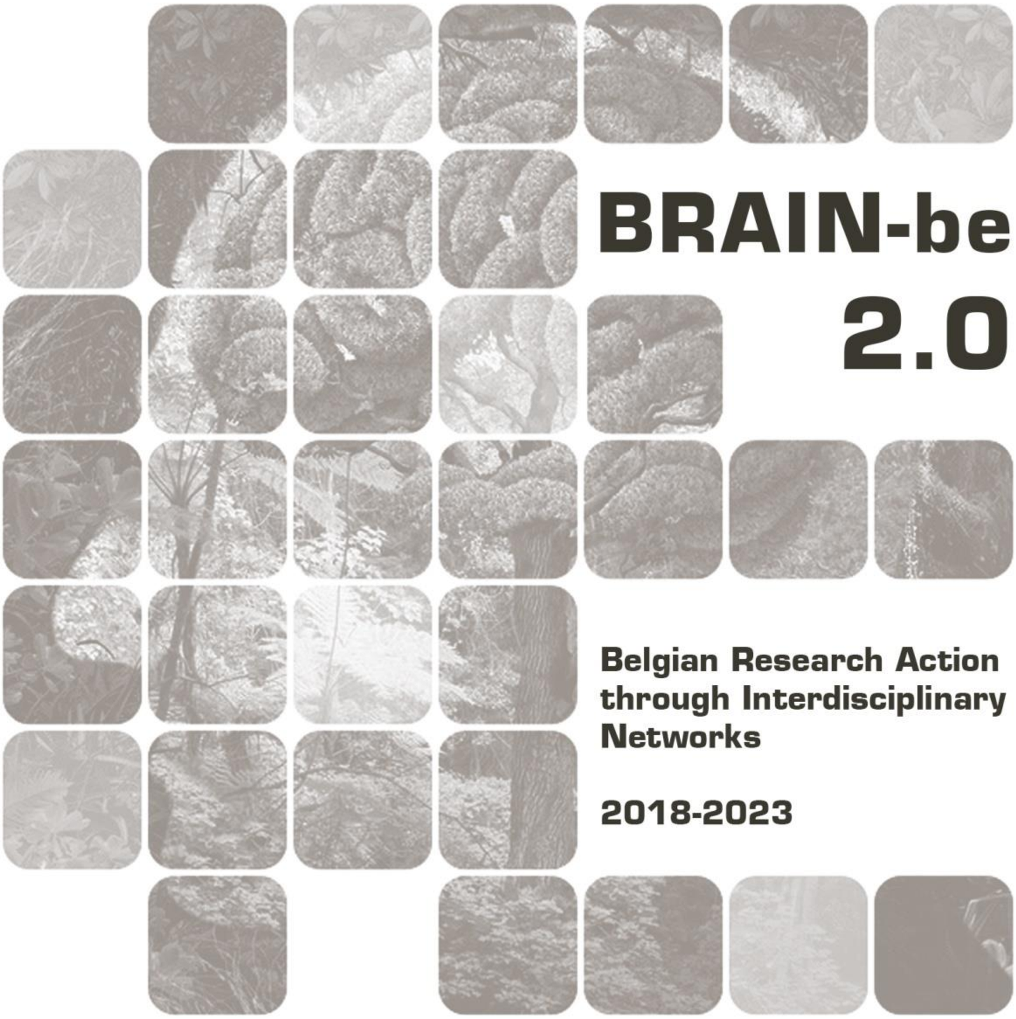


BRAIN-be
2.0

Belgian Research Action
through Interdisciplinary
Networks

2018-2023

BELSHAKE

Earthquake ground-motion database and modeling in Belgium

Kris Vanneste & Mahsa Onvani (Royal Observatory of Belgium)

Pillar 1: Challenges and knowledge of the living and non-living world

BOTTOM-UP PROJECT

BELSHAKE
Earthquake ground-motion database and modeling in Belgium

Contract - B2/202/P1

FINAL REPORT

PROMOTORS: Kris Vanneste (Royal Observatory of Belgium)

AUTHORS: Kris Vanneste (Royal Observatory of Belgium)
Mahsa Onvani (Royal Observatory of Belgium)

DOI: XXX

Published in 2026 by the Belgian Science Policy Office
WTCIII
Simon Bolivarlaan 30 bus 7
Boulevard Simon Bolivar 30 bte 7
B-1000 Brussels
Belgium
Tel: +32 (0)2 238 34 11
http://www.belspo.be
http://www.belspo.be/brain-be

Contact person: Koen Lefever
Tel: +32 (0)2 238 35 51

Vanneste K. & Onvani M. ***BELSHAKE – Earthquake ground-motion database and modeling in Belgium***. Final Report. Brussels: Belgian Science Policy Office 2026 – 138 p. (BRAIN-be 2.0 - (Belgian Research Action through Interdisciplinary Networks))

TABLE OF CONTENTS

ABSTRACT

1.1. Context

Belgium, and NW Europe in general, is situated in a stable continental or intraplate region, characterized by low to moderate seismic activity. Nevertheless, damaging earthquakes have occurred during the 20th century. The most important ones are the 1938 Zulzeke-Nukerke earthquake (magnitude $M = 5.0$), the 1983 Liège earthquake ($M = 4.6$) and the 1992 Roermond (NL) earthquake ($M = 5.3$). Historical records indicate that larger earthquakes have occurred in the past, such as the 1580 earthquake in the English Channel ($M = 6.0$) and the 1692 Verviers earthquake ($M = 6.3$). This evidence shows that damaging earthquakes, although infrequent, do occur in our region, and hence that seismic hazard is not negligible. European seismic hazard maps depict Belgium as a region with elevated hazard compared to adjacent areas in France, the United Kingdom (UK), Germany and The Netherlands. Seismic hazard assessments (SHAs) provide essential input for engineers to design earthquake-resistant buildings and infrastructure. Examples include country-wide hazard maps in support of the European building code Eurocode 8 and site-specific SHAs for critical infrastructure such as nuclear facilities, SEVESO classified industries, etc. A crucial component of SHA are ground-motion models (GMMs), which relate earthquake ground motion to magnitude, distance, and possibly other source, path and site properties. Strong variations in ground motion are observed between different regions of the world, mainly due to different attenuation characteristics. However, very few GMMs are available for intraplate areas and none specifically for Belgium. Consequently, SHAs in Belgium so far had to rely on GMMs developed for other parts of the world. Although the selection of GMMs has grown more sophisticated over the years, their validity for our regions has not been tested and their weighting remains subjective. Considering that in probabilistic SHA, the ground motion branch of the logic tree is one of the main contributors to uncertainty, there is clearly a need for ground-motion data from lower-seismicity regions.

Earthquake ground-motion databases are fundamental resources for the development of GMMs. In the past two decades, several international efforts have collected strong-motion data, such as the Next Generation Attenuation (NGA) project in the United States, and the follow-up projects NGA-West and NGA-East, the KiK-net database in Japan, the ITACA database in Italy, the RESORCE database for Europe and the Middle East, and the worldwide Engineering Strong Motion database. These databases predominantly contain data from the most active regions of the world, where earthquakes are more abundant. A large number of GMMs have been derived from these databases. In contrast, intraplate areas, where seismicity is much lower, are much less represented. Although the vast majority of recorded data in these regions concern weak-motion records from earthquakes with magnitudes lower than the magnitude range relevant for SHA ($M_W \geq \sim 4.5$), recent successful projects to compile ground-motion databases in France, Switzerland and Australia have demonstrated that it is not only feasible, but also valuable, to conduct similar studies for Belgium.

1.2. Objectives

The above considerations provided the rationale to launch a dedicated project, BELSHAKE, which intends to bridge a gap in know-how and in data collection in the field of earthquake ground-motion with respect to the international seismological community. The currently available seismic data in Belgium are mainly used to detect and locate regional earthquakes, and to determine their main source characteristics (e.g., focal depth, local magnitude, focal mechanism), but a systematic inventory of these data for the purpose of ground-motion modelling has not been undertaken yet.

The main objectives of the project are twofold: (1) to build a database of recorded earthquake ground motion in Belgium and adjacent regions, and (2) to develop the capacity to model ground motion due to earthquakes in Belgium based on this database. This will allow to evaluate the validity of existing GMMs in Belgium and ultimately to develop a model specifically for our country. The used methodology is well-established, but has not, or only in part, been applied in the Belgian context. Thus, both the topic and the methodology are situated within the international state-of-the-art.

1.3. Results

The most important result of the BELSHAKE project is the development of a new earthquake ground-motion database for Belgium and surrounding regions, i.e. the BELSHAKE database. This database is based on digital waveform data starting from 1985 and currently contains $\sim$6400 records from 344 tectonic and induced earthquakes with $M_L \geq 2$. Considerable efforts were undertaken to ensure that both the waveform data, the metadata and the derived intensity measures (IMs) adhere to the best practices set by international ground-motion databases. All waveforms have been visually inspected, and problems were either corrected or flagged. A main challenge we encountered concerns the reconstruction of missing instrument response information for older records. This information is required to convert recorded ground motion to physical units and to correct for instrument characteristics. We applied different methods and evaluated their validity in different ways. Records for which instrument response was found to be unreliable, were flagged in the database, preventing them to be used in subsequent steps. All reliable records were processed using a semi-automated, uniform workflow. The processing steps and the order in which they are applied have been adapted with slight variations from the schemes in established ground-motion databases. Subsequently, up to 109 intensity measures were computed from the processed data, for more than 120,000 record-component-window combinations. We also determined moment magnitudes from the waveform data for 294 events in the database and confirmed their reliability by comparing the results of different methods and comparing with values in other seismic catalogs. Finally, we performed various consistency checks to ensure the quality and reliability of IMs, including evaluation of different component ratios, residual analysis with respect to a generic GMM, comparison of co-located stations, and comparison with common events in the French RESIF/EPOS-France database. Anomalous IMs were analysed, and problematic values were either corrected or removed. The final residuals show no discernible trend with distance or magnitude, and vary in a narrow range that is similar to or even smaller than that in the database for France. The BELSHAKE database has already proven its value, as it allowed a recalibration of the Belgian local magnitude scale in parallel to this project. This in turn provided new insights in geometrical spreading and path duration, which are important results for the project as well.

In the second part of the project, we first analysed anelastic attenuation of seismic waves in Belgium through the high-frequency attenuation factor kappa (κ), which was determined for all records in the BELSHAKE database. Analysis of the κ-distance relationship using different fitting methods allowed decomposing whole-path κ values into site-specific (κ_0) and regional (kappa gradient κ_r) components in four crustal domains. Our results show that variations in κ_r are relatively small, with an overall average corresponding to an apparent (frequency-independent) Q factor around 1750, confirming that regional attenuation in Belgium is relatively low. Comparison of site-specific κ_0 values with V_{S30}, the average shear-wave velocity in the top 30 m, shows relatively constant values over a broad V_{S30} range for stations situated on bedrock. This indicates that bedrock with high V_{S30} in Belgium does not show the low κ_0 values typical of hard rock in e.g. Eastern North America.

Subsequently, we evaluated the goodness-of-fit of more than 20 published GMMs with respect to observed data in the BELSHAKE database, using multiple statistical measures. Although no single model consistently outperforms all others, several models demonstrate good and stable performance across different measures. An interesting and unexpected finding is that low-attenuation versions of NGA-West GMMs for California, a plate boundary, match the data better than NGA-East GMMs for the intraplate region of Eastern North America, both in terms of absolute values and in spectral shape. Building on this result, we applied the relatively simple Referenced Empirical Approach to adjust some of the tested GMMs to the Belgian situation. Our analyses show that this method should be applied carefully. It should be limited to models requiring only minor adjustment and with spectral shapes that are not too different from those observed. It should also be verified that the residuals do not show any trend with magnitude to avoid bias for higher magnitudes relevant in SHA. Nevertheless, we identified three models that could be reliably adjusted. The most promising of those is a stochastic model developed for the UK, which shows the best fit to the BELSHAKE database. In addition, thanks to data exchange other researchers have determined factors to calibrate a pan-European GMM using a more sophisticated method, which was beyond the possibilities in the BELSHAKE project. Finally, we also implemented two methods, EGF and EXSIM, to simulate ground motion from higher-magnitude earthquakes. As a proof of concept, we simulated realistic ground motions for the largest earthquake in the database, the 1992 Roermond earthquake. This exercise benefited from other results obtained in the project, such as the path and site components of κ, V_{S30} values for Belgian seismic stations, along with the newly derived geometrical spreading and path duration models. Thus, we have paved the way for stochastic simulation of accelerograms for hypothetical high-magnitude earthquakes that are in agreement with the average source, path and site characteristics in and around Belgium.

1.4. Conclusions

Available seismic data in Belgium have so far been underexploited in terms of engineering seismological applications. In the BELSHAKE project, we established the first ground-motion database serving this purpose. Thanks to its quality-checked records, integrated instrument response information and comprehensive metadata, the BELSHAKE database is a high-quality resource for earthquake ground motion in Belgium, as attested by the small uncertainties on epicentral locations and magnitudes, and the low residuals. Although the magnitude coverage is relatively limited, it contains important information on source, path and site characteristics of ground motion in our regions. The database complements existing databases by increasing the coverage of low-seismicity zones and lower magnitudes. In addition, it will serve as a reference dataset for many other seismological applications. An example is the recalibration of the Belgian local magnitude scale performed alongside this project.

Using the BELSHAKE database, we have characterised both geometrical spreading and anelastic attenuation in Belgium in more detail than before. It also allowed evaluating which published GMMs show the best agreement with observations, and we managed to calibrate a few of these models. Thus, and for the first time, we developed GMMs specifically for Belgium based on local data. This will allow better informed and data-driven selection of GMMs in SHA. In parallel, we developed the capacity to simulate ground motion from higher-magnitude earthquakes. Together, these advances lay a solid foundation for further developments in ground-motion modelling in Belgium. Our results will be very useful for new SHAs in Belgium, such as will be needed for the revision of Eurocode 8, as well as for the prolongation of existing or the siting of new nuclear facilities.

1.5. Keywords

Earthquakes, ground-motion model, database, seismic hazard, seismic attenuation

LIST OF ABBREVIATIONS

DSHA:	deterministic seismic hazard assessment
GMM:	ground-motion model
IM:	intensity measure
M:	magnitude
M_L:	local magnitude
M_W:	moment magnitude
NGA:	Next Generation Attenuation (name of American project)
PGA:	Peak Ground Acceleration
PGV:	Peak Ground Velocity
PSHA:	probabilistic seismic hazard assessment
ROB:	Royal Observatory of Belgium
SA:	Spectral Acceleration
SHA:	seismic hazard assessment

1. INTRODUCTION

Earthquake ground motion is the shaking of the ground (usually at the Earth's surface) by seismic waves generated by earthquakes. While most people consider seismic shaking to be identical with an earthquake, seismologists regard it as an effect caused by an earthquake, which is the source in the Earth's crust where the seismic waves are radiated from. Although earthquakes can also cause other damaging effects (e.g., surface rupturing, landslides, liquefaction, tsunamis), the most significant effect worldwide is ground shaking, causing collapse of buildings, which is in turn responsible for the majority of fatalities due to earthquakes (e.g., Cobum et al., 1992; Kenny, 2009). For this reason, seismic hazard is often considered to be synonymous with earthquake shaking hazard, and ground-motion models are an important ingredient in seismic hazard assessments: in probabilistic assessments (PSHA), the occurrence probability of earthquakes of different sizes (magnitudes) is combined with the probability of exceeding a given level of ground motion at a site for the given magnitudes and distances, while in deterministic assessments (DSHA), shaking at a site is computed for a particular earthquake scenario by direct application of a ground-motion model. The study of earthquake ground motion is thus at the interface of seismology and engineering, the former concerning the location, frequency and source parameters of earthquakes, and the latter concerning the design of buildings that are able to resist shaking due to earthquakes. A practical result of this interaction is the drafting of building codes, which define the level of shaking that a building must withstand. In Europe, the building code EN-1998 ("Eurocode 8", CEN (Comité Européen de Normalisation), 2004), which is applicable in all member states since 2011, is currently under revision.

The most important properties of ground motion for engineering seismology are amplitude, duration and frequency content. The factors influencing these properties can be separated into differences in (Boore, 2003):

- the earthquake <u>source</u>: this includes obvious parameters like earthquake magnitude and focal depth, but also focal mechanism, stress drop, rupture speed and rupture direction;
- the travel <u>path</u> between source and site: as seismic waves propagate through the crust, they lose energy with increasing distance due to a combination of:
    - geometrical spreading: distribution of energy over an expanding spherical wavefront;
    - elastic attenuation: frequency-dependent attenuation caused by scattering due to crustal heterogeneities;
    - anelastic attenuation: frequency-dependent attenuation caused by internal friction or changes in viscosity, resulting in the gradual absorption of elastic energy by the medium and conversion to heat.
- local <u>site</u> conditions: this includes amplification effects due to site topography and site geology (sediment cover). As seismic waves travel upward from bedrock to surface through layers of unconsolidated sediments, their frequency content and amplitude are changed as a function of sediment thickness and softness, respectively. This effect can be quite dramatic: for the same earthquake, shaking at one site can easily be ten times stronger than at another site, even when their distance from the earthquake rupture is comparable!

2. STATE OF THE ART AND OBJECTIVES

2.1. State of current knowledge at international level

Ground motion can be modeled in different ways as a function of source, path and site parameters (see overview in Douglas & Aochi, 2008). The easiest and most widely used way is using ground-motion models (GMMs), which express the level of a scalar intensity measure (commonly peak ground acceleration and response spectral acceleration) as a function of magnitude, distance and optionally other parameters. Many relations have been published in literature. A compendium compiled by Douglas {Formatting Citation} lists over 850 GMMs, and the list is still rapidly growing. Many are empirical relations, obtained by regression of recorded strong-motion data, but some are based on stochastic simulation or on a combination of both. Most GMMs are representative of more active regions near plate boundaries (Figure 1), where ground-motion data cover a wide range of magnitudes and distances. Very few GMMs are valid for intraplate areas like Belgium and more generally NW Europe, where ground motions are comparatively stronger, but seismic activity is much lower and data hence sparser. Alternatively, ground motion can be modeled more elaborately, based on explicit physical models of the earthquake process and crustal attenuation. Examples are spectral modeling (Boore, 2003), in which the source, path and site components of the Fourier amplitude spectrum (FAS) are modeled, and kinematic rupture modeling (e.g., Graves & Pitarka, 2016), in which ground motion is simulated based on rupture propagation on a rough fault plane. In contrast to GMMs, physics-based methods explicitly model regional dependence through the choice of input parameters.

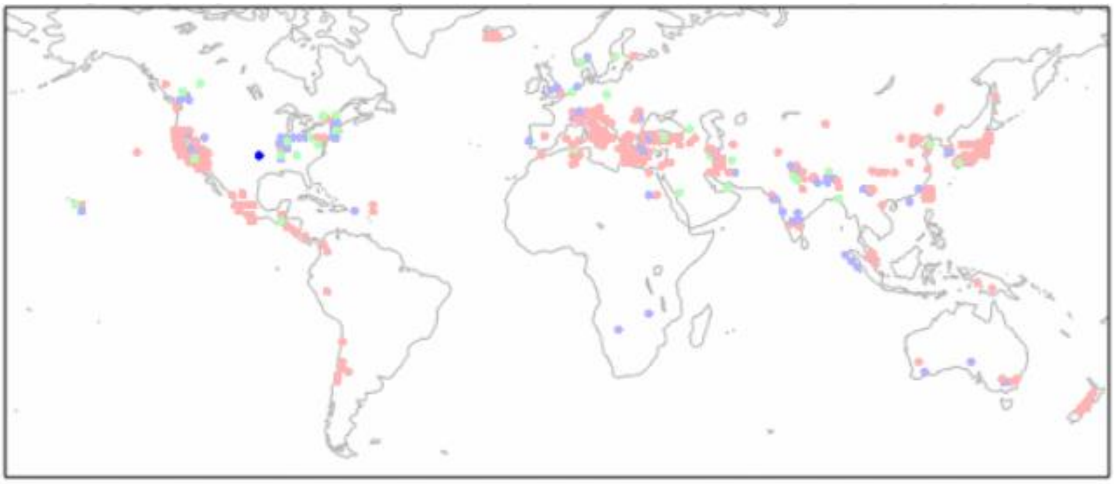


Figure 1 – Map showing regions in the world for which GMMs have been developed (Douglas, 2019). Red dots: empirical GMMs ; blue: simulation ; green: other

In the past 10 – 15 years, there have been several international efforts to collect and disseminate strong-motion data in support of earthquake ground-motion modelers, such as the Next Generation Attenuation project (NGA, Chiou et al., 2008) in the United States, and the follow-up projects NGA-West (Ancheta et al., 2014) and NGA-East (Goulet et al., 2021) in Western and Eastern North America, the KiK-net database in Japan (Dawood et al., 2016), the ITACA database in Italy (Pacor et al., 2011), the RESORCE database for Europe and the Middle East (Akkar, M. A. Sandıkkaya, et al., 2014), and the worldwide Engineering Strong Motion Database ESM (Luzi et al., 2016). More recently, a similar database has been established for France (Buscetti et al., 2025; Traversa et al., 2020), showing that it is both feasible and useful to do this exercise for lower-seismicity regions as well.

2.2. State of current knowledge at national level

Due to the lack of strong-motion data, seismic hazard assessments in Belgium so far had to rely on GMMs developed for other parts of the world. Although the selection of GMMs has grown more sophisticated over the years, their validity for our regions has not been tested. The current National Annex to Eurocode 8 (Bureau voor Normalisatie (NBN), 2011) is based on a PSHA study (Leynaud et al., 2000), in which only one GMM was considered. Growing experience acquired during several projects for the nuclear sector has resulted in a full-blown ground-motion logic tree branch, in which a distinction is made between more active areas like the Lower Rhine Graben and more stable areas (remainder of the territory), and for the latter also between standard rock with $V_{S,30}$ (average shear-wave velocity in the top 30 m) between 700 and 1200 m/s, and hard rock with $V_{S,30}$ around 1800 m/s. The most recent version of this logic tree branch (Vanneste et al., 2014) (Figure 2) consists of 11 GMMs from different regions (Europe and the Middle East, Italy, California, Japan, eastern North America), some of which were adjusted to account for differences in bedrock characteristics between the host and target regions. These GMMs were weighted according to their supposed applicability to our regions based on various criteria (e.g., tectonic region type, applied methodology, stress drop value, magnitude and distance range, standard deviation), but this remains theoretical and subjective. Furthermore, sensitivity tests show that the highest dispersion in PSHA generally comes from the ground-motion branch of the logic tree, highlighting the need for a dedicated study on this topic.

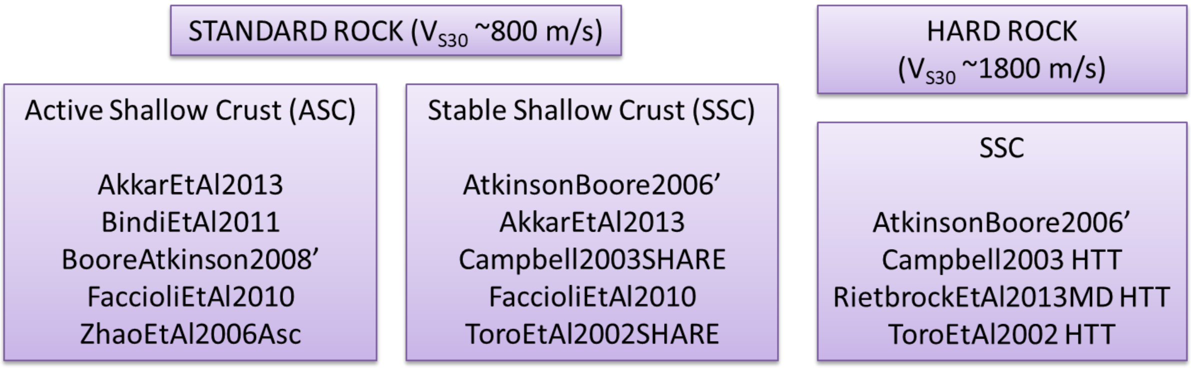


Figure 2 - GMM selection applied in recent seismic hazard studies in Belgium

2.3. Position of the project within the state of the art and the strategic scientific objectives of the FSI

BELSHAKE aims to bridge the gap in know-how and in data collection with respect to the international state of the art. As one of the core missions of the Royal Observatory of Belgium (ROB) is to monitor seismic activity in and around the Belgian territory, ROB has been routinely and continuously recording ground motions over the past decades using a network of digital seismometers and accelerometers. These data are mainly used to detect and locate regional earthquakes, and to determine their source properties (magnitude, focal mechanism, …). A systematic inventory of these data for the purpose of ground-motion modeling has not been undertaken so far. However, although the vast majority of recorded data concern weak-motion recordings of ground velocity from earthquakes with magnitudes lower than the range of interest in SHA (M≥4.5), recent successful projects in other low-seismicity regions to model earthquake ground motion in Switzerland (Edwards & Fäh, 2013) and the United Kingdom (Rietbrock et al., 2013; Rietbrock & Edwards, 2019) or to compile ground-motion databases (Ghasemi et al., 2026 in Australia; e.g., Traversa et al., 2020 in France) give us confidence that it is not only feasible, but also valuable and even essential, to conduct similar studies for Belgium. It is clear that the available ground-motion data in Belgium have so far been underexploited. Any research leading to a better characterization of earthquake ground motion in Belgium and valorizing these data is a strategic scientific objective for the ROB. Furthermore, this research also serves national and international earthquake engineering applications. It also stimulates collaboration with international research teams, for instance in neighboring regions of Germany, France and The Netherlands, which share many ground-motion characteristics with Belgium.

2.4. Research objectives

The main objective of BELSHAKE is to build a solid base for earthquake ground-motion modeling in Belgium. This is achieved by (1) building a database of earthquake ground motions for Belgium and adjacent areas, and (2) developing the capacity to model ground motion due to earthquakes in Belgium. The first part of the project is devoted to a systematic inventory of short-period, broadband and accelerometer records of local and regional earthquakes by stations of the Belgian seismic network, implementation of a semi-automatic workflow to uniformly process selected seismograms and determine relevant ground-motion parameters according to internationally established standards, and collection of these parameters in a structured database. This database is an indispensable tool for the second part of the project, in which the collected data are used to characterize seismic attenuation in the main crustal domains in Belgium, to evaluate the validity of existing ground-motion models and finally to guide modeling of the source, path and site components of ground-motion spectra. Ultimately, this should lead to the development of a stochastic ground-motion model that is valid for Belgium, or else to the calibration of regionally adaptable (so-called “backbone”) ground-motion models, which will complement an informed selection of existing models that can be applied in our region.

3. METHODOLOGY

The overall methodological approach consists of two main steps. In the first step, we compile a database of earthquake ground motions in and around Belgium based on digital broadband and accelerometer recordings of the Belgian seismic network. This is achieved by a systematic inventory of the available data, selection according to well-defined criteria, application of a uniform processing flow to the selected waveforms adhering to the current best practices, quality control to exclude poor-quality records, computation of a broad range of useful intensity measures, and association with reliable and comprehensive metadata. It is the explicit purpose to produce a database that is compatible with international databases and that meets the same high standards and quality criteria.

In the second step, we develop the capacity to model earthquake ground motions in Belgium using the database compiled in the previous step. This step involves different methodologies: first we characterize crustal attenuation by analysing the high-frequency spectral decay parameter κ (kappa) from acceleration Fourier amplitude spectra for the main crustal domains in Belgium. Secondly, we compare ground-motion values computed from the database with values predicted by a selection of published GMMs that are assumed to be valid for stable regions like Belgium. This comparison is based on different statistical measures. Thirdly, we investigate how existing GMMs can be adjusted or calibrated to the crustal attenuation properties in Belgium. Finally, we implement methods to stochastically simulate ground motion from higher-magnitude earthquakes.

3.1. Compilation of a ground-motion database

The data required for this project are digital waveforms of regional earthquakes recorded by short-period and broadband seismometers and accelerometers of the Belgian seismic network, and the associated event and station information. These are existing data that have been collected over the past decades by ROB in the scope of its strategic mission to monitor seismic activity on the Belgian territory. The Belgian seismic network (http://seismologie.oma.be/en/the-service/seismic-network) has evolved strongly over time. Digital recording started in 1985, but only for a few stations and data were only stored for detected events. In subsequent years, the network was gradually expanded and more instruments were equipped with a digital recording system. From 2005 onward, increasing storage capacity allowed storing continuous data streams, first for select stations and later for all of them. Starting in 1999, the seismic network was complemented with accelerometers, which were mainly installed in the most active areas to record strong ground motions. Initially, these stations only recorded when they were triggered, but they were replaced with instruments capable of continuous recording from 2021 onward. The Belgian seismic network currently consists of ~30 seismometers and ~20 accelerometers. As a result of this network development, the ROB waveform archive is heterogeneous. All waveform data are stored on a server in different formats and folder structures, and are accessible through different network services or protocols, depending on sensor type (seismometer or accelerometer) and the year and mode of recording (event records trimmed from originally continuous recordings, triggered records or continuous streams).

3.1.1. LITERATURE STUDY

Before embarking on this exercise, we first studied existing ground-motion databases such as the Next Generation Attenuation project (NGA, Chiou et al., 2008) in the United States, and the follow-up projects NGA-West (Ancheta et al., 2014) and NGA-East (Goulet et al., 2021), the KiK-net database in Japan (Dawood et al., 2016), the ITACA database in Italy (Pacor et al., 2011), the

RESORCE database for Europe and the Middle East (Akkar, M. A. Sandıkkaya, et al., 2014), and the worldwide Engineering Strong Motion Database ESM (Luzi et al., 2016). This ensures that we determine relevant ground-motion parameters and required metadata according to internationally established standards, and that our database is compliant with these databases, permitting data exchange. In particular the recent RESIF/EPOS-France database for metropolitan France (Buscetti et al., 2025; Traversa et al., 2020) is a useful example, as it concerns a low-seismicity region very similar to Belgium and also mostly contains weak-motion ground velocity data.

3.1.2. DATA SELECTION

The next step consisted of the systematic inventory of all available digital waveform data in the ROB archive. We selected all tectonic and induced earthquakes since 1985 from the ROB earthquake catalog which satisfy two selection criteria: (1) $M_L \geq 2.5$ in a wide rectangular region around Belgium (-1° - 9° E and 49° - 52° N), and (2) $M_L \geq 2.0$ in a more narrow region around Belgium (2.5 - 7° E and 49.25 - 51.5° N). A small number of lower-magnitude events ($M_L \geq 1.0$) were also included because these were recorded at short distances (< 10 km). The corresponding waveform data were mostly retrieved from the permanent Belgian seismic network, but for some key events we also obtained data from neighbouring networks in Germany (code BQ; Department of Geosciences, Bensberg Observatory, 2016) and The Netherlands (code NL; KNMI, 1993). In addition, we also included data from temporary networks deployed by ROB during some seismic sequences, which provide additional short-distance records. For each selected event-station combination, we calculated a suitable time window to capture the complete event, including pre-event noise and coda wave, and we imported the corresponding raw waveform data into the waveform library. The associated event and station metadata were extracted from the ROB seismic routine database, the corresponding source-to-site parameters were computed, and all these metadata were imported in the BELSHAKE database. Missing P- and S-wave arrival times were computed from the Belgian velocity model (Camelbeeck, 1993), and if necessary manually adjusted during visual inspection.

We also attempted relocating the selected earthquakes, using hand-picked phase arrival times (from multiple agencies) only, by taking the average of 100 runs with the randomly perturbed catalog location as initial location, weighted inversely by the standard deviation of their arrival-time residuals (RMSE). New solutions for the epicentral locations were only accepted if they yielded significantly lower RMSE (at least 0.05 s) and lower location uncertainties. Thus, we were able to improve the location of $\sim$100 events. Overall, location quality in the BELSHAKE database is very good, with only four events with a total RMSE on the arrival times higher than 2 s. Horizontal and vertical location uncertainties are quite low, with median values of 0.9 and 1.2 km, respectively, and 95th percentile values of 6.7 and 5.7 km, respectively. For six events (four of which came from catalogs of neighbouring networks), location quality could not be assessed due to insufficient phase picks. Finally, we also recomputed local magnitudes according to the Belgian M_L scale (Camelbeeck, 1985), using the available vertical S-wave displacement amplitudes in the ROB seismic routine database. This resulted in slightly lower magnitudes ($\Delta M_L = -0.2$) for only two events. For $\sim$25 events with insufficient amplitudes, we copied the M_L values from the catalog of the most authoritative neighbouring network (depending on location) or used the average of two or more catalogs. This resulted in magnitude differences of $\Delta M_L < -0.3$ for 11 events and $\Delta M_L > 0.3$ for 1 event. The source for each magnitude solution was included in the database.

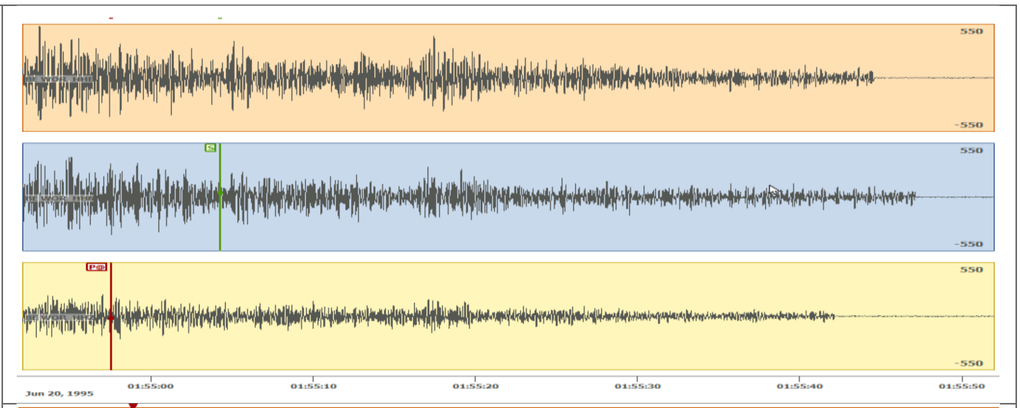
3.1.3. QUALITY CONTROL

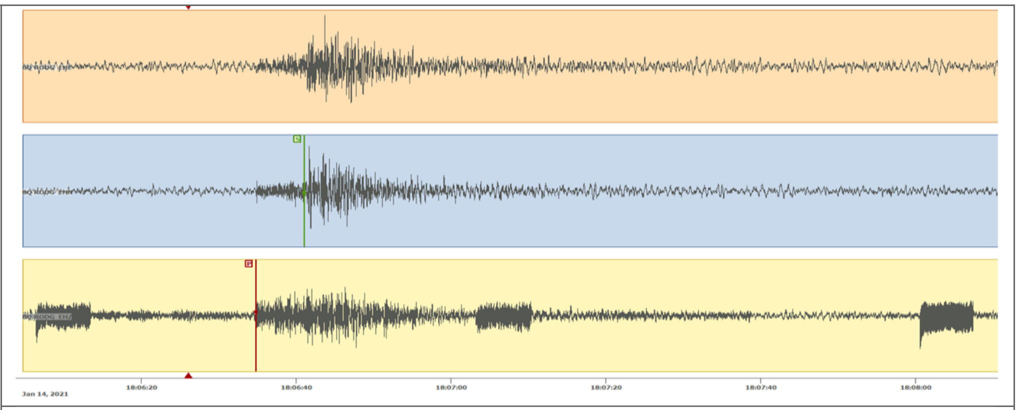
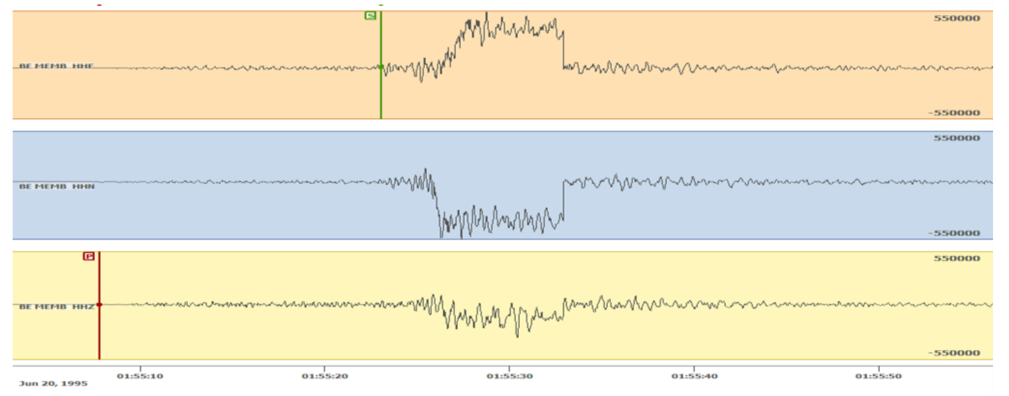
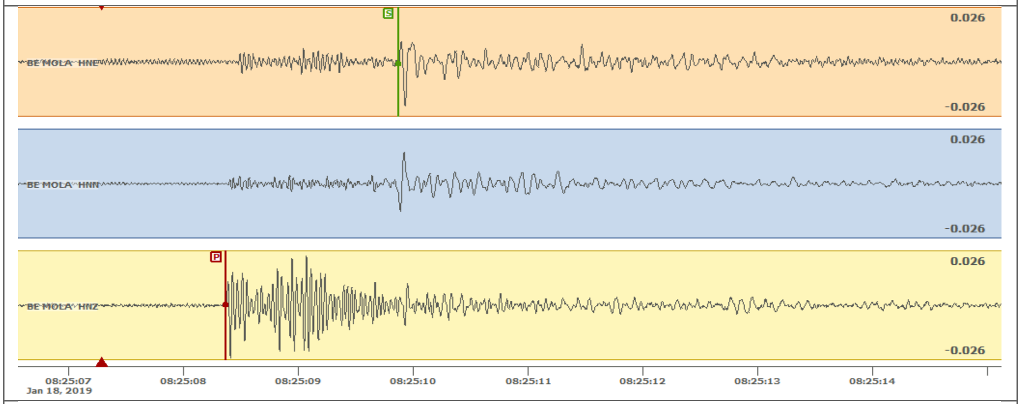
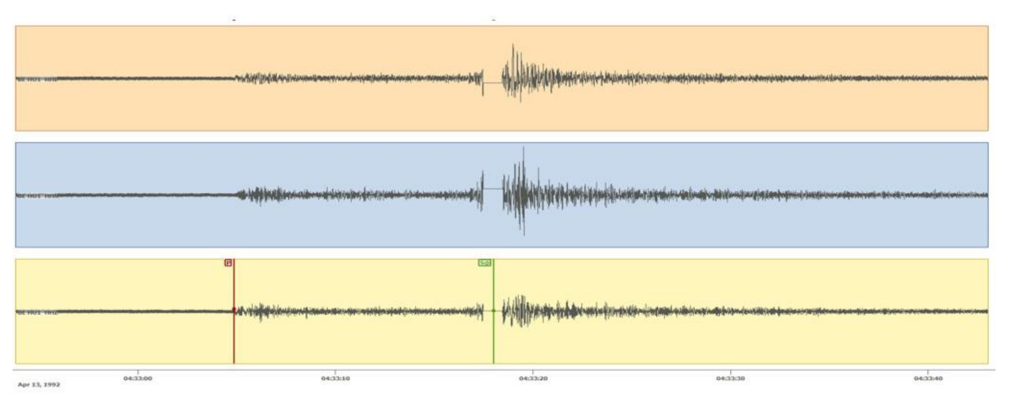
After data selection and before data processing, the waveforms imported in the BELSHAKE database were visually inspected for any of the following problems: timing problems, low signal-to-noise ratio (SNR), presence of spikes, disturbances, data gaps, baseline distortion, resonance, clipping, missing or incomplete phase windows, or presence of fore- or aftershocks in the noise and/or signal windows. Examples of each category are shown in Figure 3. In addition, we also evaluated the presence of 50 Hz hum from Fourier amplitude spectra.

Some of these problems can be, and were repaired when possible:

- Most timing problems could be fixed by determining the time shift between the P-wave onset and the expected P-wave arrival, reimporting the data with the time shift applied, and storing the time shift in the database;
- Spikes could be removed by replacing them with interpolated values;
- Some distortions could be removed by interactive baseline removal or highpass filtering;
- Disturbances which are not affecting the main signal phases (P and S) could be removed by trimming;
- 50 Hz hum could be removed by applying a narrow notch filter.

All remaining problems were considered to be irrepairable. In all cases (except for the timing problems), the original waveform data were kept and the corrected data were stored with a different tag in the ASDF files. Irrepairable records were left as such. The identified problems, and whether or not these have been corrected, are stored in the database, such that they can be queried for any component and signal window.

Problem category	**Waveform example**
Timing problem	
Spikes	

Problem category	**Waveform example**
Disturbance	
Baseline distortion	
Resonance	
Data gap	

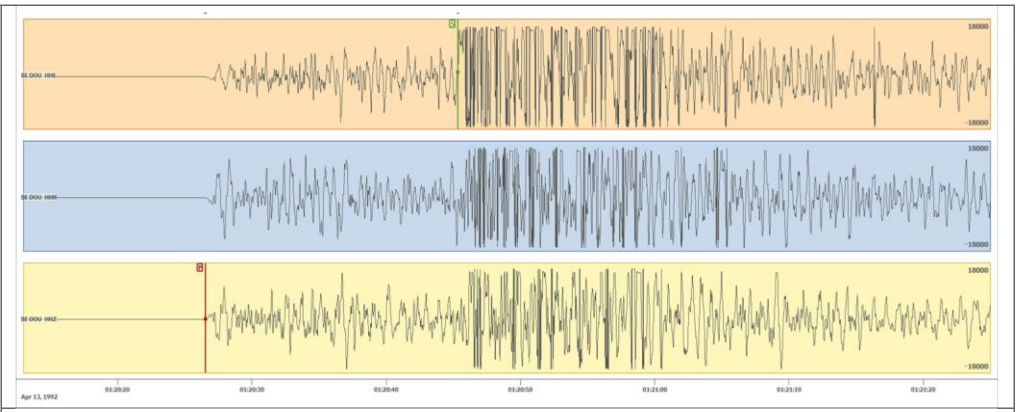
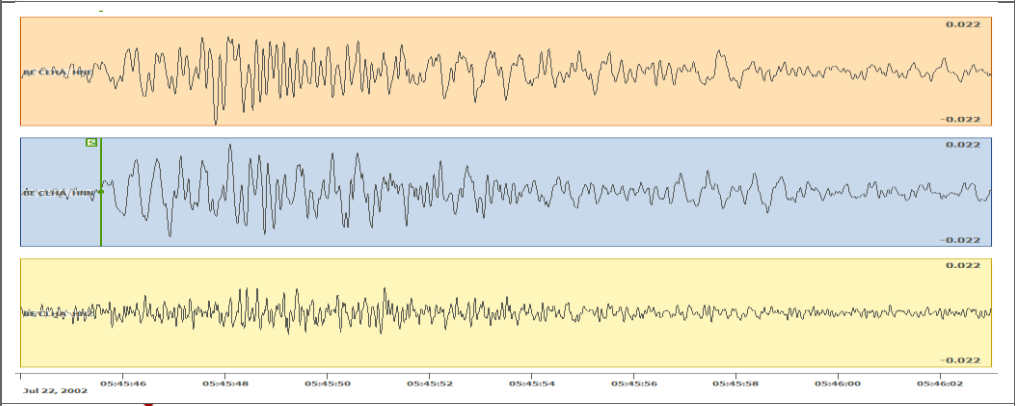
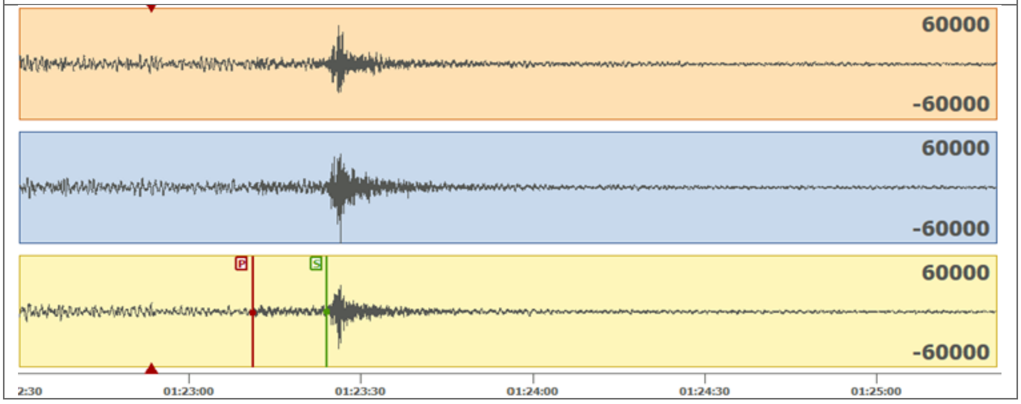
Problem category	**Waveform example**
Clipping	
Incomplete (P-window missing)	
Presence of foreshock	

Figure 3 - Illustration of waveform problem categories. For each category, a 3-component seismogram is shown using the Snuffler program of Pyrocko (Heimann et al., 2017). Red vertical lines correspond to P-wave arrival times, green lines to S-wave arrival times

3.1.4. EVALUATION OF INSTRUMENT RESPONSE

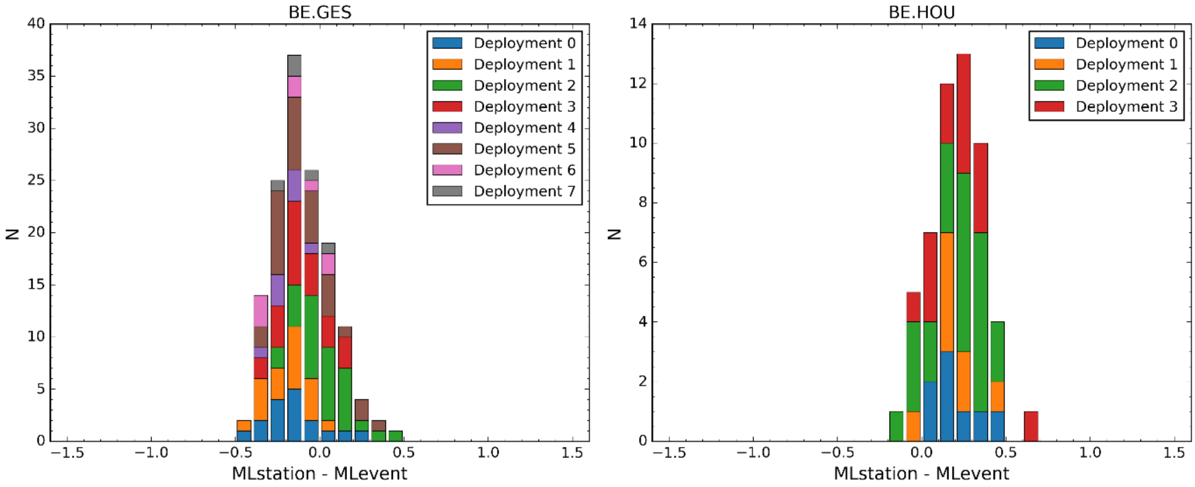
A major challenge encountered during the compilation of the BELSHAKE database was the lack of instrument response information for older waveform data, which is essential to reconstruct actual ground motion in physical units from seismic records. Since 2005 or later, depending on the seismic station, this information has been systematically collected and maintained in dataless SEED files. However, for data recorded before that time, instrument response information was not readily available. Fortunately, some important parameters are available in the ROB seismic routine database and/or in the headers of the original waveform files. It is possible to combine this information with the response information from similar instruments and/or acquisition systems from other stations or from the Nominal Response Library (NRL, https://ds.iris.edu/ds/nrl/) in order to construct missing instrument response information.

We evaluated the validity of all instrument responses based on different criteria:

- Evolution of energy levels of the noise window in different frequency bands through time;
- Consistency of power spectral densities (PSDs) of the noise window within each station deployment (= particular combination and/or configuration of instrument and acquisition system) and between different deployments of the same station;
- Consistency of station-M_L residuals (= M_L values calculated for single stations minus the average M_L values for the corresponding events) between different deployments of the same station.

If consistent anomalies were found, i.e. anomalously low/high station-M_L residuals along with anomalously low/high noise levels, we concluded that there was a problem with the instrument response for that particular station deployment. In such a case, we also evaluated the station-M_L residuals and PSDs that would be obtained with the instrument responses constructed from the other information sources mentioned above. If one of these yielded station-M_L residuals and PSDs that are in better agreement with the other deployments of the station, we replaced the instrument response in the BELSHAKE waveform library with this alternative response information. The source of the response information of each record is included in the database. In a few cases, problems with total gain were found with instrument responses in the authoritative archive maintained by ROB. We were able to solve most of these issues by examining documentation and hardware in collaboration with our data engineer, Giovanni Rapagnani. The corresponding dataless files were corrected, and if necessary the ROB seismic routine database as well, and the affected records in the BELSHAKE waveform library were updated. Finally, a number of cases remained where it was not possible to fix the instrument response. These records were flagged as unreliable in the database, preventing their use in subsequent steps of the project, but keeping them available for possible correction in the future.

It is not possible to document the instrument response evaluation for each station in this report, but we illustrate it below with a few selected examples. Figure 4 shows histograms of station-M_L residuals for 4 stations of the Belgian seismic network that show no obvious problems: (1) distributions are relatively narrow (± 0.5 magnitude units); (2) there are no obvious changes between different deployments; (3) the mean station-M_L residuals are relatively small (< ±0.5). The instrument responses can be considered to be reliable for the whole period of operation of these stations.



