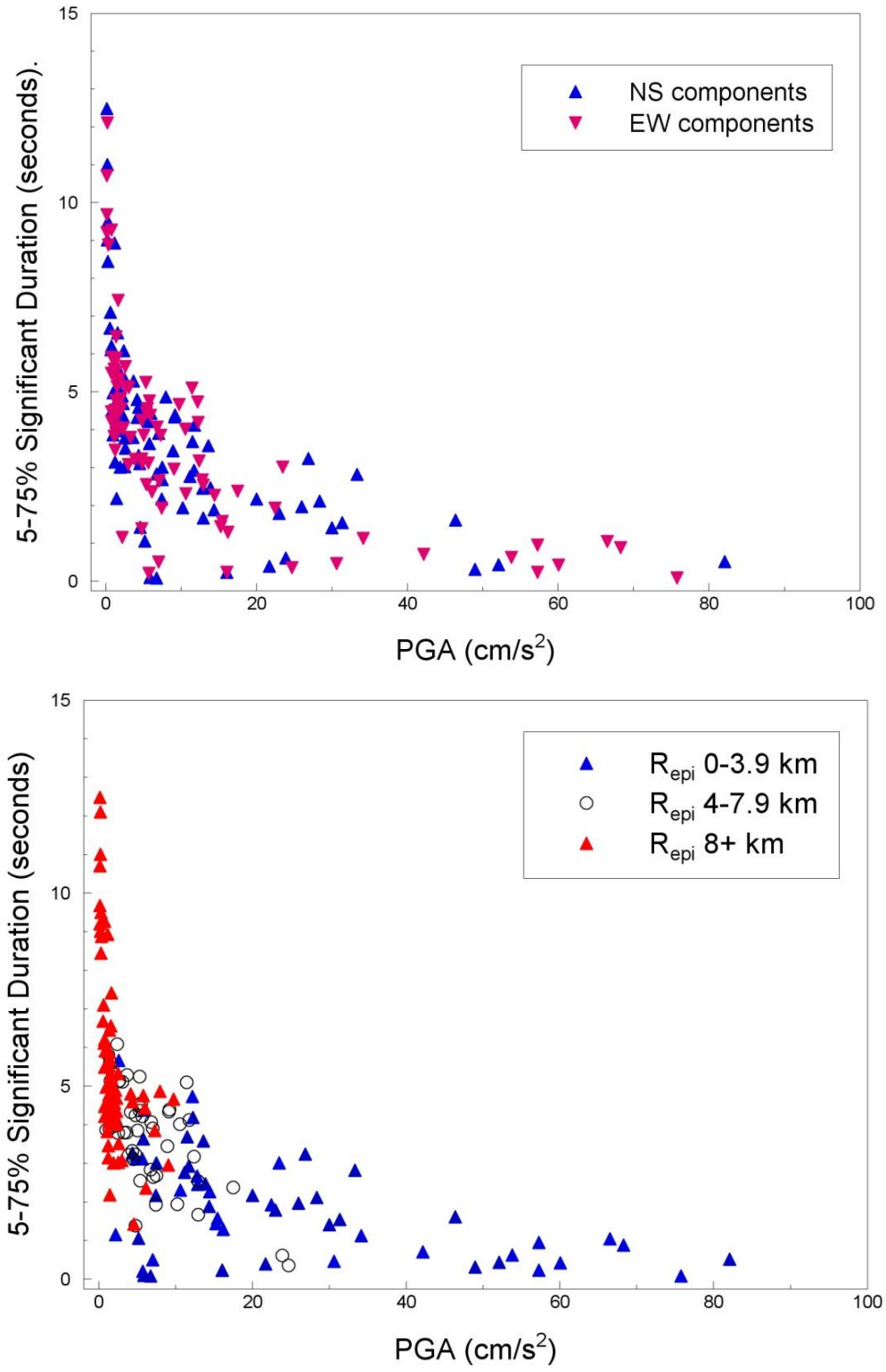


Figure 7.18. Relationship of D_{S5-75} and PGA values for individual horizontal components of motion of the Groningen records (*upper*) and the same data classified by distance (*lower*)

For modelling purposes, however, what is needed is not the general correlation of duration and acceleration, but rather the correlation of the residuals with respect to the predicted median values of the two parameters. In other words, what is required is the correlation coefficient of the residuals of duration and acceleration from each record, which will

indicate the expectation regarding the duration when the acceleration is exceptionally high or low with respect to the median prediction. If, for example, the correlation is strongly negative, then it means that a high PGA value, resulting from a large positive epsilon (ε) value is likely to be associated with a negative epsilon on duration: the PGA value will be higher than the median prediction and the duration shorter than the median prediction.

In order to explore the correlations between predicted accelerations and predicted durations, rather than use the final models—which in most cases represent some degree of compromise in terms of how exactly they explain the current Groningen database—we select the best model for each parameter. For the spectral accelerations, the preferred model is the empirical regression on Eq.(5.9) and the corresponding coefficients in Table 5.1. For the durations, there are two models, the modification of the Bommer *et al*. (2009) GMPE, presented as Eq.(7.3), and the specific deployment of the Kempton & Stewart (2006) GMPE as presented in Eq.(7.7). The analyses of correlations is performed using the individual components—since the records are observed to often be highly polarised—even though the equations are all derived for the geometric mean component, something not explicitly stated but assumed for the case of Kempton & Stewart (2006). For the Bommer *et al*. (2009) GMPE, component-to-component variability was calculated that enables the random component variability to be computed, but by using the geometric mean sigmas in all three cases, the calculations are internally consistent. The procedure followed is very simple: for each oscillator period, the predicted median values of spectral acceleration and duration are calculated for the magnitude-distance combination of the recording, then the residual calculated for each component and normalised by the total sigma associated with each equation. The 170 pairs of normalised residuals for each case are shown in Figures 7.19 to 7.23.

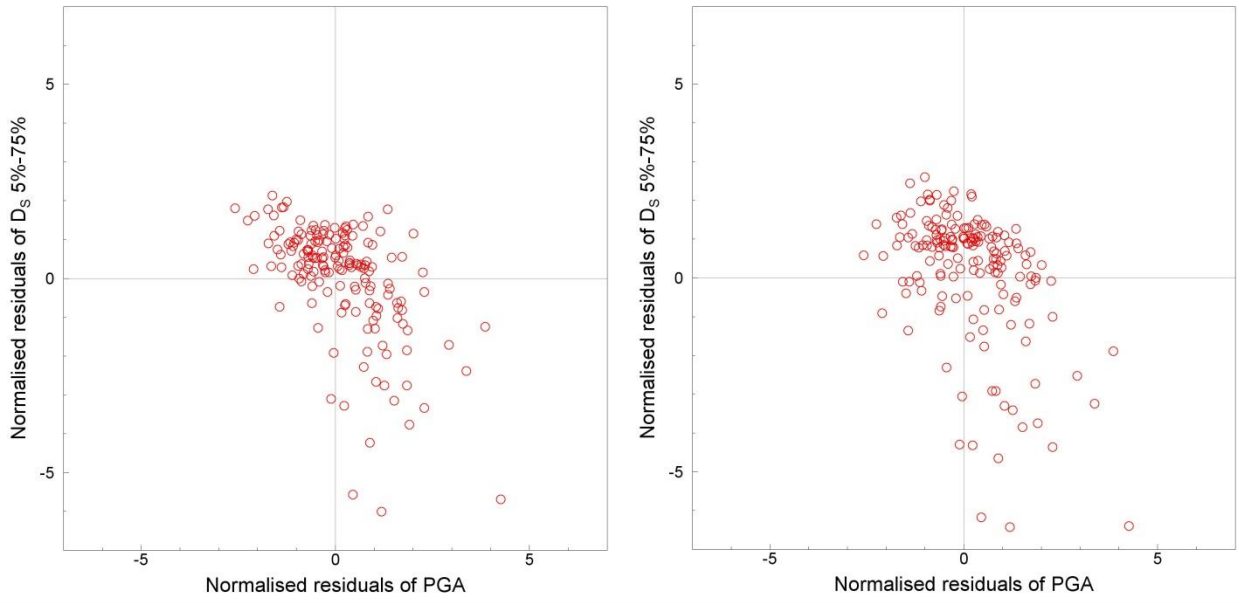


Figure 7.19. Normalised residuals of $D_{S5\text{-}75}$ and PGA values for individual horizontal components of motion of the Groningen records with the durations calculated using the modified Bommer *et al*. GMPE (*left*) and the Kempton & Stewart GMPE (*right*).

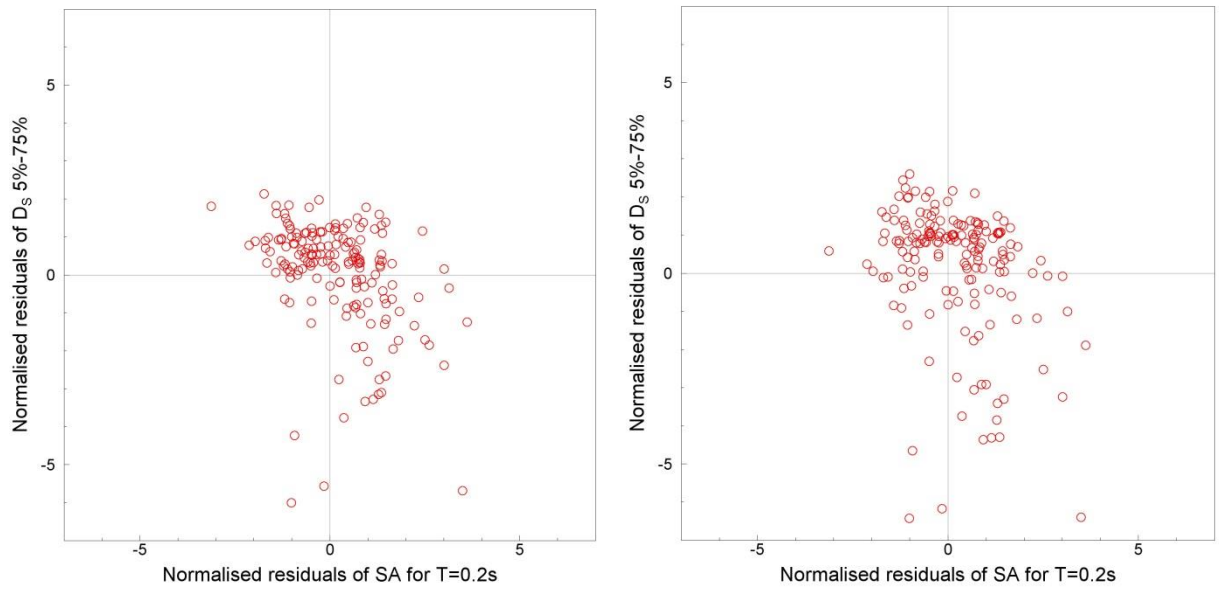


Figure 7.20. Normalised residuals of D_{S5-75} and Sa(0.2s) values for individual horizontal components of motion of the Groningen records with the durations calculated using the modified Bommer *et al*. GMPE (*left*) and the Kempton & Stewart GMPE (*right*).

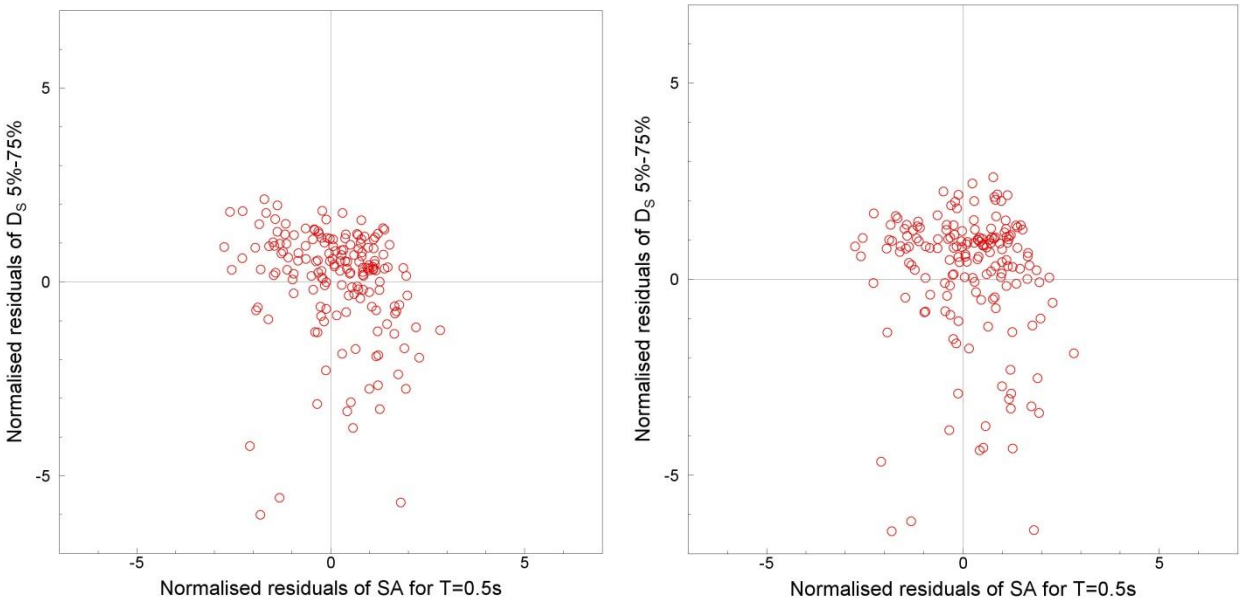


Figure 7.21. Normalised residuals of D_{S5-75} and Sa(0.5s) values for individual horizontal components of motion of the Groningen records with the durations calculated using the modified Bommer *et al*. GMPE (*left*) and the Kempton & Stewart GMPE (*right*).

The plots, for both duration prediction equations, show strong and consistent patterns of negative correlation although the strength of the correlation reduces with increasing response period. For PGA, there is a fairly consistent pattern of positive residuals of acceleration (under-predicted PGA) being associated with very strongly negative residuals of duration (grossly over-predicted values), consistent with observations that the few values of high PGA in the database are generally associated with records of extremely

short duration. Similarly, even though the maximum residuals have smaller absolute values, positive duration residuals (records of longer-than-expected duration) are often—although by no means always—associated with lower-than-average accelerations. From a risk assessment perspective, however, the clear pattern of high PGA values being associated with very short durations is the key observation and most important finding.

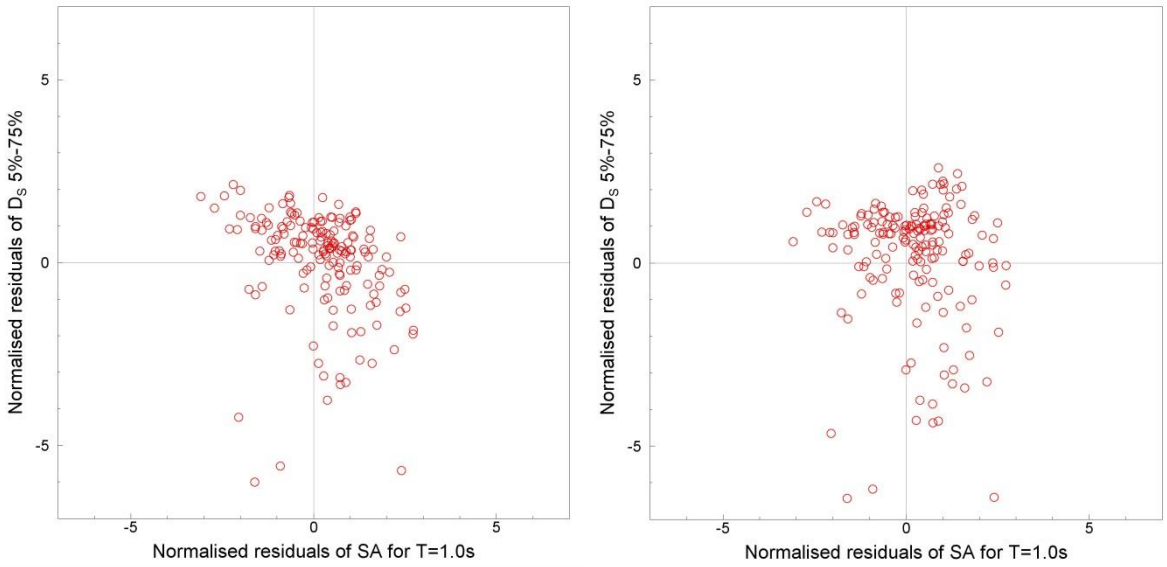


Figure 7.22. Normalised residuals of $D_{S5\text{-}75}$ and Sa(1.0s) values for individual horizontal components of motion of the Groningen records with the durations calculated using the modified Bommer *et al*. GMPE (*left*) and the Kempton & Stewart GMPE (*right*).

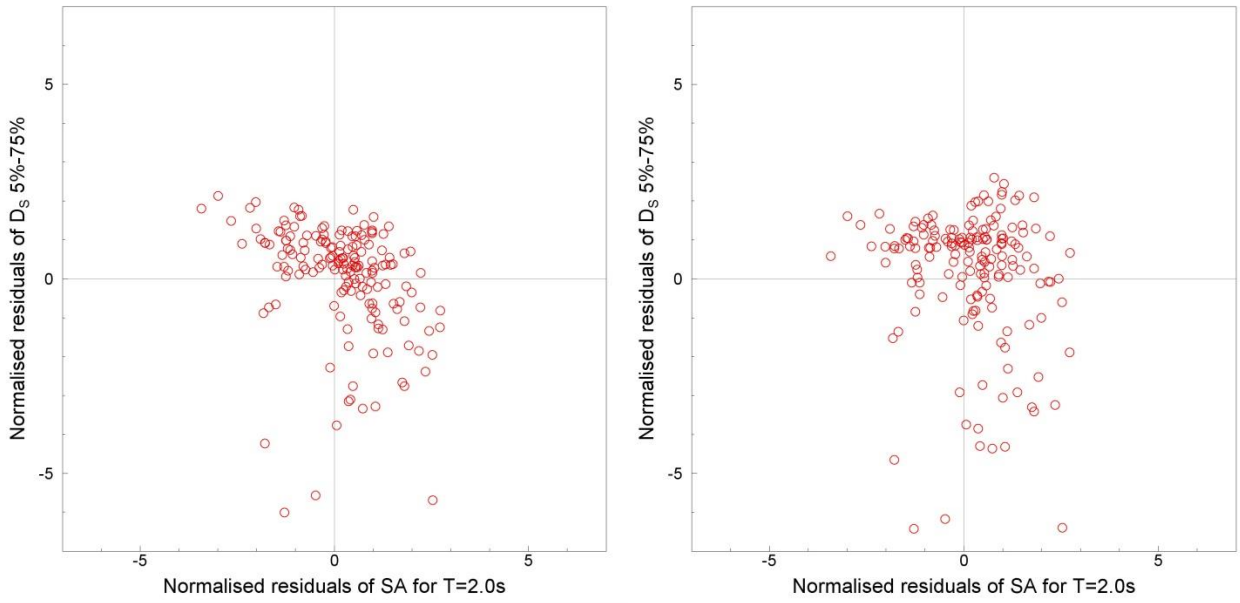


Figure 7.23. Normalised residuals of $D_{S5\text{-}75}$ and Sa(2.0s) values for individual horizontal components of motion of the Groningen records with the durations calculated using the modified Bommer *et al*. GMPE (*left*) and the Kempton & Stewart GMPE (*right*).

For all 10 pairs of residuals (5 spectral accelerations and two alternative duration GMPEs) the Pearson's sample correlation coefficient is calculated; these are reported in Table 7.2 and plotted in Figure 7.24 together with the model for correlation between spectral accelerations and D_{S5-75} of Bradley (2011). Neither duration model produces correlations that closely match the model of Bradley (2011) across the period range, although the Kempton & Stewart (2006) model results in broader agreement than the modified Bommer *et al*. (2009) equation. However, given that the study by Bradley (2011) did not consider such small magnitudes as those covered by the current Groningen database, close agreement may not be expected. The modified Bommer *et al*. (2009) model produces consistently smaller correlation coefficients, implying a stronger negative correlation between the two parameters.

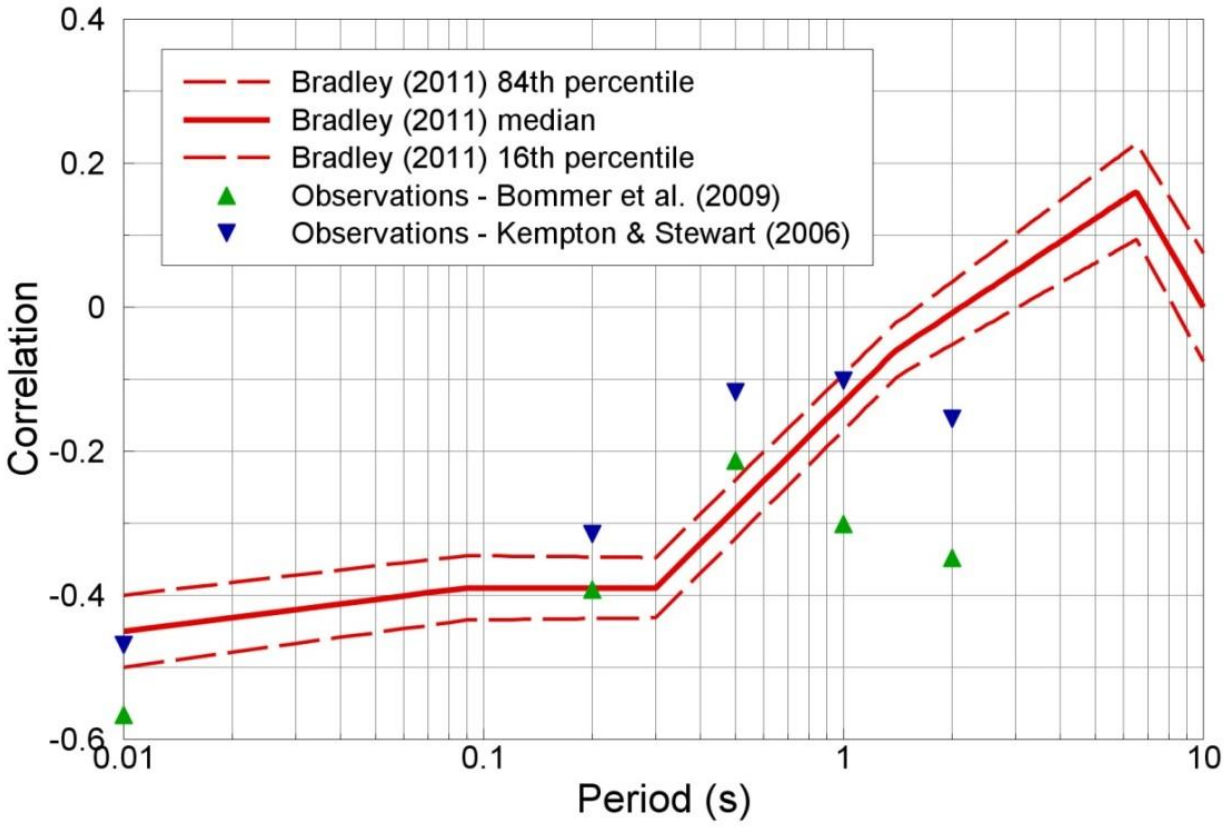


Figure 7.24. Correlations coefficients between predicted value of D_{S5-75} and Sa(T) calculated for the Groningen database using the two alternative duration GMPEs and compared with the model of Bradley (2011)

Table 7.2. Correlation coefficients of the predictions of Sa and D_S obtained using the two alternative duration prediction models

Model	PGA	Sa(0.2s)	Sa(0.5s)	Sa(1.0s)	Sa(2.0s)
Kempton & Stewart (2006)	-0.4690	-0.3159	-0.1183	-0.1023	-0.1550
Bommer *et al*. (2009)	-0.5669	-0.3920	-0.2131	-0.3011	-0.3482

7.4. Selection of Version 1 GMPEs for duration

The two models are both nothing more than first approximations, the modified Bommer et al. (2009) equation modelling strong distance dependence (out to 15 km) while the Kempton & Stewart (2006) model has stronger magnitude dependence. The latter model, however, provides greater flexibility and allows more meaningful predictions over a wide range of magnitudes and distances, as well as leading to correlations with Sa that are broadly consistent with the Bradley (2011) model over the period range covered. For these reasons, the fragility derivation team has selected the Kempton & Stewart (2006) model—as summarised in Eq.(7.7)—for use in the Version 1 risk model.

8. Next Steps for GMPE Development

The ground-motion model presented in this report provides GMPEs, calibrated to recordings obtained in the Groningen field, for the estimation of horizontal geometric mean values of PGA and 5%-damped response spectral accelerations at 0.2, 0.5, 1.0 and 2.0 seconds. The equations include coefficients for the median values of these accelerations and the associated aleatory variability defining the distribution of residuals, decomposed into between-event and within-event sigmas. Three alternatives suites of these GMPEs have been derived, and assigned relative weightings, in order to reflect the epistemic uncertainty in the ground-motion predictions. Additionally, guidance has been developed regarding the appropriate response spectral shapes in view of the fact that GMPEs have only been derived for five oscillator periods (assuming the spectral acceleration at very short periods is equivalent to PGA). An alternative version of the central GMPE has also been developed using an extend fault rupture rather than point-source distance metric. Finally, a very preliminary model for the estimation of ground-motion durations in Groningen has also been developed and used to derive a correlation model for the prediction of durations conditional on predicted spectral accelerations.

All of the above outputs correspond to the immediate requirements of the Version 1 risk model, as well as providing many of the requirements needed for the development of the Version 1 fragility functions. However, as has been noted earlier, the Version 1 risk model, and all of its components, represents just a snapshot of the ongoing development towards the submission of the 2016 Winningsplan. This chapter briefly discusses many of the activities that will be incorporated in the ongoing development of the GMPEs for the Groningen field; this effectively replaces the White Paper on GMPE Development that originally mapped out the envisaged work to produce the required ground-motion prediction models.

8.1. Point-source *vs*. extended-source risk sensitivity

In the development of the seismic hazard and risk model to date—including in Version 1—the seismic source model represents earthquakes as points (hypocentres) within the gas reservoir. This modelling choice was made on the basis of allowing greater computational efficiency and by using GMPEs based on hypocentral or epicentral distance, the hazard and risk computations are internally consistent. However, it is recognised that while the point-source approximation is perfectly acceptable for the smaller earthquakes that have occurred in the field until now, for the larger earthquakes envisaged in the hazard and risk calculations, the model becomes increasingly unrealistic as the magnitude grows. This is simply because the earthquake source (fault rupture) will tend to have dimensions that are clearly incompatible with the point-source assumption for sites at short distances. To address this issue, a second version of the GMPEs was derived using the Joyner-Boore distance metric, R_{JB}, which is defined as the shortest horizontal distance to the projection of the fault rupture on the Earth's surface. The R_{JB}-based model was presented in detail in Section 6.6.

The use of R_{JB} rather than R_{epi} leads to lower median predictions for larger magnitudes, especially close to the earthquake source (Figure 8.1), as well as lower sigma values. As noted in Section 6.6 and in Appendix II, the latter feature was not observed strongly in the derivation of the Akkar *et al.* (2014a) GMPEs, as seen in Figure 8.2, but this is the result of the sparse data obtained at short distances from the sources of larger earthquakes.

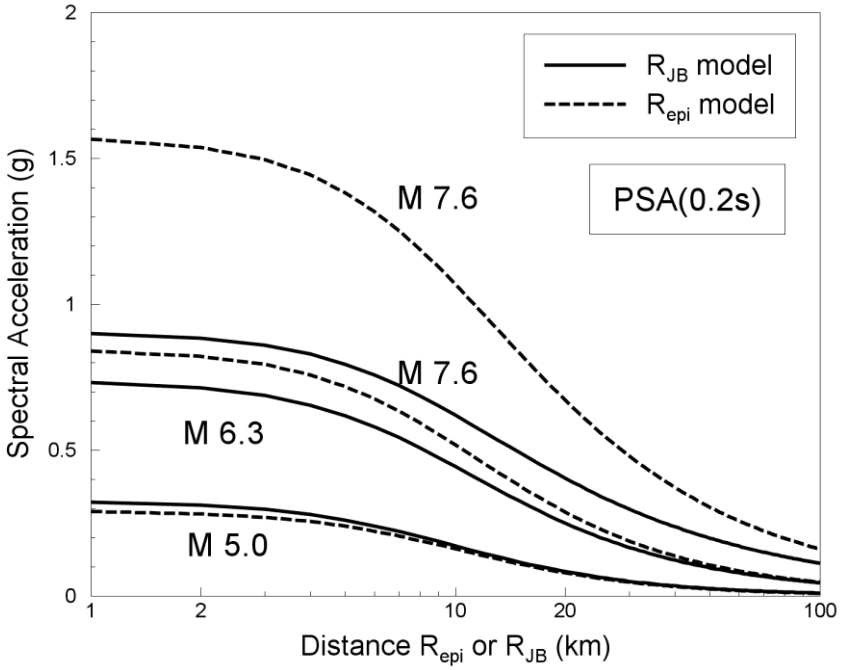


Figure 8.1. Comparison of median predicted values of spectral accelerations of GMPEs derived in terms of R_{epi} and R_{JB} (Bommer & Akkar, 2012)

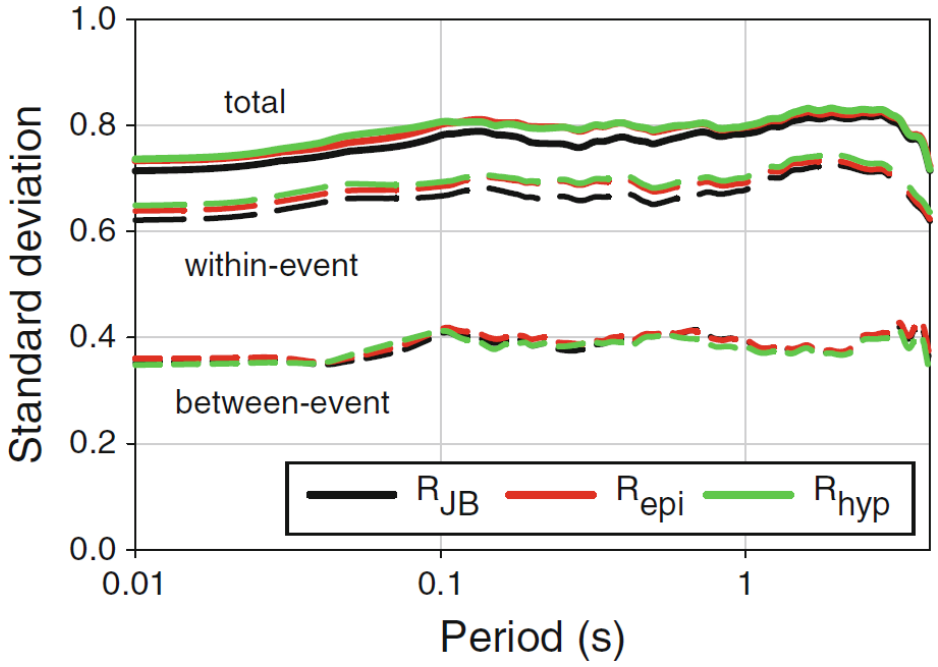


Figure 8.2. Sigma values obtained for GMPEs using different distance metrics (Akkar *et al.*, 2014a)

Although the two sets of median predictions in Figure 8.1 appear to be very different, they are actually consistent in terms of the overall 'intensity' of the predicted ground-motion field obtained in each case. The use of the R_{epi}-based model leads to circular areas of with very high-amplitude motions, whereas the R_{JB}-based model leads to a larger area—in the shape of a race track—with lower amplitudes of motion (Figure 8.3).

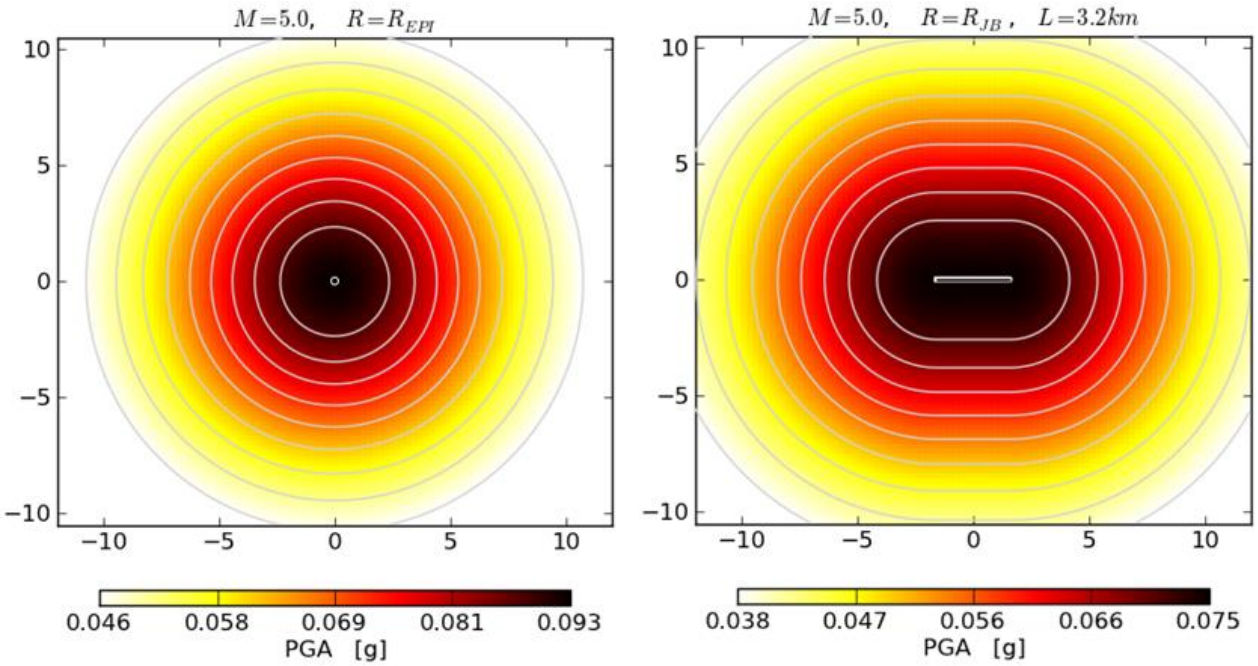


Figure 8.3. Comparison of median ground-motion fields generated for an earthquake of **M** 5 using two different versions of the Akkar *et al*. (2014a) GMPE based on point-source (*left*) and extended source (*right*) distance metrics (Courtesy of Dr Stephen Bourne).

The exercise that will be conducted to explore the sensitivity to the two options will be repeat some carefully selected part of the Version 1 risk calculations using the R_{JB}-based GMPE and modelling all earthquakes with a magnitude above a specified threshold as extended ruptures. These ruptures are likely to be assigned to the closest mapped geological fault (Figure 8.4), at least for those rupture lengths requiring a fault of sufficient size that it can be assumed not be have been missed in the mapping. The difference between the two models is that with the R_{epi}-based GMPE, fewer buildings will be hit by areas of strong shaking for each earthquake, whereas with the R_{JB}-based GMPE a larger number of buildings will experience strong shaking for each event, but the maximum amplitude of this motion will be lower. Based on the outcome of this experiment in terms of the sensitivity of the calculated risk to the two alternatives, a decision will be made regarding which approach will be used in the ongoing development of the hazard and risk estimation models. While the use of point representations of the earthquakes is computationally efficient, if the sensitivity analyses show significant differences in the risk estimations obtained using the two approaches, then consideration will need to be given to placing all larger earthquakes on mapped or hypothetical fault ruptures.

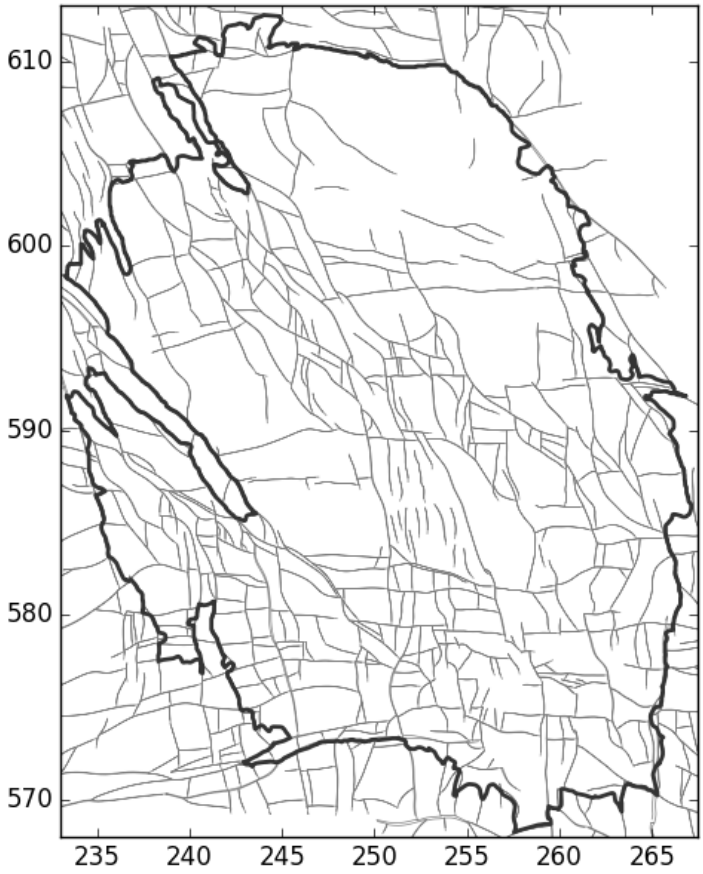


Figure 8.4. Map of known geological faults in and around the Groningen gas field (Courtesy of Dr Stephen Bourne).

8.2. Waveform modelling and refined spectral inversions

The use of stochastic simulations of ground motions has been integral to the development of the Version 1 GMPEs, and it is highly likely that this will continue to be the case in the ongoing refinement of the models. The key challenge in this approach is obtaining reliable estimates of the source, path and site parameters used in the simulations, such as stress drop, geometric spreading, Q, *etc*. The determination of these parameters from inversion of the Fourier spectra of the recorded motions is subject to various trade-offs because of parameters having either similar or opposite influences on the spectra. Therefore, if any of the parameters can be independently constrained prior to the inversions—as was done, for example, with the estimates of the high-frequency attenuation parameter, κ (Appendix I), then the influence of the trade-offs can be appreciably reduced. Modelling of the site response at the recording stations (see Section 8.4) and estimating stress drops using recordings from the geophones installed in the deep that penetrate the reservoir are two options for adding such constraints.

One of the most promising tools at our disposal is the very powerful full waveform simulation capability operated in Shell by Dr Alexander Droujinine and his colleagues, Drs Sara Minisini and Diego de Lazzari. These simulations make use of a finite difference approach and a detailed velocity model for the field from below the reservoir to near the

surface (Figure 8.5); the uppermost part of the velocity model is being refined with the more detailed information becoming available on the properties of the near-surface layers (Section 8.4). The approach is capable of producing reliable simulations up to the high frequencies of interest in ground-motion modelling (at least to 10 Hz) by virtue of the fine grid size used for the velocity model.

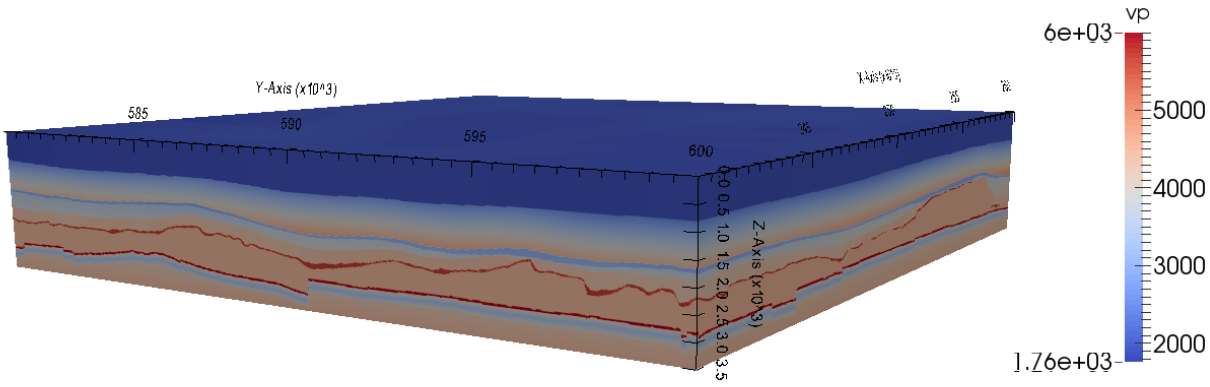


Figure 8.5. Velocity model used in the full waveform inversions; the lowermost layer of light blue is the gas reservoir (Courtesy of Dr Diego de Lazzari).

The simulations do not model the effect of anelastic attenuation, as represented by the quality factor Q, but are capable of capturing in detail the influence of the local velocity structure on the geometric spreading (Figure 8.6). These outputs will be used to define the geometric spreading function, which may be expressed in a series of segments capturing the variations that occur over rather short distances. In this regard, it is important to bear in mind that the intention here is to develop an accurate GMPE that is exclusively for application to the Groningen field; the GMPEs are very unlikely to be transportable to any other location. The geometric spreading function determined in this way can then be used in the spectral inversions to enable more reliable estimation of Q and other parameters, and similarly used in the forward simulations. Accurately capturing the path effects that occur within the first 10-15 km of the epicentre is expected to enable appreciable reduction of the within-event variability associated with the GMPEs. The trend of the predicted amplitudes shown in Figure 8.6 corresponds well with the trends of the intra-event residuals seen with respect to the Version 1 GMPE (Figure 6.31).

Another potential application of the full waveform simulations is to model the path effects on the durations of the motion, which the recordings show to undergo significant elongation over short distances from the epicentre. Estimates of the path effect on duration are needed for the stochastic simulations and therefore there is scope here for further improvements in the accuracy of predicted motions. Moreover, the accurate modelling of the elongation of the signals due to the path effects will also be useful in developing the new GMPEs for the prediction of durations (Section 8.9). Consideration will need to be given to which definitions of duration are most suitable for quantifying the length of the shaking in the Groningen field both in terms of the influence on structural response and

providing a suitable surrogate for the durations required for RVT-based stochastic simulations (*e.g.*, Atkinson, 1993; Boore & Thompson, 2014); a conversion to this duration based on DS5-95 is given by Edwards & Fäh (2013b).

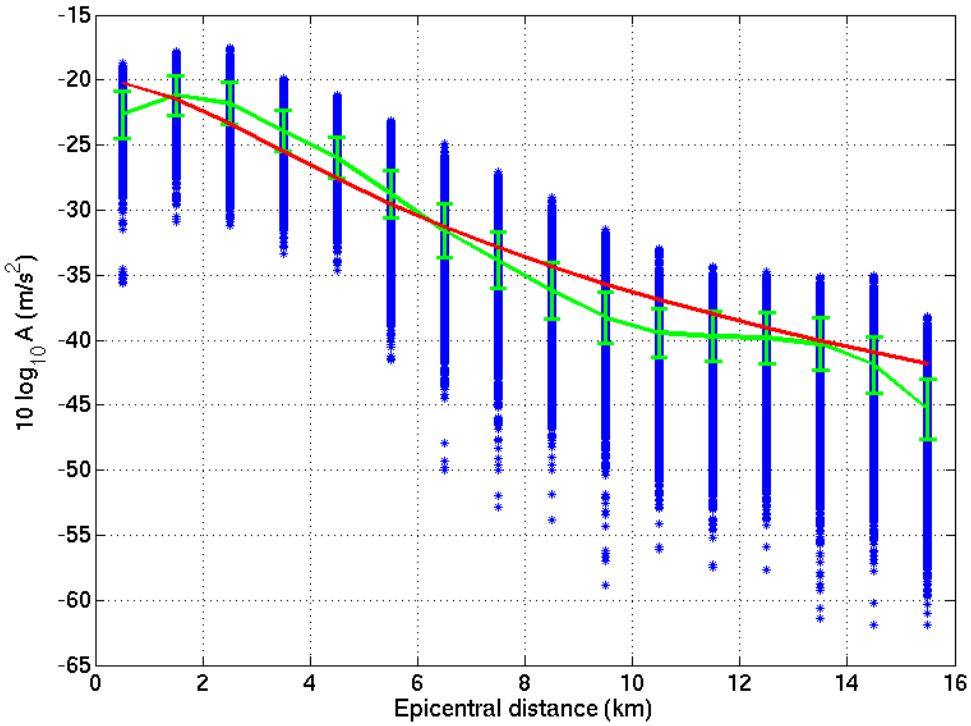


Figure 8.6. Simulated PGA values (*blue*) for an earthquake of M ~2.5 within the gas reservoir and their averages (*green*); the red curve shows median predictions from the Version 0 GMPE (Courtesy of Dr Diego de Lazzari).

8.3. Expanded database of Groningen field recordings and data processing

As was noted in Chapter 3, the installation of new and expanded seismic recording networks in the Groningen field is likely to yield large datasets for the ongoing analyses. The recordings from the existing 18 KNMI accelerographs and from the ~70 surface accelerographs installed at the locations of the 200-metre boreholes will all be directly usable; investigations are still required to ascertain which of the recordings from the TNO-installed network in buildings and from the instruments located at NAM facilities in the field are sufficiently free from the influence of structural response to be usable for purposes of calibration of the GMPEs. Even without the recordings from these other networks, the KNMI instruments can confidently be expected to yield large numbers of additional accelerograms that will be very useful in further refining the GMPEs. By way of illustration of this point, it may be noted that two earthquakes of interest have occurred in the field since the database of 85 recordings was 'frozen' for the derivation of the Version 1 GMPEs, presented in this report. These earthquakes occurred on 30th December 2014 and 6th January 2015, with magnitudes M_L 2.8 and 2.7, respectively. Invoking the assumption of equivalence between local and moment magnitudes—an issue under ongoing investigation by KNMI—the location of the data with respect the magnitude-

distance distribution of the current database is shown in Figure 8.7. The characteristics of the two earthquakes and the accelerometric recordings that they produced are summarised in Table 8.1; both events were recorded by the BFB2 station, but as before these records are not currently being used.

Table 8.1. Characteristics of most recent earthquakes contributing to the database

EQ No.	EQ Name	Date	M	Epicentral Coordinates		Records	
				X (RD)	Y (RD)	Event	Total
13	Woudbloem	30 December 2014	2.8	244,561	580,898	14	99
14	Wirdum	6 January 2015	2.7	246,987	593,800	14	113

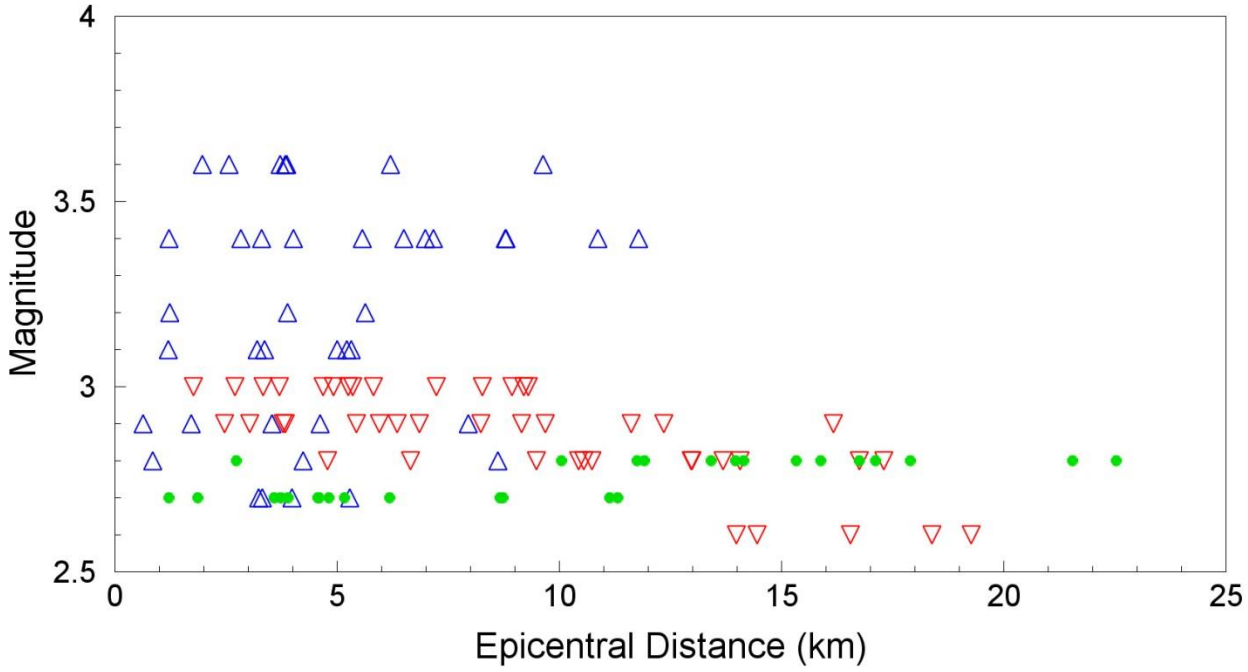


Figure 8.7. Magnitude-distance distribution of the Version 1 database (*red and blue triangles*) and the recordings from the earthquakes of 30 December 2014 and 6 January 2015 (*green circles*)

Figures 8.8-8.10 show the between- and within-event residuals of the recorded acceleration at various response periods calculated with respect to the Version 1 GMPEs; in each case, these are compared with the residuals from the 85 recordings used to derive the GMPEs. The residual plots show that the Version 1 central GMPE provides a good fit to the new data, although it is notable that the event terms are small (generally negative) in both cases, suggesting that these events were of lower than average source strength. The within-event residuals do not imply that the ϕ_{SM} value is insufficient. At some point in the coming months, the database will once again be 'frozen' for the Version 2 GMPE development, but in the meanwhile as new records are obtained, similar comparisons will be made to provide insights into the performance of the Version 1 GMPEs.