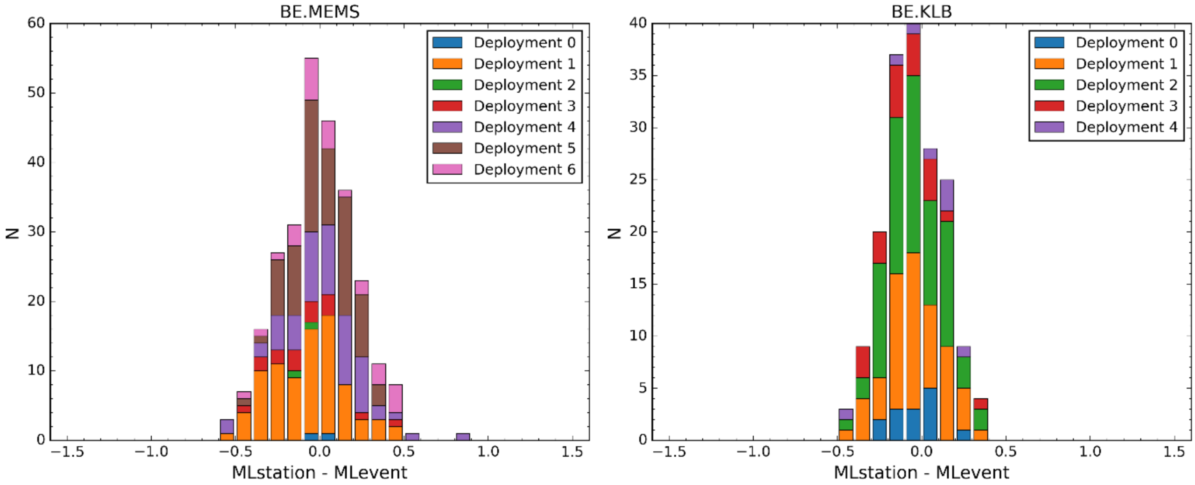
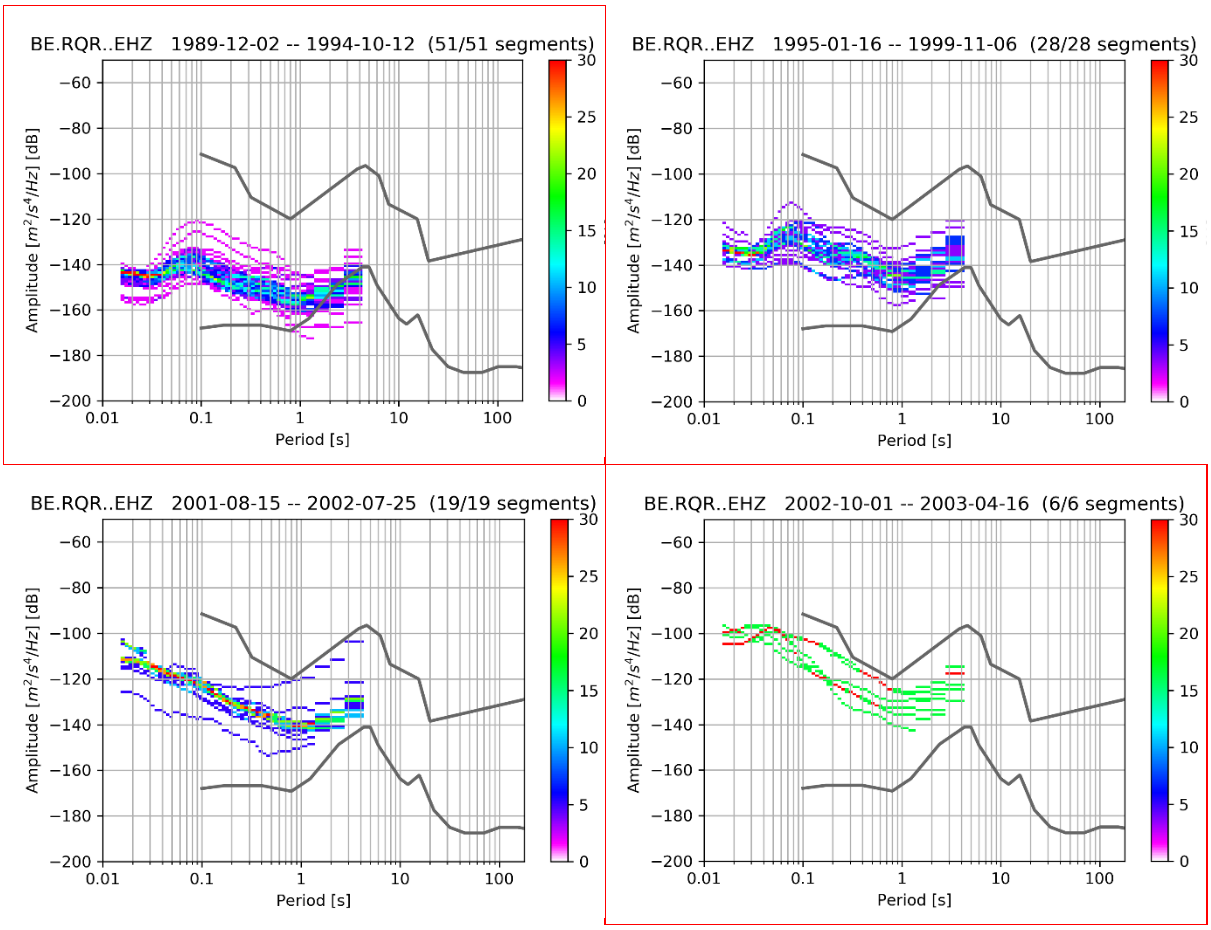


Figure 4 – Histograms of station-M_L residuals for 4 stations of the Belgian seismic network with no obvious problems. Colors correspond to different deployments of each station.

Figure 5 shows PSD plots of the noise windows for 6 successive deployments of station BE.RQR. In these plots, individual records can be recognized as lines where they are not lumped together. It can be seen that the noise levels during deployment 0 (top left) are significantly lower, and those during deployment 3 (middle right) significantly higher than in the other deployments, which are otherwise consistent. We therefore consider instrument response for these 2 deployments to be unreliable.



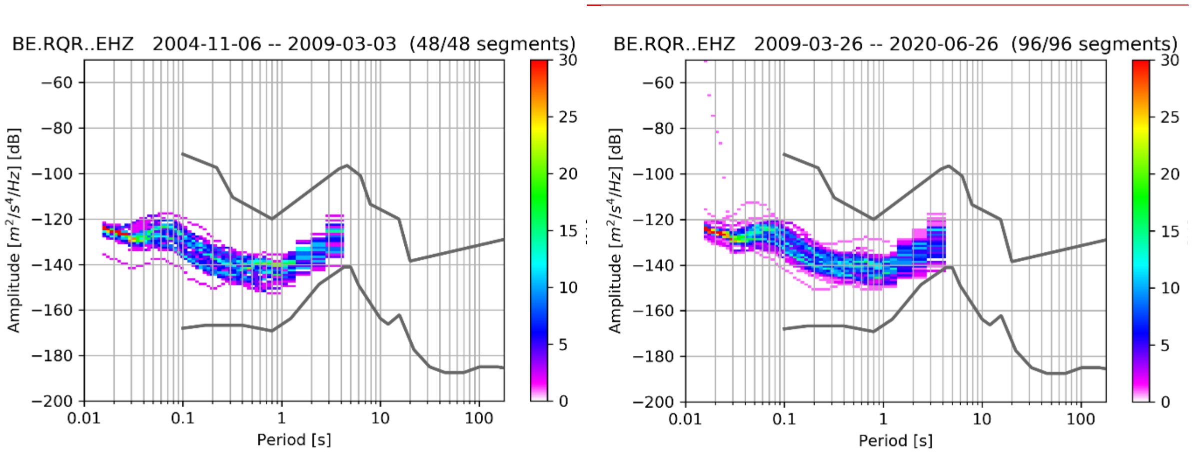


Figure 5 – Power spectral density (PSD) plots of the noise windows for 6 successive deployments of station BE.RQR. Deployments 0 and 3 (indicated with red boxes) are clearly lower and higher, respectively, than the other deployments

Figure 6 shows histograms of station-M_L residuals for one deployment of station BE.SKQ. The colors indicate results obtained with instrument responses that were constructed following different approaches. It is clear in this case that the alternative methods are in agreement with each other and give a result that is more acceptable (mean close to 0) than that obtained with the authoritative dataless information (mean close to 1.0). This points to an error in the dataless file, which we subsequently fixed.

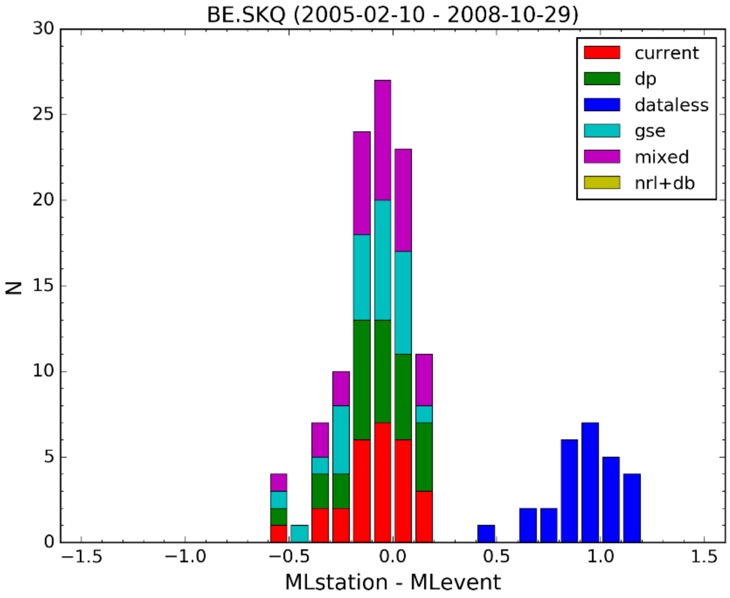


Figure 6 – Histograms of station M_L residuals obtained with instrument responses constructed using different approaches for a deployment of station BE.SKQ

Figure 7 shows the identification of a problem with the instrument response for one deployment of station BE.STI. The histogram of station-M_L residuals shows a bimodal distribution in which the residuals for deployment 6 are clearly lower (at least 0.5 magnitude units) than the distributions for the other deployments. PSDs also indicate that the noise level during this deployment is significantly lower than in the other ones. Both findings are consistent, and indicate that the instrument response for deployment 6, which represents almost 18 years of operation, was incorrect by a factor of ~2.5.

This problem appeared to be due to a wrong value of the shunt resistance and has now been corrected.

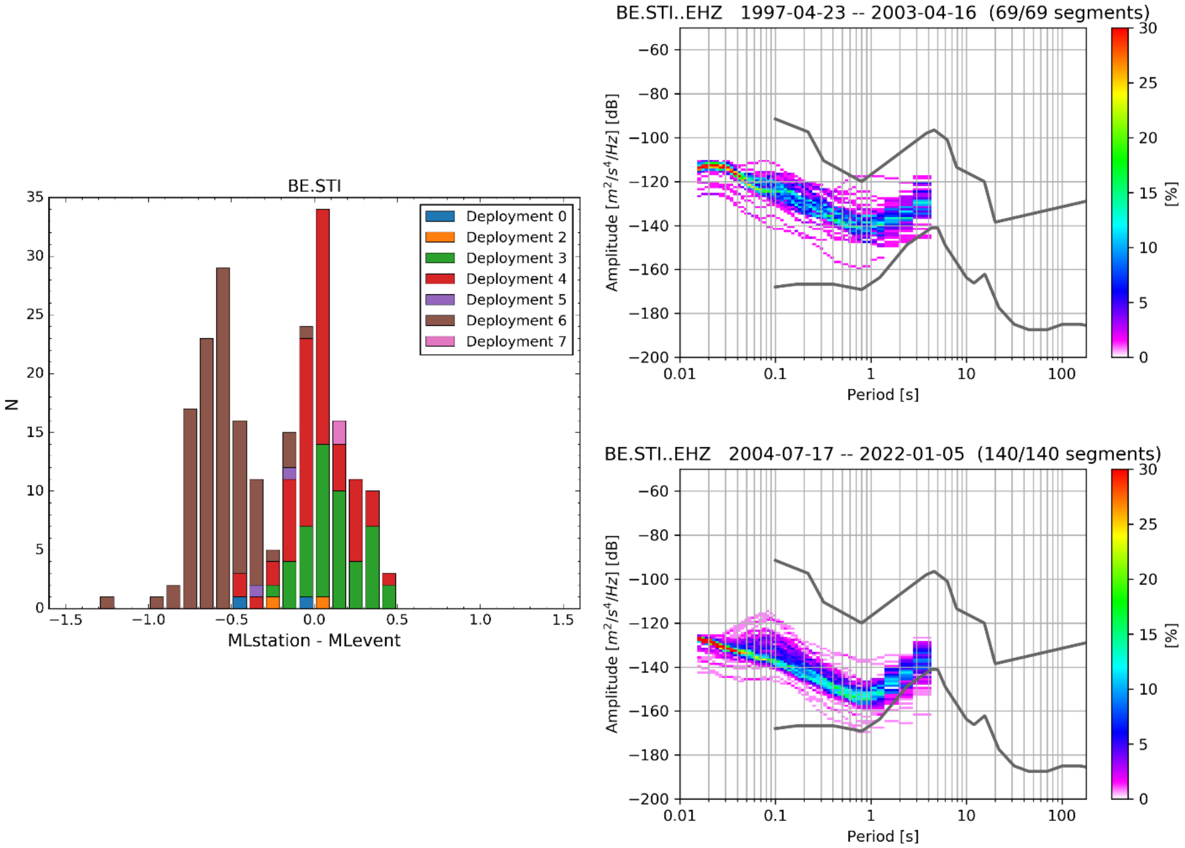


Figure 7 – Identification of instrument response problems for one deployment of station BE.STI. Left: histograms of station M_L residuals for individual deployments of this station. Right: PSD plots for deployments 5 (top) and 6 (bottom).

Finally, Figure 8 illustrates the identification of a problem with the instrument response of deployment 0 of station BE.SNF. The histogram of station-M_L residuals (bottom left) shows a distribution that is much broader (and possibly bimodal) than usual. Similar results are obtained with instrument responses constructed using other approaches, so the problem cannot be easily fixed. Considering the evolution through time of noise energies in different frequency bands (top), it can be seen that the noise energies during deployment 0 are in line with the other deployments since 1994, but show a wide spread before that time. Flagging responses for the period before 1994 as unreliable, we obtain a much narrower distribution with mean close to 0 (bottom right), which is clearly an improvement.

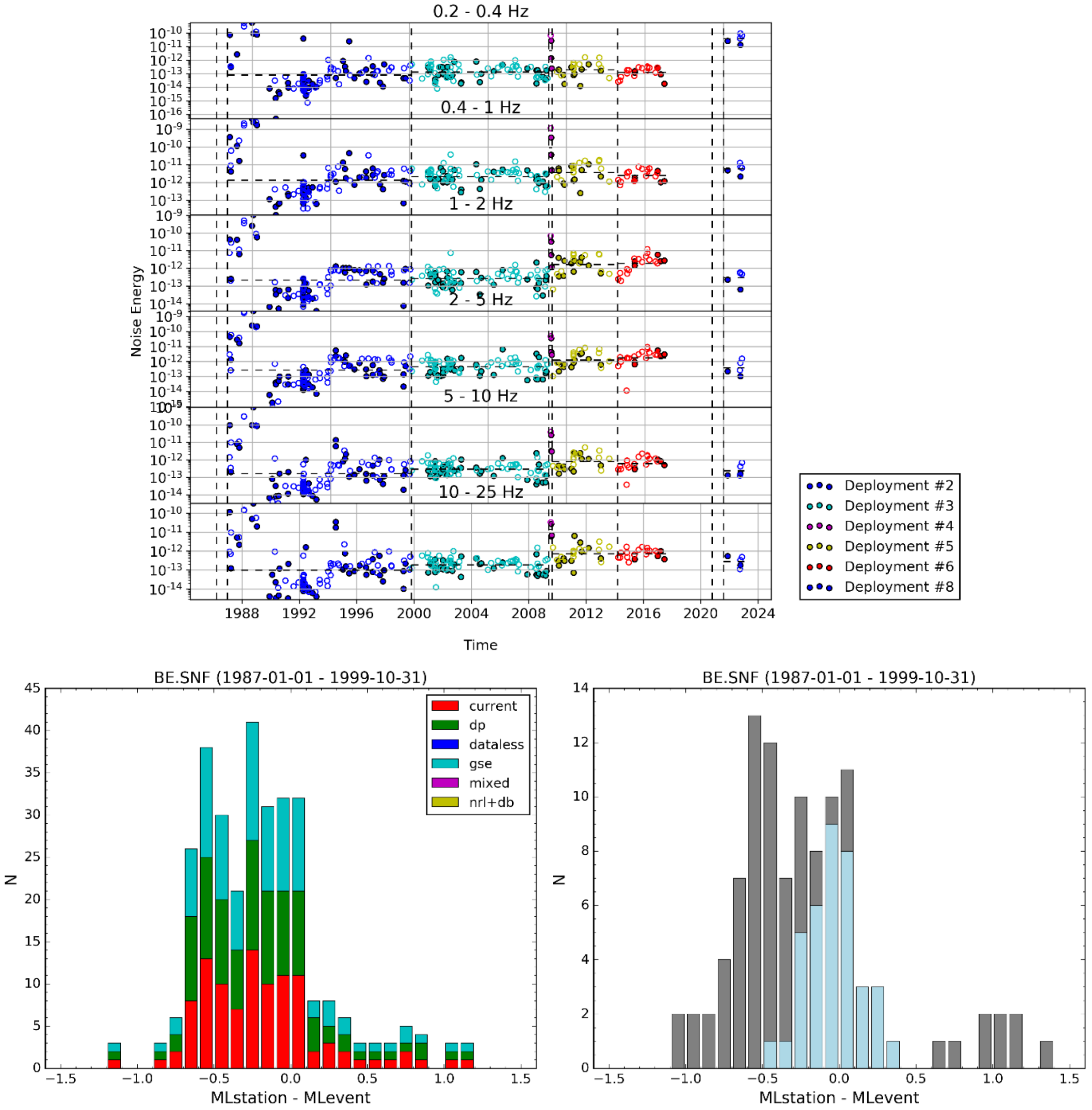


Figure 8 – Identification of instrument response problems for one deployment of station BE.SNF. Top: evolution through time of noise energies in different frequency bands; Bottom left: histograms of station-M_L residuals obtained with instrument responses constructed using different approaches for deployment 0; Bottom right: histogram of station-M_L residuals for deployment 0 after elimination of the unreliable period before 1994

3.1.5. SIGNAL WINDOWING

Following visual quality control, all records in the BELSHAKE database with reliable instrument response were processed using a semi-automated workflow. An important part of this workflow is windowing, which is the division of an earthquake record into portions corresponding to the P-wave, S-wave, coda wave and noise. Although final windowing occurs at the end of the processing flow (to compute window-specific intensity measures), it is also needed earlier in the flow to determine the valid frequency range from the signal-to-noise ratio (SNR), which is the ratio between the Fourier Amplitude Spectrum (FAS) of the signal window (P + S) and that of the noise window.

We investigated two windowing algorithms, one by Goulet et al. (2014), which was applied in the PEER NGA-East database (Goulet et al., 2014, 2021), and another by Perron et al. (2018), which was applied in the RESIF/EPOS-France database (Traversa et al., 2020). We will refer to these as the Goulet and Perron methods, respectively. Comparison of both methods applied to records in the BELSHAKE database shows that:

- The P-window is very similar in the two windowing methods;
- The S-window length in the Goulet method is based on the Lg-wave duration, while in the Perron method it is based on the S-P interval; this generally results in longer S-windows with the Goulet method;
- The Perron method sometimes results in separation between S-window and coda window, while in the Goulet method these windows are always contiguous;
- The Goulet method uses fixed tapering margins, whereas in the Perron method the tapering margins are variable because they are defined as a percentage of the window length.

Based on this comparison, we selected the Goulet method for the signal windows. However, the noise-window selection is more sophisticated in the Perron method, and more adapted to records with only short pre-event windows, as is the case in the KiK-net dataset used by Perron et al. (2018). This is also true for the BELSHAKE database, where many records have incomplete or missing pre-event noise windows due to trimming during data archival of older records, late triggering, disturbances and occasionally the presence of foreshocks. To take advantage of their respective strengths, we decided to combine both methods, and apply the Goulet method for signal (P, S) and coda windows, and the Perron method for the noise window.

The Perron method considers two duration thresholds for noise window selection: the target duration (D_t), which is the ideal length equal to the longest seismic signal requested (in general the S-window), and the minimum duration (D_{min}) to ensure a good spectral resolution according to the ratio N/f_{min}, where f_{min} corresponds to the minimum reliable FAS frequency, and N to the number of wavelengths required. In practice, D_{min} is chosen as 10 s, which corresponds to $f_{min} = 0.3$ Hz for $N = 3$. In addition to the pre-event noise window (N1), the Perron algorithm evaluates two post-event noise windows, a long one (N3, with length at least equal to D_t) and a shorter one (N2, length at least equal to D_{min}). Selection of one of these windows is based on their duration and energy level, according to the diagram in Figure 9 (blue boxes). The algorithm also returns a noise flag, indicating which noise window has been selected and how. Seven noise flags are defined, which are listed (in black) in Table .

Noise flag	Description
0	no noise window selected
0b	no noise window selected, but no possibility to evaluate post-event noise windows
1	pre-event noise window selected
1b	pre-event noise window selected, but shorter than target length
2	short post-event noise window selected
3	long post-event-noise window selected
-1	pre-event noise window selected without possibility to take post-event window
-1b	pre-event noise window selected, but shorter than the minimum duration to ensure good spectral resolution
-2	short post-event noise window selected without possibility to take pre-event window
-3	long post-event noise window selected without possibility to take pre-event window

Table I - Noise flags and their definitions. Red rows correspond to additional flags defined for BELSHAKE

For application in BELSHAKE, we refined the Perron procedure (red boxes in Figure 9) in order to minimize the number of rejected records (without noise window) and to favor the pre-event noise window. These modifications involved the definition of three additional noise flags, which are indicated in red in Table :

- If N1 has the required target length (D_t) and its energy is lower than or comparable to that of N3 ($E_{N1} < F5.E_{N3}$, with $F5 = 1.25$), then it is always selected;
- Noise flag 0b: window N1 is shorter than the minimum duration (1 s) required to compute a FAS; as a result, it is not possible to compare its energy with the post-event noise windows N2 and N3, so that these cannot be selected either. However, because these may still be acceptable and could be selected based on visual inspection, we defined this separate category, keeping 0 only for those cases where the post-event noise windows are also too short ($< D_{min}$);
- Noise flag 1b: N1 is selected, even though it is shorter than D_t (but still longer than D_{min}); this occurs when N2 and/or N3 have the required duration, but their energy is too high compared to N1;
- Noise flag -1b: In order to further reduce the number of records with noise flag 0 or 0b, we defined an additional duration threshold (D_{noise_min}), which is the absolute minimum noise-window duration allowed (2 s). This only applies to N1, when N2 and N3 cannot be selected, and the duration of N1 is shorter than D_{min}. For these records, it is not possible to compute SNR over the entire frequency range that is required, but they may still be useful for computing ground-motion values.

We applied the modified Perron noise windowing procedure to all BELSHAKE records and compared the final distribution of selected noise windows with that obtained for the KiK-net dataset by Perron et al. (2018). Figure 10 shows that the pre-event noise window (noise flag 1) is selected for most records and the long post-event noise window (noise flag 3) is the second largest category in both cases. The percentages are slightly different: compared to the KiK-net dataset, BELSHAKE has less records with noise flags 1 and 2, and more records with noise flags 0 and 3.

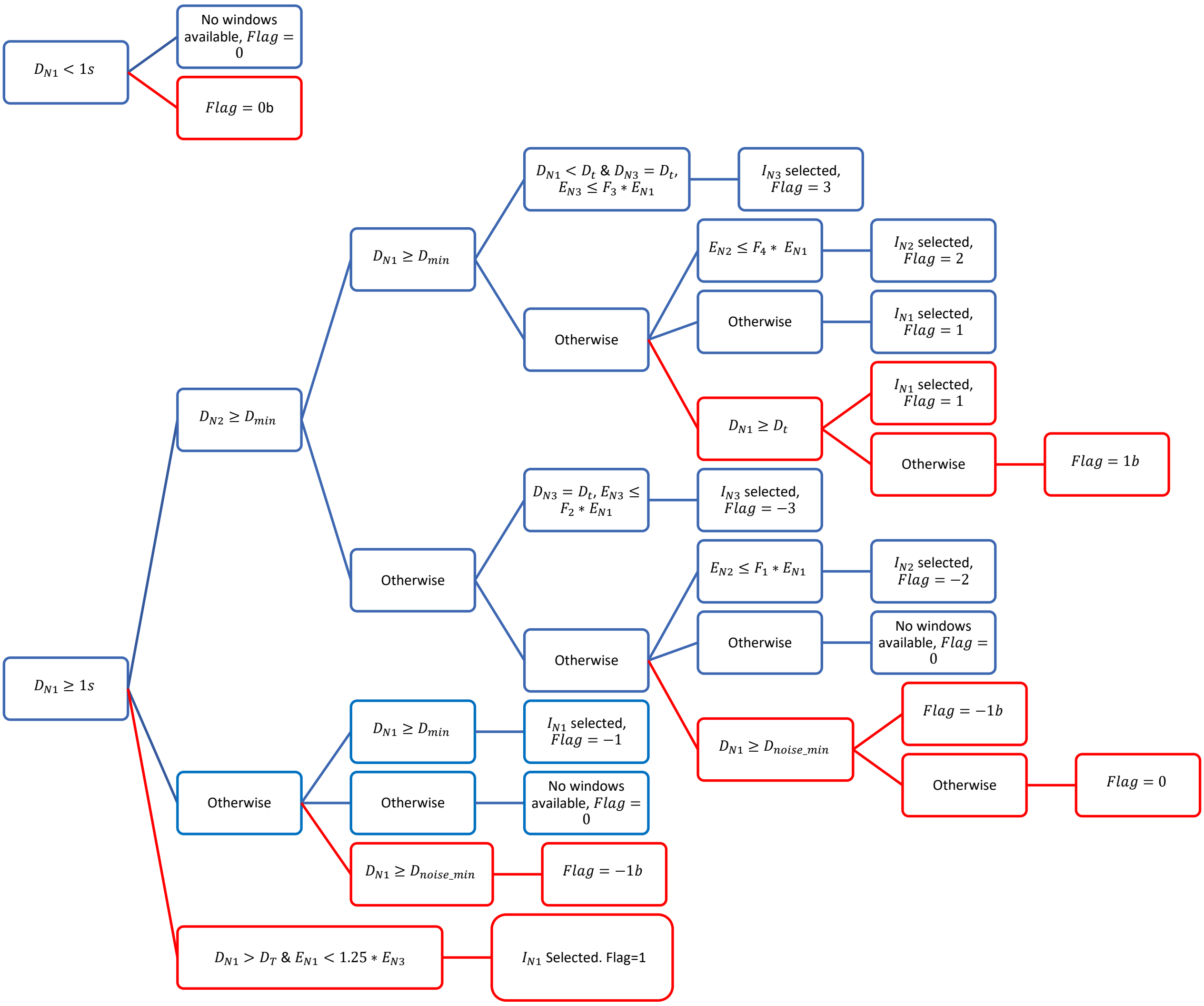


Figure 9. Scheme of the noise window selection algorithm, modified from Perron et al.(2018). Blue boxes correspond to original scheme, red boxes indicate modifications made in the frame of BELSHAKE. Legend: D_{Nn} = duration of noise window n; E_{Nn} = energy of noise window n

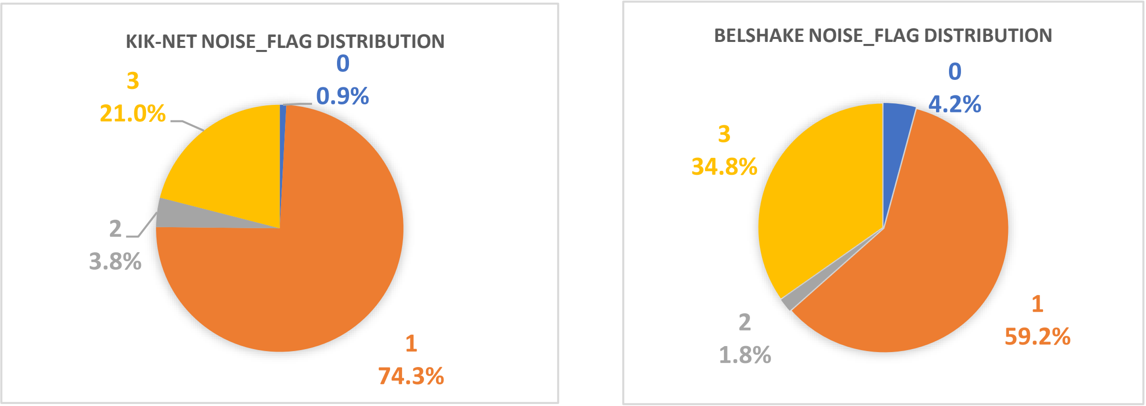


Figure 10. Comparison of noise flag distributions for the KiK-net dataset used by Perron et al. (2018) and for the BELSHAKE database

3.1.6. WAVEFORM PROCESSING

Before ground-motion parameters can be computed, the waveform data must be processed in a uniform way to obtain correct acceleration time series. First, we conducted a literature review of the processing schemes applied in existing ground-motion databases, such as RESIF/EPOS-France (Traversa et al., 2020), PEER NGA-East (Goulet et al., 2014, 2021), ITACA (Paolucci et al., 2011; Puglia et al., 2014) and ESM (Luzi et al., 2016). These processing schemes are relatively similar, but show differences in the details. One of the main differences is that in RESIF the windowing occurs at an early stage (after removal of the instrument response) and all further processing steps are applied to each window separately, whereas in PEER NGA-East windowing is only used to compute FAS and determine the SNR; in ITACA and ESM windowing is not mentioned specifically. Other differences are smaller, e.g. the order of the bandpass filter, the post-processing to ensure that pad-stripped time series give the same results as zero-padded time series, etc.

Based on this comparison, we designed the processing scheme that is summarized in Table . We developed a semi-automated workflow in Python to process all waveform data in the BELSHAKE archive according to this scheme. The code is based on the widely used open-source ObsPy toolkit for processing seismological data (Megies et al., 2011). This processing flow has been applied to all records with reliable response information in the BELSHAKE database. All processing parameters are stored in a separate table in the database. In addition, the main intermediate processing steps are stored with a particular tag in the ASDF files for each event and station. The processing flow can easily be reapplied for a particular event and station, should the original data or the instrument response have to be corrected or should a problem be discovered in the procedure. It is also possible to rerun only part of the processing flow, starting from a particular intermediary step (tag).

Step	Processing	ASDF waveform tag
0	Raw data (optional trimming, timeshift, …)	00raw
[1]	Manual correction (baseline correction, spike removal, …)	[01corr]
2	Window definition (combination of Goulet method for P-, S- and coda windows, and Perron method for noise window(s))	
3	Demeaning / detrending (linear)	
4	Tapering (using fixed lengths instead of taper rates)	
5	Restitution to native ground-motion type (velocity or acceleration)	02rest
[6]	Differentiation (freq. domain) to acceleration	
7	Compute signal/noise FAS to determine SNR and fmin/fmax	
8	Tapering	
9	Zero-phase bandpass filtering with zero padding	

Step	Processing	ASDF waveform tag
10	Re-establish initial time scale (trimming)	03filt
11a	Integration (time domain) to velocity and displacement, and fit 6th-order polynomial with 0th and 1st coefficients constrained to 0; subtract fit from displacement	04dis
11b	Subtract 1st derivative of polynomial fit from velocity	04vel
11c	Subtract 2nd derivative of polynomial fit from acceleration in step 10	04acc
12	Windowing, tapering and zero-padding to length of P- or S-window (whichever is longest)	05acc_noise, 05acc_p, 05acc_s, 05acc_coda, 05acc_signal, 05acc_full

Table II - Summary of processing steps, associated parameters, and corresponding waveform tags in BELSHAKE. Square brackets denote optional steps

The different processing steps are described in more detail in Vanneste & Onvani (2024). We illustrate the most important steps using the waveform data recorded by station BE.MEMH for the $M_L = 3.7$ earthquake of 2007-08-03 near Ochtendung (Germany), shown in Figure 11 to Figure 17.

The processed data, consisting of acceleration, velocity and displacement time histories and corresponding Fourier spectra, are also stored in the ASDF files for each event and station.

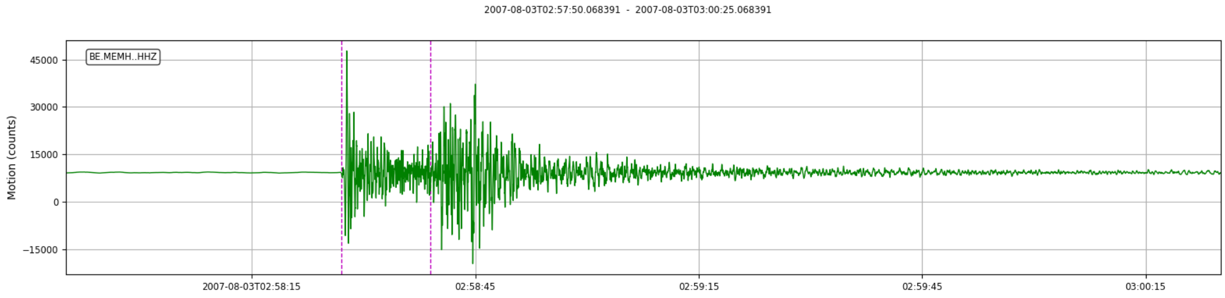


Figure 11 – Z component of raw waveform recorded by station BE.MEMH for the earthquake with event_ID = 2648 in BELSHAKE. Note offset with respect to zero line

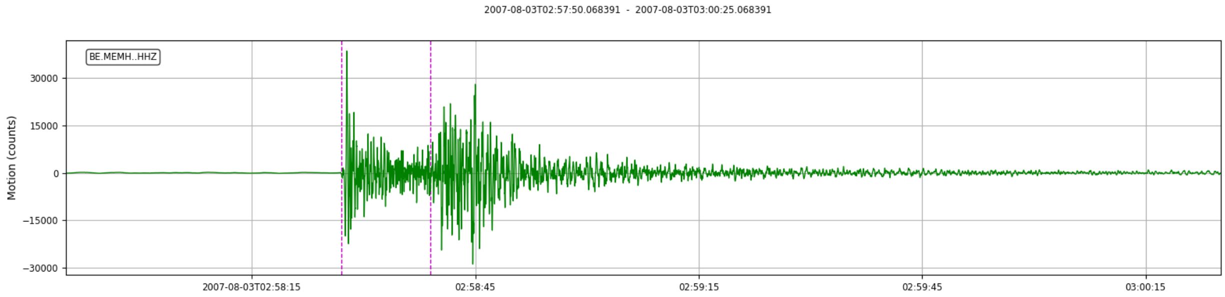


Figure 12 – Mean removal and linear detrending applied to the example waveform (step 3)

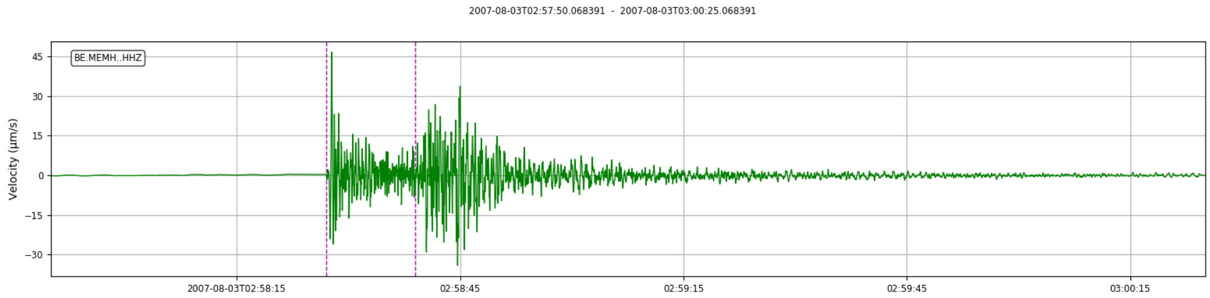


Figure 13 – Example waveform after restitution (removal of instrument response, step 5)

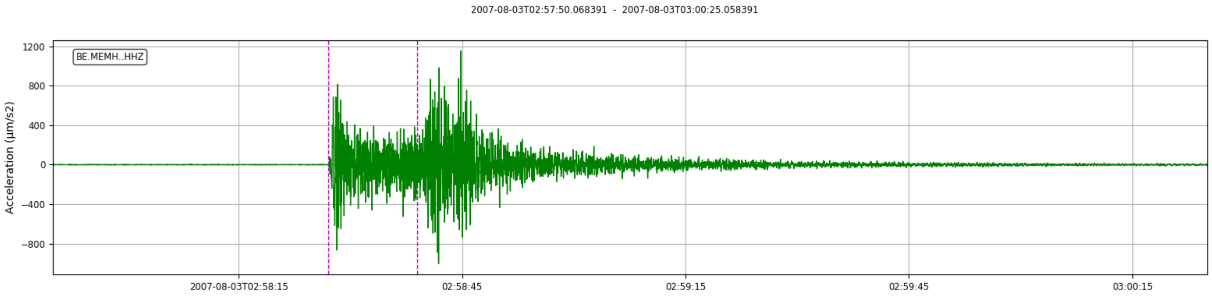


Figure 14 – Acceleration record of the example waveform transformed from velocity (step 6)

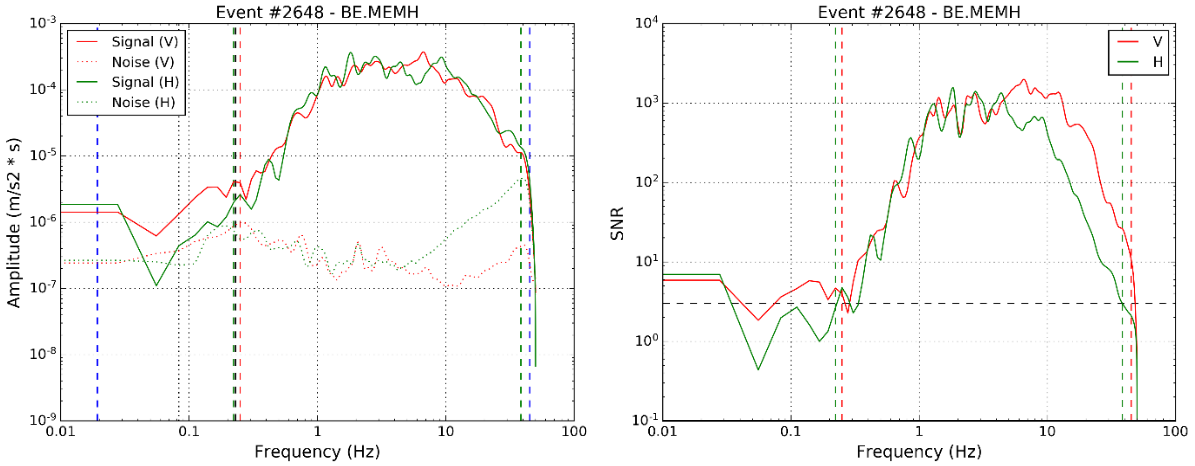


Figure 15 – Left: Fourier amplitude spectra of the vertical and horizontal components of the example record. Right: corresponding SNR curves (step 7). Red and green vertical dashed lines correspond to frequency bounds where SNR ≥ 3, blue vertical dashed lines indicate minimum and maximum reliable frequencies of the instrument, black vertical dashed line corresponds to minimum frequency that is not affected by microseismic noise, horizontal black line indicates SNR=3.

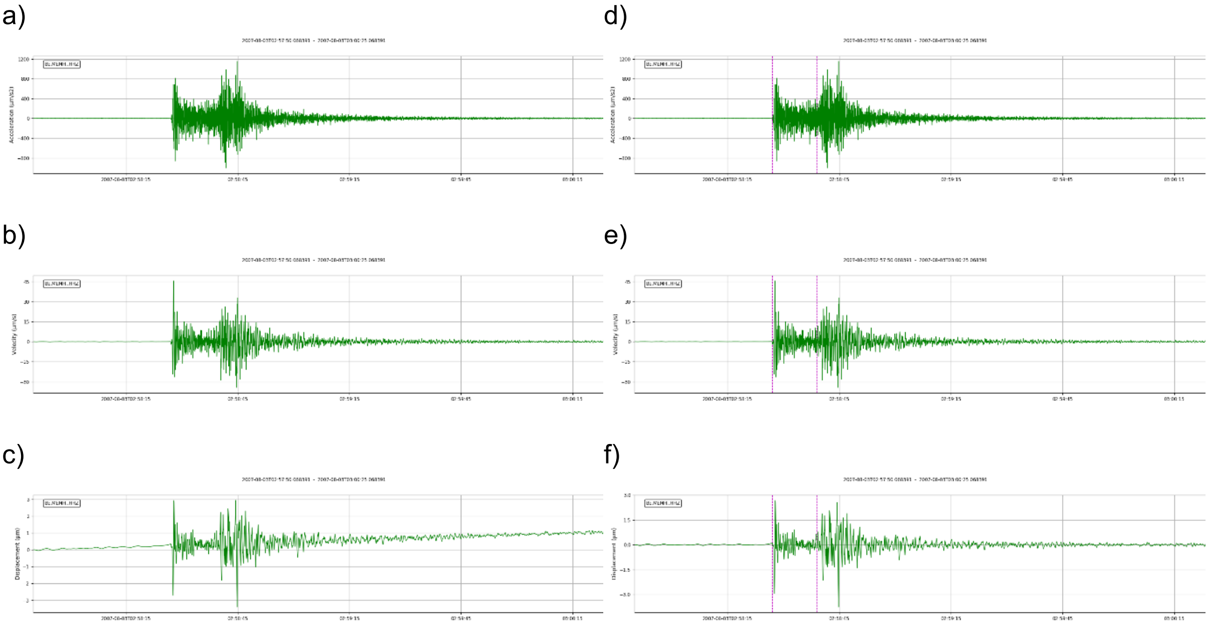


Figure 16. Integration of pad-stripped, acausally filtered acceleration (top) to velocity (middle) and displacement (bottom). Left (a-c): before post-processing; Right (d-f): after post-processing (step 11)

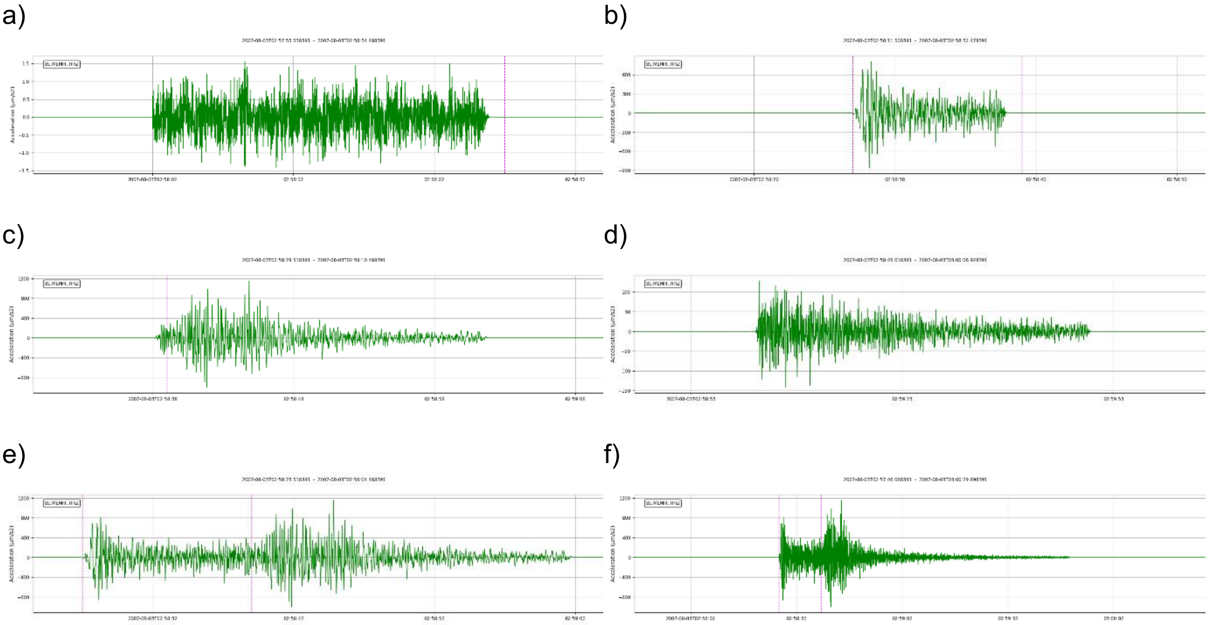


Figure 17. Tapered and zero-padded windows of the acceleration waveform (step 12): a) noise window; b) P-window; c) S-window; d) coda window; e) signal window (P+S); f) full window (P+S+coda)

3.1.7. CALCULATION OF INTENSITY MEASURES

The next step is to compute various ground-motion intensity measures (IMs) in a uniform way from the processed data. These IMs are computed for individual components. The two orthogonal horizontal components can be combined into a single intensity measure in different ways. A common way is to take the geometric mean (GM) of the intensity measures for each horizontal component H_n:

$$GM = \sqrt{IM(H_1)IM(H_2)}$$

However, this geometric mean is not invariant to the orientation of the instrument. To avoid this potentially important drawback, Boore (2010) proposed a new measure of the spectral intensity (denoted as $RotDpp$) that is independent of sensor orientation and encompasses the full range of spectral amplitudes over all possible rotation angles. This new measure is obtained by combining the two orthogonal horizontal components into a single time series H_{ROT} corresponding to a rotation angle θ using the equation:

$$H_{ROT}(\theta) = H_1 \cos\theta + H_2 \sin\theta$$

The response spectrum for $H_{ROT}(\theta)$ is computed, and the process is repeated for a range of rotation angles from 0° up to, but not including, 180°. From the resulting set of spectral values, the *pp*th percentile can then be computed at each oscillator frequency. The rotations corresponding to a given $RotDpp$ thus depend on frequency. We computed two percentiles, $RotD50$ and $RotD100$, which have also been considered in NGA-East (Boore, 2010).

3.1.8. CALCULATION OF MOMENT MAGNITUDES

If we want to compare the ground-motion parameters derived from the processed waveforms in the BELSHAKE database with existing ground-motion databases, we also need to determine the moment magnitude (M_W) of the corresponding earthquakes. Although it was not initially foreseen in our project, the consensus evaluation of our project proposal recommended that “It would be worth contacting representatives of neighbouring countries to discuss methodologies related to moment magnitude estimates and conversion schemes, ...”, highlighting the importance of this parameter. M_W is indeed the standard magnitude scale in ground-motion databases and in ground-motion prediction equations, but it is not routinely determined at ROB. It is only available for a few (mainly the largest) events which have been studied in more detail, but in most cases hitherto we had to rely on empirical magnitude scale conversions (based on M_L or surface-wave magnitude M_S) published in literature.

We conducted a study of the different approaches to determine M_W, and of the available software. The most widely used method is based on fitting of displacement spectra of the P- or S-wave. This method is illustrated in *Figure 18*. The displacement spectrum is computed for the initial part of the P- or S-wave, and a theoretical spectrum given by a source and attenuation model is fitted to that. The shape of this theoretical spectrum is defined by 3 parameters:

$$u(f) = \frac{\Omega_0}{[1 + (f/f_c)^{\gamma n}]^{1/\gamma}} e^{(-\pi f t/Q)}$$

with:

- Ω_0: displacement corresponding to the horizontal plateau of the spectrum at low frequency;
- f_c: corner frequency
- Q: anelastic attenuation factor
- n (usually 2) and γ (1 for Brune (1970) source model, 2 for Boatwright (1980) source model) describe the curvature of the theoretical spectrum

The most important parameter is Ω_0, which is directly related to the seismic moment M_0:

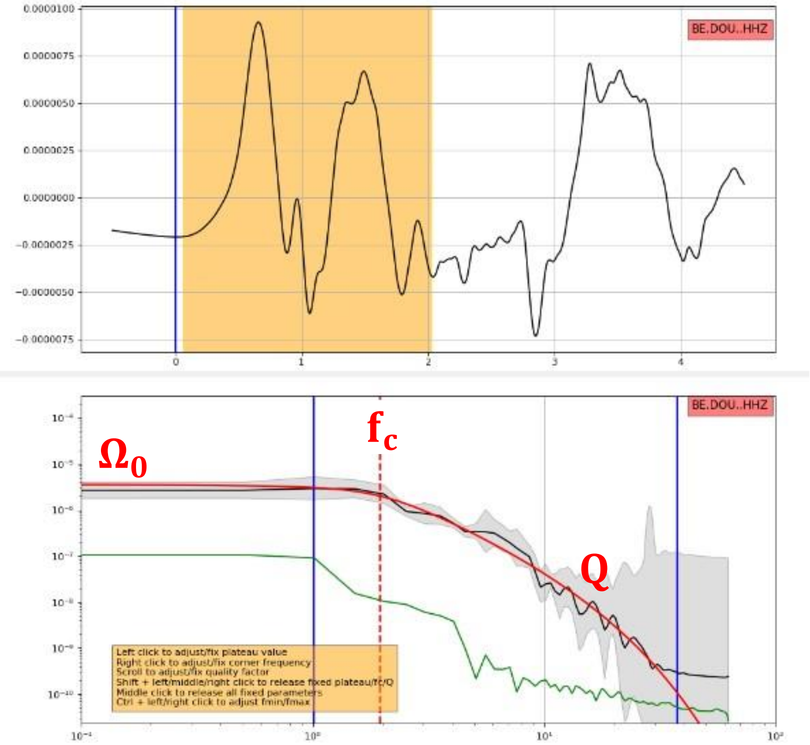
$$M_0 = \frac{4\pi\rho_h {v_h}^3 \Omega_0}{S_a |R_{\theta\varphi}|} \, r \sqrt{\frac{\rho_s v_s}{\rho_h v_h}}$$

with:

- $\rho_h, \ v_h$: density and seismic velocity (P or S) at the hypocenter
- $\rho_s, \ v_s$: density and seismic velocity at the station
- r: hypocentral distance
- S_a: surface amplification (usually 2 for surface stations)
- $R_{\theta\varphi}$: radiation pattern (if focal mechanism is known, otherwise averaged over the focal sphere)

Finally, seismic moment can be converted to moment magnitude M_W with the formula:

$$M_W = {}^2/_3 \log_{10} M_0 - 6.0633$$



Left click to adjust/fix plateau value
Right click to adjust/fix corner frequency
Scroll to adjust/fix quality factor
Shift + left/middle/right click to release fixed plateau/fc/Q
Middle click to release all fixed parameters
Ctrl + left/right click to adjust fmin/fmax

Figure 18 – Illustration of M_W estimation based on fitting of displacement spectra. Top: first 5 seconds of the P-wave on the vertical component, restituted to displacement. The orange interval corresponds to the portion of the signal for which the analysis is performed. Bottom: displacement spectrum of the signal (black) and noise (green), and fitted theoretical source spectrum (red). Blue vertical lines indicate reliable frequency range of the instrument.

We evaluated two Python packages implementing the spectral M_W estimation method and that could be integrated in our work flow: SourceSpec by Claudio Satriano (https://github.com/SeismicSource/sourcespec) and moment magnitude calculator (mmc) by Lion Krischer (https://github.com/krischer/moment_magnitude_calculator). The former is an automated script that requires data to be prepared beforehand and stored in a particular format, but the results can be evaluated from the optional plots. An example is shown in Figure 19. The latter is a program with a graphical user interface, in which the time window used to fit the spectrum can be manually selected (as opposed to SourceSpec, which uses fixed windows), and the spectral fitting is done interactively. Both programs have their pros and cons, but ultimately we preferred SourceSpec because it is actively developed and has more potential to run in a semi-automated way. For that purpose, we wrote Python code to run SourceSpec directly from data in the BELSHAKE database without the need for storing intermediary data in a different format and creating a configuration file. We also modified the source code of this program to improve the success rate with our specific data, which contain single-component records, records with short or missing pre-event noise windows, etc. These improvements included:

- Bug fix in smoothing of spectra;
- Bug fix for SV waves ;
- Allowance for different hypocentral and station velocities;
- Allow using P waves instead of only S waves ;
- Implementation of a run ID to discern multiple runs for the same input data;
- Relaxation of the requirement to have a sufficiently long noise window (see noise windowing in §3.1.5);
- Variable window length for calculating displacement spectra as a function of travel time (equivalent to distance);
- Implementation of a segmented geometrical spreading function instead of the simple $1/R$ model;
- various other changes.

These code changes were shared with the author and merged in the main development branch. As a result, one of the authors of this report (KV) has become an official co-developer of the SourceSpec package. Thanks to all these developments, it has become possible to compute moment magnitudes for all events with sufficient records in BELSHAKE in a semi-automated way, taking advantage of the metadata (component/window quality, (station-specific) reliable frequency range, response reliability, etc.).

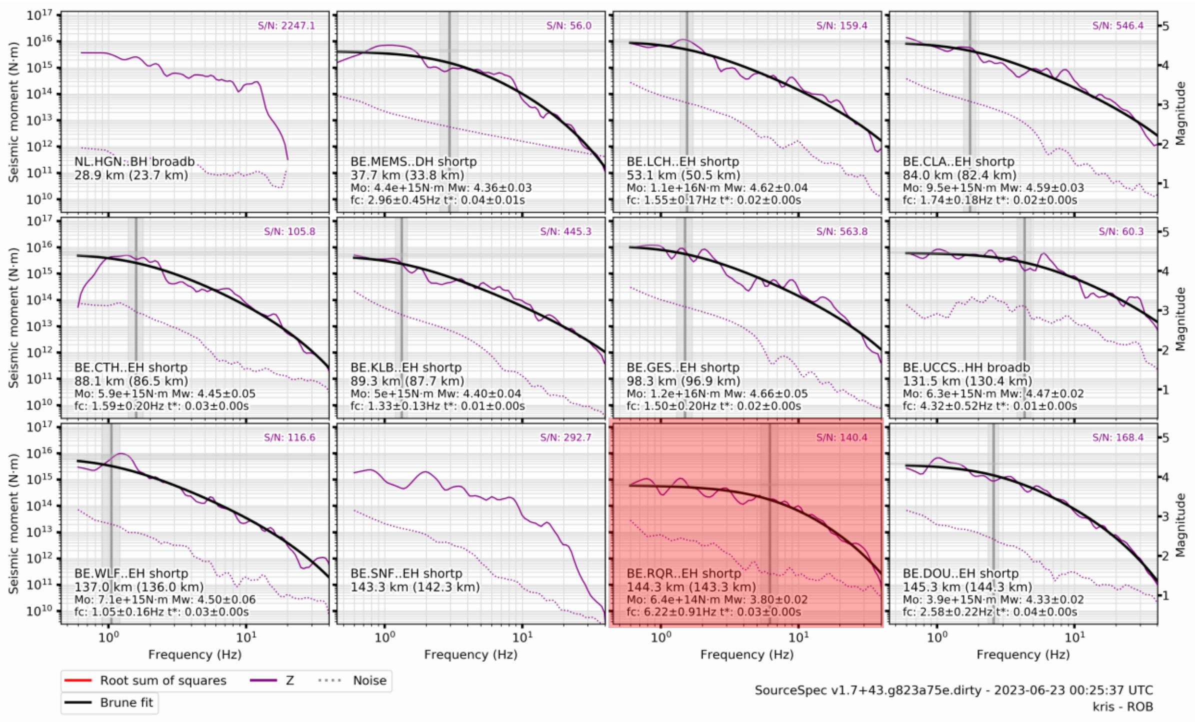


Figure 19 - Example of fitted displacement spectra for a particular earthquake in the BELSHAKE database. Note that the spectra of 2 stations (NL.HGN and BE.SNF) could not be fitted. Red shading indicates station yielding a M_W value that is too low and is discarded

3.1.9. DATA VALIDATON

To ensure the quality and reliability of ground-motion intensity measures in the BELSHAKE database, we applied different consistency checks to identify anomalous records. Outliers were examined, and those considered as problematic were either corrected and reprocessed or removed from the database. The following evaluations were performed: (1) evaluation of horizontal component ratios; (2) evaluation of horizontal/vertical ratios; (3) analysis of total residuals with respect to a generic GMM; (4) residual analysis based on mixed-effects regression; (5) comparison of co-located stations; (6) comparison with common events in the RESIF database. Part of the residual analysis based on mixed-effects regression is described in the results section (§4.1.8). Below, we provide more details about the other analyses.

3.1.9.1. Comparison of different horizontal components

First, we evaluated PGV ratios of different horizontal components, e.g. GM/RotD100, for the different signal windows, as shown in Figure 20. All records with GM/RotD100 ratio < 0.4 were inspected. The outliers in the S, signal and full windows were mostly related to records where one of the horizontal components has an anomalously low amplitude. We removed the derived IMs from the database and flagged these components as problematic, so that they will be skipped when IMs are recomputed. Some outliers in the P window correspond to records where the vertical component was swapped with one of the horizontal components. This appeared to be the case for several events recorded by station BE.HRK. For these records, we swapped the vertical and horizontal components, reprocessed them and recalculated the IMs.

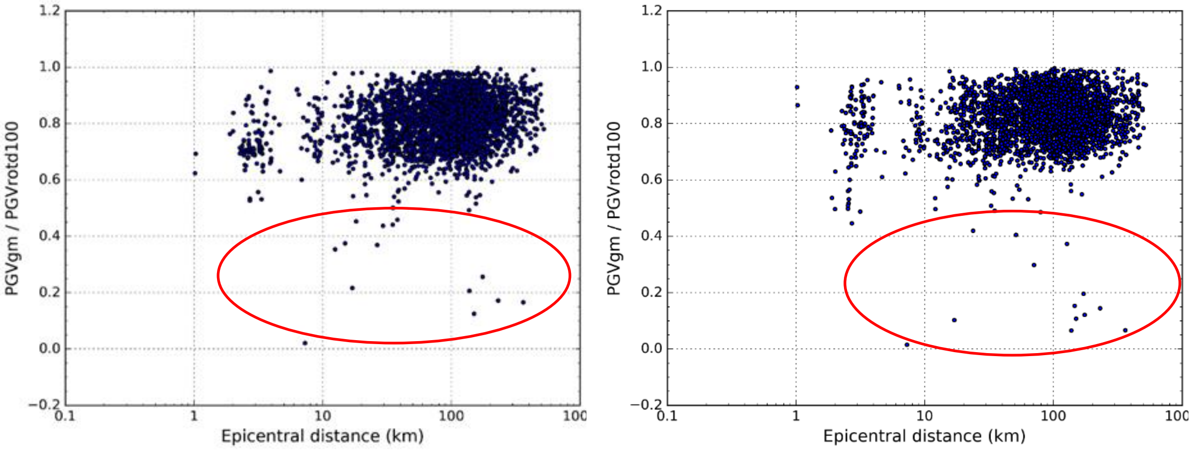


Figure 20 - $PGV_{GM}/PGV_{RotD100}$ ratios for records in the BELSHAKE database, computed for different windows. Left: P-window; Right: S-window. Outliers are indicated by red ovals

3.1.9.2. *Comparison of horizontal and vertical components*

Secondly, we calculated the PGV ratio between horizontal and vertical components, e.g. GM/Z, for the different signal windows, as shown in Figure 21. We examined records featuring large and small H/V ratios. The majority of records associated with high H/V ratios appear to be short-distance records ($R_{epi} < 20\ km$). These were not removed because high H/V ratios may be a characteristic of short-distance records. Fewer records show anomalously small H/V ratios. Most of them correspond to records where one of the horizontal components has an anomalous amplitude, and these were already identified in the previous step. A suspicious case is station BE.MOLA where a number of records have very small H/V ratios in the P-window. This appeared to be an effect of resonance, caused by bad emplacement, as suggested by a peak between 20 and 30 Hz in the FAS of the vertical component. These records were flagged as problematic and the associated IMs were removed.

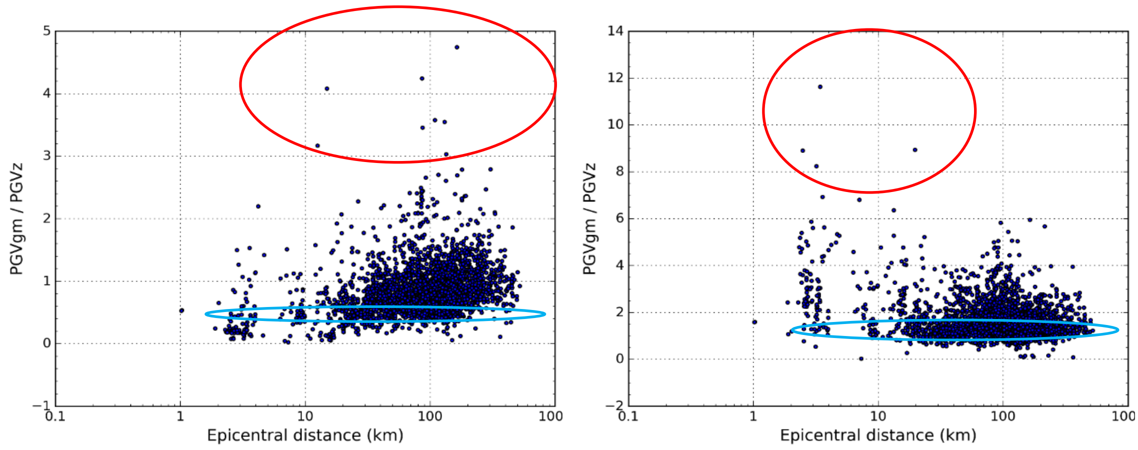


Figure 21 - PGV_{GM}/PGV_{Z} ratios for records in the BELSHAKE database, computed for different windows. Left: P window; Right: S-window. Outliers with high ratios are indicated by red ovals, outliers with low ratios by blue ovals

3.1.9.3. Analysis of total residuals with respect to an ad hoc GMM

This analysis involved a comparison between PGA/PGV in the database and PGA/PGV predicted by a generic GMM fitted to the observations. The advantage of an ad-hoc GMM compared to a published model is that its median is centered on the dataset, which allows identifying inconsistencies with higher sensitivity (Traversa et al., 2020). We used the same functional form as Traversa et al. (2020) in their analysis of the French RESIF/EPOS-France database:

$$log_{10}Y = F_D(R,M) + F_M(M) + e_1 \tag{1}$$

where:

$$F_D(R,M) = \left(c_1 + c_2\left(M - M_{ref}\right)\right)\log_{10}\left(\frac{\sqrt{R^2+h^2}}{R_{ref}}\right) - c_3\left(\sqrt{R^2+h^2} - R_{ref}\right) \tag{2}$$

is the distance scaling component;

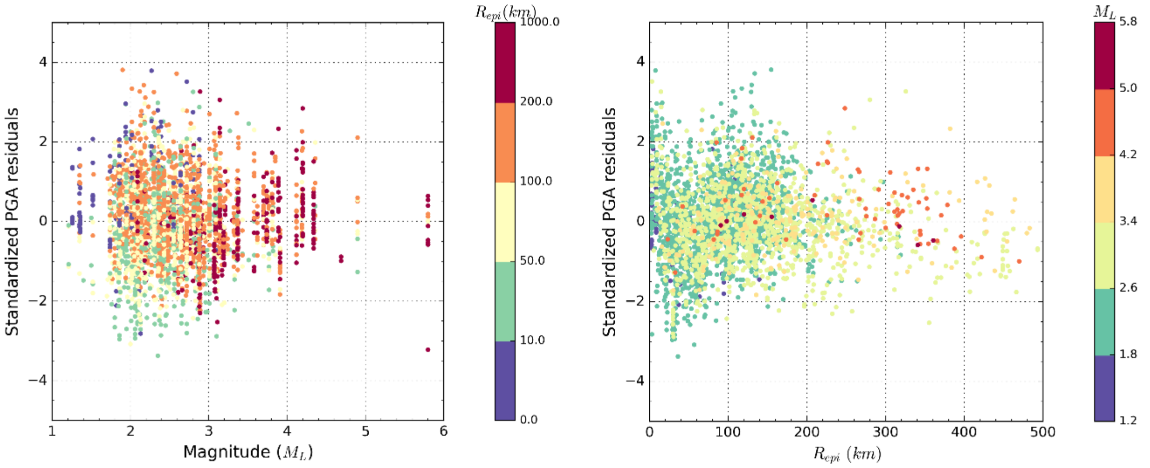
$$F_M(M) = b_1\left(M - M_{ref}\right) \tag{3}$$

is the magnitude scaling component;

and e_1 is the global offset or total residual.

The coefficients b_1, c_1, c_2 and c_3 are determined from the regression. We used the same magnitude scale (M_L), reference distance ($R_{ref} = 1\ km$) and reference magnitude ($M_{ref} = 3.7$) as Traversa et al. (2020).

In a first step, we computed total residuals using ordinary least-squares regression for the full window and different components, including GM, RotD50 and RotD100. This serves as a basic check for identifying outliers. Figure 22 shows the standardized residuals as a function of M_L and epicentral distance for PGA and PGV, in both cases for the GM component. The majority of standardized residuals fall within the range of ±2 standard deviations, with nearly all of them between ±3. The presence of outliers (residuals > 3) is minimal. A few of them could be attributed to poor-quality recordings and have been deleted.



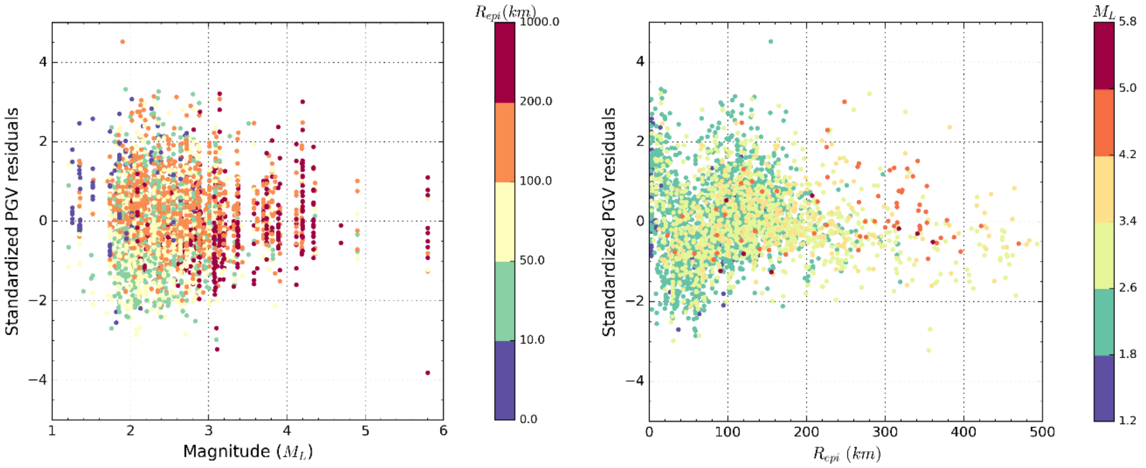


Figure 22 – Standardized total residuals with respect to ad hoc GMM as a function of M_L (color-coded by distance, left) and distance (color-coded by M_L, right). Top: PGA, GM component, full window; Bottom: PGV, GM component, full window

Further insight could be gained by analyzing the temporal evolution of the total residuals for each station individually. During this analysis, we identified anomalies for three specific stations, namely BE.UCCS, BE.HRK, and BE.HUM, which exhibited notable shifts within a specific time period (Figure 23). In the case of BE.UCCS, the affected period corresponds to a particular station deployment and the shift is due to a problem with the instrument response. We have corrected and reprocessed these records. For the two other stations, the outliers appear to be associated with a particular earthquake source, i.e. events from a seismic sequence or swarm originating in the same region close to the station. These records have not been corrected or removed.

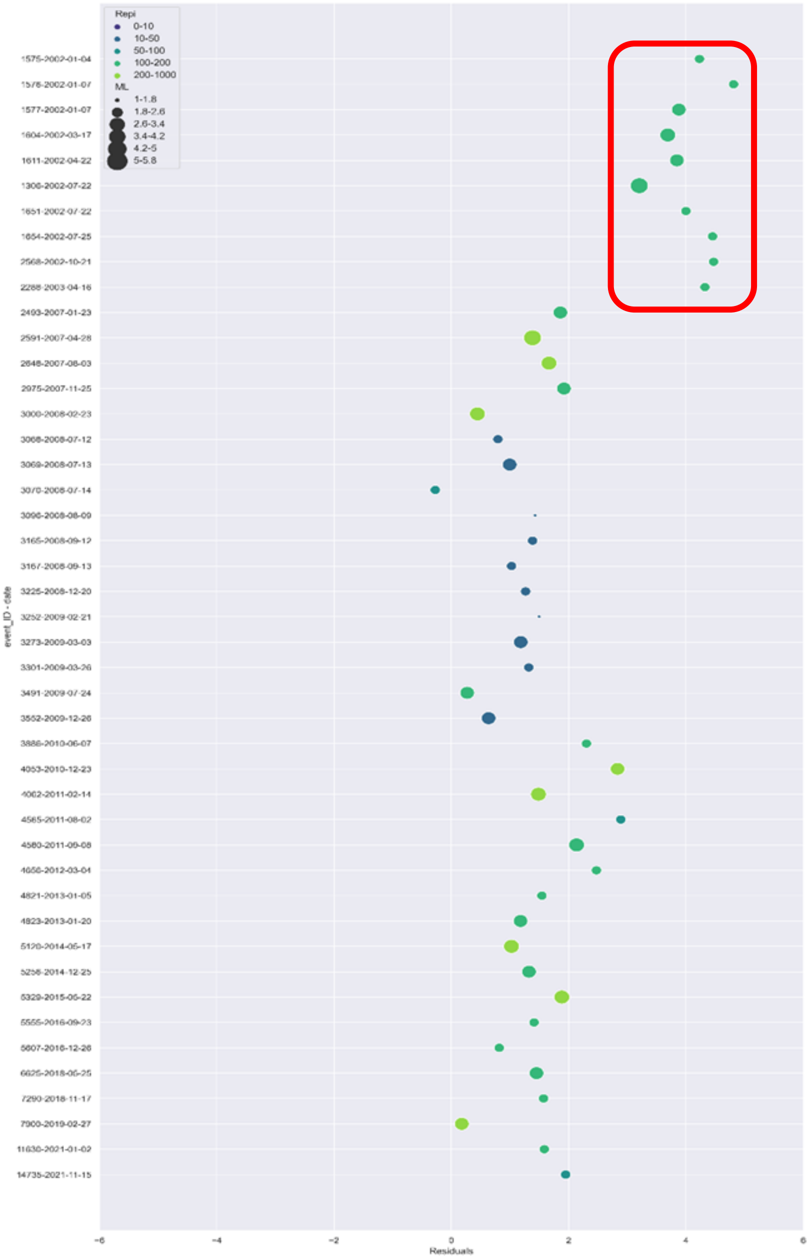


Figure 23 - Evolution of total residual for PGA (GM component) recorded at BE.UCCS. Colors are function of the record epicentral distance and circle sizes are function of event M_L magnitude. Red squares indicate periods with notable shifts

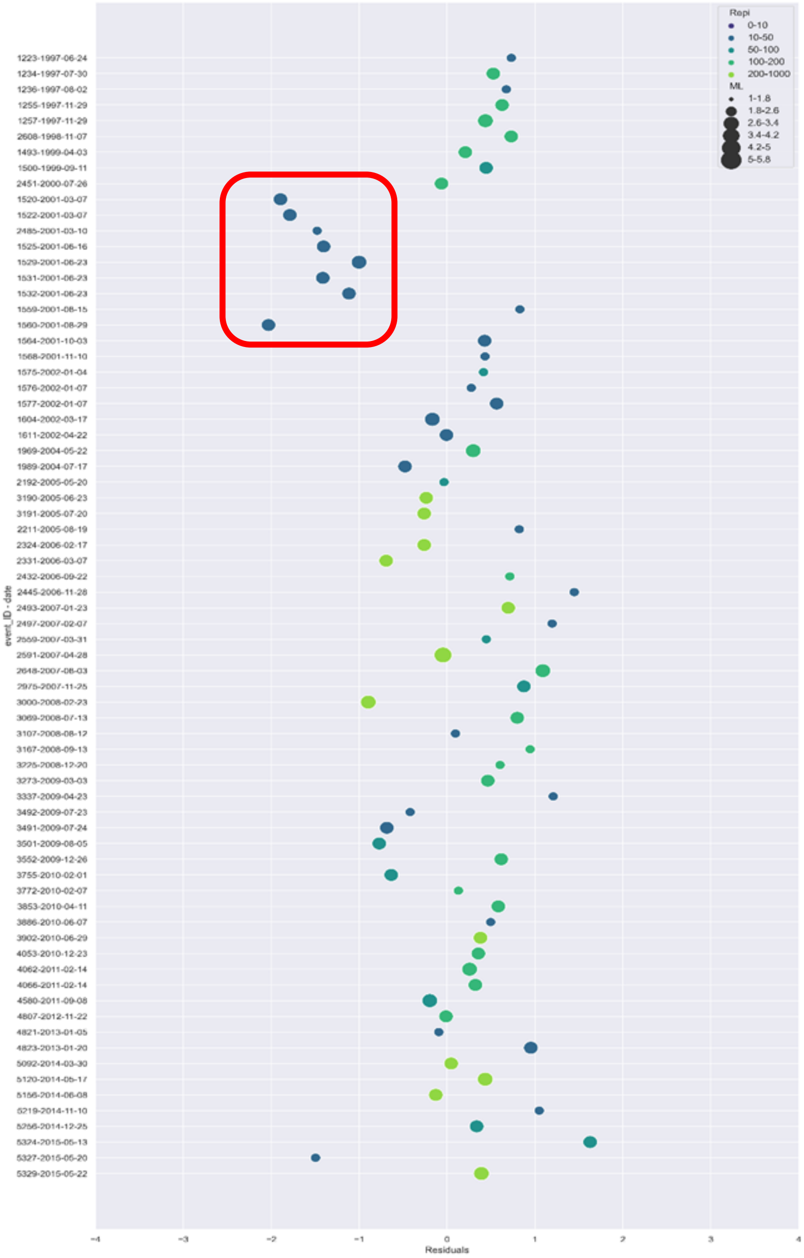


Figure 24 - Evolution of total residual for PGA (GM component) recorded at BE.HRK. Colors are function of the record epicentral distance and circle sizes are function of event M_L magnitude. Red squares indicate periods with notable shifts

3.1.9.4. Residual analysis using mixed-effects regression

In a next step, we performed the GMM fit using mixed-effects regression (using the python statsmodels module, Seabold & Perktold, 2010), which allows to decompose the total residual into different components:

$$log_{10}Y = F_D(R, M) + F_M(M) + \delta B_e + \delta S2S_s + \epsilon \tag{4}$$

Y is the ground-motion parameter, in this case PGA or PGV; $F_D(R, M)$ and $F_M(M)$ are the distance and magnitude scaling components of the GMM (see previous section), which are considered as the fixed effects; δB_e and $\delta S2S_s$ are random effects corresponding to between-event and between-station residuals, respectively; ϵ is a lognormally distributed residual term, representing the unexplained or event- and station-corrected residual. This separation gives better insight in the origin of possible inconsistencies (e.g., related to source parameters, station/site parameters, …). Note that we used M_W instead of M_L in this analysis.

The between-station, between-event and unexplained residuals are discussed in the results section (§4.1.8). In addition, plots of the time evolution of unexplained residuals for each station allowed identifying anomalies that are consistent over a particular period of time. These may indicate problems with the instrument response or important changes of the instrument emplacement. Two examples are shown in Figure 25: station BE.LCH (left panel) shows a clear negative bias for the Z component during the first deployment. We have not been able to find the cause of this problem, so we flagged the instrument response as unreliable for this period. For station BE.BRQ (right panel), all residuals during the first deployment show a positive bias with respect to subsequent deployments. This shift can be explained by the migration of the station into a 30-m deep borehole from the second deployment onwards. We solved this by assigning different location codes to the surface and borehole emplacements.

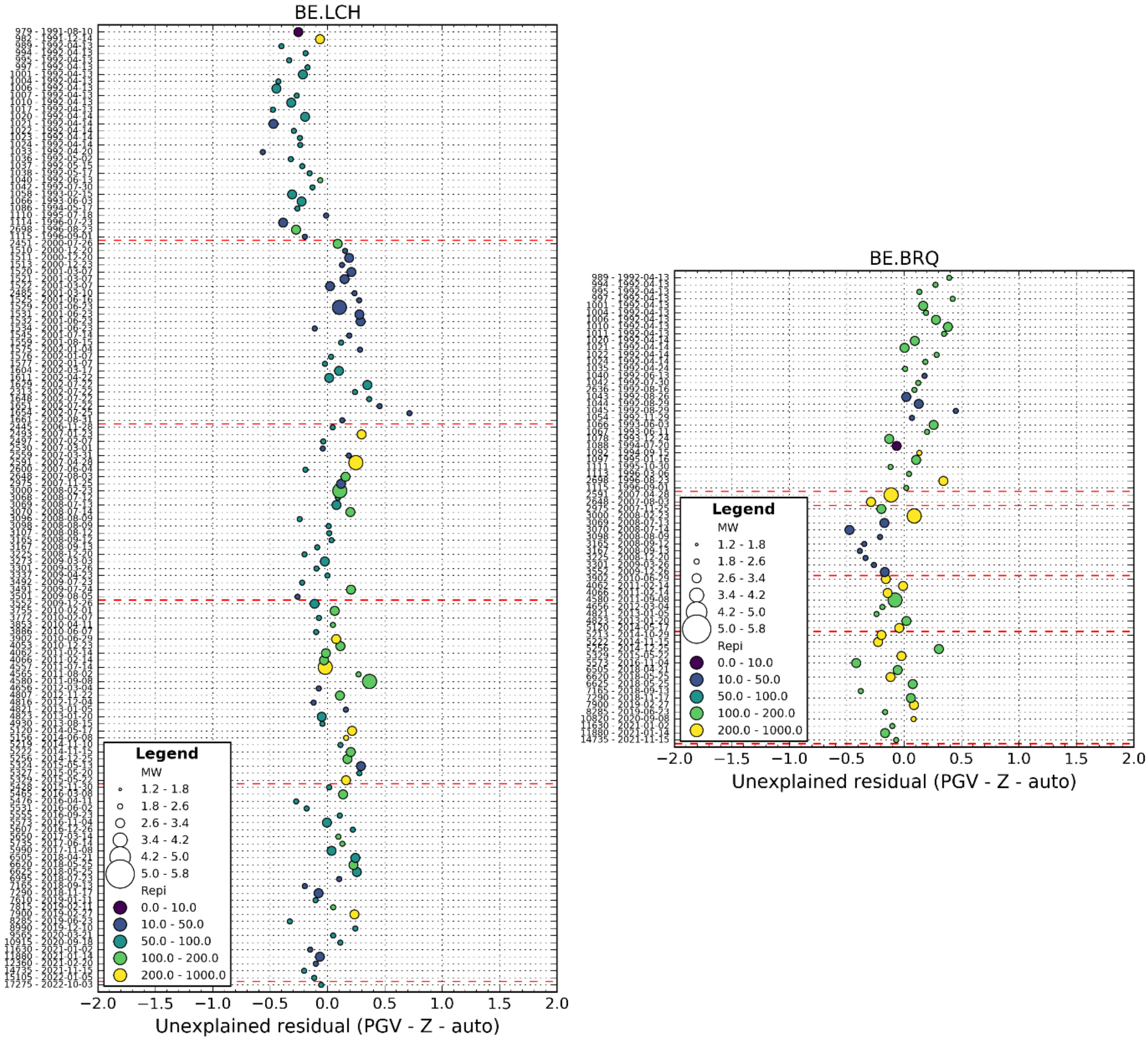


Figure 25 - Time evolution of unexplained residuals (PGV, Z component, "auto" window) for particular stations. Left: BE.LCH; Right: BE.BRQ. Red dashed lines correspond to changes in the station configuration. Note shift of residuals between first (counting from top to bottom) and second periods in both cases

3.1.9.5. Comparison of co-located stations

In a few places, different instruments are located next to one another. We can check the consistency of these co-located stations by comparing the IM values for common events and by evaluating the spectral ratio of Fourier amplitude spectra from both instruments, for the same component.

Stations with co-located broadband seismometers and accelerometers

In three locations, a broadband seismometer is situated in proximity to an accelerometer. This is the case in Uccle (BE.UCCS and BE.UCCA), in Mol at the surface (BE.MOLS and BE.MOLA) and in an underground tunnel (BE.MOL5 and BE.MOLB). For the latter pair, which are only 1 m apart, agreement is quite good for all intensity measures (Figure 26, top row). This is confirmed by the spectral ratios (Figure 27, left panel), which are close to 1 over the entire frequency range, except for the first two records, which appear to be disturbed by high-frequency noise. This is no longer the case for more recent events, suggesting that the emplacement of one of the instruments has been improved in between. The older records were flagged as unreliable and the corresponding IMs were removed. Between the surface stations BE.MOLS and BE.MOLA, agreement for PGA and PGV is poor Figure 26, bottom row). Spectral ratios (Figure 27, right panel) are close to 1 between 1 and 10 Hz, but above 10 Hz, BE.MOLS shows clearly higher amplitudes than BE.MOLA. This is probably because the two instruments are too far apart (135 m) and in proximity to different buildings.

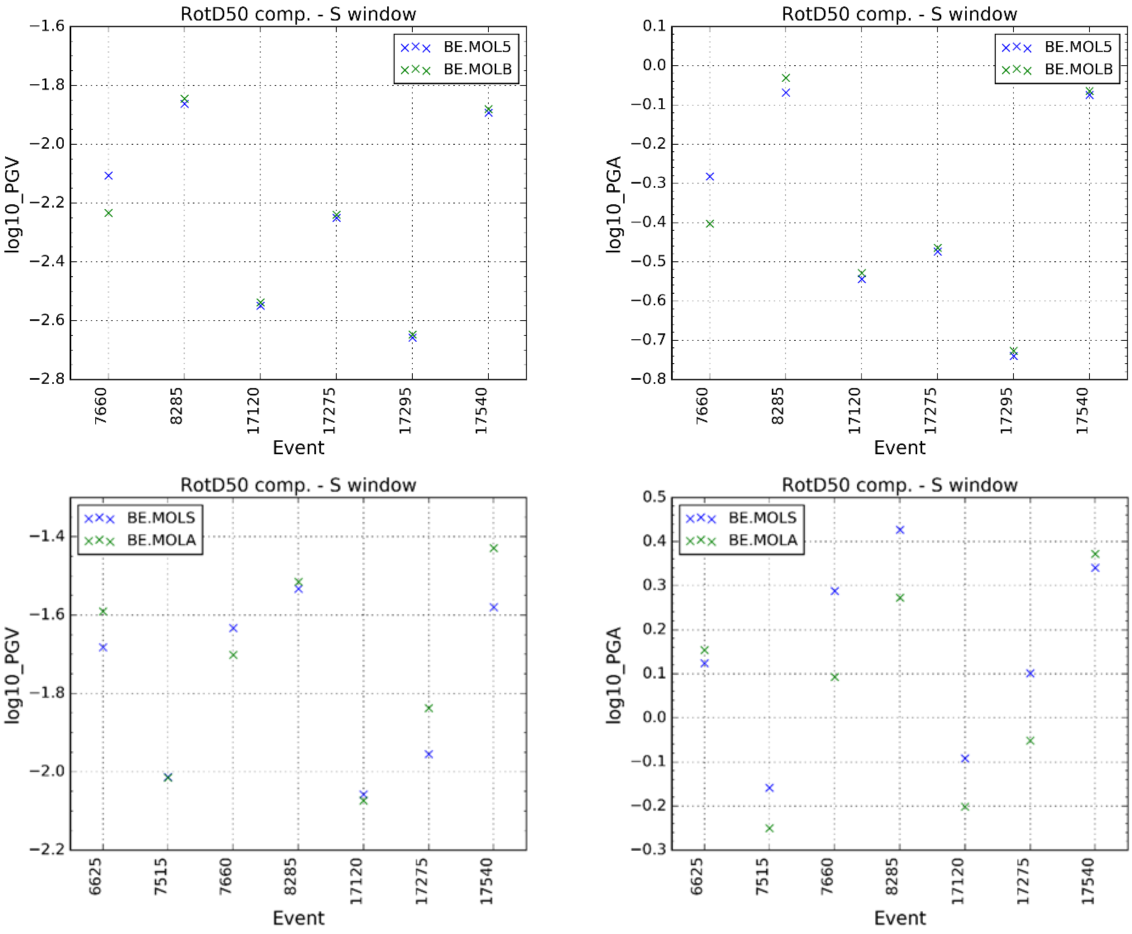


Figure 26- Comparison of IMs (RotD50 component, S window) computed for common events recorded by co-located station pairs Top row: BE.MOL5 (broadband) versus BE.MOLB (accelerometer); Bottom row: BE.MOLS (broadband) versus BE.MOLA (accelerometer). Left column: PGV; Right column: PGA

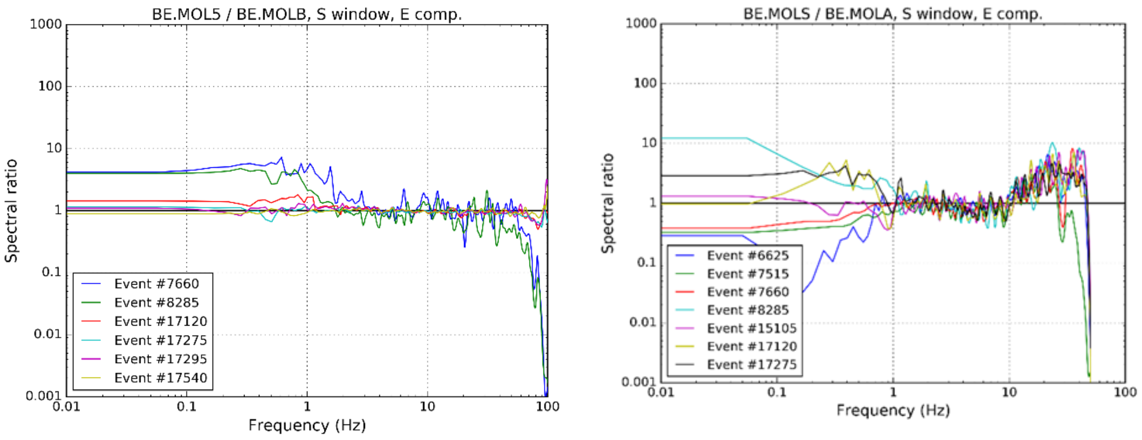


Figure 27– Spectral ratios (S window, E component) for events recorded by co-located station pairs. Left: BE.MOL5/BE.MOLB; Right: BE.MOLS/BE.MOLA

Finally, at Uccle the agreement between accelerometer BE.UCCA and broadband seismometer BE.UCCS is quite good for PGV, but the differences are quite large for PGA (Figure 28, top panel). Inspection of the waveform data shows that the accelerometer records are affected by high-frequency disturbances. Spectral ratios (Figure 28, bottom panel) are close to 1 between ~2 and 45 Hz, above which there is a steep increase. Similar to BE.MOLS/BE.MOLA, there is also a divergence at lower frequencies. The high-frequency disturbance is due to emplacement of BE.UCCA on a pillar. We solved the issue by reprocessing these data with an imposed maximum frequency of 45 Hz.

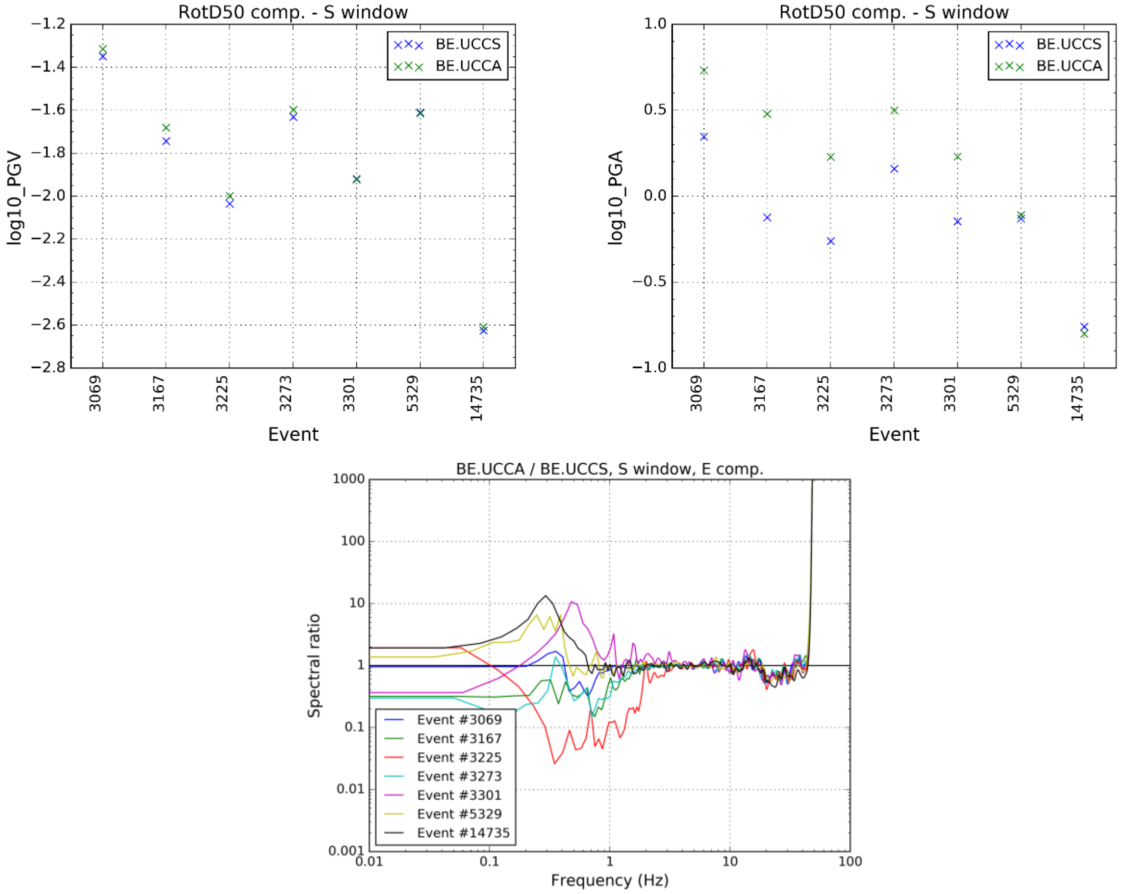


Figure 28 – Comparison of co-located station pair BE.UCCA/BE.UCCS. Top row: Comparison of IMs (RotD50 component, S window) computed for common events recorded. Left: PGV; Right: PGA. Bottom: spectral ratios (S window, E component) for the same events

In conclusion, agreement between broadband seismometers and accelerometers in the same station should be quite good, but only if they are close to each other and their emplacements are similar.

3.1.9.6. Comparison of broadband and short-period seismometers

In the Membach underground gallery, a short-period seismometer (BE.MEMS) has been operating next to two broadband seismometers: BE.MEMB (1994-2005) and BE.MEMH (2006-present). The agreement between BE.MEMS and BE.MEMH is quite good (Figure 29, top panel). This is an important result, demonstrating that short-period and broadband seismometers provide equivalent results over a relatively broad frequency range, at least for the moderate-magnitude seismicity in Belgium. The agreement between BE.MEMS and BE.MEMB is mostly good (Figure 29, bottom panel), except for some events (for all IMs). There likely is a problem with the instrument response of BE.MEMB for these events, so we flagged them as unreliable and removed the corresponding IMs from the database. These findings are confirmed by the spectral ratio plots (Figure 30), which are close to 1 up to 10-20 Hz for the BE.MEMH/BE.MEMS pair, but show larger divergence outside the 1-10 Hz range for BE.MEMB/BE.MEMS.

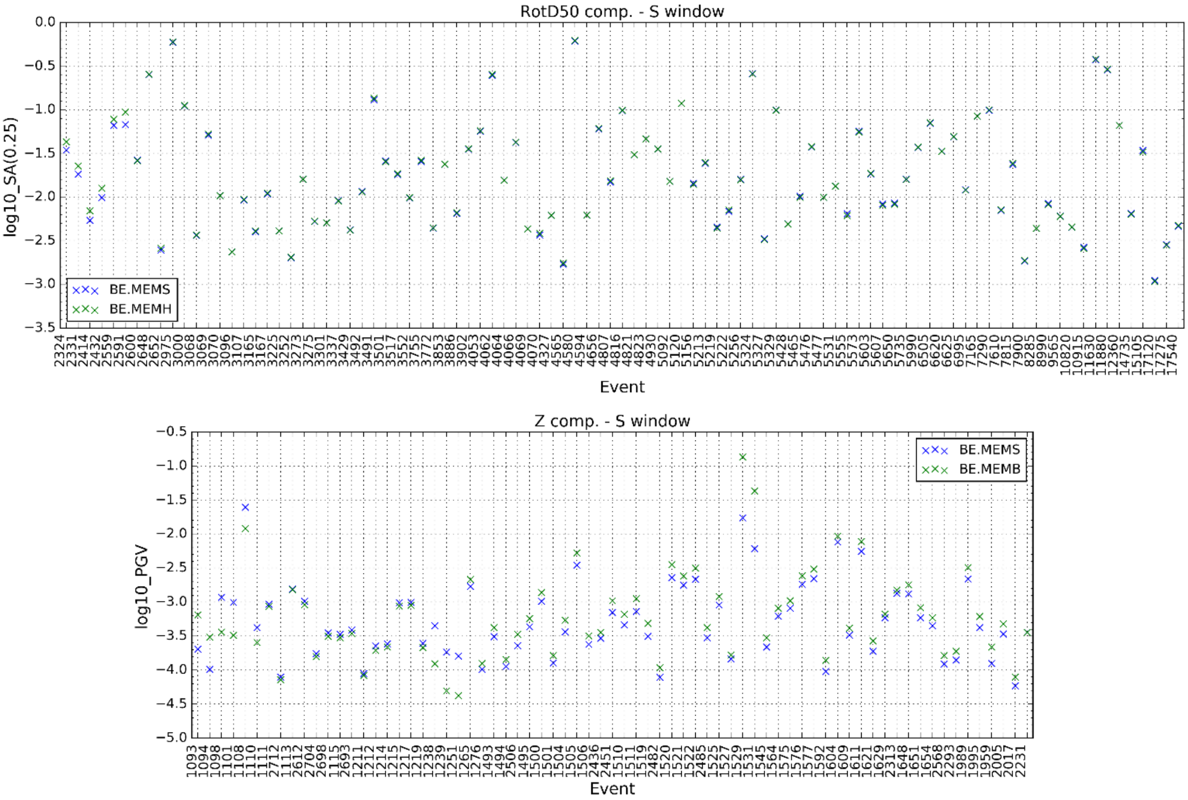


Figure 29– Comparison of IMs calculated for common events recorded by different instruments in the Membach underground gallery. Top panel: 0.25 s spectral acceleration (RotD50 component, S window) for stations BE.MEMS and BE.MEMH; Bottom panel: PGV (Z component, S window) for stations BE.MEMS and BE.MEMB

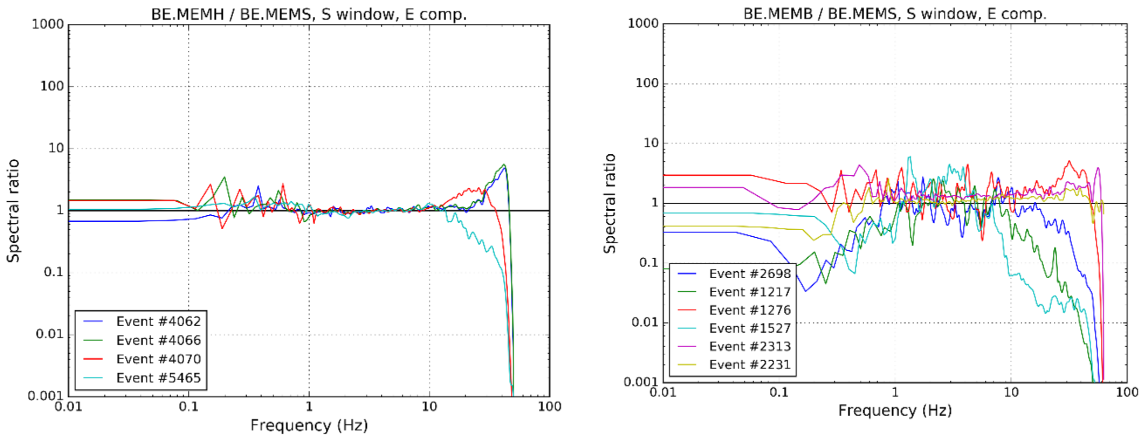


Figure 30– Spectral ratios (S window, E component) for events recorded by co-located station pairs. Left: BE.MEMH/BE.MEMS; Right: BE.MEMB/BE.MEMS

3.1.9.7. Comparison with the French RESIF/EPOS-France database

The BELSHAKE database also has 14 events in common with the RESIF/EPOS-France database (Buscetti et al., 2025; Traversa et al., 2020). For each event, we have plotted IM values versus hypocentral distance for records in BELSHAKE and RESIF (flatfile version 2024-04-02). An example ($M_L = 4.1$ earthquake near Ramsgate, UK, in 2015) is shown in Figure 31. In the majority of cases, the agreement is quite good and both datasets exhibit the same trends with distance. This indicates that there are no major problems with the instrument response and the processing procedure in BELSHAKE. There are, however, some important differences between the two databases in terms of source parameters (location, focal depth, magnitude) for these common events, but these are not important when we only make a direct comparison of their IM values.

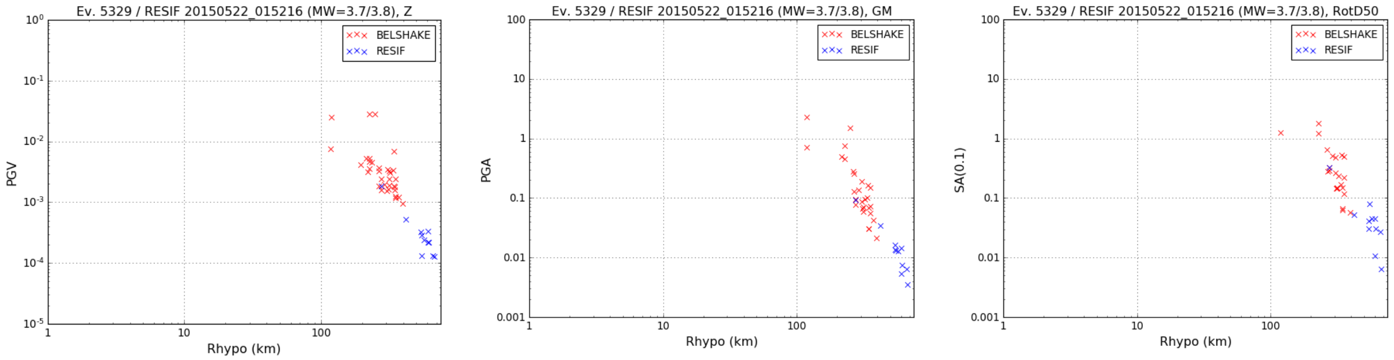


Figure 31- Comparison of IMs computed for a common event in the BELSHAKE (red) and RESIF/EPOS-France (blue) databases. Left column: PGV, Z component; Middle column: PGA, GM component; Right column: SA(T=0.25), RotD50 component

3.2. Building the capacity to model earthquake ground motions in Belgium

3.2.1. STUDY OF REGIONAL CRUSTAL ATTENUATION

The waveform data collected in the BELSHAKE database can be used to study crustal attenuation in the main crustal domains in Belgium, more specifically anelastic attenuation represented by the high-frequency spectral decay parameter κ. This parameter was introduced by Anderson & Hough (1984) to explain the exponential decay observed at high frequencies in acceleration FAS of shear waves, compared to the FAS corresponding to a simple ω-square source model (Brune, 1970). This decay is expressed by:

$$A(f) = A_0 e^{-\pi\kappa f}$$

and is observed for frequencies f between a lower bound f_1 (which should be higher than the corner frequency f_c) and an upper bound f_2 (where spectral amplitude of the signal approaches that of noise). Anderson & Hough (1984) also observed that κ values increase with epicentral distance R, and concluded that it can be written as the combination of a path-dependent term due to predominantly horizontal S-wave propagation through the crust and a site-dependent term (κ_0) representing the attenuation of S waves traveling vertically through the local bedrock beneath the station:

$$\kappa(R) = \kappa_0 + \kappa_r R = \kappa_0 + \kappa_{path}$$

Here, we describe the methods used to (1) determine whole-path κ values for individual records in the BELSHAKE database, and (2) estimate the site-specific component κ_0 and the regional kappa gradient κ_r for stations and crustal domains in Belgium.

3.2.1.1. Whole-path κ calculation

We calculated κ values for individual records in the BELSHAKE database as the slope of a linear fit of S-wave acceleration FAS in log-linear space (Anderson & Hough, 1984; Ktenidou et al., 2014). First, we calculated FAS from the processed acceleration time series in the BELSHAKE database. For all three components, vertical (Z), east (E), and north (N), the lower and upper bounds of the fitting frequency range (f_1 and f_2) were manually selected for each record. These frequencies were chosen within the reliable frequency range, determined in a prior analysis based on a minimum signal-to-noise ratio (SNR) of 3. The frequency range was selected carefully to exclude resonance peaks and to ensure it remained above the corner frequency. This visual selection was done on a record-by-record basis to ensure appropriate fitting and to maintain consistency across the dataset. The S-wave FAS was smoothed, a linear ordinary least-squares (OLS) regression was applied over the selected frequency range, and κ was calculated by dividing the fitted slope by π. For the horizontal component, κ was calculated as the arithmetic average of the values derived from the east and north components. Figure 32 illustrates this process for an example spectrum, showing the fit within the selected frequency bounds. This procedure yielded a robust dataset of κ values for ~7800 records involving 359 events and 132 stations.

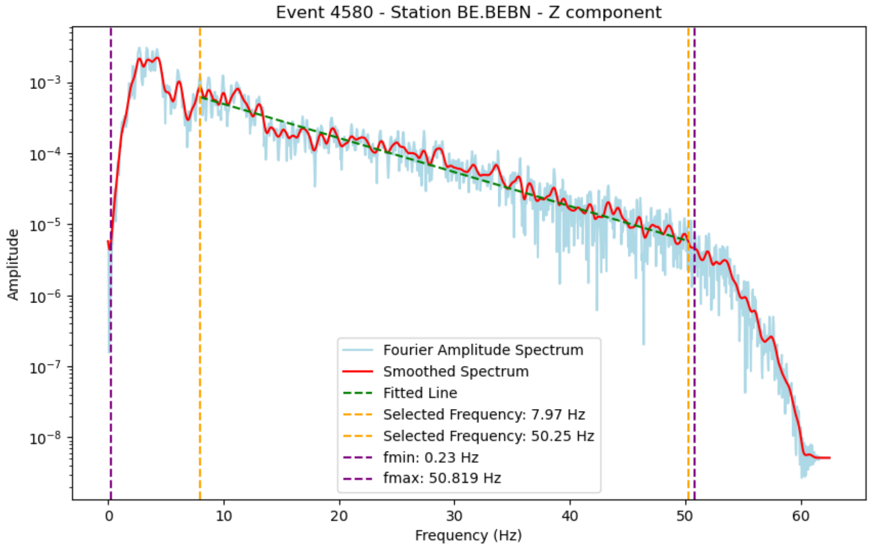


Figure 32 - κ determination for an individual record (M_L=4.3 Goch, Germany, earthquake in 2011, recorded by station BE.BEBN in ARD crustal domain). Blue curve: Fourier amplitude spectrum of the S-window for Z component. Red curve: smoothed Fourier amplitude spectrum. Green line: fitted line to the smoothed Fourier amplitude spectrum by OLS method. Dashed purple lines show the minimum (f_{min}) and maximum (f_{max}) frequency established based on minimum signal to noise ratio equal to 3. Dashed yellow lines show selected frequencies for the linear portion to compute κ

To evaluate the consistency of our κ estimates, we compared the κ values obtained for the horizontal component with equivalent t^* (t-star) values, which were determined independently as a byproduct of moment magnitude estimation from displacement spectra (see §4.1.5), in which anelastic attenuation is modeled as $exp(-\pi f t^*)$, with the attenuation parameter t^* defined as:

$$t^* = \int \frac{dr}{Q_S(r)\beta(r)} = \frac{t}{\overline{Q_S}}$$

Where $Q_S(r)$ is the spatially varying shear-wave quality factor and $\beta(r)$ is the shear-wave velocity along the ray path (Cormier 1982).

As the S-wave arrival time is generally known, this can be simplified to travel time t divided by the average quality factor over the path $\overline{Q_S}$. In principle, Q_S also varies with frequency. However, if Q_S is assumed to be frequency-independent, as is the case in the software used to determine moment magnitudes (SourceSpec, Satriano, 2025; Satriano et al., 2016), t^* becomes equivalent to the anelastic attenuation parameter κ; both parameters are expressed in seconds.

It should be noted that our t^* values are the result of an automated spectral fitting procedure, focused on the low-frequency part of the displacement spectrum, whereas our κ values are derived from dedicated, interactive fitting of the high-frequency part of the acceleration spectrum. Despite these methodological differences, comparison of κ and t^* values for each event recorded by a given station show a good agreement, and for most stations, the Pearson correlation coefficient (PCC) indicates a strong linear relationship. Figure 33 demonstrates this for station BE.BEBN, where a correlation coefficient of *PCC* = 0.75 demonstrates the reliability of the computed κ values in the database.