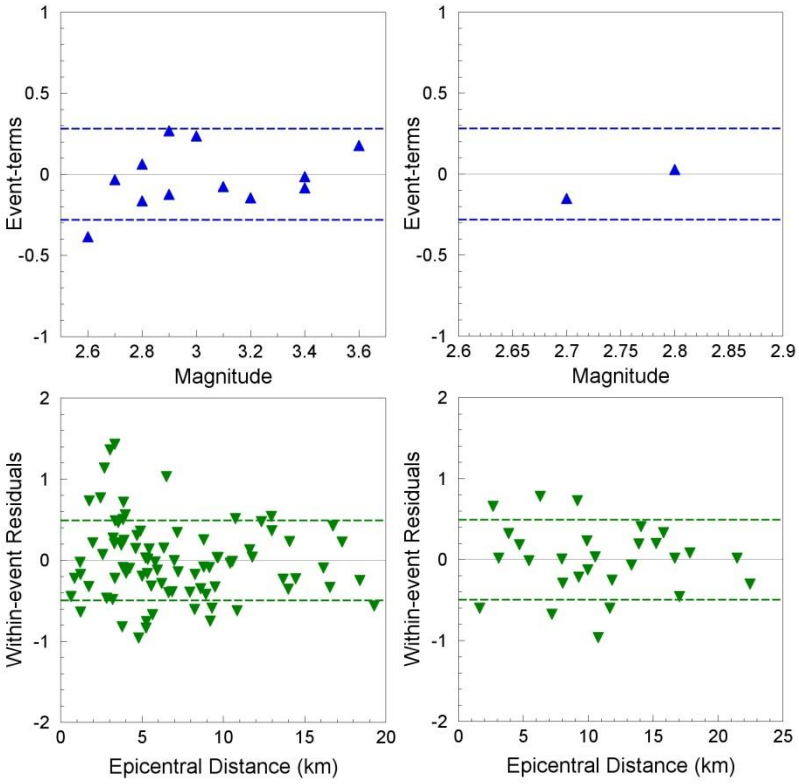


Figure 8.8. Residuals of the Groningen PGA values with respect to the Version 1 GMPE for the 85 records used in deriving the model (*left*) and for the additional recordings from the two most recent earthquakes (*right*)

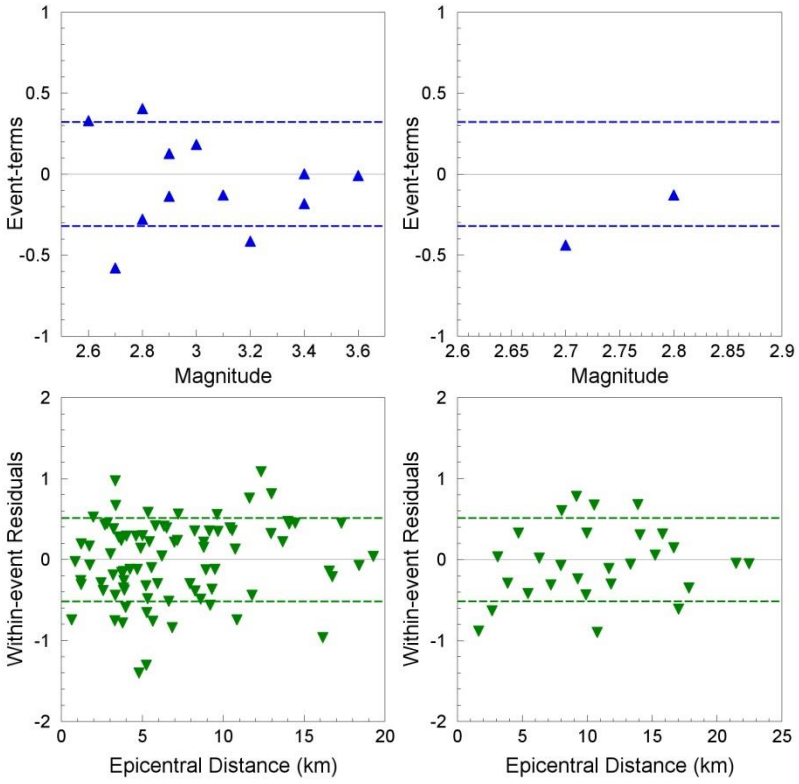


Figure 8.9. Residuals of the Groningen Sa(0.5s) values with respect to the Version 1 GMPE for the 85 records used in deriving the model (*left*) and for the additional recordings from the two most recent earthquakes (*right*)

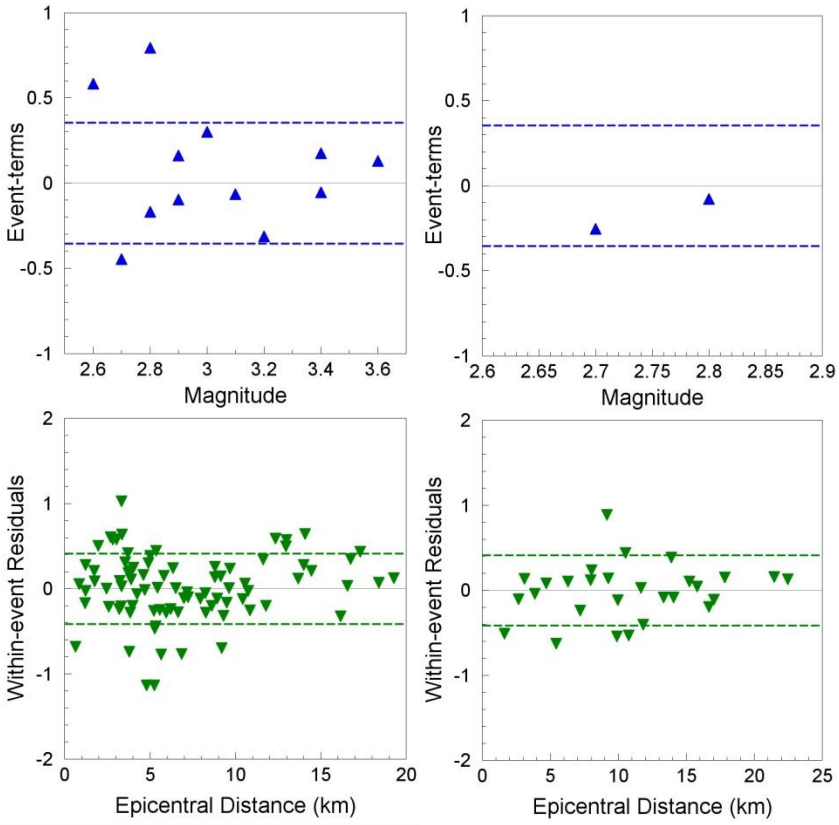


Figure 8.10. Residuals of the Groningen Sa(2.0s) values with respect to the Version 1 GMPE for the 85 records used in deriving the model (*left*) and for the additional recordings from the two most recent earthquakes (*right*)

In addition to the obvious benefits of an expanded database, another important feature is that several of the accelerograph stations have now recorded a number of earthquakes (Table 8.2), and they are likely to accumulate additional recordings in the coming months. Such multiple recordings at individual sites can facilitate the direct calculation of the so-called single-station sigma (Atkinson, 2006; Rodriguez-Marek *et al*., 2011, 2013). In essence, it has long been recognised that the sigma values associated with empirical GMPEs are probably over-estimated with respect to the estimation of variability over time (*i.e.,* over multiple earthquakes) at the site of interest, since this is modelled using the variability of motions recorded at many sites, often in many regions; this substitution of spatial variability for temporal variability is referred to as the ergodic assumption (Anderson & Brune, 1999). By decomposing the total variability, sigma, into its component parts (*e.g*., Al Atik *et al*., 2010), we can identify those that are treated as being random but are actually repeatable, and can therefore be removed from the sigma. One such element is the site-to-site variability, which can be modelled through site response analyses (Section 8.4), but single-station sigma—the variability that remains when the influence of repeatable site effects has been removed—can also be estimated directly from multiple recordings at a site. This is actually more exactly referred to as single-station phi (ϕ_{SS}) since it refers only to the intra-event variability, but others have extended the same concept to looking at the contributions to the total variability (sigma) coming from repeatable source and path effects (Lin *et al*., 2011). Since the Groningen earthquakes are currently assumed to all occur on

fault ruptures within the gas reservoir and propagate upwards through geological layers that are rather uniform across the field (up to within the uppermost few hundreds of metres), there should be ample scope of reductions of the variability associated with the GMPEs.

Table 8.2. Numbers of earthquakes from current database of 14 events recorded at KNMI accelerograph stations yielding five or more recordings

Station	Number of EQs	Station	Number of EQs
BWSE (WSE)	12	BOWW	6
BZN2 (ZAN2)	10	BSTD (STDM)	6
BGAR (GARST)	8	BZN1 (ZAN1)	6
BHKS (HKS/HOEK)	8	BFB2 (FRB2)	5
BMD2 (MID3)	8	BLOP	5
BWIN (WIN)	8	BONL	5
BAPP	6	BUHZ	5
BMD1 (MID1)	6	BWIR	5

As the Groningen ground-motion database continues to grow, choices will also need to be made regarding the processing of the accelerograms. Since the piece-wise linear corrections for baseline shifts discussed in Section 3.2 (Figure 3.11) probably cannot be automated and is a time-consuming procedure, it will probably be necessary to apply high-pass filters to the recordings, but some sensitivity studies will be performed to explore the different outcomes obtained with the two procedures. Ultimately, the purpose of the filtering will be to ascertain for each record the maximum response periods for which the response spectral ordinates may be considered reliable (*e.g.*, Akkar & Bommer, 2006). As noted in Section 3.2, these limits will be decided in conjunction with the processing of the records for inversion of the Fourier spectral ordinates.

8.4. Characterisation of recording stations and site response model

The Version 1 GMPEs represent a major advance with respect to the Version 0 GMPEs, in providing predictions for PGA and response spectral ordinates in a way that more faithfully represents the local data and simultaneously captures the epistemic uncertainty in the extrapolations to larger magnitudes. However, while the Version 1 GMPEs are well calibrated to local data, there are two aspects in which the models could be significantly improved to lead to both more robustly constrained median predictions and reduced sigma values:

1. The models represent the local site response as represented by the motions recorded at the KNMI accelerograph stations. In other words, the site amplification factors implicit in the Version 1 predictions represent the profiles at the recording locations and it is not clear at this stage how representative these are of the profiles across the entire field.
2. The site amplifications implicit in the recorded data probably represent the linear response of the near-surface layers at the recording sites, as a result of the small

magnitudes of the earthquakes. The extrapolation to larger magnitudes using stochastic simulations has simply maintained the site term constant at each frequency, with the result that the site response is still modelled as linear even for larger magnitude events. Since the near-surface layers in the Groningen field have very low shear-wave velocities, it is reasonable to assume that the response will be non-linear under the motions generated by larger earthquakes.

These two shortcomings of the Version 1 GMPEs are among the main focuses of the envisaged developments of the Version 2 equations, and several activities are already underway towards improving these features. Firstly, there is a campaign to obtain detailed characterisation of the near-surface layers (at least the top 30 m) at each of the recording stations. This work is being carried out by Deltares and will begin in the coming weeks with intensive measurements at a number of selected sites, subject to permission from landowners for access. The original plan was for several measurement techniques to be applied at the BHAR, BOWW, BWIN and BZN2 stations but permitting issues prevented access at these sites at the current time; the initial set of measurements will now be made for the BOPP station. At these pilot locations, shear-wave velocity measurements will be made using all of the following techniques:

- Active and passive MASW
- Seismic CPT, with offsets
- Downhole direct measurements
- Cross-hole direct measurements
- PS suspension logging

The purpose of applying this range of measurement techniques at these sites is both to explore the applicability of each of them to this environment (although some, such as seismic CPT, have already been extensively used) and to compare the results obtained. On the basis of these comparisons, a subset of the measurement approaches will be applied to the remaining 14 KNMI accelerographs and subsequently at the additional ~70 accelerographs being installed in conjunction with the geophones in 200-metre boreholes (Figure 3.2). The V_S profiles obtained in this way will clearly enable better characterisation and improved interpretation of the recorded motions.

The ultimate objective is to develop GMPEs that include non-linear site amplification functions for the Groningen field. The aim is to condition the functions at each response period on the spectral acceleration at the same period in the underlying baserock, which is preferable to the more widely-used approach of conditioning the non-linear response on PGA, as recommended by Bazzurro & Cornell (2004a) and implemented by Chiou & Youngs (2008). While it is tempting to integrate fully probabilistic site response into the hazard and risk calculations following the method of Bazzurro & Cornell (2004b), especially in view of the relatively simple (but computationally intense) implementation of this approach within a Monte Carlo framework, this is more appropriate for site-specific studies (*e.g.*, Rodriguez-Marek *et al.*, 2014).

To incorporate the non-linear site response characteristics of the Groningen field into the ground-motion predictions, the first stage of work is the development of a velocity model for the entire field (Kruiver *et al*., 2015a). The characterisation of the near-surface layers is being led by Deltares, who have constructed a geological model for the uppermost part of the field in which each layer was then be assigned a shear-wave velocity using correlations established for the field. A preliminary version of a V_{S30} map for a pilot study area is shown in Figure 8.11.

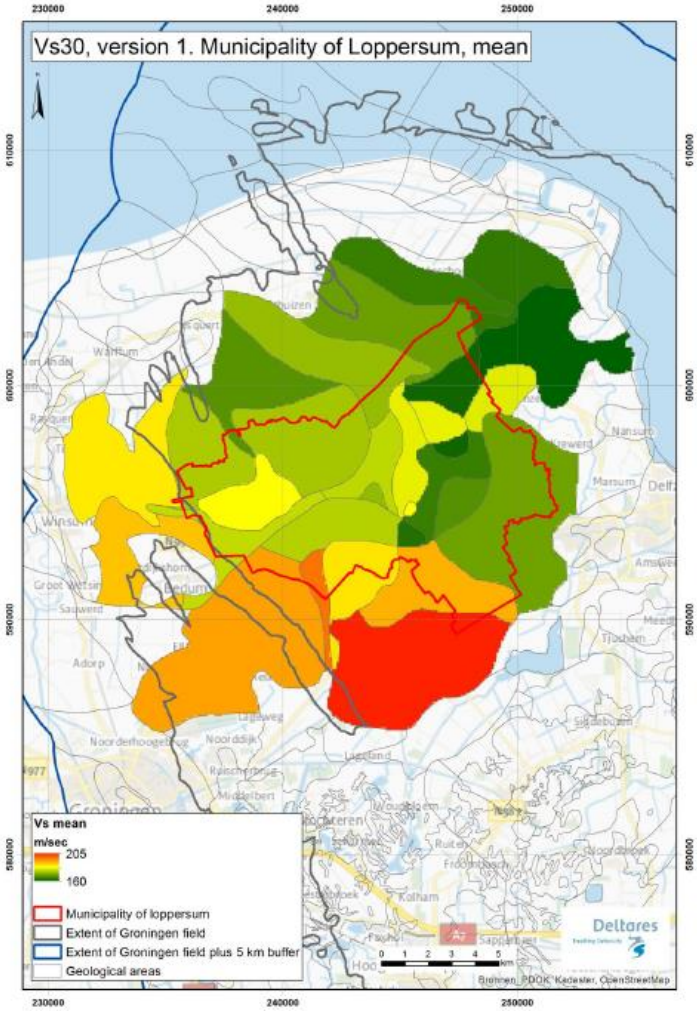


Figure 8.11. Version 1 V_{S30} map for the Loppersum municipality based on geological characterisation of the near-surface layers and assignment of V_S values to each lithological unit (Kruiver *et al*., 2015).

The near-surface velocity model developed by Deltares was then extended to greater depths through combination with the velocity model being produced by Ewoud van Dedem from Shell, who is re-processing the surface-wave information from seismic imaging of the gas reservoir conducted in the 1980s. The method being used, which is called MEIDAS (developed by F. Ernst) is akin to MASW, and its application to this legacy data is yielding V_S profiles down to 120 m (Figure 8.12).

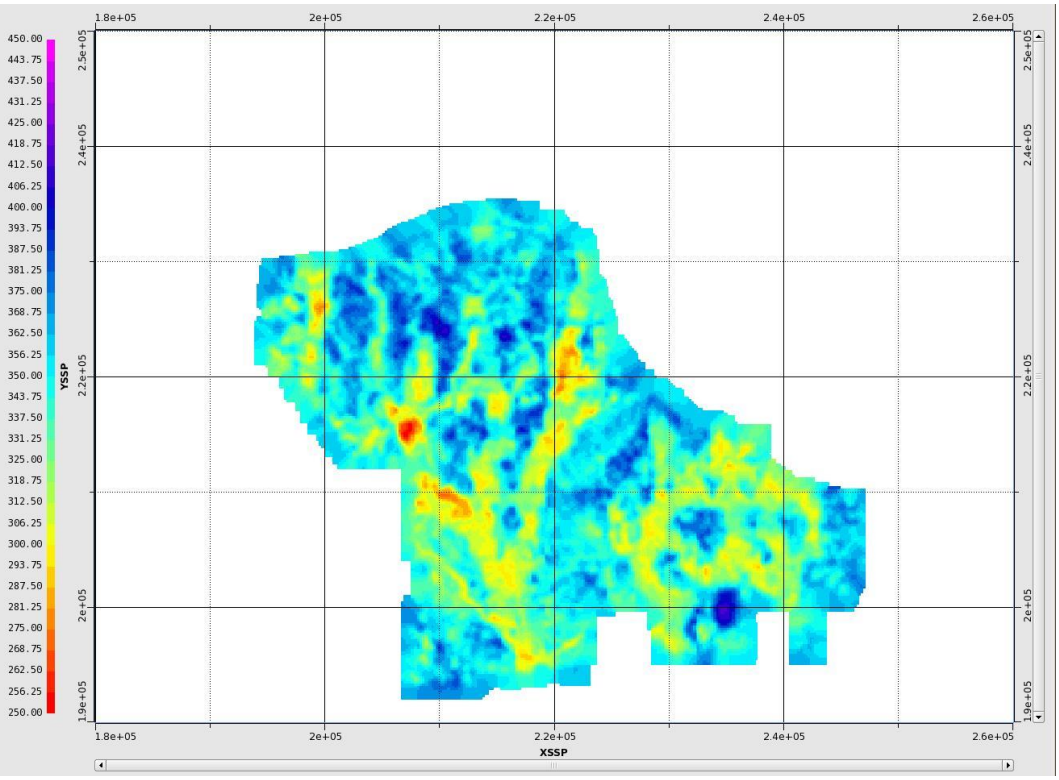


Figure 8.12. Shear-wave velocities at 40 m depth obtained from analysis of surface-wave data obtained during seismic imaging of the gas reservoir (Courtesy of Ewoud ven Dedem)

Below the ~120 m limit of the V_S model obtained through the MEIDAS model, the velocity profiles is linked with the field-wide model recently updated by Remco Romijn and his team at Shell. On the basis of the complete velocity profile from the ground surface to the gas reservoir located at a depth of about 3 km, a horizon will then be selected that will be taken as the top of the elastic half-space in the site response analyses. At the time of writing, this horizon is likely to be at the base of the Upper North Sea formation, which is encountered at a depth of about 300 m and where V_S values reach at least 1,000 m/s.

The site response calculations will be performed by Deltares, with advice provided by Professor Adrian Rodriguez-Marek of Virginia Tech who will be working with the GMPE development team on the incorporation of site amplification factors into the predictions. As a result of the very large numbers of site response calculations that need to be performed, it is likely that Deltares will take advantage of the RVT-based approach to site response analysis (Rathje & Ozbey, 2006) using the program STRATA (Kottke & Rathje, 2008), which do not require time-series inputs. However, since it is well established that RVT-based and time-domain site response analyses yield different results (Kottke & Rathje, 2013), it is envisaged that time-history analyses will also be conducted. Input to these analyses may be provided by stochastic simulations. An even more challenging part of the site response analyses is the definition of the other properties (other than V_S) required for the full dynamic characterisation of the profiles: the densities of the materials can be estimated with some confidence but the selection of appropriate damping curves and

stiffness degradation curves, particularly for the organic soils (peats) that are encountered in many parts of the field. A campaign of field sampling and laboratory testing is being developed by Deltares to provide this information.

A White Paper outlining the planned programme of work to develop the field-wide site response characterisation model has been issued by Deltares in collaboration with the GMPE development team (Kruiver *et al*., 2015b). The envisaged approach to developing a ground-motion prediction model incorporating these non-linear amplification functions is as follows. The final objective will be to divide the entire field into a number of zones, for each of which a representative non-linear site amplification function can be defined. These zones will be based on the geological areas defined in Chapter 3 of Kruiver *et al.* (2015a) but they may be sub-divisions of these geologically-delineated zones. The site amplification zones will be defined such that there is a limited range of variability in the calculated amplification characteristics within each zone and distinct variations in the amplification functions of adjacent zones. Once the field is divided into N site amplification zones, the GMPE for predicting values of spectral acceleration for zone i, Sa_i, at different periods, T, will have the following functional form:

$$\ln[Sa_i(T)] = \ln[Sa_{NU\_B}(T)] + [AF_i(T) + \delta S2S_i] + \delta B + \delta WS \quad (8.1)$$

where $Sa_{NU\_B}(T)$ is the predicted median spectral acceleration at period T in the reference baserock horizon, which is the base of the Upper North Sea Formation, designated as NU_B, $AF_i(T)$ is the median amplification function for the spectral acceleration at period T for the i^{th} zone, and $\delta S2S_i$ is the variability of the zone-specific amplification function and is described in more detail below.

In terms of the variability in the predictions, the term δB is the earthquake-to-earthquake residual (i.e., a random sample from the distribution of between-event variability) and δWS is randomly sampled from the distribution of single-station within-event variability. Both δB and δWS are assumed to be zero-mean Gaussian random variables with standard deviation τ and Φ_{ss}, respectively. These two components of variability will have distributions that are constant across the field. The term $\delta S2S_i$ is the randomly sampled residual from the site-to-site variability for zone *i*, which is assumed to be a zero-mean Gaussian random variable with standard deviation Φ_{S2Si}; for more background on these terms and the decomposition of the ground-motion variability, see Al Atik *et al*. (2010) and Rodriguez-Marek *et al*. (2014). In this formulation, which is specific to the Groningen field and not intended in any way to create equations that would be transportable to any other environment, there is no requirement for the inclusion of simplified site classification parameters such as V_{S30}. The amplification functions will represent the nonlinear response of the overlying layers (of thickness varying from about 80 to 630 m across the field, with a mean value of about 350 m for much of the area) above the NU_B horizon. Amplification factors will be determined for each response period for which equations are developed to predict spectral accelerations. Randomised profiles will be used to represent the lateral variability of V_S and other parameters across the specific zone for which they are derived. The amplification functions will be defined as a median value from the analyses of both the

randomised profiles and varying dynamic input as a function of the baserock acceleration at the same period, and will include an estimate of the variability (Figure 8.13).

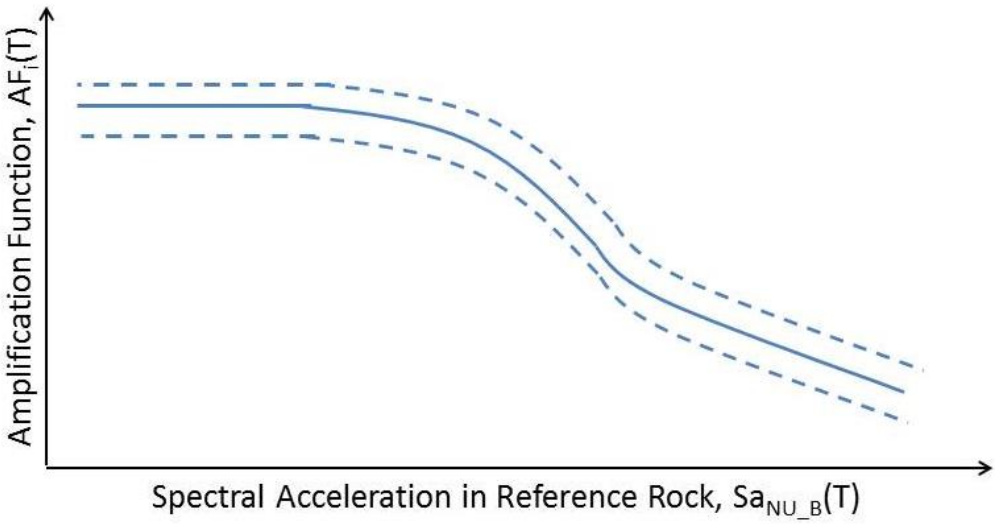


Figure 8.13. Schematic illustration of the form of the amplification function for one zone and one response period; the solid line represents the median estimates, the dashed lines the uncertainty as indicated by +/- one standard deviation

As noted above, the measures of between-earthquake (inter-event) variability, τ, and single-station within-earthquake (intra-event) variability, Φ_{ss}, will be determined from the recordings in the field and other considerations, and will apply at all locations. The site-to-site variability, $\delta S2S_i$, however will be specific to each zone and will reflect the variability in the site amplification functions obtained for the zone. If the variability of AF_i in a particular zone is very large, consideration may be given to sub-dividing the zone into smaller areas over which the profiles are less variable. Similarly, for those zones where there is additional information—for example from the 200-m geophone boreholes and/or from site-specific V_S measurements—the benefit of the improved constraint on the local site response characteristics will be reflected in smaller values of $\delta S2S_i$.

In order to be able to make full use of the accelerograms recorded by surface instruments in the field (currently about 30 but ultimately at least 88 and perhaps in excess of 100), the amplification functions at each recording station will be required. For the weak levels of motion recorded to date, these response functions are expected to be linear and therefore their application to the records to transform them to baserock and then their subsequent application to bring the predicted baserock motions back to the surface have a zero net effect. The surface recordings can therefore be used to estimate the variability in the GMPEs, following broadly the same approach used for the Version 1 GMPEs but also informed by global estimates for single-station within-event variability (*e.g.*, Rodriguez-Marek *et al.*, 2013) and for single-path-single station variability (*e.g.*, Lin *et al.*, 2011). Using site-specific information at the recording stations, non-linear site amplification functions, AFS_j, will be calculated, where *j* is the number of the recording station site (likely

to be much fewer than the number of site amplification zones) in the same way as the amplification functions for each zone. If the amplification function for a given station, AFS, does not fall within the uncertainty bounds on AF for the zone in which the station is located, then the uncertainty range for that zone will need to be expanded or the AF re-calculated taking into account the site-specific information at the station. Using the station-specific amplification function, AFS_j, the residuals for each station will be calculated as the difference between the recorded surface acceleration and that predicted by the following prediction of the unbiased median (with δS2S set to zero):

$$\ln[Sa_i(T)] = \ln[Sa_{NU\_B}(T)] + AFS_j(T) \quad (8.2)$$

The components of variability (τ and Φ_{ss}) will then be calculated from these residuals, which as in the Version 1 GMPE will be calculated relative to the parametric functional form regressed on the stochastically-simulated ground motions.

As for the Version 1 GMPEs, stochastic simulations that will be used to generate the accelerations used to derive the model for predicting the median spectral accelerations, which will now be referenced to the baserock, $Sa_{NU\_B}(T)$. The first stage will be to invert the Fourier spectra of the final database of surface recordings to determine the source (moment, stress drop), path (geometric spreading, Q) and site (κ, amplification function) parameters; it is hoped that several of these factors will be independently constrained.

In the inversions for the Version 1 GMPEs, which provided direct predictions of motions at the ground surface, the amplification function obtained was a network average. For the Version 2 model, the site amplification element of the inversions will have two components, one representing the path from the source (reservoir) to the NU_B horizon, and the other the layers from the NU_B horizon to the ground surface. Using the station-specific estimates of the attenuation parameter kappa and linear amplification functions estimated from the V_S profiles above the NU_B horizon at each station, the inversions will in effect be performed on the Fourier spectra transformed from the surface to the NU_B. The inversions will therefore yield a field-wide estimate of the amplification function and kappa value for the NU_B horizon with respect to the reservoir.

In the forward modelling, the same parameters will be used for the source and path—with the signal duration hopefully constrained by full waveform simulations—but with the amplification function and κ value corresponding to the top of the assumed elastic half-space for the site response analyses, *i.e.*, the NU_B horizon. For the station-specific amplification functions that will be removed before the inversions are performed, these will initially be based on the V_S profiles at the stations inferred from the GSG model. As V_S measurements at the recording stations progress, the values of AFS_j will be updated using the improved V_S profiles.

The site response amplification function will be based on the V_S and density profiles established for the path from the reservoir to the NU_B horizon (see Chapter 8); the

inferred value of κ assigned to the NU_B horizon may need to be revised iteratively such that the combined effect of the Q in the travel path, the κ_{NU_B} value and the damping assigned to the overlying layers modelled in the site response analyses, is consistent with the κ values estimated for the surface locations. Immediately here there will be a challenge to be addressed if, as in the derivation of the Version 1 GMPEs, a single field-wide value of κ is assumed, since there will be a common field-wide path from reservoir to NU_B but varying site response layer models.

Finally, checks will need to be made on the effect of decoupling the wave propagation and amplification up to the NU_B horizon from the amplification and attenuation in the uppermost layers considered in the site response analyses. This will be done by comparing the results of the decoupled model with the results obtained by modelling the wave propagation directly from the reservoir to the ground surface.

8.5. Spatial correlation of predicted ground motions

Several studies have noted that the variability of ground-motion amplitudes at closely-spaced accelerograph stations is lower than that expected from empirical GMPEs, indicating that there is a degree of spatial correlation in the seismic shaking (*e.g*., Boore *et al*., 2003; Wang & Takada, 2005; Jayaram & Baker, 2009; Goda & Atkinson, 2010; Esposito & Iervolino, 2011). Examples of spatial correlation functions for PGA are shown in Figure 8.14. The effect of the spatial correlation of ground motions is to produce higher estimates of losses in risk modelling for geographically-distributed exposure, since it leads to pockets of higher acceleration that result in greater damage when these coincide with concentrations of weak buildings (*e.g*., Crowley *et al.,* 2008).

In the Version 0 and Version 1 risk models, the exposure is grouped into 3 x 3 km squares and the ground-motion amplitudes calculated at the centre of each square applied to all buildings within the grid cell. This is a computational convenience, since sensitivity analyses showed that using a smaller grid size (such as 1 x 1 km) resulted in a tremendous computational penalty, while the coarser grid does not result in great loss of accuracy, consistent with findings from other studies (*e.g*., Bal *et al*., 2010). The assumption of uniform motions across each grid cell also conveniently serves as a surrogate for including spatial correlation. However, it must be recognised that the correlation lengths vary with spectral response period (*e.g*., Esposito & Iervolino, 2012), so the approximation becomes even cruder when spectral ordinates at multiple oscillator frequencies are being considered. The grid cell sizes—which need not necessarily remain uniform across the field—may be modified to accommodate any marked spatial variation of site amplification effects (Section 8.4). The final choices regarding spatial aggregation of the risk calculations and the degree to which spatial correlation of motions is explicitly modelled will be decided as part of the development of the Version 2 risk model. If spatial correlation of ground motions is to be included, then the Groningen data may be used to ascertain if any of the existing models can be adopted or adapted for this application.

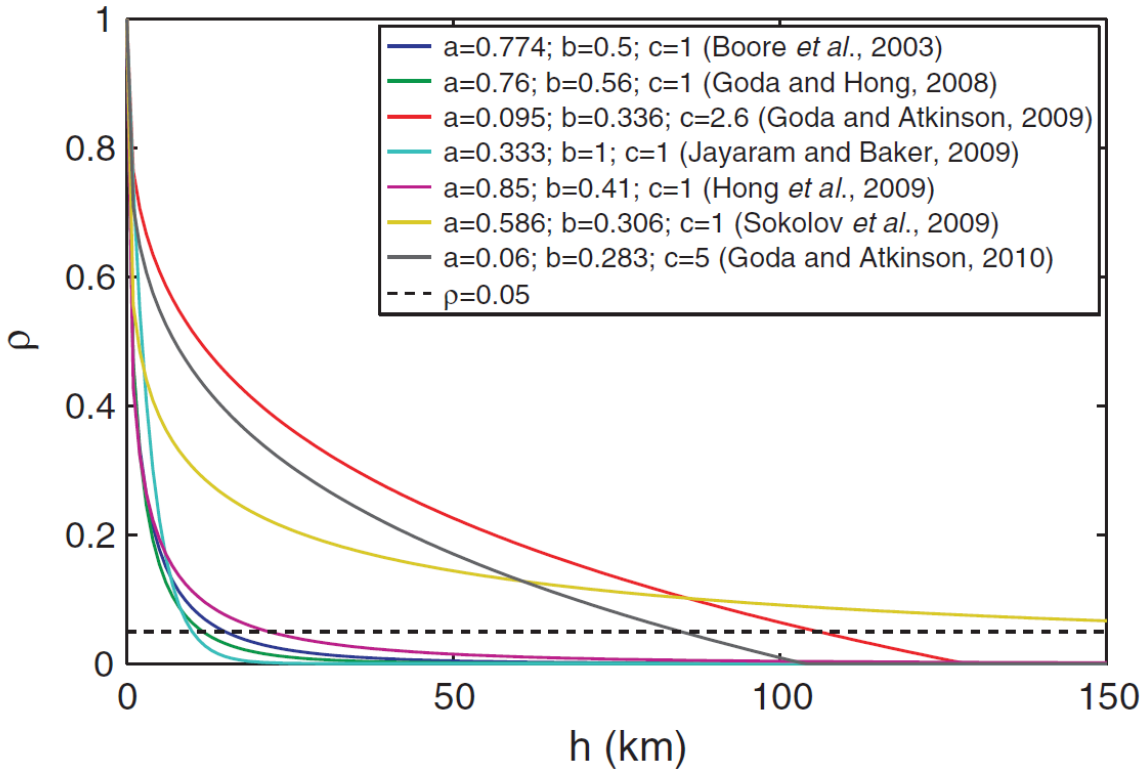


Figure 8.14. Comparison of published correlation functions for PGA as a function of separation distance, *h*; the dashed black line represents the correlation coefficient of 0.05, which may be considered as the level at which all correlation is effectively lost (Esposito & Iervolino, 2011)

8.6. Period-to-period correlations of variability in spectral ordinate predictions

Using GMPEs for the spectral ordinates at multiple oscillator periods to calculate each spectral ordinate independently in PSHA in effect treats the variability at all periods as being perfectly correlated. Various studies have shown that this is not the case and to avoid this unintended conservatism it has been proposed to generate response spectra taking account of the decreasing correlation with increasing separation of the periods (Baker & Cornell, 2006). A response spectrum generated in this way, referred to as a conditional mean spectrum, CMS (Baker, 2011), is the approach being used by the group responsible for the development of the fragility functions. The concept of the CMS is illustrated in Figure 8.15.

The key requirement to be able to generate a CMS for a given earthquake scenario, in addition to the GMPEs for the spectral accelerations at multiple periods, is a model for the period-to-period correlation of the variability in the predicted accelerations. For the Version 1 risk model, the fragility group have made use of the correlations derived by Akkar *et al.* (2014c) using European data. As part of the Version 2 GMPE development, which will include the prediction of spectral ordinates at a much greater number of oscillator periods, the derivation of a Groningen-specific correlation function.

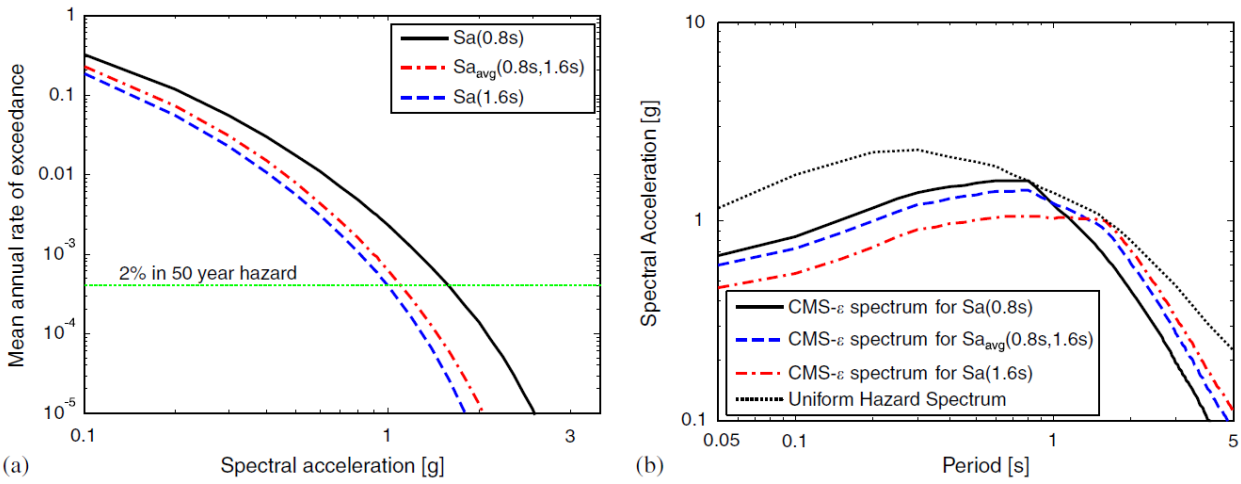


Figure 8.15. (a) Hazard curves for spectral accelerations at a site, and (b) the uniform hazard spectrum for the site for a 2% probability of exceedance in 50 years and the CMS conditioned on the spectral accelerations at two periods or on their average (Baker & Cornell, 2006)

8.7. Component-to-component ratios and variability

One of the striking features of many of the stronger ground-motion recordings from the Groningen field is the marked polarisation of the horizontal components, which is very evident in the spectral ordinates, particularly when displayed in terms of relative displacements (Figure 8.16).

In keeping with conventional practice, the GMPEs will continue to be developed for the geometric mean of the horizontal component of motion, but there are several reasons why predictions may also be required for the randomly-oriented component of motion, including the fact that durations may be better predicted using both components of horizontal motion rather than geometric mean. While the median predictions obtained using both horizontal components from each record or their geometric mean are identical, the variability associated with the former is larger due to the component-to-component variability. Therefore, this additional component of variability will be calculated for the Groningen recordings in order to allow the adjustment between geometric mean and randomly-oriented components, following Boore (2005b).

Additionally, for purposes such as generating acceleration time-series for 3D structural analyses, it would be necessary to be able to transform the geometric mean response spectral ordinates into two orthogonal components, with appropriate ratios of their ordinates. While it would be possible to simply apply published response spectral ratios between the different horizontal component definitions—such as those of Beyer & Bommer (2006) or Watson-Lamprey & Boore (2007)—it would be make sense to first check if these are consistent with the Groningen data, and if not then to modify the ratios accordingly.

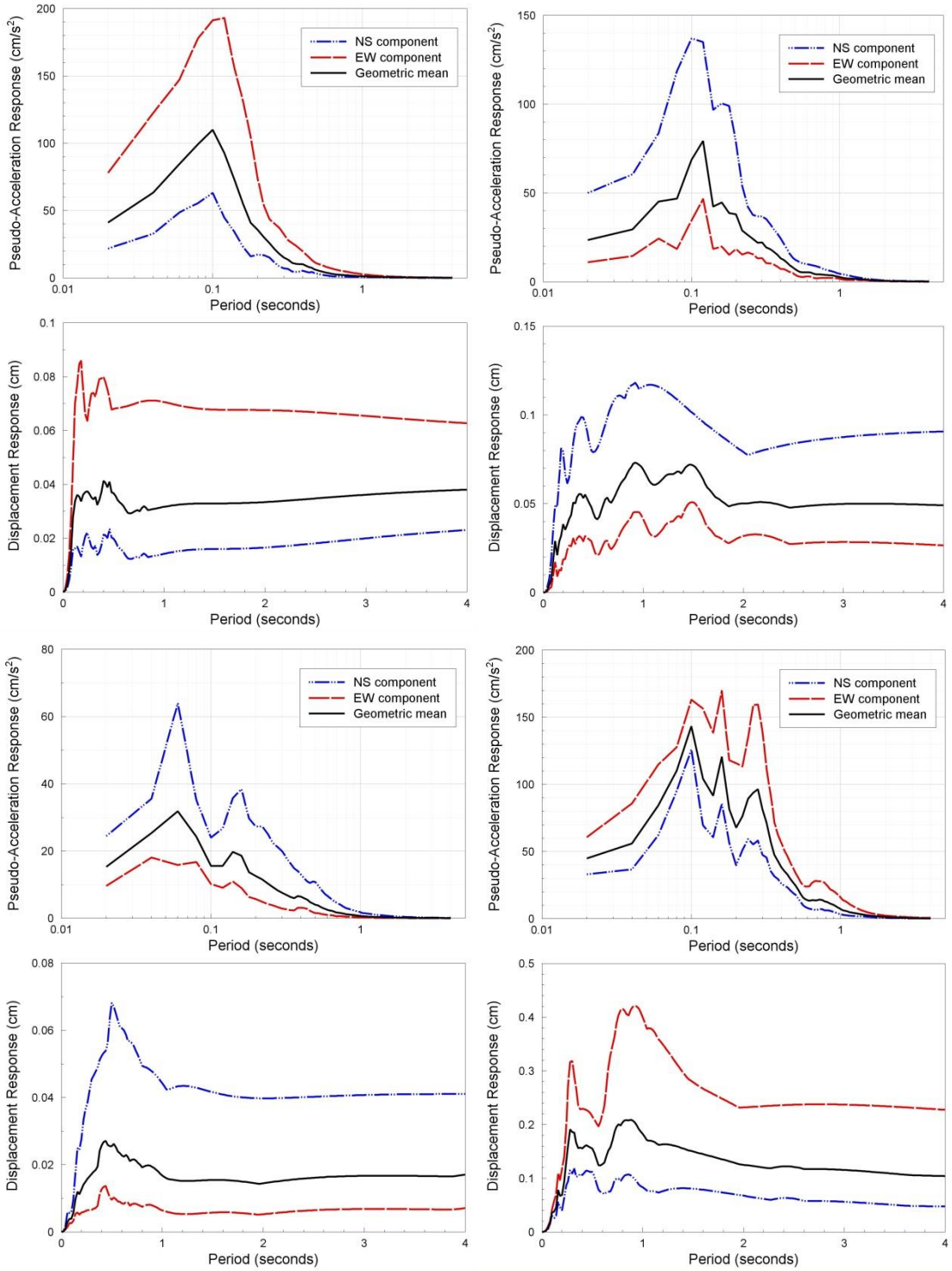
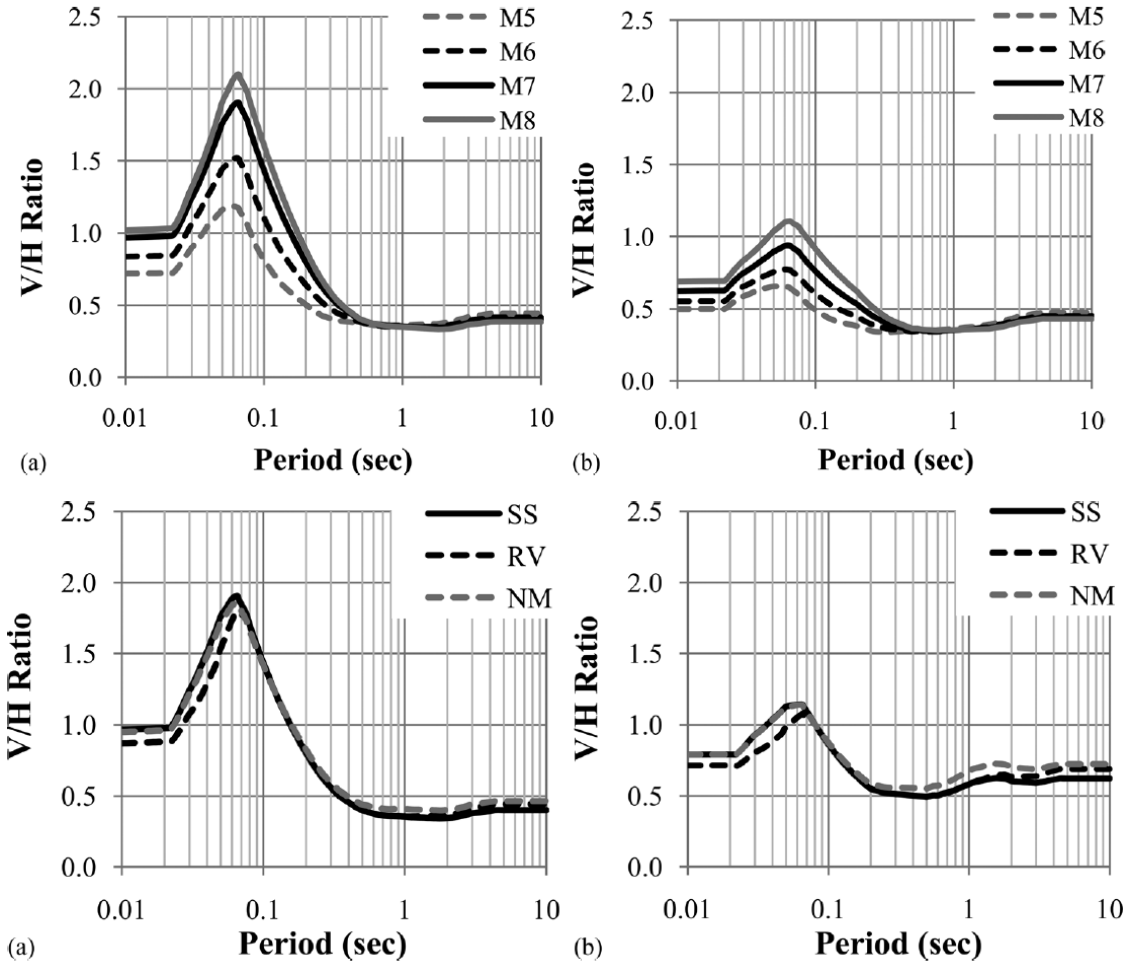


Figure 8.16. Examples of highly-polarised recordings from the Groningen field

8.8. Vertical-to-horizontal ratios of Groningen ground motions

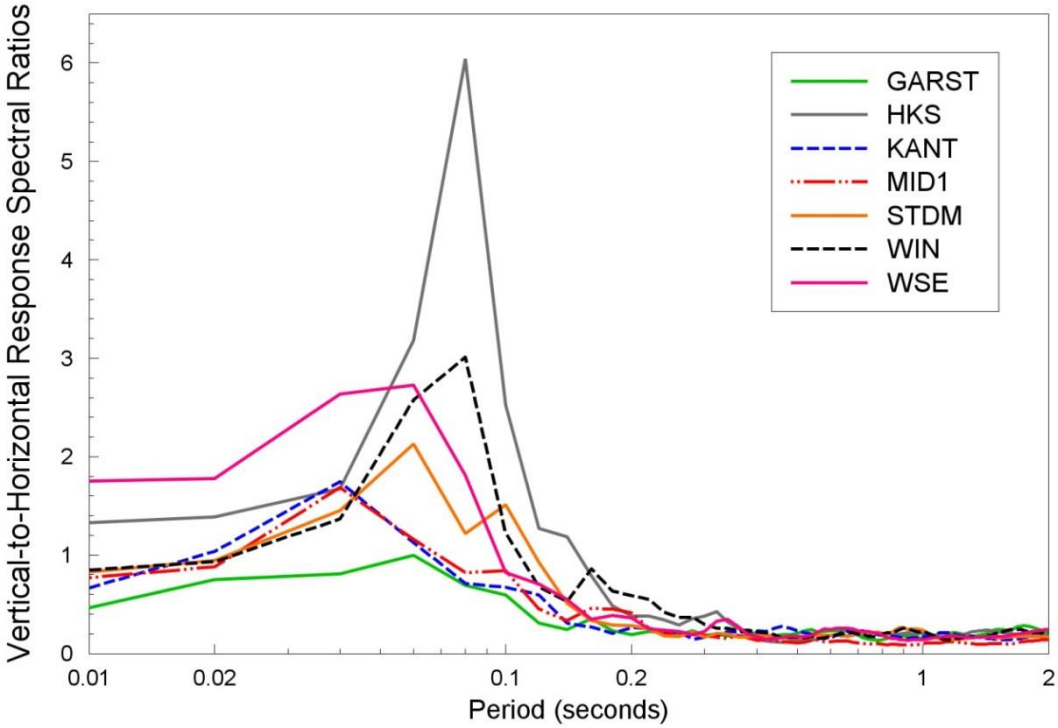
The fragility functions for buildings are being defined only in terms of horizontal ground motions and therefore it has initially been assumed that there is no requirement for direct predictions of the vertical component of motion. However, for the masonry and pre-cast concrete structures, it is thought that vertical motion may play a significant influence on the response, so three-dimensional dynamic input to the analyses of the structures will be

required. In other words, the horizontal motions on the x-axis of the fragility functions would represent the coupled effect of the combined horizontal and vertical motions. This approach is based on the assumption that the V/H ratio is reasonably stable over the magnitude-distance ranges of interest, and the variations from those ratios do not exert a strong influence on the structural response. The validity of this assumption needs to substantiated by the structural modelling group. In order to ensure that the vertical components are appropriately selected and scaled, V/H response spectral ratios consistent with the seismicity and ground conditions need to be defined. Past practice has often defined the vertical spectrum as simply a scalar product—the factor usually being on the order of ⅔--of the horizontal spectrum, but it is now recognised that the V/H ratio varies with response period and that the vertical spectrum has a distinct shape. Moreover, it is recognised that V/H ratio varies with magnitude, style-of-faulting, distance and site classification (Figure 8.17).