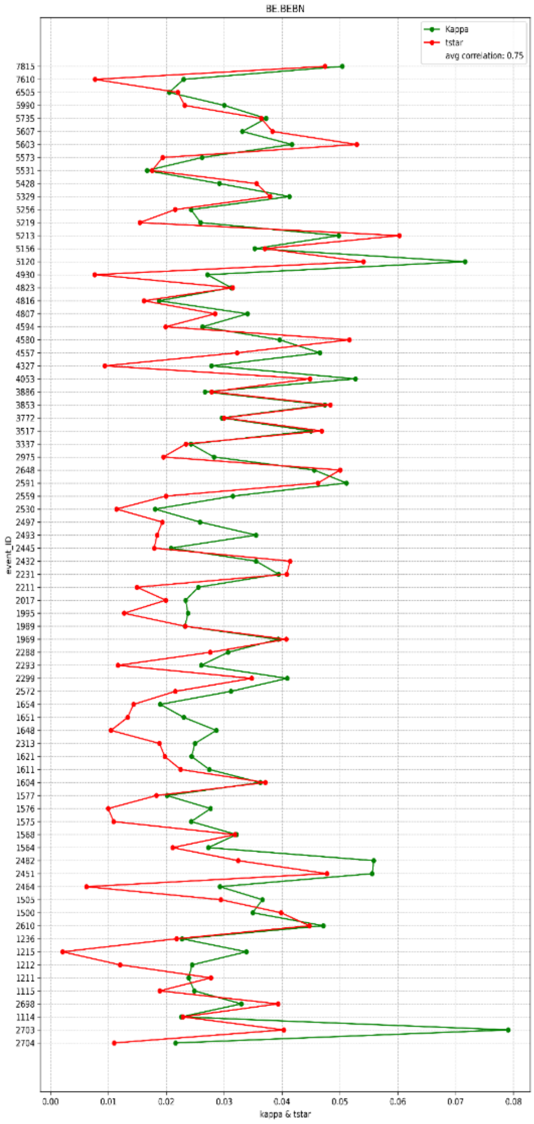


Figure 33. Correlation between horizontal component κ and t^* values for different events (vertical axis) recorded by station BE.BEBN

3.2.1.2. Estimation of site and path components of κ

In a second step, we analyze the variation of κ with distance to estimate the site-specific and path-dependent components κ_0 and κ_r. As illustrated in Figure 34 (left panel), these are determined as the intercept and slope, respectively, of a linear regression through a set of κ values obtained for a given seismic station or crustal domain as a function of distance. As κ_r represents regional attenuation, data from different stations within the same crustal domain are often pooled together to obtain a more robust estimate (Figure 34, right panel). In this study, we denote the station-specific kappa gradient as κ_{r_s} and the domain-specific kappa gradient as κ_r.

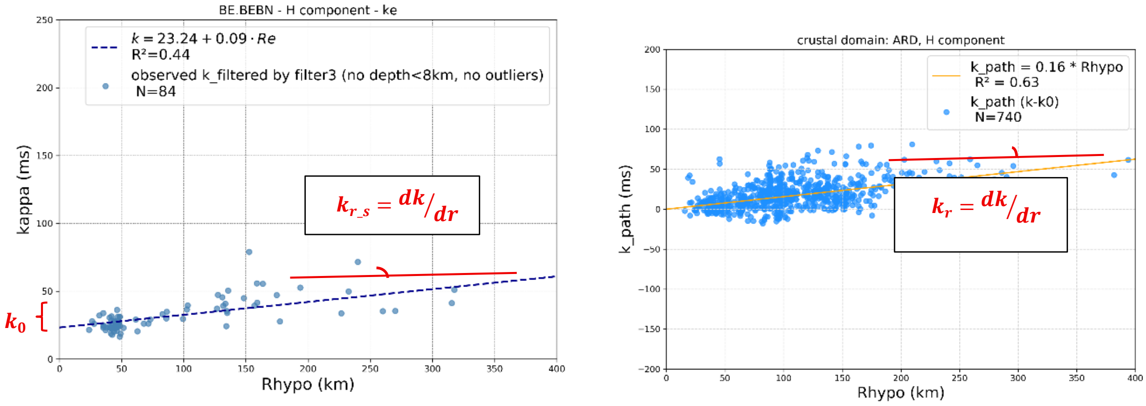


Figure 34. Left: Regression line showing the relationship between κ values and hypocentral distance for station BE.BEBN (in ARD). Extrapolating the regression line to zero hypocentral distance yields the site-specific κ_0 value (intercept), and the slope corresponds to the site-specific kappa gradient (κ_{r_s}). Right: Regression line between κpath values and hypocentral distance in the ARD crustal domain where station BE.BEBN is located. The slope (κr) is the kappa gradient of the ARD crustal domain

Different ways can be conceived to determine the regional kappa gradient in a crustal domain, along with the station-specific κ_0 values, by combining kappa data from different stations in that domain. We have conducted an in-depth analysis of three methods, which we named the free kappa gradient method, the joint kappa gradient method, and the mixed effects method. Each method has its own advantages and challenges.

The free kappa gradient (FKG) method is a relatively straightforward approach that estimates κ_0 and κ_r values for individual stations. In this method, the individual κ_0 values are subtracted from the κ values at each station, leaving the residual κ_{path}. The κ_{path} values from all stations within a crustal domain are then pooled together to derive regional estimates of the slope, kappa gradient κ_r (Figure 34, right). While computationally efficient, it is sensitive to data quality and availability, often resulting in unreliable estimates when stations have insufficient or missing data.

The joint kappa gradient (JKG) method offers a more integrated approach by applying a joint fit of the kappa gradient to the observed kappa values across all stations within a given crustal domain. It minimizes the root-mean square error between observed and predicted κ values by imposing a common κ_r to all stations within the domain, while allowing for individual κ_0 values at each station. As a result, all stations of a crustal domain have a common κ_{r_s} which is equal to the κ_r of the crustal domain. Our implementation of this approach requires predefined bounds for model parameters, which can limit its flexibility and accuracy in certain scenarios. Despite these limitations, it offers a more sophisticated and robust framework for estimating κ_r in more data-rich environments.

The mixed effects regression (MER) method adopts a statistical framework that incorporates both fixed and random effects, making it well suited for grouped datasets with uneven data distribution. In this context, the fixed or global effects represent the overall slope and intercept of the κ–distance relationship within each crustal domain. The random effects account for deviations in slope and intercept at individual stations from the global values. The sum of the global intercept and a station's random intercept corresponds to the station-specific κ_0 value, while the sum of the global slope (κ_r)

and a station's random slope represents the station-specific kappa gradient (κ_{r_s}). The general form of the mixed effects model used in this study is:

$$\kappa_{j,e} = \left(global\ intercept_i + random\ intercept_j\right) + \left(global\ slope_i + random\ slope_j\right) * \mathrm{R}_{hypoj,e} + \epsilon_j$$

where:

- $\kappa_{j,e}$: The kappa value predicted for station j in the crustal domain i and for event e.
- global intercept i (fixed effect): representing the average baseline κ_0 across all stations in the crustal domain i.
- random intercept for station j: reflecting the station-specific deviation from the global intercept i.
- global slope i (fixed effect): representing the overall gradient of kappa with respect to hypocentral distance (κ_r) across the crustal domain i.
- random slope for station j: representing the deviation in slope from the global slope specific to station j. This term is unique for each station and allows the κ_{r_s} to vary for each station.
- $\mathrm{R}_{hypoj,e}$: The hypocentral distance associated with station j and event e.
- ϵ_j: The residual error term for station j, capturing any remaining unexplained variation in kappa at that station.

This method is particularly advantageous for complex datasets where there is varying data availability across stations. It optimizes the model to account for differences among stations while also capturing the overall trend across each crustal domain. The key outputs are κ_0 and κ_r, which are obtained by minimizing the combined variance that includes both the random effects variance (station-specific deviations) and the residual error variance (unexplained variation within stations).

To better characterise and quantify the station-specific variance and associated standard deviation, we explicitly calculated the random-effects variance following the extended approach proposed by Johnson (2014) which generalises the widely used framework of Nakagawa & Schielzeth (2013) to models with both random intercepts and random slopes. Rather than approximating random-effects variance from a simplified intercept-only model, we computed it directly by incorporating the full covariance structure of the random effects and averaging over the random-effects design matrix. This approach captures not only the variance of the random intercepts and slopes but also their covariance, ensuring a more complete representation of between-station variability.

3.2.1.3. Data filtering

For each station, the relationship between κ and distance was examined, generally showing increasing κ with distance, as illustrated for the horizontal component at station BE.BEBN (Figure 35a). However, vertically aligned clusters in the κ–distance plots indicate anomalous behavior, primarily associated with induced earthquakes and shallow focal depths. A likely explanation for the strong variability of κ values for shallow events at similar distances is additional attenuation of downgoing waves near the earthquake source, where complex scattering and interaction with near-surface heterogeneities may amplify the high frequency attenuation.

To improve the robustness of the κ–distance relationship, a three-stage filtering procedure was applied. Filter 1 removed induced earthquake records. Filter 2 additionally excluded events with epicentral distances < 50 km and focal depths < 7 km, along with visual outliers. Filter 3 applied the strictest criteria, excluding all events with focal depths < 8 km and remaining outliers, regardless of distance. The effectiveness of each filter was evaluated using regression fits and visual inspection. While Filter 1 yielded limited improvement and Filter 2 reduced shallow-distance clustering, Filter 3 produced the most consistent κ estimates by eliminating nearly all vertical clusters (Figure 35.b-d). Although the strict filtering reduced the number of records at some stations, the gain in reliability outweighed the loss of data. Consequently, Filter 3 was adopted for all stations, providing a consistent and robust dataset for subsequent estimation of site κ_0 and κ gradients.

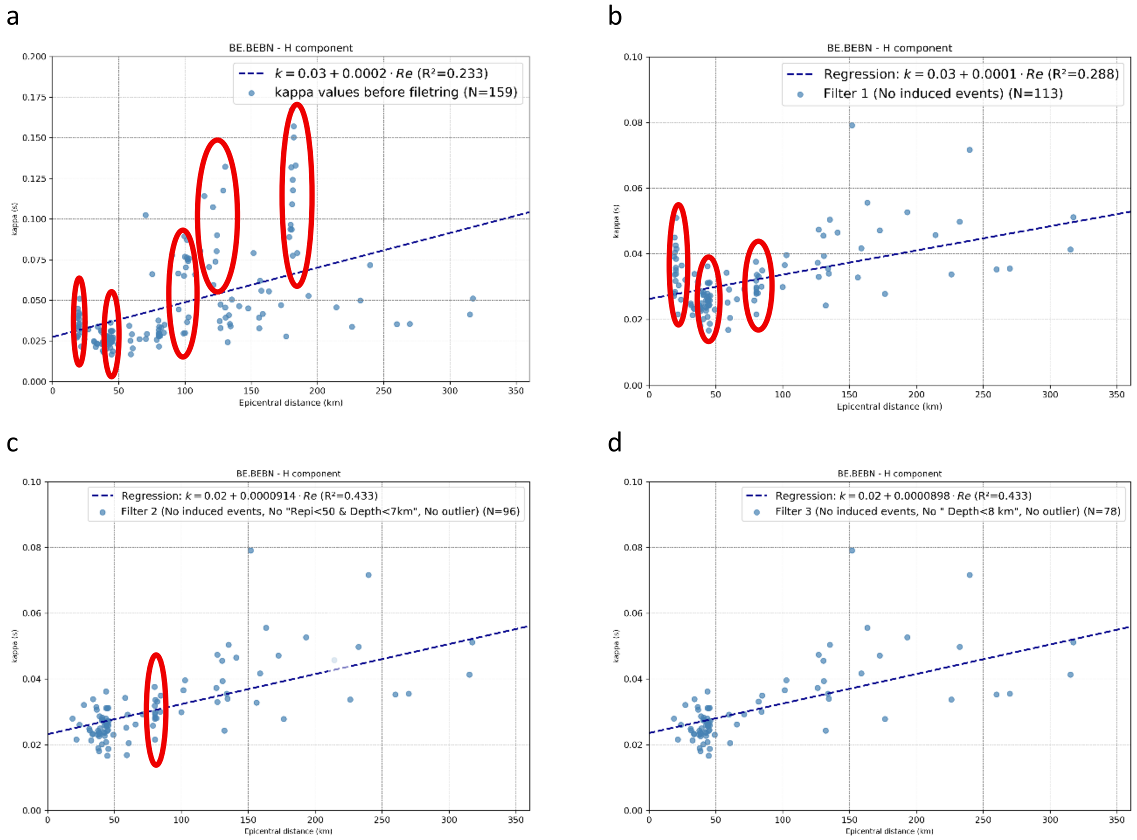


Figure 35. Comparison of κ -distance plots for the H component of station BE.BEBN, with three filtering approaches. (a) without any filter, (b) after applying filter 1, (c) after applying filter 2; (d) after applying filter 3. Some vertical alignments and outliers have been highlighted in red

3.2.2. SELECTION OF GROUND-MOTION MODELS VALID FOR BELGIUM

In order to make a better informed and up-to-date selection of ground-motion models that are appropriate for use in SHA in Belgium, we compare the ground-motion intensities (mainly response spectral accelerations, PGA and PGV) in the BELSHAKE database (i.e., “observed” ground motions) with those predicted by a number of recent candidate GMMs.

In the past two decades, several metrics have been proposed to evaluate the goodness of fit between a GMM and observed ground motion:

- The LH (likelihood) measure, proposed by Scherbaum et al. (2004), is based on the absolute standardized residuals. It is intended to quantify the model fit, but also how well the underlying statistical assumptions are met: if both the mean and variance of the observed and predicted data match, LH is evenly distributed between 0 and 1. LH can be represented as a histogram but is often reduced to the median LH, with a value of 0.5 corresponding to a perfect match, whereas lower or higher values indicate that either the model is biased or that there is a difference between model variance and sample variance.

- The LLH (log-likelihood) measure, introduced by Scherbaum et al. (2009), is based on information theory. It is calculated from the probabilities given by the GMM to predict the observations. Similar to Villani et al. (2019), we compute LLH with the base-2 logarithm and change the sign of the results to obtain positive values. Again, the LLH is represented as a histogram, but this time with higher degrees of skewness towards lower values, indicating a better match. Differences in average LLH between two models are a good estimate of the relative information loss and can be used to rank GMMs, which is an advantage over the LH measure. However, LLH doesn't solve the issue that it is not possible to distinguish between model bias and different variance of model and observations. It has also been criticised (Kale & Akkar, 2013) for focusing on models that can accurately represent the aleatory variability of the dataset, and hence for favouring GMMs with large standard deviations.
- The EDR (Euclidean-based distance ranking) measure was proposed by Kale & Akkar (2013) to remedy the shortcomings of the LH and LLH measures. It is composed of three indexes: (1) MDE (modified Euclidean distance), which represents the total probability of the differences between the observed and predicted data over a range of standard deviations, and generally lies between 0.5 and 3.2; (2) Kappa (κ), not to be confused with the high-frequency attenuation parameter with the same name (§3.2.1), which measures the level of bias between the median predictions and the overall trend (with ground-motion amplitude) in the observed data, and the optimum value of which is 1.0; and (3) EDR, which is the overall ranking based on both the normalized MDE and $\sqrt{\kappa}$, providing a balanced evaluation of both dispersion and bias.
- The Rtrend (Residual trend with distance) meausre was introduced by Vanneste et al. (2024), who noted that none of the above ranking methods considers the trend of residuals with distance, i.e. the attenuation term of the GMM. As the κ score measures the trend with increasing observed ground motion, and ground motion generally decreases with distance, this score could be considered as a proxy for the distance trend, but it does not indicate whether this trend is positive or negative. Rtrend explicitly captures the distance trend by computing the weighted average of the trends shown by ground-motion data from individual events. The weights are based on a combination of the distance covered by the observations (97.5th percentile) and the number of observation. At least 10 observations are required to contribute. An upward/downward trend indicates that the evaluated GMM attenuates faster/slower than the observed intensities. Rtrend is reported as ground-motion units per 10 km.
- The MBE (Mean Bias Error) measure is a standard statistical measure to quantify the average deviation between model predictions and observations. Unlike LLH or EDR, which incorporate both probabilistic and dispersion-related considerations, MBE directly evaluates systematic bias in the model. It is calculated as the mean of the residuals (log-scale differences between observed and predicted ground motion). Lower absolute MBE values indicate better agreement between predictions and observations, while the sign indicates overprediction (negative) or underprediction (positive). In this study, MBE is applied

specifically as the single metric to assess model performance for SA, serving as a simple yet effective indicator of bias.

MDE (and hence EDR) is hard to visualize, but we show the other measures in Figure 36 for an ideal dataset obtained by random simulation of ground motions with the GMM of Akkar et al. (2014) for a range of magnitudes, epicentral distances and focal depths. The subplots in Figure 36 show histograms of standardized residuals (related to MBE) and of LH and LLH, as well as a plot of observed versus predicted ground motions and the fitted $\sqrt{\kappa}$ value. These provide a visual means to evaluate the goodness of fit according to these different measures. Figure 37 illustrates Rtrend for a real-world case, but using macroseismic intensities as ground-motion values.

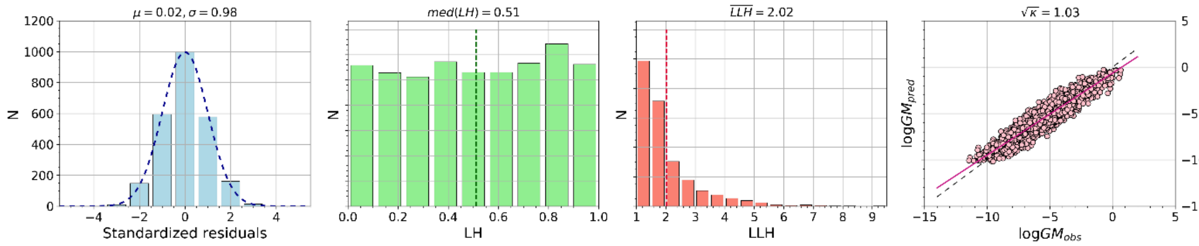


Figure 36 – Goodness-of-fit measures for an ideal dataset obtained by random simulation (2500 samples). From left to right: histogram of standardized residuals compared to the standard normal distribution; LH histogram (Scherbaum et al., 2004); LLH histogram (Scherbaum et al., 2009); $\sqrt{\kappa}$ (Kale & Akkar, 2013). See text for an explanation of the LH, LLH and Kappa measures

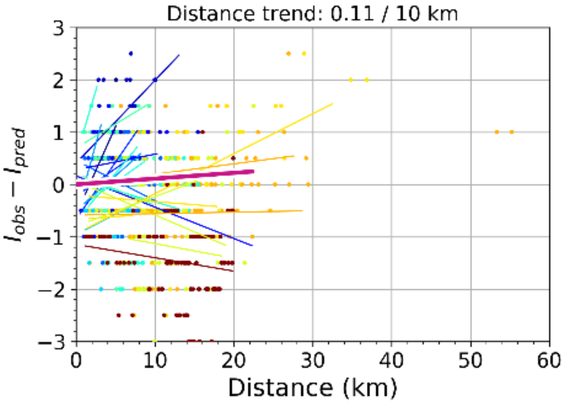


Figure 37 – Example of Rtrend for residuals between observed macroseismic intensities and an intensity prediction equation (Vanneste et al., 2024). Note that residuals have been rounded to the nearest half degree. Magenta line corresponds to the average distance trend, other coloured lines indicate the distance trend for individual earthquakes

3.2.3. ADJUSTMENT OF GROUND-MOTION MODELS

In low-to-moderate seismicity regions like Belgium, the available observations are usually restricted to smaller earthquakes with magnitudes lower than the range of interest in seismic hazard assessment. However, these data still contain information on the source, path and site characteristics of the region. By taking this information into account, regionalized models can be more accurate and more precise. Several methods have been proposed in literature to adjust existing ground-motion models using local observations from a low-seismicity region. We discuss two such methods: the “Referenced Empirical Approach” introduced by Atkinson (2008) and applied by Ramadan et al. (2024) in France, and Bayesian updating (Kotha & Traversa, 2024).

3.2.3.1. Referenced Empirical Approach

In this method, spectral adjustment factors are determined by analyzing the differences between the ground-motion dataset of a target region, in this case the BELSHAKE database, with respect to a reference ground-motion model.

The total residuals between the BELSHAKE data and GMM predictions are calculated in log space:

$$RES(Y) = \log_{10} Y_{obs} - \log_{10} Y_{pred}$$

where Y corresponds to a particular intensity measure (e.g., PGA, PGV, spectral acceleration at different periods), Y_{obs} represents the observed intensity measure (IM) values and Y_{pred} those predicted by the chosen GMM.

To analyze the dependence of the total residuals on explanatory variables, these are decomposed using mixed-effect regression (Al Atik et al., 2010; Stafford, 2014), similar to §4.1.8:

$$RES(Y) = \delta_0 + \delta c_3 R_{es} + \delta B_e + \delta S2S_s + [\delta L2L_z] + \delta W_{es}$$

Thus, total residuals are separated into median bias (δ_0) and distance-dependent ($\delta c_3 R$) terms (fixed effects), between-event (δB_e), site-to-site ($\delta S2S_s$) and locality-to-locality ($\delta L2L_z$) terms (random effects), and event-and-site-corrected or unexplained (W_{es}) residuals:

- δc_3 is a fixed-effect correction of the coefficient accounting for anelastic attenuation (commonly named c_3) with respect to event-station distance R_{es} (in the metric used by the GMM);
- δB_e represents the systematic deviation of the observed ground motion due to earthquake e from the median predicted by the GMM;
- $\delta S2S_s$ represents the systematic deviation of the observed ground motion at station s;
- $\delta L2L_z$ is an optional mixed-effect term representing the systematic bias for events originating in source zone z;
- δW_{es} is the leftover residual after systematic event and site residuals are removed.

Based on this analysis, an adjustment factor Δ can be calibrated for each spectral period, such that GMM predictions for the target region can be obtained:

$$\log_{10} Y_{target} = \log_{10} Y_{GMM} + \Delta$$

with

$$\Delta = \delta_0 + \delta c_3 R_{es} + \overline{\delta B_e} + \overline{\delta S2S_s} + [\delta L2L_z]$$

where

$\overline{\delta B_e}$ and $\overline{\delta S2S_s}$ are the average between-event and station-to-station residuals, weighted by the number of events and stations, respectively. In contrast, $\delta L2L_z$ is usually applied as a separate term for each considered source zone, thus resulting in a partially non-ergodic GMM.

The total standard deviation σ of the adjusted GMM can be expressed as:

$$\sigma = \sqrt{\tau^2 + \phi^2 + [\tau_{L2L}^2] + \sigma_0^2}$$

where τ, ϕ and τ_{L2L}refer to the standard deviations of the between-event, station-to-station and locality-to-locality residual terms, respectively; σ_0 is the standard deviation of the unexplained residuals. To calculate total σ for a single source zone, i.e. the uncertainty associated with the prediction including the zone correction, τ_{L2L} should be excluded.

3.2.3.2. Bayesian updating

Bayesian updating is a more sophisticated approach overcoming the limitations of other GMM adjustment methods. Kotha & Traversa (2024) presented a first application of this method based on the KothaEtAl2020 GMM (Kotha et al., 2020), a variation of which was adopted in the European Seismic Hazard Model 2020 (ESHM20, Danciu et al., 2024), in order to adjust the model based on the French RESIF/EPOS-France database (Buscetti et al., 2025; Traversa et al., 2020). Bayesian regression provides the full posterior distributions for each individual parameter, as well joint distributions of all mixed-effects parameters of a GMM. These joint distributions allow robust uncertainty propagation and assessing which terms are relatively better constrained, and which ranges of the dataset may require better sampling. More important for the application of updating a GMM, they can be used as informative priors instead of performing a new regression on an extended dataset (containing both the original dataset used to develop the GMM and the dataset of the target region).

The functional form of the KothaEtAl2020 GMM is:

$$\ln Y = e_1 + f_M(M_W) + f_{R,g}(M_W, R) + f_{R,\alpha}(R) + \delta L2L_z + \delta B_e + \delta S2S_s + \varepsilon$$

In this equation, the fixed-effects terms are:

- e_1 is the offset, bias or intercept of the GMM median;
- f_M represents the magnitude scaling, modeling the increase of ground motion as a function of moment magnitude M_W;
- $f_{R,g}$ represents the geometric spreading attenuation as a function of distance R (Joyner-Boore distance in this case) and M_W;
- $f_{R,\alpha}$ models the anelastic attenuation as a function of distance R. This term contains a coefficient c_3, representing the pan-European average apparent anelastic attenuation at far-source distances estimated from the ESM database. This parameter is regionalized by adding a region-specific $\delta c_{3,r}$ factor, which can be estimated from the regional dataset. However, one could also consider attenuation variability around the source by adding a separate $\delta c_{3,e}$ factor (Georges, 2025).

The random-effects terms δB_e, $\delta S2S_s$ and $\delta L2L_z$ model the natural source variability, the variability between stations due to site effects, and regional source variability, as explained earlier (see §3.2.3.1).

The Bayesian approach allows previous models to be used as informative priors and simply updating the existing model using new data. In this case, one needs to examine carefully which coefficients can be updated and which coefficients should be kept fixed, as shown in Kotha & Traversa (2024).

More recently, Georges (2025) performed a Bayesian regression using the complete dataset from which the original GMM was derived, extended with regional data, along with non-informative priors. In that study, regional adjustments to the global model were made in 0.5 by 0.5 degree cells in a grid covering Europe, thus deriving a partially non-ergodic GMM, which improves the uncertainty and precision of the global model in each region (grid cell) where there is enough data.

3.2.4. SIMULATION OF GROUND-MOTION FROM HIGH-MAGNITUDE EVENTS

As the BELSHAKE database contains only few events with magnitude above 4, there is a need to extend the database with ground-motion data from higher-magnitude events if we want to develop ground-motion models covering the magnitude range used in seismic hazard assessments. There are two ways to do this: 1) include data from other, “comparable” regions in Europe or elsewhere that are available in other ground-motion databases; 2) generate synthetic ground-motion records for hypothetical higher-magnitude earthquakes. Different methods are available for the second option: physics-based simulation (PBS), Empirical Greens’ functions (EGF) and stochastic simulation (EXSIM). As PBS is computationally very expensive and is not able to simulate high frequencies due to the limited resolution of geological/velocity models, it will not be considered in this project. Both EGF and EXSIM are based on discretization of an extended fault rupture into a number of subfaults, followed by time-delayed summation of accelerograms associated with each subfault (Figure 38). In EGF, the subfault accelerograms correspond to a recorded aftershock, while in EXSIM, they are stochastically simulated (cf. Boore, 2003) for each subfault with a different “dynamic” corner frequency, which depends on the number of simultaneously rupturing subfaults (called the “pulsing fraction”). An advantage of EGF is that the aftershock record already includes the path and site attenuation effects, while these have to be estimated in EXSIM. On the other hand, EXSIM can easily account for random variation because accelerograms are simulated for each fault separately, while EGF only uses a single aftershock record.

We first reimplemented the EGF-method that was applied by Gariel et al. (1994) to generate acceleration records from the 1992 Roermond mainshock based on accelerograms from aftershocks recorded by a temporary network in its epicentral area (see §3.1.2). As the precise scaling method was insufficiently documented, we continued to investigate the scaling factors applied in EXSIM, a well-known fortran program to simulate ground motion generated by extended fault sources developed more than a decade later, most recently by Boore (2009). As we already implemented stochastic simulation for point sources prior to the project, it was possible to develop new code implementing both EGF and EXSIM. Based on this exercise, and by exploring the differences between both methods, we realized that it is possible to correct the aftershock record used as subfault accelerogram in EGF to make it similar to the simulated subfault accelerogram(s) generated in EXSIM. This correction involves:

- constant scaling by the ratio of target (mainshock) slip and aftershock slip;
- scaling by the spectral shape ratio between the aftershock and the subfault records modelled in EXSIM. This requires knowledge of the aftershock corner frequency and has to be done for each subfault separately, because the dynamic corner frequency is different for each;
- in addition, the Fourier phase of the aftershock record can be randomized for each subfault, to obtain different (but equivalent) accelerograms for each subfault, as suggested by Ben Edwards (pers. comm. 2025).

The effect of the two first correction factors is illustrated in Figure 39. When these corrections are applied, we should obtain subfault accelerograms that are comparable to EXSIM, following which the same scaling factors as in EXSIM can be applied before the time-delayed summation.

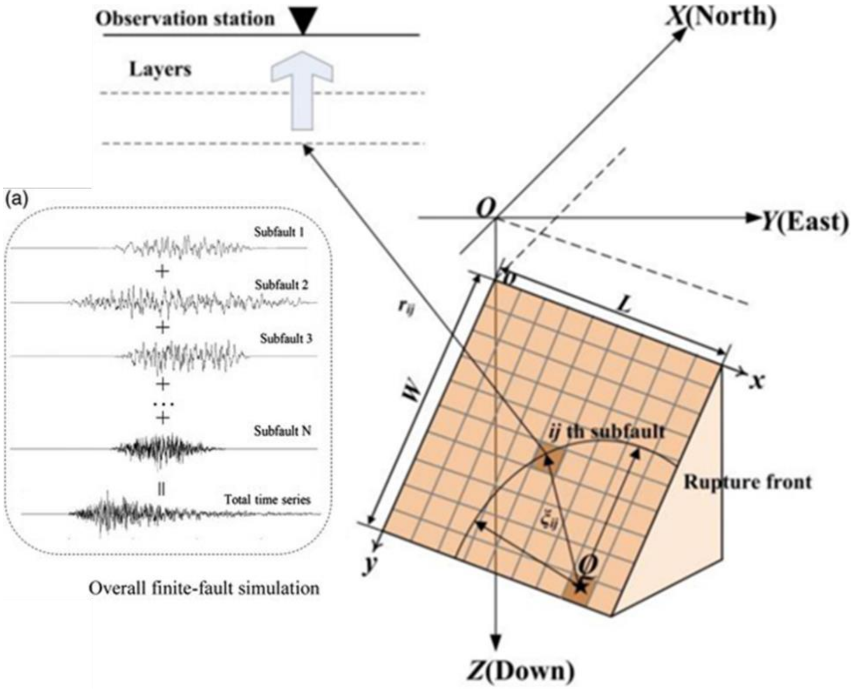


Figure 38 – Diagram showing principle of time-delayed summation of subfault accelerograms

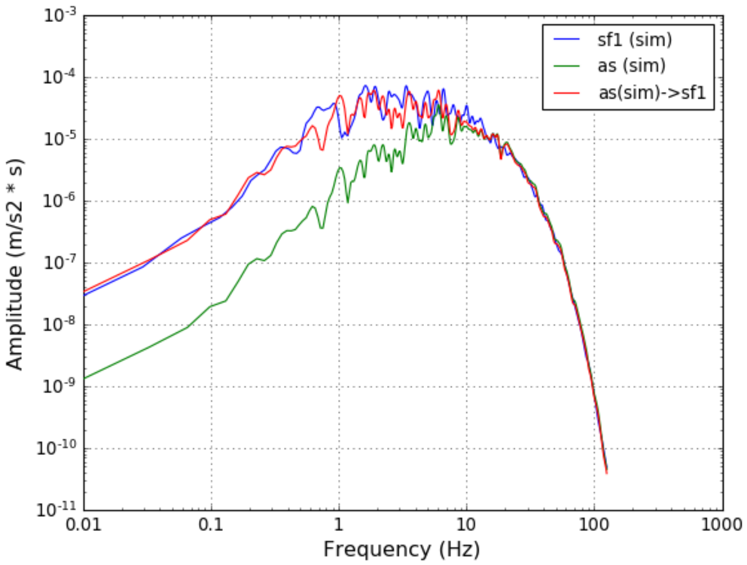


Figure 39 - Illustration of spectral scaling to be applied to an aftershock accelerogram in order to make it similar to the accelerogram modeled for a particular subfault in EXSIM. The green line shows the spectrum of a (simulated in this case) aftershock, which we want to make similar to the spectrum of the subfault accelerogram shown in blue. The red line depicts the scaled aftershock spectrum, which is very similar to the target spectrum (apart from random variability)

4. SCIENTIFIC RESULTS AND RECOMMENDATIONS

4.1. The BELSHAKE database

The most important result of the BELSHAKE project is the development of an earthquake ground-motion database for Belgium and surrounding regions, which we refer to as the BELSHAKE database. This database contains digital records from tectonic and induced earthquakes with $M_L \geq 2$, recorded since 1985 with short-period, broadband and accelerometric sensors operated by the Royal Observatory of Belgium.

4.1.1. DATABASE STRUCTURE

The database consists of two main components:

- a waveform archive containing raw and processed waveform data and the associated instrument response information for each event-station pair;
- a relational database containing the associated metadata as well as the derived ground-motion parameters.

The waveform library is structured in folders corresponding to each event, containing one file for each station. We opted for the flexible ASDF format (Krischer et al., 2016), which can contain any number of waveforms in a single file, along with metadata, instrument response information and auxiliary (non-waveform) data, and is based on existing formats and standards. In our case, we use it to store waveform data, including intermediate processing steps (accessible through different "tags"), along with the instrument response, relevant event and station metadata, and corresponding Fourier and response spectra in a single file for each event-station pair.

The relational SQL database consists of seven main tables:

- event metadata : origin and source parameters;
- station metadata: physical properties of a seismic station that do not change with time, such as the location, housing, emplacement and local site conditions;
- event-station metadata: distances according to different metrics, (back-)azimuth, P- and S-wave arrival times and incidence angles;
- waveform metadata that are common to all components : instrument information, start and end time of different signal windows, identified problems;
- component-specific waveform metadata: start and end time, sample rate, reliable frequency range, problems;
- processing parameters;
- calculated ground-motion intensity measures.

We also developed a dedicated Python library (based on ObsPy, Megies et al., 2011), which provides a programming interface with the database and waveform archive, allows reading and writing of metadata, and streamlines data import.

4.1.2. DATA SELECTION

The BELSHAKE database currently contains ~6400 digital seismograms, ~5840 from 327 tectonic earthquakes and ~580 from 17 induced earthquakes. The location of these events is shown in Figure 40. Table lists the most notable events (with $M_L \geq 4.0$) in the database. The largest and hence most important event is the $M_W = 5.3$ Roermond, The Netherlands, earthquake on 13 April 1992, which caused light damage corresponding to intensity VII on the European Macroseismic Scale EMS98

(Haak et al., 1994) in the border region between The Netherlands, Belgium and Germany. Unfortunately, most records from this event were clipped, because seismometers at the time had limited resolution and were tuned to detect small earthquakes. Furthermore, there were no accelerometers operational at the time. We therefore collected additional waveform data from networks in France, Germany, the United Kingdom and Czech Republic for this particular event and its aftershocks. In total, we imported 221 records from 10 different networks in our database.

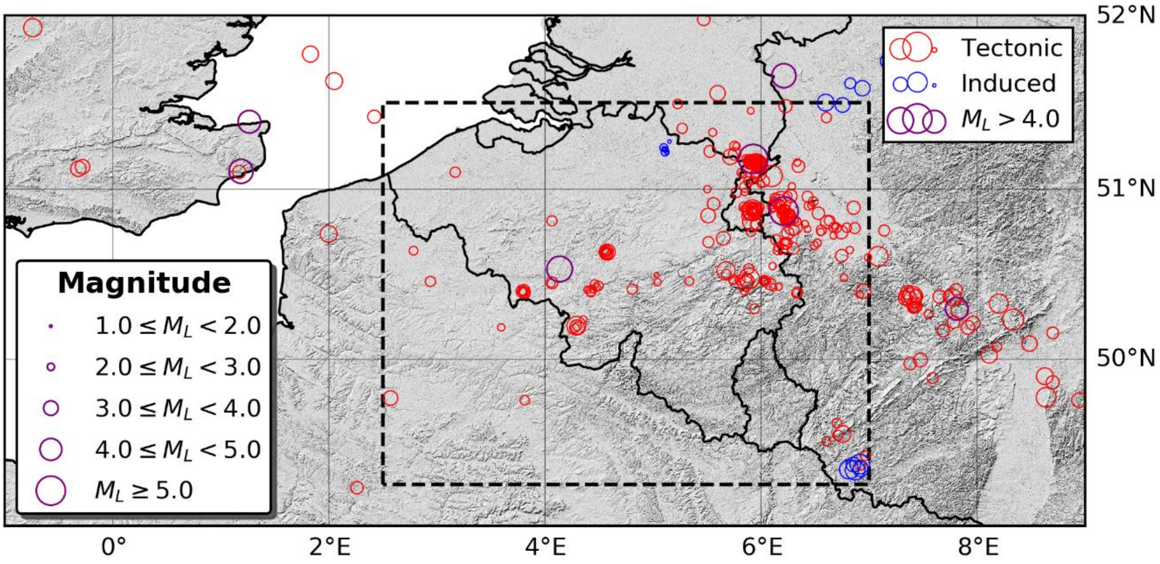


Figure 40 - Map showing epicenters of earthquakes included in the BELSHAKE database. Dashed rectangle corresponds to narrow selection area with lower magnitude threshold ($M_L \geq 2.0$)

Date	Location	Lon (°)	Lat (°)	Depth (km)	M_L	M_W	Imax	Num. records
1992-04-13	Roermond (NL)	5.9278	51.1758	16.4	5.8	5.3	VII [1]	221
1995-06-20	Le Roeulx (BE)	4.1359	50.5330	26.2	4.5		V [2]	22
2002-07-22	Alsdorf (DE)	6.2113	50.8731	13.4	4.9	4.6	VI [2]	25
2007-04-28	Folkestone (UK)	1.1870	51.1040	8.8	4.3		VI [3]	28
2011-02-14	Koblenz (DE)	7.8112	50.2992	13.3	4.1		V [4]	29
2011-09-08	Goch (DE)	6.2116	51.6557	17.7	4.3		IV [5]	31
2015-05-22	Ramsgate (UK)	1.2639	51.3910	14.5	4.1		V [6]	36

Table III – Origin time, hypocentral location, magnitude, maximum macroseismic intensity (Imax) and number of records for main events in the BELSHAKE database. Imax sources: 1 (Haak et al., 1994); 2 (Neefs et al., 2025); 3 (Sargeant et al., 2008); 4 (Hinzen, 2019); 5 (Van Noten et al., 2017); 6 (Baptie, 2017)

4.1.3. DATA QUALITY

As described in §3.1.3, all waveform data have been visually inspected to identify, and if possible correct, problematic records. The identified problems, and whether or not these have been corrected, are stored in the database, such that they can be queried for any component and signal window. Based on these problem categories and solutions, we defined three different quality classes, and developed a method in our programming interface to select only records having a specific minimum quality, at the level of component and window. These quality classes are:

- Good: windows that have no problems, are only slightly curtailed, or with spikes that have been successfully removed or other disturbances that have been successfully trimmed or filtered;
- Intermediate: windows that are noisy or have uncorrected 50 Hz hum;
- Bad: windows that are very noisy, clipped, missing, incomplete, disturbed, interrupted, have no waveform or contain spikes, distortions, fore- or aftershocks that could not be removed.

Table lists the number of waveform records in each quality category, as well as the number of records with unreliable instrument response, for different event types and components.

Event type	Comp.	Good	Intermediate	Bad	Unreliable	Total
tectonic	Z	3816	869	1156	574	6415
tectonic	H	5894	1316	1776	422	9408
induced	Z	300	89	190	16	595
induced	H	534	158	336	30	1058
all	Z	4116	958	1346	590	7010
all	H	6428	1474	2112	452	10466

Table IV – Number of records in the BELSHAKE database for different event types, components (H=horizontal; Z=vertical) and quality categories

4.1.4. DATA COVERAGE

Figure 3 (left panel) illustrates the magnitude-distance coverage of the subset of the BELSHAKE database having horizontal components and acceptable quality of the S-window. The magnitude coverage is relatively good up to $M_L = 4.0$, but significantly decreases towards higher magnitudes. Hypocentral distances range mainly between 20 and 300 km, but there are also several records at short distance (< 10 km) for lower-magnitude ($M_L < 3.0$) events. Compared to the French RESIF/EPOS-France database (Buscetti et al., 2025; Traversa et al., 2020) in Figure 3 (right panel), the magnitude coverage is sparser in the range $M_L = 4.0 - 5.8$, but the coverage at short distances appears to be better. This is also illustrated by the histograms in Figure 42. The first histogram (panel A) shows the number of events as function of magnitude, indicating that most of the selected events have $M_L < 2.5$ and that there are only few events with $M_L > 4.0$. The second histogram (panel B) shows the number of records as function of hypocentral distance, indicating that there are only a few records with hypocentral distance less than 10 km and most of the records have a hypocentral distance between 50 and 150 km. Panel C shows that many events have been recorded by 10 or more stations, while panel D shows that most stations have recorded less than 25 events, but a few stations have recorded more than 100 events.

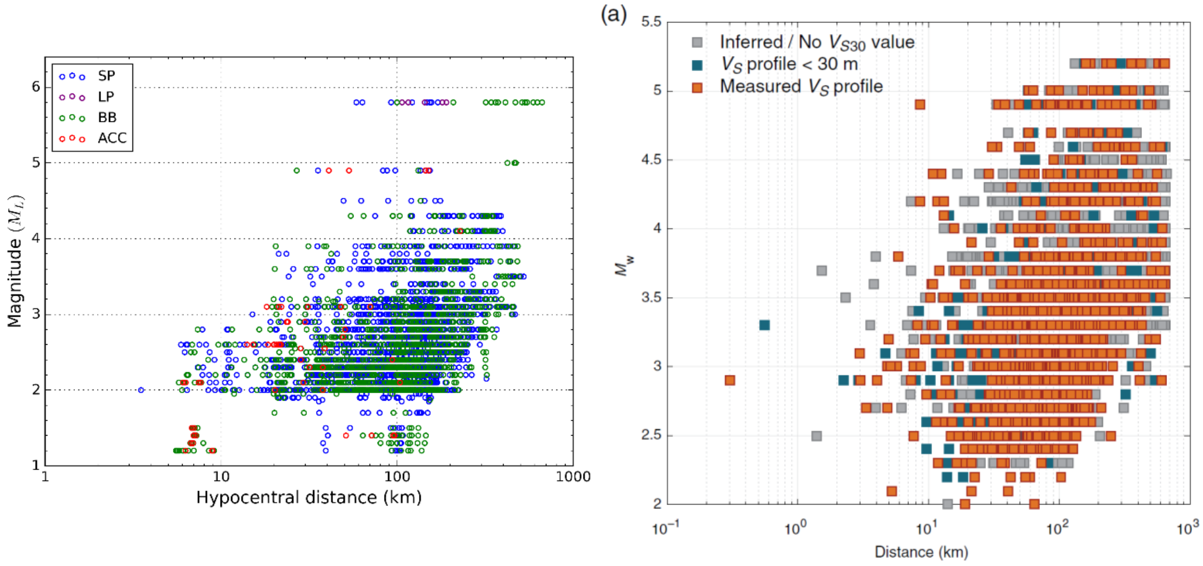


Figure 41 – Magnitude-distance coverage of records with horizontal components and acceptable quality of the S-window in the BELSHAKE database (left panel) compared to the RESIF/EPOS-France database (right panel; Buscetti et al., 2025; Traversa et al., 2020). Colors of the BELSHAKE dataset correspond to instrument types: short-period seismometer (SP), long-period seismometer (LP), broadband seismometer (BB) and accelerometer (ACC)

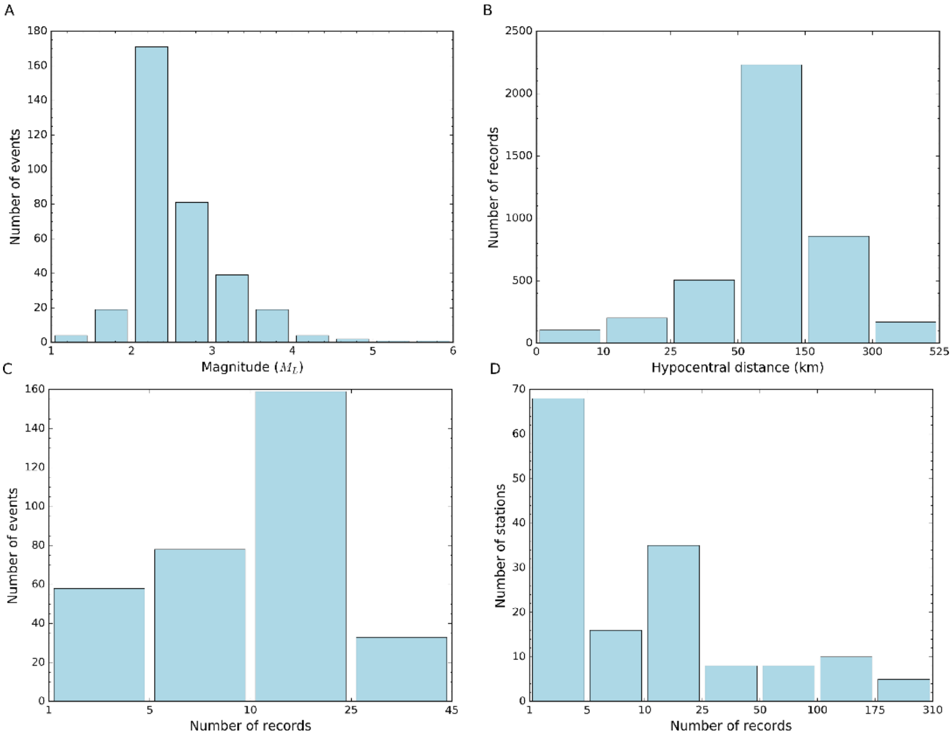


Figure 42 - Histograms illustrating the coverage of the BELSHAKE database: A) Magnitude histogram; B) Hypocentral distance histogram; C) Histogram of number of records per event; D) Histogram of number of records per station. Note that horizontal axes of histograms are not linear

4.1.5. CALCULATION OF MOMENT MAGNITUDES

Moment magnitudes (M_W) are only available for a few events in the ROB earthquake catalog, from which earthquake source parameters were imported in BELSHAKE. As it is essential to have reliable M_W values to compare with other ground-motion databases and GMMs, we calculated those with the Python package SourceSpec (Satriano, 2025; Satriano et al., 2016), from displacement spectra derived from the BELSHAKE waveform data. . For each earthquake, we performed up to five computations, depending on the available components and information on the focal mechanism:

- 2 wave types: P-wave (on the Z component) and SH-wave (on the T component);
- 2 radiation patterns: average radiation factor (0.42 for P-waves and 0.59 for S-waves, cf. Kaneko & Shearer, 2014) and focal mechanism (if available);
- 2 geometrical spreading functions: simple $1/R$ and a segmented powerlaw function derived from a recalibration of the local magnitude scale (Vanneste & Camelbeeck, n.d.). The latter was only applied for the SH-wave and average radiation pattern.

All results are stored in two database tables, one containing the parameters determined for individual stations and the other the averaged solutions for each event. Comparison of these different computations shows that results based on the P-wave and on the SH-wave are very consistent for average radiation pattern and simple $1/R$ geometrical spreading (Figure 7, left panel): the overall difference is small ($\Delta M_{W_{P-SH}} = 0.10$), the variability is quite low ($\sigma = 0.10$), and there are practically no outliers. For the SH-wave, the difference between simple $1/R$ geometrical spreading and the segmented spreading function is also very small (Figure 7, right panel): ($M_W(1/R) - M_W(segm) = 0.03 \pm 0.05$), but the segmented model yields a lower mean station variability ($\overline{\sigma_{segm}} = 0.136$) than the simple model ($\overline{\sigma_{1/R}} = 0.149$), indicating a better fit. The differences between values based on focal mechanisms and those based on average radiation pattern (Figure 44) are small when P-waves are used ($\Delta M_{W_{fm-avg}} = 0.10$), but significantly higher for the SH-wave ($M_{W_{fm-avg}} = 0.21$). In both cases, the station variability is significantly higher when using focal mechanisms ($\overline{\sigma_{P,fm}} = 0.244, \overline{\sigma_{SH,fm}} = 0.275$) than when using the average radiation factor ($\overline{\sigma_{P,avg}} = 0.181, \overline{\sigma_{SH,avg}} = 0.174$), which indicates that solutions based on focal mechanisms are less reliable.

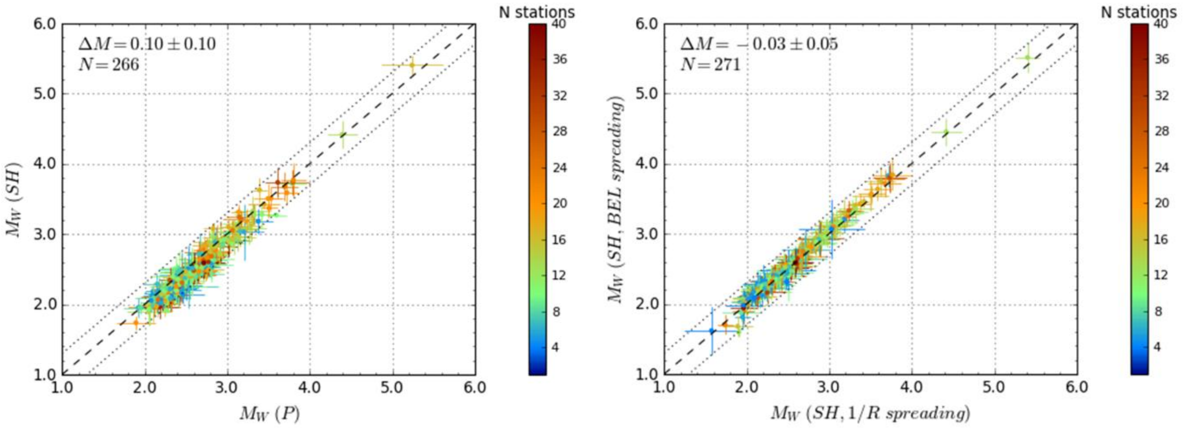


Figure 43 - Comparison of different methods to compute M_W. Left: comparison of M_W values based on the P-wave and on the SH-wave; Right: comparison of M_W values based on simple $1/R$ geometrical spreading and based on newly derived segmented spreading model (BEL), using SH waves in both cases. Note that only earthquakes with at least 4 contributing stations are included. Horizontal and vertical bars indicate standard deviations and symbols are colored according to number of contributing stations

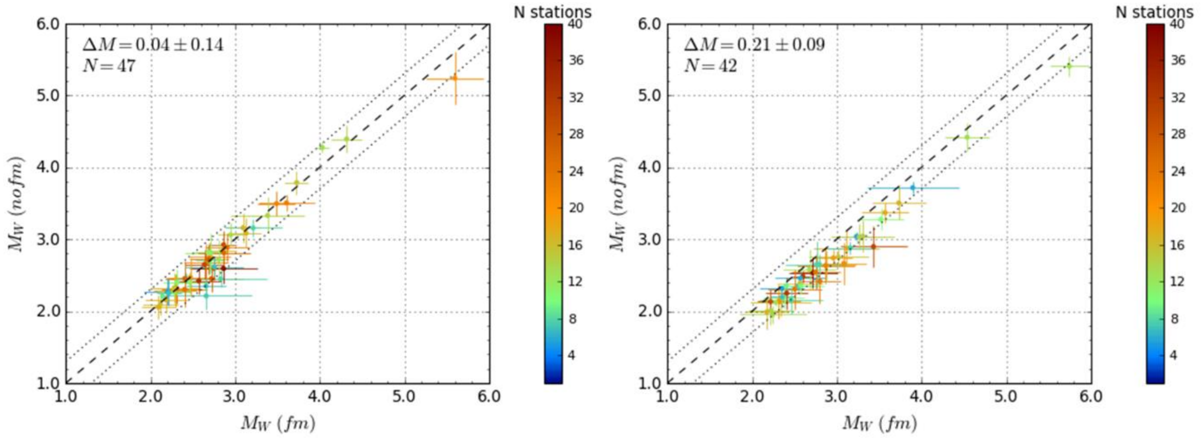


Figure 44 - Comparison of M_W values obtained using focal mechanism (fm, X axis) and using average radiation factor (nofm, Y axis). Left: P-wave solutions; Right: SH-wave solutions. Legend as in Figure 44

Based on this analysis, we calculate final M_W values for BELSHAKE events as the average (weighted by the number of contributing stations) of the P-wave solution based on simple $1/R$ geometrical spreading and the SH-wave solution based on the segmented spreading model, in both cases using the average radiation factor only and requiring at least 4 contributing stations. This resulted in direct M_W values for 294 events. The use of both P- and S-waves should minimize any bias related to the radiation pattern. The median and 95th percentile standard deviations are 0.15 and 0.27, respectively, indicating that these M_W values are well constrained. Our average M_W values are also in good agreement with those in the catalogs of neighboring agencies, such as the HRF2020 catalog (110 common events) of the Bensberg network for the northern Rhine area (Hinzen et al., 2020) and FCAT-17 (17 common events) in France (Manchuel et al., 2018), which have been computed from different waveform data. The comparison with the HRF-2020 catalog is shown in Figure 45 (left panel). The overall difference ($M_W(BELSHAKE) - M_W(HRF2020) = 0.05$) is small, the standard deviation ($\sigma = 0.12$) is also relatively small, and there is only one minor outlier: a small earthquake near Heerlen (The Netherlands) on 20 Dec. 2000, for which we obtain $M_W = 2.65 \pm 0.14$ (based on 11 stations) compared to $M_W = 2.2$ in HRF2020. We did not find an obvious explanation for this outlier, but it is consistent with the M_L value. Thus, moment magnitudes in the BELSHAKE database are consistent both internally and with respect to independent catalogs.