8.17. Predicted median ratios of vertical-to-horizontal response spectral ordinates obtained from the GMPEs of Gülerce & Abrahamson (2011). *Upper*: V/H ratios at a site with V_{S30} of 760 m/s for different magnitude earthquakes at a distance of (a) 5 km and (b) 30 km. *Lower*: V/H ratios for an **M**7 earthquake at 5 km from sites with V_{S30} of (a) 270 m/s and (b) 760 m/s, for different styles-of-faulting. Figures from Gülerce & Abrahamson (2011)

In the first instance, before beginning work to derive a V/H model specifically for Groningen, the approach will be to review the many available models for applicability to the Groningen data (Figure 8.18). A review of V/H ground-motion models is provided by Bommer *et al*. (2011). Among those currently available, the most promising are likely to be those by Gülerce & Abrahamson (2011) and by Akkar *et al*. (2014c) because they both include soil non-linearity, which is likely to be important in the Groningen field. Additionally, both of these studies provide correlations between V/H ratios and the period-to-period variability in the horizontal motions to enable the generation of vertical CMS. Since we will have V_S profiles (Section 8.4), in addition to the standard GMPE based V/H models we could also use the V_{sQWL} approach of Poggi *et al*. (2012) that can account for the resonant frequencies due to impedance contrasts.



8.18. Ratios of the vertical spectral accelerations to the geometric mean horizontal spectral ordinates for the recordings of the **M** 3.6 Huizinge earthquake of 16 August 2012

8.9. Vector predictions of durations of Groningen ground motions

In Chapter 7 of this report, a preliminary model for the 5-75% AI-based significant duration was developed through comparison of existing GMPEs with the durations of the recordings in the Groningen database. Using this model and the Version 1 GMPEs for the prediction of spectral accelerations, and invoking some assumptions regarding the variability of the predicted durations, the correlation coefficients between residuals of Sa(T) and $D_{S5\text{-}75}$ were calculated and compared with the model of Bradley (2011). While all of this was an acceptable and fit-for-purpose solution for the Version 1 risk model, a key part of the Version 2 GMPE development will be the derivation completely new GMPEs for the vector prediction of durations conditioned on spectral accelerations. As was noted in Chapter 7,

the first stage of this work is to find a definition of the ground-motion duration that is best suited to the Groningen records, which may or may not be one of the standard definitions currently in widespread use. Since there is no requirement for the Groningen GMPEs to be transportable, a duration measure may be selected that is appropriate only for this application.

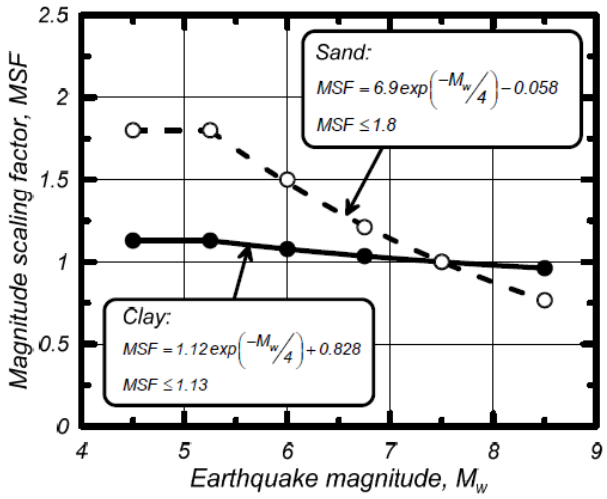
Once the duration definition is selected, the next step will be to develop a GMPE for the prediction of this parameter, as a function of magnitude, distance, site classification and any other parameters found to exert an influence. The development of the GMPE for duration is likely to benefit appreciably from the full waveform modelling, which can capture the elongating effect on the motions of the travel paths related to refraction and reflection by the high-velocity layers above, and maybe also below, the gas reservoir. Once the duration GMPE has been developed, the correlation function between the residuals of the duration and of the spectral accelerations at various periods will be calculated in order to enable the vector prediction of acceleration-duration pairs.

8.10. GMPEs for numbers of equivalent cycles of motions

In addition to quantifying the hazard due to ground shaking, and the influence on site amplification on this hazard, the Groningen seismic hazard and risk model will also include a quantification of the hazard from earthquake-induced liquefaction. This work, like the quantification of site response effects (Section 8.4), is being developed in collaboration with Deltares, with additional expert input and guidance provided by Professor Russell Green of Virginia Tech. Liquefaction hazard assessment approaches currently in wide use (*e.g*., Boulanger & Idriss, 2014) are generally based on the calculation of a factor of safety defined by the ratio of the seismic demand of the imposed shaking to the seismic resistance of the soil deposits. The seismic demand is represented by the cyclic stress ratio (CSR), which is a function of the PGA and the number of cycles of motion. The number of cycles of motion enters the calculation via a magnitude scaling factor (MSF) that adjusts the CSR for the difference in the expected number of cycles of motion at the reference magnitude of **M** 7.5 to the magnitude of the earthquake being considered. The number of cycles is generally based on the approach proposed by Seed *et al.* (1975). One of the key challenges in the development of a liquefaction hazard model for the Groningen field is related to the extension of the current approaches to the small-to-moderate earthquake magnitudes expected to dominate the hazard and risk assessments. This is because little attention has been paid to liquefaction due to small earthquakes—although when this has been observed it has attracted some attention (*e.g*., Holzer *et al*., 2010)—and it has often been assumed that the MSF can be held constant for events smaller than **M** 5.5 (Figure 8.19).

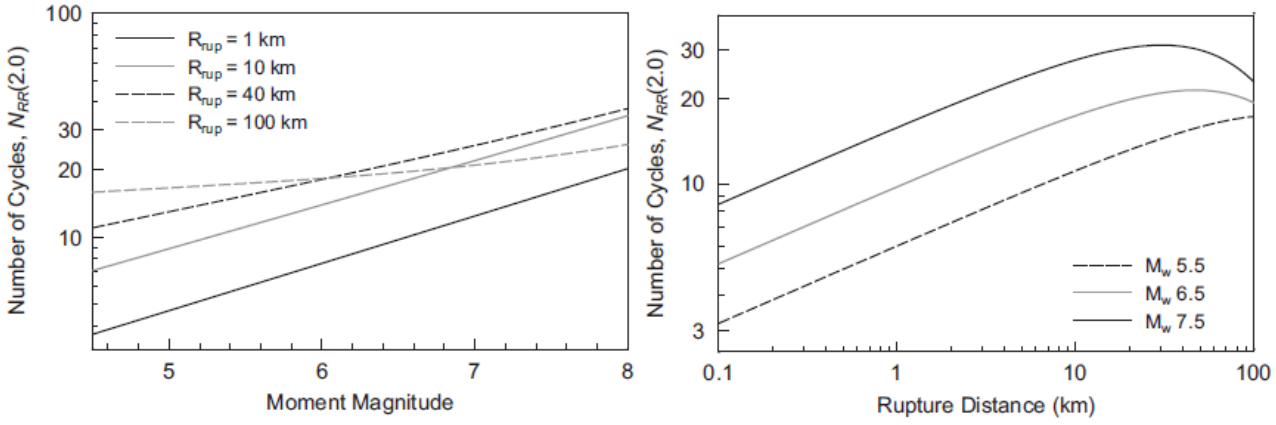
There are actually a large number of different approaches that may be used to estimate the equivalent number of effective cycles in an acceleration time-series (Hancock & Bommer, 2005), all of which are surprisingly poorly correlated with measures of duration (Bommer *et al*., 2006). As with durations, the first stage of the work may be to explore

which of these definitions is best suited both to the Groningen records and to the assessment of liquefaction.



8.19. Magnitude scaling factors for use in liquefaction assessment (Boulanger & Idriss, 2014)

Once the most suitable measure of effective number of cycles is selected, a key part of adapting liquefaction assessment techniques to the small-magnitude Groningen earthquakes is likely to be the development of GMPEs for the prediction of the number of cycles as a function of magnitude, distance and other independent variables (Figure 8.20). There could also be significant value in developing vector predictions of PGA and the number of cycles; the correlation of the residuals of the two parameters is generally not considered in current liquefaction assessment methods.



8.20. Predicted median values of number of cycles of motion using the so-called rainflow-counting definition (Stafford & Bommer, 2009)

References

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

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

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

Aki, K. (1967). Scaling law of seismic spectrum. *Journal of Geophysical Research* **72**, 1217-1231.

Akkar, S. & J.J. Bommer (2006). Influence of long-period filter cut-off on elastic spectral displacements. *Earthquake Engineering & Structural Dynamics* **35**(9), 1145-1165.

Akkar, S., M.A. Sandıkkaya & J.J. Bommer (2014a). Empirical ground-motion models for point- and extended-source crustal earthquake scenarios in Europe and the Middle East. *Bulletin of Earthquake Engineering* **12**(1), 359-387. *Erratum*: **12**(1), 389-390.

Akkar, S., M.A. Sandıkkaya, M. ¸Senyurt, A. Azari Sisi, B.Ö. Ay, P. Traversa, J. Douglas, F. Cotton, L. Luzi, B. Hernandez & S. Godey (2014b). Reference database for seismic ground-motion in Europe (RESORCE). *Bulletin of Earthquake Engineering* **12**(1), 311-339.

Akkar, S., M.A. Sandıkkaya & B.Ö, Ay (2014c). Compatible ground-motion prediction equations for damping scaling factors and vertical-to-horizontal spectral amplitude ratios for the broader European region. *Bulletin of Earthquake Engineering* **12**(1), 517-547.

Akkar, S., Ö. Kale, E. Yenier & J.J. Bommer (2011). The high-frequency limit of usable response spectral ordinates from filtered analogue and digital strong-motion accelerograms. *Earthquake Engineering & Structural Dynamics* **40**(12), 1387-1401.

Al Atik, L., N.A. Abrahamson, J.J. Bommer, F. Scherbaum, F. Cotton & N. Kuehn (2010). The variability of ground-motion prediction models and its components. *Seismological Research Letters* **81**(5), 783-793.

Allen, T.I. (2012). *Stochastic ground motion prediction equations for southeastern Australian earthquakes using updated source and attenuation parameters*. Record 2012/69, Geoscience Australia, Canberra, Australia.

Ancheta, R.D. R.B. Darragh, J.P. Stewart, E. Seyhan, W.J. Silva, B.S.-J. Chiou, K.E. Wooddell, R.W. Graves, A.R. Kottke, D.M. Boore, T. Kishida & J.L. Donahue (2014). NGA-West2 database. *Earthquake Spectra* **30**(3), 989-1005.

Anderson, J.G. (2000). Expected shape of regressions for ground-motion predictions for ground-motion parameters on rock. *Bulletin of the Seismological Society of America* **90**(6B), 543-552.

Anderson, J.G. & J.N. Brune (1999). Probabilistic seismic hazard assessment without the ergodic assumption. *Seismological Research Letters* **70**, 19-28.

Anderson, J.G. & S.E. Hough (1984). A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies. *Bulletin of the Seismological Society of America* **74**, 1969-1993.

Atkinson, G.M. (1993). Notes on ground motion parameters for eastern North America – duration and h/v ratio. *Bulletin of the Seismological Society of America* **83**, 587-596.

Atkinson, G.M. (2006). Single-station sigma. *Bulletin of the Seismological Society of America* **96**, 446-455.

Atkinson, G.M. & I. Beresnev (1997). Don't call it stress drop. *Seismological Research Letters* **68**(1), 3-4.

Atkinson, G.M., J.J. Bommer & N.A. Abrahamson (2014). Alternative approaches to modeling epistemic uncertainty in ground motion in probabilistic seismic-hazard analysis. *Seismological Research Letters* **85**(6), 1141-1144.

Atkinson, G. M. & D. M. Boore (2006). Earthquake ground-motion prediction equations for eastern North America. *Bulletin of the Seismological Society of America* **96**, 2181-2205.

Atkinson, G.M. & M. Morrison (2009). Observations on regional variability in ground-motion amplitude for small-to-moderate magnitude earthquakes in North America. *Bulletin of the Seismological Society of America* **99**(4), 2393-2409.

Baker, J.W. (2011). Conditional mean spectrum: tool for ground motion selection. *ASCE Journal of Structural Engineering* **137**, 322-331.

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

Bal, I.E, J.J. Bommer, P.J. Stafford, H. Crowley & R. Pinho (2010). The influence of geographical resolution of urban exposure data in an earthquake loss model for Istanbul. *Earthquake Spectra* **26**(3), 619-634.

Baltay, A.S. & T.C. Hanks (2014). Understanding the magnitude dependence of PGA and PGV in NGA-West 2 data. *Bulletin of the Seismological Society of America* **104**(6), 2851-2865.

Bazzurro, P. & C.A. Cornell (2004a). Ground-motion amplification in nonlinear soil sites with uncertain properties. *Bulletin of the Seismological Society of America* **94**, 2090-2109.

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

Beyer, K. & J.J. Bommer (2006). Relationships between median values and aleatory variabilities for different definitions of the horizontal component of motion. *Bulletin of the Seismological Society of America* **94**(4A), 1512-1522. *Erratum*: 2007, **97**(5), 1769.

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

Bommer, J. J. & S. Akkar (2012). Consistent source-to-site distance metrics in ground-motion prediction equations and seismic source models for PSHA. *Earthquake Spectra* **28**(1), 1-15.

Bommer, J.J., S. Akkar & Ö. Kale (2011). A model for vertical-to-horizontal response spectral ratios for Europe and the Middle East. *Bulletin of the Seismological Society of America* **101**(4), 1783-1806.

Bommer, J.J., K.J. Coppersmith, R.T. Coppersmith, K.L. Hanson, A. Mangongolo, J. Neveling, E.M. Rathje, A. Rodriguez-Marek, F. Scherbaum, R. Shelembe, P.J. Stafford & F.O. Strasser (2015). A SSHAC Level 3 probabilistic seismic hazard analysis for a new-build nuclear site in South Africa. *Earthquake Spectra* **31**(2), 661-698.

Bommer J. & B. Dost (2014). *Version 0 ground-motion prediction equations*. 16 June 2014, 17 pp.

Bommer, J.J., B. Dost & M. Ntinalexis (2014). *Ground-motion records from the Groningen field*. 28 November 2014, 195 pp.

Bommer, J.J., J. Douglas & F.O. Strasser (2003). Style-of-faulting in ground motion prediction equations. *Bulletin of Earthquake Engineering* **1**(2), 171-203.

Bommer, J.J., J. Hancock & J.E. Alarcón (2006). Correlations between duration and number of cycles of earthquake ground motion. *Soil Dynamics & Earthquake Engineering,* **26**(1), 1-13.

Bommer, J.J., G. Magenes, J. Hancock & P. Penazzo (2004). The influence of strong-motion duration on the seismic response of masonry structures. *Bulletin of Earthquake Engineering* **2**(1), 1-26.

Bommer, J.J. & A. Martinez-Pereira (1999). The effective duration of earthquake strong motion. *Journal of Earthquake Engineering* **3**, 2, 127-172.

Bommer, J.J., P.J. Stafford & J.E. Alarcón (2009). Empirical equations for the prediction of the significant, bracketed and uniform duration of earthquake ground motion. *Bulletin of the Seismological Society of America* **99**(6), 3217-3233.

Bommer, J.J., P.J. Stafford, J.E. Alarcón & S. Akkar (2007). The influence of magnitude range on empirical ground-motion prediction. *Bulletin of the Seismological Society of America* **97**(6), 2152-2170.

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

Boore, D.M. (2001). Effect of baseline corrections on displacements and response spectra for several recordings of the 1999 Chi-Chi, Taiwan, earthquake. *Bulletin of the Seismological Society of America* **91**(5), 1199–211.

Boore, D.M. (2003). Simulation of ground motion using the stochastic method. *Pure & Applied Geophysics* **160**, 635-676.

Boore, D.M. (2005a). *SMSIM – Fortran programs for simulating ground motions from earthquakes: Version 2.3—A revision of OFR 96-80*. US Geological Survey Open-File Report 00-509, 55 pp.

Boore, D.M. (2005b). Erratum to 'Equations for estimating horizontal response spectra and peak acceleration from Western North American earthquakes: a summary of recent work'. *Seismological Research Letters* **76**, 368-369.

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

Boore, D.M. (2010). Orientation-independent, nongeometric-mean measures of seismic intensity from two horizontal components of motion. *Bulletin of the Seismological Society of America* **100**(4), 1830-1835.

Boore, D.M. & J.J. Bommer (2005). Processing strong-motion accelerograms: needs, options and consequences. *Soil Dynamics & Earthquake Engineering* **25**(2), 93-115.

Boore, D.M., J.F.Gibbs, W.B. Joyner, J.C. Tinsley & D.J. Ponti (2003). Estimated ground motion from the 1994 Northridge, California, earthquake at the site of the Interstate 10 and La Cienega

Boulevard bridge collapse, West Los Angeles, California. *Bulletin of the Seismological Society of America* **93**(6), 2737-2751.

Boore, D.M. & E.M. Thompson (2014). Path durations for use in the stochastic-method simulations of ground motions. *Bulletin of the Seismological Society of America* **104**(5), 2541-2552.

Boore, D.M., Watson-Lamprey, J. & Abrahamson, N.A. (2006). Orientation-independent measures of ground motion. *Bulletin of the Seismological Society of America* **94**(4A), 1502-1511.

Boulanger, R.W. & I.M. Idriss (2014). *CPT and SPT based liquefaction triggering procedures*. Report No. UCD/CGM-14/01, Center for Geotechnical Modeling, Department of Civil and Environmental Engineering, University of California at Davis, April, 138 pp.

Bourne, S.J., S.J. Oates, J. van Elk & D. Doornhof (2014). A seismological model for earthquakes induced by fluid extraction from a subsurface reservoir. *Journal of Geophysical Research Solid Earth* **119**, doi: 10.1002/201JB011663.

Bourne, S.J., S.J. Oates, J.J. Bommer, B. Dost, J. van Elk & D. Doornhof (2015). A Monte Carlo method for probabilistic hazard assessment of induced seismicity due to conventional natural gas production. *Bulletin of the Seismological Society of America* **105**(3), 1721-1738.

Boyd, O.S., D. McNamara, S. Hartzell & G. Choy (2015). A comparison of stress drop between tectonic and potentially induced earthquakes in the CEUS (abstract). *Seismological Research Letters* **86**(2B), 690.

Bradley, B.A. (2011). Correlation of significant duration with amplitude and cumulative intensity measures and its use in ground motion selection. *Journal of Earthquake Engineering* **15**(6), 809-832.

Brune, J.N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes. *Journal of Geophysical Research* **75**(26), 4997-5009. Correction (1970), **76**(20), 5002.

Campbell, K.W. (1997). Empirical near-source attenuation relationships for horizontal and vertical components of peak ground acceleration, peak ground velocity, and pseudo-absolute acceleration response spectra. *Seismological Research Letters* **68**(1), 154-179.

Chandramohan, R., J.W. Baker & G.G. Deierlein (2015). Quantifying the influence of ground motion duration on structural collapse capacity using spectrally equivalent records. *Earthquake Spectra*, *in press.*

Chiou, B., R. Darragh, N. Gregor & W. Silva (2008). NGA project strong-motion database. *Earthquake Spectra* **24**(1), 23-44.

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

Chiou, B., R. Youngs, N. Abrahamson & K. Addo (2010). Ground-motion attenuation model for small-to-moderate shallow crustal earthquakes in California and its implications on regionalization of ground-motion prediction models. *Earthquake Spectra* **26**(4), 907-926.

Cotton, F., R. Archuleta & M. Causse (2013). What is sigma of the stress drop? *Seismological Research Letters* **84**(1), 42-48.

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.

Crowley, H., P.J. Stafford & J.J. Bommer (2008). Can earthquake loss models be validated using field observations? *Journal of Earthquake Engineering* **12**(7), 1078-1104.
Cramer, C.H. (2015). Are ENA potentially induced earthquakes different from natural earthquakes? (abstract). *Seismological Research Letters* **86**(2B), 690.

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

Douglas, J., B. Edwards, V. Convertito, N. Sharma, A. Tramelli, D. Kraaijpoel, B. Mena Cabrera, N. Maercklin & C. Troise (2013). Predicting ground motion from induced earthquakes in geothermal areas. *Bulletin of the Seismological Society of America* **103**(3), 1875–1897.

Douglas, J. & D.M. Boore (2010). High-frequency filtering of strong-motion records. *Bulletin of Earthquake Engineering* **9**, 395-409.

Douglas, J. & P. Jousset (2011). Modeling the difference in ground-motion magnitude-scaling in small and large earthquakes. *Seismological Research Letters* **82**(4), 504-508.

Edwards, B. & D. Fäh (2013a). Measurements of stress parameter and site attenuation from recordings of moderate to large earthquakes in Europe and the Middle East. *Geophysical Journal International* **194**, 1190-1202.

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

Edwards, B., A. Rietbrock, J. J. Bommer & B. Baptie (2008). The acquisition of source, path, and site effects from micro-earthquake recordings using *q* tomography: Application to the United Kingdom. *Bulletin of the Seismological Society of America* **98**, 1915-1935.

EPRI (2006). *Truncation of the Lognormal Distribution and Value of the Standard Deviation for Ground Motion Models in the Central and Eastern United States*. EPRI Report 1013015, Electric Power Research Institute, Palo Alto, California.

Esposito, S. & I. Iervolino (2011). PGA and PGV spatial correlation models based on European multievent datasets. *Bulletin of the Seismological Society of America* **101**(5), 2532-2541.

Esposito, S. & I. Iervolino (2012). Spatial correlation of spectral acceleration in European data. *Bulletin of the Seismological Society of America* **102**(6), 2781-2788.

Fukushima, Y. (1996). Scaling relations for strong ground motion prediction models with M^2 terms. *Bulletin of the Seismological Society of America* **86**(2), 329-336.

Goda, K. & G.M. Atkinson (2010). Intraevent spatial correlation of ground-motion parameters using SK-net data. *Bulletin of the Seismological Society of America* **100**(6), 3055-3067.

Gregor N., N.A. Abrahamson, G.M. Atkinson, D.M. Boore, Y. Bozorgnia, K.W. Campbell, B.S.-J. Chiou, I.M. Idriss, R. Kamai, E. Seyhan, W. Silva, J.P. Stewart & R. Youngs. (2014). Comparison of NGA-West2 GMPEs. *Earthquake Spectra* **30**(3), 1179-1197.

Grünthal, G., R. Wahlström & D. Stromeyer (2009). The unified catalogue of earthquakes in central, northern, and northwestern Europe (CENEC)—updated and expanded to the last millennium. *Journal of Seismology* **13**, 517-541.

Gülerce, Z. & N.A. Abrahamson (2011). Site-specific design spectra for vertical ground motion. *Earthquake Spectra* **27**(4), 1023-1047.

Hancock, J. & J.J. Bommer (2005). The effective number of cycles of earthquake ground motion. *Earthquake Engineering & Structural Dynamics* **34**(6), 637-664.

Hanks, T.C. & R.K. McGuire (1981). The character of high-frequency ground motion. *Bulletin of the Seismological Society of America* **71**(6), 2071-2095.

Herrmann, R. B. (1985). An extension of random vibration theory estimates of strong ground motion to large distances. *Bulletin of the Seismological Society of America* **75**, 1447-1453.

Holzer, T.L., A.S. Jayko, E. Hauksson, J.P.B. Fletcher, T.E. Noce, M.J. Bennett, C.M. Dietel, K.W. Hudnut (2010). Liquefaction caused by the 2009 Olancha, California (USA), **M**5.2 earthquake. *Engineering Geology* **116**, 184-188.

Hough, S.E. (2014). Shaking from injection-induced earthquakes in the Central and Eastern United States. *Bulletin of the Seismological Society of America* **104**(5), 2619-2626.

Iervolino, I., G. Manfredi & E. Cosenza (2006). Ground motion duration effects on nonlinear seismic response. *Earthquake Engineering & Structural Analyses* **35**(1), 21-38.

Jayaram, N. & J.W. Baker (2009). Correlation model for spatially distributed ground motion intensities. *Earthquake Engineering & Structural Dynamics* **38**(15), 1687-1708.

Joyner, W. B. & D. M. Boore (1981). Peak horizontal acceleration and velocity from strong motion records including records from the 1979 Imperial Valley, California, earthquake. *Bulletin of the Seismological Society of America* **71**, 2011-2038.

Kempton, J.J. & J.P. Stewart (2006). Prediction equations for significant duration of earthquake ground motions considering site and near-source effects. *Earthquake Spectra* **22**(4), 985-1013.

Kottke, A.R. & E.M. Rathje (2008). *Technical Manual for Strata*. PEER Report 2008/10, Pacific Earthquake Engineering Research Center, University of California at Berkeley, February, 84 pp.

Kottke, A.R. & E.M. Rathje (2013). Comparisons of time series and random-vibration theory site-response methods. *Bulletin of the Seismological Society of America* **103**(3), 2111-2127.

Kruiver, P.P., G. de Lange, A. Wiersma, P. Meijers, M. Korff, J. Peeters, J. Stafleu, R. Harting, R. Dambrink, F. Busschers, J. Gunnink (2015a). *Geological schematisation of the shallow subsurface of Groningen - For site response to earthquakes for the Groningen gas field*. Deltares Report No. 1209862-005-GEO-0004-v5-r, 16 March 2015.

Kruiver, P.P., P. Meijers, G. de Lange, M. Huber, M. Korff, J. Bommer, R. Romijn, E. van Dedem, B. Edwards & A. Rodriguez-Marek (2015b*). Site response Groningen – white paper*. Deltares Report No. 1209862-008-GEO-0001, version 5, 22 May 2015.

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

Miller, A.C. & T.R. Rice (1983). Discrete approximations of probability-distributions. *Management Science* **29**, 352-362.

Neighbors, C., E.S. Cochran, G.M. Atkinson & K.M. Keranen (2015). Low stress drops observed for the 2011 **M**5.7 Prague, Oklahoma, earthquake sequence (abstract). *Seismological Research Letters* **86**(2B), 725.

Papaspiliou, M., S. Kontoe & J.J. Bommer (2012). An exploration of incorporating site response into PSHA – Part I: Issues related to site response analysis methods. *Soil Dynamics & Earthquake Engineering* **42**, 302-315.

Pinheiro, J.C. & D.M. Bates (2004). *Mixed-effects models in S and S-PLUS*. Springer.

Poggi, V., B. Edwards & D. Fäh (2012). Characterizing the vertical-to-horizontal ratio of ground motion at soft-sediment sites. *Bulletin of the Seismological Society of America* **102**, 2741-2756.

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

Rietbrock, A., F. Strasser & B. Edwards (2013). A stochastic earthquake ground-motion prediction model for the United Kingdom. *Bulletin of the Seismological Society of America* **103**, 57-77.

Rodriguez-Marek, A., F. Cotton, N.A. Abrahamson, S. Akkar, L. Al Atik, B. Edwards, G.A. Montalva & H. Dawood (2013). A model for single-station standard deviation using data from various tectonic regions. *Bulletin of the Seismological Society of America* **103**, 3149-3163.

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

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

Sandıkkaya, M.A., S. Akkar & P.-Y. Bard (2013). A nonlinear site-amplification model for the next pan-European ground-motion prediction equations. *Bulletin of the Seismological Society of America* **103**(1), 19-32.

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

Seed, H.B., I.M. Idriss, F. Makdisi & N. Banerjee (1975). *Representation of irregular stress time histories by equivalent uniform stress series in liquefaction analyses*. Report No. EERC 75-29, Earthquake Engineering Research Center, College of Engingeering, University of California at Berkeley, 40 pp.

Stafford, P.J. & J.J. Bommer (2009). Empirical equations for the prediction of the equivalent number of cycles of earthquake ground motion. *Soil Dynamics & Earthquake Engineering* **29**(11/12), 1425-1436.

Stafford, P.J. & J.J. Bommer (2012). *Overview of approaches to derive GMPEs in low-seismicity regions (stochastic, hybrid-empirical, reference empirical).* Project Report for Council for Geoscience, Pretoria, South Africa.

Stafford, P.J., R. Mendis & J.J. Bommer (2008). The dependence of spectral damping ratios on duration and number of cycles. *ASCE Journal of Structural Engineering* **134**(8), 1364-1373.

Strasser, F.O., N.A. Abrahamson & J.J. Bommer (2009). Sigma: issues, insights, and challenges. *Seismological Research Letters* **80**(1), 40-56.

Toro, G.R., N.A. Abrahamson & J.F. Schneider (1997). Model of strong ground motions from earthquakes in central and eastern North America: Best estimates and uncertainties. *Seismological Research Letters* **68**(1), 41-57.

Trifunac, M.D. & A.G. Brady (1975). A study on duration of earthquake ground motion. *Bulletin of the Seismological Society of America* **65**, 581-626.

Viegas, G., R.E. Abercrombie & W.-Y. Kim (2010). The 2002 **M**5 Au Sable Forks, NY, earthquake sequence: Source scaling relationships and energy budget. *Journal of Geophysical Research* **115**, B07310, doi:10.1029/2009JB006799.

Viegas, G., T. Urbanic & A. Baig (2015). Are source characteristics of fluid driven hydraulic fracture induced earthquakes distinct from natural tectonic earthquakes? (abstract). *Seismological Research Letters* **86**(2B), 689.

Wang, M. & T. Takada (2005). Macrospatial correlation model of seismic ground motions. *Earthquake Spectra* **21**(4), 1137-1156.

Watson-Lamprey, J. & D.M. Boore (2007). Beyond Sa_{GMRotI}: Conversion to Sa_{Arb}, Sa_{SN} and Sa_{MaxRot}. *Bulletin of the Seismological Society of America* **97**, 1511-1524.

Wells, D. L. & K. J. Coppersmith (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bulletin of the Seismological Society of America* **84**, 974-1002.

Wong, I., J. Bott, M. Dober & P. Thomas (2015). Examination of the strong motion and broadband data for injection-induced earthquakes in the US (abstract). *Seismological Research Letters* **86**(2B), 692.

Yenier, E. & G.M. Atkinson (2014). Equivalent point-source modeling of moderate-to-large magnitude earthquakes and associated ground-motion saturation effects. *Bulletin of the Seismological Society of America* **104**(3), 1458-1478.

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

APPENDIX I

Spectral Analysis of Groningen Ground Motion Data

Spectral Analysis of Ground-Motion Records from the Groningen Field

Benjamin Edwards

Version: 12th March 2015

Introduction

This document details spectral analysis of recordings from surface accelerographs operated by KNMI in the Groningen field. The spectral analysis of the recordings aims to parameterise, in terms of seismological models, the source characteristics (e.g., how energetic the sources were) of Groningen field earthquakes in addition to the associated wave-field propagation and local-site effects. These seismological models can subsequently be used to calibrate stochastic ground-motion simulations for the prediction of pseudo-spectral acceleration (PSA) and peak acceleration, velocity and displacement (PGA, PGV and PGD respectively) for chosen earthquake scenarios. The simulated ground motions can then be used for guiding or deriving predictive equations (GMPEs) for use in hazard and risk calculations.

The earthquake recordings and recording network are detailed in Bommer et al. (2014). Stations are referred to using the numbers in Table A1.1. At the time of writing KNMI currently recommend not to use records from station FRB2, since they display an unusual monochromatic high-frequency component that is not explained.

Table A1.1: Station numbers and corresponding codes.

Station Number	Station Code	Station Number	Station Code
0001	BAPP	**0017**	MID1
0002	BLOP	**0018**	STDM
0003	BMD2	**0019**	WIN
0004	BONL	**0020**	WSE
0005	BOWW	**0021**	MID3
0006	BUHZ	**0022**	ZAN1
0007	BWIR	**0023**	ZAN2
0008	BWSE	**0024**	BGAR
0009	BZN1	**0025**	BHAR
0010	BZN2	**0026**	BHKS
0011	BMD1	**0027**	BSTD
0013	HKS	**0028**	MD2
0014	FRB2	**0029**	APP
0015	GARST	**0030**	BFB2
0016	KANT	**0031**	BWIN

Seismological Model

The radiation of energy from a seismic source, principally as seismic waves (ground motion), is highly dependent on a large number of variables related to the fault rupture. The subsequent propagation of seismic waves through the crust and near-surface is complex, reflecting the physical properties and geometry of the subsurface. Given perfect knowledge of the source (e.g., described by the moment tensor) and the propagation medium (Green's functions), the resulting ground motion can be determined at any given receiver location. However, such approaches typically involve significant uncertainty and are limited to low frequency motion due to the heterogeneity of small-scale crustal structure (particularly in the upper layers of rock and soil) in addition to computational limitations.

For the purpose of the simulation of surface ground-motion over a range of frequencies important to engineering we can instead rely on the observation that recorded accelerations can be approximated by band-limited Guassian signal with frequency dependent amplitude and of specified duration (Hanks and Mcguire, 1981). This well-observed phenomenon forms the basis of the stochastic ground-motion simulation method. The challenge, and that addressed in this report, is to define a model to describe this process as a function of parameters used to estimate seismic hazard: e.g., earthquake magnitude and source-receiver distance. In practice there are two components that form the input for the stochastic simulation approach: models for (i) predicting the Fourier Acceleration Spectrum (FAS) used to control the frequency content of simulated ground motions and (ii) the shaking duration of earthquake recordings. Modelling the FAS of Groningen earthquake records is the subject of this report.

FAS Regression Approach

Stage 1 – Fitting the high-frequency decay

Initially the Fourier acceleration spectra of the Groningen earthquake recordings are fit in the log-linear acceleration-frequency domain to estimate the slope of the FAS' high-frequency decay, termed kappa (κ) (Anderson and Hough, 1984). From borehole analyses (e.g., Abercrombie and Leary, 1993) it is apparent that the bulk of this observed decay is due to attenuation, and therefore equivalent to the t* parameter (a measure of attenuation along the whole path). Assuming attenuation is due to a path and near-surface component (since the near surface is significantly more heterogeneous than the deeper layers) we can write:

$$t^* = \kappa = \frac{T}{Q} + \kappa_0 \qquad \text{(A1.1)}$$

where Q defines the attenuation in the homogeneous structure, and T is the propagation time. κ_0 is then the path-independent site specific attenuation. Conceptually this defines a layer-over-halfspace model, with the layer depth not defined. The component of t* (or κ) that increases with distance from the source is attributed to Q, while the 'zero-distance' part is attributed to propagation in the upper layers, where body wave paths are mostly vertical due to the velocity reduction.

Two approaches are used to estimate the t* term. The first method used to determine the attenuation along the wave path is a broadband inversion approach (e.g., Masuda and Suzuki, 1982, De Natale *et al.*, 1987, Scherbaum, 1990). The approach we are using is detailed in Edwards et al. (2008), and aims to fit the spectral bandwidth with an earthquake far-field point-source model (Brune, 1970), defined by its source-corner frequency (f_c) and seismic moment (M_0), along with the t* parameter to account for attenuation.

The fitting bandwidth is defined based on the measured signal exceeding the pre-event noise by a factor 3 (i.e., signal-plus-noise to noise ratio, SNR > 3).

The second method we use for estimating t* is the high-frequency approach introduced by Anderson and Hough (1984). The method uses a least-squares approach to fit a line with gradient equal to $-\pi t^*$ to the high-frequency part of the Fourier acceleration spectrum, in lin-log space. In this case the frequency range over which t* is measured is from f_1 (which should lie above the source corner frequency, here 10 Hz) and f_2, which should lie below the frequency at which the noise floor begins (a SNR of 3 is used to define f_2).

Stage 2 – Fitting the broadband FAS

Using the high-frequency decay term, t*, defined in Stage 1 (values are taken from the broadband approach) we refit the FAS in the log-log space to (re-)determine the long-period spectral displacement plateau (related to the seismic moment, M_0) and the source corner-frequency. From the long-period spectral displacement we can determine the seismic moment and average site amplification by correcting for geometrical spreading. However, in the current analysis M_w are provided by KNMI, therefore we only search for the amplification function.

In this report geometrical spreading is accounted for by assuming spherical spreading with amplitudes proportional to 1/R while M_w values are provided by KNMI. Later refinements may take advantage of the full-waveform modelling of the 3D-velocity model for the Groningen area and updates to the magnitudes provided by KNMI. Provided such updates the values determined in this report may change. However, the aim is for a self-consistent model that predicts FAS of Groningen recordings based on currently available data – which approach reported here delivers. Improved information will require repeated analysis to determine parameters consistent with any revised information, such that a fully consistent model is again obtained. This means that mean that FAS predictions (and subsequent ground motion simulations) will not change significantly, but the uncertainty associated to those predictions should be reduced through the inclusion of additional data and information.

Modelling Results

The fitting procedure was applied to the earthquakes detailed in Bommer et al. (2014), adopting the same nomenclature.

Earthquake 1

The 2006 Westeremden earthquake (2003-08-08 05:04) had reported magnitudes of M_L = 3.5 and M_W = 3.4 (Bommer et al., 2015). All four stations that recorded the event were processed. An example of the initial broadband and high-frequency fitting to FAS (geometrical mean of the two horizontal components) in the frequency domain is shown in Figure A1.1. The resulting t^* ($\kappa(R)$) values for the source station pairs are given in Table A1.2 and Figure A1.2. Due to the very short path lengths and the size of the values, they can be assumed to approximate station specific κ_0. The best-fitting corner-frequency for this event was 3.2 Hz, which corresponds to a stress-drop of 23 bar. An example of the log-domain best-fitting models (using fixed t^*, Table A1.2) are shown in Figure A1.3. All FAS and corresponding spectral models are shown in Section 'Additional Material: Spectral Models'.

Table A1.2: t^* values from the broadband and high-frequency fit for EQ-1.

Station #	Station Code	Hypocentral Distance (km)	Epicentral Distance (km)	t^* (s)	lower error (s)	upper error (s)	t^* High Freq. (s)
13	HKS	9.3	8.8	0.0877	0.0086	0.0113	0.0666
21	MID3	4.5	3.3	0.0766	0.0057	0.0075	0.0804
22	ZAN1	6.3	5.6	0.0716	0.0071	0.0091	0.0493
23	ZAN2	5.0	4.0	0.0770	0.0073	0.0093	0.1064

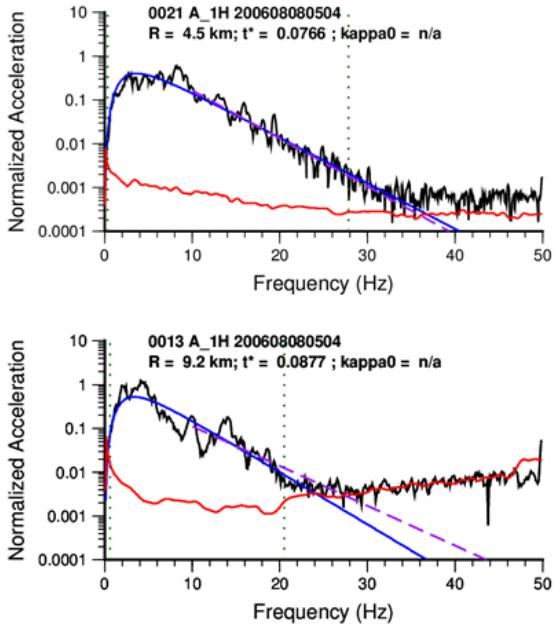


Figure A1.1: best lin-log fit models for geometrical mean of horizontal component recordings for EQ-1. Red: noise, black: earthquake, blue: broadband model; dashed purple: high-frequency model.

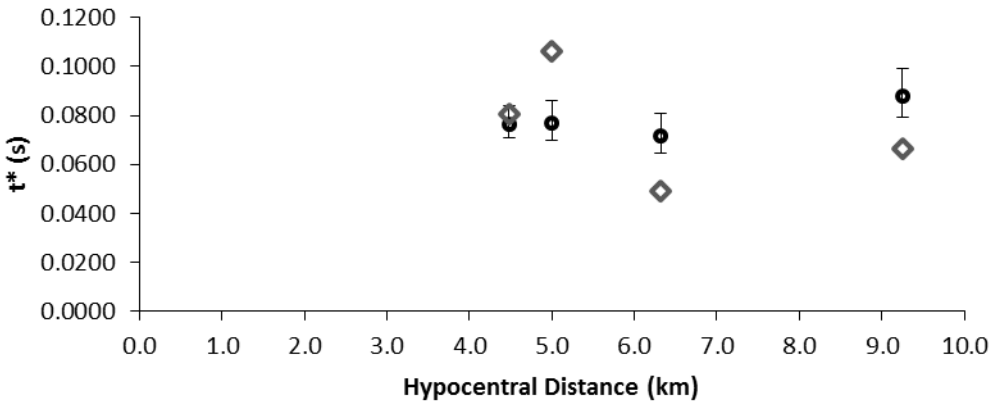


Figure A1.2: t^* values from the broadband (circles) and high frequency fit (diamonds) plotted against distance. Error bars for the broadband fit indicate a 5% tolerance of the minimum misfit.

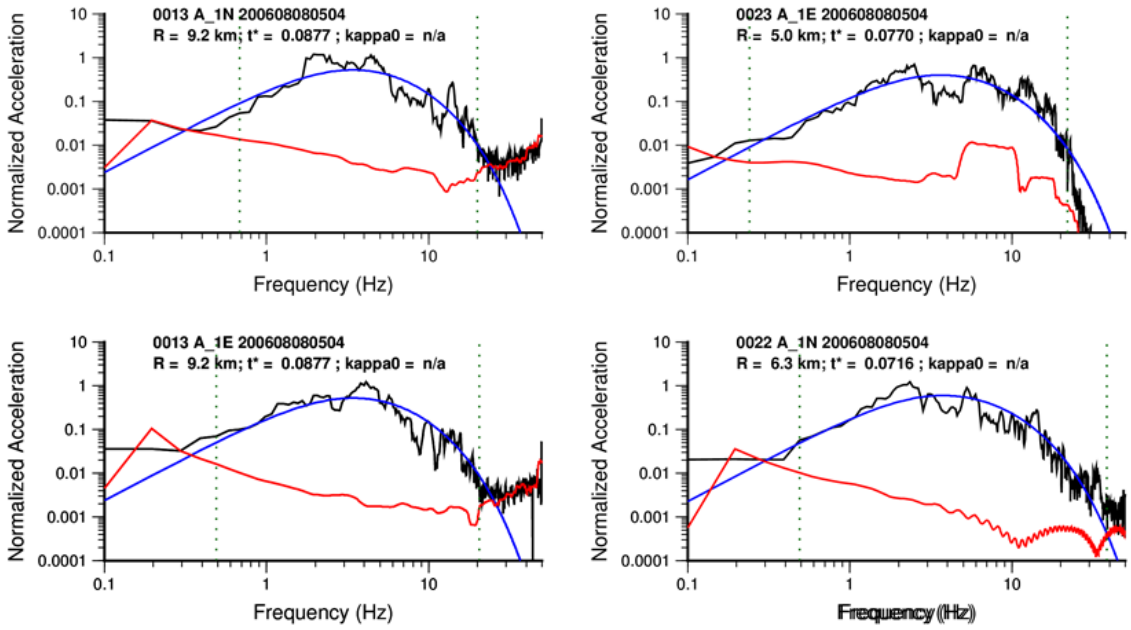


Figure A1.3: log-fit broad-band models for individual horizontal component recordings of EQ-1. Whole path attenuation (t*) is fixed based on the lin-log fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 2

The 2008 Westeremden earthquake (2008-10-30 05:54) had reported magnitudes of M_L = 3.2 and M_W = 3.1 (Bommer et al., 2015). All six stations that recorded the event were processed. An example of the initial broadband and high-frequency fitting to FAS (geometrical mean of the two horizontal components) in the frequency domain is shown in Figure A1.4. The resulting t^* ($\kappa(R)$, Figure A1.5) values for the source station pairs are given in Table A1.3. Due to the very short path lengths and the size of the values, they can be assumed to approximate κ_0. The best-fitting corner-frequency for this event was 3.1 Hz, which corresponds to a stress-drop of 7.4 bar. An example of the log-domain best-fitting models (using fixed t^*, Table A1.3) are shown in Figure A1.6. All FAS and corresponding spectral models are shown in Section 'Additional Material: Spectral Models'.

Table A1.3: t^* values from the broadband and high-frequency fit for EQ-2.

Station #	Station Code	Hypocentral Distance (km)	Epicentral Distance (km)	t^* (s)	lower error (s)	upper error (s)	t^* High Freq. (s)
17	MID1	6.1	5.3	0.0746	0.0081	0.0097	0.1009
21	MID3	6.0	5.2	0.0612	0.0042	0.0053	0.0524
19	WIN	4.4	3.2	0.0475	0.0033	0.0041	0.0340
20	WSE	3.2	1.2	0.0571	0.0037	0.0049	0.0661
22	ZAN1	5.8	5.0	0.0629	0.0047	0.0057	0.0405
23	ZAN2	4.5	3.3	0.0447	0.0041	0.0052	0.0452

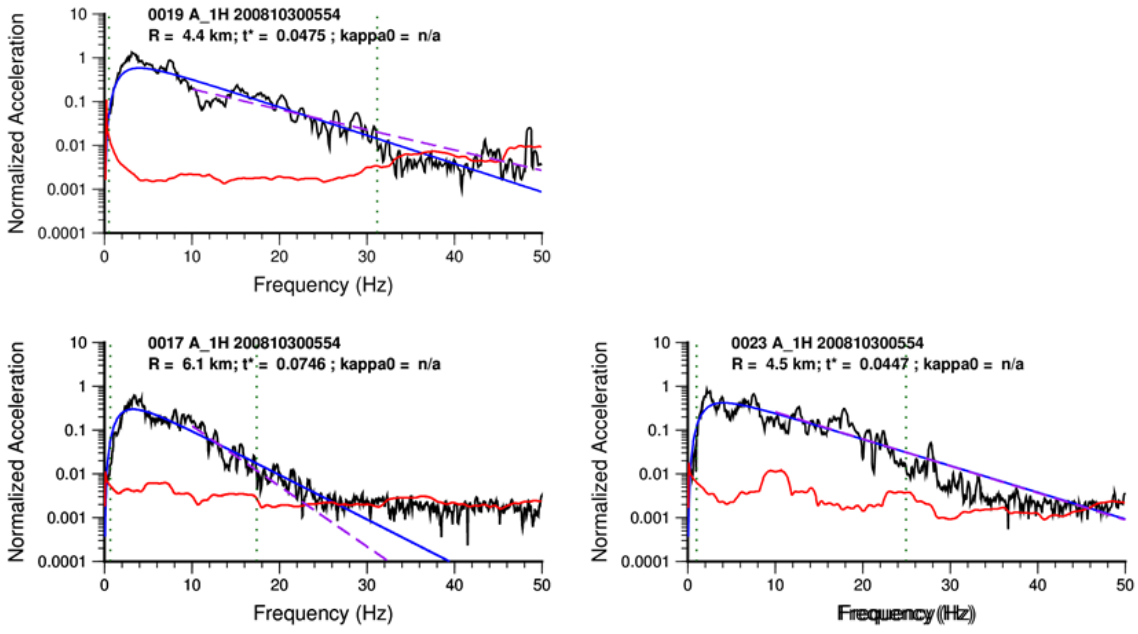


Figure A1.4: best lin-log fit models for geometrical mean of horizontal component recordings for EQ-2. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

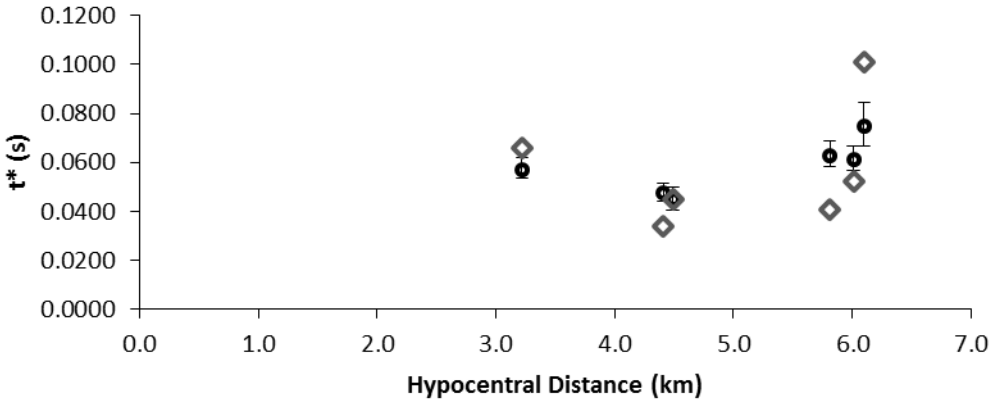


Figure A1.5: t^* values from the broadband (circles) and high frequency fit (diamonds) plotted against distance. Error bars for the broadband fit indicate a 5% tolerance of the minimum misfit.