The M_W computations also allowed establishing a new empirical relation with the Belgian M_L scale (Vanneste & Camelbeeck, n.d.). Figure 45 (right panel) shows a density plot compared to two fitted relations: 1) a quadratic (second-order polynomial) relation with the same form as a relation based on the HRF2020 catalog (Hinzen et al., 2020), but in which the quadratic term is almost zero, and 2) a linear fit of the form $M_W = {}^2/_3\, M_L + C$, which is a theoretical relation based on random vibration theory that should be valid for small earthquakes with $M_W < \sim 4$ (Munafò et al., 2016). The difference between both is small, indicating that our results are in good agreement with theoretical considerations. This also demonstrates that these empirical relations are robust and can be applied to earthquakes for which we have determined M_L (using the same method and scale!), as long as M_L is not higher than ~4.5. Thus, we have used the quadratic fit to estimate M_W for all earthquakes in the BELSHAKE database with insufficient number of stations to determine M_W directly.

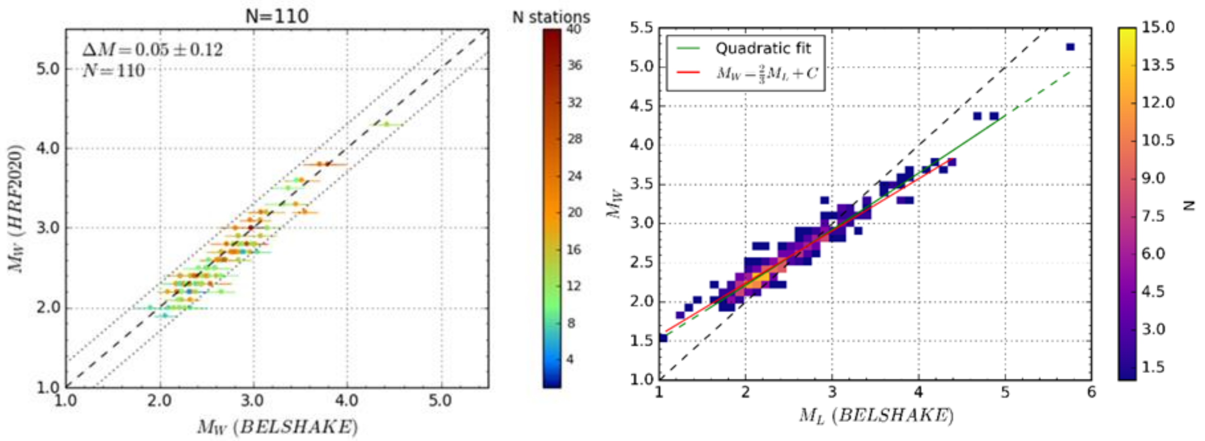


Figure 45 – Left : Comparison of average M_W values determined from the BELSHAKE database and M_W values in the HRF2020 catalog. Legend as in Figure 44. Right : Density plot of local magnitudes versus moment magnitudes in the BELSHAKE database plot, compared to two fitted relations

4.1.6. SITE CHARACTERIZATION OF BELGIAN STATIONS

Station metadata are an important aspect of ground-motion databases, e.g. to link ground motion to site conditions. However, collecting this information is not so straightforward, because it has not been systematically maintained in the ROB seismic routine database and many stations have been discontinued since their initial deployment. We tried to complete this information as much as possible for stations of the Belgian seismic network, including: Eurocode 8 (EC8) soil class (CEN (Comité Européen de Normalisation), 2004), time-averaged shear-wave velocity for the top 30 m (V_{S30}), depth to bedrock (h), natural resonance frequency (f_0), lithology, stratigraphic formation and geological age. For the geotechnical parameters, we included a field to describe the method that was used to determine them.

For this purpose, we could benefit from the results of the EPOS-BE project “Upgrade of Belgian federal research infrastructure for EPOS” (Bruyninx et al., 2023), in which ten key Belgian stations have been characterized using array measurements. For 21 other stations, quantitative information was available from well logging or geophysical investigations, but most of the current and all discontinued stations lacked specific characterization. For those stations, we had to fall back on (1) geological correlation with measured stations, (2) estimation based on other parameters, e.g. f_0 converted to bedrock depth h following powerlaw relations designed for specific regions (Van Noten et al., 2022), (3) Rayleigh-wave single-station methods (RayDec; Hobiger et al., 2009) or (4) average values associated with Eurocode 8 soil classes. Table lists the different methods used to determine V_{S30}, which is the most important site parameter for ground-motion modelling. For more than half of the stations (68 out of 129), representing 3147 out of 5936 ground-motion records , V_{S30} is based on the Eurocode 8 (EC8) soil class, which is the least reliable method.

We also compiled information on the installation conditions following the classification in the ESM database (Luzi et al., 2016), assigning 3 codes to each station (whenever possible): housing code (indicating the type of structure the instrument is situated in), installation code (related to the contact surface on which the instrument is installed) and proximity code (indicating proximity to structures). These factors may influence recorded motion compared to free-field conditions.

V_{S30} method	**Number of stations**	**Number of records**
Array measurements	10	1327
Well logging	11	444
Seismic refraction (Camelbeeck, 1993)	5	245
SASW (Camelbeeck, 1993)	3	166
Active seismics using hammer shots	2	181
RayDec	1	55
Estimated from f_0 and h	16	138
Geological correlation with measured station	8	197
Extrapolation based on geological model (Stafleu et al., 2024)	4	30
EC8 soil class	69	3153
TOTAL	129	5936

Table V - Current distribution of methods used to determine V_{S30} for stations of the Belgian seismic network in the BELSHAKE database

The EC8 soil class of stations of the Belgian network (having ≥ 10 records) is shown in Figure 46, superposed on the Belgian geological outcrop map. An interesting result of our analysis so far is that the mean and median V_{S30} for stations situated on rock (soil class A), not including soil-class-derived values, are 1440 and 1100 m/s, respectively. This is significantly higher than the value of 800 m/s that is generally considered for soil class A. Accordingly, we attribute the median V_{S30} value to stations for which no direct information is available and that are situated on soil class A.

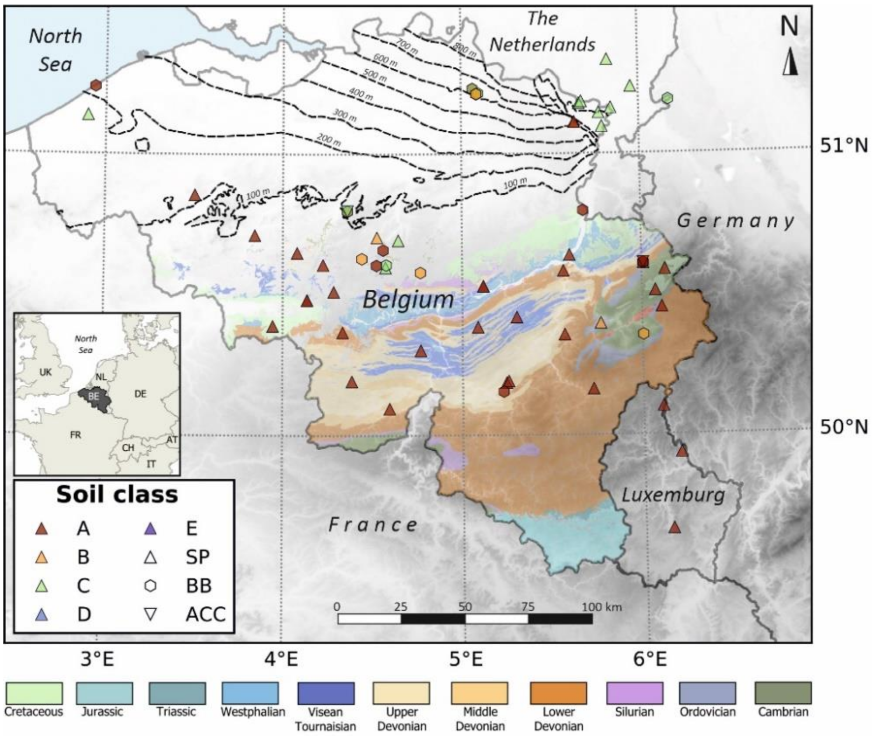


Figure 46 – Location of seismic stations with ≥10 records in the BELSHAKE database. Symbol colors correspond to Eurocode 8 soil class (A-E) and symbol shapes to instrument type (SP=short-period; BB=broadband; ACC=accelerometer). Background image depicts simplified surface geology, with bedrock in the south and NNE-ward thickening sediment layers (100 m contour interval) in the north. Note that some of these stations were deployed in temporary networks, and some stations of the permanent network have been discontinued

4.1.7. PROCESSING AND CALCULATION OF INTENSITY MEASURES

Following visual quality control, all records in the BELSHAKE database with reliable instrument response were processed using a uniform, semi-automated workflow, which is described in more detail in §3.1.6. This is necessary to obtain correct acceleration time series, from which intensity measures (IMs) can be calculated. Various IMs have been defined over time in literature. Their purpose is to provide a simplified estimate of the damage potential of earthquake ground motion. This is more relevant than just earthquake magnitude. Damage depends on many factors, including the source (magnitude, stress drop, directivity), path (source-to-site distance, crustal structure) and site (soil properties) properties, as well as the frequency content of the radiated waves and the natural period of the affected construction. Table lists and describes all intensity measures that we calculated for records in the BELSHAKE database. The most widely used IMs are peak ground motions and response spectra. The other IMs can be categorized as integrals over time series, integrals over part of the response spectrum, averages over part of the response spectrum and durations. In total, 109 IMs have been calculated, including 92 response spectral periods between 0.01 and 4 s.

Category	Code	Intensity measure	Formula	Unit
Peak ground motion	PGA	Peak Ground Acceleration	$max(\|a(t)\|)$	cm/s^2
	PGV	Peak Ground Velocity	$max(\|v(t)\|)$	cm/s
	PGD	Peak Ground Displacement	$max(\|d(t)\|)$	cm
Response spectra	PSA	Acceleration response spectrum	0.01 – 4 s / 0.25 – 100 Hz, 5% damping	cm/s^2
	PSV	Pseudo-spectral velocity	$PSA * (2\pi f)^{-1}$	cm/s
	SD	Spectral displacement	$PSA * (2\pi f)^{-2}$	cm
Integrals over time series	RMSa	Root-Mean-Square Acceleration	$\sqrt{{}^{1}/_{T_d} \int a(t)^2 \, dt}$	cm/s^2
	AI	Arias Intensity	$(\pi/2g) \int^{T_d} a(t)^2 \, dt$	m/s
	CAV	Cumulative Absolute Velocity	$\int \|a(t)\| \, dt$	$g.s$
	sCAV	Standardized CAV	CAV of 1-s windows where PGA ≥ 0.025 g	$g.s$
	bCAV	Bracketed CAV	CAV of portion of $a(t)$ ≥ 0.05 g	$g.s$
Integrals over part of response spectrum	HI	Housner Intensity	$\int_{T=0.1}^{2.5} PSV(T) \, dT$	cm
	ASI	Acceleration Spectral Intensity	$\int_{T=0.1}^{0.5} PSA(T) \, dT$	cm/s
Average over part of response spectrum	EPA	Effective Peak Acceleration	$\overline{PSA[0.1 \leq T \leq 0.5]}/2.5$	cm/s^2
	EPV	Effective Peak Velocity	$\overline{PSV[0.7 \leq T \leq 2]}/2.5$	cm/s
	EPD	Effective Peak Displacement	$\overline{SD[2.5 \leq T \leq 4.0]}/2.5$	cm

Category	Code	Intensity measure	Formula	Unit
Duration	D5_75	Significant (Arias) duration	T corresponding to buildup from 5% to 75% of cumulative AI	s
	D5_95	Significant (Arias) duration	T corresponding to buildup from 5% to 95% of cumulative AI	s
	D5_95’	Effective duration	2xT corresponding to buildup from 20% to 80% of cumulative AI (Boore & Thompson, 2014)	s
	Db5PcG	Bracketed duration	T of portion between first and last occurrence of $a(t) \geq 0.05$ g	s

Table VI - Definition of intensity measures included in the BELSHAKE database

All IMs were computed for different component and window combinations. In addition to the physical Z (vertical), E (east) and N (north) components, we also computed IMs for the T (transverse) component (equivalent to SH-wave motion), the geometric mean (GM) of the horizontal components, and the orientation-independent RotD50 and RotD100, corresponding to the 50th and 100th percentiles, respectively, over all rotation angles of the horizontal components (Boore, 2010). Most ground-motion databases provide IMs for the entire record, but this is not always available due to problems in the P or coda windows. We therefore computed IMs for different windows: P, S, signal (P + S) and full (P + S + coda). As most damage caused by earthquakes is related to the S/Lg wave, we also define an “auto” window, which corresponds to either the full, the signal or the S window, whichever is available in the given order of preference. Considering all combinations of components and windows, there can be up to 3052 IM values (109 IMs x 7 components x 4 windows) for each record in the database.

We did not compute IMs for components/windows with bad quality. We also did not compute PGA and PGV for records with low sample rate (< 20 Hz). Furthermore, response spectra and derived IMs (HI, ASI, EPA, EPV and EPD) were not computed for:

- records for which the full instrument correction was not available (only the total sensitivity);
- records which have been manually corrected;
- components/windows with intermediate quality.

In total, the BELSHAKE database contains IMs for more than 120,000 record/component/window combinations. It is impossible to show all calculated IMs in a few plots, but as a representative example, Figure 5 shows the distribution of PGA values (RotD50 component, “auto” window) as a function of distance and magnitude. In a first approximation, these results appear quite comparable with the RESIF/EPOS-France database (Buscetti et al., 2025; Traversa et al., 2020), showing a consistent relation with respect to distance and magnitude.

Similar to the RESIF/EPOS-France database, we limit the range over which response spectra are computed by considering an inward margin of 20% with respect to the reliable frequency range of the FAS. As a result, the number of records contributing to each response spectral frequency is different (Figure 48). At low frequencies, coverage quickly drops below ~0.7 Hz. As the corner frequency of the source spectrum decreases with magnitude, higher magnitudes ($M_L \geq 4$) contribute relatively more to the lower frequencies, but these are fewer. At high frequencies, coverage drops above ~30 Hz, which is related to the dominant sampling rate of 100 Hz and the associated anti-alias filters.

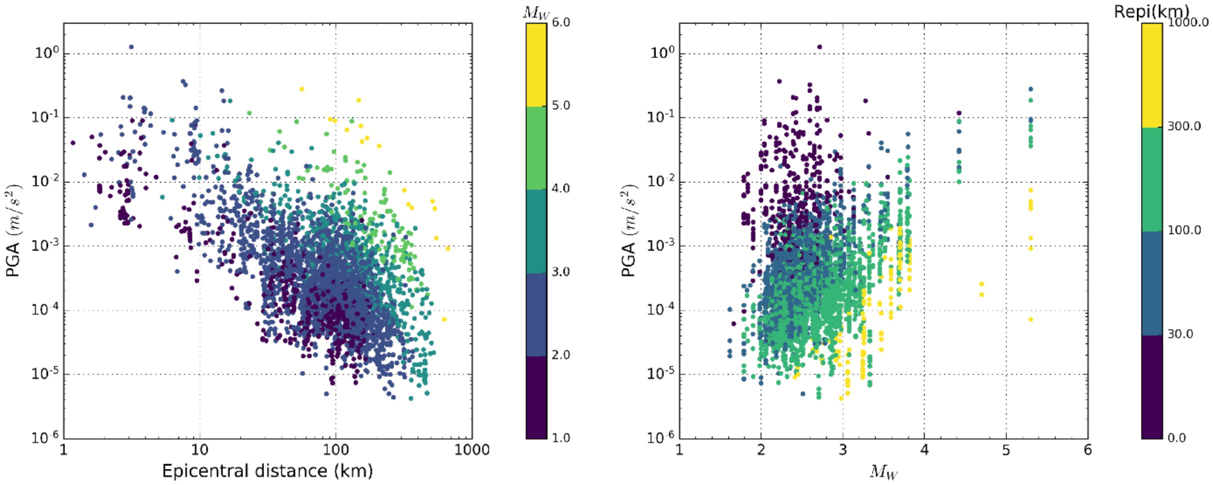


Figure 47 - Distribution of PGA values computed for the RotD50 component and "auto" window in the BELSHAKE database. Left: as a function of epicentral distance, colors represent M_W classes; Right: as a function of magnitude M_W, colors represent epicentral distance classes

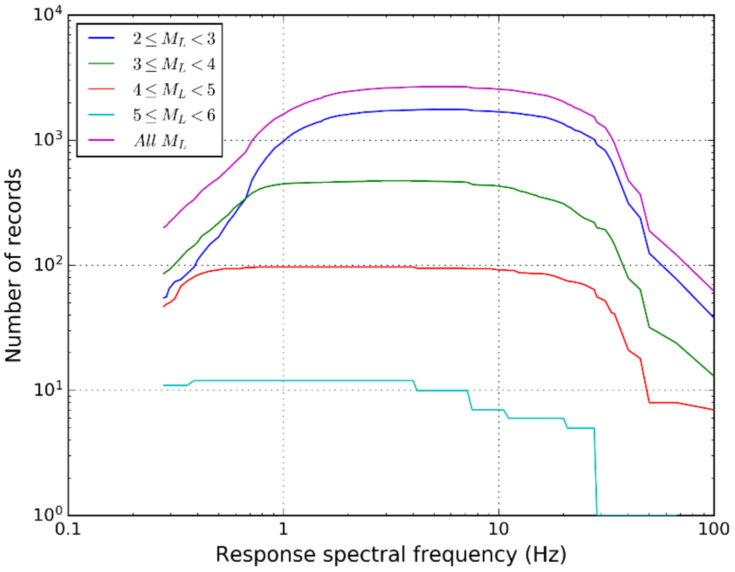


Figure 48 - Number of reliable records for each response spectral frequency (GM component, S window) and different magnitude ranges

4.1.8. DATA VALIDATION

To ensure the quality and reliability of IMs in the BELSHAKE database, we applied different consistency checks to identify anomalous records. Outliers were examined, and those considered as problematic were either corrected and reprocessed or removed from the database. The following evaluations were performed: (1) evaluation of different component ratios; (2) residual analysis with respect to a generic ground-motion model; (3) comparison of co-located stations; (4) comparison with events that are also included in the RESIF/EPOS-France database.

These different methods and their results are described in more detail in §3.1.9. Here, we only show the results of residual analysis based on mixed-effects regression of the final dataset, after several problems which came to light during the validation were corrected. This analysis involved a comparison between IMs (PGA/PGV) in the database and those predicted by a generic ground-motion model (GMM) fitted to the observations. The advantage of an ad-hoc GMM compared to a published model is that its median is centered on the dataset, which allows identifying inconsistencies with higher sensitivity (Traversa et al., 2020). We used the same functional form as Traversa et al. (2020) in their analysis of the French RESIF/EPOS-France database, and we also used the same reference distance ($R_{ref} = 1\ km$) and reference magnitude ($M_{ref} = 3.7$), but a different magnitude scale (M_W instead of M_L). Mixed-effects regression (using the Python statsmodels module of Seabold & Perktold, 2010) allows to decompose the total residual into different components:

$$log_{10}Y = F_D(R,M) + F_M(M) + \delta B_e + \delta S2S_s + \epsilon \qquad (5)$$

Y is the ground-motion parameter, i.e. PGA or PGV; $F_D(R,M)$ and $F_M(M)$ are the distance and magnitude scaling components of the GMM, which are considered as the fixed effects (see §3.1.9.3); δB_e and $\delta S2S_s$ are random effects corresponding to between-event and between-station residuals, respectively; ϵ is a lognormally distributed residual term, representing the unexplained or event- and station-corrected residual. This separation gives better insight into the origin of possible inconsistencies (e.g., related to source parameters, station/site parameters, …).

We computed residuals that are similar to or smaller than those in the French RESIF/EPOS-France database (Traversa et al., 2020). The results for PGV, RotD50 component and “auto” window are shown in Figure 8. The top panel displays the between-station residuals. These show a relatively good correlation with the time-averaged shear-wave velocity for the top 30 m (V_{S30}). Stations with $V_{S30} > 1200\ m/s$ show consistently negative residuals, whereas stations with $V_{S30} < 360\ m/s$ show generally positive residuals. The between-event residuals (bottom left panel) are all within the range -0.5 to +0.5. Data are sparser at higher magnitudes, but there is no obvious bias or trend with respect to magnitude. Outliers may be linked to earthquakes characterized by imprecise magnitude or location estimates but also to unique source characteristics, such as particularly high or low stress drop, which would require further investigation. Considering the good behavior of the between-event residuals, it may be concluded that there are no major problems in that respect in the BELSHAKE database. Finally, the unexplained residuals (bottom right panels) fall mostly between -0.5 and +0.5, and entirely between -1 and +1. They show no discernible trend with magnitude, while the slight variations with distance are likely related to divergence from the simplified geometrical spreading term in the fitted model.

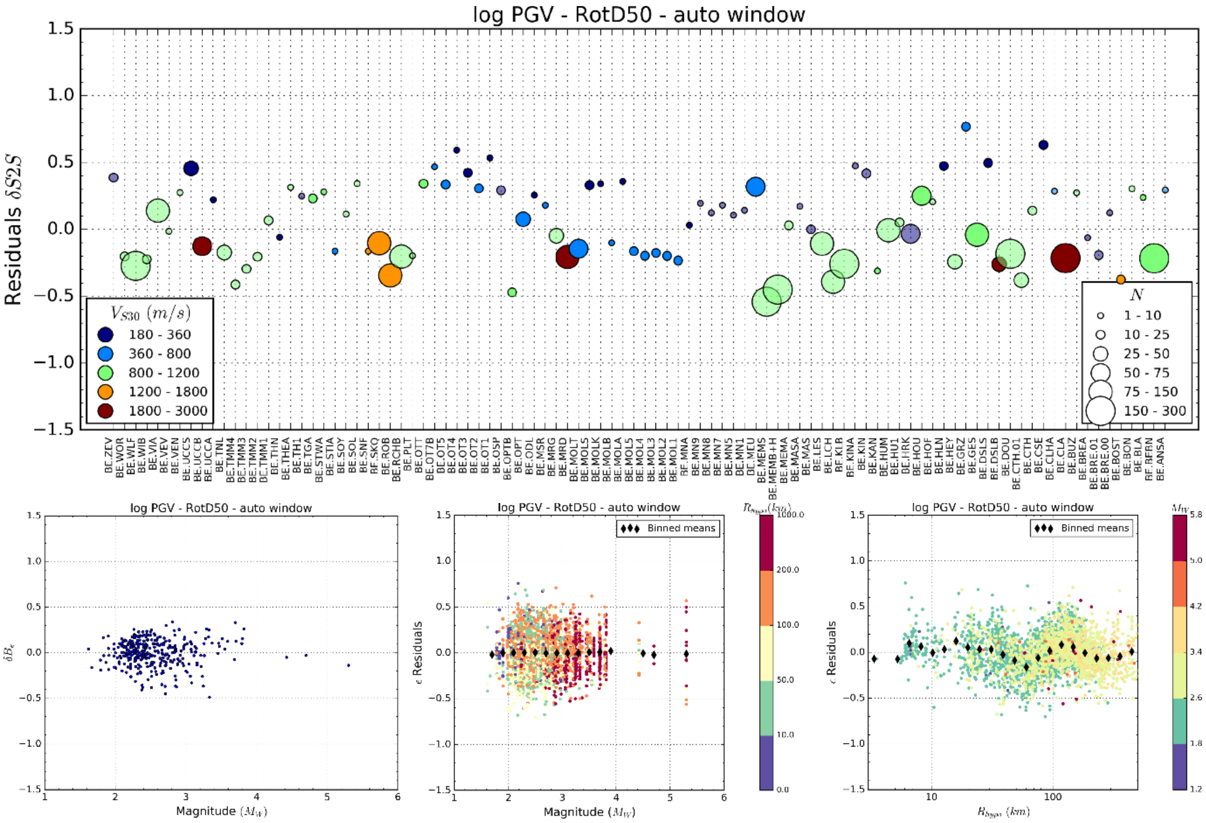


Figure 49 – Results of mixed-effects regression for PGV, RotD50 component, "auto" signal window. Top: station-to-station residuals ($\delta S2S_s$) for stations of the BE network (sorted alphabetically). Colors correspond to V_{S30} and marker sizes to number of records (N); transparent colors correspond to V_{S30} values estimated from Eurocode 8 soil class. Bottom left: between-event residuals (δB_e) as a function of M_W; Bottom center: unexplained residuals (ϵ) as a function of M_W, color-coded by hypocentral distance; Bottom right: ϵ residuals as a function of hypocentral distance, color-coded by M_W

4.1.9. CONCLUSION

For the first time, earthquake ground-motion data from Belgium and surrounding regions have been systematically collected in the BELSHAKE database. This database currently contains ~6400 records from 344 tectonic and induced earthquakes with $M_L \geq 2$, recorded since 1985 by a mix of single- and three-component short-period instruments, broadband seismometers and (to a lesser extent) accelerometers. Considerable efforts were undertaken to ensure that both the waveform data, the metadata and the derived ground-motion data adhere to the standards set by international ground-motion databases. All waveforms and phase picks have been visually inspected, and problems were either corrected or flagged. A unique feature of the BELSHAKE database is that waveform quality of each record can be assessed at the component and signal window level. Depending on the quality, different ground-motion intensity measures could be computed for each component-window combination. Reconstructing instrument response for older (pre-2005) records proved to be challenging, but different methods were used to check their validity. Records for which instrument response was found to be unreliable, were flagged in the database, preventing them to be used in subsequent steps. All reliable records were processed using a semi-automated, uniform workflow, which was adapted with slight variations from the processing schemes in existing ground-motion databases. Subsequently, up to 109 IMs (including 92 response spectral frequencies) were computed from the processed data, for more than 120,000 records and component-window combinations. In addition, we determined moment magnitudes from the waveform data for 294 events in the database. Their reliability was confirmed by comparing the results of different methods (using P or SH-waves) and comparing with values in other catalogs for a subset of the data. Finally, we performed various consistency checks to ensure the quality and reliability of IMs, including evaluation of different component ratios, residual analysis with respect to a generic ground-motion model, comparison of co-located stations, and comparison with common events in the French RESIF/EPOS-France database (Traversa et al., 2020). Anomalous records were analysed, and those found to be problematic were either corrected or removed. Overall, residuals show no discernible trend with distance or magnitude, and vary in a narrow range that is similar to or even smaller than that in RESIF/EPOS-France.

Although magnitude coverage of the BELSHAKE database is limited due to the relatively low seismic activity around Belgium, the small uncertainties on epicentral locations and magnitudes and low residuals testify its high quality. The BELSHAKE database will thus allow evaluating existing ground-motion models for application in Belgium. It will also contribute to a better assessment of seismic attenuation, which will in turn allow calibrating regionally adaptable models and guide stochastic simulation of earthquake ground motion. It will also contribute to a better assessment of seismic attenuation, which will in turn allow calibrating regionally adaptable models and guide stochastic simulation of earthquake ground motion. On a broader level, it will complement existing databases by increasing the coverage of low-seismicity zones and lower magnitudes. In addition, the processed waveform data in the BELSHAKE database will serve as a reference dataset for many other seismological applications in Belgium. To that end, we will continue to improve and extend the database beyond the project.

4.2. Building the capacity to model earthquake ground motions in Belgium

4.2.1. ANALYSIS OF THE SITE-SPECIFIC AND REGIONAL COMPONENTS OF THE CRUSTAL ATTENUATION FACTOR κ IN BELGIUM

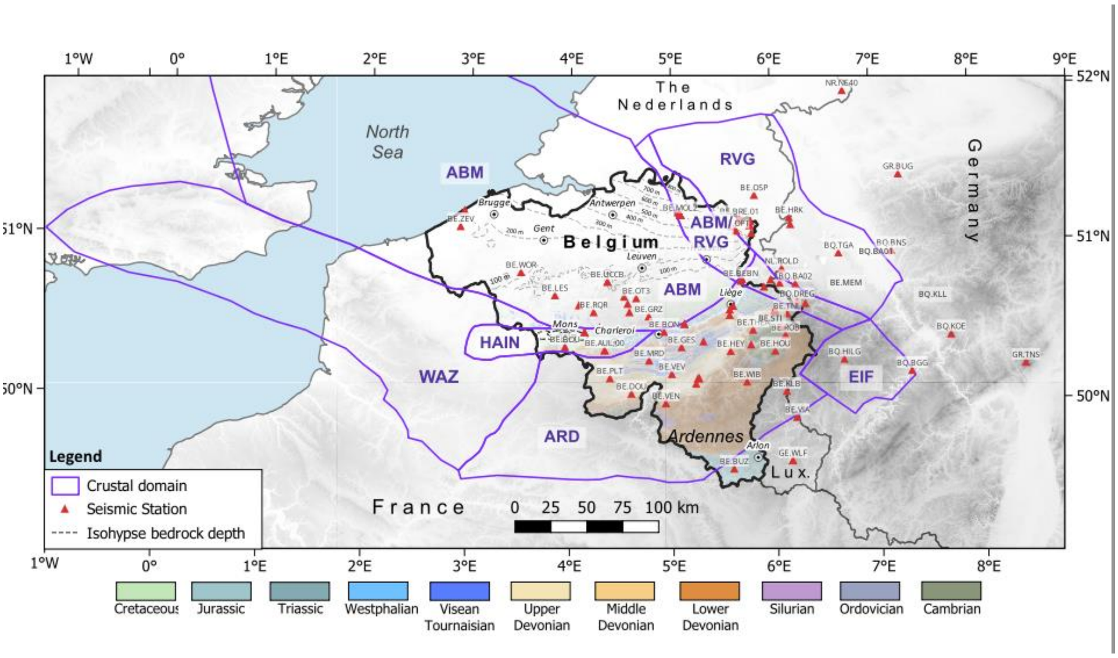
4.2.1.1. Crustal domains

We studied anelastic attenuation in four main crustal domains in and around Belgium, which were simplified from a more detailed seismotectonic model (Verbeeck et al., 2009): the Anglo-Brabant Massif (ABM), Ardenne (ARD), Hainaut (HAIN) and Roer Vally Graben (RVG). These domains are shown in Figure 50. Stations in transition zones between domains (ABM/RVG and ARD/RVG) were included in both relevant regions:

- ABM is a tectonic block connecting the Brabant Massif in Belgium with the Anglia Basin in England. In Belgium, it mainly consists of Early Paleozoic shales and metasandstones, which were strongly folded during the Brabantian phase of the Caledonian orogeny. These rocks are capped by Cretaceous limestones and marls, and a cover of Neogene sands and clays that becomes thicker towards the NNE, reaching > 1000 m at the Dutch border.
- ARD is a tectonic block that was deformed during the Upper Paleozoic Variscan orogeny and consists of an E-W oriented tectonic structure where Devonian and Carboniferous sandstones, shales and limestones are exposed in synclines and anticlines. Several Lower

Paleozoic metamorphic inliers are present. To the south, ARD is overlain by Triassic and Jurassic sediments. Quaternary deposits are only present in incised river valleys and are usually thin.

- HAIN is a narrow (only 15 km wide) basin at the northern edge of the Ardenne, which probably formed as a pull-apart basin connecting with the Weald-Artois shear zone (WAZ) to the west. In this basin, Cretaceous chalk and limestones are at the surface in the west, and Carboniferous coal, mud- and sandstones in the east. HAIN is limited in depth up to max. 6 km (Camelbeeck et al. 2025) and was overthrusted on top of the ABM during the Variscan orogeny.
- RVG is a tectonic zone characterized by active normal faulting since the Late Oligocene. This graben is filled with thick layers of loose Tertiary marine sands and clays, covered by Quaternary alluvial gravels and sands. The bedrock, composed of Late Cretaceous to Early Neogene limestone, is situated at a depth of more than 1000 m in the center of the graben, compared to only a few hundreds of meters on its flanks. ABM/RVG marks a transition zone between both crustal domains. Here, the top of the basement quickly plunges to the NE and is overlain by Upper Devonian and Carboniferous sand- and limestones deposited during early development of the RVG. Similarly, ARD/RVG indicates the zone where faulting in the RVG affected the ARD. Both ARD/RVG and ABM/RVG are characterized by predominantly NW-SE- to NNW-SSE-trending normal faulting due to RVG extension.



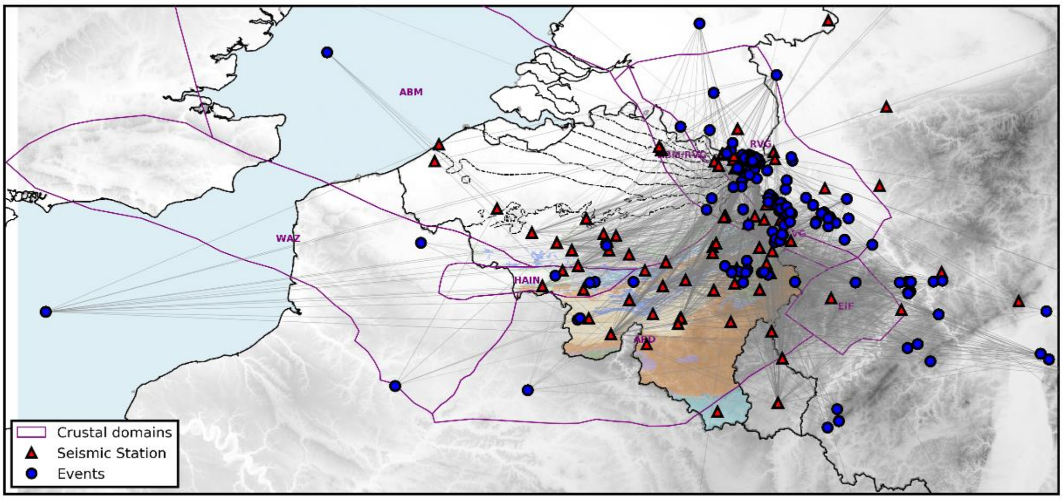


Figure 50. Top: Map showing main crustal domains and stations where kappa was determined in Belgium. Background image depicts surface geology, with bedrock in the south and NNE-ward thickening sediment layers (100 m contour interval) in the north; Bottom: all pairs of seismic events (blue circles) and recording stations (red triangles) used for κ_0 and κ_r estimation. Gray lines connect events to stations, illustrating the coverage density across Belgium. The figure highlights clear differences in event–station coverage between the main crustal domains.

4.2.1.2. Estimation of station κ_0 and regional κ_r values based on different fitting methods

We estimated the site-specific and path-dependent components κ_0 and κ_r from whole-path κ values determined from the BELSHAKE database using three different fitting methods: the free kappa gradient (FKM) method, the joint kappa gradient (JKR) method and the mixed-effects regression (MER) method. These methods are explained in §3.2.1.2.

Table shows the result of κ_r for different crustal domains computed by the three methods, together with the final selected κ_r value. The corresponding κ_0 values for each station, computed using the same three methods along with the final selected κ_0, are provided in Table (H component) and Table (Z component).

The FKG method provides a station-by-station assessment of κ_0 and the crustal-domain κ_r. Stations with sufficient data yield reliable estimates, evaluated using the following criteria: stations with positive κ_0 and κ_{r_s} are considered reliable, whereas stations with negative κ_0 are classified as unreliable. For stations with negative κ_{r_s}, reliability depends on data availability: results are deemed unreliable when fewer than five data points are available, typically accompanied by large κ_0 values and low R^2, while results with mildly negative κ_{rs} ($-1 < \kappa_{r\_s} < 0$) and more than five data points are retained as acceptable. A second phase of analysis only used a smaller subset of stations with reliable results. Compared to the full dataset, the R^2 values improved, indicating that removing unreliable stations enhances model performance, despite reducing data quantity. Importantly, κ_r estimates remained stable, confirming the reliability of domain-level results derived through this method.

Crustal domian	Component code	FKG method		JKG method			MER method					Final adopted κ_r	Q_i
		κ_r	R^2	κ_r	R^2	σ	κ_r	Marginal R^2	Conditional R^2	R^2	σ		
ABM	H	0.14	0.59	0.16	0.54	0.40	0.17	0.18	0.60	0.41	0.03	**0.17**	1634
ARD	H	0.16	0.64	0.15	0.62	0.15	0.16	0.29	0.55	0.46	0.01	**0.16**	1736
RVG	H	0.11	0.37	0.15	0.24	0.46	0.10	0.09	0.70	0.66	0.02	**0.13**	2137
ABM	Z	0.07	0.42	0.05	0.24	0.24	0.09	0.15	0.24	0.26	0.02	**0.08**	3968
ARD	Z	0.12	0.51	0.10	0.49	0.09	0.10	0.15	0.45	0.43	0.01	**0.10**	2778
HAIN	Z	0.13	0.56	0.06	0.19	0.4	0.03	0	0.04	0.07	0.06	**0.05**	-
RVG	Z	0.08	0.23	0.08	0.21	0.2	0.10	0.09	0.64	0.59	0.02	**0.09**	**3086**
All	H	-	-	-	-	-	0.16	0.24	0.58	0.47	0.02	**0.16**	1736
All	Z	-	-	-	-	-	0.09	0.09	0.59	0.45	0.01	**0.09**	3086

Table VII. Estimated κ_r values, obtained with the three different fitting methods for both H and Z components in different crustal domains of Belgium. The table lists κ_r values along with the final adopted κ_r value, Q, R^2 values and standard deviation (σ) for different methods/crustal domains. Note that standard deviations for the JKG method are based on an approximation of the inverse Hessian matrix in the optimizer, which does not appear to be stable and often results in unrealistically high values. The two last rows contain values obtained with the MER method for H and Z components with pooled data of all crustal domains in Belgium

Crustal domain	**Station code**	**N**	**FKG method κ0 (ms)**	**JKG method κ0 (ms)**	**MER method κ0 (ms)**	**Final selected κ0 (ms)**
ABM	BE.BOST	3	-0.66	0	26,63	
ABM	BE.CSE	7	-33.46	13,37	21,56	**22**
ABM	BE.DSLB	8	29.08	24,01	24,91	**25**
ABM	BE.DSLS	2	39.52	42,09		**42**
ABM	BE.GRZ			35,04		
ABM	BE.LES	10	99.94	50	22,15	**22**
ABM	BE.MOL1			24,88		
ABM	BE.MOL2	3	10.26	21,79	23,78	**24**
ABM	BE.MOL3	2	27.76	32,54	24,09	**24**
ABM	BE.MOL4	2	28.95	34,25	24,11	**24**
ABM	BE.MOL5	2	34.06	23,38	24,36	**24**
ABM	BE.MOLS			50		**-**
ABM	BE.MOLT	20	28.75	25,99	25,98	**26**
ABM	BE.OT2			8,33		
ABM	BE.OT3			17,04		
ABM	BE.OTT			0		
ABM	BE.SKQ	48	27.08	19,11	23,7	**24**
ABM	BE.UCCB	5	-0.44	2,8	25,14	**25**
ABM	BE.UCCS	11	103.51	50	25,37	**25**
ABM	BE.WOR	3	0.10	50	21,29	**21**
ABM	BE.ZEV	4	34.05	21,95		
ARD	BE.ANSA			22,64		
ARD	BE.BEBN	85	23.09	18,5	22,4	**22**
ARD	BE.CLA	96	35.45	25,56	33,66	**34**
ARD	BE.CTH	5	31.82	38,1	34,77	**35**
ARD	BE.CTH.01	11	-9.15	22,16	14,66	**15**
ARD	BE.DOU	78	19.90	16,16	21,46	**21**
ARD	BE.GES	75	21.82	26,42	21,34	**21**
ARD	BE.HEY	9	-2.74	18,54	14,83	**15**
ARD	BE.HOU	17	22.90	23,72	18,07	**18**
ARD	BE.HU1	6	-17.56	25,33	23,4	**23**
ARD	BE.HUM	58	58.85	36,43	50,19	**50**
ARD	BE.KLB	90	6	19,27	15,73	**16**

Crustal domain	Station code	N	FKG method κ_0 (ms)	JKG method κ_0 (ms)	MER method κ_0 (ms)	Final selected κ_0 (ms)
ARD	BE.LCH	38	20.86	25,28	21,27	**21**
ARD	BE.MEMA	8	31.37	33,02	32,68	**33**
ARD	BE.MEMB	36	12.41	17,54	15,71	**16**
ARD	BE.MEMH	25	34.02	21,29	32,21	**32**
ARD	BE.MEMS	68	26.83	26,09	18,7	**19**
ARD	BE.MEU	25	59.44	50	57,84	**58**
ARD	BE.MRD	36	17.10	18,45	17,57	**18**
ARD	BE.MRG	11	27.62	23,53	15,01	**15**
ARD	BE.PLT	6	21.29	33,86	26,9	**27**
ARD	BE.RCHB	17	16.13	7,6	17,77	**18**
ARD	BE.ROB	18	48.07	25,83	27,06	**27**
ARD	BE.STWA			30,12		
ARD	BE.THEA			23,51		
ARD	BE.TMM1	8	9.81	19,46	16,54	**17**
ARD	BE.TMM2	8	-5.24	20,97	14,56	**15**
ARD	BE.TMM3	8	4.69	18,13	14,08	**14**
ARD	BE.TMM4	8	-2.07	13,53	8,88	**9**
ARD	BE.TNL	4	99.15	29,32		
ARD	BE.VEV			17,19		
ARD	BE.VIA	74	15.65	20,36	16	**16**
ARD	BE.WIB	3	20.00	4,29	11,73	**12**
HAIN	BE.CLHA			25,69		
RVG	BE.BRE.00	5	35.68	43,67	45,39	**45**
RVG	BE.BRE.01	4	151.93	50	58,61	**59**
RVG	BE.BREA			37,9		
RVG	BE.HRK	28	53.32	46,61	48,02	**48**
RVG	BE.KAN	2	55.38	50		
RVG	BE.KIN	7	76.28	50	73,61	**74**
RVG	BE.KINA			50		
RVG	BE.MASA	3	88.91	50	60,79	**61**
RVG	BE.OPT	15	26.16	26,87	30,63	**31**
RVG	BE.OPTB	4	8.54	23,44	29,32	**29**
RVG	BE.OSP	5	38.25	50	52,77	**53**

Table VIII - Comparison of κ_0 values on the H component for all stations of the Belgian seismic network in the BELSHAKE database computed with the FKG, JKG and MER methods (after excluding problematic stations), along with the final selected κ_0 values. Stations without a final selected κ_0 value are those for which no reliable estimate could be obtained using any of the three methods

Crustal domain	Station code	N	FKG method κ_0 (ms)	JKG method κ_0 (ms)	MER method κ_0 (ms)	Final selected κ_0 (ms)
ABM	BE.CSE	9	4.62	23,52	18,08	**18**
ABM	BE.DSLB	11	6.65	16,05	15,36	**15**
ABM	BE.DSLS	2	3.13	25,49	18,57	**19**
ABM	BE.GRZ			16,41		
ABM	BE.LES	37	52.86	33,62	-	**34**
ABM	BE.MOL1			26,67		
ABM	BE.MOL2	3	-0.34	15,57	17,27	**17**
ABM	BE.MOL3			17,96		**18**
ABM	BE.MOL4			12,03		**12**
ABM	BE.MOL5			12,43		**12**
ABM	BE.MOLT	4	37.06	19,73	16,3	**16**
ABM	BE.OT2			23,68		
ABM	BE.OT3			50	-	
ABM	BE.OTT	2	-11.33	13,59	17,28	**17**
ABM	BE.RQR	57	26.35	32,1	23,46	**23**

Crustal domain	**Station code**	**N**	**FKG method κ_0 (ms)**	**JKG method κ_0 (ms)**	**MER method κ_0 (ms)**	**Final selected κ_0 (ms)**
ABM	BE.SKQ	38	26.82	23,38	-	**23**
ABM	BE.UCCA			31,34		
ABM	BE.UCCB	7	49.79	33,46	-	**33**
ABM	BE.UCCS	16	19.04	26,58	19,65	**20**
ABM	BE.WOR	3	150.44	50	-	
ABM	BE.ZEV	6	39.22	31,55	19,8	**20**
ARD	BE.BEBN	125	22.71	20,71	22,68	**23**
ARD	BE.BON	2	2.39	20,68		
ARD	BE.CLA	104	29.62	25,69	28.05	**28**
ARD	BE.COI	2	69.32	31,36		
ARD	BE.CTH	15	33.64	36,04	32,29	**32**
ARD	BE.CTH.01	10	10.23	29,4	26,35	**26**
ARD	BE.DOU	117	15.91	22,23	20,34	**20**
ARD	BE.GES	63	24.42	18,72	23,52	**24**
ARD	BE.HEY	10	-0.64	18,68	19,89	**20**
ARD	BE.HOU	12	20.43	41,4	27,49	**27**
ARD	BE.HU1	7	-18.88	28,52	26,51	**27**
ARD	BE.HUM	67	43.04	29,92	34,09	**34**
ARD	BE.KLB	69	18.54	12,96	17,02	**17**
ARD	BE.LCH	42	11.16	17,63	12,41	**12**
ARD	BE.MEMA	9	21.43	21,08	22,2	**22**
ARD	BE.MEMB	51	26.13	21,33	24,07	**24**
ARD	BE.MEMH	29	22.08	18,95	20,47	**20**
ARD	BE.MEMP			30,09		
ARD	BE.MEMS	141	21.70	22,71	20,81	**21**
ARD	BE.MEU	45	61.26	48,06	49,25	**49**
ARD	BE.MRD	45	7.09	19,61	14,95	**15**
ARD	BE.MRG	5	18.82	7,48	19,19	**19**
ARD	BE.PLT	7	53.13	17,73	21,03	**21**
ARD	BE.RCHB	25	16.83	24,78	20,92	**21**
ARD	BE.ROB	32	32.41	27,8	28	**28**
ARD	BE.STI	95	38.87	42,1	38,28	**38**
ARD	BE.THEA			22,63		
ARD	BE.TMM1	6	30.18	20,48	21,63	**22**
ARD	BE.TMM2	6	5.00	22,55	21,92	**22**
ARD	BE.TMM3	9	12.44	12,35	16,56	**17**
ARD	BE.TMM4	6	10.18	12,79	16,46	**16**
ARD	BE.TNL	7	26.72	7,27	20,52	**21**
ARD	BE.VEN			20,69		
ARD	BE.VIA	77	17.39	24,42	23,44	**23**
ARD	BE.WIB	2	23.57	6,65	14,22	**14**
HAIN	BE.AUL	21	69.80	41,61	52,85	**53**
HAIN	BE.AUL.00			50		
HAIN	BE.BOU	44	26.88	41,12	35,18	**35**
HAIN	BE.BRQ.00	14	82.89	50	74,86	**75**
HAIN	BE.BRQ.01	8	37.88	28,57	46,76	**47**
HAIN	BE.SNF	84	71.21	50	64,09	**64**
RVG	BE.BRE.01	5	22.53	22,63	-	
RVG	BE.BREA			37,19		
RVG	BE.HRK	28	45.71	43,84	42,57	**43**
RVG	BE.KAN	2	29.53	25,31		
RVG	BE.KIN	9	77.87	50	78,15	**77**
RVG	BE.KINA			50		
RVG	BE.MASA	3	104.82	50	73,14	**73**
RVG	BE.OPT	7	20.64	23,54	26,77	**26**
RVG	BE.OPTB	3	16.87	19,73		**20**
RVG	BE.OSP	6	37.09	42,25	42,5	**42**

Table IX - Comparison of κ_0 values on the Z component for all stations of the Belgian seismic network in the BELSHAKE database computed with the FKG, JKG and MER methods (after excluding problematic stations), along with the final selected κ_0 value. Results for problematic stations are removed; consequently, no mixed-effects κ_0 values are available for them, and these entries are indicated with a "–". The κ_0 values for the two stations in HAIN that have negative station-specific slopes were retained in the table, and their results are underscored. Stations without a final selected κ_0 value are those for which no reliable estimate could be obtained using any of the three methods

The other two methods improve robustness by jointly fitting all stations within each crustal domain. The JKG method performs a single κ_r fit to all κ values in a domain, but requires predefined bounds for κ_0 and κ_r, which we selected based on the FKG results. Initial tests showed convergence to unrealistic κ_r values, particularly for the Z component, prompting a sensitivity analysis. Optimal bounds of 0–0.3 ms/km for κ_r and 0–50 ms for κ_0 yielded the best fit and closest agreement with the FKG method. Horizontal components show relatively uniform κ_r values (0.15–0.16 ms/km), whereas vertical components vary more (0.05–0.10 ms/km). κ_0 values reach the imposed upper limit at a few stations, especially in RVG and HAIN, reflecting limited data (≤4 records).

Figure 51 presents JKG results for station BE.BEBN in crustal domain ARD. The left panel shows the station-level results for BE.BEBN (H component), with κ_0 as the intercept and the domain-level κ_r as the slope. The right panel displays the corresponding crustal domain plot (e.g., κ_r = 0.15 ms/km for the ARD domain, H component). To assess fit quality, R^2 values are computed for each station. The results demonstrate that stations with sufficient data generally yield high R^2 values, indicating that the domain-level κ_r provides a robust and consistent fit for each station. However, although the method assigns a station-specific κ_0 to all stations, the reliability of the results for some stations with very few data points may still be questionable.

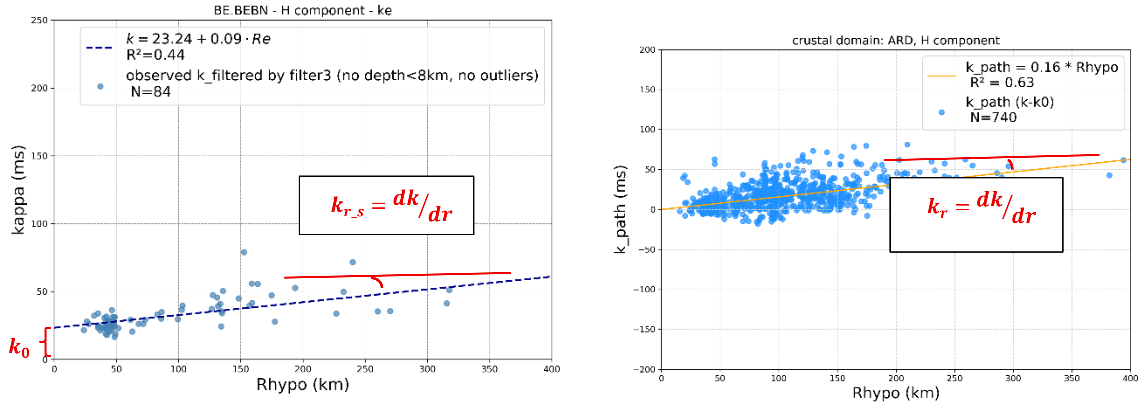


Figure 51 - Example of fitting results obtained with the JKG method. Left: κ values versus hypocentral distance (Rhypo) for the H component of station BE.BEBN. The brown line represents the fit, drawn with the station-specific κ_0 as the intercept and the joint crustal-domain κ_r as the slope; Right: κ_{path} versus Rhypo for the H component of the ARD crustal domain. The slope of the green line corresponds to the fitted κ_r. It is drawn with zero intercept, as the fitted station-specific κ_0 values are incorporated in the κ_{path} data.

As a third approach, we applied mixed-effects regression to capture both crustal-domain trends and station-specific variability in κ. This framework allows all stations to be modeled simultaneously while accounting for hierarchical structure and uneven data coverage. The results are shown in four different types of plots in Figure 52 to Figure 55:

a) Station-specific slopes (κ_{r_s}) and intercepts (κ_0). Clear station-to-station variability is evident in both slope and intercept, reflecting the interplay between regional and local effects and underscoring the need to account for station-specific attenuation characteristics. This type of plot allows identifying stations that are not consistent with the overall gradient in the crustal domain. For ARD-H, all stations generally exhibit positive slopes (κ_{r_s}= 0.1-0.24 ms/km), supporting the robustness of the method. However, a few stations deviate from this pattern. e.g. BE.HUM ($\kappa_{r\_s} = 0.03\ ms/km$) but were retained for further analysis. For all domains, stations with very limited data (often from neighboring networks) and stations with negative slopes were excluded to ensure reliable domain characterization. The analysis was then repeated using only stations with consistent behavior. All results shown here are based on this filtered dataset. After filtering, station-specific slopes (κ_{r_s}) are generally consistent within each crustal domain and component, indicating stable path attenuation effects, with the exception of HAIN-Z, which remains problematic (discussed in more detail later). In contrast, κ_0 values vary between stations, reflecting expected site effects. These results demonstrate that the MER method reliably estimates both the site κ_0 and the κ gradients (κ_{r_s}), in agreement with theoretical expectations.

b) Random intercepts for each station. Positive values (above the dashed red line) indicate stations with higher κ_0 than the global value, while negative values indicate lower κ_0. In most domains, stations cluster within a narrow range around zero (±10 ms), suggesting limited deviation from the global κ_0. The HAIN region exhibits the highest κ_0 values, which is consistent with previous studies in this area (Camelbeeck et al., 2022).

c) Predicted versus observed whole-path κ values. These are plotted along with a linear regression fit (blue line) and the Pearson correlation coefficient (PCC), which quantifies how closely the predicted values align with the observed ones. Good agreement between observed and predicted κ values was observed in most domains, with high PCC scores for RVG-H (0.78) and RVG-Z (0.82), and moderate correlations in ABM-H (0.65), ARD-H (0.68), and ARD-Z (0.66). However, weaker correlations were observed in ABM-Z (0.48) and HAIN-Z (0.32).

d) Residuals vs. fitted κ values. The residuals have a balanced distribution around zero (red dashed line), with no significant bias, and only a few larger residuals attributable to some data-poor stations, suggesting that the model predictions are generally robust.

The mixed-effects model offers a more nuanced representation of κ variability by accounting for both regional trends and site-specific deviations simultaneously. Compared to the FKG method, it yields more reliable estimates for stations with limited data by drawing strength from global domain level trends. In contrast to the JKG method, it allows for more precise modelling by assigning each station its own κ_0 and κ_{r_s}, thereby capturing local variations more effectively. However, the model's performance declines in the HAIN-Z domain, primarily due to limited data availability. In this region, two of the five stations exhibit negative slopes, and the global κ_r (≈0.01 ms/km) was very low before excluding problematic stations from this domain. Even after removing these stations, the results remain unsatisfactory; the global κ_r (≈0.03 ms/km) is still low and one of the remaining stations which initially showed a positive slope subsequently exhibits a negative slope (BE.AUL in Figure 52). Beyond the issue of sparse data, the HAIN domain is relatively small and has been shown in previous studies to exhibit stronger attenuation than other crustal domains in Belgium (Camelbeeck et al., 2022). Most events recorded by its stations originate outside the domain, with seismic waves passing through one or two other domains before arrival. The inconsistency likely reflects the combined effects of strong attenuation (high κ_0) at short distances within HAIN and much lower attenuation at greater distances. As these signals propagate into HAIN, abrupt changes in attenuation properties may cause significant signal loss, which can result in near-zero or even negative κ_r for this crustal domain.

Since the κ_r values obtained from the JKG and MER methods showed minimal variation across different crustal domains, we assigned the average of both as a final κ_r value to each domain, as shown in Table . The final adopted κ_r values range from 0.13 to 0.17 ms/km (H component) and from 0.05 to 0.10 ms/km (Z). The standard deviations on κ_r obtained with MER are generally small across all crustal domains, indicating stable and reliable estimates. The lowest standard deviations are observed for the ARD and RVG domains (0.006–0.016 ms/km), while slightly higher values appear for HAIN and ABM (up to 0.056 ms/km). These results are consistent with the ray coverage of the different domains (Figure 50): domains with denser data coverage, such as ARD and RVG, yield smaller standard deviations. R^2 values are generally higher for H than Z components, whereas the FKG and JKG methods yield similar R^2 values for both components (e.g., ABM-H: 0.59 vs. 0.53; ARD-H: 0.64 vs. 0.61; Z components mostly between 0.2–0.5 for both), except for HAIN-Z where the divergence is more noticeable. This indicates that the two approaches explain the variability in the data to a comparable degree. The MER model reveals the added value of incorporating station-specific random effects: while marginal R^2 values remain low (≤0.3), indicating that domain-level fixed effects alone explain only a small portion of the variance, the conditional R^2 values increase substantially (0.24–0.77). This clear improvement shows that most of the variance in κ is captured when local, station-specific effects are included.

The final κ_0 values for each station (Table and Table) were adopted from the MER estimates. Standard deviations range from 3.89 to 23.21 ms, indicating substantial site-to-site variability across the Belgian crustal domains. The relatively large uncertainties reflect the strong influence of local site conditions, and possibly instrumentation or emplacement differences, all of which cause certain stations to deviate markedly from the global trends. This finding also explains why conditional R^2 values substantially exceed marginal R^2. For instance, for the H component the increase from marginal to conditional R^2 is evident: ABM (0.18 → 0.60), ARD (0.29 → 0.55), and RVG (0.09 → 0.70). For the Z component, the marginal R^2 values are lower than for H component, showing that vertical estimates are considerably more difficult to capture with only global trends. Although accounting for random effects leads to substantial relative improvements in some cases for the Z component, such as RVG-Z (0.07 → 0.77) and ARD-Z (0.15 → 0.45), the final R^2 values remain low in poorly constrained domains like ABM-Z (0.15 → 0.24) and HAIN-Z (0 → 0.04). This indicates that much of the variance remains unaccounted even after including random effects in these cases. It can be concluded that the MER framework can provide meaningful improvements, but it is still not always sufficiently powerful to fully overcome the limitations imposed by scarce data in some of the crustal domains or components.

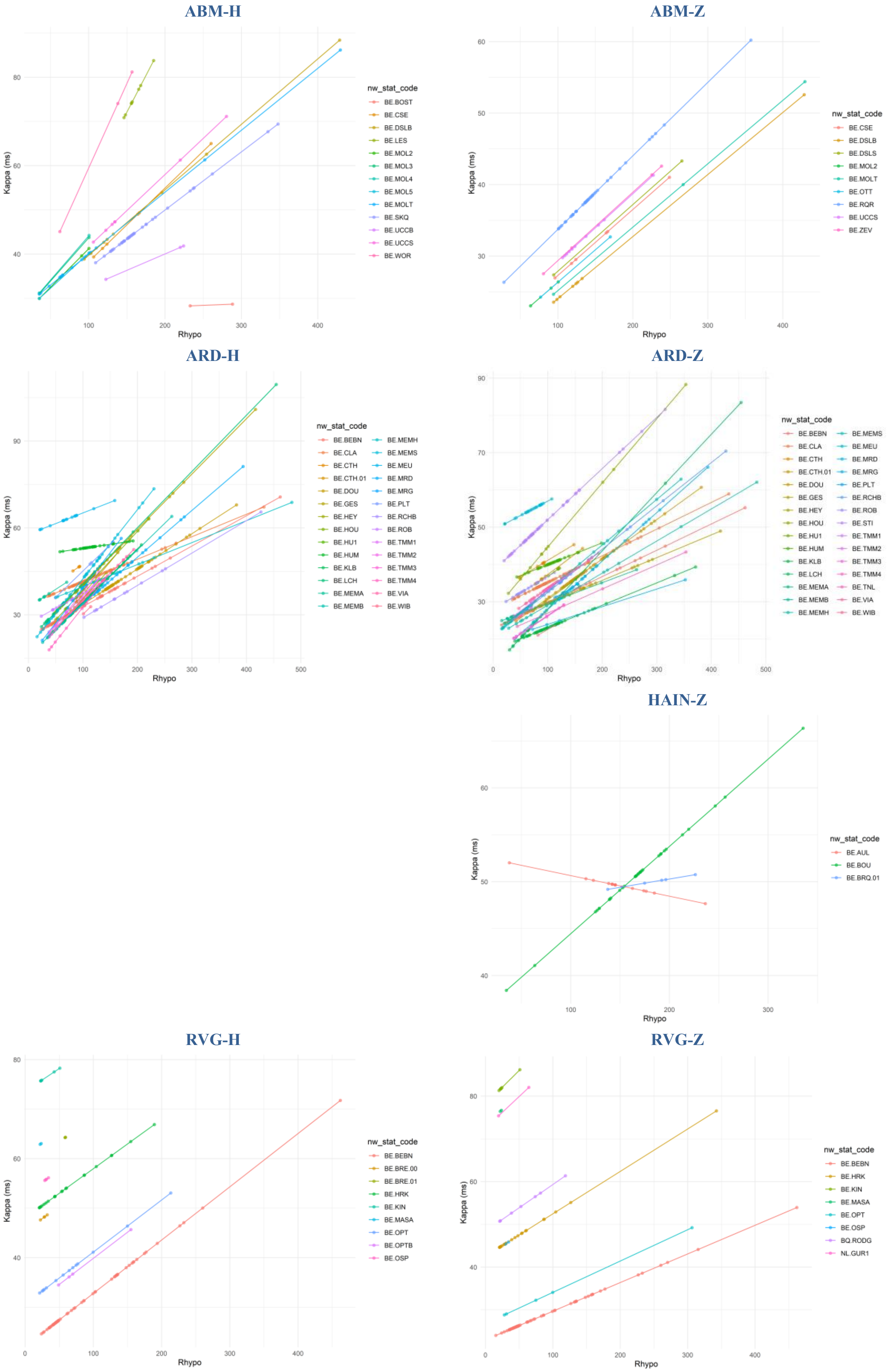


Figure 52 - Station-specific (= random + fixed) slopes (κ_{r_s}) and intercept κ_0 for each station within each crustal domain (rows) and component (columns) derived with the MER method after excluding problematic stations

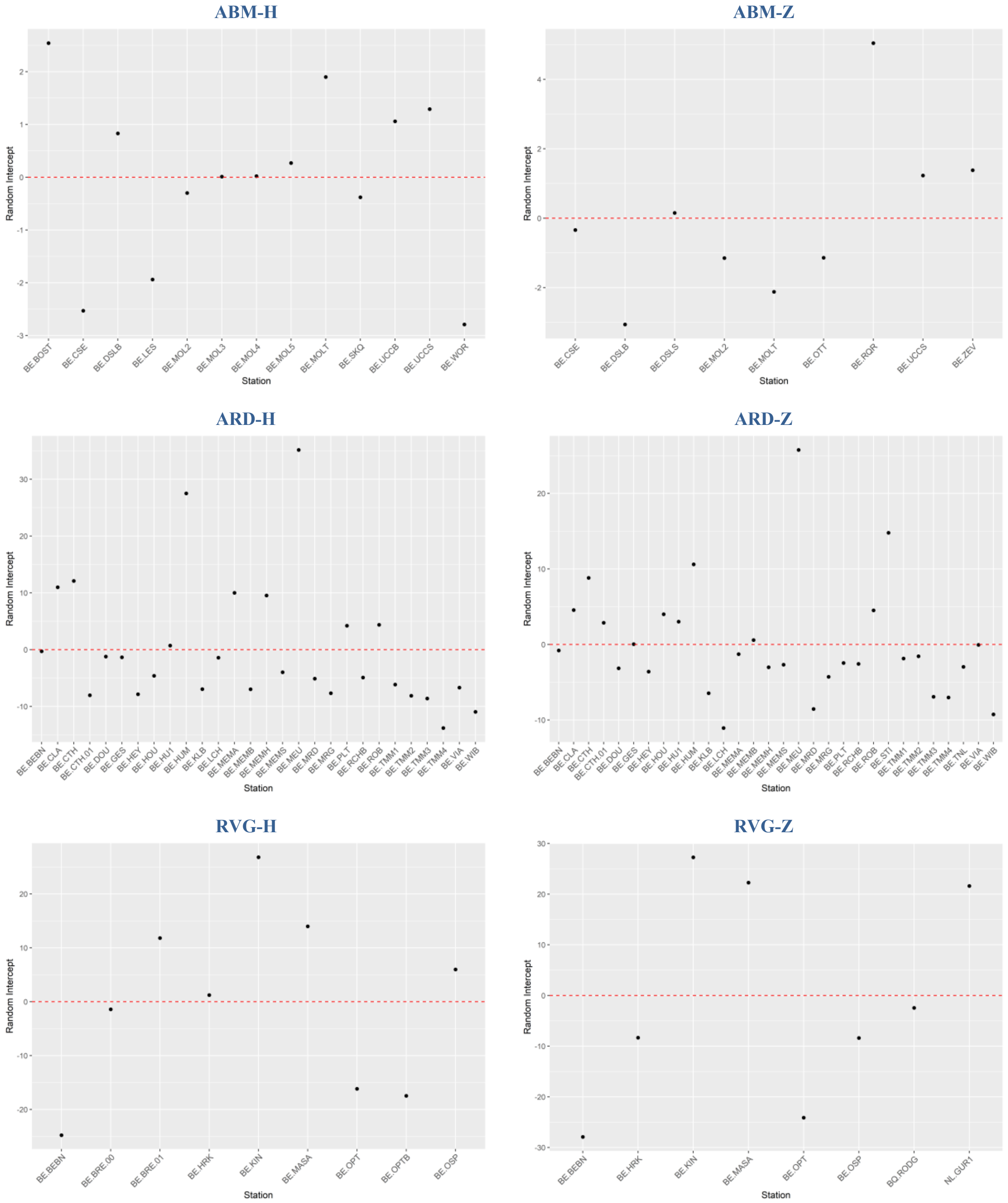
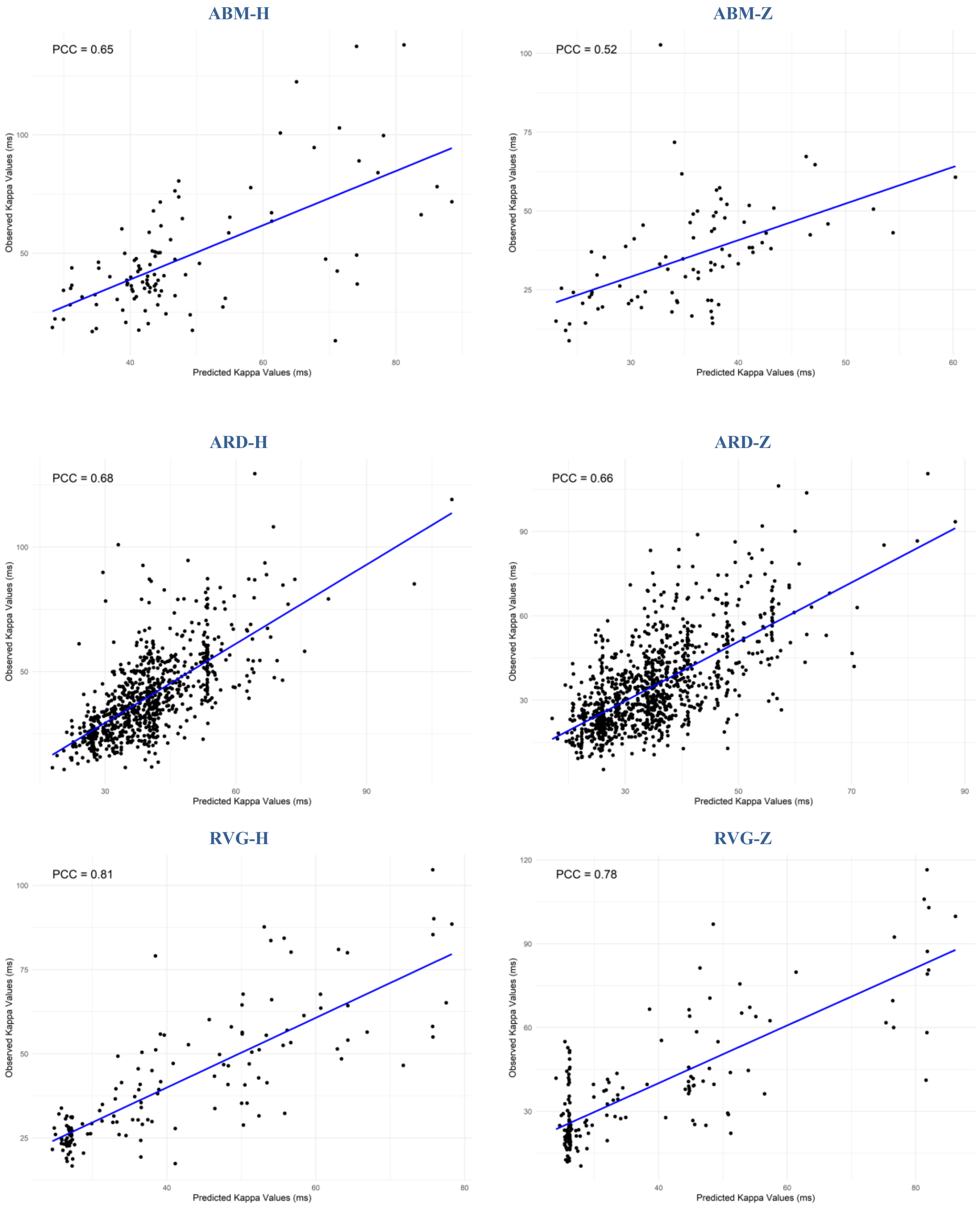


Figure 53 - Random intercept for all stations within each crustal domain except HAIN (rows) and component (columns) derived with the MER method after excluding problematic stations. Dashed red line indicates zero line



ABM-H

ABM-Z

ARD-H

ARD-Z

RVG-H

RVG-Z

Figure 54 - Observed versus predicted κ values along with PCC score derived using the MER method after excluding problematic stations. Blue line indicates the linear regression fit

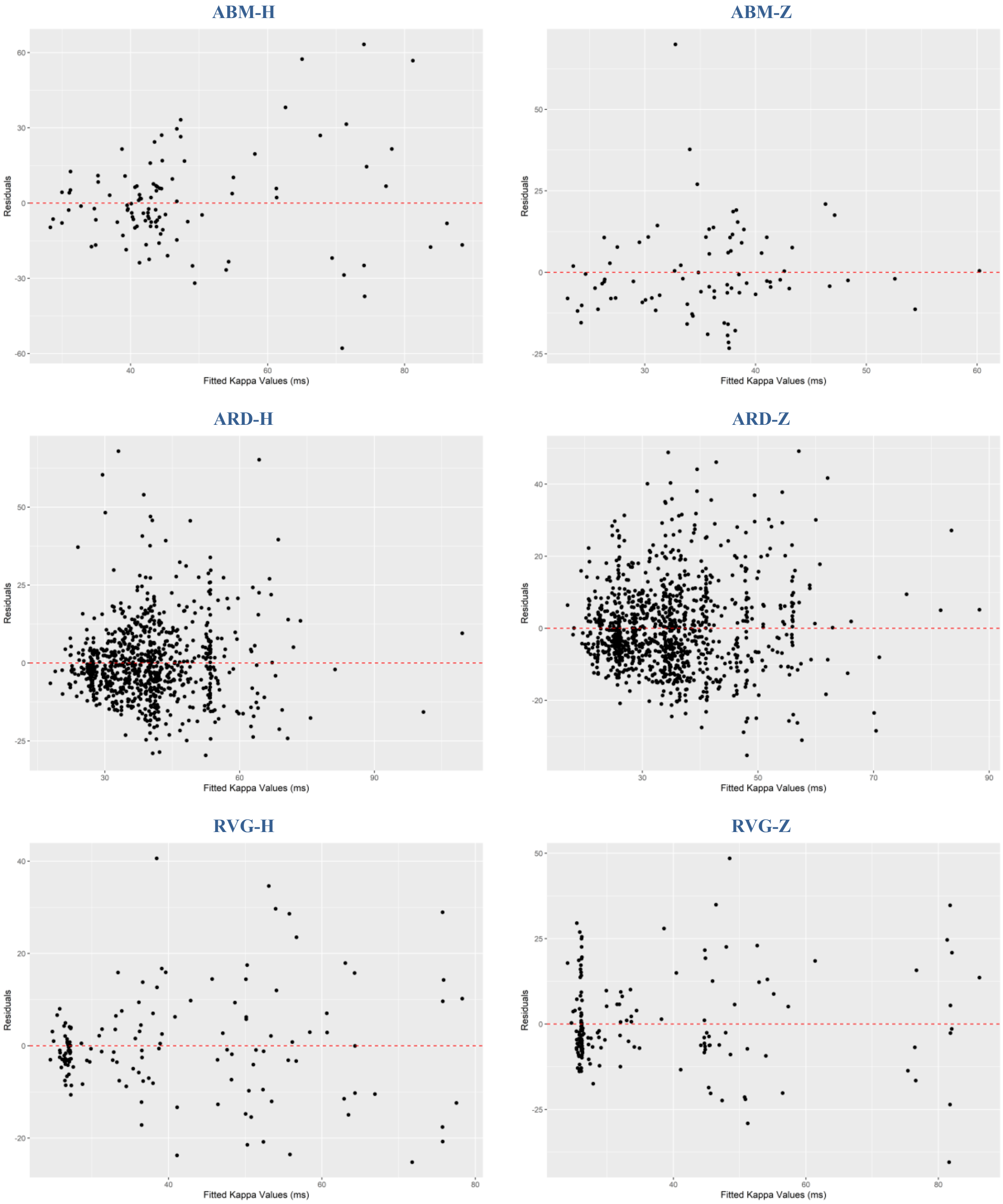


Figure 55 - Residuals versus fitted κ values obtained with the MER method for each crustal domain except HAIN (rows) and component (columns) after excluding problematic stations. Red dashed line indicates zero line

To overcome these limitations, we pooled all domains and applied the MER method to the combined dataset. This allows assessing the overall attenuation characteristics of the Belgian crust while still accounting for site-specific variability through random station effects. We thus obtain κ_r values of 0.16 ms/km for the H component and 0.09 ms/km for the Z component. The results in Table show that pooling enhances model performance compared to individual domains with limited data, such as HAIN and ABM, at the cost of a small reduction of fit quality compared to well-sampled domains like RVG and ARD. This indicates that results for domains with abundant data are best interpreted from their individual analyses, whereas pooled results offer a more robust representation for domains with sparse observations. The κ_0 estimates obtained from the pooled dataset (not shown here) are nearly identical to those derived from the individual crustal domains. In addition, two stations (BE.OSP and BE.WLF), now have κ_0 estimates from the pooled data, whereas they previously had none because they are not assigned to any specific crustal domain.

4.2.1.3. Relation between κ_r and Q

As described in §3.2.1.1, κ_r is inversely proportional to the apparent (frequency-independent) quality factor Q_i through the equation:

$$Q_i = \frac{1}{V_s \kappa_r}$$

Higher Q_i corresponds to lower attenuation, while lower Q_i indicates stronger attenuation. For the Belgian crust, we adopt a representative shear-wave velocity of V_s=3600 m/s, which is an average value based on the velocity model of Camelbeeck (1993). Applying this relation, κ_r values estimated for each domain were converted into Q_i, and the results are summarized in last column of Table . The obtained values range from about 1600 to 2100 (H component) and from 2700 to 5500 (Z component). However, as mentioned earlier, the κ_r estimations for the HAIN–Z domain are not reliable; therefore, the corresponding Q_i value is excluded from interpretation. Considering this, the Q_i values for the Z component range more realistically from about 2700 to 3900. The Q_i values based on the pooled data lie between the corresponding Q_i ranges for the H and Z components, reflecting the stabilizing effect of pooling across different crustal domains.

4.2.1.4. Relation between κ_0 and Vs_{30}

The κ_0 estimates for the H component from the three fitting methods were compared with globally expected κ_0 values based on V_{S30}, considering only stations with measured V_{S30} values (see §4.1.6), sufficient data for reliable estimation, and those located on Eurocode 8 soil classes A or B. As illustrated in Figure 56, the expected ranges are represented as two sets of box plots: one set based on the compilation of $\kappa_0 - V_{S30}$ values published for different regions in the world by Stanko et al. (2017) for which we considered for each station a V_{S30} range of ±100 m/s around the measured value and extracted the corresponding κ_0 values; the other set corresponds to the $\kappa_0 - V_{S30}$ model developed by Van Houtte et al. (2011) using KiK-Net and NGA data.

Figure 56 shows that all three fitting methods generally produce κ_0 estimates within these expected global ranges for stations with low V_{S30} values (associated with soft soil sites). This indicates that all these methods are effective under such conditions; however, it should be noted that the expected global ranges are wider for low V_{S30} sites, making it more likely for estimates to fall within them.

For bedrock sites (V_{S30} > 800 m/s), the results show that κ_0 values remain relatively similar across low (800–1200 m/s), intermediate (1200–1800 m/s), and high (>1800 m/s) V_{S30} values. The obtained values are in line with global expectations up to about 1500 m/s (apart from two outliers with the FKG method for stations BE.OPTB and BE.ROB), but not for higher V_{S30} values. Instead, they appear to remain constant, consistent with the observation by Ktenidou et al. (2015) that κ_0 tends to stabilize at high V_{S30}, suggesting the existence of a regional hard-rock asymptotic κ_0 value. This would imply that bedrock with high V_{S30} in Belgium does not show the low κ_0 values typical of hard rock in other regions of the world (e.g., κ_0 ≈ 0.005–0.006 s at CENA sites with V_{S30} ≈ 2000 m/s, based on Ktenidou et al. (2016). Instead, attenuation along short travel paths in Belgium may be influenced more by shallow crustal heterogeneity and regional geological context than by the near-surface shear-wave velocity alone.

The MER method achieves full station coverage (100%), meaning it estimates the site kappa for all stations with available records, even those with very few recordings. Of these, 80% fall within the expected range, with a mean deviation of 8.17 ms. The JKG method has a slightly lower coverage of 90%, with 77.8% of values falling within the expected range and a mean deviation of 6.25 ms. In contrast, the FKG method also has 90% coverage, but only 44.4% of estimates fall within the expected range, with the highest mean deviation of 13.0 ms. In summary, the MER method demonstrates the most consistent performance across the full range of V_{S30} values, particularly for sites with V_{S30} between 800 and 1500 m/s., while the JKG method produces the most accurate estimates in terms of average deviation. The FKG method also performs reasonably well, but shows limitations.

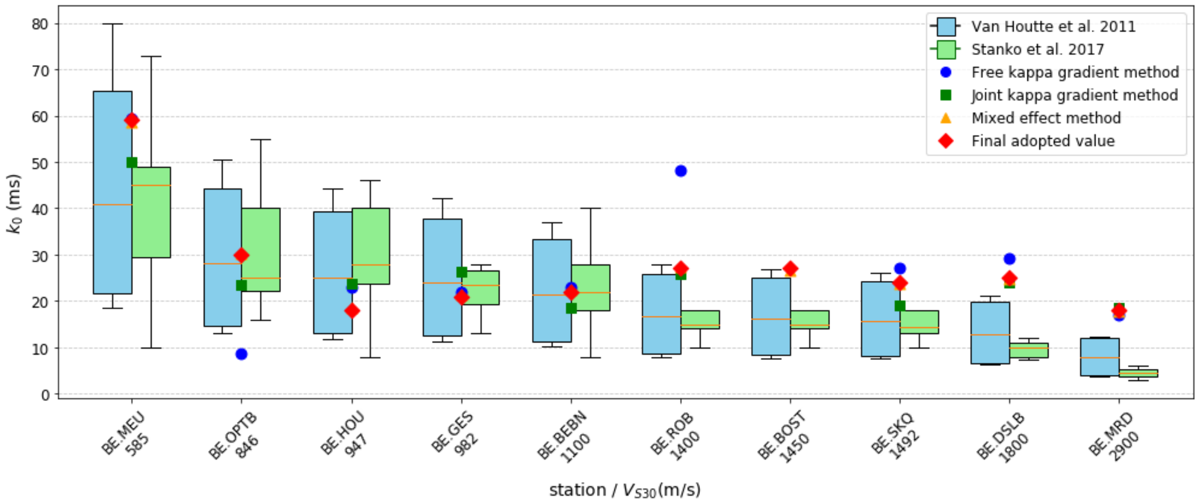


Figure 56. Comparison of κ_0 estimates for stations with measured V_{S30} and sufficient data for reliable estimation. The blue and green boxes show the ranges reported by Van Houtte et al.(2011) and Stanko et al. (2017), respectively, for each station. The boxes represent the interquartile range (IQR, 25th–75th percentile) and the whiskers indicate the full range of reported values. The orange line shows the median value. The markers indicate method-specific estimates: blue circles for the FKG method, green squares for the JKG method, and orange triangles for the MER method; the red diamonds represent the final adopted κ_0 value for each station. Stations are ordered by increasing Vs_{30} value

4.2.1.5. Comparison with previous studies in Belgium

Previous attenuation studies in different crustal domains in Belgium have highlighted notable regional variations. These studies were mainly based on macroseismic data and are therefore not directly comparable to the present results in terms of resolution and quantity. In the ABM crustal domain, anisotropic attenuation has been reported, with faster N–S decay and slower E–W decay (Van Noten et al., 2017). However, travel paths in our study are mostly E-W and do not allow us to corroborate this anisotropy. For the HAIN domain, fast attenuation for shallow events has been observed, which is attributed to low-velocity coal layers and a highly fractured crust (Camelbeeck et al., 2022). Our findings confirm that κ_0 values in HAIN stations are systematically higher than in other domains. However, these results must be interpreted with caution, as HAIN is poorly covered by the travel paths we used (Figure 50) and few records were available for the vertical component only. The regional kappa gradient could not be determined. Previous studies reported rapid attenuation for low-magnitude events in the RVG, attributed to site effects (Van Noten et al., 2017). In our study, the vertical component in the RVG exhibits the highest κ_r, whereas the horizontal components show the lowest values.

We compared our κ_0 and κ_r values with those from Goraj (2000), who analyzed data from permanent and temporary stations for the period 1985–1992 using a method similar to the FKG method in this study. This earlier work for 65 earthquakes (1056 usable data points) reported whole-path κ trends with distance and depth, revealing that κ generally decreased over the first 50 km distance and increased between 50 and 200 km, in addition to regional differences, with higher κ in HAIN than in ARD, and station-specific anomalies at Membach (BE.MEM). The latter two findings are consistent with our results. In a more direct comparison, we matched results from individual stations in the earlier study to our estimates from the three fitting methods (Figure 57). For κ_0, the JKG and MER methods yield more stable and consistent station estimates than the earlier study, which exhibits greater variability, particularly at BE.LES, BE.KIN, and BE.MEU, similar to the FKG method in our study. For κ_r the same pattern emerges; our JKG and MER approaches produce stable values, while the earlier study includes pronounced outliers (e.g., BE.HUM, BE.VIA) and the FKG method shows frequent negative or near-zero values, indicating instability (e.g. BE.LES, BE.ROB). These results confirm that robust statistical frameworks (JKG or MER) provide more reliable κ parameter estimation across stations and regions.

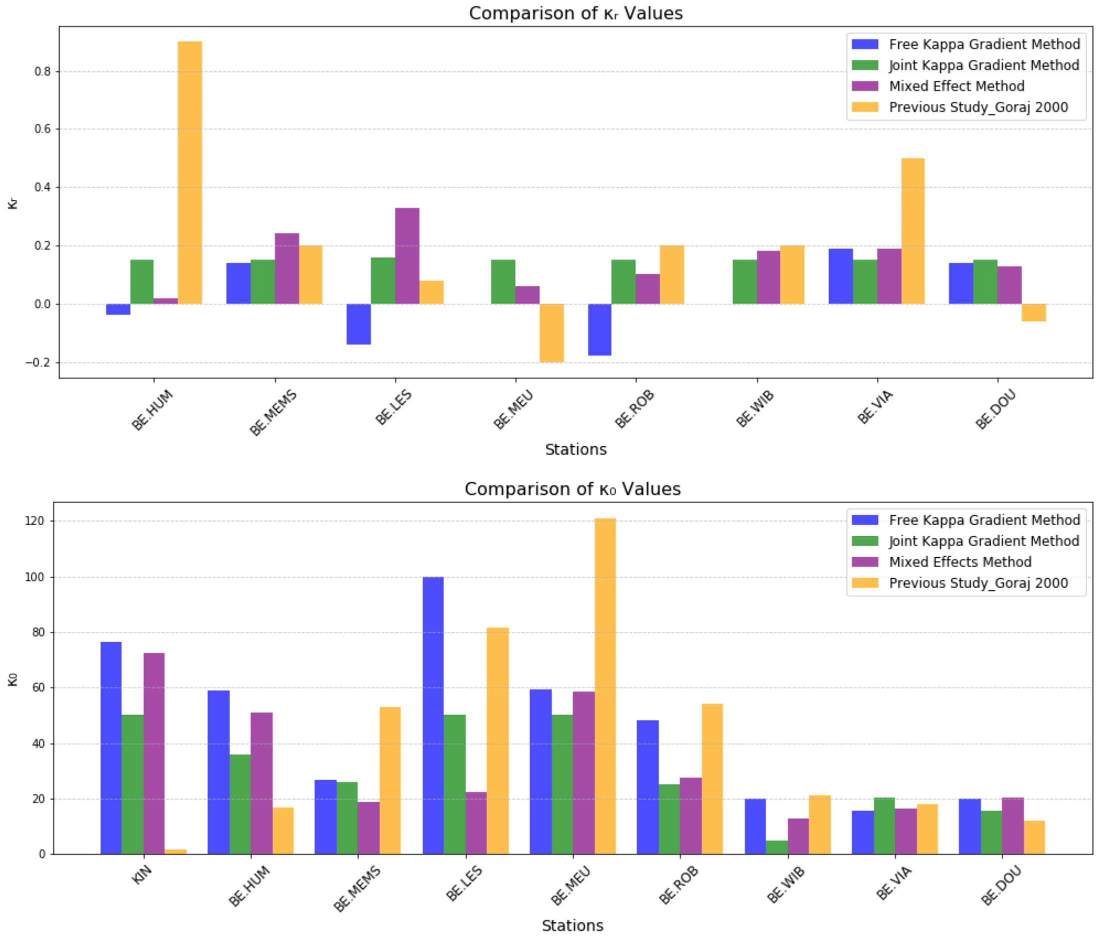


Figure 57 - Comparison of κ_0 (top) and κ_r (bottom) values for similar stations between a previous study (Goraj 2000) and the three fitting methods used in this study

Another study (El Bouch et al., 2002) provided a quantitative assessment of seismic attenuation in Belgium based on vertical-component coda-wave analysis of 86 local earthquakes recorded between 1985 and 1992. They derived frequency-dependent relationships of the form $Q_c = Q_0 f^{\alpha}$ for six crustal regions, with Q_0 ranging from 17 to 31 and α between 1.07 and 1.54 across the 1–32 Hz frequency band. Their result for all stations in eastern Belgium gives $Q_c(26 \text{ Hz}) \approx 1704$, which agrees well with our pooled horizontal Q_i value (1736). However, as the study of El Bouch et al. (2002) is based on the Z component, this value should more appropriately be compared to our vertical Q_i estimate, which is higher, and therefore the agreement with Q_i should be interpreted with caution.

In a probabilistic seismic hazard assessment (PSHA) of the Belgian nuclear power plant sites, Vanneste et al. (2015) conducted a limited study of the κ_0 value for the Cretaceous rocks outcropping at station BE.BEBN for the purpose of host-to-target adjustment of GMMs (Al Atik et al., 2014). They determined whole-path κ on the horizontal components for 68 events in a distance range between 4.2 and 227 km using a similar method as used in this study. Based on simple OLS regression between κ and epicentral distance (equivalent to the FKG method), they obtained $\kappa_0 = 27 \pm 30\ ms$. The kappa gradient was not reported, however. In our study, the final κ_0 value for the horizontal component of BE.BEBN is 22 ms, in good agreement with this earlier estimate.

4.2.1.6. Comparison with neighbouring regions

In this section, we compare our results in terms of site kappa (κ_0) and frequency independent regional quality factor Q_i with those from neighboring regions, mainly from France, where several studies are available.

Site kappa ($\boldsymbol{\kappa_0}$) estimations

In mainland France, Douglas et al. (2010) estimated κ_0 from 263 accelerograms recorded by the French RAP network and found that values are intermediate between western North America (≈40 ms on rock) and eastern North America (≈6 ms on rock). Rock sites in the Alps yield κ_0 ≈ 25 ms, while overall values for the French regions range from 20 to 30 ms, somewhat higher than Switzerland (12–15 ms) and lower than central Europe, mainly Germany (≈50 ms). These values compare well with our Belgian rock stations BE.GES (V_{S30}≈982m/s), BE.BEBN (V_{S30}≈1100m/s), BE. ROB (V_{S30}≈1400m/s) and BE.SKQ (V_{S30}≈1492m/s) where we obtained κ_0 = 21, 22, 27, 24 ms, respectively, indicating similar high-frequency attenuation at (standard) rock sites. In southeastern France (Provence), an area with low to moderate seismicity, Perron et al. (2017) carried out a site-specific study of κ. They showed that site amplification significantly biases κ_0 estimates, even for hard-rock sites, and leads to large discrepancies at stiff-soil sites, where κ_0 values determined following the classical method (Anderson & Hough, 1984) and values determined from displacement spectra (Biasi & Smith, 2001) differed by a factor of two. Despite this, both approaches gave consistent results at the two hard-rock sites (V_{S30} ≈ 2000 m/s), with κ_0 ≈ 30 ms (uncorrected for site amplification). This value is at the higher end of $\kappa_0 - V_{S30}$ correlations and is consistent with the relatively strong attenuation previously reported in Provence and the Alps. These results are also consistent with κ_0 values for our hard rock sites where we obtained κ_0 = 25 ms for BE.DSLB (V_{S30}≈ 1800m/s) and κ_0 = 34 ms for BE.CLA (V_{S30}≈ 2900m/s).

Another study by Ktenidou et al. (2021) used a method based on Fourier transfer functions to determine κ_0 for eight stations in France. Figure 58 shows the $\kappa_0 - V_{S30}$ relationship in log scale for several Belgian stations (the same as in Figure 56) and for different French stations with κ_0 values obtained with various methods. Belgian sites cluster in the 20–30 ms range across V_{S30} ≈ 800-2000 m/s, while French datasets diverge more strongly. The Ktenidou et al. (2021) values show good consistency with Belgium in the V_{S30} ≈ 1000–2000 m/s range. The Douglas et al. (2010) values are also close to the Belgian estimates, particularly for V_{S30} between ≈800–2000 m/s, although two French κ_0 estimates in this range can be considered outliers. Overall, French κ_0 span ≈ 6-30 ms, with higher values in the Alps and Provence, and lower values in the Pyrenees.

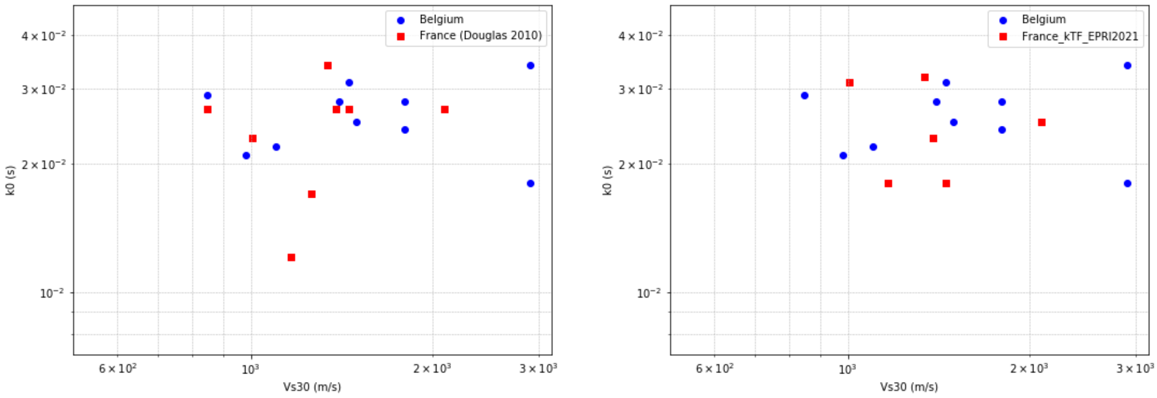


Figure 58. Comparison of V_{S30}–κ_0 relationships between Belgium and two studies in France. Left: Douglas et al., 2010; Right: Ktenidou et al., 2021

Quality factor (Q)

In a study by Drouet et al. (2010), the attenuation properties of the different French regions were quantified using the frequency-dependent quality factor $Q(f) = Q_0 \cdot f^{\alpha}$, where Q_0 represents the quality factor at 1 Hz, and α describes the frequency dependence of attenuation. Based on their inversion results, Q_0 was found to vary across different French regions: 336 ± 15 in the Alps, 1163 ± 247 in the Upper Rhine Graben, and 790 ± 31 in the Pyrenees. These values indicate that attenuation is decreasing from the Alps, the Pyrenees and the Central Massif towards the Rhine Graben. The study by Ktenidou et al. (2021) found very similar values for Q_0 across different French regions: 347 in the Alps, 1374 in the Rhine Graben, and 928 in the Pyrenees. Table summarizes the $Q(f)$ relationships reported for different regions of France by different studies. Since most studies in France concern frequency-dependent Q models, we compare our (frequency-independent) pooled Q_i estimate for Belgium with the values obtained with the French regional $Q(f)$ functions for the frequency range used in our analysis (Table). The median minimum and maximum frequencies (f_1 and f_2) we considered for determining κ are 9 Hz and 27 Hz, respectively, based on all available H-component records. The average Q_i pooled for the H component (1736) lies above the upper limit of the Pyrenees range and below the range estimated for the Upper Rhine Graben, indicating that Belgian crust exhibits lower attenuation than the Alps and Pyrenees, but stronger attenuation than the Upper Rhine Graben.

Study \ Region	Alps	Pyrenees	Upper Rhine Graben	Armorican, Central Massif	Aquitaine, Brabant, Southeast
Ktenidou et al. (2021)	$347f^{0.31}$	$928f^{0.15}$	$1374f^{0.17}$		
	686-964	1290-1521	1996-2406		
Calvet and Margerin (2013)	-	$400\ f^{0.4}$	-	-	-
		963-1495			
Drouet et al. (2010)	$336f^{0.32}$	$790f^{0.15}$	$1163f^{0.19}$	-	-
	679-965	1098-1295	1766-2175		
Campillo and Plantet (1991)	$320\ f^{0.5*}$				
	960-1663				

Table X. Summary of frequency-dependent and frequency-independent quality factor relations, reported for various tectonic regions of France. For each Q relation, the minimum and maximum Q values were calculated using $f_{\min} = 9$Hz and $f_{\max} = 27$Hz, corresponding to the median frequencies of the range employed to estimate kappa for horizontal components in this study. Legend: *Includes intrinsic + scattering effects

In another study, Mayor et al. (2017) estimated Q_i in and around France (including Belgium) from vertical-component coda waves in five frequency bands (1–2, 2–4, 4–8, 8–16, and 16–32 Hz), presenting maps of normalized values Q_i/Q_m, where Q_m is the spatial average of Q_i for each band. Q_i exhibits strong lateral variations, of the order of ±50% around Q_m, with low Q_i in sedimentary basins and higher Q_i in crystalline massifs. The average Q_m values increase with frequency, from about 190 at 1–2 Hz, 400 at 2–4 Hz, 720 at 4–8 Hz, 1,220 at 8–16 Hz, up to 2,250 at 16–32 Hz, with the high-frequency range (16–32 Hz) being most comparable to our study. Since Mayor et al. (2017) used the vertical component, the appropriate comparison is with our pooled Q value for the Z component. Based on their maps, Belgium exhibits a rather high Q_i value (exceeding 2600), which is in good agreement with our pooled value (3086). These maps also show that Q_i in Belgium is higher than in the Alps, Pyrenees, and Upper Rhine Graben, but comparable to central western France.

4.2.1.7. Conclusion

This study presents a comprehensive evaluation of κ, the high-frequency anelastic attenuation parameter, in Belgium, based on the BELSHAKE database. Whole-path κ values were computed from S-wave acceleration Fourier amplitude spectra for both horizontal and vertical components, following the classic Anderson and Hough (1984) approach. The obtained results were first compared with independent estimates of the equivalent t^* attenuation parameter, showing strong correlations and confirming the consistency of the results. These κ values were then decomposed into site-specific components (κ_0) and regional kappa gradients (κ_r) in four distinct crustal domains, allowing for a detailed characterization of both local and regional attenuation behavior. Three fitting methods were tested; the free kappa gradient (FKG), joint kappa gradient (JKG), and mixed effects (MER) methods. While the FKG method is simple and computationally efficient, it is highly sensitive to data availability and cannot always provide accurate estimates. The JKG method offers a more sophisticated approach but can suffer from optimization range issues that limit its effectiveness. In contrast, the MER method is better equipped to handle the complexities of crustal domain data, making it the most dependable option for estimating κ_r for a crustal domain and κ_0 for each station within that domain simultaneously. Comparison with earlier Belgian studies revealed consistency in regional trends and station anomalies, while also demonstrating the advantages of modern statistical approaches in reducing outliers and stabilizing estimates.

Final station-specific κ_0 values are adopted from the MER method and range from $\sim$14 to 59 ms for the H component and from $\sim$12 to 53 ms for the Z component. κ_0-V_{S30} comparisons with global trends indicate that κ_0 values are quite similar for all bedrock sites, both with low (800-1200 m/s), intermediate (1200-1800 m/s) and high (> 1800 m/s) V_{S30} values. This would imply that bedrock with high V_{S30} in Belgium does not show the low κ_0 values typical of hard rock in other regions of the world. Instead, the observed κ_0 values reflect the heterogeneity of the Belgian crust, supporting the asymptotic κ_0 model proposed by Ktenidou et al. (2015).

The regional kappa gradients, κ_r, adopted for the different crustal domains were derived by averaging the results of the JKG and MER methods. They range from 0.13 to 0.17 ms/km for the H component and from 0.05 to 0.1 ms/km for the Z component. Considering S-wave velocities of 3.6 km/s for all crustal domains, the corresponding frequency-independent quality factor (Q_i) ranges from $\sim$1600 to 2100 for the H component and from $\sim$2700 to 3900 for the Z component. Overall, the results are more robust for the H component than for the Z component.

Using pooled data from all crustal domains, the horizontal component shows $\kappa_r \approx$ 0.09 ms/km and $Q_{iH} \approx$ 1736, while the vertical component shows $\kappa_r \approx$ 0.16 ms/km and $Q_{iZ} \approx$ 3086. Although our Q_i estimates for Belgium are frequency-independent, when compared with the frequency-dependent Q models for France over the same frequency range that we used to compute κ_r, our pooled Q_{iH} (1736) lies above the Pyrenees range and within the Upper Rhine Graben range, indicating that the Belgian crust occupies an intermediate position between these two French regions. For the vertical component, studies by Mayor et al. (2017) consistently identify Belgium as a high Q region, with Q_{iZ} exceeding 2600, in agreement with our pooled vertical estimate (3086).

Overall, while this study provides a robust framework for quantifying κ and its decomposition into κ_0 and κ_r across Belgium, certain limitations must be acknowledged. The relatively modest R^2 values highlight that the κ-distance relationship cannot be fully captured by a linear gradient, suggesting that additional factors such as complex scattering and frequency dependence of κ may play an important role. Moreover, the estimated κ_r values are likely most reliable for the ARD domain, which is best covered, whereas coverage in other crustal domains, in particular HAIN, remains partial and requires further data to ensure representativeness. Future work should therefore focus on refining site-specific κ_0 estimates using alternative approaches, such as ambient noise analysis (e.g., Moratto et al., 2025), and on exploring frequency-dependent quality factors through generalized inversion methods (e.g., Oth, 2013), which would enable a more comprehensive description of anelastic attenuation.

These results have been integrated into the BELSHAKE database and will support improved ground-motion modeling and probabilistic seismic hazard assessment in Belgium, enabling more accurate host-to-target adjustments for ground-motion models and better accounting for regional and site-specific attenuation effects.

4.2.2. SELECTION AND EVALUATION OF GMMS

We selected more than 20 ground-motion models to evaluate against the BELSHAKE database. This selection contains updated versions of GMMs applied in site-specific PSHA studies in Belgium during the past decade (Vanneste et al., 2014), GMMs used in recent national PSHA studies in neighbouring countries like France (Drouet et al., 2020), Germany (Grünthal et al., 2018) and the United Kingdom (Mosca et al., 2022; Villani et al., 2019), as well as a number of GMMs developed specifically for induced earthquakes. These GMMs can be organized in different "families", depending on the underlying ground-motion dataset: Europe (n=4), NGA-West (n=6), NGA-East (n=4), Global and Japan (n=2), local GMMs developed for specific countries or regions (n=4), and induced-seismicity GMMs (n=2). It is worth noting that the Douglas et al. (2013) GMM in the latter family actually contains 36 different relations, corresponding to different combinations of stress drop (1, 10 and 100 bar), attenuation quality factor Q (200, 600 and 1800) and high-frequency kappa attenuation factor (0.005, 0.02, 0.04 and 0.06 s). In some national PSHA studies, host-to-target adjustments were applied to the selected GMMs, but these were not considered here, except for the host-to-target adjustments of the NGA-East GMMs of Campbell (2003) and Toro (2002). The latter were developed for site-specific PSHA studies in Belgium (Vanneste et al., 2014).

We applied the standard versions of these GMMs as implemented in the open-source hazard engine OpenQuake (Pagani et al., 2014; https://github.com/gem/oq-engine). In several cases, different implementations are available, corresponding to different distance metrics (e.g., hypocentral or Joyner-Boore distance), assumed stress drop, whether style of faulting is taken into account, or which horizontal components are used. We tested all available implementations but show only the one giving the best result. For the Kotha et al. (2020) GMM, we considered both the general form and an adapted form for application in the 2020 European Seismic Hazard Model (Weatherill et al., 2020). The latter contains calibrations for five main regions in Europe, but we only considered the default region (0), as most of Belgium is not located in one of the main zones, and no regionalization is defined for PGV anyway. As indicated in Table , some GMMs are not defined for PGV, but all of them are defined for PGA and spectral acceleration.