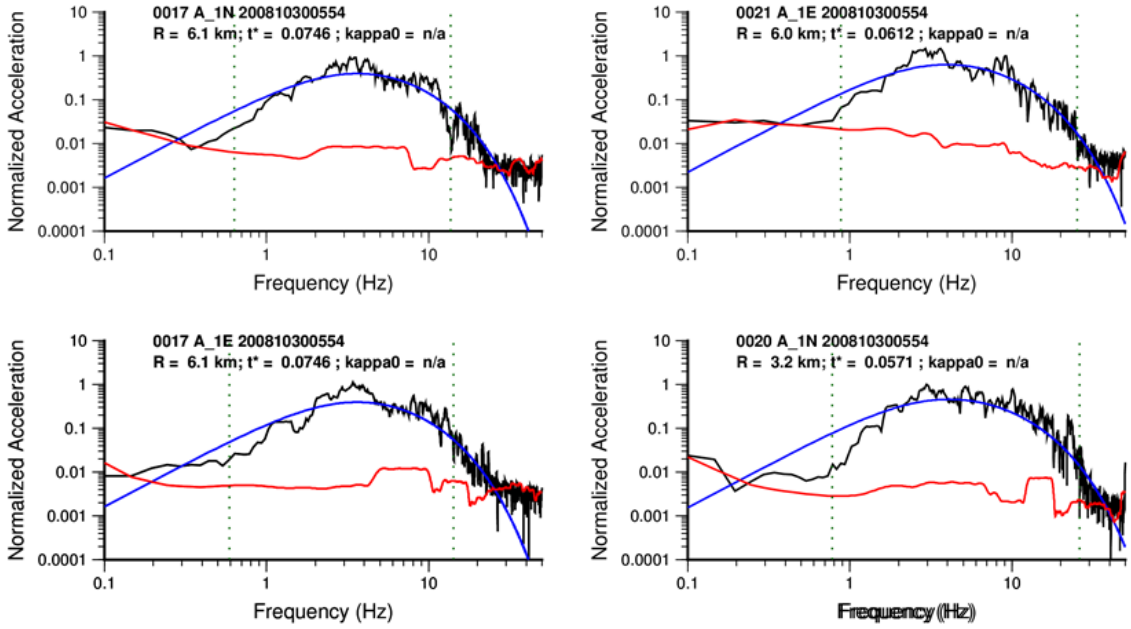


Figure A1.6: log-fit broad-band models for individual horizontal component recordings of EQ-2. Whole path attenuation (t*) is fixed based on the lin-log fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 3

The 2009 Zeerjip earthquake (2009-05-08 05:23) had reported magnitudes of M_L = 3.0 and M_W = 2.9 (Bommer et al., 2015). The event was recorded on 5 stations, all of which were processed. One record (MID1) could not be processed using the high frequency fit due to insufficient signal-to-noise ratio at high-frequency. An example of the initial broadband and high-frequency fitting to FAS (geometrical mean of the two horizontal components) in the frequency domain is shown in Figure A1.7. The resulting t^* ($\kappa(R)$) values for the source station pairs are given in Table A1.4 and in Figure A1.8. Due to the very short path lengths and the size of the values, they can be assumed to approximate station specific κ_0. The best-fitting corner-frequency for this event was 3.6 Hz, which corresponds to a stress-drop of 5.6 bar. An example of the log-domain best-fitting models (using fixed t^*, Table A1.4) are shown in Figure A1.9. All FAS and corresponding spectral models are shown in Section 'Additional Material: Spectral Models'.

Table A1.4: t^* values from the broadband and high-frequency fit for EQ-3.

Station #	Station Code	Hypocentral Distance (km)	Epicentral Distance (km)	t^* (s)	lower error (s)	upper error (s)	t^* High Freq. (s)
17	MID1	8.5	8.0	0.0923	0.0152	0.0160	#N/A
19	WIN	5.5	4.6	0.0511	0.0046	0.0052	0.0539
20	WSE	4.7	3.6	0.0540	0.0049	0.0058	0.0534
22	ZAN1	3.5	1.7	0.0535	0.0067	0.0075	0.0425
23	ZAN2	3.1	0.6	0.0412	0.0036	0.0042	0.0497

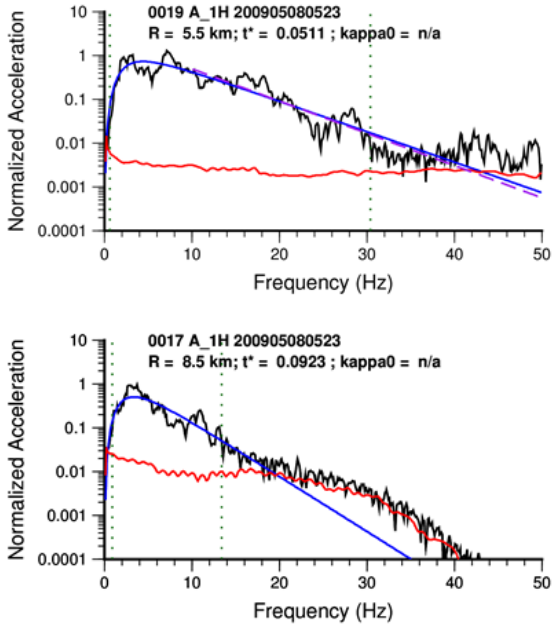


Figure A1.7: best lin-log fit models for geometrical mean of horizontal component recordings for EQ-3. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

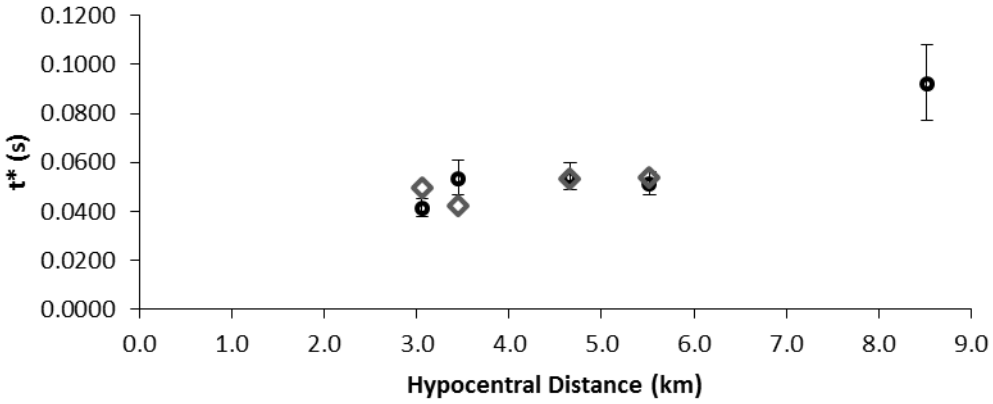


Figure A1.8: t^* values from the broadband (circles) and high frequency fit (diamonds) plotted against distance. Error bars for the broadband fit indicate a 5% tolerance of the minimum misfit.

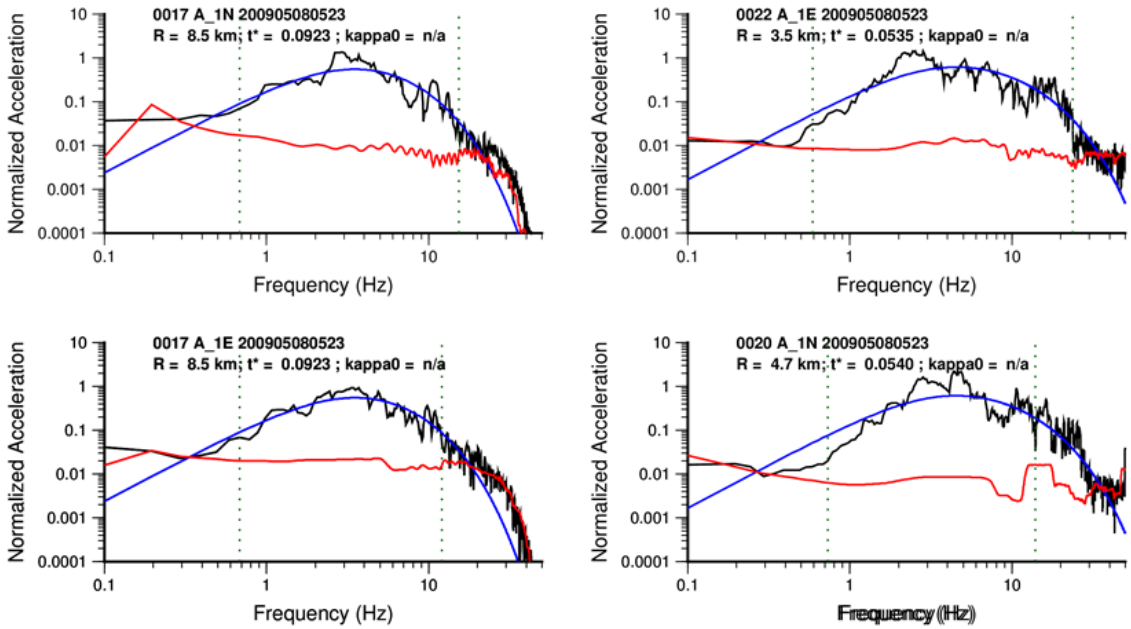


Figure A1.9: log-fit broad-band models for individual horizontal component recordings of EQ-3. Whole path attenuation (t*) is fixed based on the lin-log fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 4

The 2011 Westerwijtwerd earthquake (2011-01-19 19:39) had reported magnitudes of M_L = 2.4 and M_W = 2.7 (Bommer et al., 2015). The event was recorded on 4 stations, all of which were processed. One record (WSE) could not be processed using the high frequency fit due to insufficient signal-to-noise ratio at high-frequency. However, the low SNR seems to be due to an overestimation of the noise-level. An example of the initial broadband and high-frequency fitting to FAS (geometrical mean of the two horizontal components) in the frequency domain is shown in Figure A1.10. The resulting t^* (equivalent to $\kappa(R)$) values for the source station pairs are given in Table A1.5 and Figure A1.11. Due to the very short path lengths and the size of the values, they can be assumed to approximate station specific κ_0. The best-fitting corner-frequency for this event was 9.6 Hz, which corresponds to a stress-drop of 55 bar. An example of the log-domain best-fitting models (using fixed t^*, Table A1.5) are shown in Figure A1.12. All FAS and corresponding spectral models are shown in Section 'Additional Material: Spectral Models'.

Table A1.5: t^* values from the broadband and high-frequency fit for EQ-4. *bound not constrained.

Station #	Station Code	Hypocentral Distance (km)	Epicentral Distance (km)	t^* (s)	lower error (s)	upper error (s)	t^* High Freq. (s)
17	MID1	4.5	3.3	0.0564	0.0094	0.1345*	0.0349
21	MID3	5.0	3.9	0.0466	0.0104	0.1159*	0.0439
18	STDM	4.4	3.2	0.0947	0.0139	0.2142*	0.0793
20	WSE	6.1	5.3	0.0825	0.0184	0.1902*	#N/A

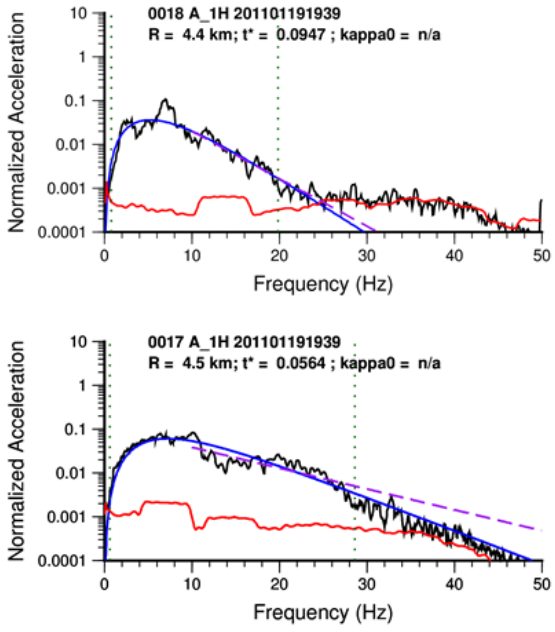


Figure A1.10: best lin-log fit models for geometrical mean of horizontal component recordings for EQ-4. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

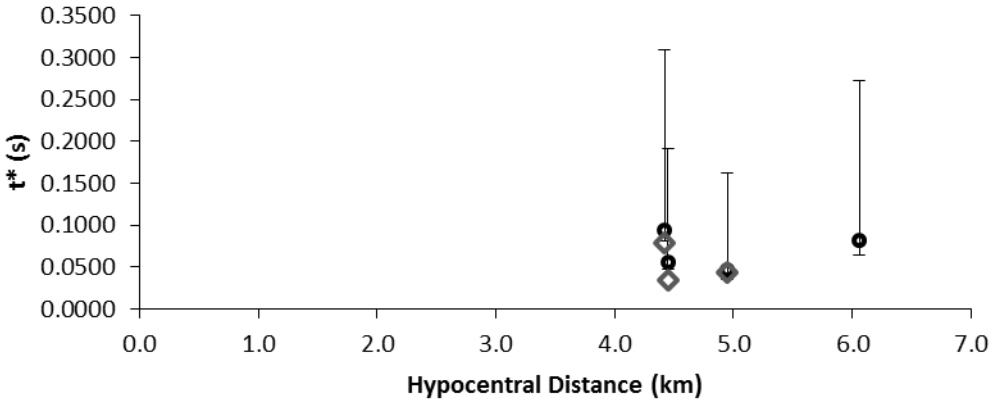


Figure A1.11: t^* values from the broadband (circles) and high frequency fit (diamonds) plotted against distance. Error bars for the broadband fit indicate a 5% tolerance of the minimum misfit. *upper bound not constrained.

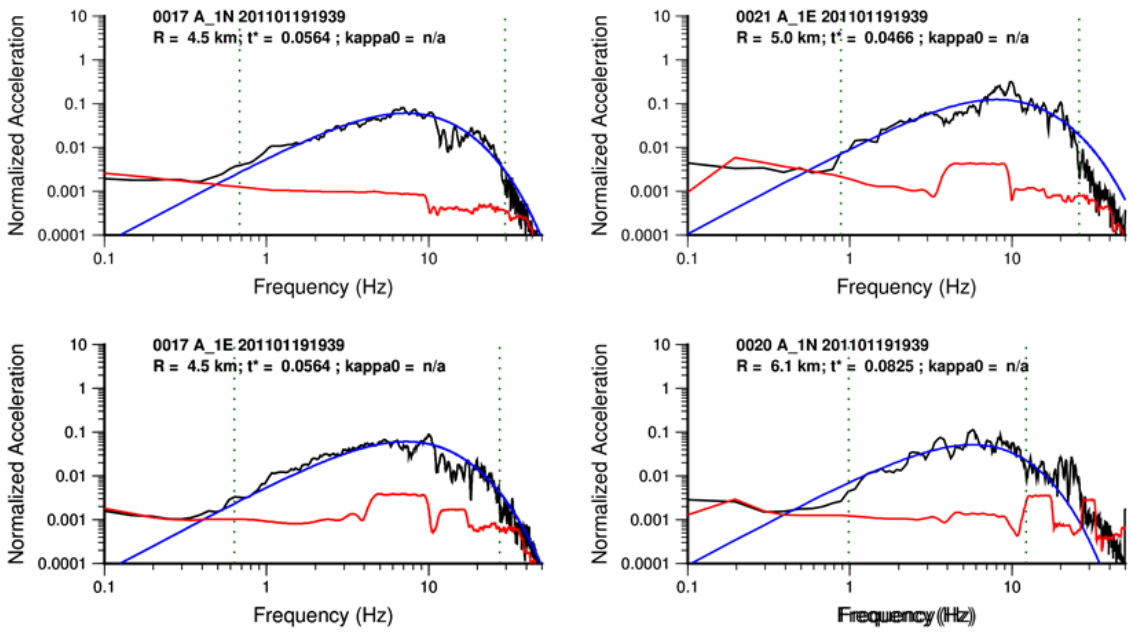


Figure A1.12: log-fit broad-band models for individual horizontal component recordings of EQ-4. Whole path attenuation (t*) is fixed based on the lin-log fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 5

The 2011 Garrelsweer earthquake (2011-06-27 15:48) had reported magnitudes of M_L = 3.2 and M_W = 3.4 (Bommer et al., 2015). The event was recorded on 8 stations, all of which were processed. Two records (WSE and MID1) did not provide results using the high frequency fit due to insufficient signal-to-noise ratio at high-frequency. The low SNR may be due to an overestimation of the noise-level due to the short pre-event data. An example of the initial broadband and high-frequency fitting to FAS (geometrical mean of the two horizontal components) in the frequency domain is shown in Figure A1.13. In this case significant differences are apparent at some stations (e.g., HSK) which are due to a strong peak in the spectrum around 10 – 15 Hz adversely affecting the high-frequency fit. The resulting t^* ($\kappa(R)$, Figure A1.14) values for the source station pairs are given in Table A1.6. The best-fitting corner-frequency for this event was 2.4 Hz, which corresponds to a stress-drop of 9.9 bar. An example of the log-domain best-fitting models (using fixed t^*, Table A1.6) are shown in Figure A1.15. All FAS and corresponding spectral models are shown in Section 'Additional Material: Spectral Models'.

Table A1.6: t^* values from the broadband and high-frequency fit for EQ-5.

Station #	Station Code	Hypocentral Distance (km)	Epicentral Distance (km)	t^* (s)	lower error (s)	upper error (s)	t^* High Freq. (s)
15	GARST	9.3	8.8	0.0558	0.0039	0.0067	0.0673
13	HKS	3.3	1.4	0.0990	0.0066	0.0119	0.0239
16	KANT	12.2	11.8	0.0661	0.0097	0.0152	0.1137
17	MID1	11.3	10.9	0.0828	0.0107	0.0162	#N/A
20	WSE	7.6	7.0	0.0829	0.0125	0.0188	#N/A
22	ZAN1	7.8	7.2	0.0822	0.0049	0.0083	0.0888
23	ZAN2	7.1	6.5	0.0608	0.0037	0.0067	0.1045

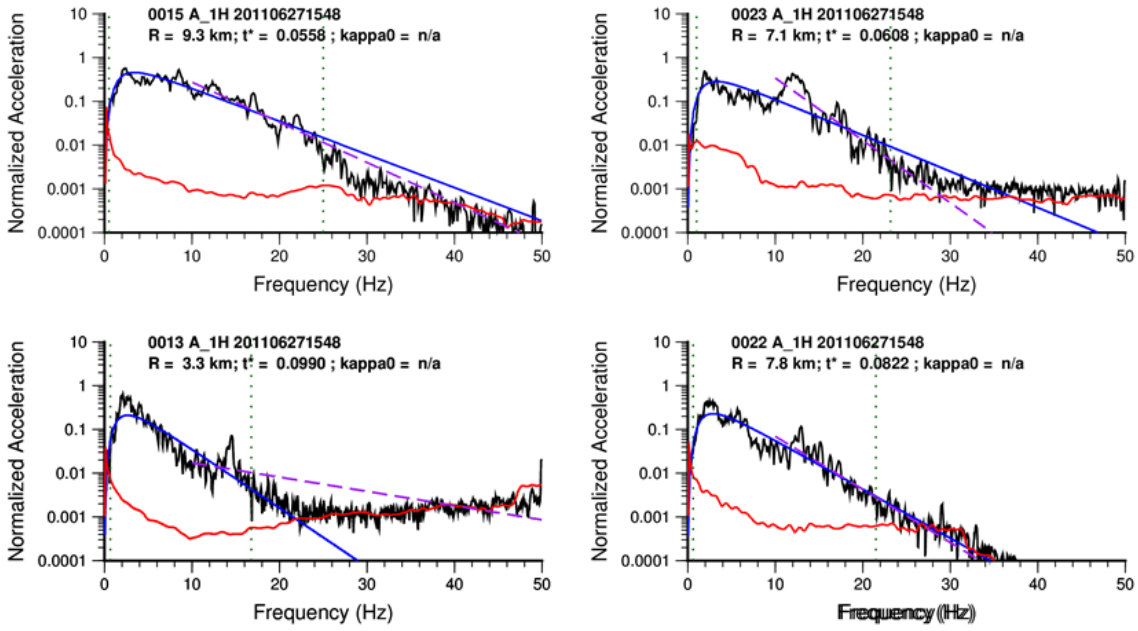


Figure A1.13: best lin-log fit models for geometrical mean of horizontal component recordings for EQ-5. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

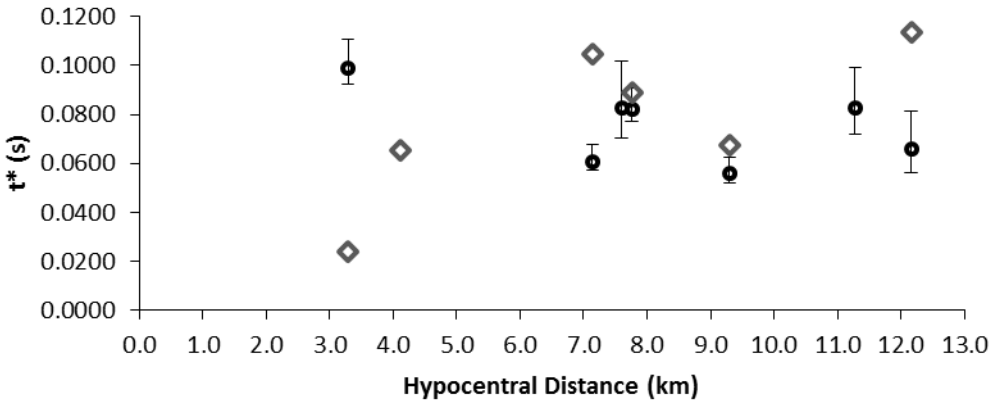


Figure A1.14: t^* values from the broadband (circles) and high frequency fit (diamonds) plotted against distance. Error bars for the broadband fit indicate a 5% tolerance of the minimum misfit. *bound not constrained.

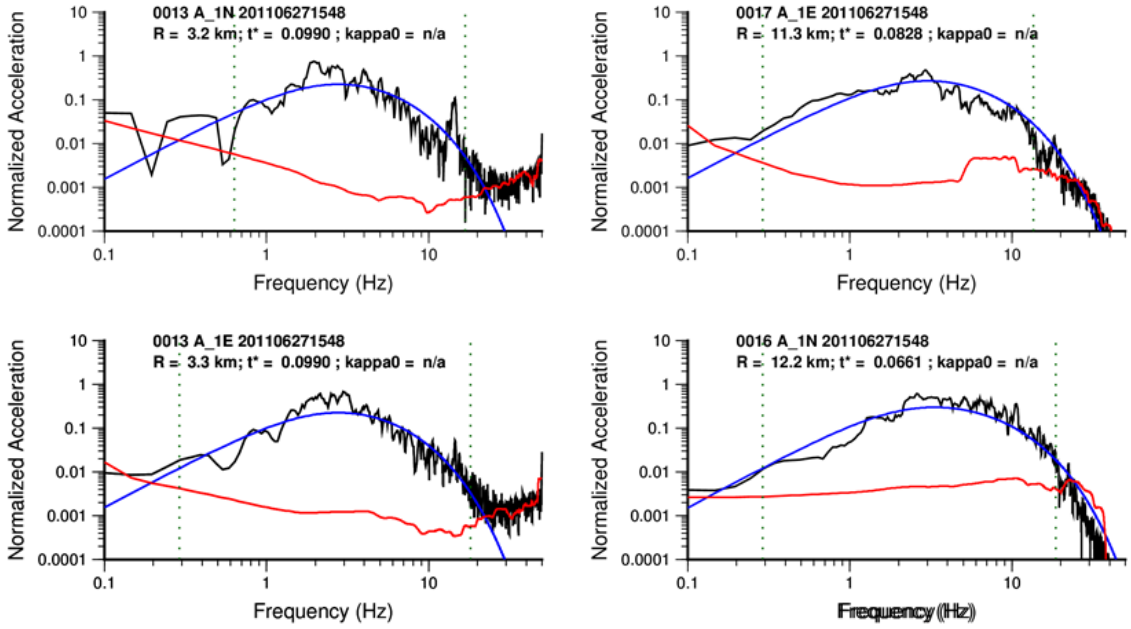


Figure A1.15: log-fit broad-band models for individual horizontal component recordings of EQ-5. Whole path attenuation (t*) is fixed based on the lin-log fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 6

The Huizinge earthquake (2012-08-16 20:30) had reported magnitudes of M_L = 3.6 and M_W = 3.6 (Dost and Kraaijpoel, 2013; Bommer et al, 2015). Seven stations out of the total eight that recorded the event were processed. The excluded station (FRB2) had too short pre-event signal to define a noise estimate and has been previously reported to have an unexplained high-frequency disturbance. An example of the initial broadband and high-frequency fitting to FAS (geometrical mean of the two horizontal components) in the frequency domain is shown in Figure A1.16. The resulting t^* ($\kappa(R)$, Figure A1.17) values for the source station pairs are given in Table A1.7. Due to the very short path lengths and the size of the values, they can be assumed to approximate station specific κ_0. The best-fitting corner-frequency for this event was 3.4 Hz, which corresponds to a stress-drop of 55 bar. An example of the log-domain best-fitting models (using fixed t^*, Table A1.7) are shown in Figure A1.18. All FAS and corresponding spectral models are shown in Section 'Additional Material: Spectral Models'.

Table A1.7: t^* values from the broadband and high-frequency fit for EQ-6.

Station #	Station Code	Hypocentral Distance (km)	Epicentral Distance (km)	t^* (s)	lower error (s)	upper error (s)	t^* High Freq. (s)
15	GARST	4.9	3.9	0.0527	0.0054	0.0072	0.0675
13	HKS	10.1	9.6	0.1074	0.0105	0.0132	0.0702
16	KANT	4.8	3.8	0.0751	0.0056	0.0076	0.0718
17	MID1	3.6	2.0	0.1027	0.0087	0.0114	0.1129
18	STDM	4.8	3.8	0.0798	0.0061	0.0083	0.0769
19	WIN	6.9	6.2	0.0939	0.0070	0.0093	0.0800
20	WSE	3.9	2.6	0.0806	0.0055	0.0076	0.0910

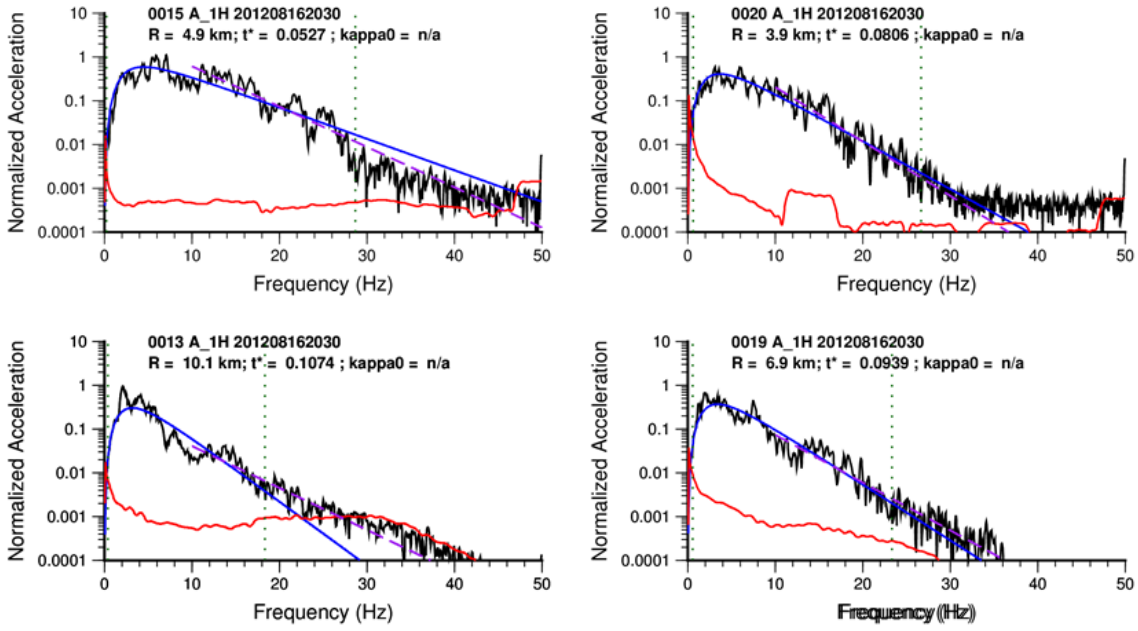


Figure A1.16: best lin-log fit models for geometrical mean of horizontal component recordings for EQ-6. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

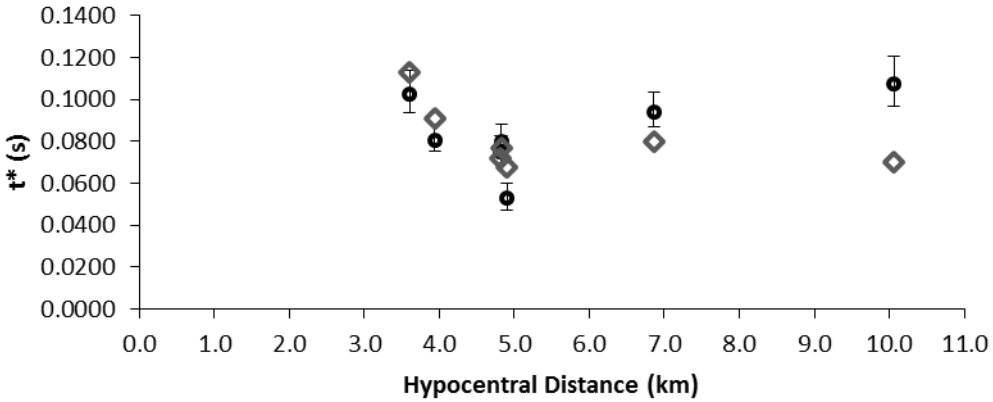


Figure A1.17: t^* values from the broadband (circles) and high frequency fit (diamonds) plotted against distance. Error bars for the broadband fit indicate a 5% tolerance of the minimum misfit.

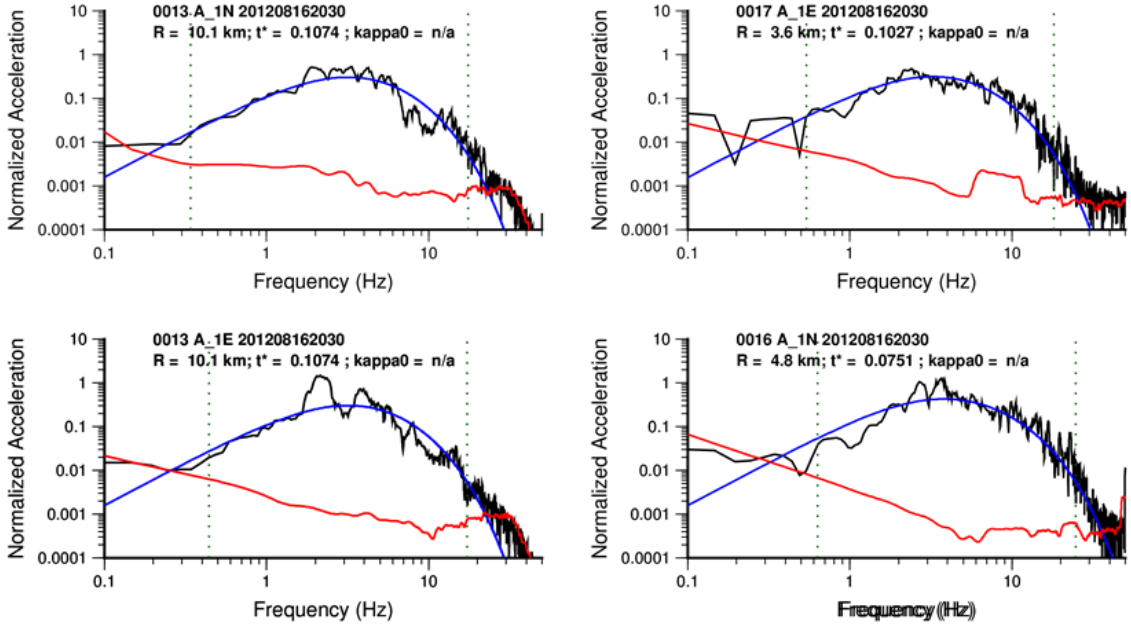


Figure A1.18: log-fit broad-band models for individual horizontal component recordings of EQ-6. Whole path attenuation (t*) is fixed based on the lin-log fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 7

The 2013 Overschild earthquake (2013-01-19 20:10) had reported magnitudes of M_L = 2.4 and M_W = 2.8 (Bommer et al., 2015). The event was recorded on 3 stations, all of which were processed. An example of the initial broadband and high-frequency fitting to FAS (geometrical mean of the two horizontal components) in the frequency domain is shown in Figure A1.19. The resulting t^* ($\kappa(R)$, Figure A1.20) values for the source station pairs are given in Table A1.8. The large difference at station HKS is due to a low-frequency peak in the spectrum affecting the broadband fit. Due to the very short path lengths and the size of the values, they can be assumed to approximate station specific κ_0. The best-fitting corner-frequency for this event was 3.3 Hz, which corresponds to a stress-drop of 3.2 bar. An example of the log-domain best-fitting models (using fixed t^*, Table A1.8) are shown in Figure A1.21. All FAS and corresponding spectral models are shown in Section 'Additional Material: Spectral Models'.

Table A1.8: t^* values from the broadband and high-frequency fit for EQ-7.

Station #	Station Code	Hypocentral Distance (km)	Epicentral Distance (km)	t^* (s)	lower error (s)	upper error (s)	t^* High Freq. (s)
13	HKS	3.1	0.8	0.0660	0.0038	0.0063	0.0413
19	WIN	5.2	4.2	0.0432	0.0031	0.0051	0.0474
20	WSE	9.1	8.6	0.0544	0.0049	0.0076	0.0563

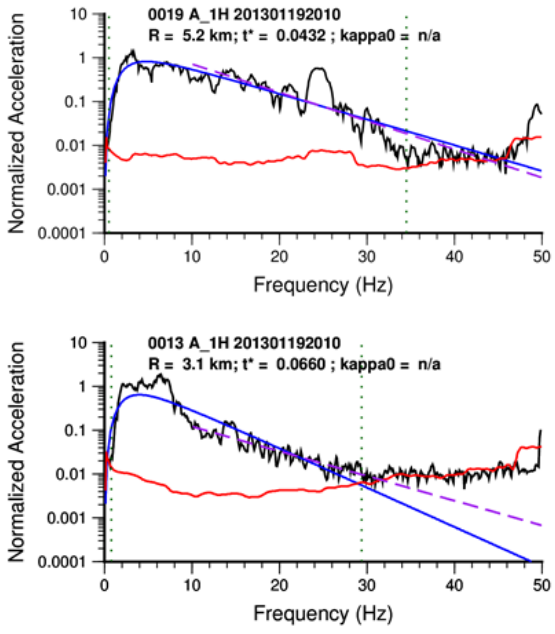


Figure A1.19: best lin-log fit models for geometrical mean of horizontal component recordings for EQ-7. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

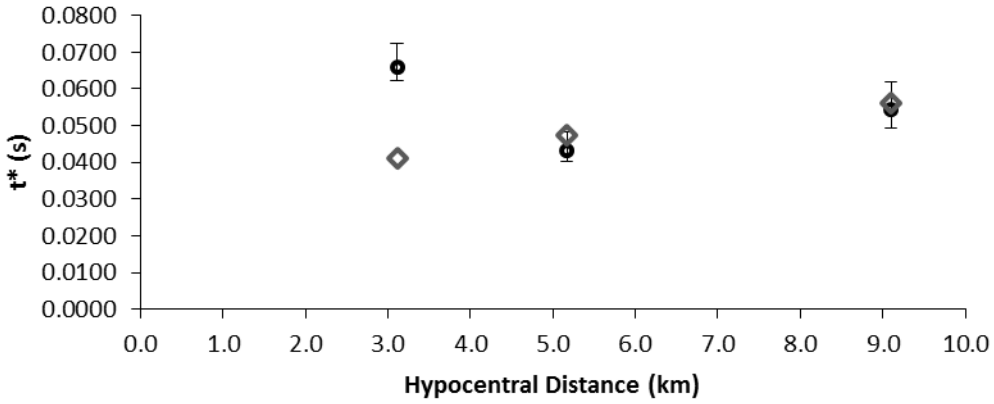


Figure A1.20: t^* values from the broadband (circles) and high frequency fit (diamonds) plotted against distance. Error bars for the broadband fit indicate a 5% tolerance of the minimum misfit.

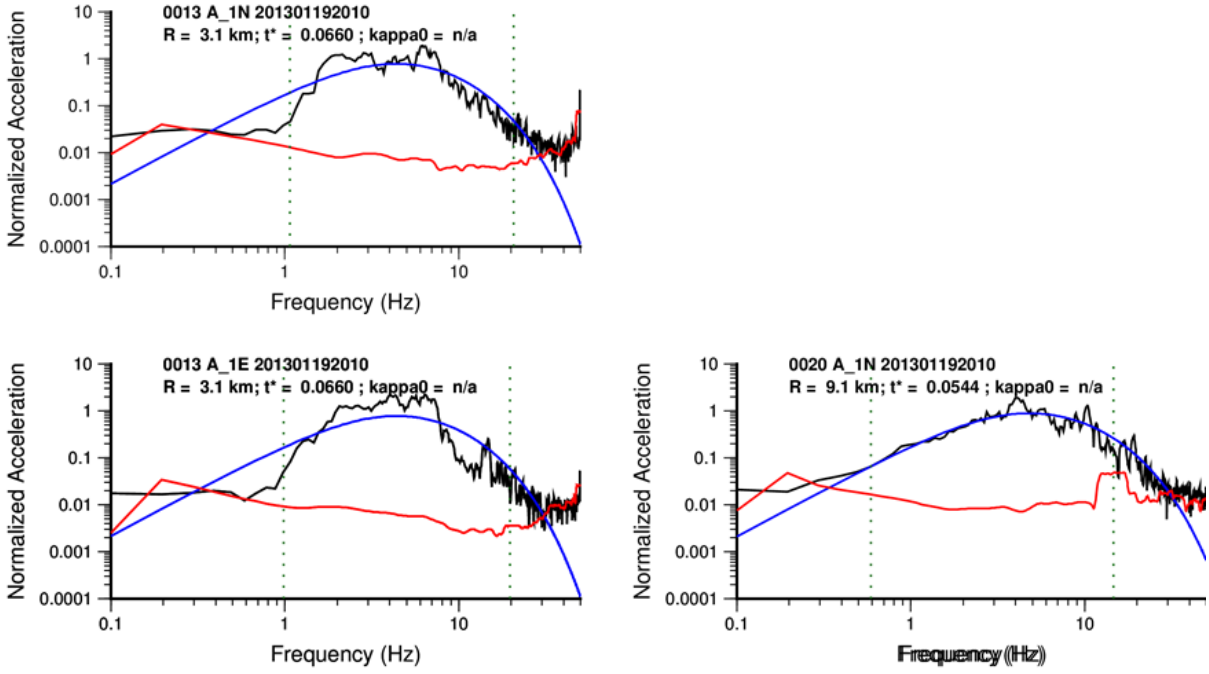


Figure A1.21: log-fit broad-band models for individual horizontal component recordings of EQ-7. Whole path attenuation (t*) is fixed based on the lin-log fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 8

The 2013 Zandeweer earthquake (2013-02-07 23:19) had reported magnitudes of M_L = 3.2 and M_W = 3.2 (Bommer et al., 2015). The event was recorded on 3 stations, all of which were processed. An example of the initial broadband and high-frequency fitting to FAS (geometrical mean of the two horizontal components) in the frequency domain is shown in Figure A1.22. The resulting t^* ($\kappa(R)$, Figure A1.23) values for the source station pairs are given in Table A1.9. Due to the very short path lengths and the size of the values, they can be assumed to approximate station specific κ_0. The best-fitting corner-frequency for this event was 4.1 Hz, which corresponds to a stress-drop of 23 bar. An example of the log-domain best-fitting models (using fixed t^*, Table A1.9) are shown in Figure A1.24. All FAS and corresponding spectral models are shown in Section 'Additional Material: Spectral Models'.

Table A1.9: t^* values from the broadband and high-frequency fit for EQ-8.

Station #	Station Code	Hypocentral Distance (km)	Epicentral Distance (km)	t^* (s)	lower error (s)	upper error (s)	t^* High Freq. (s)
15	GARST	4.9	3.8	0.0553	0.0051	0.0063	0.0602
16	KANT	3.2	1.2	0.0795	0.0053	0.0064	0.0736
20	WSE	6.4	5.6	0.0667	0.0055	0.0068	0.0513

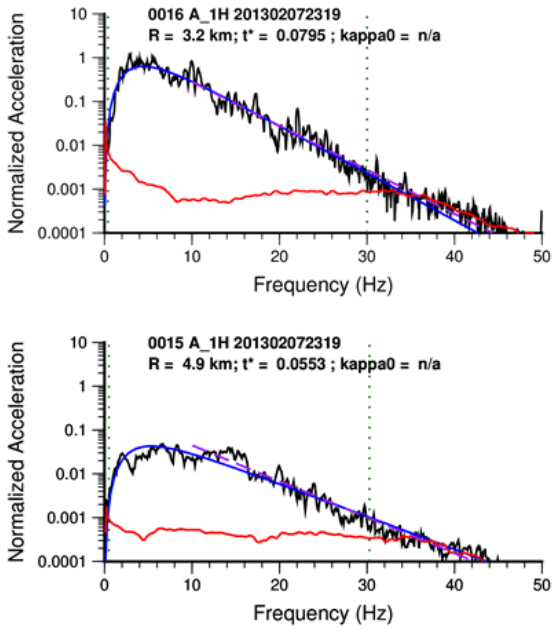


Figure A1.22: best lin-log fit models for geometrical mean of horizontal component recordings for EQ-8. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

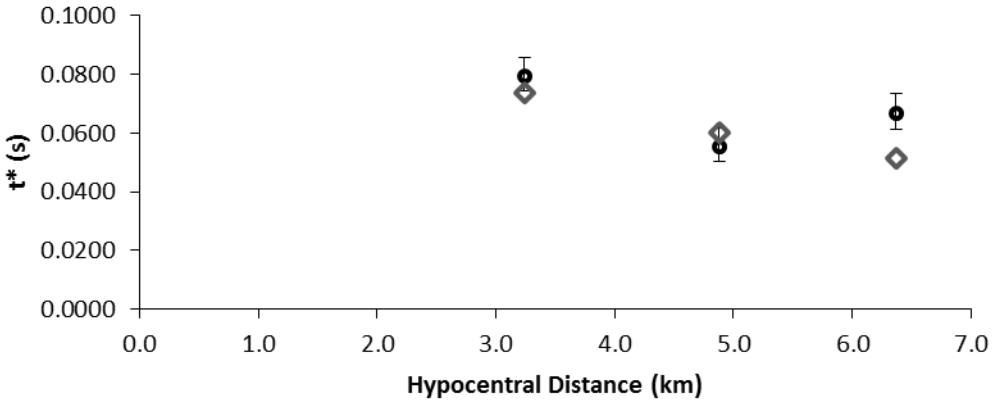


Figure A1.23: t^* values from the broadband (circles) and high frequency fit (diamonds) plotted against distance. Error bars for the broadband fit indicate a 5% tolerance of the minimum misfit.

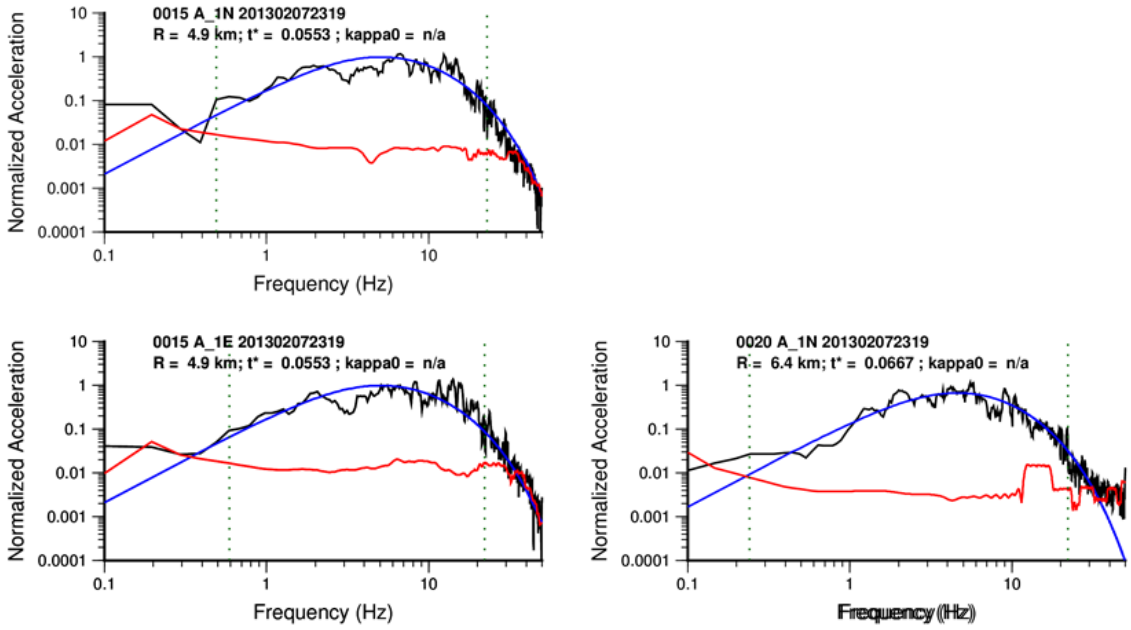


Figure A1.24: log-fit broad-band models for individual horizontal component recordings of EQ-8. Whole path attenuation (t*) is fixed based on the lin-log fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 9

The 2014 Leermens earthquake (2014-02-13 02:13) had reported magnitudes of M_L = 3.0 and M_W = 3.0 (Bommer et al., 2015). The event was recorded on 14 stations, all of which were processed. One record could not be fit using the high frequency method due to insufficient SNR at high frequency. An example of the initial broadband and high-frequency fitting to FAS (geometrical mean of the two horizontal components) in the frequency domain is shown in Figure A1.25. The resulting t^* ($\kappa(R)$, Figure A1.26) values for the source station pairs are given in Table A1.10. Due to the very short path lengths and the size of the values, they can be assumed to approximate station specific κ_0. The best-fitting corner-frequency for this event was 3.5 Hz, which corresponds to a stress-drop of 7.4 bar. An example of the log-domain best-fitting models (using fixed t^*, Table A1.10) are shown in Figure A1.27. All FAS and corresponding spectral models are shown in Section 'Additional Material: Spectral Models'.

Table A1.10: t^* values from the broadband and high-frequency fit for EQ-9.

Station #	Station Code	Hypocentral Distance (km)	Epicentral Distance (km)	t^* (s)	lower error (s)	upper error (s)	t^* High Freq. (s)
1	BAPP	6.6	5.9	0.0617	0.0022	0.0026	0.0583
2	BLOP	4.6	3.5	0.0417	0.0014	0.0017	0.0389
11	BMD1	9.8	9.4	0.0736	0.0047	0.0055	#N/A
3	BMD2	9.5	9.0	0.0537	0.0032	0.0038	0.0446
4	BONL	6.0	5.1	0.0509	0.0023	0.0029	0.0559
5	BOWW	4.1	2.7	0.0570	0.0011	0.0021	0.0541
6	BUHZ	9.6	9.2	0.0599	0.0038	0.0046	0.0490
7	BWIR	4.9	3.8	0.0481	0.0024	0.0029	0.0456
8	BWSE	5.8	5.0	0.0623	0.0027	0.0034	0.0573
10	BZN2	3.5	1.8	0.0466	0.0017	0.0020	0.0521
15	GARST	5.6	4.7	0.0552	0.0049	0.0058	0.0571
19	WIN	6.2	5.4	0.0887	0.0063	0.0075	0.0519

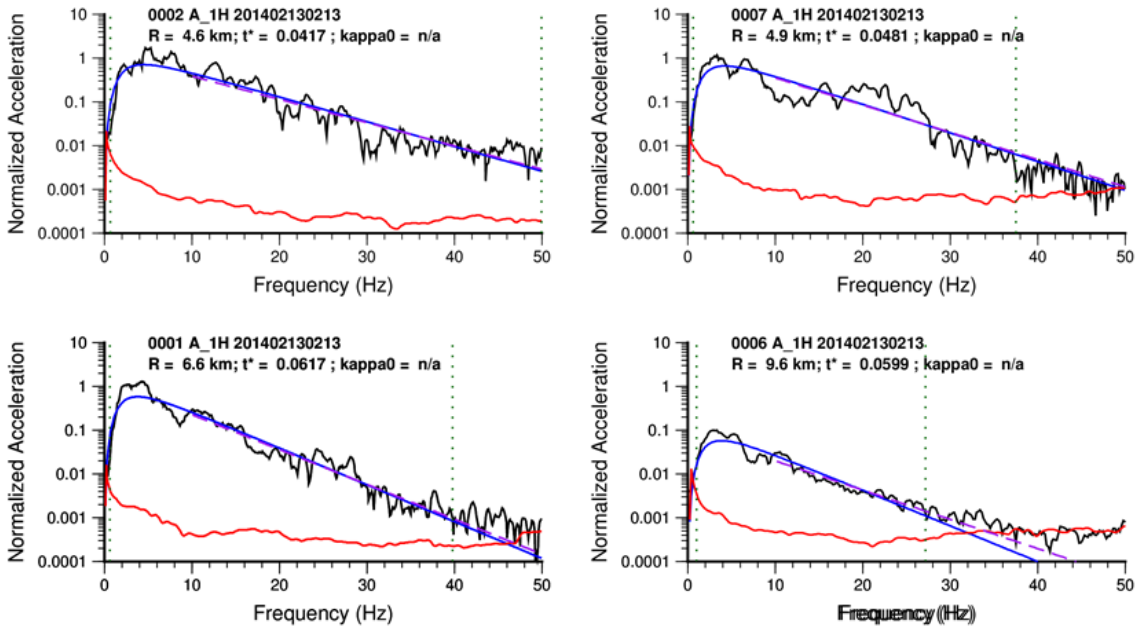


Figure A1.25: best lin-log fit models for geometrical mean of horizontal component recordings for EQ-9. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

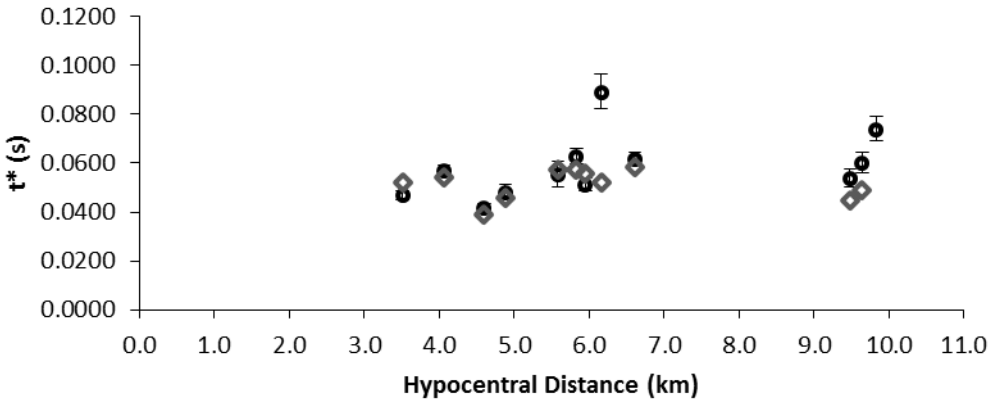


Figure A1.26: t^* values from the broadband (circles) and high frequency fit (diamonds) plotted against distance. Error bars for the broadband fit indicate a 5% tolerance of the minimum misfit. *bound not constrained.