To analyze the performance of these GMMs, we limited the observed ground-motion values from the BELSHAKE database to those recorded on bedrock (soil class A) and considered only the S-wave window of the records. To compute predicted ground-motion values with the selected GMMs, we consider appropriate site conditions for GMMs containing a site term with Vs_{30}=1100 m/s, corresponding to the median value for Belgian bedrock stations (see §4.1.6), we automatically select the appropriate distance metric and type of horizontal components supported by each GMM, and we take into account the focal mechanism of the associated events (if supported by the GMM), assuming strike-slip faulting when this is not determined.

	Belgium (Vanneste et al., 2014)	**UK (Mosca et al., 2022; Villani et al., 2019)**	**Germany (Grünthal et al., 2018)**	**France (Drouet et al., 2020)**	**Other**
Europe	Akkar et al. (2014) Bindi et al. (2014)	Akkar et al. (2014) Bindi et al. (2014)	Akkar et al. (2014) Bindi et al. (2014) Derras et al. (2014)		Kotha et al. (2020)
NGA-West	Boore & Atkinson (2008)	Boore et al. (2014) Chiou & Youngs (2014) Campbell & Bozorgnia (2014)	Bindi et al. (2017)*	Abrahamson et al. (2014)	
NGA-East	Atkinson & Boore (2006) Campbell (2003, SHARE) * Toro (2002, SHARE) *	Atkinson & Boore (2006)			Yenier & Atkinson (2015)
Japan (crustal)	Zhao et al. (2006)*				
Global	Cauzzi et al. (2015)	Cauzzi et al. (2015)	Cauzzi et al. (2015)	Cauzzi et al. (2015)	
Local	Rietbrock et al. (2013)	Rietbrock et al. (2013) Rietbrock & Edwards (2019)		Ameri et al. (2017)* Drouet & Cotton (2015)	
Induced					Atkinson (2015) Douglas et al. (2013)

Table XI - ***Overview of selected GMMs, based on recent PSHA studies in Belgium (Vanneste et al. 2022) and neighbouring countries, and organized in GMM "families" (depending on regional database). Legend: *Not defined for PGV***

The goodness of fit of the selected GMMs with respect to the BELSHAKE database was determined using the various statistical measures described earlier (§3.2.2). In a first step, we selected records with a minimum magnitude (M_{min}) of 2.5. This value was chosen in order to obtain a sufficient number of records (1468 in this case). To allow comparing the performance of different GMMs, we plot their goodness-of-fit measures in spider diagrams, grouped per GMM family. These spider plots show the relative scores of different GMMs for the different metrics along the axes, normalized such that scores improve outward from the centre of the plot. Figure 59 shows the results for PGA and Figure 60 those for PGV.

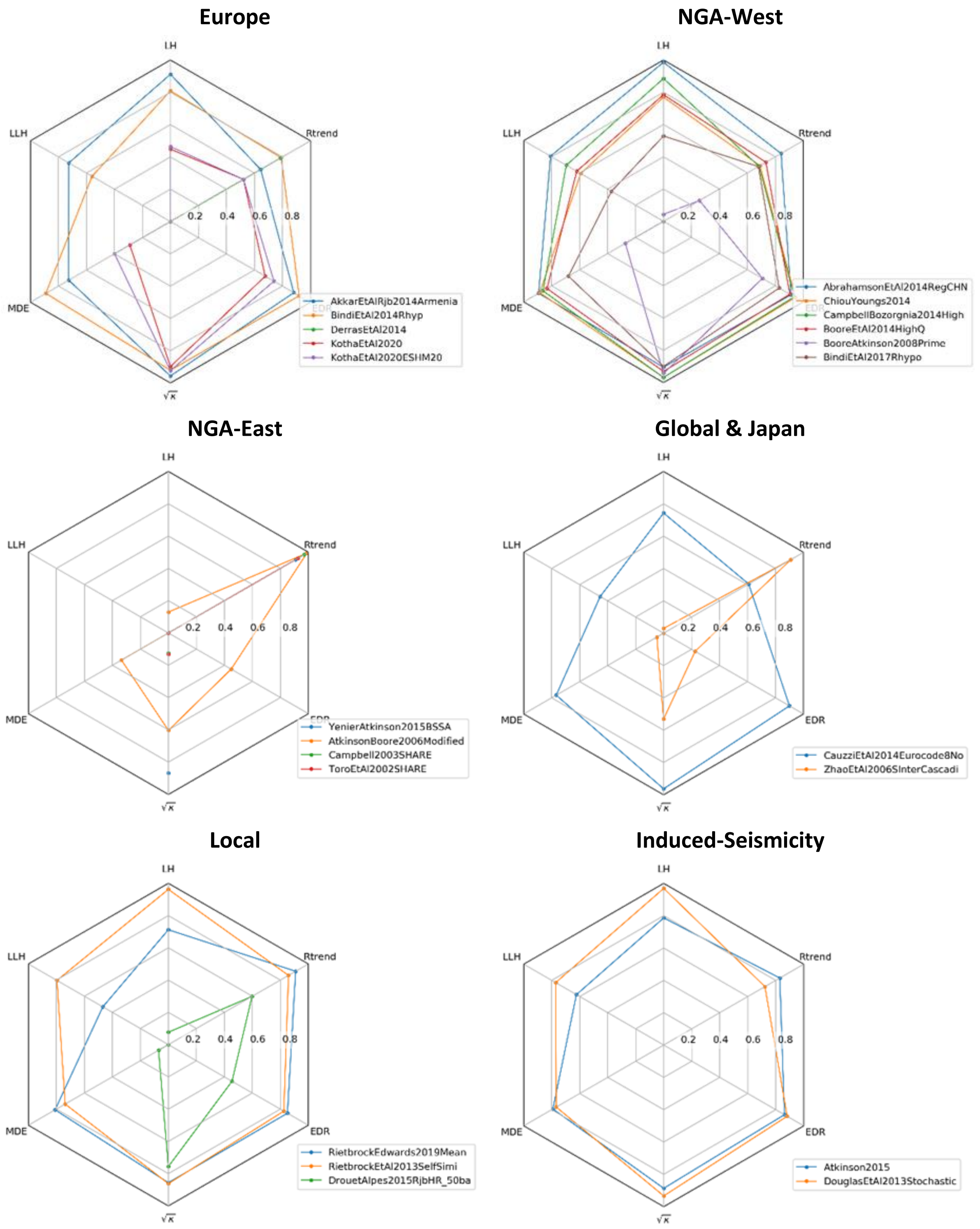


Figure 59 - Spider diagrams showing goodness-of-fit measures for PGA predicted by different GMMs grouped by family, with respect to data from events with $M_W \geq 2.5$ in the BELSHAKE database. Goodness-of-fit measures are normalized to different minimum and maximum values (LH: $|med(LH) \;- 0.5|$ * 2; $\overline{LLH}$: 1.75 – 3.75; MDE: 1 –2; $\sqrt{\kappa}$: 1 – 2.5; EDR: 1 – 4)

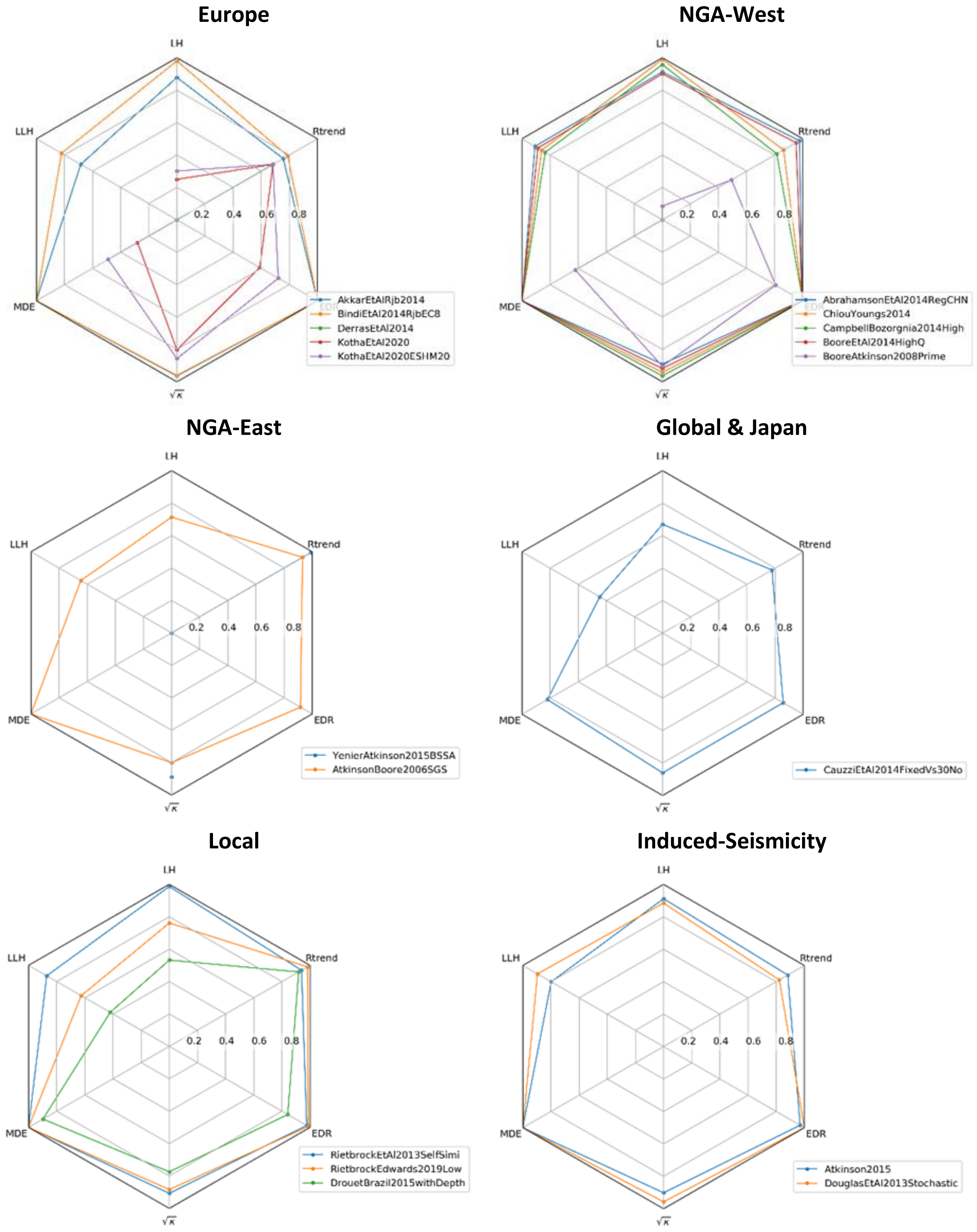


Figure 60- Spider diagrams showing goodness-of-fit measures for PGV predicted by different GMMs grouped by family, with respect to data from events with $M_W \geq 2.5$ in the BELSHAKE database

We evaluated GMM performance for SA at various periods from 0.015 to 1.2 s using several statistical measures, including LH, LLH, EDR, and MBE. While these measures generally provide consistent results, some exhibit fluctuations that make it difficult to clearly identify the best-performing GMM. Considering these issues, and the clearer differentiation provided by MBE, we focus on MBE for the evaluation presented here. Therefore, for simplicity, only MBE-based results for different GMM families are shown in the SA line plots in Figure 61.

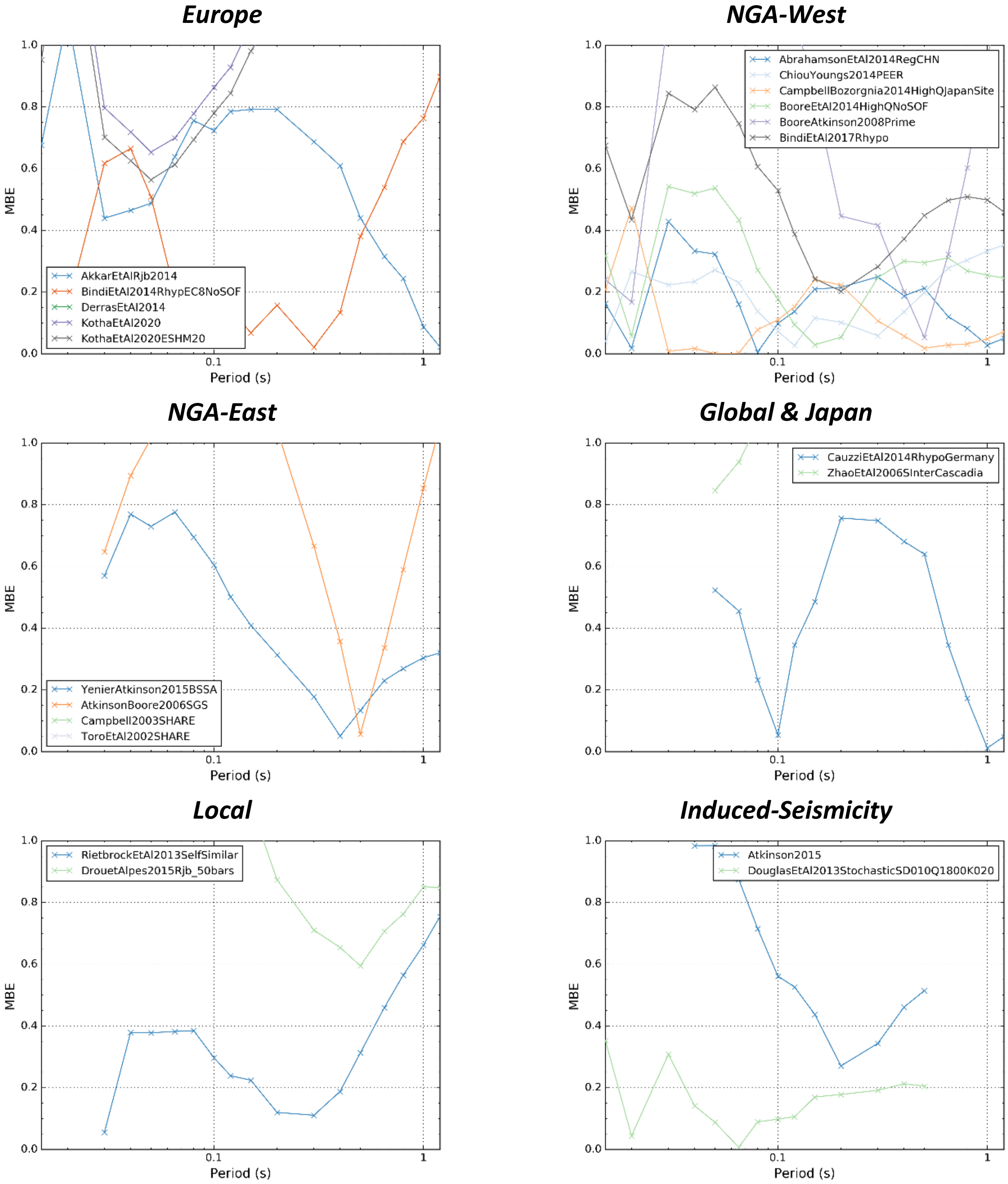


Figure 61. Plots showing Mean Bias Error as a goodness of fit measure for SA predicted by different GMMs grouped by family, with respect to data from events with $M_W \geq 2.5$ in BELSHAKE database

To rank the GMMs, we prioritized EDR as the most important measure, followed by $\sqrt{\kappa}$, with the remaining metrics considered thereafter. The best performing GMMs for PGA, PGV and SA are compared in Figure 62, and ranked based on their overall performance. The selected list remains nearly identical for all three intensity measures. GMMs for different intensity measures remain the same, although their specific versions may vary. Some GMM families have multiple models represented in the selected list, while others are not included.

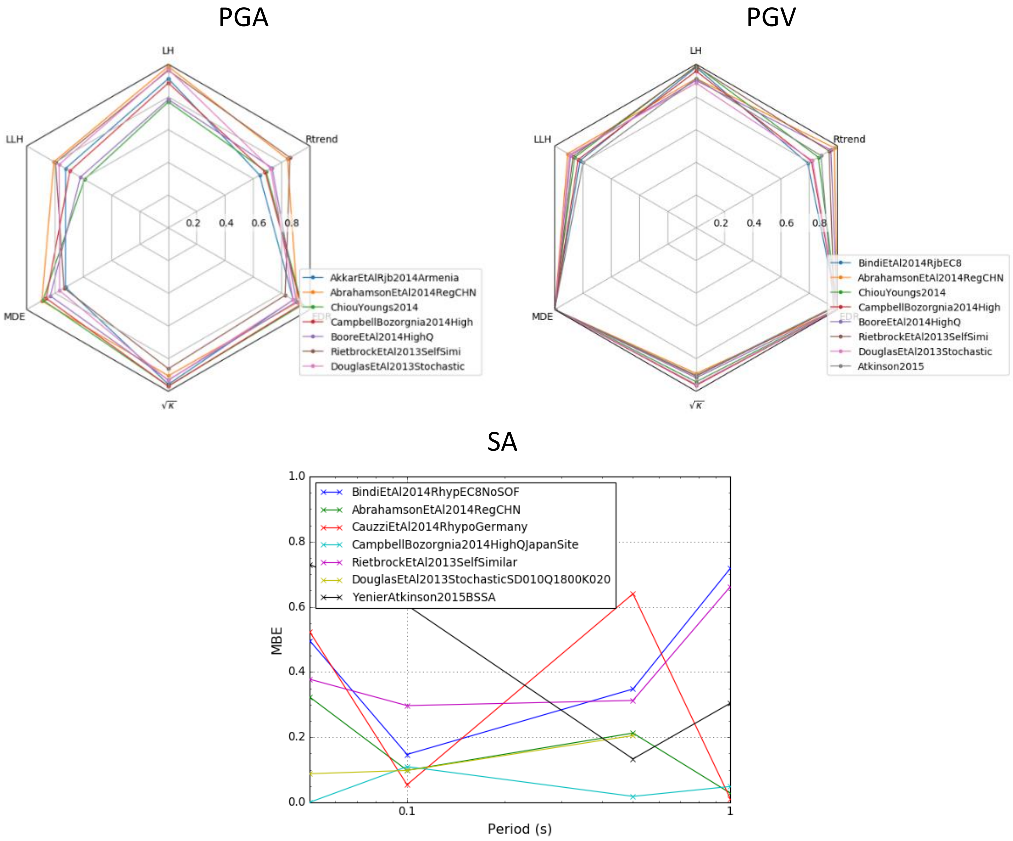


Figure 62. Spider diagrams showing the goodness-of-fit measures for the best-performing GMMs for PGA and PGV, and line plots presenting MBE as the sole goodness-of-fit measure for SA, with respect to data from events with $M_W \geq 2.5$ in the BELSHAKE database

For PGA, all the selected GMMs perform well, with minimal differences in their overall scores. ChiouYoungs2014, CampbellBozorgnia2014, and AbrahamsonEtAl2014 rank the highest in terms of EDR, though the remaining GMMs follow closely behind. RietbrockEdward2013 ranks the lowest for EDR and $\sqrt{\kappa}$, but it has high scores in LH, LLH, and Rtrend. Conversely, ChiouYoungs2014 has the lowest LLH score among the models. Overall, the most consistent GMMs for PGA are AbrahamsonEtAl2014, CampbellBozorgnia2014, and DouglasEtAl2013Stochastic, as they maintain strong performance across multiple key measures.

For PGV, the scores are even closer to one another, indicating a more uniform performance among the selected models. Similar to PGA, the EDR scores for PGV are consistently high across all models. DouglasEtAl2013Stochastic, BooreEtAl2014, and AbrahamsonEtAl2014 achieve the highest EDR scores, along with top rankings in the LLH measure, reinforcing their reliability across different ground motion parameters.

For SA, CampbellBozorgnia2014HighQJapanSite, DouglasEtAl2013Stochastic, and AbrahamsonEtAl2014RegCHN demonstrate the most consistently low MBE across all periods, indicating higher predictive accuracy. In contrast, models such as CauzziEtAl2014RhypoGermany, RietbrockEtAl2013SelfSimilar, BindiEtAl2014RhypoEC8NoSOF, and YenierAtkinson2015BSSA exhibit more variable results, with higher MBE values at certain periods. This reflects greater deviations from observed ground motion and suggests reduced reliability across different periods for these GMMs.

The plots illustrating different goodness-of-fit measures for AbrahamsonEtAl2014RegCHN and DouglasEtAl2013StochasticSD100Q1800K040 as examples are shown in *Figure 63* for PGA and *Figure 64* for PGV. Both GMMs exhibit very good $\sqrt{\kappa}$ scores (close to 1) for both PGA and PGV. In both GMMs, the median LH values are close to 0.5, indicating good LH scores for both PGA and PGV. Their LLH values are lower than the other GMMs, leading to high scores. Additionally, a comparison of the histogram of normalized residuals against a standard normal distribution shows no significant bias. The mean (μ) remains close to zero, while the standard deviation (σ) is close to 1 in most cases, further confirming the robustness of these GMMs in predicting ground motion.

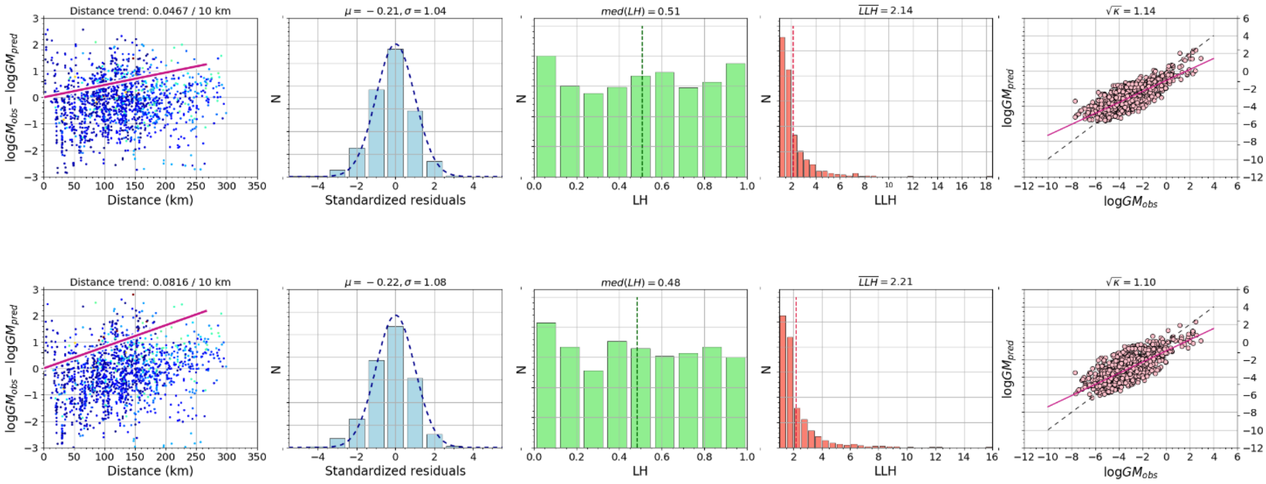


Figure 63. Goodness-of-fit measures for PGA predictions by the top-performing GMMs; Top: AbrahamsonEtAl2014RegCHN, Bottom: DouglasEtAl2013StochasticSD100Q1800K040

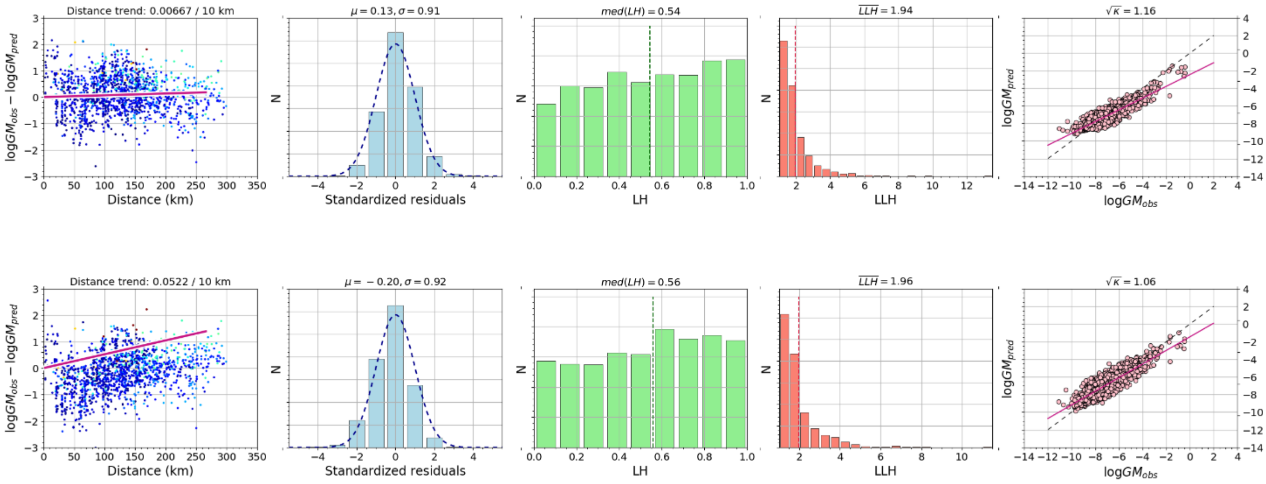
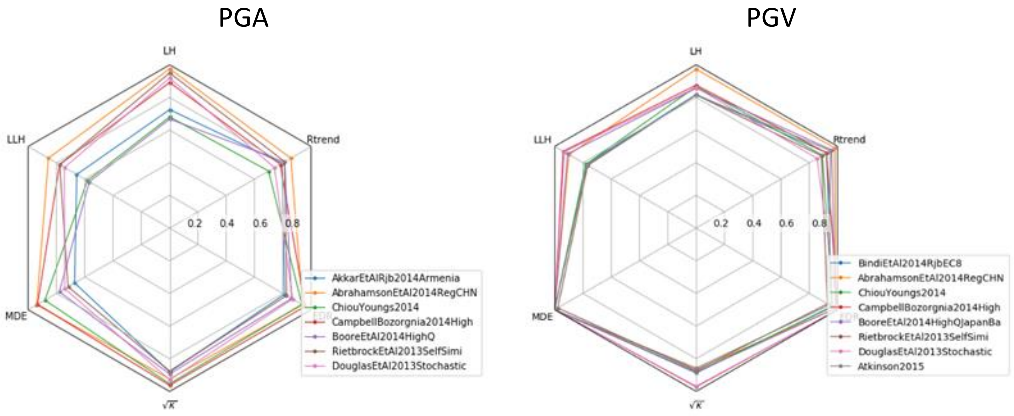


Figure 64. Goodness-of-fit measures for PGV predictions by the top-performing GMMs; Top: AbrahamsonEtAl2014RegCHN, Bottom: DouglasEtAl2013StochasticSD100Q1800K040

In a second step, we considered a higher magnitude threshold ($M_{min} = 3.0$), closer to the valid magnitude range of most GMMs. This reduced the number of records from 1468 to 390. Figure 65 presents the results for PGA, PGV and SA. In general, the LLH scores of some GMMs are reduced for both PGA and PGV. However, other measures such as Rtrend, $\sqrt{\kappa}$, MDE, and EDR remain largely unchanged, and the results for these measures are very similar to those for $M_{min} = 2.5$. For SA, while some GMMs continue to perform well, others show reduced stability or fit with fewer data points. Overall, increasing the magnitude threshold leads to a higher variability of goodness-of-fit scores. This is likely caused by the significantly lower number of test data. By and large, the results, and in particular the ranking order, remain in relatively good agreement with those based on $M_{min} = 2.5$. Thus, we will maintain the selection based on the lower magnitude threshold, as it is based on more data.



SA

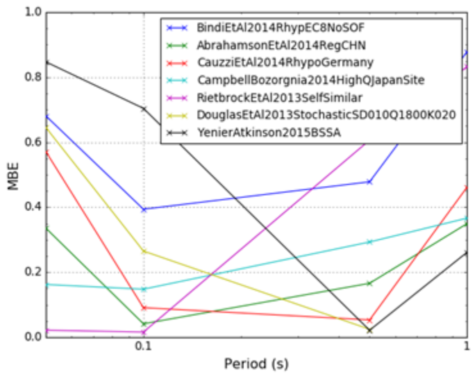


Figure 65. Spider diagrams showing the goodness-of-fit measures for the best-performing GMMs for PGA and PGV, and line plots presenting MBE as the sole goodness-of-fit measure for SA, with respect to data from events with $M_W \geq 3.0$ in the BELSHAKE database

Finally, we extended the test data with ground-motion data from NE France in the RESIF/EPOS-France database (Traversa et al., 2020). This region was selected because of its similar attenuation characteristics to Belgium (Traversa, pers. comm. 2025). This resulted in the addition of 238 records, primarily in the magnitude range of 3.0–4.0, and 5.0 to 5.5, slightly improving the representation of moderate-to-higher magnitudes in the analysis. However, the goodness-of-fit results for the combined dataset were nearly identical to those for the BELSHAKE data alone, both in terms of numerical scores and the relative rankings of GMMs within each family. This consistency is not unexpected, given that the number of selected RESIF records represents only about 1% of the records from the BELSHAKE database, making strong differences in model evaluation unlikely.

In conclusion, we applied multiple statistical measures to quantify the goodness-of-fit between observed ground motions and GMM predictions. The results enabled a comparative assessment of GMM performance and informed the ranking of models based on their predictive capabilities.

Although no single model consistently outperforms all others across all measures and scenarios, a subset of GMMs demonstrate good and consistent performance. Notably, variants of the models of Abrahamson et al. (2014), Campbell & Bozorgnia (2014), Boore et al. (2014), and Douglas et al. (2013) perform well for Belgian bedrock conditions. Interestingly, all of these models have a lower valid magnitude of $M_W = 3.0$, which is close to the magnitude threshold of our data set. Another observation is that three of these models are low-attenuation variants of NGA-West GMMs, whereas no NGA-East GMM has been selected. This is surprising, as the former represent a plate boundary (California), whereas the latter are derived for an intraplate region (Eastern North America), which is more similar to the Belgian setting.

The selected models are recommended as a starting point for application in future seismic hazard assessments in Belgium, particularly as part of logic-tree frameworks to account for epistemic uncertainty. However, it should be kept in mind (Beauval et al., 2012) that rankings based on testing against lower magnitudes may be different from rankings that would be obtained with the magnitude range considered in PSHA ($M_W \geq 4.5$). Furthermore, it should be ensured (Bommer & Scherbaum, 2008) that GMMs selected in PSHA are mutually exclusive (not too similar to each other) and commonly exhaustive (covering the full range of ground motion that is technically defensible).

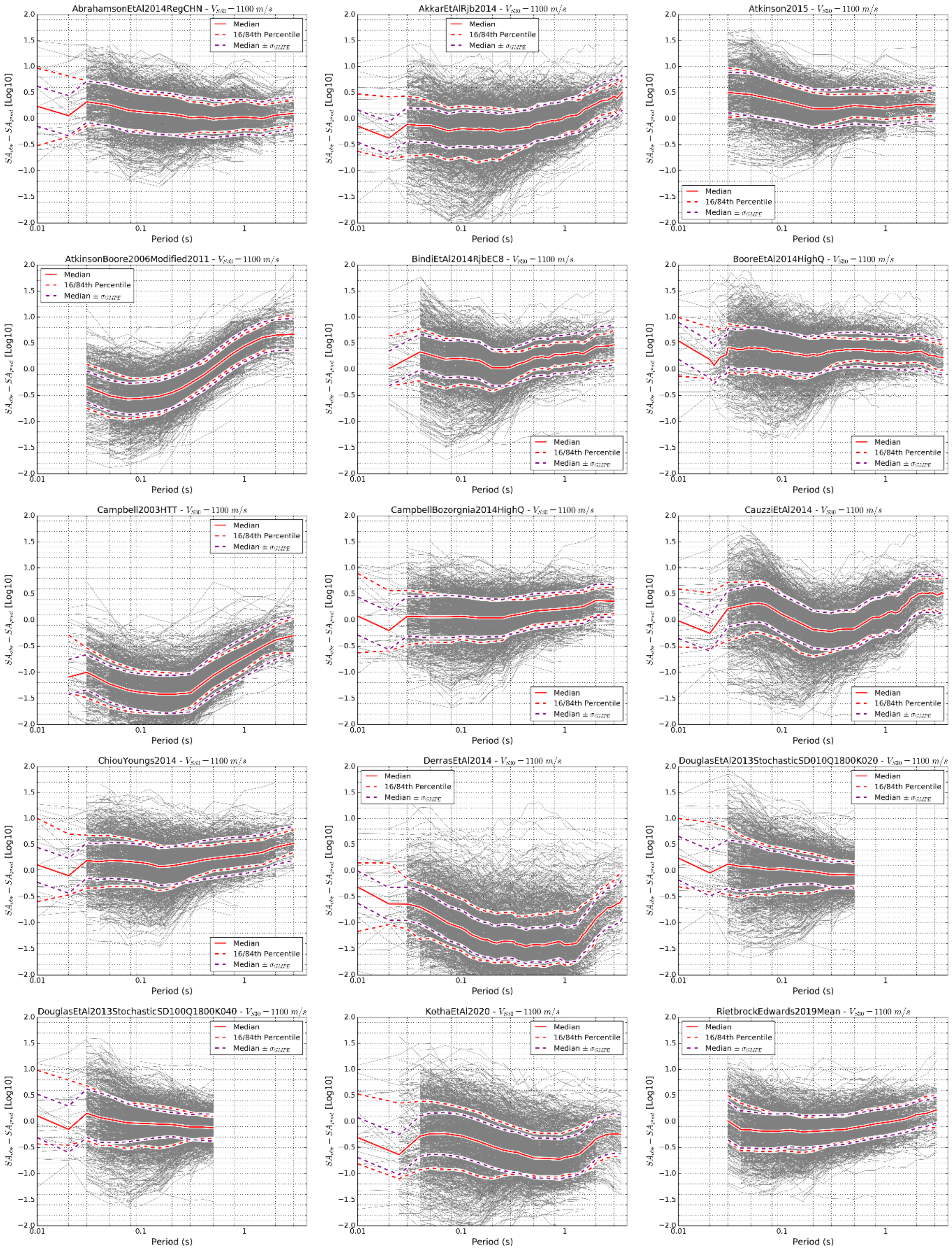
4.2.3. ADJUSTMENT OF GROUND-MOTION MODELS

4.2.3.1. Referenced Empirical Approach

We tested the Referenced Empirical Approach, as applied by Ramadan et al. (2024), against the BELSHAKE database. Similar to the evaluation of GMMs in the previous section (§4.2.2), we restricted the observed ground-motion records to those with $M_W \geq 2.5$ recorded on bedrock (soil class A) and considering only the S-wave window. In addition, we limited the analysis to natural earthquakes.

In a first step, we calculated total spectral residuals with respect to several GMMs, mainly those that showed good performance based on statistical tests, extended with some GMMs of particular interest. Similar to the previous section, we consider appropriate site conditions for GMMs containing a site term with Vs_{30}=1100 m/s, corresponding to the median value for Belgian bedrock stations (see §4.1.6), and automatically select the appropriate distance metric and type of horizontal components supported by each GMM.

The results are shown in Figure 66, and allow evaluating which models yield the smallest mean bias and dispersion, and the best agreement in spectral shape. These models need the least amount of adjustment, and are therefore preferable. Note that in these plots, positive residuals indicate underprediction and negative values overprediction by the GMM.


AbrahamsonEtAl2014RegCHN - $V_{S30} - 1100\ m/s$
AkkarEtAlRjb2014 - $V_{S30} - 1100\ m/s$
Atkinson2015 - $V_{S30} - 1100\ m/s$
AtkinsonBoore2006Modified2011 - $V_{S30} - 1100\ m/s$
BindiEtAl2014RjbEC8 - $V_{S30} - 1100\ m/s$
BooreEtAl2014HighQ - $V_{S30} - 1100\ m/s$
Campbell2003HTT - $V_{S30} - 1100\ m/s$
CampbellBozorgnia2014HighQ - $V_{S30} - 1100\ m/s$
CauzziEtAl2014 - $V_{S30} - 1100\ m/s$
ChiouYoungs2014 - $V_{S30} - 1100\ m/s$
DerrasEtAl2014 - $V_{S30} - 1100\ m/s$
DouglasEtAl2013StochasticSD010Q1800K020 - $V_{S30} - 1100\ m/s$
DouglasEtAl2013StochasticSD100Q1800K040 - $V_{S30} - 1100\ m/s$
KothaEtAl2020 - $V_{S30} - 1100\ m/s$
RietbrockEdwards2019Mean - $V_{S30} - 1100\ m/s$
Median
16/84th Percentile
Median $\pm\ \sigma_{GMPE}$
$SA_{obs} - SA_{pred}$ [Log10]
Period (s)
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
0.01
0.1
1

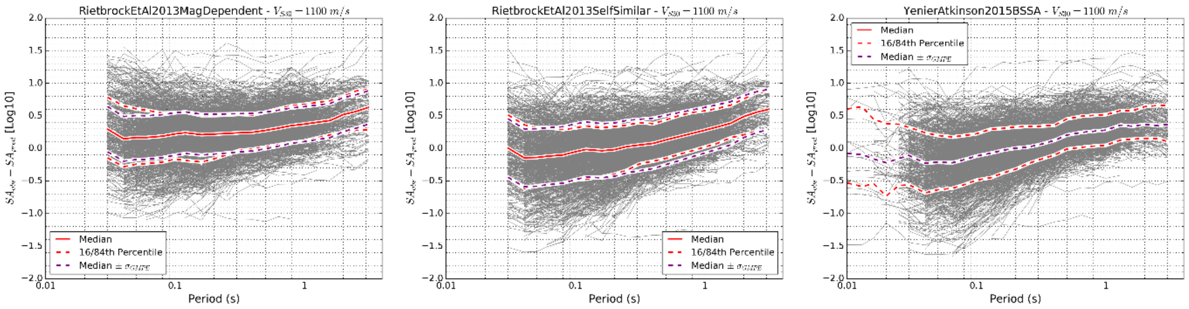


Figure 66 – Spectral residuals between observed ground motion data in the BELSHAKE database and different GMMs. Grey lines correspond to individual records. Red lines indicate the median and 16th and 84th percentiles of the residuals. Purple lines correspond to the median ± 1 standard deviation of the GMM

Some observations:

- AbrahamsonEtAl2014RegCHN: slight underprediction, particularly at short periods; residual dispersion in good agreement with standard deviation of GMM;
- AkkarEtAlRjb2014: very slight overprediction at short periods, changing to underprediction at longer periods; dispersion of residuals higher than GMM sigma at short to intermediate periods;
- Atkinson2015: slight underprediction, becoming stronger at short periods; residual dispersion in agreement with GMM sigma at short periods, becoming more narrow at longer periods;
- AtkinsonBoore2006Modified2011: strong overprediction at short periods, strong underprediction at long periods, pointing to strongly different spectral shape;
- BindiEtAl2014RjbEC8: slight underprediction at short periods, becoming stronger for longer periods; residual dispersion in agreement with GMM sigma at short periods, becoming more narrow at longer periods;
- BooreEtAl2014HighQ: moderate underprediction, which is similar at all spectral periods, indicating very similar spectral shape; residual dispersion in agreement with GMM sigma at short periods, becoming more narrow at longer periods;
- Campbell2003HTT: strong overprediction, becoming less at longer periods;
- CampbellBozorgnia2014HighQ: very good match at short periods, moderate underprediction at longer periods; residual dispersion in good agreement with GMM sigma;
- CauzziEtAl2014: slight underprediction at short periods, slight overprediction at intermediate periods, and again slight underprediction at longer periods, pointing to differences in spectral shape; dispersion of residuals higher than GMM sigma at short to intermediate periods;
- ChiouYoungs2014: slight underprediction at short periods, becoming stronger for longer periods; residual dispersion larger than GMM sigma at short periods, smaller than GMM sigma at long periods;
- DerrasEtAl2014: very strong overprediction, becoming less at short and long periods; dispersion of residuals larger than GMM sigma at all but the longest periods;
- DouglasEtAl2013StochasticSD010Q1800K020: close to zero bias, but only available up to 0.5 s! Residual dispersion larger than GMM sigma;

- DouglasEtAl2013StochasticSD100Q1800K040: close to zero bias, but only available up to 0.5s! Residual dispersion closer to GMM sigma;
- KothaEtAl2020: strong overprediction at all periods, but variable (strongest at intermediate periods); residual dispersion larger than GMM sigma;
- KothaEtAl2020ESHM20: shape similar to KothaEtAl2020, but less strong overprediction;
- RietbrockEdwards2019Mean: slight overprediction at short periods, slight underprediction at long periods, but overall close to zero bias; residual dispersion comparable to GMM sigma;
- RietbrockEtAl2013MagDependent: slightly overpredicting at short periods, becoming stronger for longer periods; residual dispersion comparable to GMM sigma;
- RietbrockEtAl2013SelfSimilar: close to zero bias at short periods, but significantly underpredicting at intermediate and longer periods, resulting in a stronger difference in spectral shape than RietbrockEtAl2013MagDependent; residual dispersion comparable to GMM sigma;
- YenierAtkinson2015BSSA: close to zero bias at short periods, but significantly underpredicting at intermediate and longer periods; residual dispersion relatively low, but GMM sigma not available.

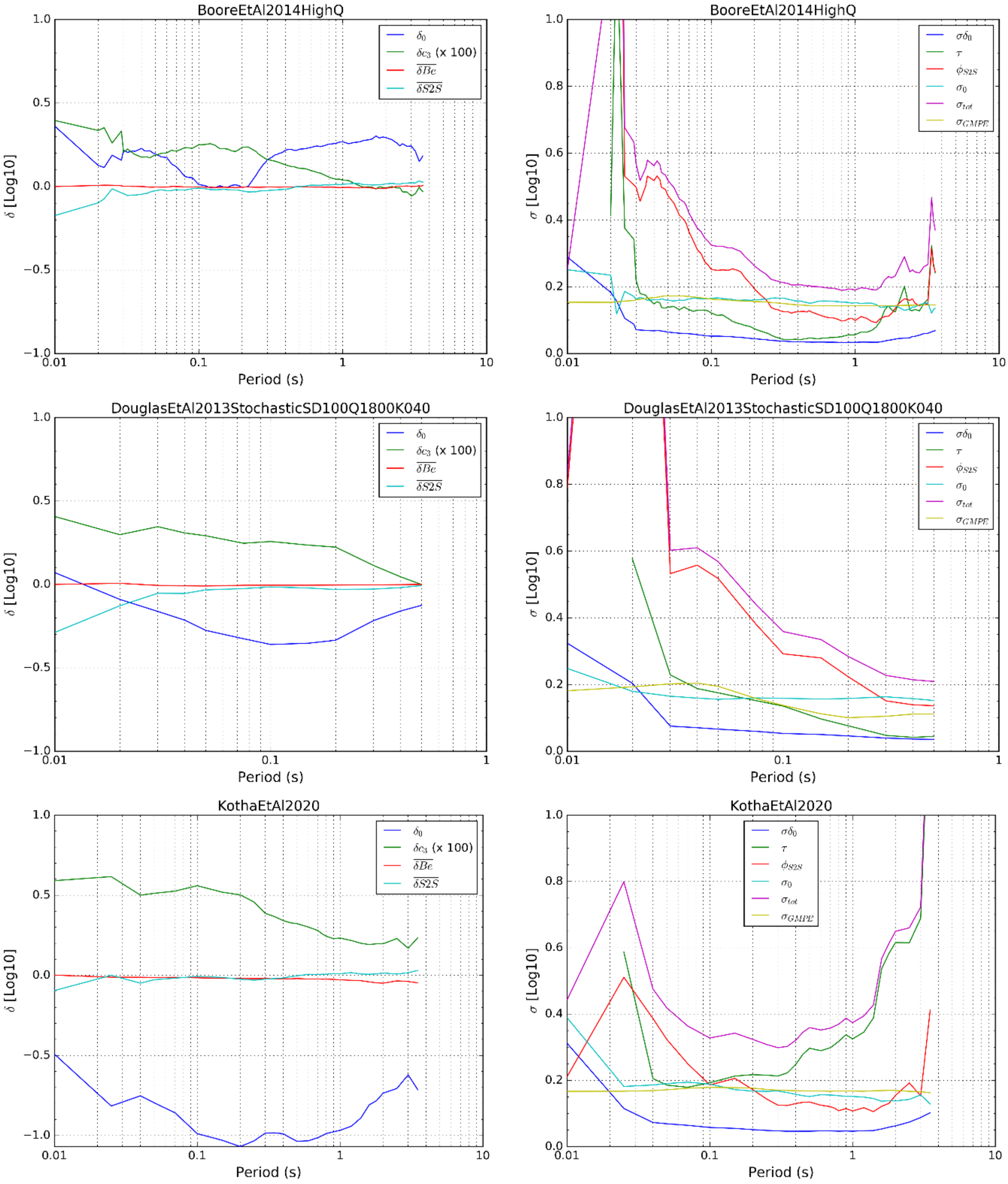
This first analysis shows that only a few GMMs show sufficient agreement with the observed ground motions in the BELSHAKE database, and can be considered for adjustment. Notably, low-attenuation (high-Q) variants of several NGA-West models show a good match, whereas NGA-East models strongly overpredict, even after host-to-target adjustment, or have a very different spectral shape. Some European models perform well too, but others are more variable. Based on this initial test and also taking into account the observed dispersion, we narrow the selection for residual decomposition and calculation of adjustment factors down to the following models:

- BooreEtAl2014HighQ;
- DouglasEtAl2013StochasticSD100Q1800K040 (preferred over the other tested variant because of smaller residual dispersion);
- RietbrockEdwards2019Mean

AbrahamsonEtAl2014RegCHN and CampbellBozorgnia2014HighQ fit relatively well too, but these models use multiple distance metrics and include terms related to fault geometry (hanging-wall term, depth to rupture, …), which is probably too complex for our purposes. We also tested KothaEtAl2020, although it shows a large discrepancy with the observed data, because this model is also used in the second adjustment method (see §4.2.3.2). For these calculations, we limited the data to response spectra for events recorded by at least 5 stations, and for stations that recorded at least 5 events.

Figure 67 shows the calculated adjustment factors and uncertainty components (see §3.2.3.1) calculated for each GMM as a function of spectral period. It is clear that the RietbrockEdwards2019Mean GMM needs the smallest mean bias adjustment and distance correction of the tested models, whereas KothaEtAl2020 needs the largest corrections. The two other models need slightly larger mean bias adjustment, but their distance correction is larger, though in both cases decreasing with spectral period. These distance corrections are positive, indicating that attenuation in Belgium is “slower” than in these models.

On the other hand, the uncertainties are relatively large. In most cases, the uncertainties related to site correction (ϕ_{S2S}) are higher at short periods (high frequency), while the uncertainties on the and between-event correction (τ) are higher at long periods (low frequency). These drive the total uncertainty, which approaches that of the original model only between ~0.25 s (4 Hz) and ~2 s (0.5 Hz). The larger uncertainty is most likely due to the low amount of data in the BELSHAKE database compared to the data (stochastic data or empirical databases) used to develop the models. In a few cases, the uncertainties increase even more strongly near the lower and/or upper limits of the period range; the periods where this occurs should be considered as the limits of the reliable range for adjustment.



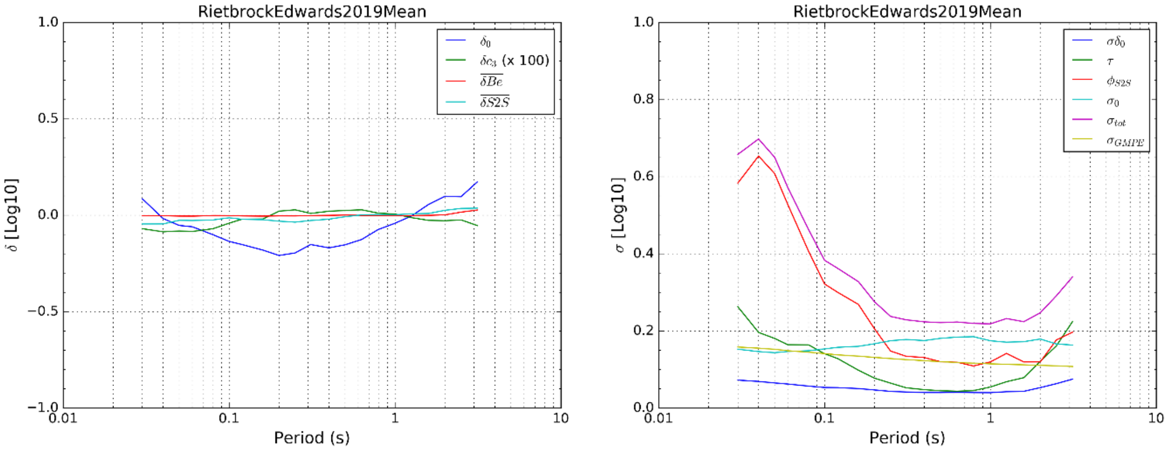
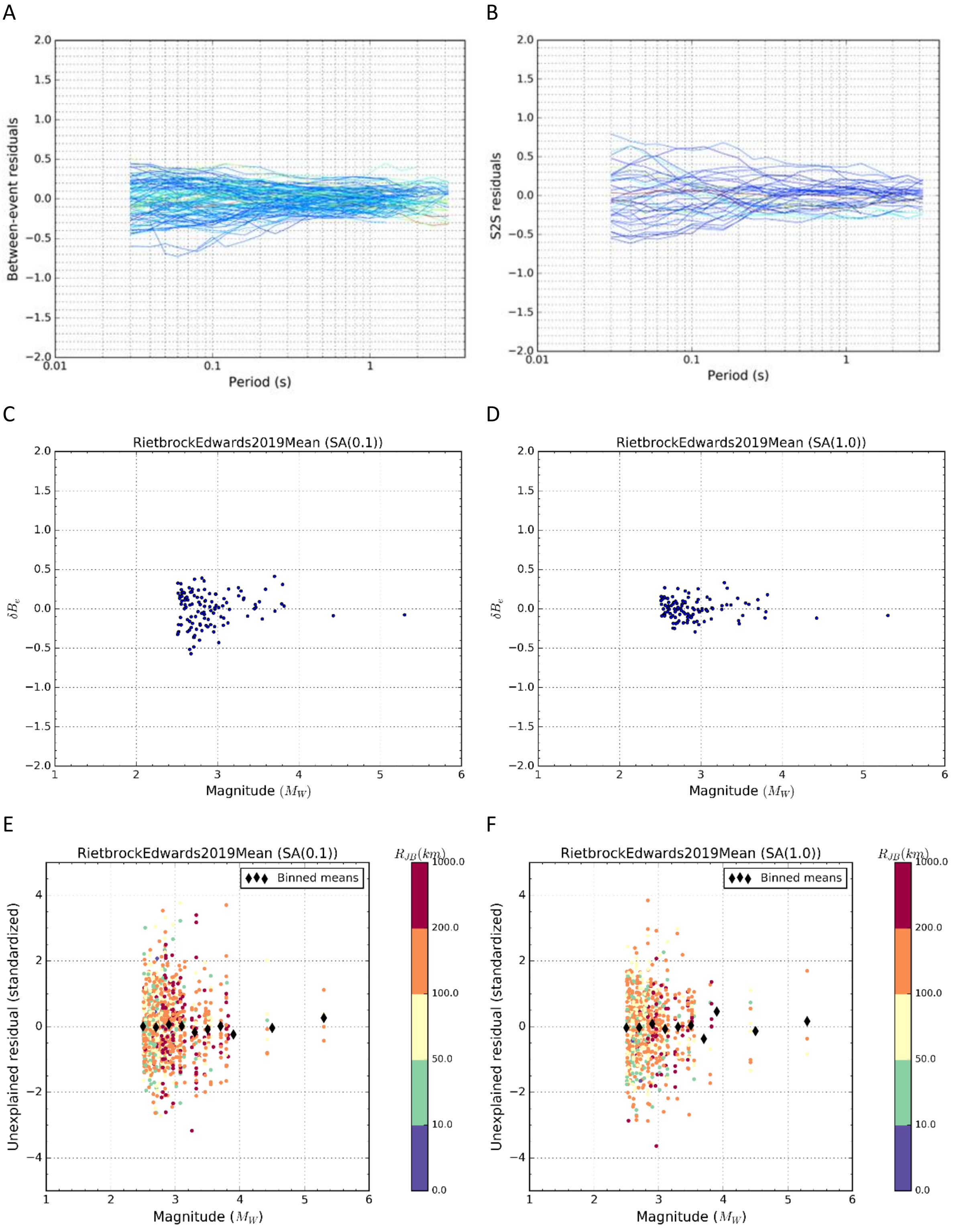


Figure 67 – Spectral adjustment factors (left panel) and uncertainty components (right panel) obtained by residual decomposition for 4 different GMMs (rows). δ_0: median bias adjustment, δc_3: distance-dependent adjustment (multiplied by 100, which corresponds to a distance of 100 km), $\overline{\delta B_e}$: weighted average between-event correction, $\overline{\delta S2S}$: weighted average site-to-site correction, $\sigma\delta_0$: standard deviation related to δ_0, τ: standard deviation related to δB_e, ϕ_{S2S}: standard deviation related to $\delta S2S$; σ_0: standard deviation of leftover residuals; σ_{tot}: total standard deviation of the adjustment; σ_{GMPE}: standard deviation associated with GMM

We illustrate more detailed results for the best performing model, the RietbrockEdwards2019Mean GMM. In Figure 68, the top panel shows between-event and station-to-station residuals versus spectral period, indicating a larger spread of the latter at short periods. The second panel shows between-event residuals versus magnitude for two spectral periods, T=0.1 s (left) and T=1 s(right). No obvious trend with magnitude can be observed, regardless of period, indicating that the magnitude scaling in the BELSHAKE database is well captured by the model. This can also be inferred for the other GMMs, except for KothaEtAl2020, which yields between-event residuals showing a strong magnitude bias between $M_W = 2.0$ and $M_W \cong 4.0$, beyond which residuals are stable (Figure 69). This is in agreement with findings by the author (Kotha, pers. comm. 2025) that the model behaves poorly for lower magnitudes (for which it is not calibrated). The third panel shows unexplained residuals versus magnitude. No obvious trend can be seen, for any spectral period nor for any GMM, including KothaEtAl2020. The fourth panel shows unexplained residuals versus distance. For most GMMs, coherent deviations can be seen, which are stronger for PGV and PGA. We attribute this to deviations in the geometrical spreading assumed in the models compared to the situation in Belgium (see §4.3.1). Interestingly, RietbrockEdwards2019Mean is the only tested model that incorporates a segmented spreading model rather than a simple powerlaw.