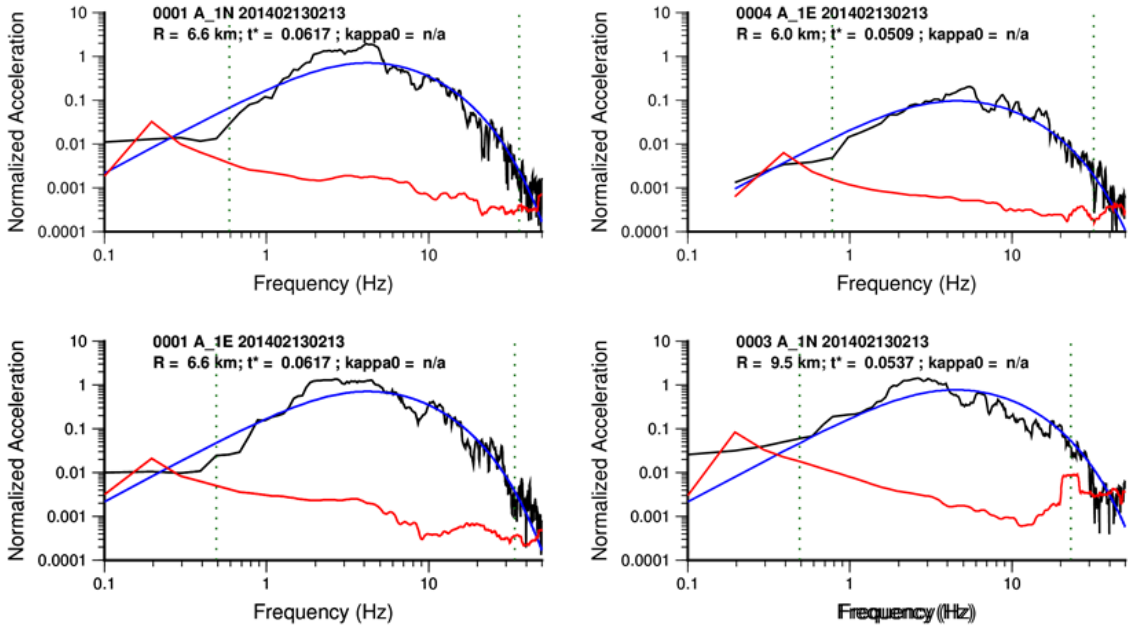


Figure A1.27: log-fit broad-band models for individual horizontal component recordings of EQ-9. Whole path attenuation (t*) is fixed based on the lin-log fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 10

The 2014 Froombosch earthquake (2014-09-01 07:17) had a reported magnitude of M_W = 2.6 (Bommer et al., 2015). The event was recorded on 5 stations, all of which were processed. An example of the initial broadband and high-frequency fitting to FAS (geometrical mean of the two horizontal components) in the frequency domain is shown in Figure A1.28. The resulting t^* ($\kappa(R)$, Figure A1.29) values for the source station pairs are given in Table A1.11. The best-fitting corner-frequency for this event was 1.5 Hz, which corresponds to a relatively low stress-drop of 0.14 bar. An example of the log-domain best-fitting models (using fixed t^*, Table A1.11) are shown in Figure A1.30. All FAS and corresponding spectral models are shown in Section 'Additional Material: Spectral Models'.

Table A1.11: t^* values from the broadband and high-frequency fit for EQ-10.

Station #	Station Code	Hypocentral Distance (km)	Epicentral Distance (km)	t^* (s)	lower error (s)	upper error (s)	t^* High Freq. (s)
1	BAPP	14.2	13.8	0.0950	0.0108	0.0111	0.0713
3	BMD2	20.2	20.0	0.0642	0.0099	0.0107	#N/A
7	BWIR	14.6	14.3	0.1020	0.0128	0.0130	#N/A
10	BZN2	18.5	18.2	0.0446	0.0088	0.0092	#N/A

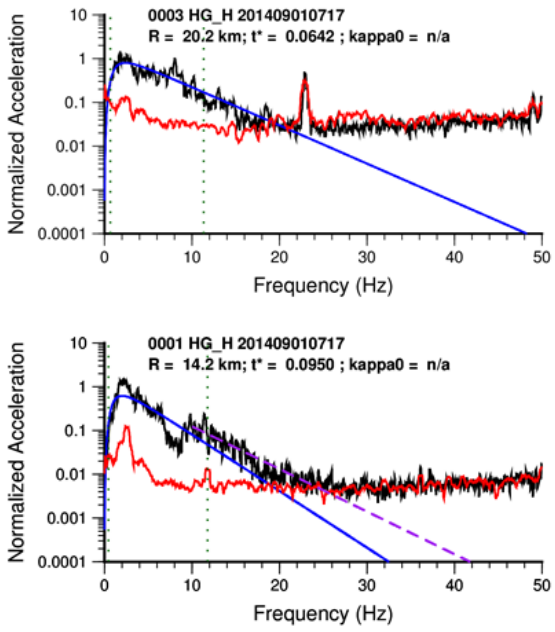


Figure A1.28: best lin-log fit models for geometrical mean of horizontal component recordings for EQ-10. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

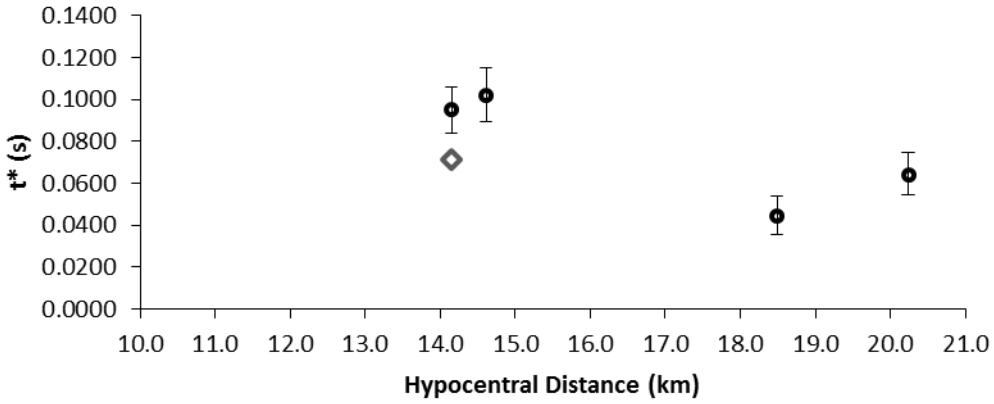


Figure A1.29: t^* values from the broadband (circles) and high frequency fit (diamonds) plotted against distance. Error bars for the broadband fit indicate a 5% tolerance of the minimum misfit. *bound not constrained.

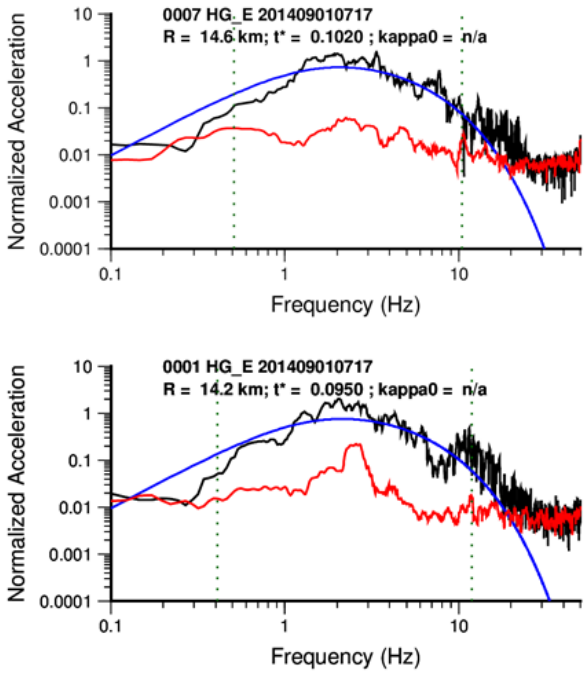


Figure A1.30: log-fit broad-band models for individual horizontal component recordings of EQ-10. Whole path attenuation (t*) is fixed based on the lin-log fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 11

The 2014 Garmerwolde earthquake (2014-09-30 11:42) had a reported magnitude of M_W = 2.8 (Bommer et al., 2015). The event was recorded on 13 stations, 11 of which were processed (one record was missing the station location, one was too noisy). An example of the initial broadband and high-frequency fitting to FAS (geometrical mean of the two horizontal components) in the frequency domain is shown in Figure A1.31. The resulting t^* ($\kappa(R)$, Figure A1.32) values for the source station pairs are given in Table A1.12. The best-fitting corner-frequency for this event was 1.9 Hz, which corresponds to a stress-drop of 0.57 bar. An example of the log-domain best-fitting models (using fixed t^*, Table A1.12) are shown in Figure A1.33. All FAS and corresponding spectral models are shown in Section 'Additional Material: Spectral Models'.

Table A1.12: t^* values from the broadband and high-frequency fit for EQ-11.

Station #	Station Code	Hypocentral Distance (km)	Epicentral Distance (km)	t^* (s)	lower error (s)	upper error (s)	t^* High Freq. (s)
1	BAPP	14.0	13.7	0.1035	0.0061	0.0082	0.0978
24	BGAR	13.3	13.0	0.0618	0.0053	0.0071	0.0803
25	BHAR	5.7	4.8	0.0473	0.0044	0.0060	0.0426
2	BLOP	11.0	10.6	0.0415	0.0071	0.0093	0.0353
4	BONL	17.6	17.3	0.0698	0.0089	0.0115	#N/A
5	BOWW	14.4	14.1	0.0965	0.0045	0.0061	0.0907
27	BSTD	7.3	6.6	0.0670	0.0068	0.0090	0.1034
6	BUHZ	17.0	16.8	0.0767	0.0081	0.0107	#N/A
10	BZN2	13.3	13.0	0.0787	0.0072	0.0094	0.0749
14	FRB2	11.1	10.7	0.0275	0.0028	0.0047	0.0260
28	MD2	11.2	10.7	0.0724	0.0055	0.0074	0.0750

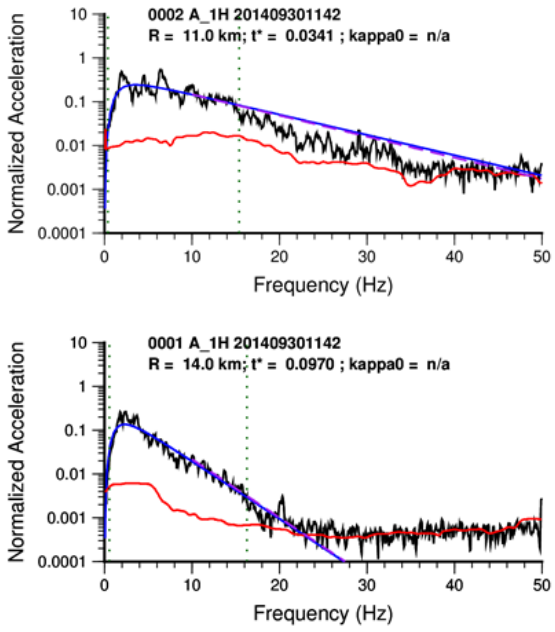


Figure A1.31: best lin-log fit models for geometrical mean of horizontal component recordings for EQ-11. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

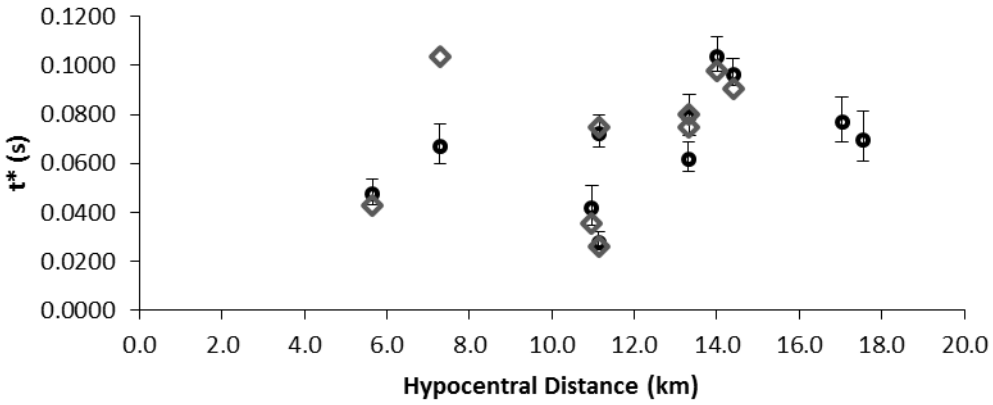


Figure A1.32: t^* values from the broadband (circles) and high frequency fit (diamonds) plotted against distance. Error bars for the broadband fit indicate a 5% tolerance of the minimum misfit. *bound not constrained.

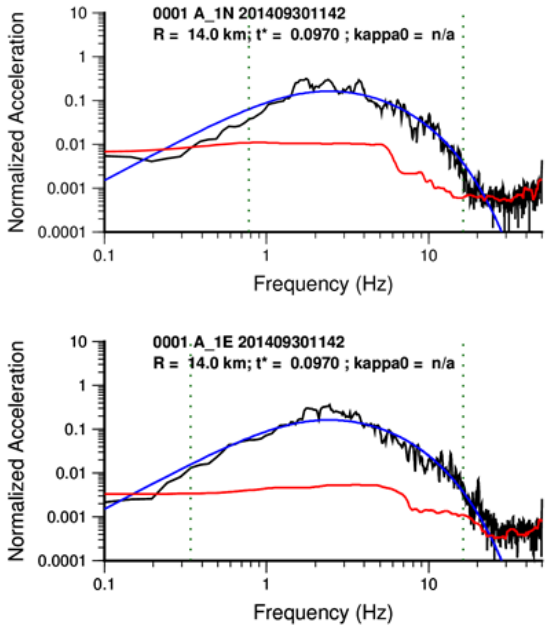


Figure A1.33: log-fit broad-band models for individual horizontal component recordings of EQ-11. Whole path attenuation (t*) is fixed based on the lin-log fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 12

The 2014 Zandeweer earthquake (2014-11-05 01:12) had reported magnitudes of M_L = 2.9 and M_W = 2.9 (Bommer et al., 2015). The event was recorded on 15 stations, all of which were processed. An example of the initial broadband and high-frequency fitting to FAS (geometrical mean of the two horizontal components) in the frequency domain is shown in Figure A1.34. The resulting t^* ($\kappa(R)$, Figure A1.35) values for the source station pairs are given in Table A1.13. Due to the very short path lengths of some records and the size of the values, they can be assumed to approximate station specific κ_0. The best-fitting corner-frequency for this event was 5.7 Hz, which corresponds to a stress-drop of 23 bar. An example of the log-

domain best-fitting models (using fixed t^*, Table A1.13) are shown in Figure A1.36. All FAS and corresponding spectral models are shown in Section 'Additional Material: Spectral Models'.

Table A1.13: t^* values from the broadband and high-frequency fit for EQ-12.

Station #	Station Code	Hypocentral Distance (km)	Epicentral Distance (km)	t^* (s)	lower error (s)	upper error (s)	t^* High Freq. (s)
29	APP	12.7	12.3	0.1340	0.0128	0.0193	0.1215
30	BFB2	21.7	21.5	0.0480	0.0088	0.0144	0.0612
24	BGAR	3.9	2.5	0.0575	0.0065	0.0119	0.0715
25	BHAR	16.4	16.1	0.0635	0.0086	0.0143	0.0718
2	BLOP	7.0	6.3	0.0748	0.0092	0.0151	0.0618
3	BMD2	4.2	3.0	0.0577	0.0054	0.0097	0.0638
4	BONL	6.7	6.0	0.0741	0.0060	0.0108	0.0765
5	BOWW	10.2	9.7	0.0928	0.0101	0.0161	0.0677
27	BSTD	7.5	6.8	0.0722	0.0068	0.0118	0.0765
6	BUHZ	4.9	3.8	0.0808	0.0089	0.0148	0.0897
31	BWIN	8.7	8.2	0.1024	0.0111	0.0176	0.0674
7	BWIR	9.6	9.2	0.0863	0.0093	0.0152	0.0248
8	BWSE	4.9	3.8	0.0666	0.0128	0.0195	0.0554
10	BZN2	6.2	5.4	0.0661	0.0071	0.0121	0.0710
13	HKS	12.0	11.6	0.1230	0.0119	0.0182	0.0695

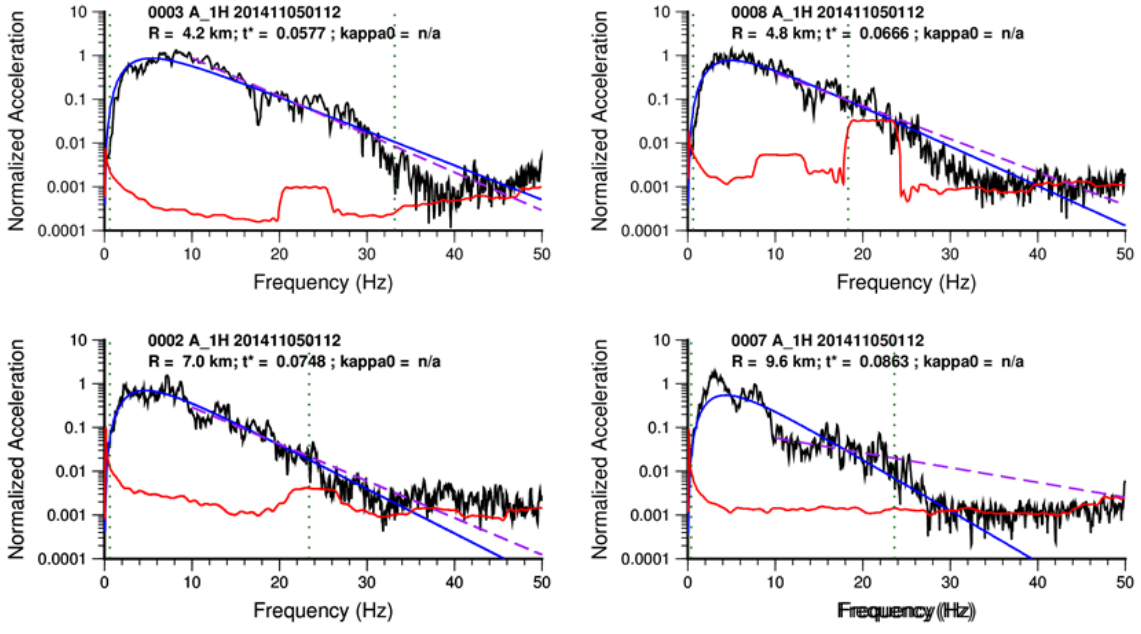


Figure A1.34: best lin-log fit models for geometrical mean of horizontal component recordings for EQ-12. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

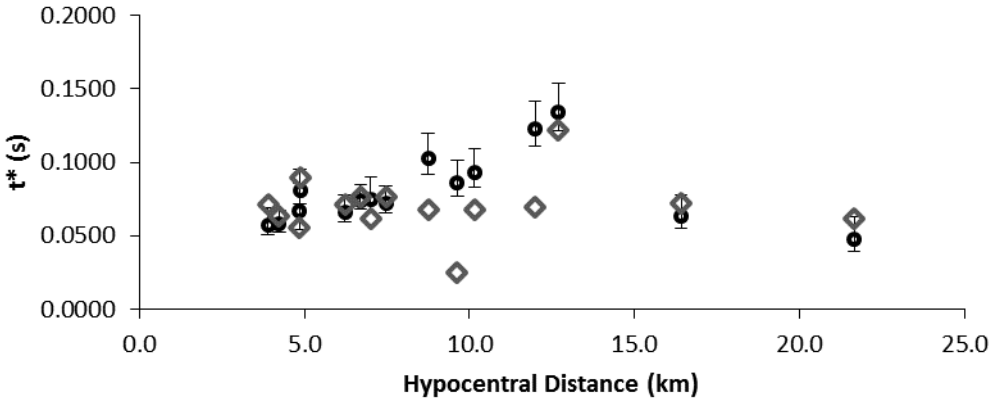


Figure A1.35: t^* values from the broadband (circles) and high frequency fit (diamonds) plotted against distance. Error bars for the broadband fit indicate a 5% tolerance of the minimum misfit.

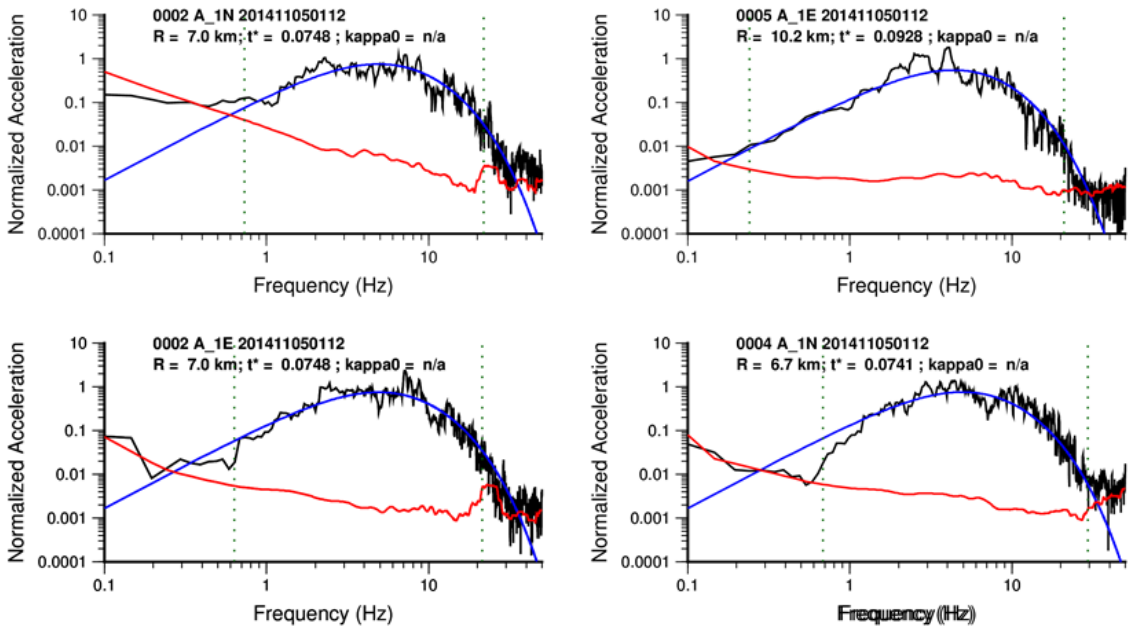


Figure A1.36: log-fit broad-band models for individual horizontal component recordings of EQ-12. Whole path attenuation (t*) is fixed based on the lin-log fit models. Red: noise, black: earthquake, blue: broadband model.

Attenuation: Q and κ_0

Using the measured high-frequency decay parameter (t*) from the processed records Q and κ_0 terms are estimated through a linear regression of equation (A1.1). Due to the limited number of recordings a bootstrap approach was applied to estimate uncertainty. 1000 random realisations of the dataset with replacement were regressed using a least-squares minimisation. For the broadband fit 72 recordings were processed. The fit of equation (A1.1) to the data is shown in Figure A1.37. T in equation (A1.1) is calculated at a reference velocity of 3.5km/s, which is in terms of physical properties, too high for this data – material property Q values will be higher than quoted. Nevertheless for the purposes of simulation the choice of the reference velocity makes no difference as long as it is consistent with the forward simulations. This Q at this reference velocity is referred to as $Q_{3.5}$. The average values for $Q_{3.5}$ and κ_0 were 264 and 0.062s respectively.

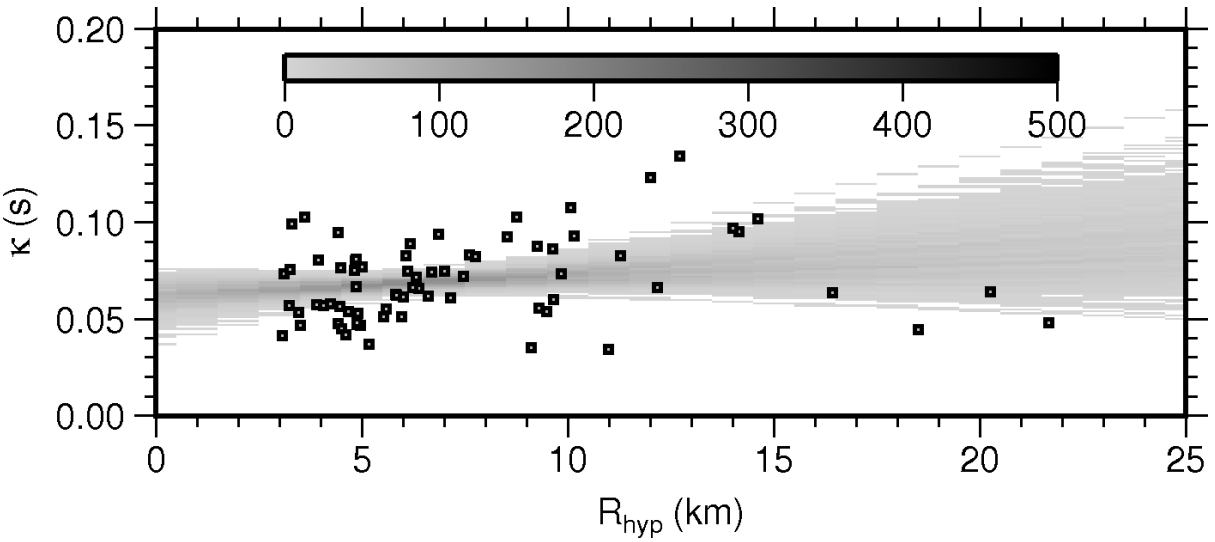


Figure A1.37: broadband fit t* plotted against hypocentral distance. Best-fis from 1000 random samples are shown with grey-shade indicating the density of solutions.

Using 58 t* values obtained from the high-frequency fit (Anderson and Hough, 1984) the average $Q_{3.5}$ was 149 and the average κ_0 was 0.051s. While the high-frequency fit t* values suggest a slightly stronger attenuation (higher Q) they lead to a lower site-specific attenuation (κ_0).

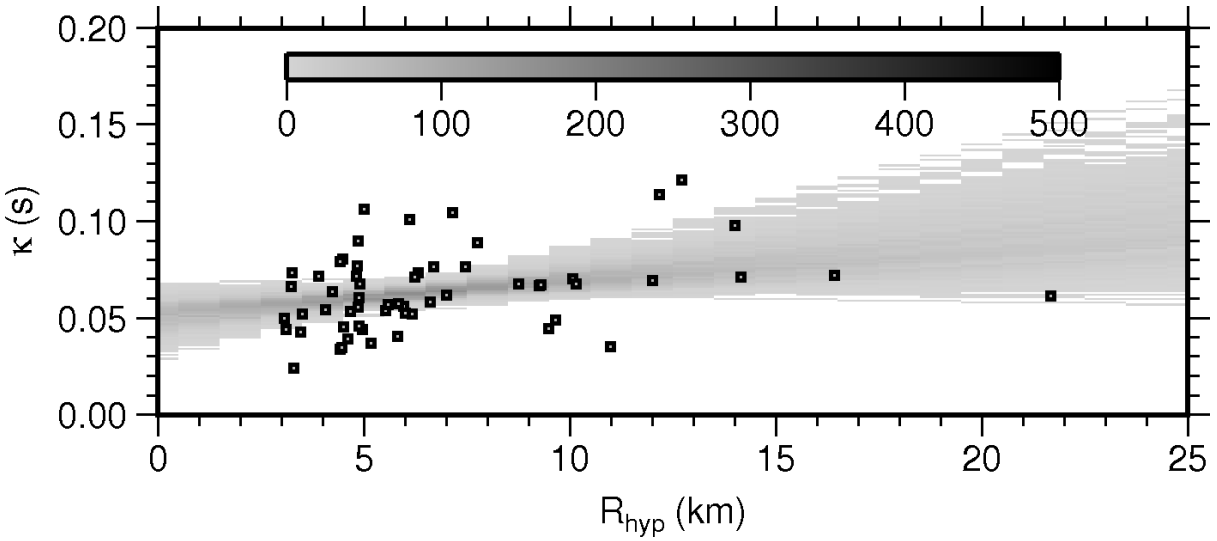


Figure A1.38: high-frequency fit t* plotted against hypocentral distance. Best-fis from 1000 random samples are shown with grey-shade indicating the density of solutions.

κ_0 at stations with three or more recordings are shown in Figure A1.39. Generally the two methods (high-frequency (Anderson and Hough, 1984) and broadband) provide values within agreement at one standard deviation. Nevertheless, due to the limited number of recordings at each station, the site specific κ_0 are not considered robust. Station 13 (HKS) shows the largest disagreement (0.040s for the high-frequency fit versus 0.090s for the broadband fit). Upon inspection of the fits, one of the 5 high frequency fits was very poor, removing this only leads to a minor increase (0.004s) in the κ_0 for this site however. Further investigation is required to establish the reason for this difference.

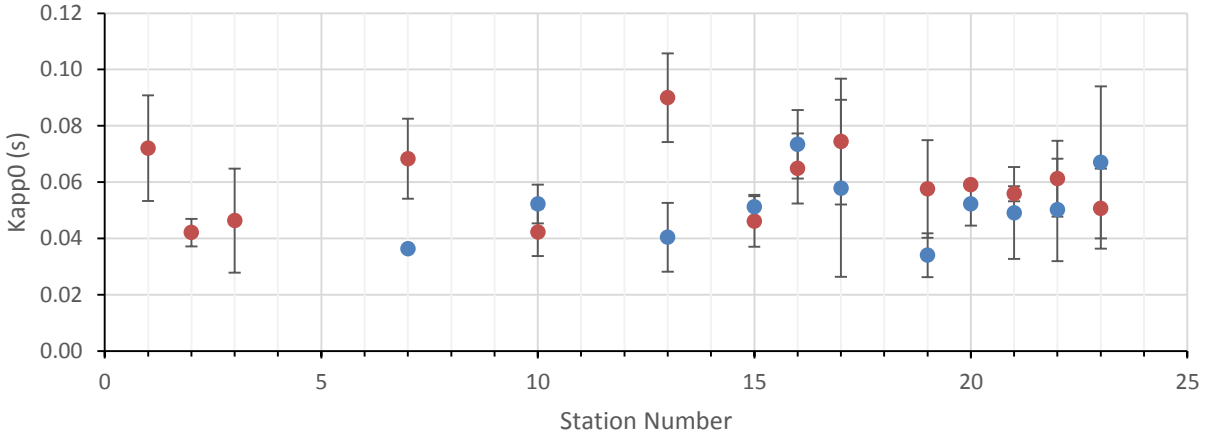


Figure A1.39: Comparison of κ_0 at stations with three or more recordings using the broadband (red, Q=264) and high-frequency fits (blue, Q=149).

Earthquake Stress Parameter

Stress parameters for the 12 analysed Groningen events are estimated using Brune's model, with:

$$\Delta\sigma = M_0 \left(\frac{f_c}{0.4906\beta}\right)^3 \quad \text{(A1.2)}$$

And:

$$M_0 = 10^{1.5(M_w+6.033)}. \quad \text{(A1.3)}$$

It is assumed that the shear-wave velocity near the source is β = 2.6km/s.

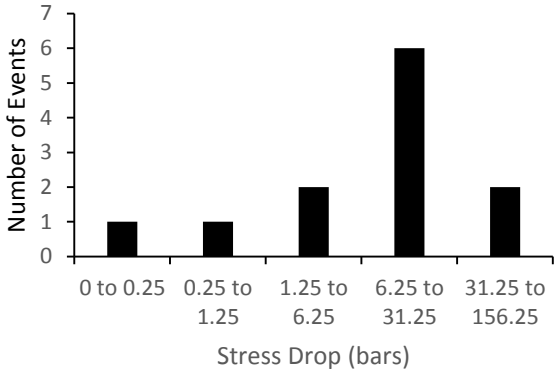


Figure A1.40: histogram of stress parameter for the analysed events.

Table A1.14: Stress parameter for the Groningen events.

Event #	Date	Time	M_w	Stress Drop (bar)
1	20060808	0504	3.4	23.3
2	20081030	0554	3.1	7.4
3	20090508	0523	2.9	5.6
4	20110119	1939	2.7	55.0
5	20110627	1548	3.4	9.9
6	20120816	2030	3.6	55.0
7	20130119	2010	2.8	3.2
8	20130207	2319	3.2	23.3
9	20140213	0213	3	7.4
10	20140901	0717	2.6	0.1
11	20140930	1142	2.8	0.6
12	20141105	0112	2.9	23.3

Amplification

Amplification is determined for the Groningen recordings based on the provided M_w and assumed geometrical decay of 1/R. This amplification includes 1D, 2D and 3D effects related to the both the amplification of individual phases (e.g., Sn) and the generation of systematically observed multiple phases

or surface waves. The average amplification (excluding the free-surface effect) over all recordings was found to be a factor 3.6, with a peak at around 2Hz of factor 5.5. The amplification below 1Hz is not estimated due to the fact that the earthquakes recorded so far did not generate sufficiently strong ground motions at these frequencies. It is therefore assumed, consistent with theoretical considerations of increasing velocity with depth, that the amplification returns to unity at 0.1s.

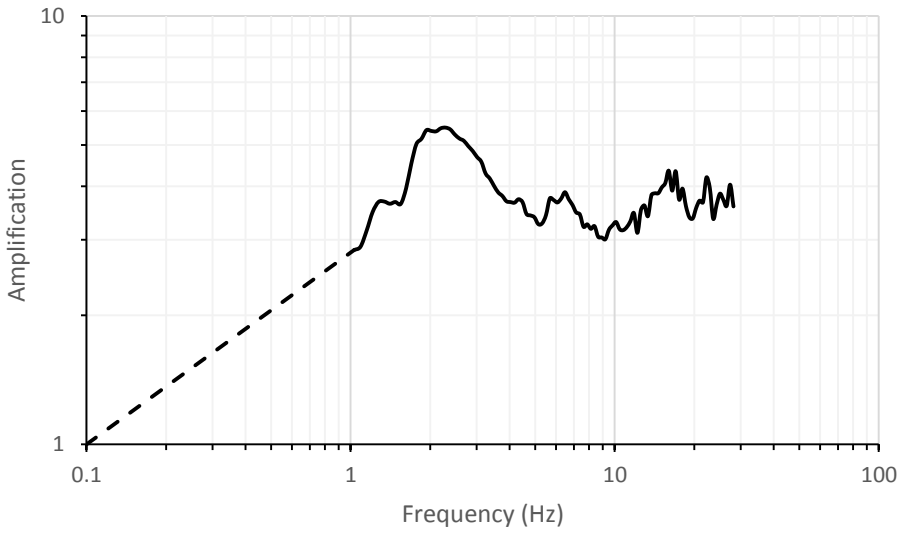


Figure A1.41: Average FAS amplification function. Solid: inverted, dashed: interpolated.

Summary and Next Steps

The analysis presented here has determined source (event specific stress-parameter), path (Q) and site effects (network-average κ_0, amplification) using recordings of 12 events on 30 surface accelerographs of the KNMI monitoring network. These estimates can be used to guide the construction of stochastic simulation models and GMPEs. Some of the determined values suffer from the limited dataset (e.g., κ_0 values require more recordings at each station - network average values must currently be relied upon). However, the dataset is rapidly expanding due to the expansion of the monitoring network and more recent earthquakes. Using the results and experience of this analysis we will be able to build more robust models using the new data. Beyond the limited data available at the point of this analysis, one issue highlighted was that a good estimate of the noise level is important in order to properly assess the fitting bandwidth. In many cases overly conservative noise estimates (due to short pre-event noise in the current dataset) meant that the fitting bandwidth was shorter than necessary. In addition to new data, the project has made significant progress in modelling the geometrical decay function using full-waveform simulations. The next analyses will therefore use this information to reduce the uncertainty in the predictive models.

In addition to the work documented here, KNMI, led by Bernard Dost, are investigating source, path and site effects of Groningen recordings. The current status of their analyses focuses on the separation of effects for the direct arrival (Sn), which allows a physical interpretation of the recovered parameters, consistent with ray-theory, to be undertaken. The primary difference in the case of this report is that the focus is on reproducing the average spectral characteristics of the entire seismogram rather than single phases – as required for the subsequent hazard and risk calculations. It is therefore noted that in order to

obtain physical material properties, which are critical, for instance, in full-waveform modelling approaches, the analysis of direct arrivals by KNMI for the quantification of attenuation effects is crucial. However, the approaches, whilst potentially providing different estimates of model parameters, should be taken within the context of their objectives.

References

Abercrombie, R. and P. Leary (1993). Source parameters of small earthquakes recorded at 2.5 km depth, cajon pass, southern california - implications for earthquake scaling, *Geophys Res Lett* **20**, 1511-1514

Anderson, J. G. and S. E. Hough (1984). A model for the shape of the fourier amplitude spectrum of acceleration at high-frequencies, *B Seismol Soc Am* **74**, 1969-1993

Brune, J. N. (1970). Tectonic stress and spectra of seismic shear waves from earthquakes, *J Geophys Res* **75**, 4997-5009

De Natale, G., R. Madariaga, R. Scarpa and A. Zollo (1987). Source parameter analysis from strong motion records of the friuli, italy, earthquake sequence (1976-1977), *B Seismol Soc Am* **77**, 1127-1146

Edwards, B., A. Rietbrock, J. J. Bommer and B. Baptie (2008). The acquisition of source, path, and site effects from microearthquake recordings using q tomography: Application to the united kingdom, *B Seismol Soc Am* **98**, 1915-1935

Hanks, T. C. and R. K. Mcguire (1981). The character of high-frequency strong ground motion, *B Seismol Soc Am* **71**, 2071-2095

Masuda, T. and Z. Suzuki (1982). Objective estimation of source parameters and local q values by simultaneous inversion method, *Phys Earth Planet In* **30**, 197-208

Scherbaum, F. (1990). Combined inversion for the three-dimensional q structure and source parameters using microearthquake spectra, *Journal of Geophysical Research: Solid Earth (1978–2012)* **95**, 12423-12438

Technical Reports

Bommer, J. J., B. Dost and M. Ntinalexis (2014). Ground-Motion Records from the Groningen Field. Technical Report.

Dost, B., & Kraaijpoel, D. (2013). The August 16, 2012 earthquake near Huizinge (Groningen). KNMI Scientific report (http://www. knmi. nl/knmilibrary/miscellaneousreport. html).

Additional Material: Spectral Models

Earthquake 1

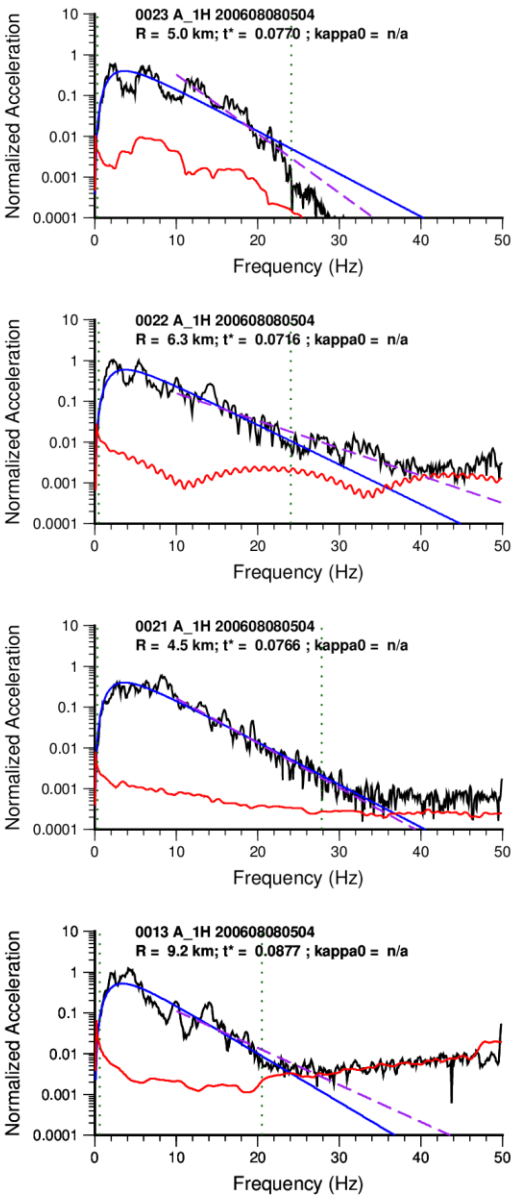


Figure A1.42: best lin-fit broad-band models for geometrical mean of horizontal component recordings for EQ-1. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

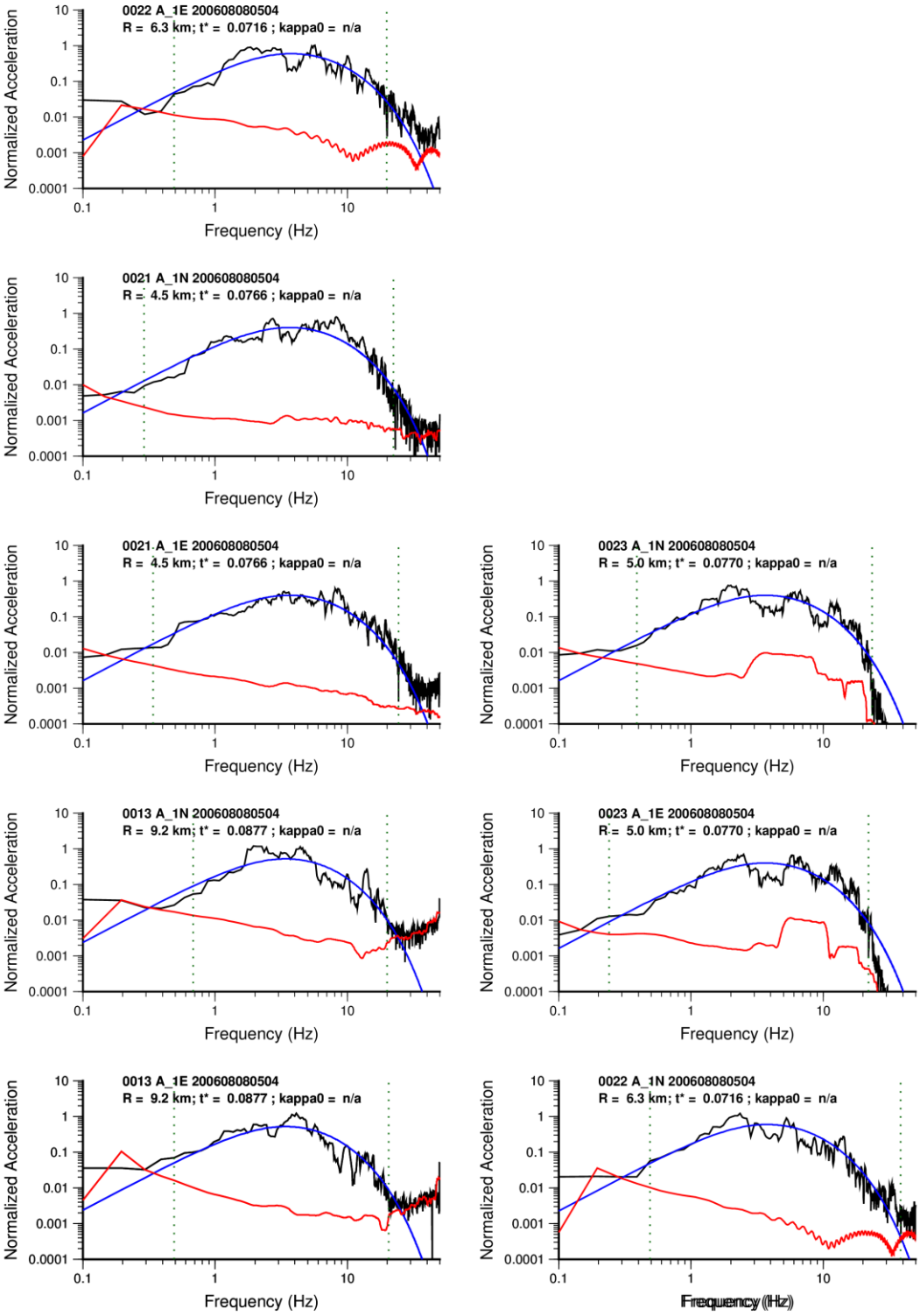


Figure A1.43: log-fit broad-band models for individual horizontal component recordings of EQ-1. Whole path attenuation (t*) is fixed based on the lin fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 2

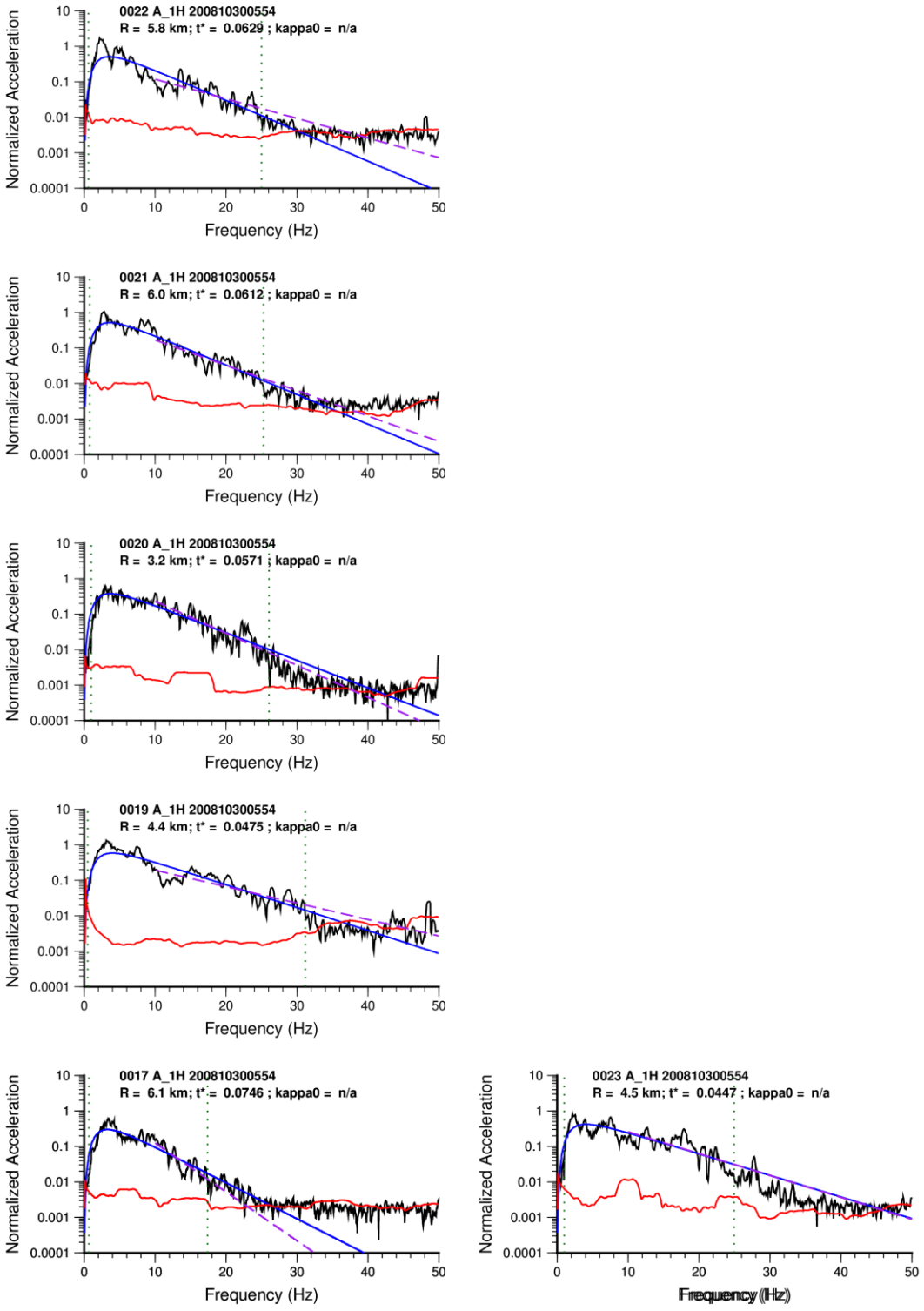


Figure A1.44: best lin-fit broad-band models for geometrical mean of horizontal component recordings for EQ-2. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

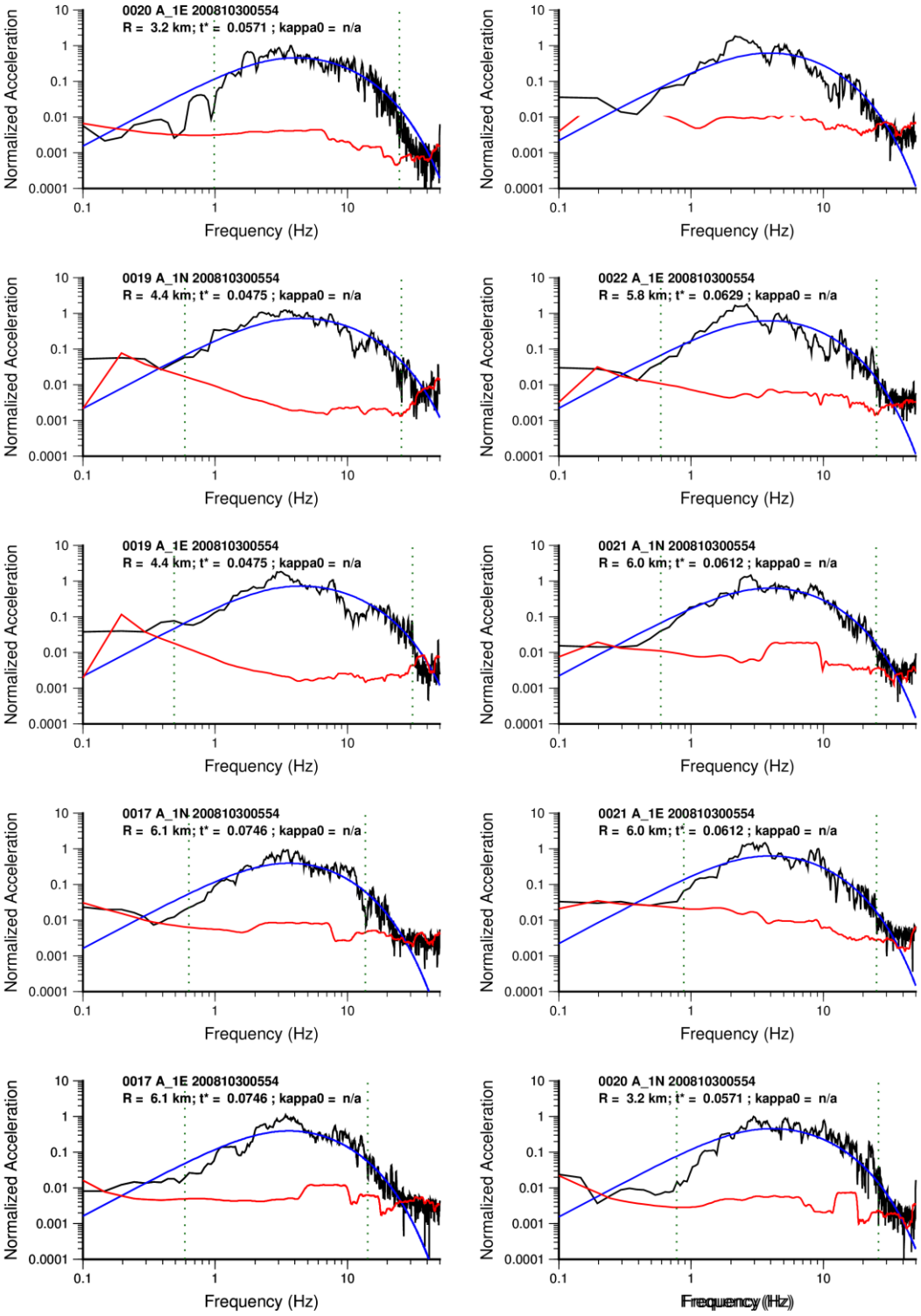


Figure A1.45: log-fit broad-band models for individual horizontal component recordings of EQ-2. Whole path attenuation (t*) is fixed based on the lin fit models. Red: noise, black: earthquake, blue: broadband model.

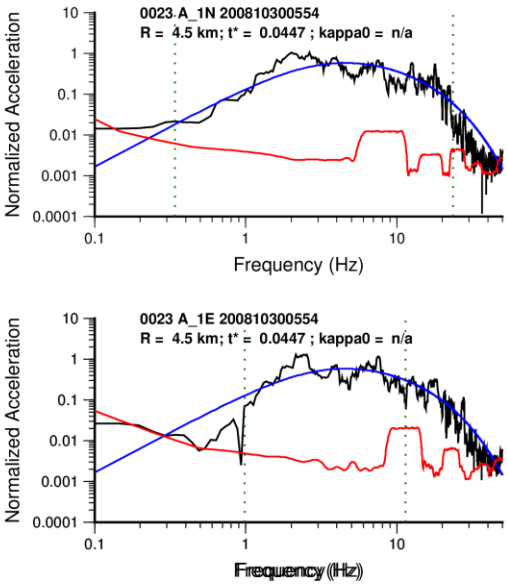


Figure A1.46: log-fit broad-band models for individual horizontal component recordings of EQ-2. Whole path attenuation (t*) is fixed based on the lin fit models. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

Earthquake 3

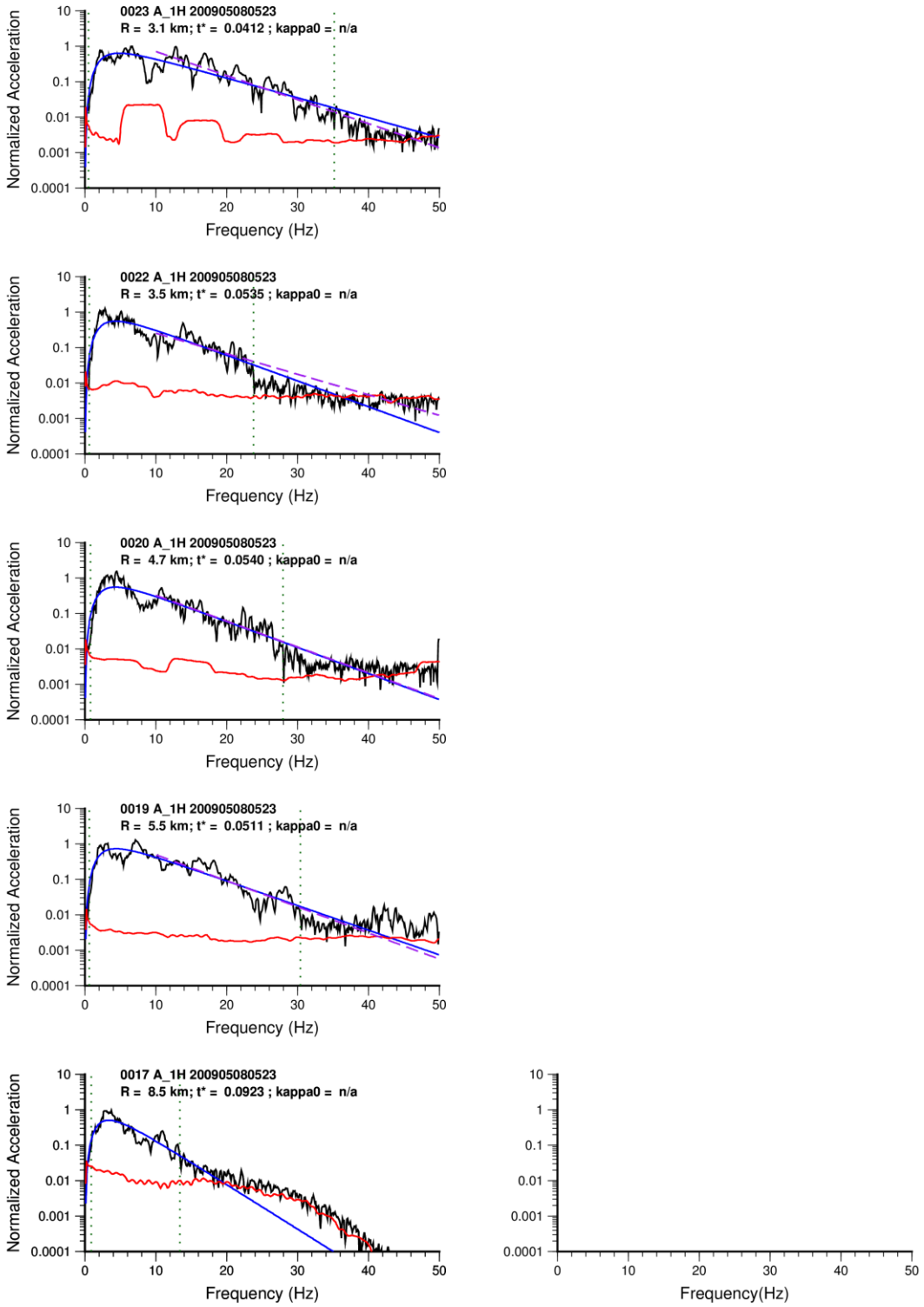


Figure A1.47: best lin-fit broad-band models for geometrical mean of horizontal component recordings for EQ-3. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

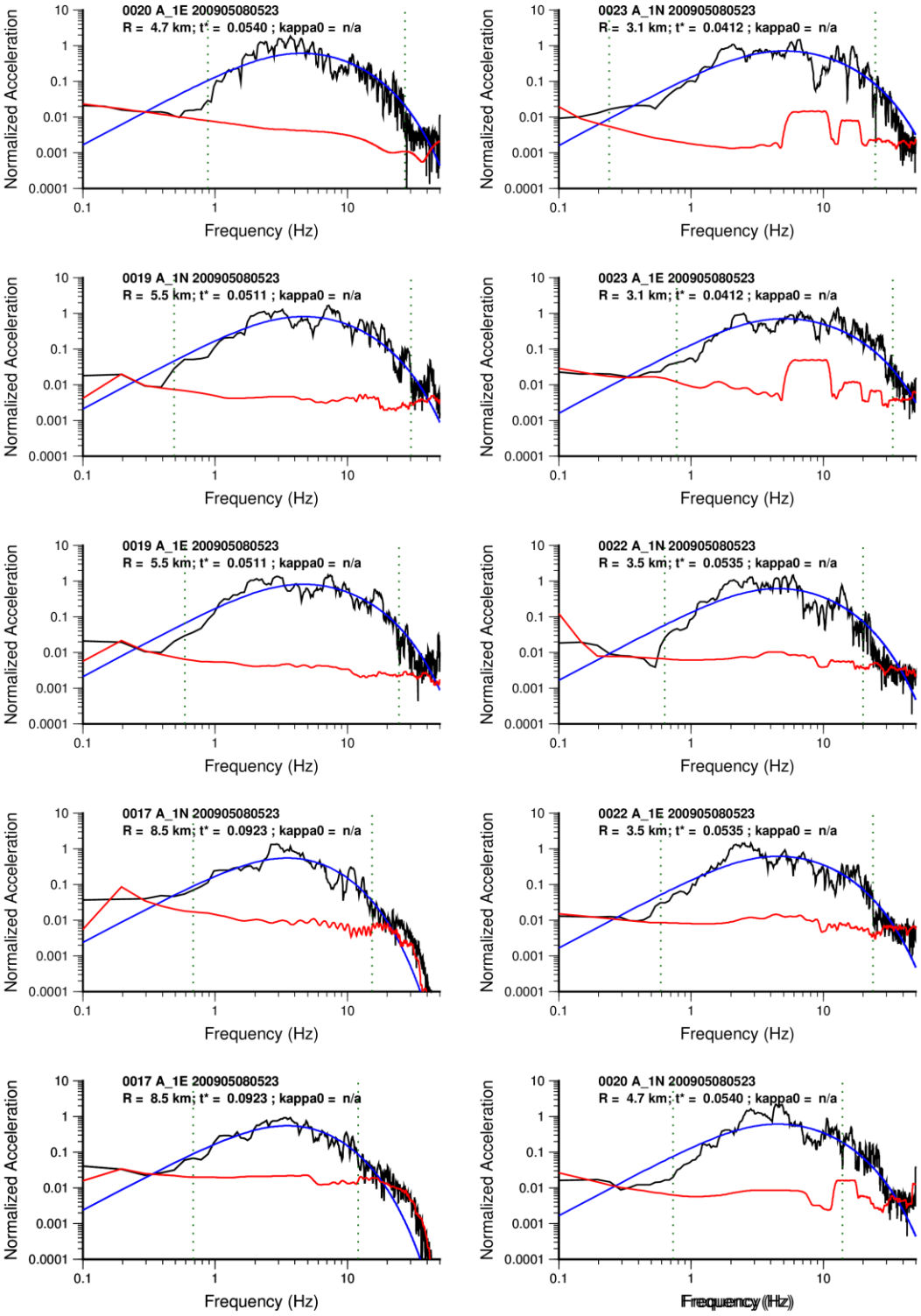


Figure A1.48: log-fit broad-band models for individual horizontal component recordings of EQ-3. Whole path attenuation (t*) is fixed based on the lin fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 4

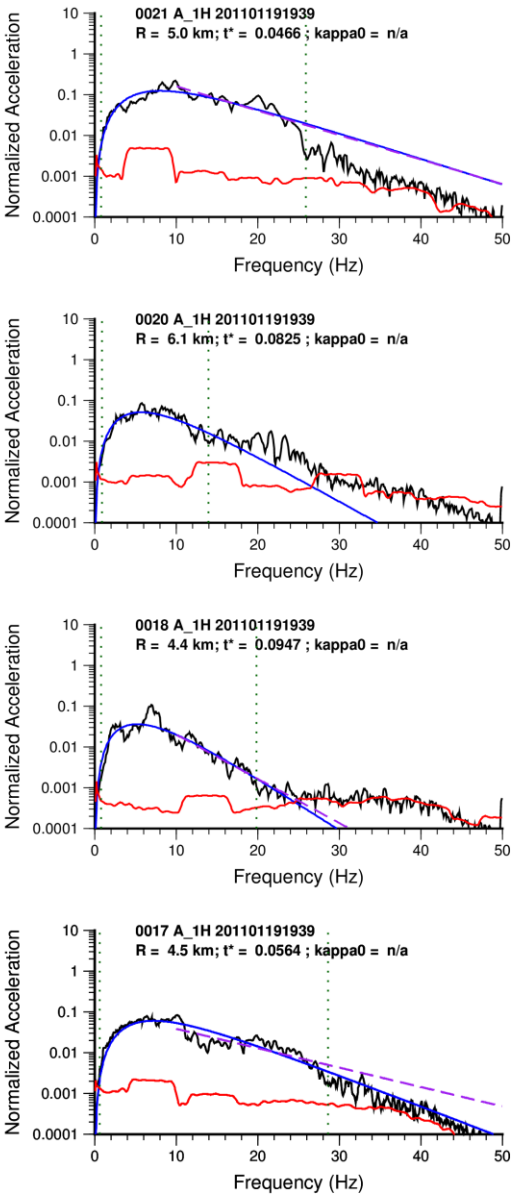


Figure A1.49: best lin-fit broad-band models for geometrical mean of horizontal component recordings for EQ-4. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

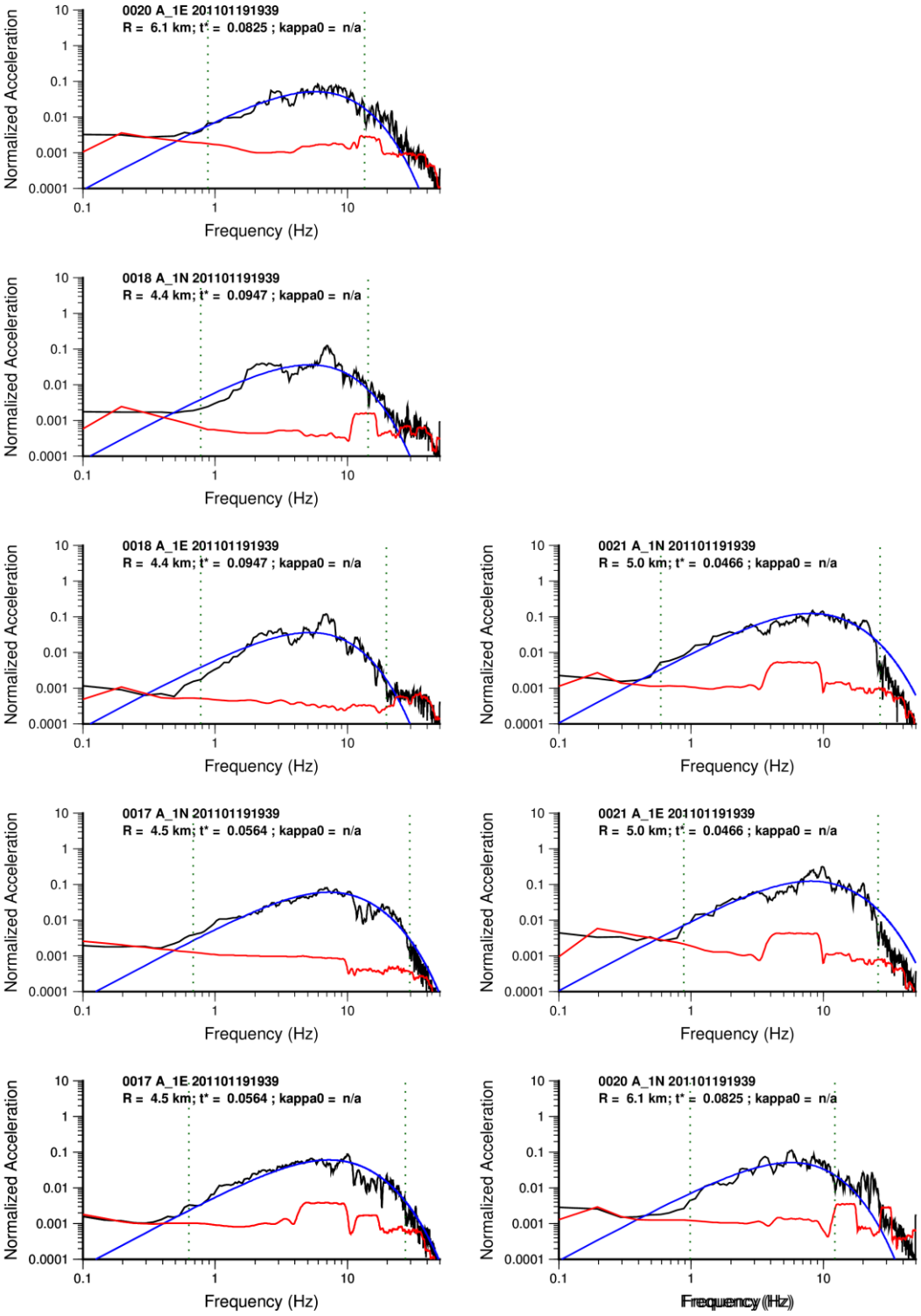


Figure A1.50: log-fit broad-band models for individual horizontal component recordings of EQ-4. Whole path attenuation (t*) is fixed based on the lin fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 5

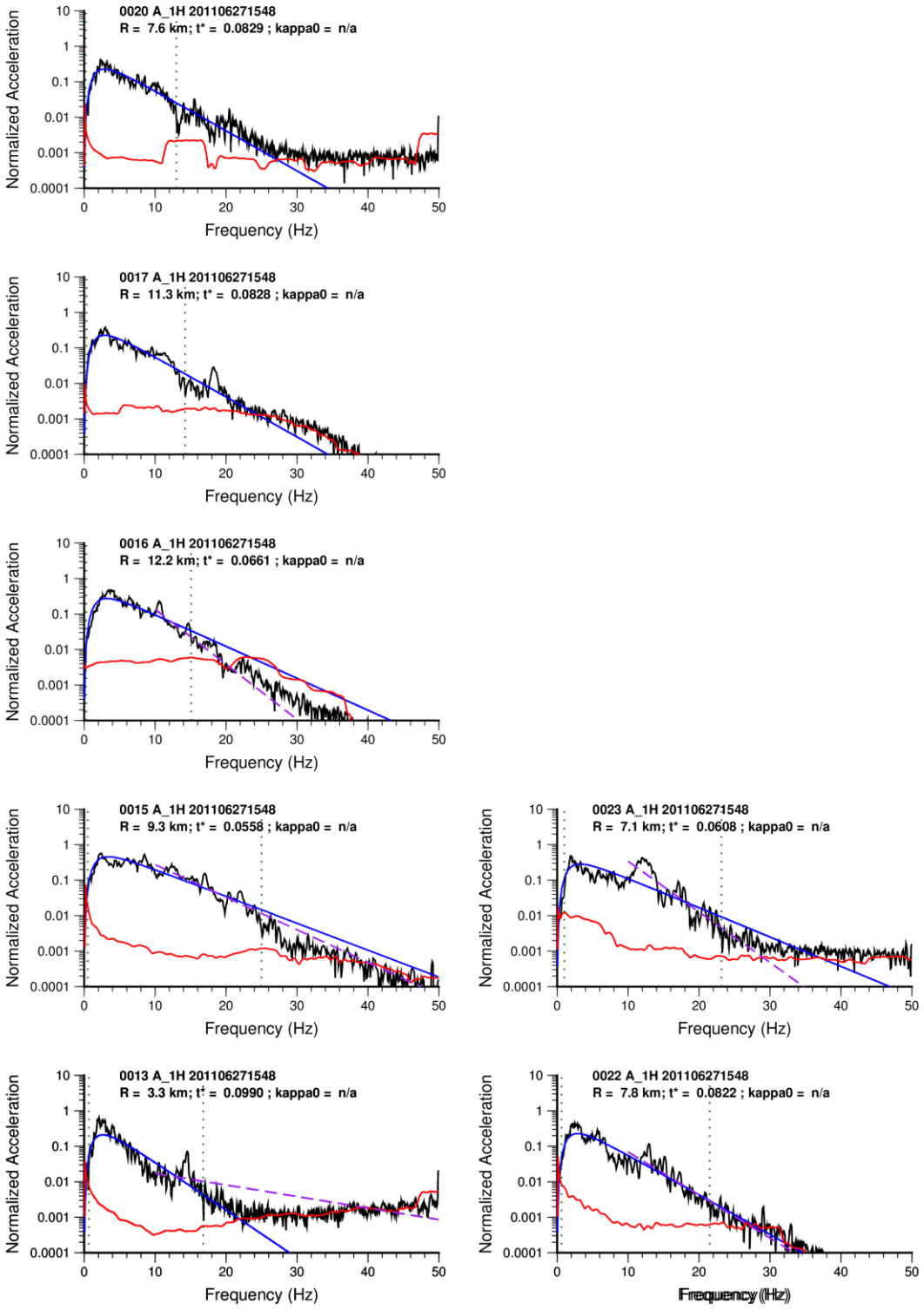


Figure A1.51: best lin-fit broad-band models for geometrical mean of horizontal component recordings for EQ-5. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

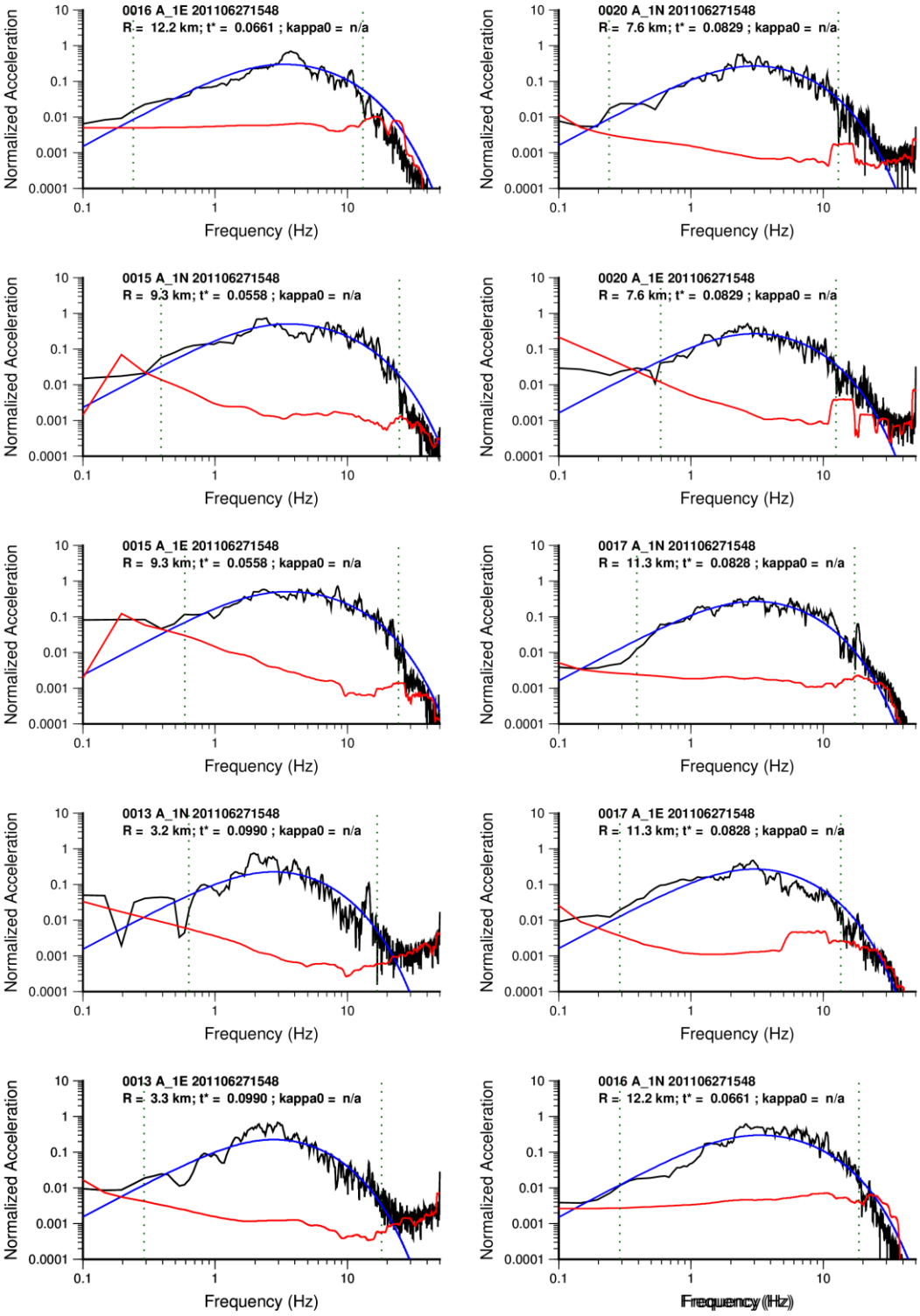


Figure A1.52: log-fit broad-band models for individual horizontal component recordings of EQ-5. Whole path attenuation (t*) is fixed based on the lin fit models. Red: noise, black: earthquake, blue: broadband model.

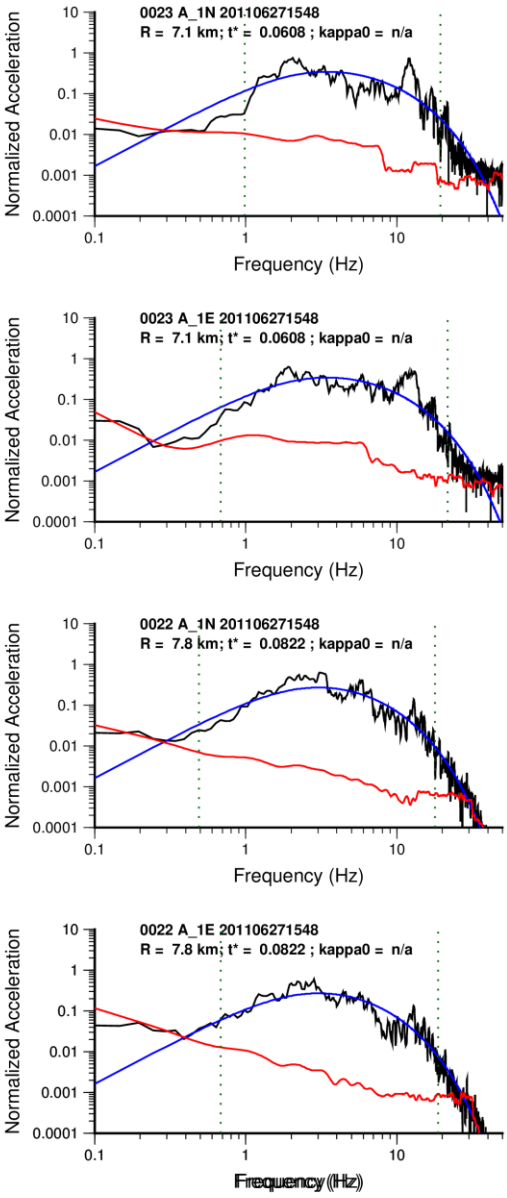


Figure A1.53: log-fit broad-band models for individual horizontal component recordings of EQ-5. Whole path attenuation (t*) is fixed based on the lin fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 6

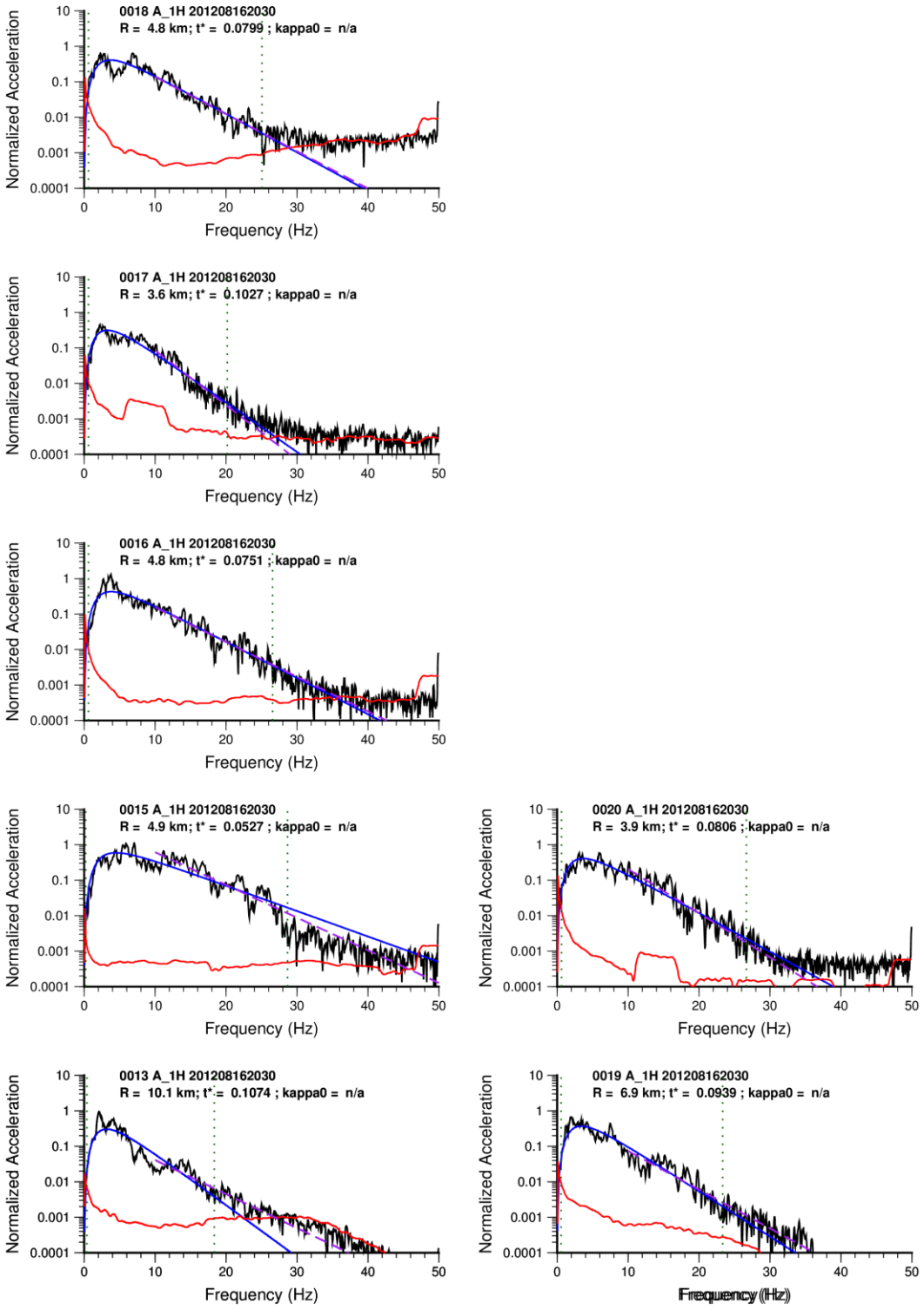


Figure A1.54: best lin-fit broad-band models for geometrical mean of horizontal component recordings for EQ-6. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

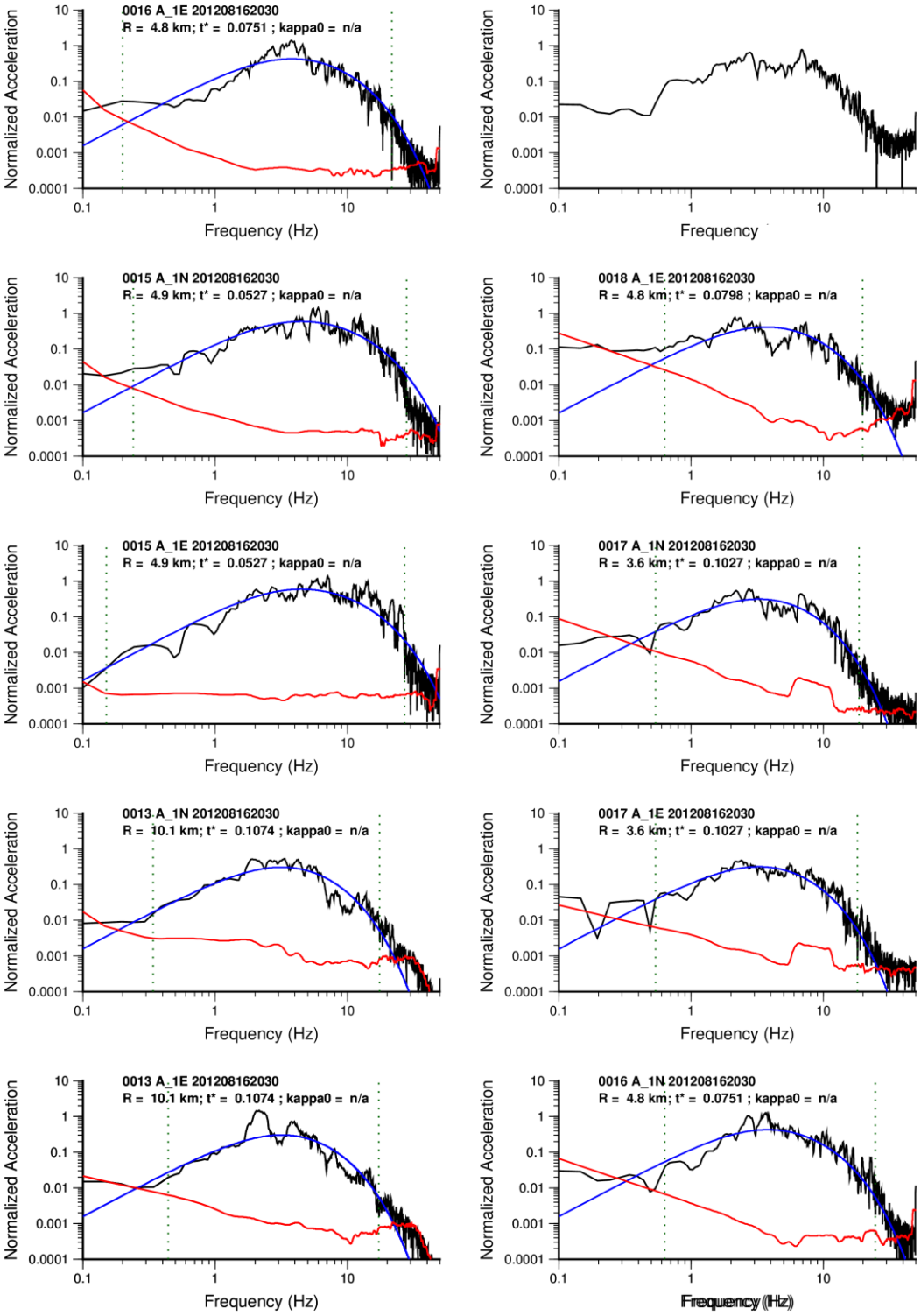
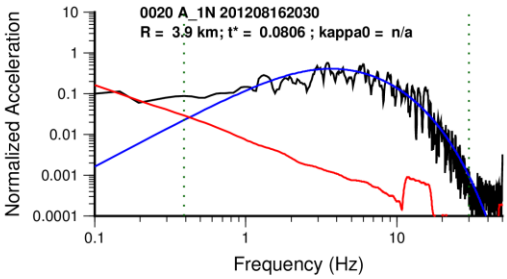
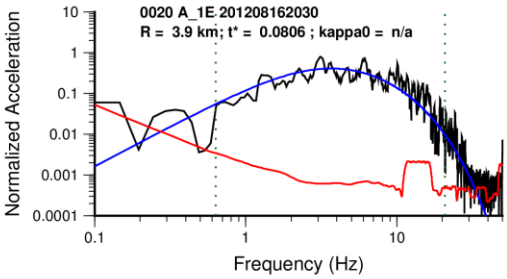
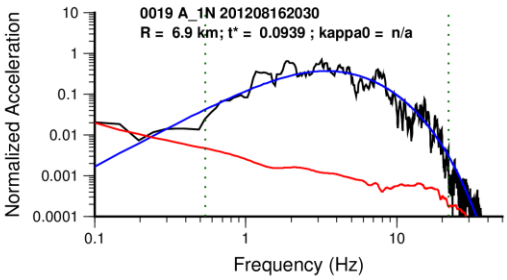


Figure A1.55: log-fit broad-band models for individual horizontal component recordings of EQ-6. Whole path attenuation (t*) is fixed based on the lin fit models. Red: noise, black: earthquake, blue: broadband model.







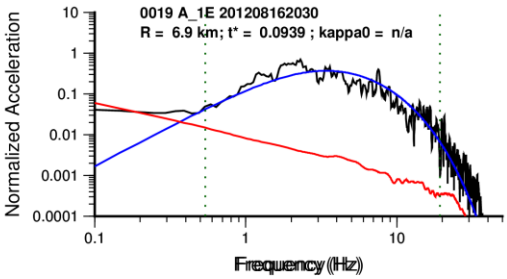


Figure A1.12: log-fit broad-band models for individual horizontal component recordings of EQ-6. Whole path attenuation (t*) is fixed based on the lin fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 7

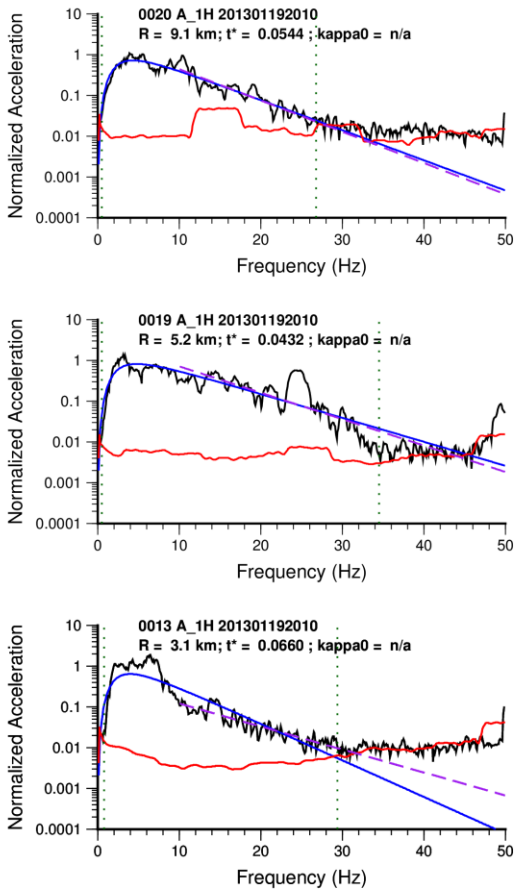


Figure A1.56: best lin-fit broad-band models for geometrical mean of horizontal component recordings for EQ-7. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

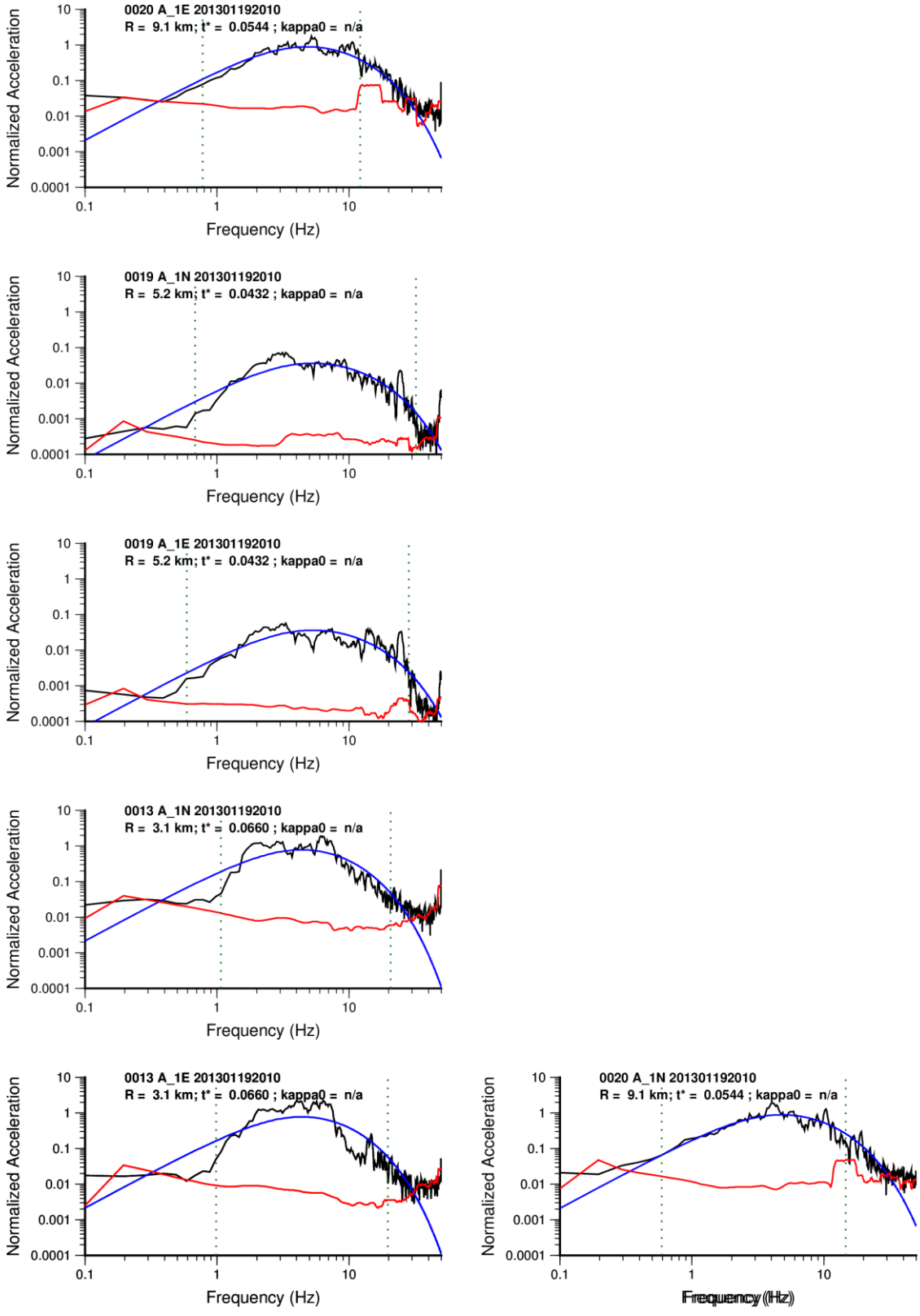


Figure A1.57: log-fit broad-band models for individual horizontal component recordings of EQ-7. Whole path attenuation (t*) is fixed based on the lin fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 8

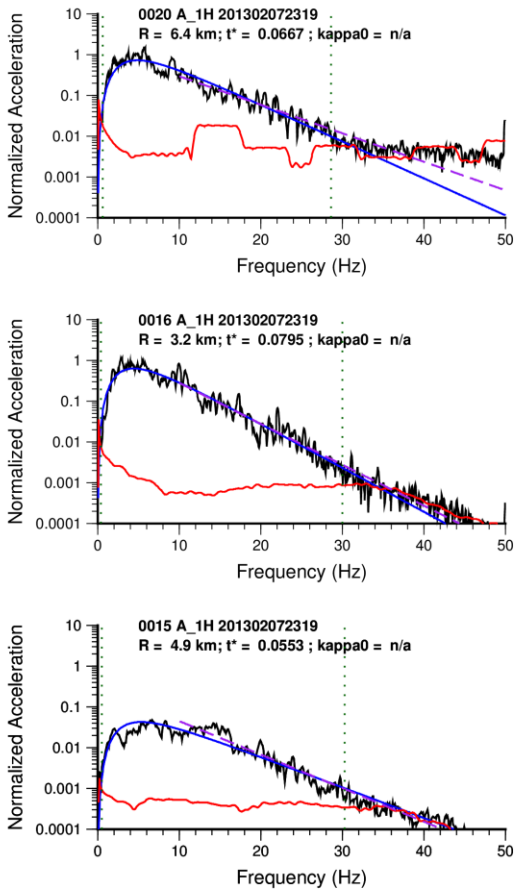


Figure A1.58: best lin-fit broad-band models for geometrical mean of horizontal component recordings for EQ-8. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

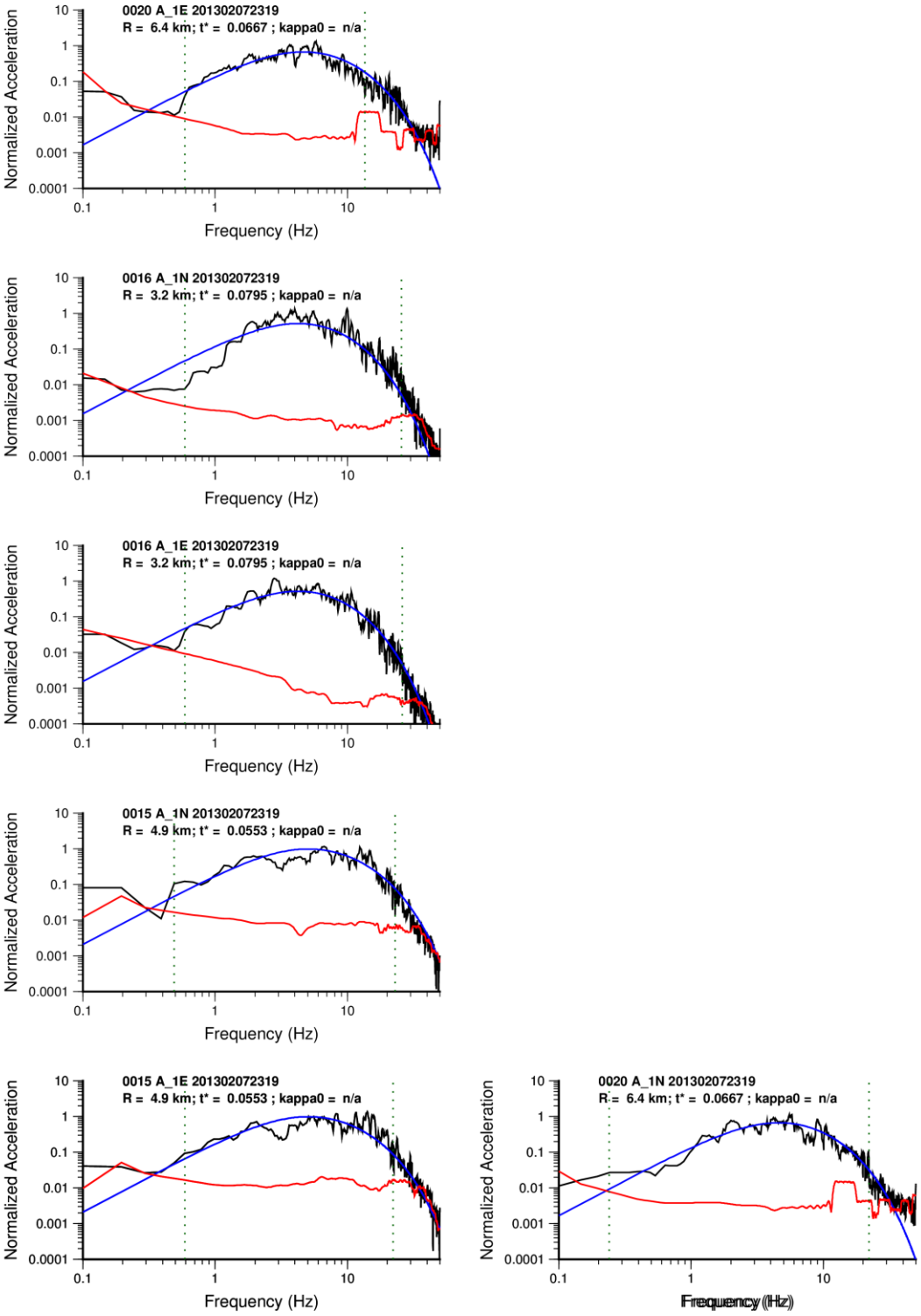


Figure A1.59: log-fit broad-band models for individual horizontal component recordings of EQ-8. Whole path attenuation (t*) is fixed based on the lin fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 9

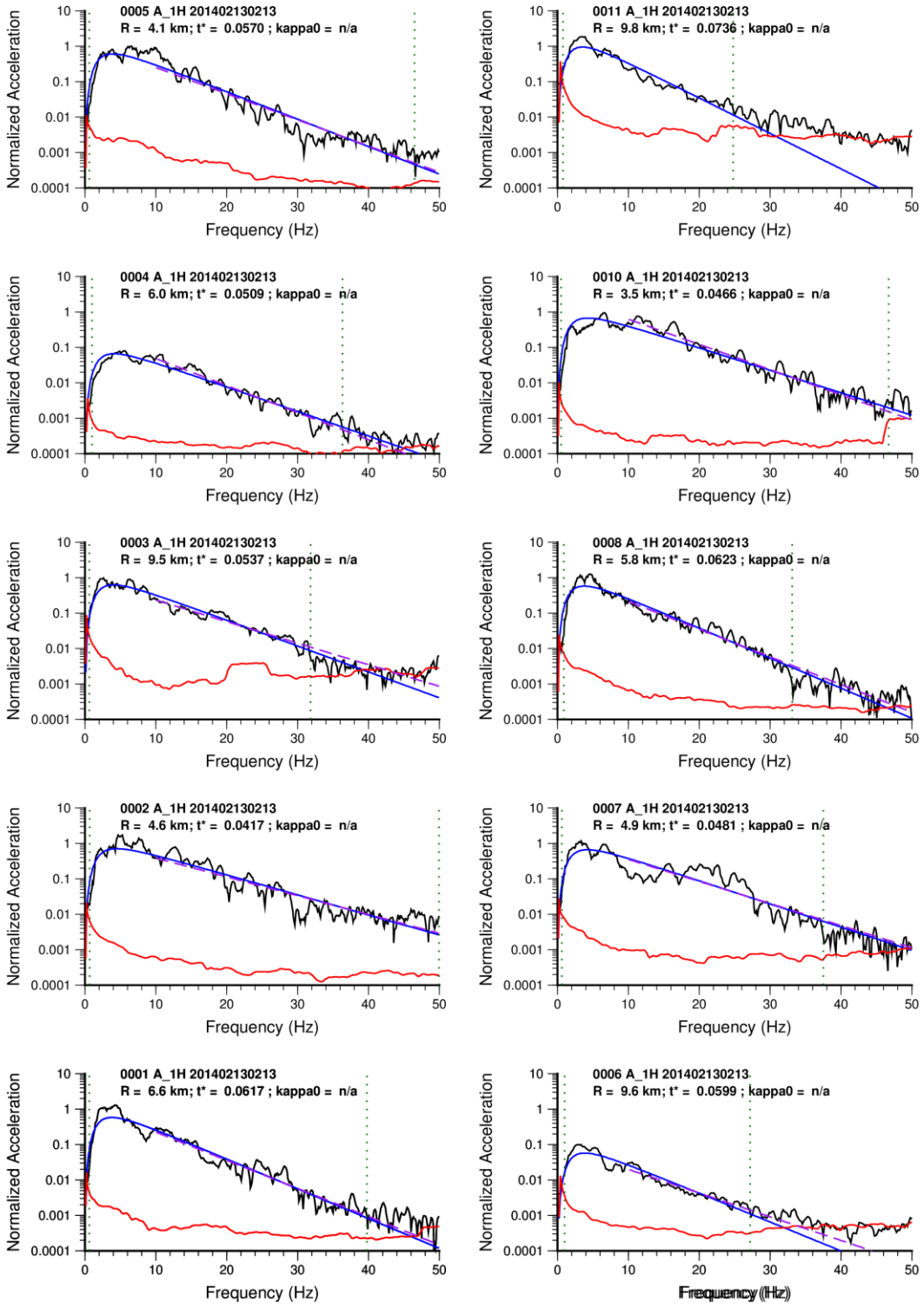


Figure A1.60: best lin-fit broad-band models for geometrical mean of horizontal component recordings for EQ-9. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

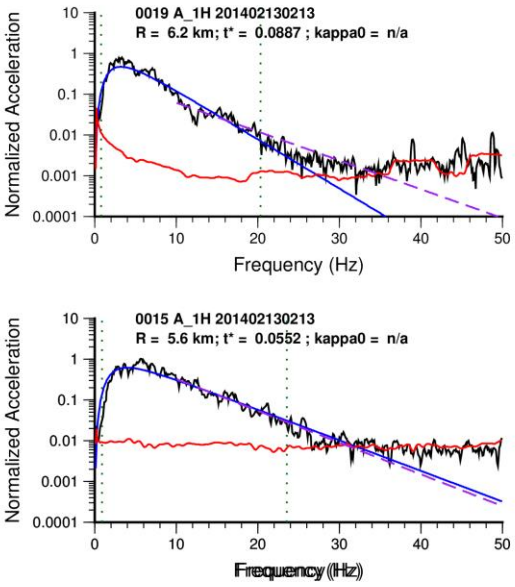


Figure A1.61: best lin-fit broad-band models for geometrical mean of horizontal component recordings for EQ-9. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

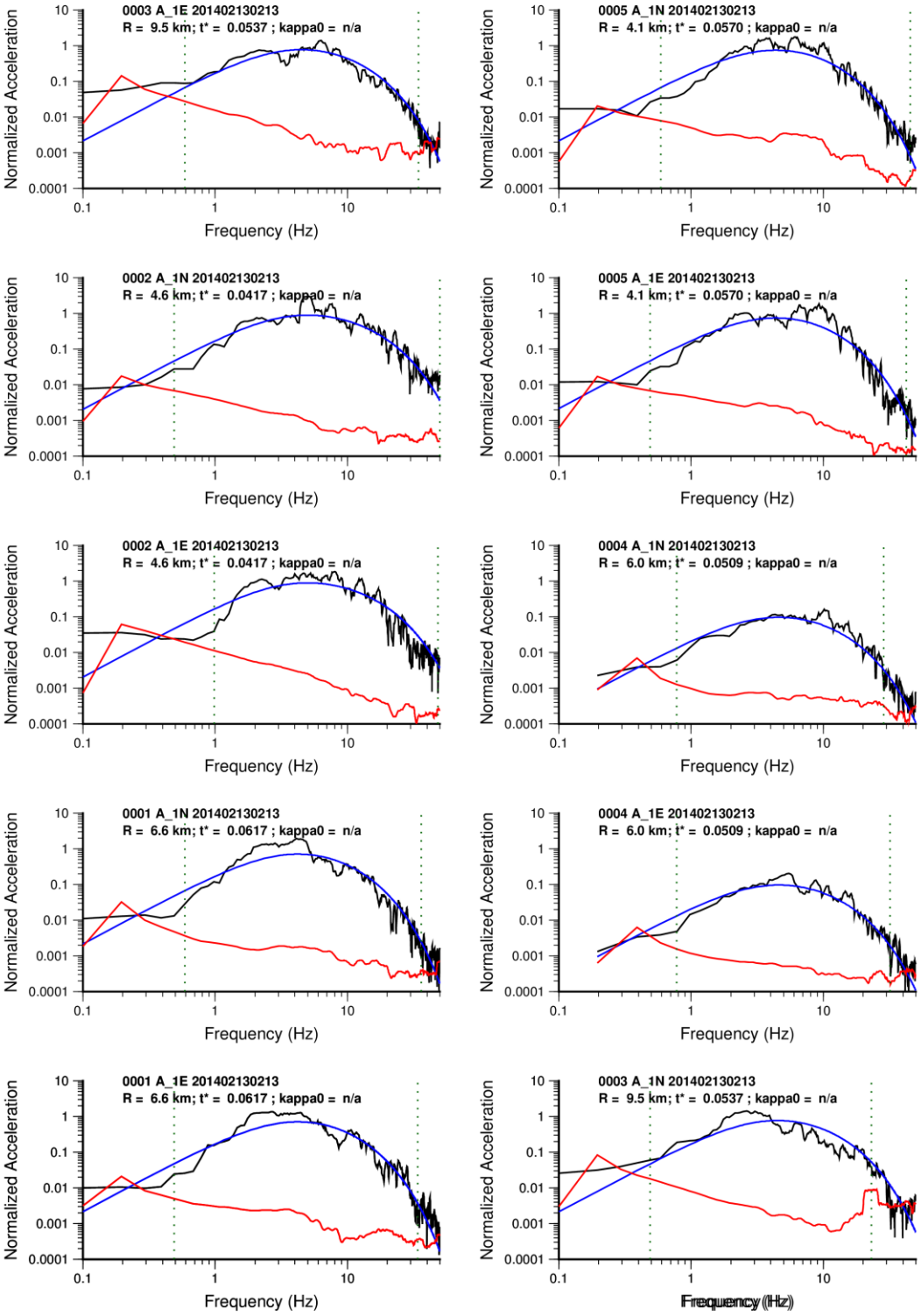


Figure A1.62: log-fit broad-band models for individual horizontal component recordings of EQ-9. Whole path attenuation (t*) is fixed based on the lin fit models. Red: noise, black: earthquake, blue: broadband model.

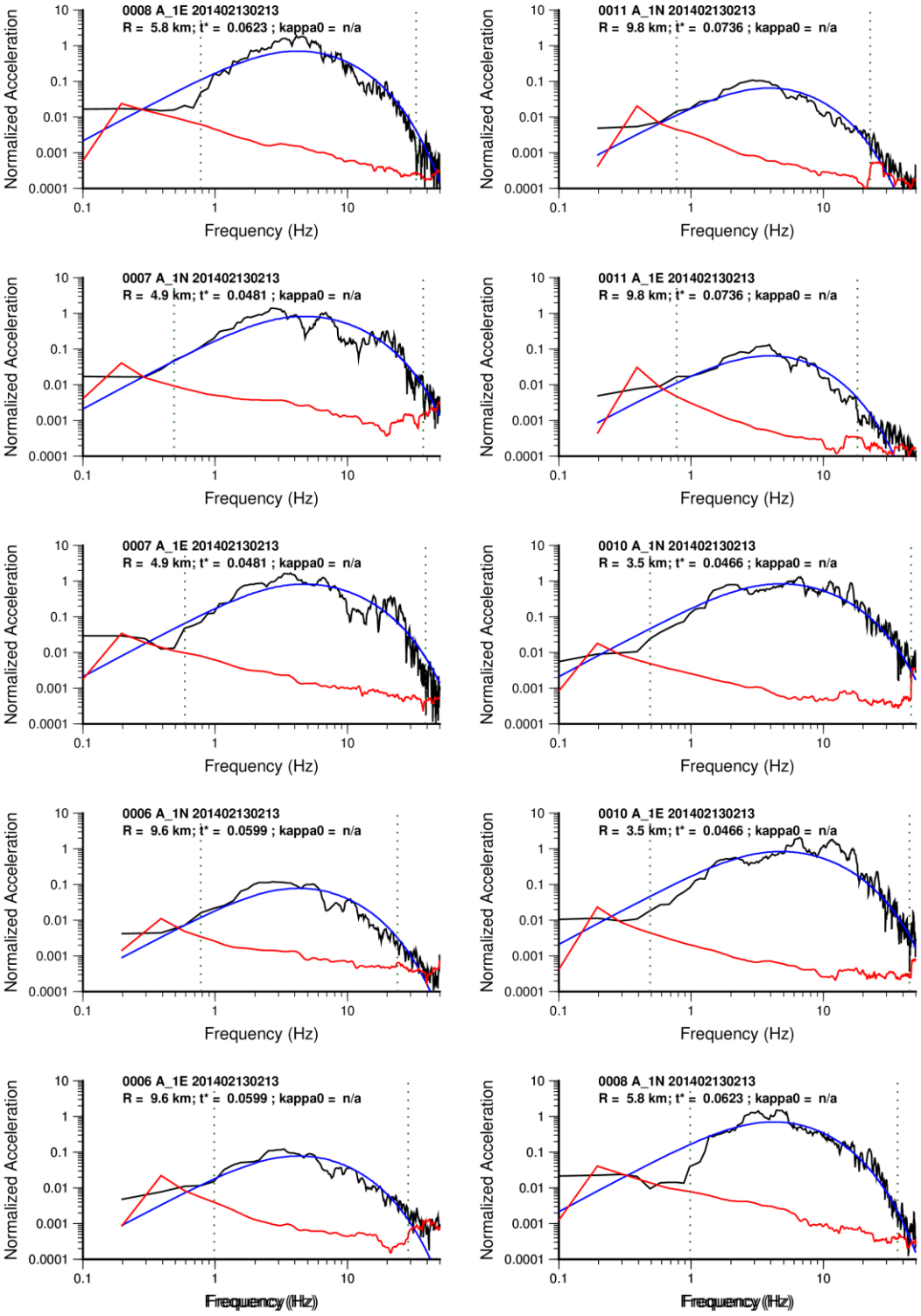


Figure A1.63: log-fit broad-band models for individual horizontal component recordings of EQ-9. Whole path attenuation (t*) is fixed based on the lin fit models. Red: noise, black: earthquake, blue: broadband model.

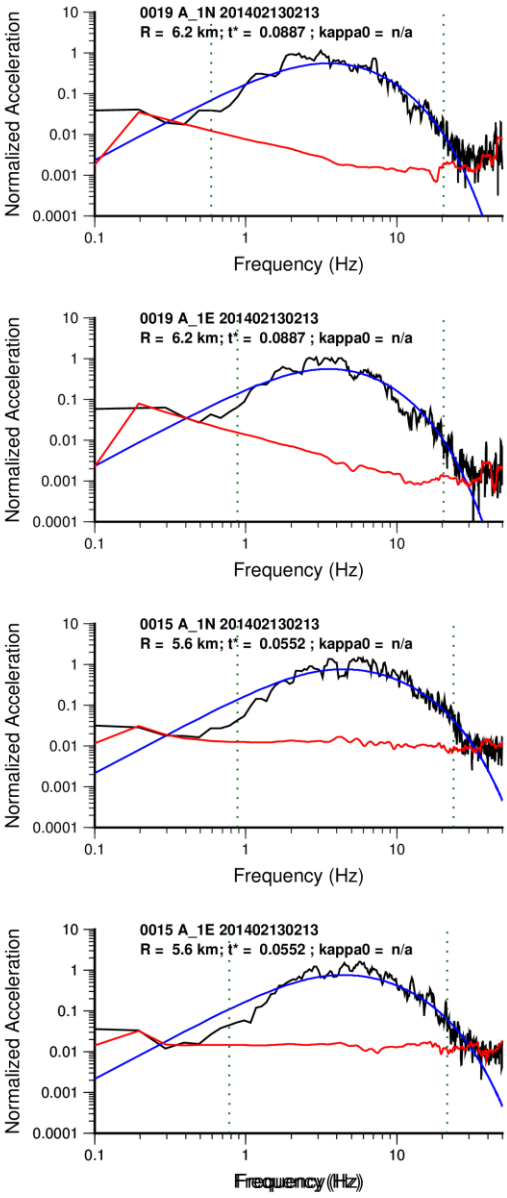


Figure A1.64: log-fit broad-band models for individual horizontal component recordings of EQ-9. Whole path attenuation (t*) is fixed based on the lin fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 10

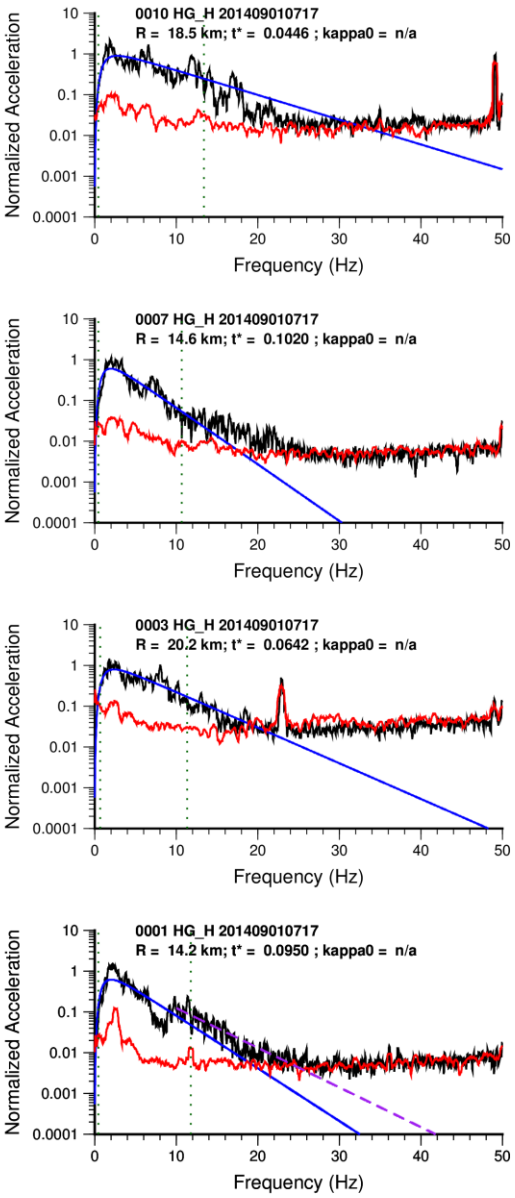


Figure A1.65: best lin-fit broad-band models for geometrical mean of horizontal component recordings for EQ-10. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

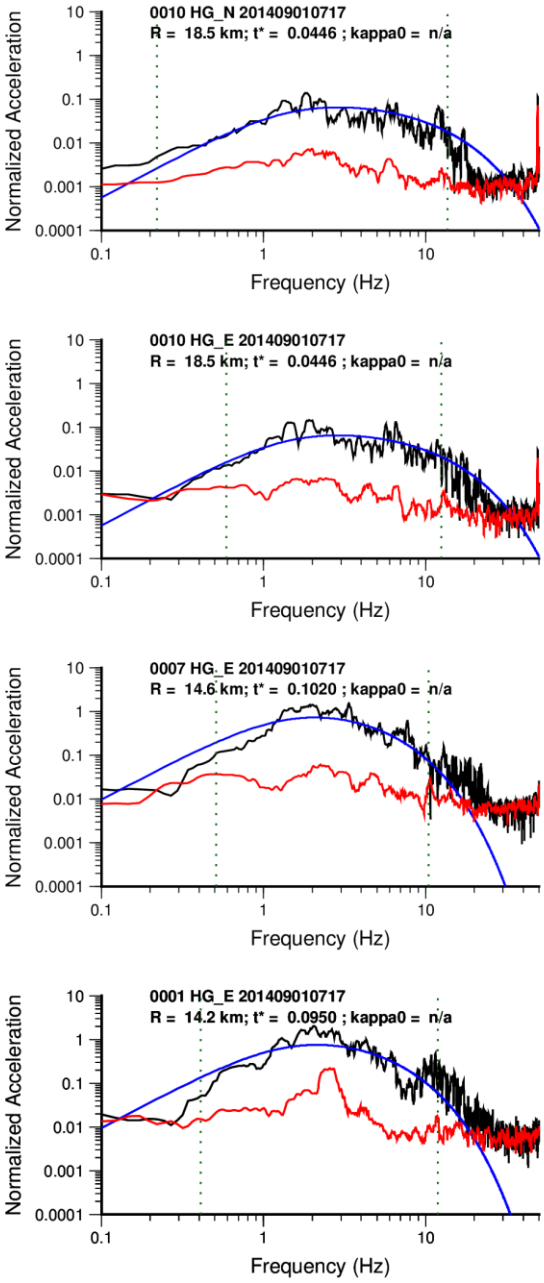


Figure A1.66: log-fit broad-band models for individual horizontal component recordings of EQ-10. Whole path attenuation (t*) is fixed based on the lin fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 11

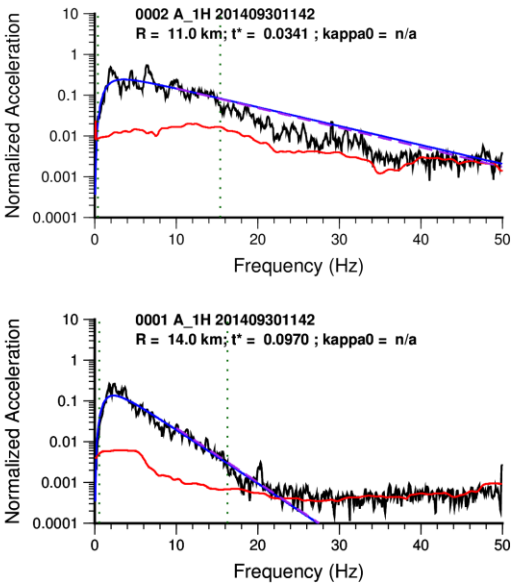


Figure A1.67: best lin-fit broad-band models for geometrical mean of horizontal component recordings for EQ-11. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

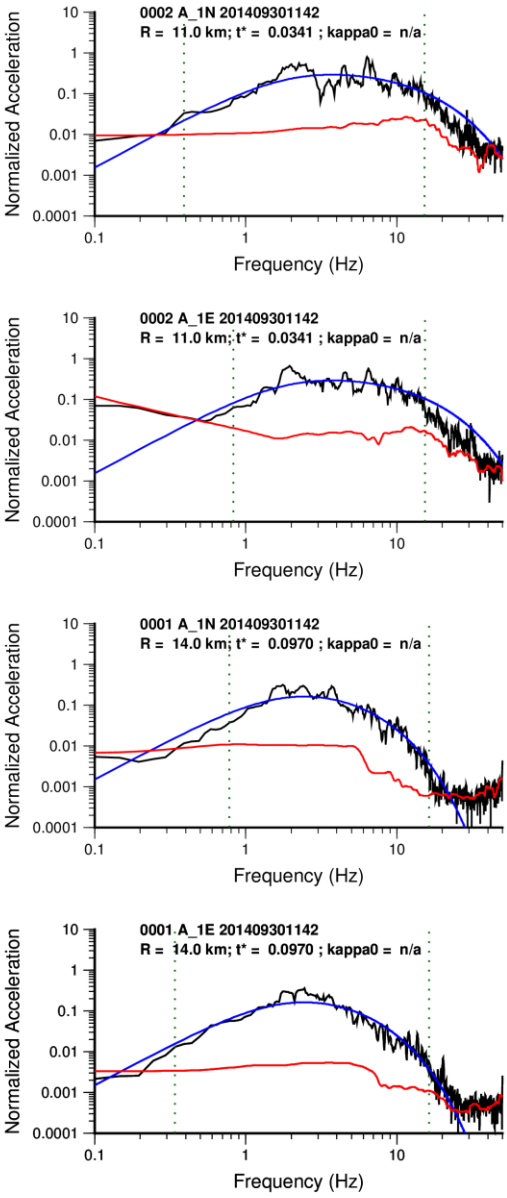


Figure A1.68: log-fit broad-band models for individual horizontal component recordings of EQ-11. Whole path attenuation (t*) is fixed based on the lin fit models. Red: noise, black: earthquake, blue: broadband model.

Earthquake 12

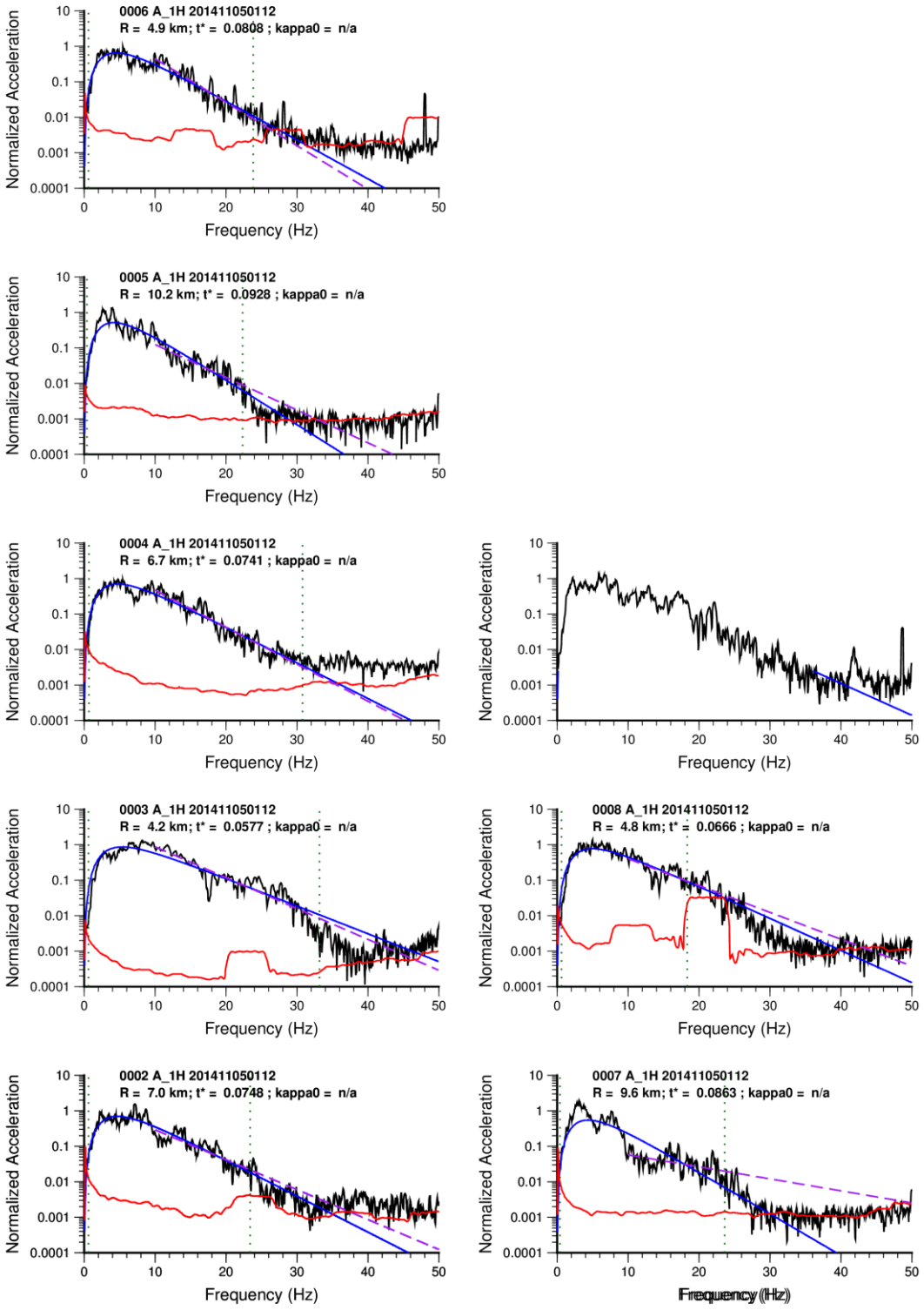


Figure A1.69: best lin-fit broad-band models for geometrical mean of horizontal component recordings for EQ-12. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

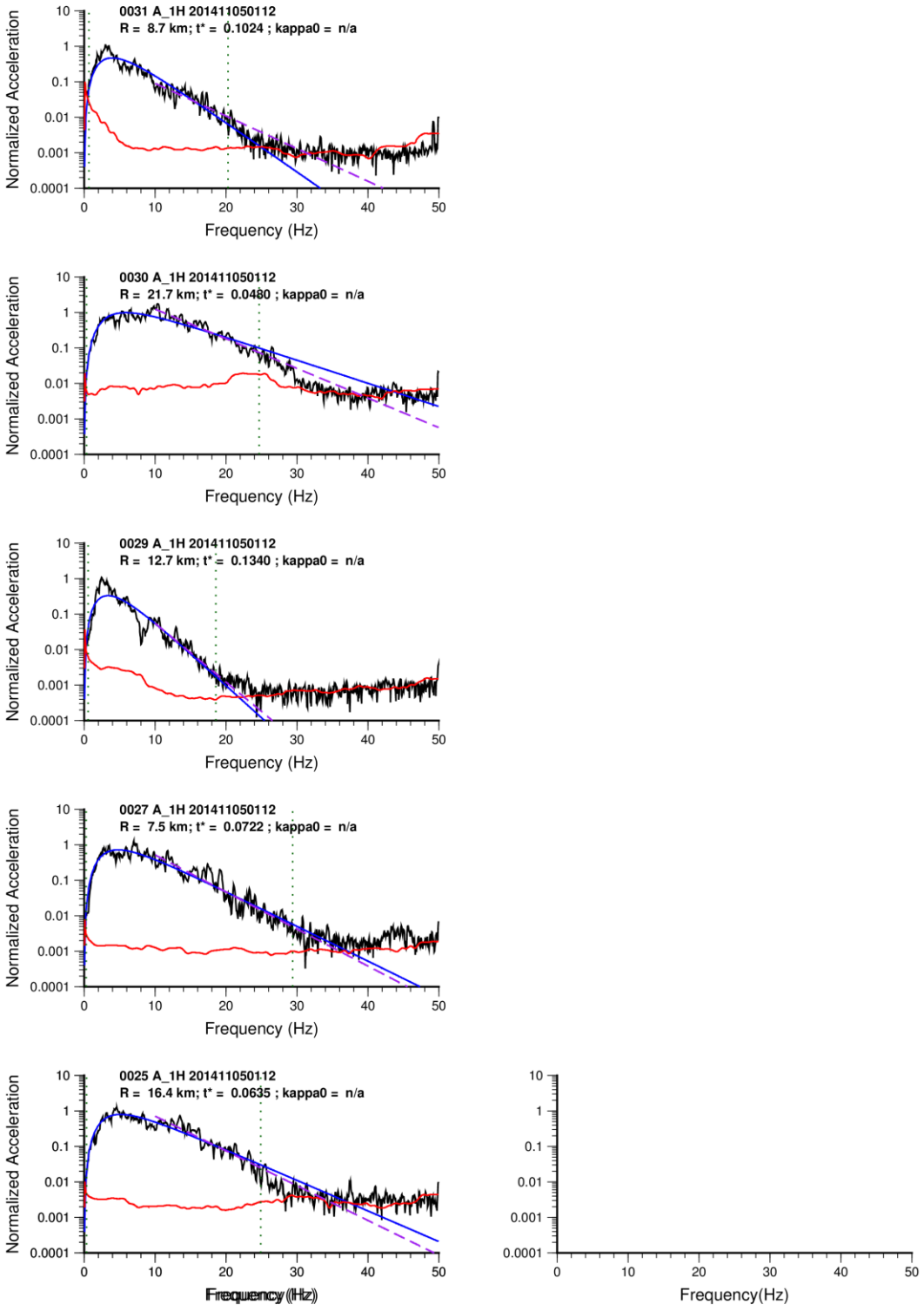


Figure A1.70: best lin-fit broad-band models for geometrical mean of horizontal component recordings for EQ-12. Red: noise, black: earthquake, blue: broadband model, dashed purple: high-frequency model.

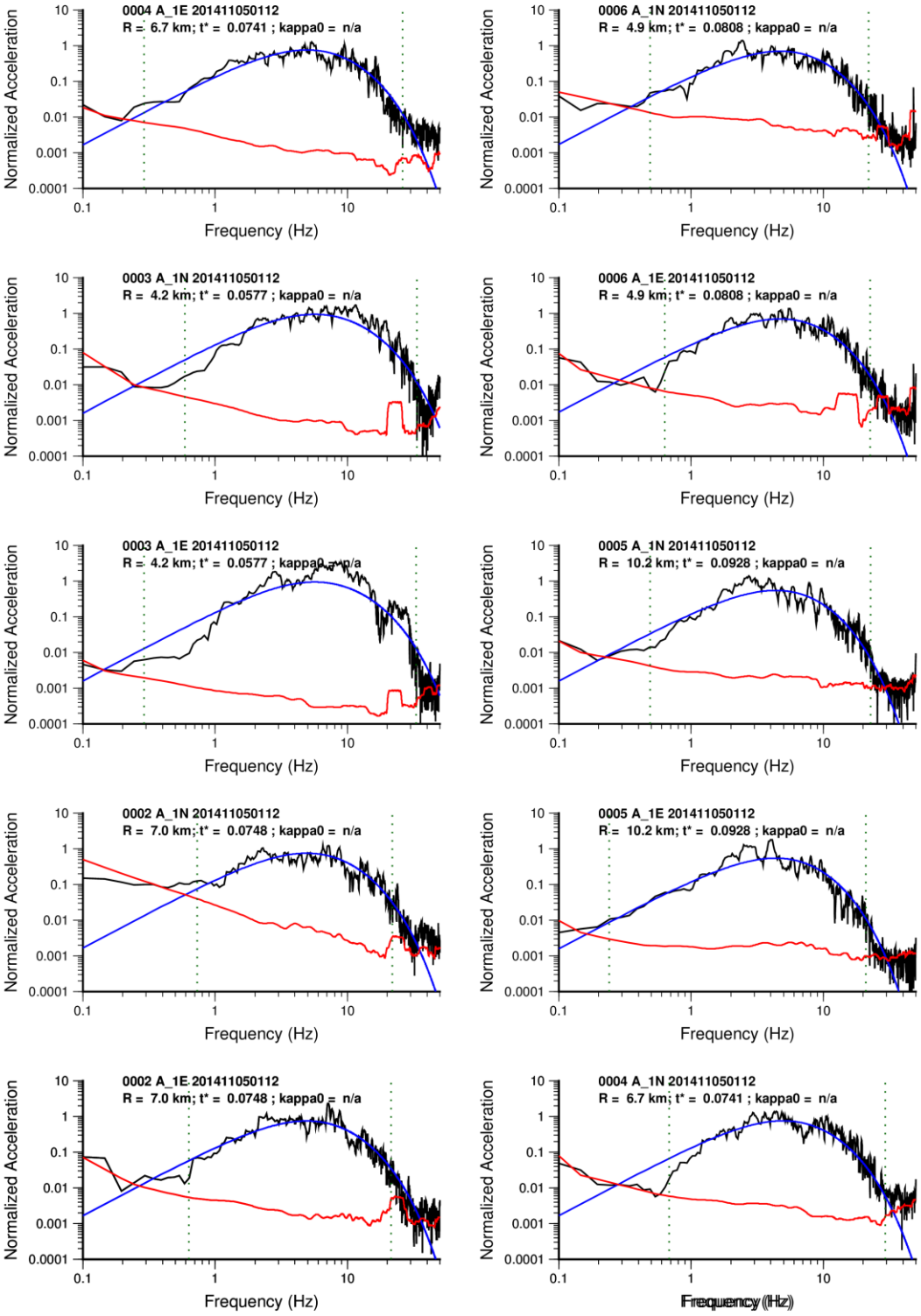


Figure A1.71: log-fit broad-band models for individual horizontal component recordings of EQ-12. Whole path attenuation (t*) is fixed based on the lin fit models. Red: noise, black: earthquake, blue: broadband model.

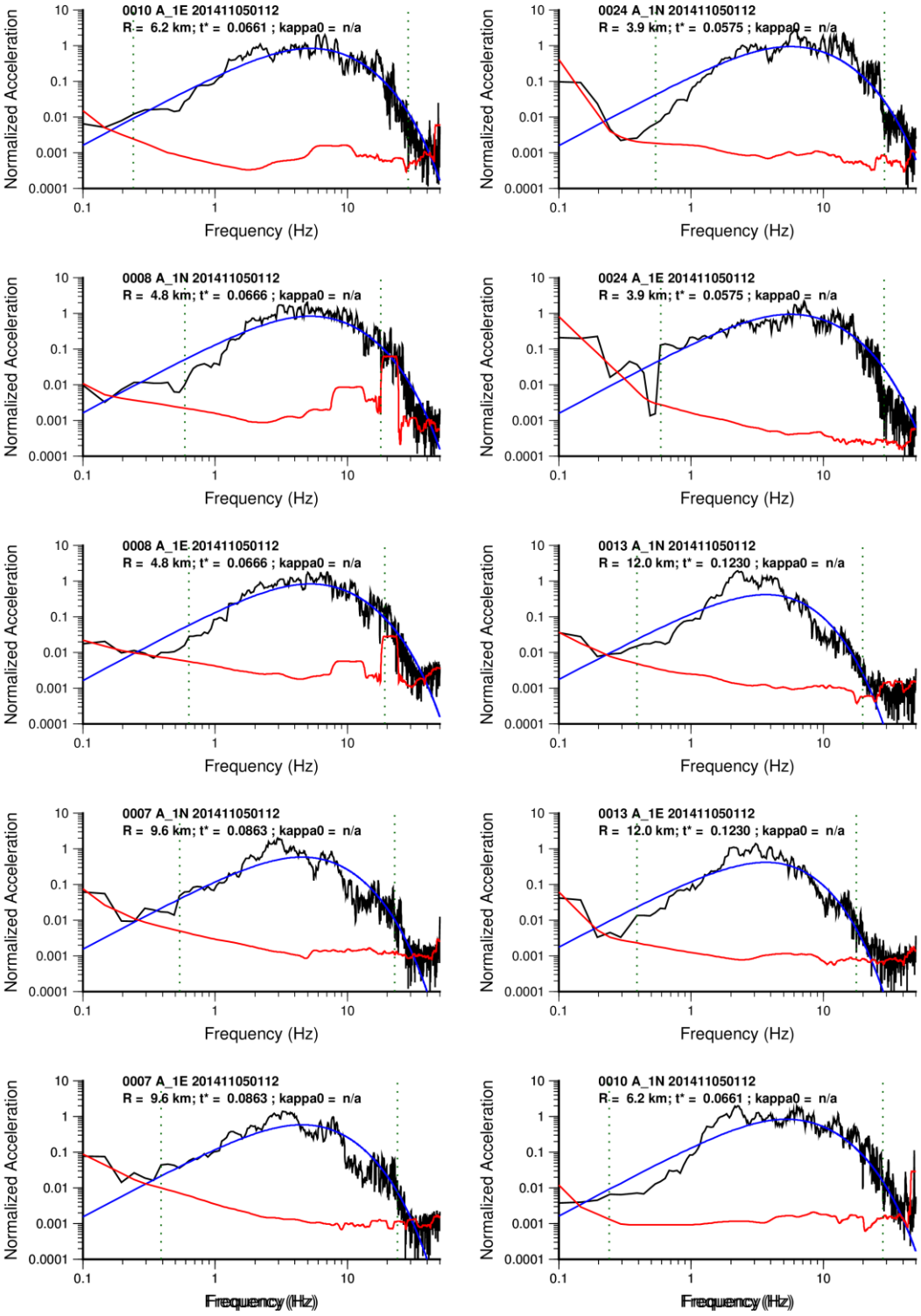


Figure A1.72: log-fit broad-band models for individual horizontal component recordings of EQ-12. Whole path attenuation (t*) is fixed based on the lin fit models. Red: noise, black: earthquake, blue: broadband model.

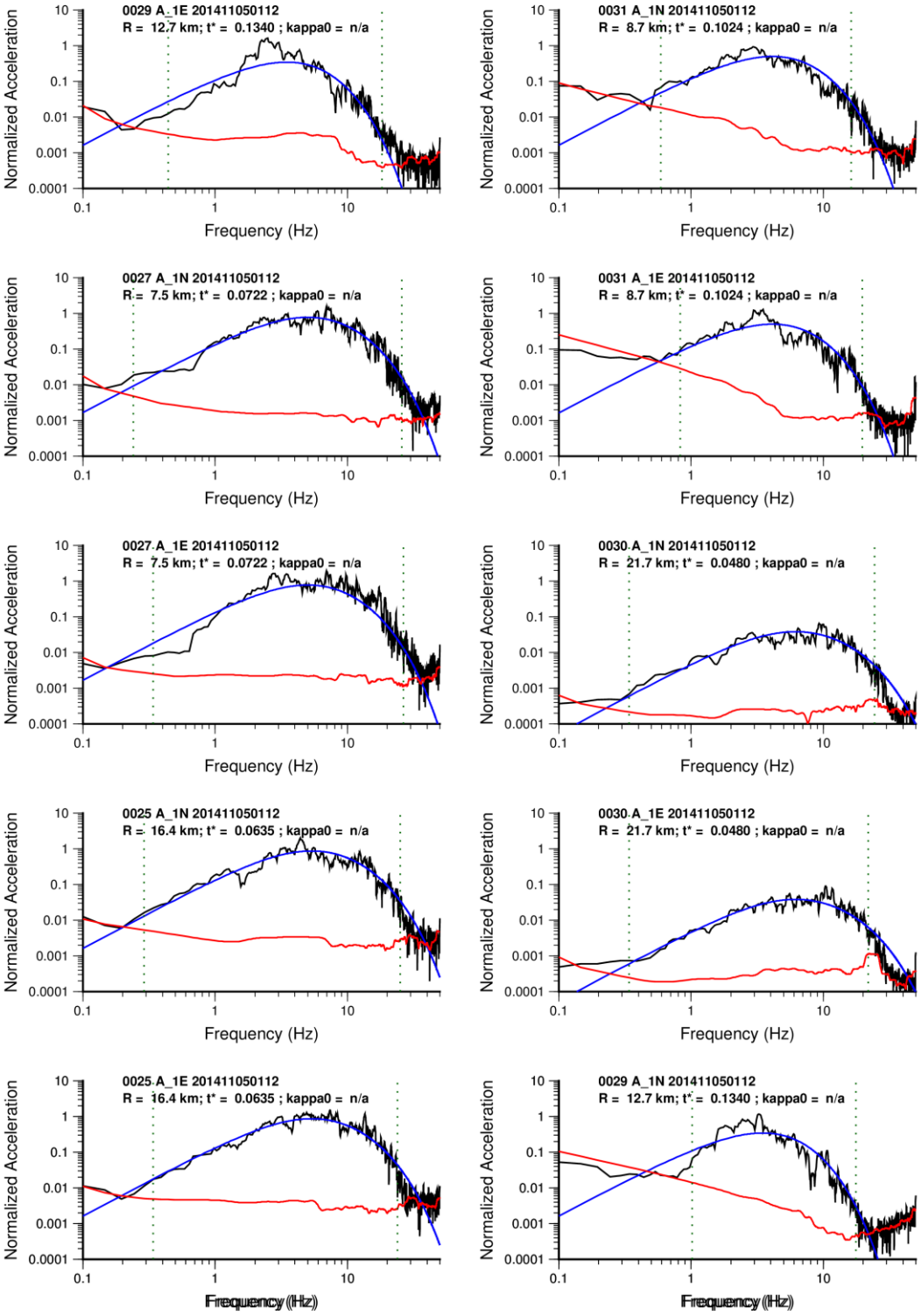


Figure A1.73: log-fit broad-band models for individual horizontal component recordings of EQ-12. Whole path attenuation (t*) is fixed based on the lin fit models. Red: noise, black: earthquake, blue: broadband model.

APPENDIX II

Sigma Penalty in GMPEs Based on Point-Source Distances

A determination of the additional variability of using point-source rather than extended-source distance metrics in GMPEs

Michail Ntinalexis, Peter J Stafford & Julian J Bommer

Version: 25th March 2015

Introduction

The quantification of seismic hazard at any location is made through the use of Probabilistic Seismic Hazard Assessment codes (PSHA). PSHAs are procedures that assess how likely it is that any ground motion value is exceeded at a particular location. This is done by estimating probability distributions of ground motion values that may be caused at that location from seismic events in nearby seismic sources and the recurrence frequency of the earthquake scenario that produces them. These distribution estimations are obtained from ground-motion prediction equations. GMPEs predict a median ground-motion value accompanied by a variation σ, which is the standard deviation of the normal distribution with that mean. The importance of using an accurate value of σ cannot be overstated; it exerts a strong influence on the PSHA, as shown in Figure A2.1 below and explained in Bommer & Abrahamson (2006).

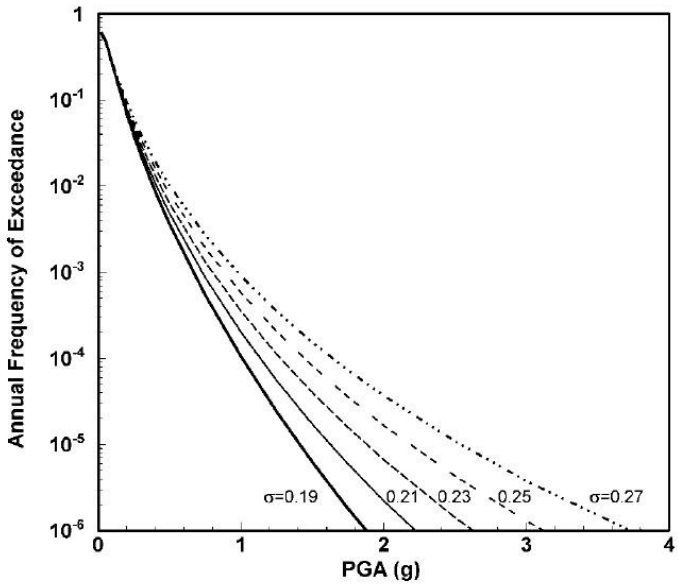


Figure A2.1. An illustration of the influence of the value of σ in PSHA. (Bommer & Abrahamson, 2006).

It is expected that point-source distance-metric based GMPEs should have larger σ values compared to extended-source distance metric based ones for larger magnitudes. However, the σ values that have been obtained for the ASB14 (Akkar *et al.,* 2014*)* GMPEs using the two types of distances are similar. It has therefore been concluded by Akkar *et al*. (2014) that the database on which their set of GMPEs is based lacks recordings from large magnitude events at short distances sufficient to quantify σ for the point-source models accurately. The scope of this paper is to examine and quantify this difference in variation between the two types of models. This will be done by filling the gap in the data artificially through a Monte Carlo simulation of vertical strike-slip ruptures, and by calibrating the predictions of the point-source models against the corresponding ones made by the extended-source ones.

Different source metrics and their use in GMPEs

Two seismic source conceptual representations are generally in use; extended- and point-source representations. In the first type, the source is represented as the entire fault rupture, while, in the second, the source is represented as a single point.

Most GMPEs that are currently employed have been derived by fitting into existing event data of extended-source representations, and therefore extended-source distance metrics are the basis for their predictions. This is a more realistic approach, as taking into account the full extension of the fault rupture reflects the fact that seismic energy is not released solely from a single point but from its full length (Bommer & Akkar, 2012). Such extended source metrics include R_{rup} and R_{JB}. The

first is a measure of the distance of the closest point of the fault rupture to the site, while the latter is a measure of the distance of the closest point of the fault rupture's projection on the Earth's surface to the site. R_{JB} stands for Joyner-Boore distance, as it was introduced by Joyner & Boore (1981).

Point-source distance metrics are easier to obtain and simplify the approach by representing the energy release as being sourced to a single point in space. Such point-source metrics are R_{hyp} and R_{epi}. The first is a measure of the distance from the seismic focus (hypocentre) to the site, while the latter is a measure of the distance from the surface projection of the focus, the epicentre, to the site (Figure A2.2).

The use of R_{rup} and R_{hyp} would produce the most accurate predictions when extended-source and point-source based GMPEs are used respectively, especially R_{hyp} (Bommer & Akkar, 2012), as these distance measures represent the earthquake scenario better in space, taking depth into account. However, R_{JB} is more commonly used for model derivation in Europe, because of lack of data for R_{rup} (Akkar *et al*., 2014). Additionally, both R_{JB} and R_{epi} also simplify the approach by one dimension and their parallel use is in that way consistent. Therefore, this is the pair that will be used in this study.

The use of point-source based GMPEs was common in early GMPEs, until it was shown that the simplification of the point-source representation would cause a severe underestimation of ground motions in some cases. A good example of this is the 1978 Tabas earthquake, where R_{epi} was equal to 57 km from the recording station, but R_{JB} was equal to only 8 km. Thus, the use of any existing R_{epi}–based equation would predict a PGA significantly smaller than the 1*g* that was recorded. More such cases have followed since, to illustrate that the use of extended-source based GMPEs yields more realistic estimates in such cases, or, in other words, to show that the problem of the variability of point-source based predictions is quite significant for larger events and smaller extended-source distances.

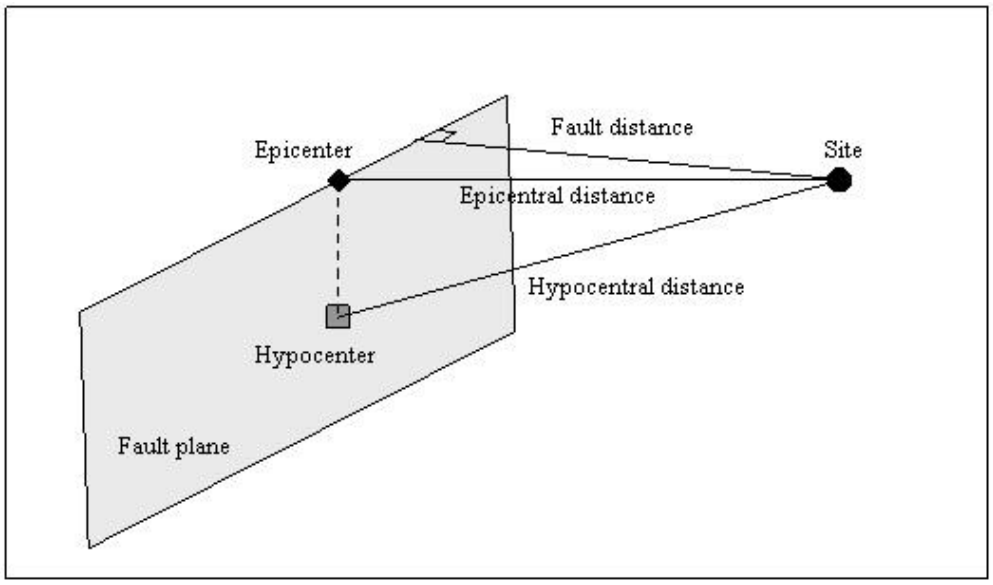


Figure A2.2. A simple representation of the definitions of the different distance metrics (Kentucky Geological Survey, 2012). Please note that in this figure, for a vertically-dipping fault place, R_{JB} is effectively the same as “fault distance”.

However, a problem arises when such, extended-source metric based GMPEs are used by PSHAs in which seismic sources are often represented as points (Bommer & Akkar, 2012). PSHA codes assess the seismic hazard in a location by assessing the ground motions that may be caused by different seismic sources around that location. For any particular site, nearby seismic sources considered by PSHAs consist of (a) known seismogenic faults and (b) areas with distributed seismic activity in them. Since the space characteristics of the first type of sources are known, a calculation of the extended-source distance to the site is possible. That distance can be used in the extended-source based GMPEs for the assessment without problem. On the other hand, in the second type of sources, the finite earthquake sources that may contribute to the seismicity of the area are represented as points within the area. The source-to-site distances calculated for these are therefore in effect point-source distances. It goes without saying that GMPEs are expected to yield accurate results only when used with the metric that they were derived with. Many PSHA codes, however, use even this second set of distances on the equations derived for extended-source distances, with which they are not compatible. A serious problem of inconsistency therefore arises.

A number of solutions to this problem have been suggested. The first solution, as described by Bommer & Akkar (2012), involves translating the point-source representations of the PSHA codes to extended-source representations that produce valid metrics for use with the GMPEs, by virtually simulating faults on that location. This simulation can be achieved through empirical length-magnitude relationships such as those presented in Wells & Coppersmith (1994) and Leonard (2010) and then by estimating or assuming the other characteristics of the rupture, such as the location of the epicentre relative to the rupture bounds, the rupture orientation,

whether the rupture is bilateral. as well as whether it can exceed the boundary of the seismic source zone. When the metric in question is R_{rup}, the depth must also be decided before the new extended-source distance is calculated. Although computationally expensive, most good PSHA codes that offer this option.

The second solution has been derived by Scherbaum *et al*. (2004) as an alternative approach to the first solution. Scherbaum *et al*. (2004) derived empirical conversion relationships between the different metrics, to avoid the need of the sigma penalty.

The third solution is proposed by Bommer & Akkar (2012) and it involves simply avoiding any conversion by deriving all future GMPEs in several versions to have both point-source and extended-source options. This recommendation was since carried out in Akkar *et al*. (2014) and Bindi *et al*. (2014). However, as Akkar *et al*. (2014) remark, the point-source based models are expected to be less accurate and have a clearly higher variability in their results due to the space and energy model simplification that their source representation is based on.

Quantifying the additional variability of point-source based GMPEs

As explained before, extended-source representations are more realistic and therefore the corresponding GMPEs yield more accurate predictions, because of their better integrated approach connecting rupture space characteristics and seismic energy radiation. In the GMPEs of Akkar *et al*. (2014), however, the R_{JB} model appears to have approximately the same variability (σ) as the R_{epi} one. The reason for this theoretical discrepancy with what is observed is attributed to the fact that the database on which the equations of Akkar *et al.* (2014) is based lacks sufficient data for large magnitude events and at short distances.

The objective is to examine the difference in the variability of the predictions of the two models, henceforth defined as Δσ, in order to obtain a more accurate estimate of the variability of the R_{epi} model, defined here as σ_{epi} as:

$$\sigma_{epi} = \sqrt{\sigma_{JB}^2 + \Delta\sigma^2} \qquad \text{(A2.1)}$$

where σ_{JB} is the variation of the R_{JB} model.

An option for attempting to explore this difference in the variabilities is to simulate virtual faults to fill the existing data gap artificially, and then assess the variability difference by comparing the predictions made in virtual recording sites by the different models in the same scenario. More specifically, since we have decided to use R_{JB} and R_{epi}, the ASB14 GMPEs for use with R_{JB} will be used as the reference point, while the predictions made by using the equivalent R_{epi} model will be

compared against it. The residuals of the R_{epi} predictions versus the R_{JB} predictions will be then used to calculate the additional variability.

Akkar *et al*. (2014) suggest that this simulation be done by creating a dense grid of recording points around the virtual rupture and obtaining the predictions at those points, as illustrated in Figure A2.3. The approach of this study is similar, but instead of a fixed rupture and multiple recording sites, we have a single fixed recording site that sits at the centre of a circular source area with a single background source, with respect to which multiple ruptures will be simulated. As shown in Figure A2.4, we simulate the different scenarios of vertical strike-slip faulting by placing the epicentre of the source to consecutive set radial distances from the site, and then generating different possible ruptures there. Different configurations of the fault rupture space characteristics are then simulated to obtain different distances and the corresponding predictions. In this way, virtually any scenario in the areal source can be simulated This approach also enables us to derive the Uniform Hazard Spectra after the σ values have been determined. This setting and the simulation procedure, executed in Matlab, are described below in detail.

Simulation procedure

The simulation of the different scenarios is carried out by determining the combinations of the various parameters that define the scenario and the space characteristics of the corresponding fault at each epicentral location. R_{JB} can then be calculated and the predictions for both models can be made. Shear-wave velocity V_{S30}, that reflects the local site conditions, will be considered constant for all scenarios and at all places of the virtual source area at the ASB14 reference value of 750m/s.

Moment magnitude, M_w, is the first defining independent parameter. While magnitude does not directly modify the space configuration, it dictates what the rupture length will be and also influences the predictions. An initial choice of magnitude range for the simulations would be from 4 to 7.6, in order to match the database which the ASB14 equations were derived from. However, it is recognised that the ASB14 database lacked sufficient data for M_w>6 events, which is the reason that it is expected that the variability of the point-source models is underestimated. Additionally, Akkar *et al.* (2014) suggest that the ASB14 GMPEs are applicable to a magnitude range of 4 to 8. For these reasons, therefore, the moment magnitude boundaries for our simulations will be 4 and 8. It must be noted here that one could argue that, since we are missing data for events larger than M_w=6, we should only simulate for that magnitude range. This, however, is not followed, as the upper end of the magnitude range should show larger variability for the point-source models and thus could bias the calculation of sigma, while the objective is producing a model

of determining the new σ, that will be compatible for use with the full range of magnitude applicability of ASB14.

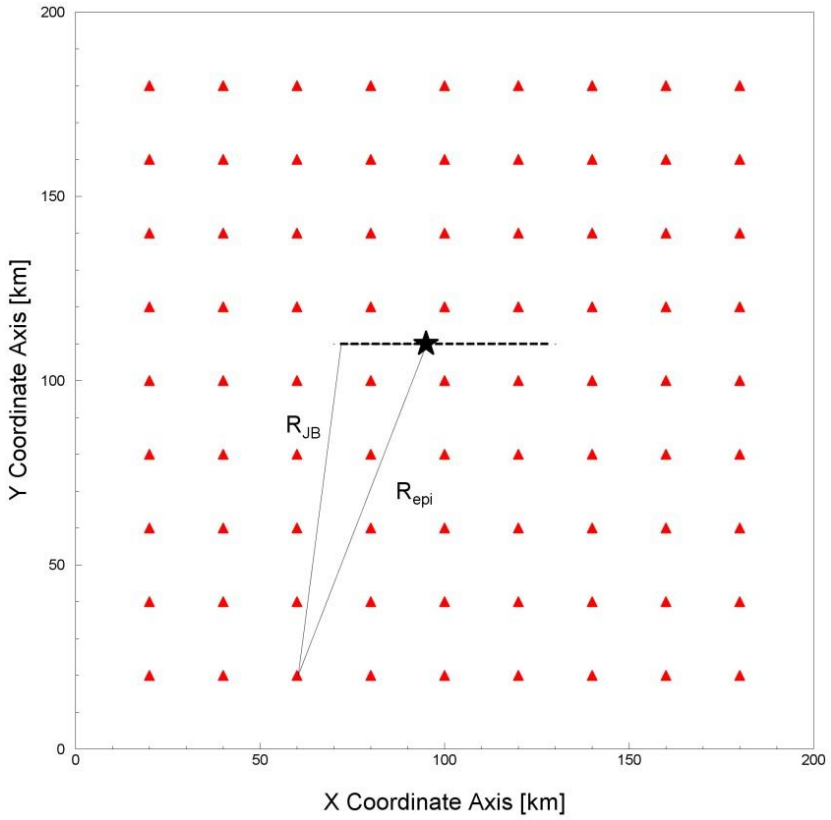


Figure A2.3. A simplified conceptual illustration of an adaption of the grid representation proposed by Akkar *et al.* (2014) with an example with the definition of R_{JB} and R_{epi} measured from one of the sites of the grid. The epicentre is marked as a black star while the sites/recording stations are marked with red triangles. Note that the fault rupture would have to be simulated by discrete elements to calculate the distances in this approach.

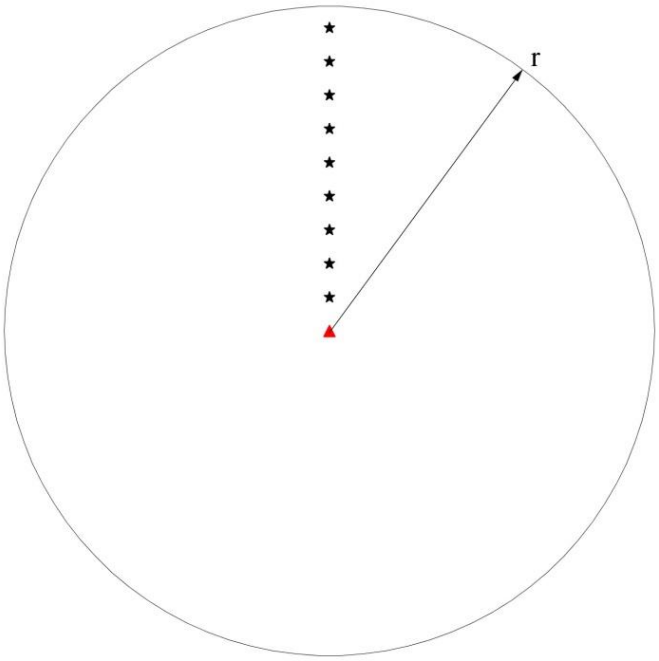


Figure A2.4. A conceptual illustration of the simulation model used in this study. Different possible epicentres are marked as black stars while the site/recording station is marked with a red triangle. Note that, by using the radial coordinate and setting uniform ground conditions we simplify the problem by representing all scenarios of the area by performing simulations over a single radius.

To determine the rupture length given the magnitude scenario, we can use the length-magnitude relationships derived by Wells & Coppersmith (1994) or Leonard (2010). Since they are simpler and more straightforward to use, we choose to use the relationships of Wells & Coppersmith (1994) and more specifically the relationship developed for subsurface rupture length (RLD) which is as follows:

$$\log(RLD) = 0.59M_w - 2.44 \quad \text{(A2.2)}$$

It must be noted that Wells & Coppersmith (1994) also provide individual equations of this form for separate styles of faulting. However, the use of a single general equation is simpler and, following their own suggestion, more exact, as it is based on a larger database than those of the three individual equations; the sum of the databases of the three. The standard deviation for Eq. (A2.2) is given by Wells & Coppersmith (1994) as 0.16 and is indeed smaller than those of the individual equations.

To simulate as many probable scenarios as possible, we must obtain a few representative length values. This is done by making a number of discrete approximations of the rupture length distribution, assuming that rupture lengths are log-normally distributed and using the method of Miller & Rice (1983) to obtain the ε values and the corresponding probabilities for the number of discrete approximations

that we choose. Epsilon ε represents the number of standard deviations (σ) that a value differs from the mean. We choose an odd number of approximations so that the mean is included.

The second defining independent parameter is the epicentral distance R_{epi}, which is pre-defined in each simulation. The only decision to make with respect to the epicentre is the range of the positions of the epicenters. For compatibility with the database of ASB14 and to avoid biasing the results, the epicentral distances and the boundary of the area will reach a radius of 200km from the site, which is the maximum distance of the ASB14 database. Akkar *et al.* (2014) suggest that the GMPEs are applicable for a range greater than 200km; however we choose to not extend our limit further than 200km. Additionally, we choose to allow for leaky boundaries of the source area so that the rupture scenarios will not be limited by excluding those that cross the boundaries.

The location of the epicentre on the rupture and the orientation of the rupture are the another two defining independent parameters that are necessary and enough to describe the location of the fault rupture within the source area and thus simulate all possible scenarios. Their definition is described below and shown in Figure A2.5.

The location of the epicentre, x_{epi}, is simply the distance of the epicentre to the rupture end. For consistency, we define that end to always be the one that is the «southern" end. To avoid complicated definitions and programming that would require re-examining the rupture length constantly, x_{epi} will be obtained as a portion of L at set intervals by defining a ratio α which varies from 0 to 1, as follows:

$$x_{epi} = a \times L \tag{A2.3}$$

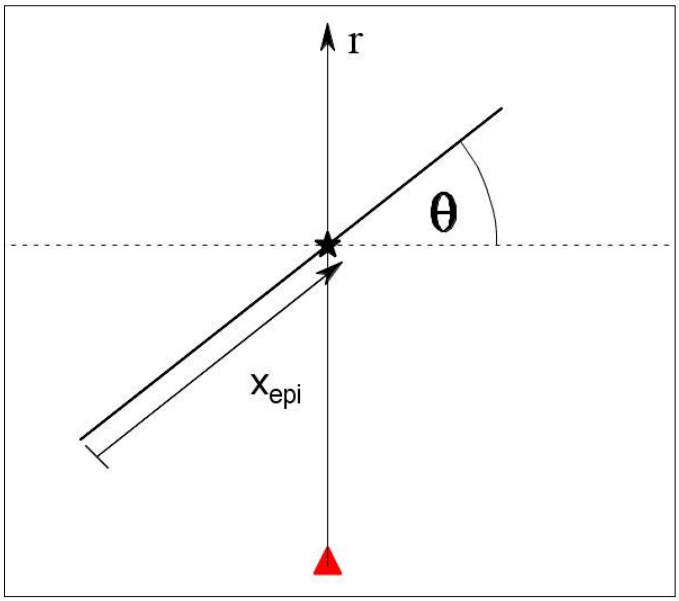


Figure A2.5. An illustration of the definitions of the epicentral location and the angle θ. X_{epi} is defined as the distance of the epicentre to the rupture end that is closer to the site (*i.e.* in the lower hemisphere of the virtual Cartesian system centred on the epicentre), while θ is the counter-clockwise angle of the rupture to the virtual x-axis.

The parameter varying in the simulations will therefore be α, from which x_{epi} will be calculated separately for each simulation.

The orientation of the rupture will be represented by an angle θ. Since the rupture is anchored at the epicenter and all other parameters are defined independently of θ, whichever axis we define θ against is not important, as the results will be the same. For the purpose of clarity of illustration and for use in our calculations later on, however, we define θ as the angle of the fault orientation with respect to a virtual x-axis counter-clockwise, as shown in Figure A2.5. This virtual x-axis is defined as the axis of a Cartesian system centered on the epicentre, with the radial coordinate axis of the areal source as the y-axis of the system, *i.e.* with the virtual x-axis perpendicular to the radial coordinate axis.

It can be easily shown that the definitions of θ and x_{epi} allow us to thus describe all possible scenarios by ranging θ from 0^o to 180^o. Additionally, it can also be shown that, as illustrated in Figure A2.6, due to symmetry, a θ ranging from 0^o to 90^o is enough to calculate the results for all scenarios. This is an important simplification that will make the simulation significantly less computationally intense.

While the basis of these calculations is simulating the fault ruptures to obtain the necessary distances, simulating the fault ruptures in terms of discrete of finite elements will not be necessary in this study. Instead, the distances are calculated through geometrical arguments as described and illustrated in Figure A2.7 and Equations (A2.4) and (A2.5).

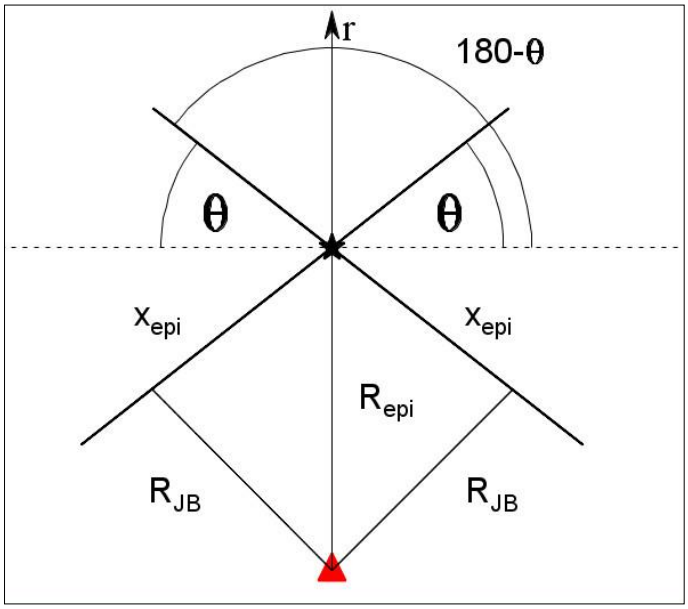


Figure A2.6. When two fault ruptures are symmetrical and have equal x_{epi}, then all other parameters, such as R_{JB}, are also equal. Therefore the results for all scenarios can be obtained by varying θ simply from 0 to 90°.

From the above definitions it follows that:

$$R_{JB}^2 = R_{epi}^2 + \gamma^2 - 2R_{epi}\,\gamma\cos\left(\frac{\pi}{2} - \theta\right) \quad \text{(A2.4)}$$

where

$$\gamma = \min[aL, R_{epi}\sin(\theta)] \quad \text{(A2.5)}$$

Simulation results

Using this method it is possible to simulate many possible scenarios, thus providing a significant range of variation of calculated R_{JB} values for all scenarios with set values of R_{epi}. The resulting R_{JB} values range from zero to the original R_{epi} value itself, which is by definition the upper bound for R_{JB}. The different ground-motion predictions of each model for each scenario have been calculated using the corresponding distance-metric-consistent models of ASB14. Since the aim is to develop a model of calculating the additional variability for the R_{epi} model, this variation-difference model has to be usable with R_{epi} as its parameter. Therefore the variation of the results has been calculated by comparing the prediction of the R_{epi} model for each given R_{epi} to the predictions of the R_{JB} model for all R_{JB} values calculated for that particular R_{epi} value, per magnitude scenario. The results for different magnitudes are presented in Figures A2.8-A2.12 below for PGA and PSA at periods of 0.2, 0.5, 1 and 2s.

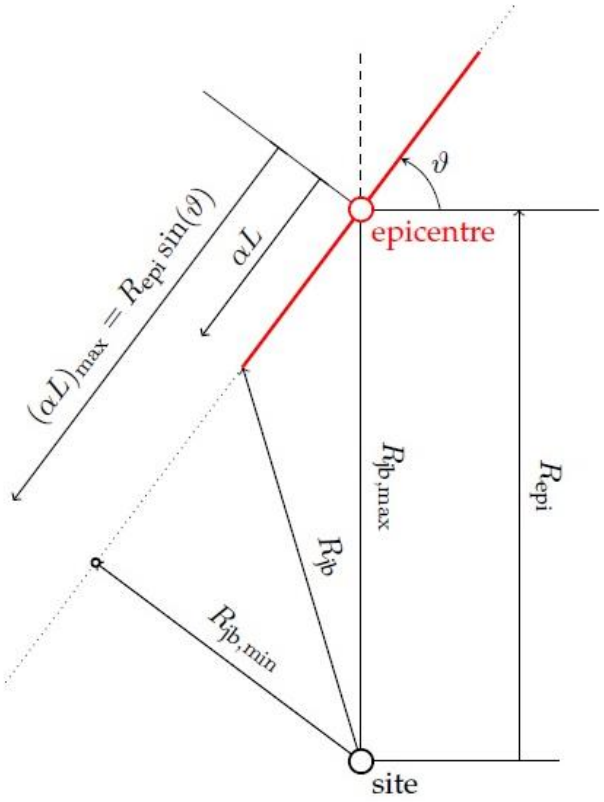


Figure A2.7. Generic geometry for simulation of vertical strike-slip ruptures.

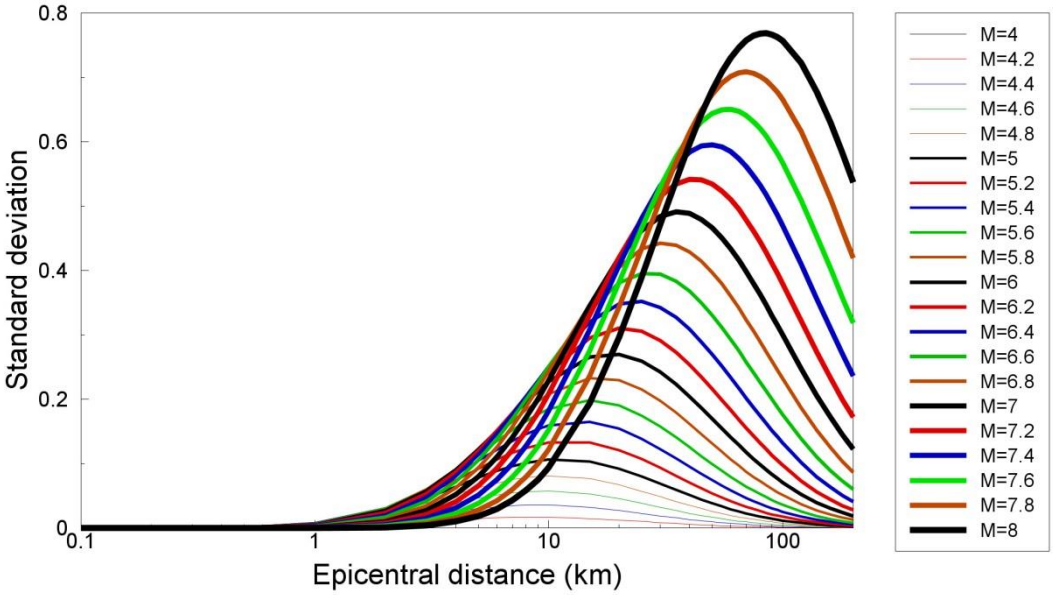


Figure A2.8. Additional variation of the R_{epi} model for PGA.

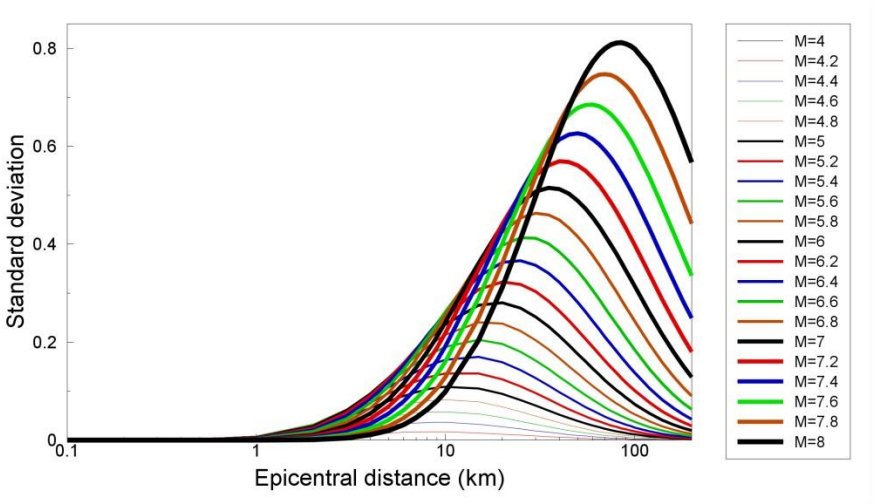

Standard deviation
0.8
0.6
0.4
0.2
0
0.1
1
10
100
Epicentral distance (km)
M=4
M=4.2
M=4.4
M=4.6
M=4.8
M=5
M=5.2
M=5.4
M=5.6
M=5.8
M=6
M=6.2
M=6.4
M=6.6
M=6.8
M=7
M=7.2
M=7.4
M=7.6
M=7.8
M=8

Figure A2.9. Additional variation of the R_{epi} model for the spectral acceleration for the period of 0.2s.

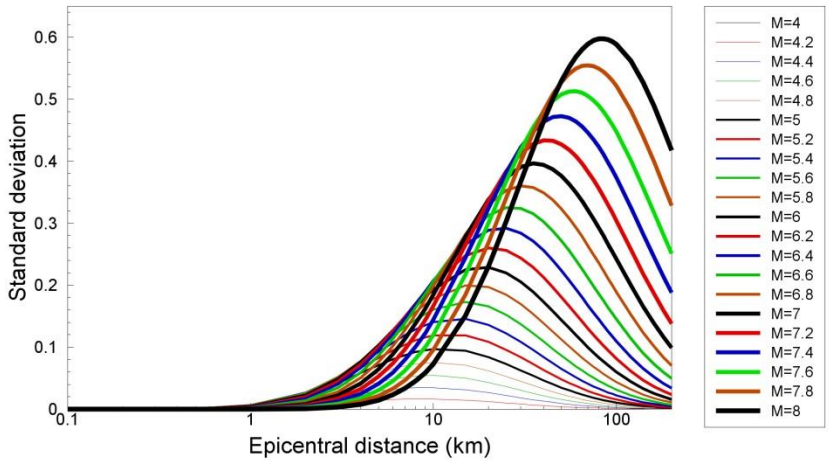


Figure A2.10. Additional variation of the R_{epi} model for the spectral acceleration for the period of 0.5s.

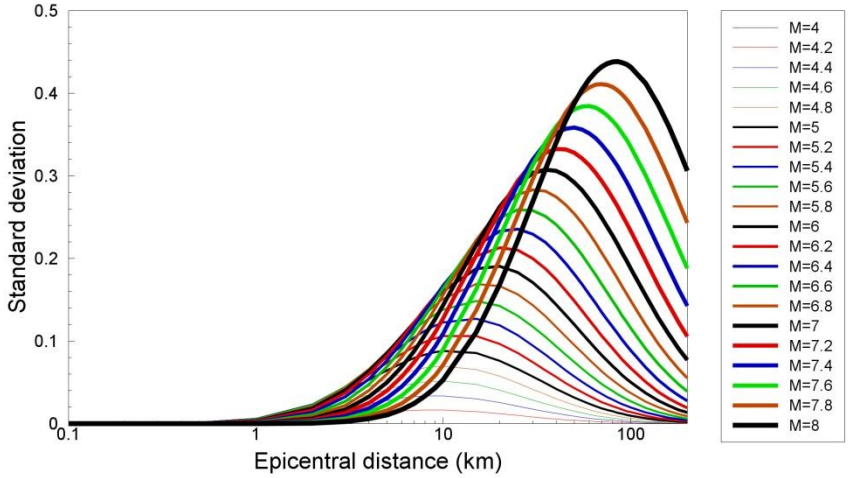


Figure A2.11. Additional variation of the R_{epi} model for the spectral acceleration for the period of 1s.

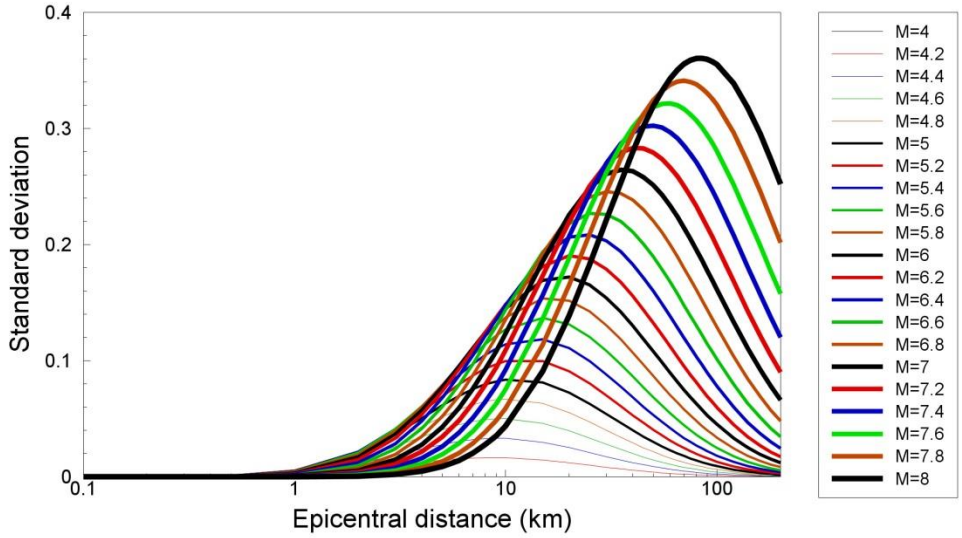


Figure A2.12. Additional variation of the R_{epi} model for the spectral acceleration for the period of 2s.

From Figures A2.8-A2.12 above, we observe, as expected, that the variation shows a strong dependence on magnitude and distance. More specifically, the curve appears to approximate the shape of the normal distribution (a bell curve) with the mean of the log-normal distribution (with distance) increasing with magnitude. On this base, a simple formulation was developed to represent the adjustment to the within-event variability resulting from the geometric effect of modelling the source of earthquakes of extended rupture as a single point. The magnitude- and distance-dependent adjustment to the intra-event variability is defined as follows:

$$\delta\phi = SF.\frac{\varphi(z)}{\sigma_Z} \qquad M \geq 4 \quad \text{and} \quad R_{epi} > 0 \qquad \text{(A2.6a)}$$

$$\delta\phi = 0 \qquad M < 4 \quad \text{or} \quad R_{epi} = 0 \qquad \text{(A2.6b)}$$

where SF is the magnitude-dependent scaling factor, expressed as follows:

$$SF = \beta_1(M-4) + \beta_2(M-4)^2 \qquad \text{(A2.7)}$$

and $\varphi[\ ]$ is the normal probability density function, which is given by the following expression:

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-z^2}{2}\right) \qquad \text{(A2.8)}$$

The argument of this expression is given by:

$$z = \frac{\ln(R_{epi}) - \mu_Z}{\sigma_Z} \quad \text{(A2.9)}$$

and the parameters of this expression are given by:

$$\mu_Z = \beta_3 + \beta_4(M - 6.75) + \beta_5(M - 6.75)^2 \quad \text{(A2.10)}$$

$$\sigma_Z = \beta_6 \quad \text{(A2.11)}$$

The coefficients of Eqs.(A2.6 to A2.11) are presented in Table A2.1. The fit of Equation (A2.6) to the observed data produced by the simulations is presented in Figures A2.13-A2.17.

Table A2.1. Coefficients of model for adjustment of within-event variability

	PGA	**Sa(0.2s)**	**Sa(0.5s)**	**Sa(1.0s)**	**Sa(2.0s)**
β_1	0.20380	0.20284	0.20761	0.21116	0.21290
β_2	0.073419	0.080624	0.044808	0.018152	0.005130
β_3	3.39511	3.39511	3.39511	3.39511	3.39511
β_4	0.70978	0.70978	0.70978	0.70978	0.70978
β_5	0.0090045	0.0090045	0.0090045	0.0090045	0.0090045
β_6	1.03275	1.03275	1.03275	1.03275	1.03275

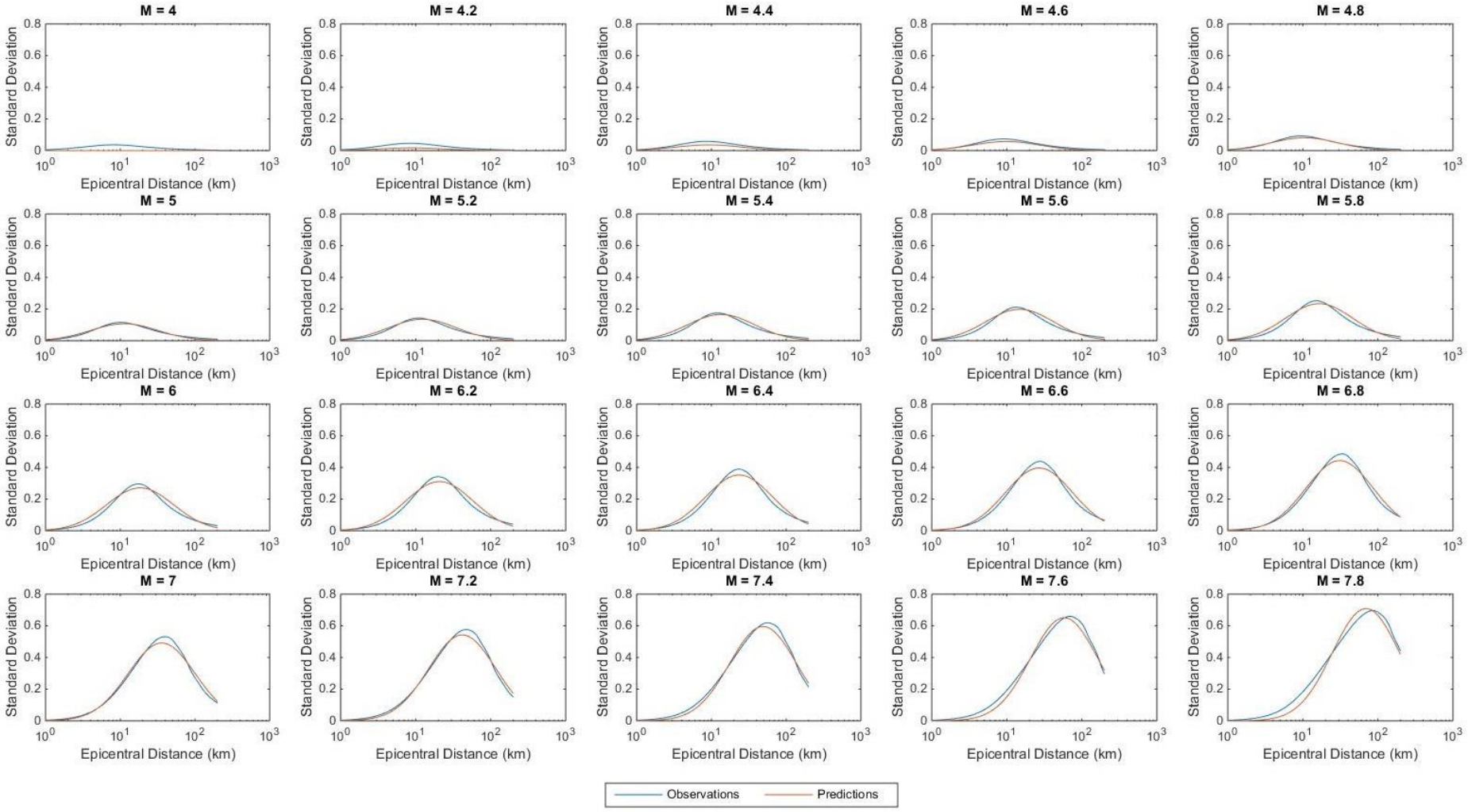


Figure A2.131. Observed data and model predictions for T=0s (PGA).

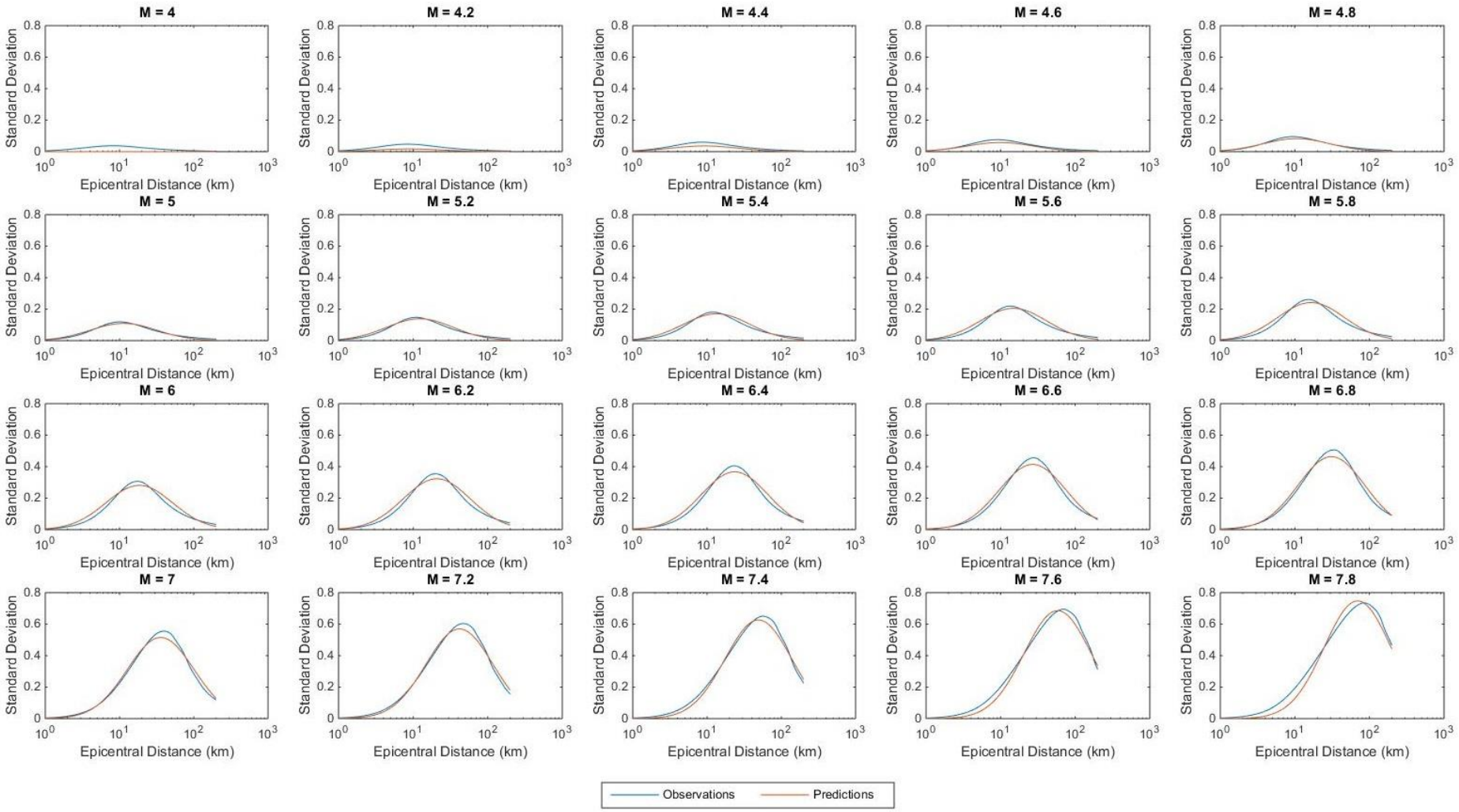


Figure A2.142. Observed data and model predictions for T=0.2s.

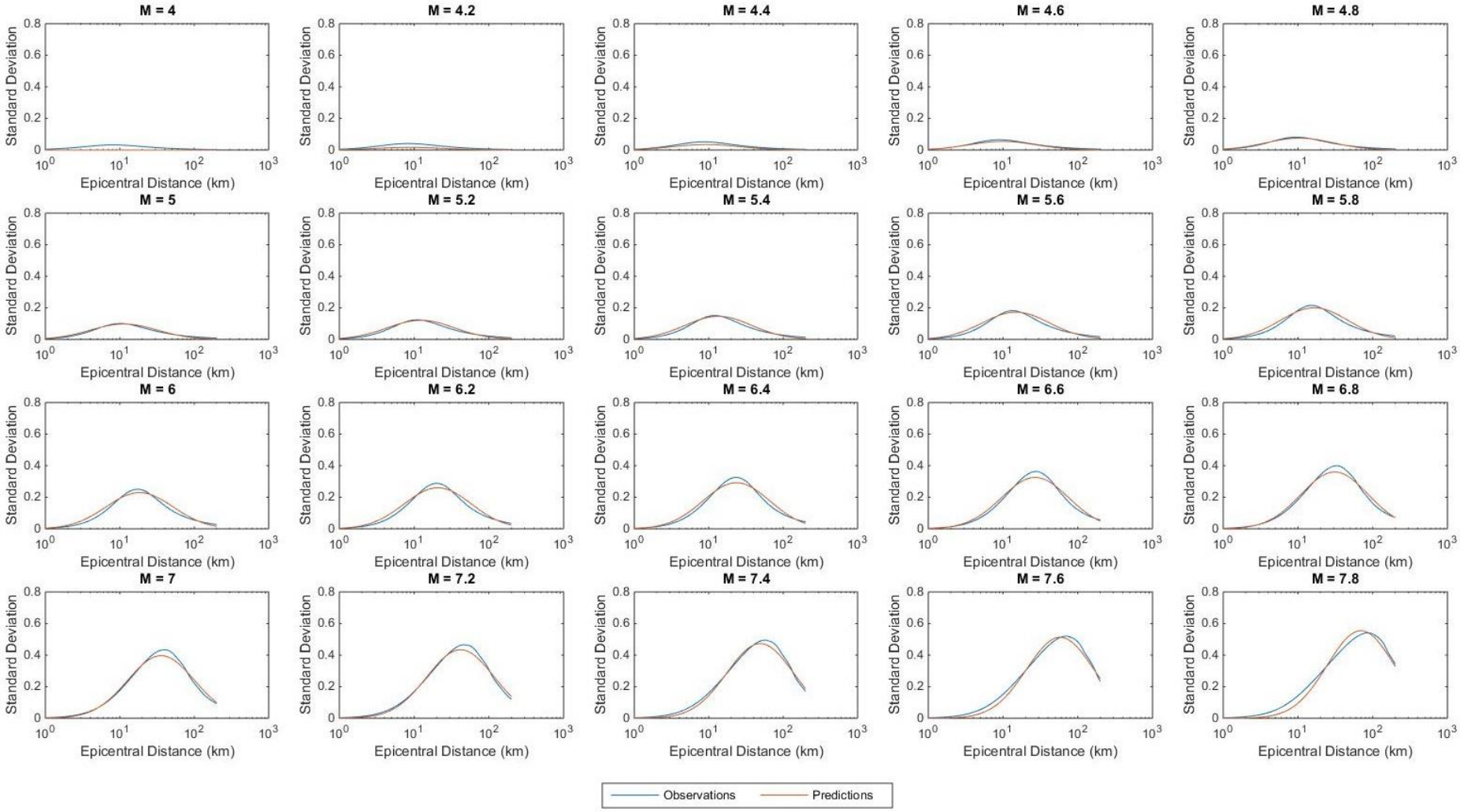


Figure A2.15. Observed data and model predictions for T=0.5s.

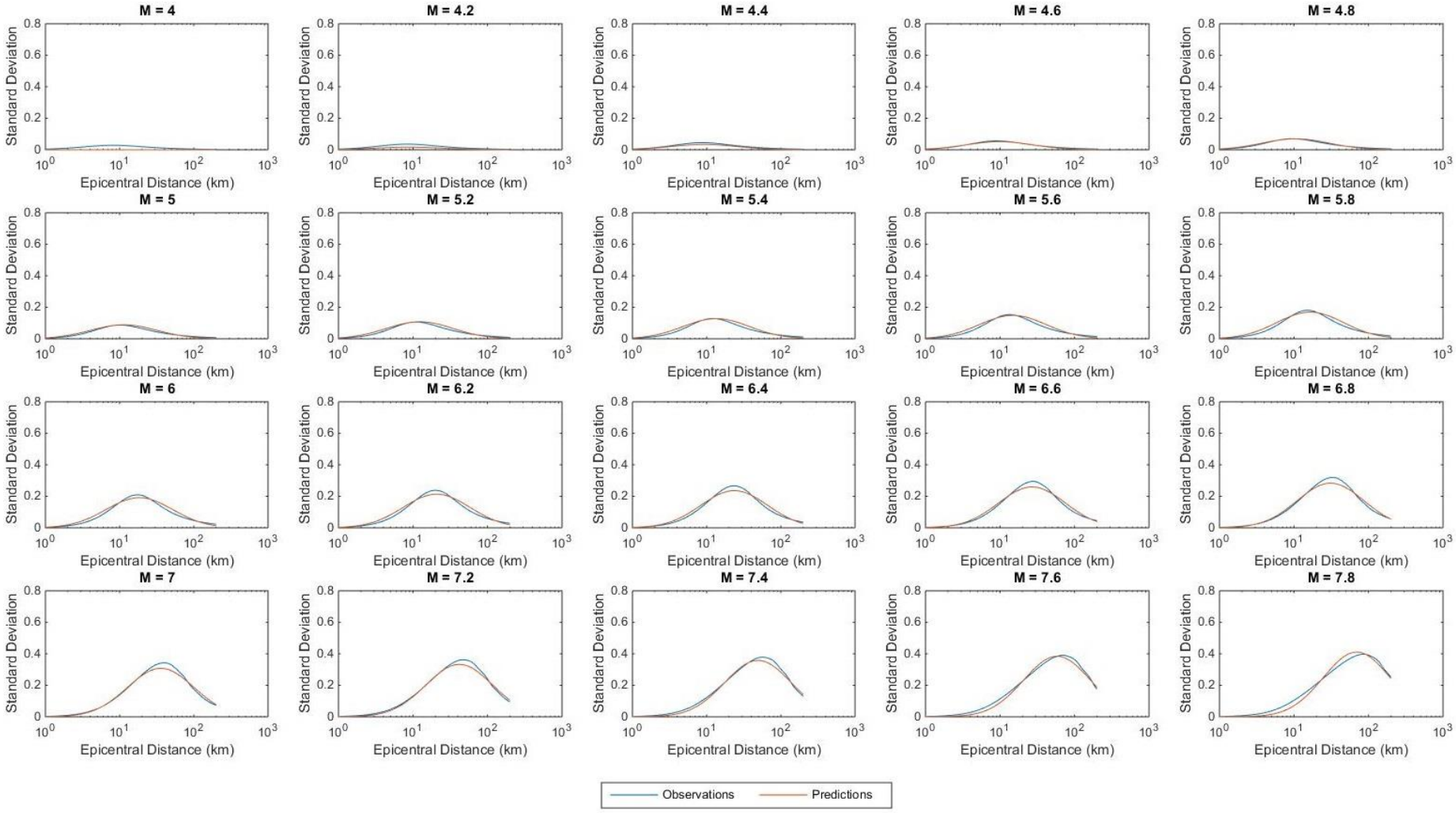


Figure A2.16. Observed data and model predictions for T=1s.

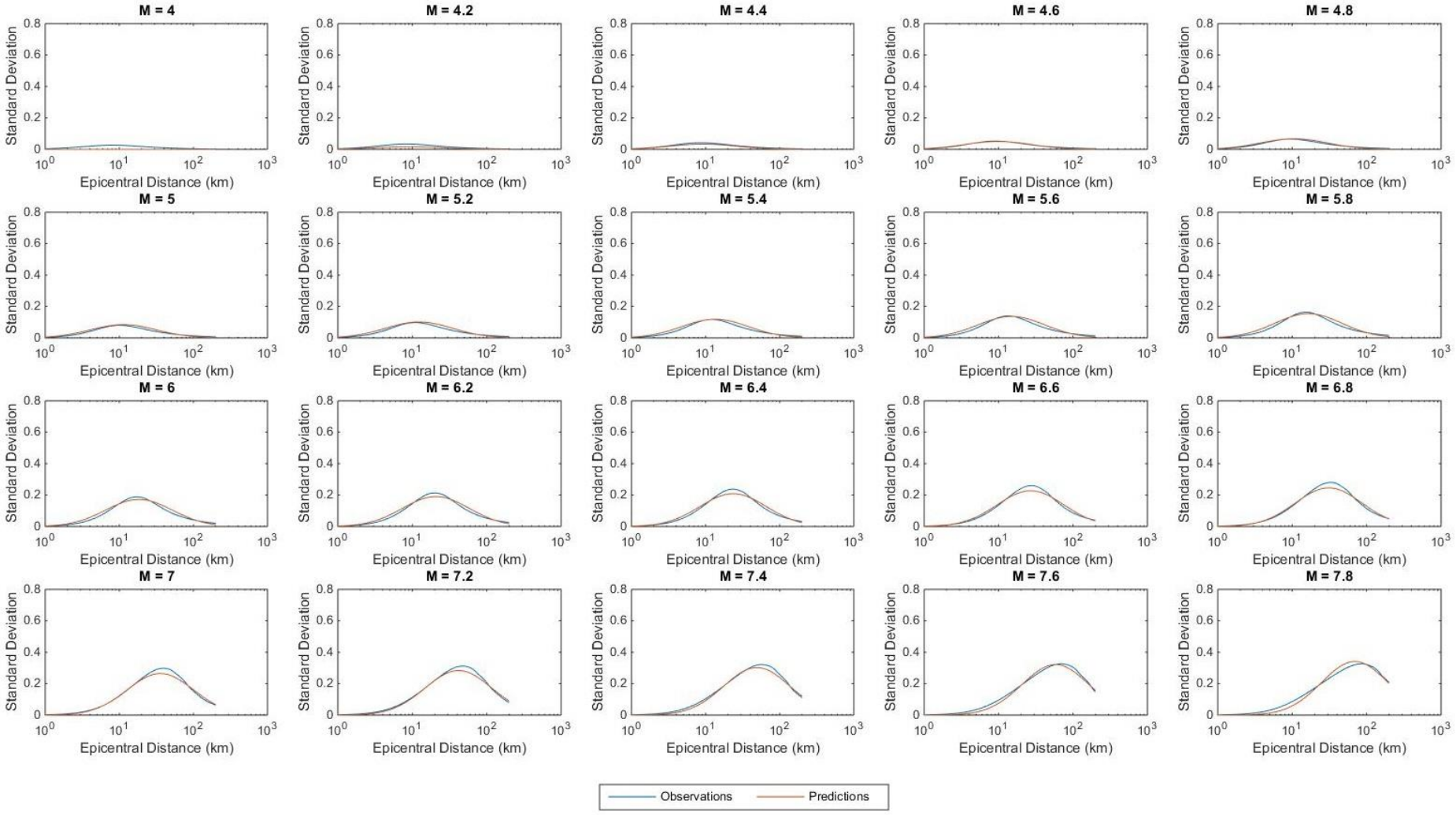


Figure A2.17. Observed data and model predictions for T=2s.

References

Akkar, S., M.A. Sandıkkaya & J.J. Bommer (2014). Empirical ground-motion models for point- and extended-source crustal earthquake scenarios in Europe and the Middle East. *Bulletin of Earthquake Engineering* **12**(1), 359-387. *Erratum*: **12**(1), 389-390.

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

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

Bommer, J. J. & S. Akkar (2012). Consistent source-to-site distance metrics in ground-motion prediction equations and seismic source models for PSHA. *Earthquake Spectra* **28**(1), 1-15.

Joyner, W. B. & D.M. Boore (1981). Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake. *Bulletin of the Seismological Society of America* **71**, 2011-2038.

Kentucky Geological Survey, U.o.K., (2012). *Kentucky Geological Survey - University of Kentucky.* [Online] Available at [Accessed 10 12 2014]: http://www.uky.edu/KGS/geologichazards/research_psha.htm.

Leonard, M. (2010). Earthquake fault scaling: Self-consistent relating of rupture length, width, average displacement, and moment release. *Bulletin of the Seismological Society of America* **100**(5A), 1971–1988.

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

Wells, D. L. & K. J. Coppersmith (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bulletin of the Seismological Society of America* **84**, 974-1002.