A
Between-event residuals
Period (s)
B
S2S residuals
Period (s)
C
RietbrockEdwards2019Mean (SA(0.1))
δB_e
Magnitude (M_W)
D
RietbrockEdwards2019Mean (SA(1.0))
δB_e
Magnitude (M_W)
E
RietbrockEdwards2019Mean (SA(0.1))
Binned means
Unexplained residual (standardized)
Magnitude (M_W)
$R_{JB}(km)$
1000.0
200.0
100.0
50.0
10.0
0.0
F
RietbrockEdwards2019Mean (SA(1.0))
Binned means
Unexplained residual (standardized)
Magnitude (M_W)
$R_{JB}(km)$
1000.0
200.0
100.0
50.0
10.0
0.0

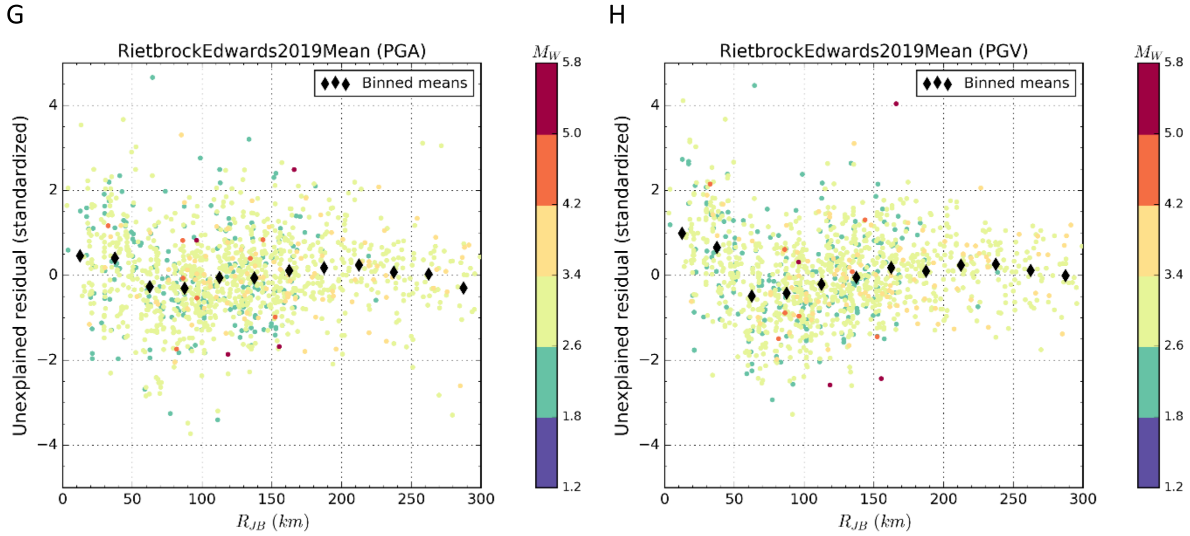


Figure 68 – Residuals obtained for RietbrockEdwards2019Mean GMM with respect to the BELSHAKE database. A: δB_e versus spectral period; B: $\delta S2S$ versus period; C: δB_e for T=0.1 s versus magnitude; D: δB_e for T = 1 s versus magnitude; δW_{es} (unexplained residual) for T=0.1 s versus magnitude; δW_{es} for T=1 s versus magnitude; δW_{es} for PGA versus distance; δW_{es} for PGV versus distance. Note that δW_{es} residuals are standardized (= divided by their standard deviation)

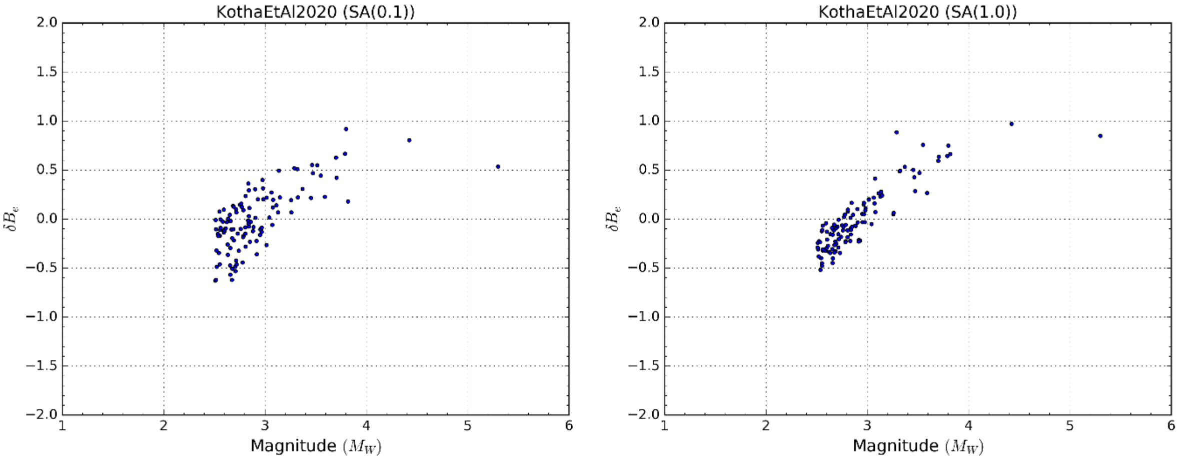


Figure 69 - δB_e versus magnitude obtained with the KothaEtAl2020 GMM. Left: for T=0.1 s; Right: for T=1 s

We evaluate the results of the adjustment by calculating adjusted residuals and comparing ground motions observed for individual earthquakes with the original and adjusted GMMs. The former is shown in Figure 70, demonstrating that the median adjusted residual is effectively centred on the dataset. We also note a slight improvement of the 16th and 84th percentiles of the adjusted residuals, which is more obvious for models where dispersion of the original residuals was larger. In the tested cases, these percentiles lie close to ±1 standard deviation of the GMM. In some cases, they lie inside the $\pm 1\sigma$ range for longer spectral periods, which we attribute to decreasing amount of data. Considering these observations, we think it is justified to maintain the original uncertainty for the adjusted GMM, rather than using the one computed from the fixed and random effects in the adjustment procedure. Otherwise, the larger standard deviation would have a huge impact in PSHA, where the standard deviation is crucial to calculate probabilities of exceedance.

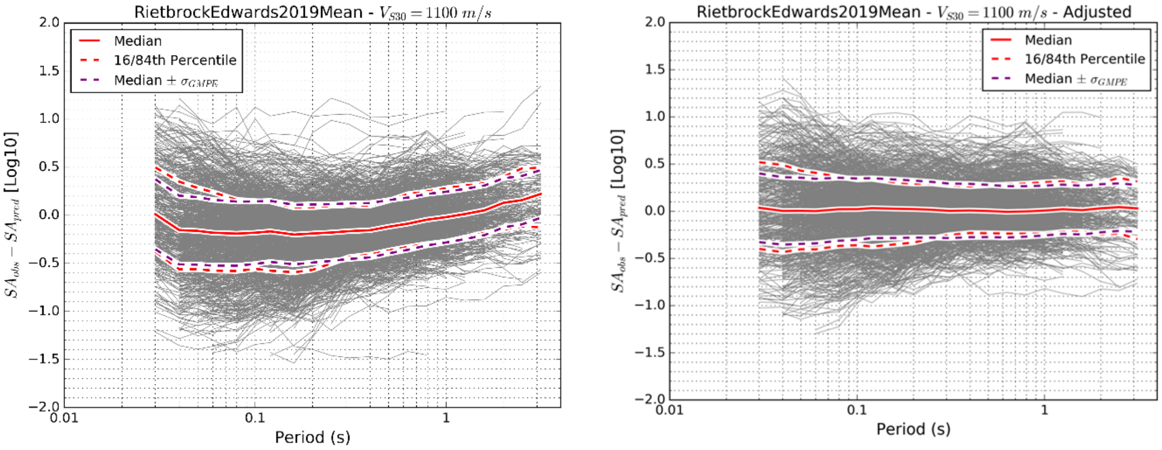
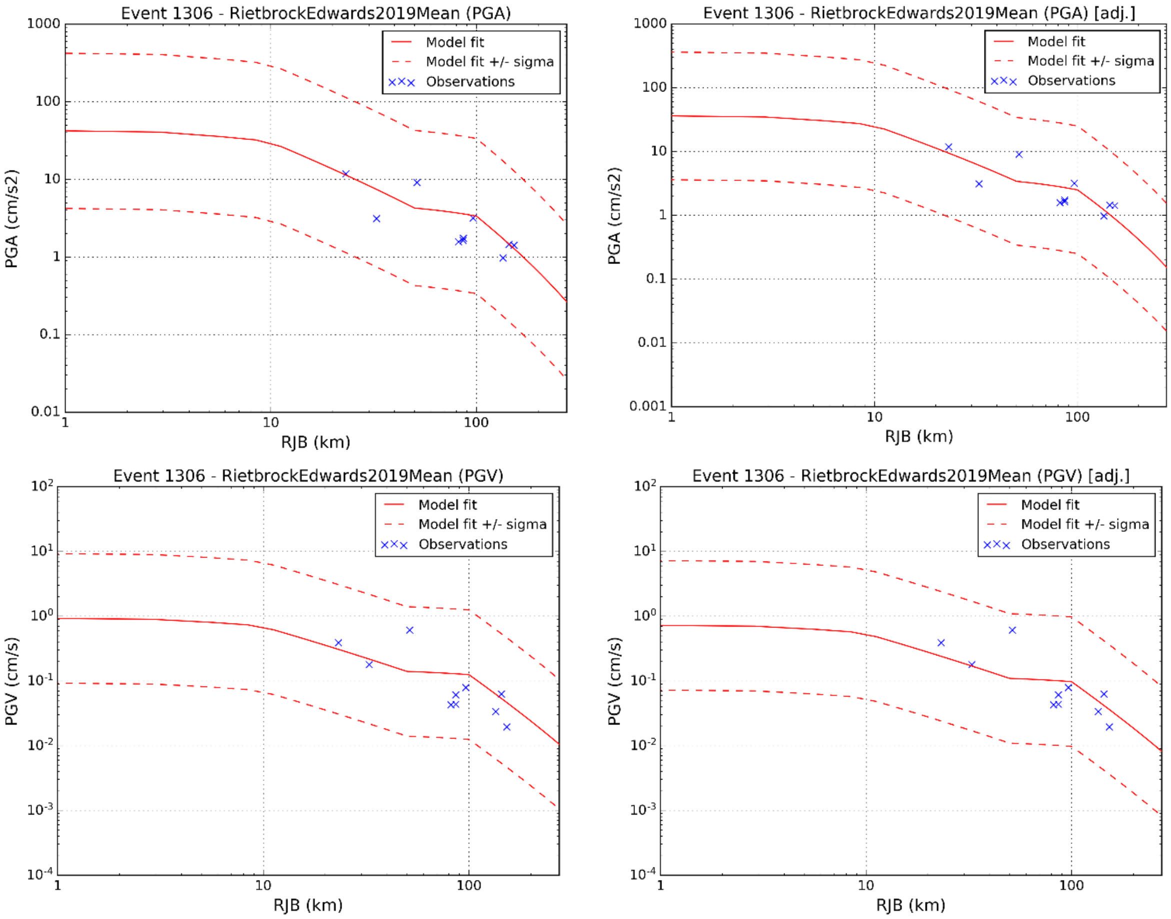


Figure 70 – Comparison of spectral residuals obtained with the original (left) and the adjusted form (right) of the RietbrockEdwards2019Mean GMM

Finally, we compare observed ground-motion data for individual earthquakes with values predicted by the original and adjusted GMMs, as a function of distance. We mainly considered larger events for this evaluation, with $M_W > 3.5$. It is impossible to show all these results. As an example, Figure 71 shows the results for the 2002 Alsdorf earthquake ($M_W = 4.4$). It can be seen that the adjusted model fits the data slightly better, indicating that the adjusted GMM performs well, not only for the lower magnitudes. A similar conclusion can be made for the other tested models, except for KothaEtAl2020 (Figure 72). In this case, the adjusted model clearly performs worse than the original model, at least for events with relatively higher (compared to the database) magnitudes.



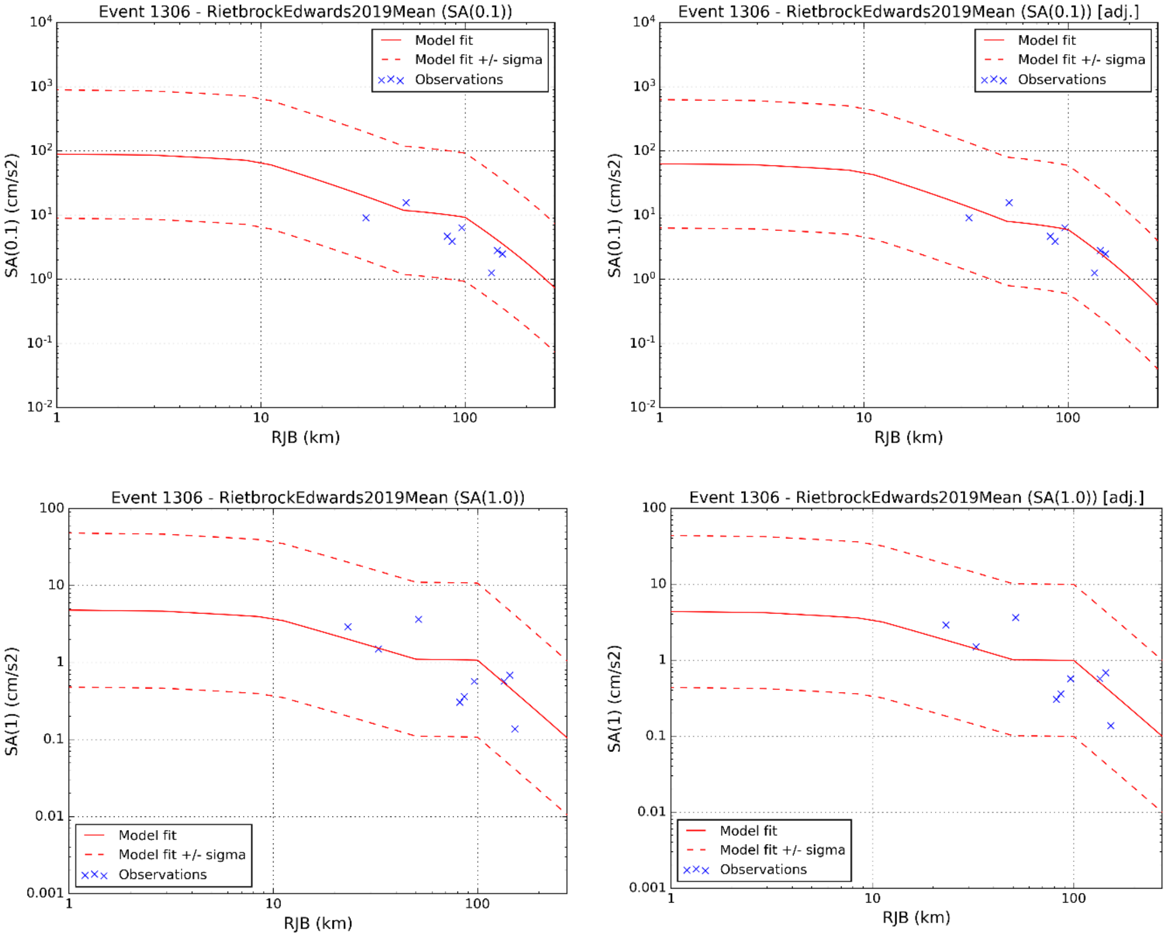


Figure 71 – Comparison of observed ground motions for the 2002 Alsdorf earthquake with values predicted by the original RietbrockEdwards2019 GMM (left column) and the adjusted GMM (right column). Rows correspond to different intensity measures, from top to bottom: PGA, PGV, SA (T=0.1 s) and SA (T=1 s)

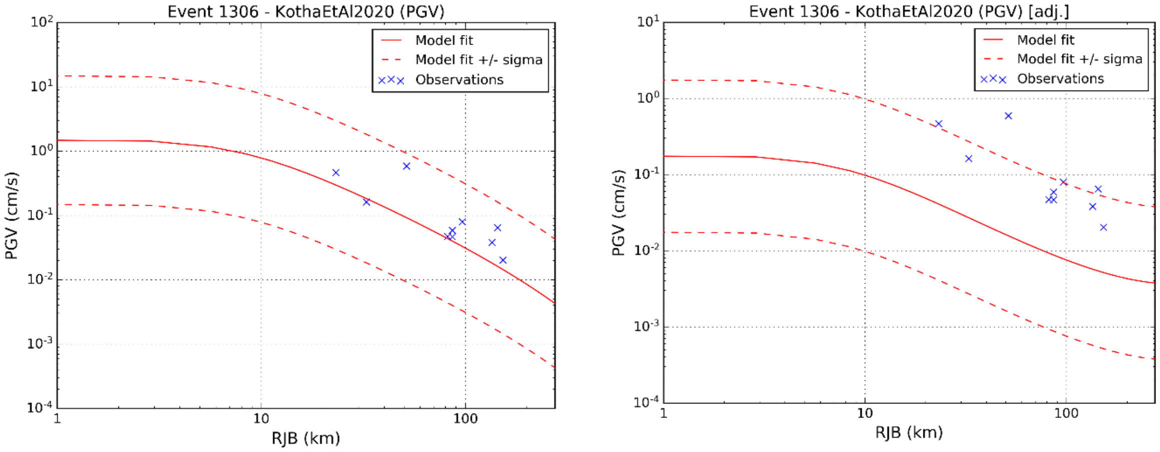


Figure 72 – Same as Figure 71, but for KothaEtAl2020 GMM (PGV only). Left: original model; Right: adjusted model

In conclusion, the Referenced Empirical Approach represents a simple method to adjust ground-motion models to observed data, typically from another region. However, it should be carefully applied to datasets containing few high-magnitude events, as is the case with the BELSHAKE database. In particular, it should be verified that the between-event residuals do not show any trend with magnitude, otherwise the adjusted model will be biased over the magnitude range of interest in PSHA. It is recommended to limit application of this method to models requiring the least amount of adjustment and with spectral shapes that are not too different from those observed. An interesting observation in that respect is that low-attenuation versions of NGA-West GMMs fit our data much better than NGA-East GMMs, despite that the crust in Belgium is classified as stable continental crust. Overall, however, our evaluation indicates that the RietbrockEdwards2019Mean GMM shows the best fit to our database and that it can be reliable adjusted over the entire period range for which it is defined, even though this model was not among the top performing models based on our statistical tests (mainly for PGA and PGV, see §4.2.2). We tentatively conclude that the BooreEtAl2014HighQ and DouglasEtAl2013StochasticSD100Q1800K040 GMMs can be adjusted as well, although both should be restricted for SA to periods above ~0.025 s (f ~40 Hz). The usefulness of the latter model may also be limited, as it is only defined up to 0.5 s (f=2 Hz) and calibrated for magnitudes up to $M_W = 5.0$. In contrast, the KothaEtAl2020 GMM cannot be adjusted using this method.

We implemented the adjusted RietbrockEdwards2019Mean GMM in OpenQuake. To validate it, we compute goodness-of-fit measures similar to §4.2.2. The results in Figure 73 show that the adjusted model indeed has a negligible bias, small distance trend, median LH scores around 0.5, small average LLH scores and small $\sqrt{\kappa}$ scores, pointing to an overall good fit.

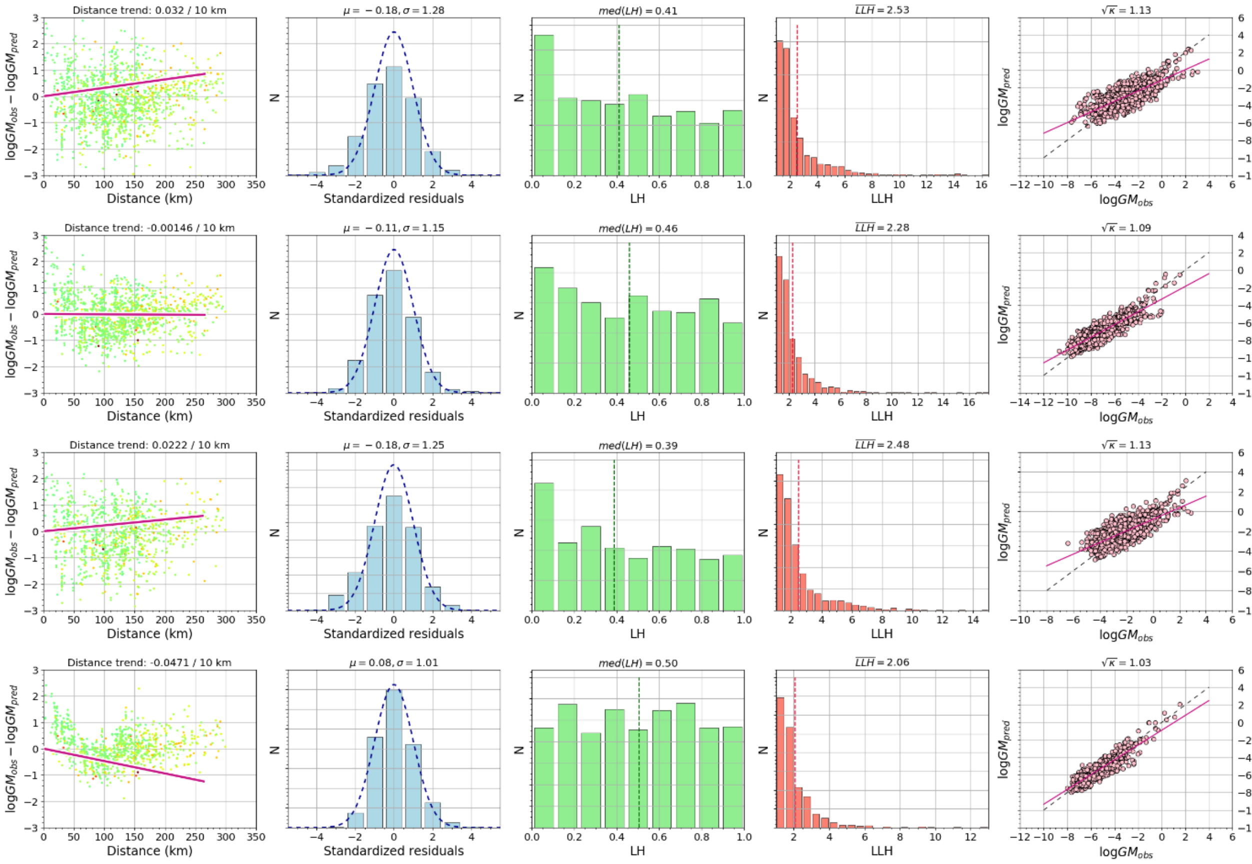


Figure 73 - Goodness-of-fit measures calculated for adjusted RietbrockEdwards2019Mean GMM for different intensity measures, from top to bottom: PGA, PGV, SA (T=0.1 s) and SA (T=1 s)

4.2.3.2. *Bayesian updating*

It has not been possible to apply this more sophisticated method during the project, because we are not sufficiently familiar with the procedure and the software tools needed, and also because informative prior distributions are currently not available. This would require substantial developments, as well as the inclusion of the entire database used to derive the original GMM, for which there was not enough time. However, Georges (2025) has applied this method to adjust the KothaEtAl2020 GMM for France, based on the ESM (Engineering Strong Moton) database (Luzi et al., 2016), complemented with the RESIF/EPOS-France database (Buscetti et al., 2025; Traversa et al., 2020). Thanks to the open-access publication of the BELSHAKE database on Zenodo, she could also include our data and determine adjusted coefficients for Belgium. These (preliminary) coefficients were made available to our project (Georges & Kotha, pers. comm. 2025), which enabled us to implement adaptations of the KothaEtAl2020 GMM in OpenQuake (Pagani et al., 2014), which can be used for PSHA in Belgium.

Below, we evaluate the performance of these adjusted GMMs with respect to the BELSHAKE database, based on the same goodness-of-fit measures that we used to evaluate published GMMs, as well as the GMM that we adjusted with the more simple adjustment method (§4.2.3.1).

Georges (2025) calculated the mean value and standard deviations of all random parameters ($\delta c_{3,r}$, $\delta c_{3,e}$, $\delta S2S_s$ and $\delta L2L_z$, see §3.2.3.2) at each spectral period using records from stations located within Belgium and within 200 km from the national borders. The locality-to-locality regionalization ($\delta L2L_z$) is optional and should preferably be considered as epistemic uncertainty and implemented as two branches in a logic tree. In addition, coefficients were calculated for two different bedrock conditions, for all stations with $V_{S30} \geq 800$ m/s, and for stations with $V_{S30} \geq 1200$ m/s. Together, this results in 4 different models, which we denote as GeorgesEtAl2025BelgiumVs800, GeorgesEtAl2025BelgiumVs800L2L, GeorgesEtAl2025BelgiumVs1200 and GeorgesEtAl2025BelgiumVs1200L2L. Here, we will restrict our evaluation to the relations for $V_{S30} \geq 800$ m/s, as this represents more observations and all other evaluations were based on the same data.

Figure 74 shows spectral residuals obtained with the original and the adjusted GMMs. The plots demonstrate that the median of the adjusted models is much better centered on the dataset. The dispersion of the residuals has also significantly decreased, with the 16th and 84th percentiles close to ±1 standard deviation of the GMM. The model including locality-to-locality regionalization shows a slightly flatter median compared to the model without, but also a slightly larger dispersion of negative residuals in the period range between 0.04 and 0.1 s.

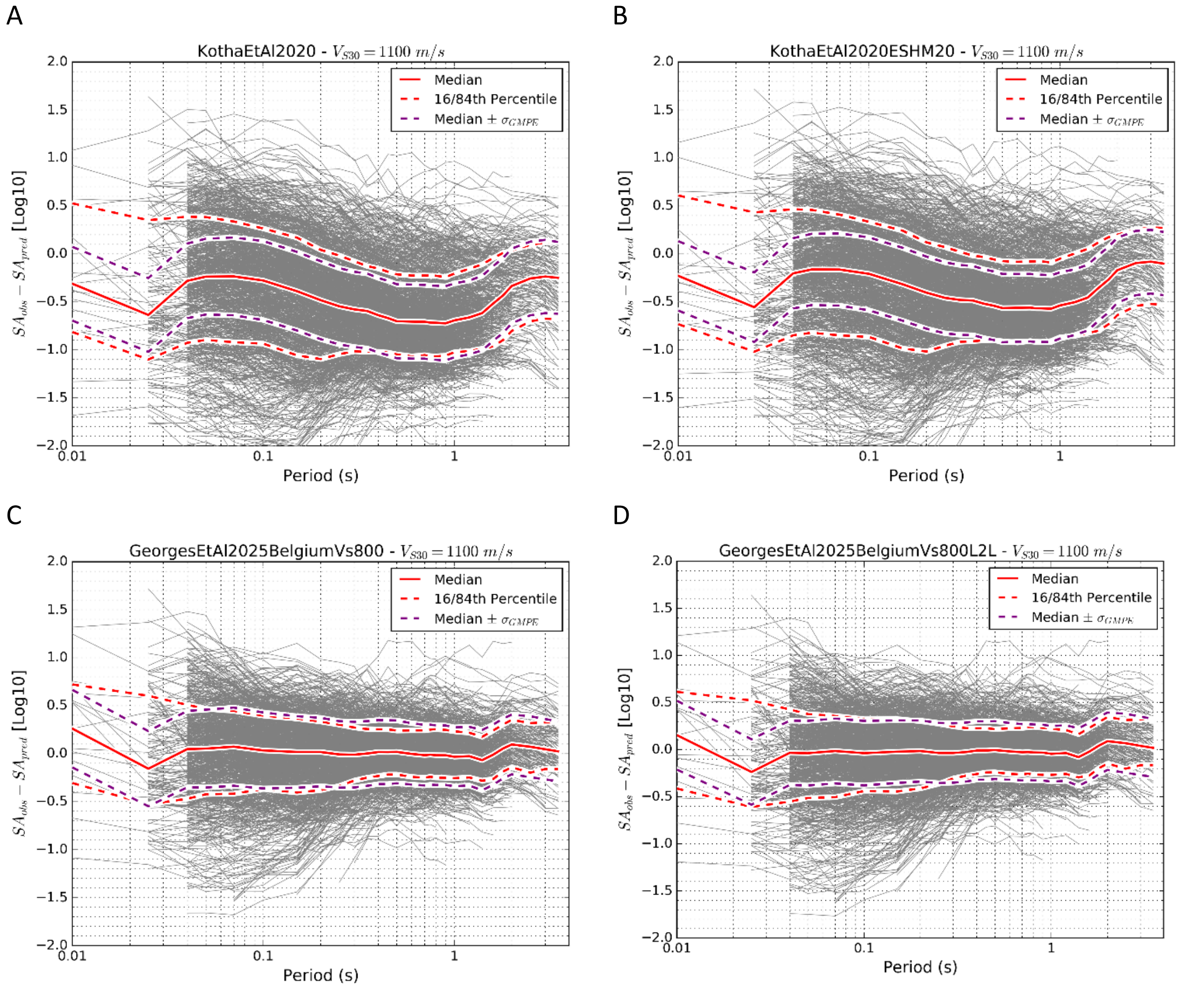


Figure 74 – Comparison of spectral residuals obtained with the original (top row) and the adjusted versions (bottom row) of the KothaEtAl2020 GMM. A: KothaEtAl2020; B: KothaEtAl2020ESHM20 (region 0); C: GeorgesEtAl2025BelgiumVs800; D: GeorgesEtAl2025BelgiumVs800L2L

The spider plots in Figure 75 show how the goodness-of-fit scores of the adjusted models have dramatically improved compared to the original GMM. Based on this comparison, the model without locality-to-locality regionalization seems to perform slightly better than the model with regionalization.

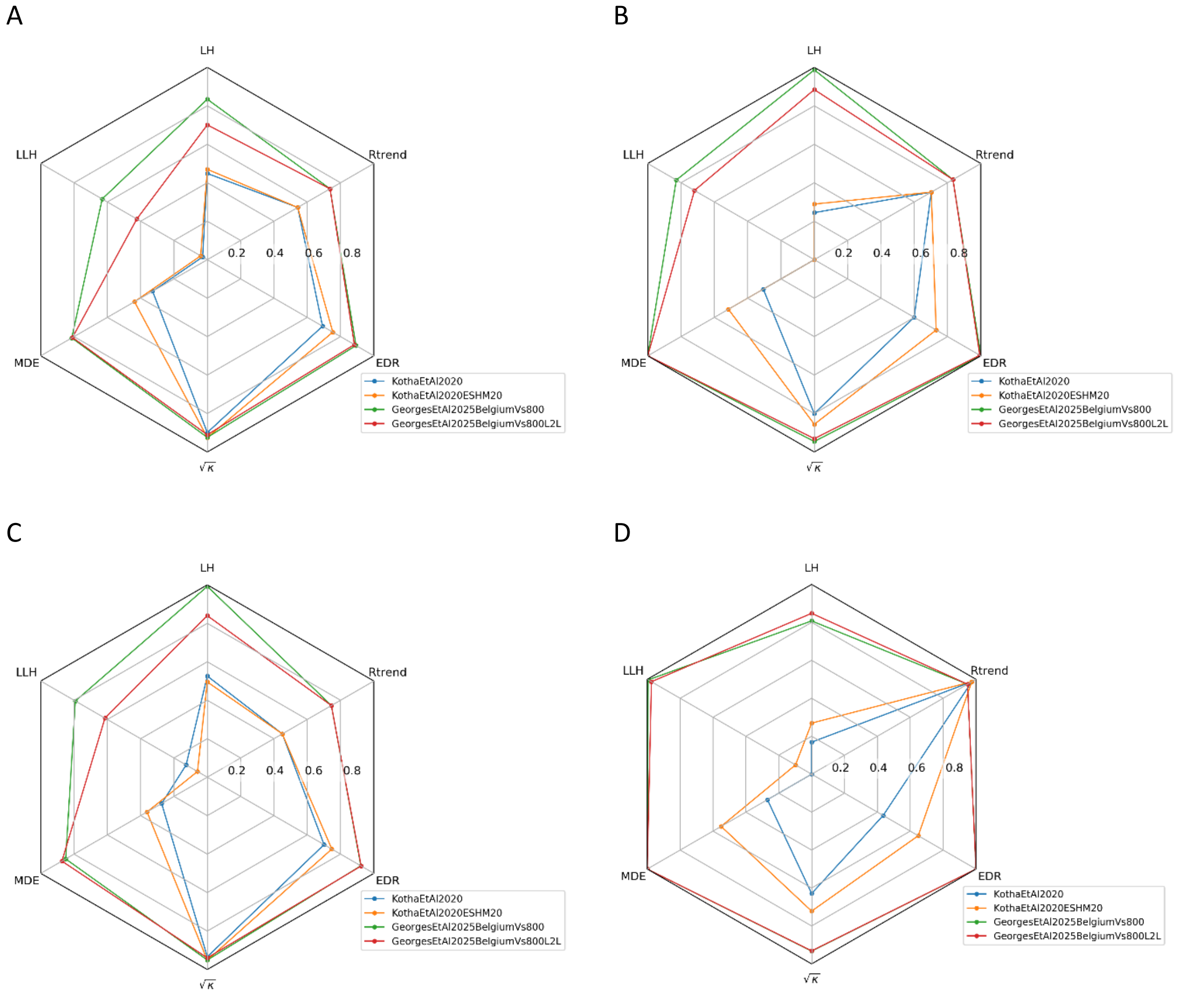


Figure 75 – Spider diagrams showing goodness-of-fit scores for different intensity measures predicted by the original and adjusted KothaEtAl2020 GMM, with respect to data from events with $M_W \geq 2.5$ in the BELSHAKE database. A: PGA; B: PGV; C: SA (T=0.1 s); D: SA (T=1 s)

Figure 76 shows that the GeorgesEtAl2025BelgiumVs800 model has a small bias, median LH scores around 0.5, small average LLH scores and small $\sqrt{\kappa}$ scores across different spectral periods, pointing to an overall good fit. Only the distance trend is still relatively strong, except for SA (T=1 s). We speculate that this may be due to the simple geometrical spreading function used in the KothaEtAl2020 model, while we have found that geometrical spreading in Belgium is better described by a segmented powerlaw model (see §4.3.1). On the other hand, the adjusted models for $V_{S30} \geq 1200$ m/s (not shown here) show smaller distance trends (when compared with the smaller corresponding subset of the data), so this needs further investigation.

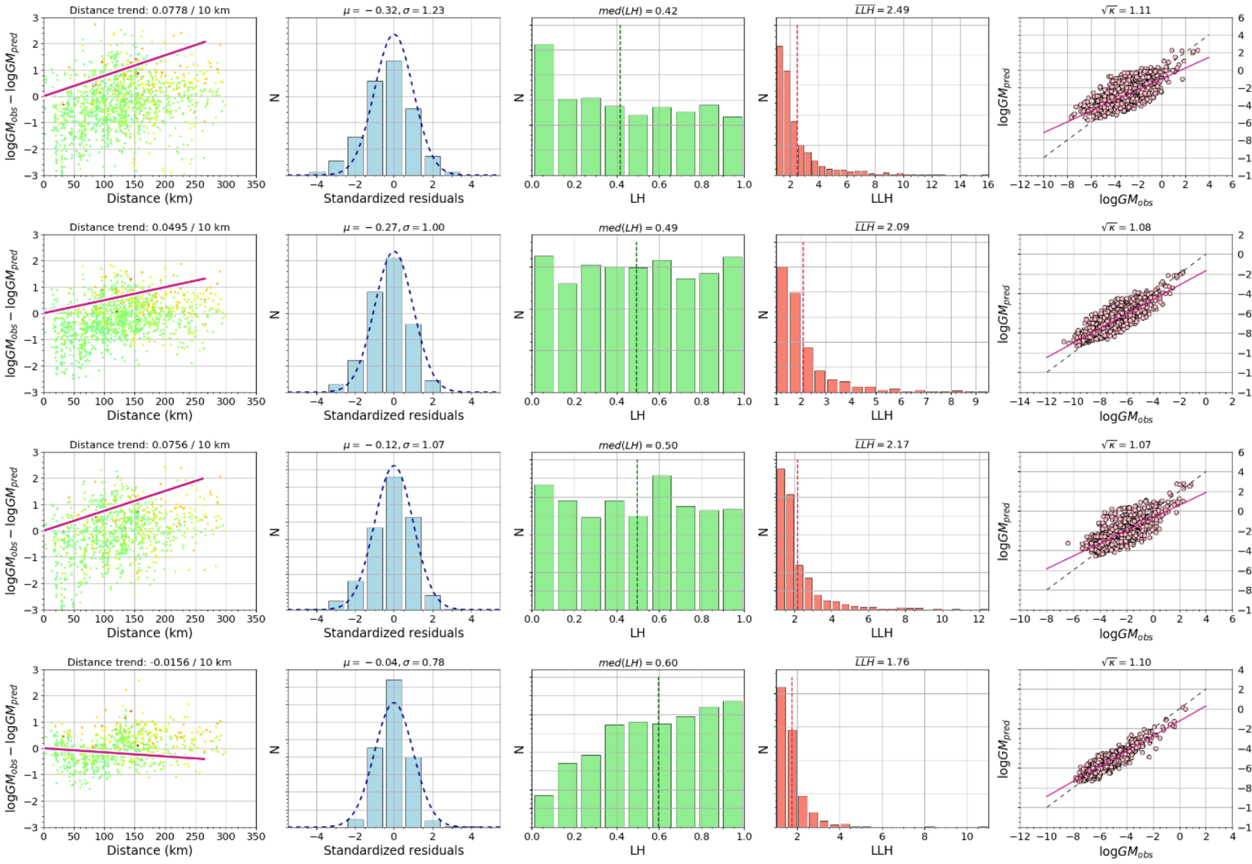


Figure 76 - Goodness-of-fit measures calculated for GeorgesEtAl2025BelgiumVs800L2L GMM adjusted for Belgium ($V_{S30} \geq 800$ m/s, with locality-to-locality regionalization) for different intensity measures, from top to bottom: PGA, PGV, SA (T=0.1 s) and SA (T=1 s)

Finally, Figure 77 compares observed ground-motion data (PGV) for individual earthquakes with values predicted by the GeorgesEtAl2025BelgiumVs800 model, as a function of distance. The selected events range in magnitude from $M_W = 3.5$ to $M_W = 5.3$. The results indicate that this adjusted model performs similar to the original KothaEtAl2020 model for moderate to high magnitudes, but much better for lower magnitudes ($M_W \leq 3.5$). This contrasts to the adjustment of the same GMM with the Referenced Empirical Approach, which deteriorated the fit for higher magnitudes, as can be seen by comparing panel B in Figure 77 with Figure 72. This demonstrates the superiority of the Bayesian updating method, which only corrects for the site and path attenuation properties of the target region, without affecting the scaling with magnitude.

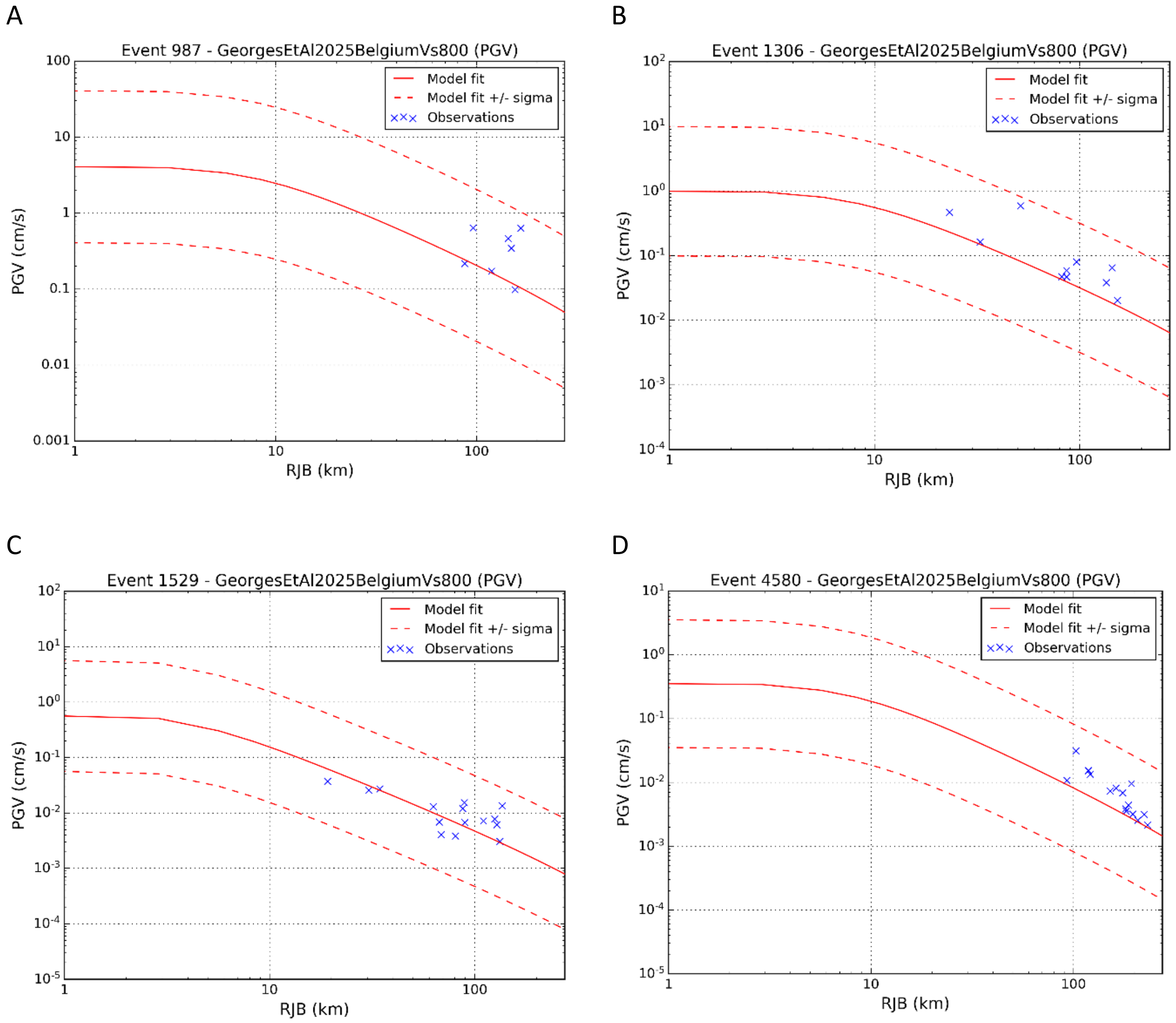


Figure 77 – Comparison of observed PGV for different earthquakes with values predicted by the GeorgesEtAl2025BelgiumVs800 GMM. A: $M_W = 5.3$ Roermond (NL) earthquake (1992); B: $M_W = 4.4$ Alsdorf (DE) earthquake (2002); C: $M_W = 3.5$ Voerendaal (NL) earthquake (2001); D: $M_W = 3.8$ Goch (DE) earthquake (2011)

In conclusion, adjustment of the KothaEtAl2020 GMM using the Bayesian updating method by Georges (2025) has resulted in 4 slightly different models that can be applied in Belgium, while retaining validity for higher magnitudes.

4.2.4. SIMULATION OF GROUND-MOTION FROM HIGH-MAGNITUDE EARTHQUAKES

The BELSHAKE database contains only few events with magnitude above 4, which is often considered as the minimum magnitude in seismic hazard assessments. The largest earthquake in the database is the M_W=5.3, 1992 Roermond earthquake. Unfortunately however, most records at short to intermediate distance (< 100 km) were clipped and can therefore not be used to derive ground-motion intensity measures. The epicentral shaking intensity of this key event remains therefore largely unknown. It is thus clear that, in order to derive meaningful conclusions about ground-motion models in Belgium, we have to extend the database to higher magnitudes. There are two ways to do this: 1) include data from other, “comparable” regions in Europe or elsewhere that are available in other ground-motion databases; 2) generate synthetic ground-motion records for hypothetical higher-magnitude earthquakes. For the latter case, we described two methods in §3.2.4 that we implemented in the frame of the project, EGF (Empirical Greens’ Functions) and EXSIM.

We aim to apply both EGF and EXSIM to reconstruct epicentral records of the 1992 Roermond earthquake, as well as to stochastically simulate accelerograms using EXSIM for hypothetical high-magnitude earthquakes that are in agreement with the average source, path and average site characteristics in and around Belgium. In that respect, the results obtained in other tasks of the project, such as the path and site components of the anelastic attenuation factor κ (§4.2.1), and the V_{S30} values at the sites of seismic stations (§4.1.6), along with the geometrical spreading and path duration models derived from recalibration of the M_L scale (§4.3.1), will prove to be very valuable. The intensity measures extracted from these generated ground-motion records can then be used to develop a ground-motion model specifically for Belgium.

As a proof of concept, we applied both the EGF and EXSIM methods to the closest station (BE.HU1 at a distance of 120 km) where the Roermond mainshock was recorded without clipping. The results in Figure 78 and Figure 79 show that realistic ground motion accelerograms and spectra are obtained with amplitudes that are in good agreement with those observed. There is also a very good match between the spectra obtained with our Python implementation of EXSIM and those obtained with the original Fortran program by Boore (Boore, 2009).

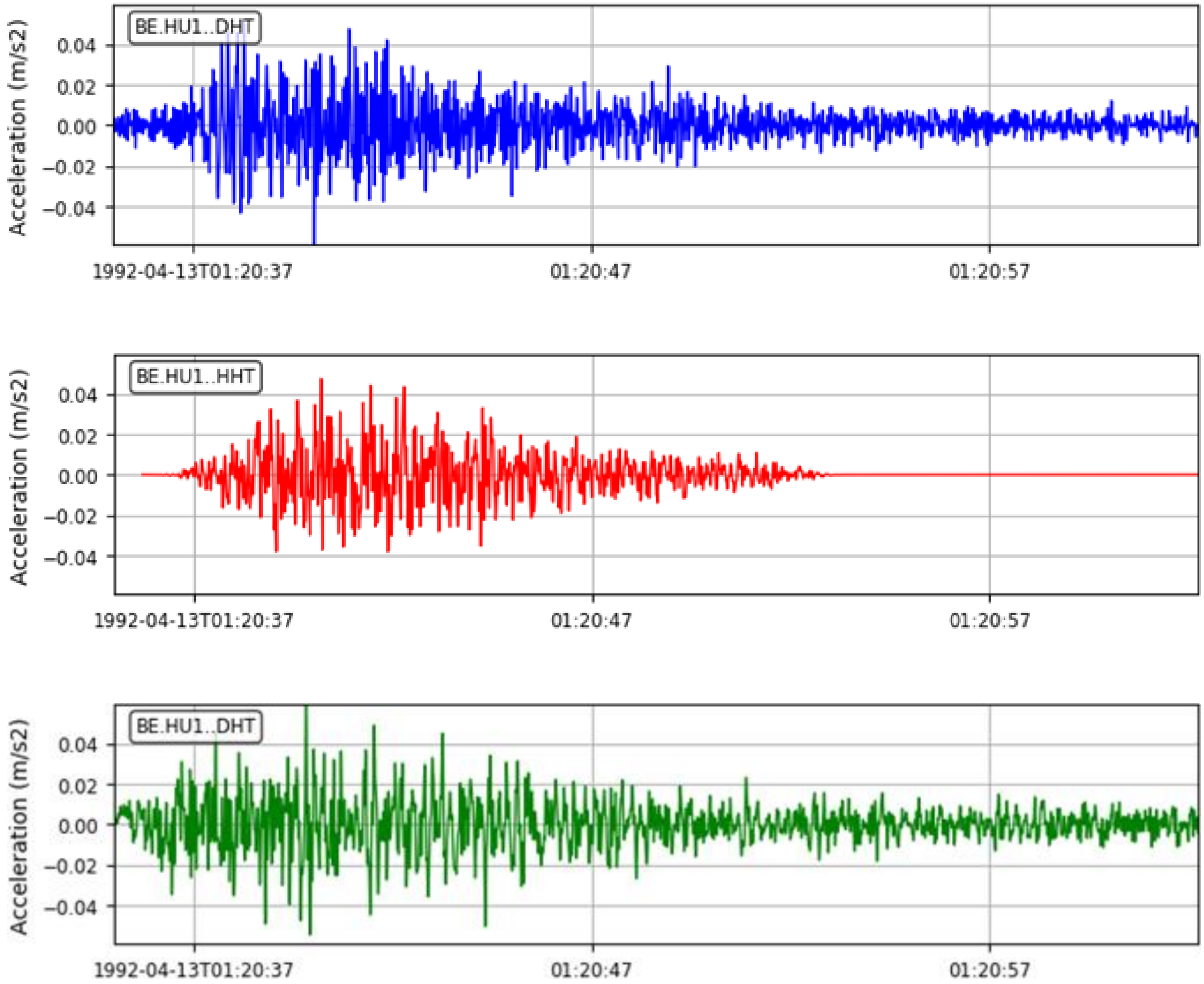


Figure 78 - Comparison of recorded and simulated accelerograms due to the 1992 Roermond earthquake at station BE.HU1 at a distance of 120 km. Blue: recorded; Red: simulated with python implementation of EXSIM; Green: simulated with python implementation of EGF. Note that vertical and horizontal scales are identical for all three panels

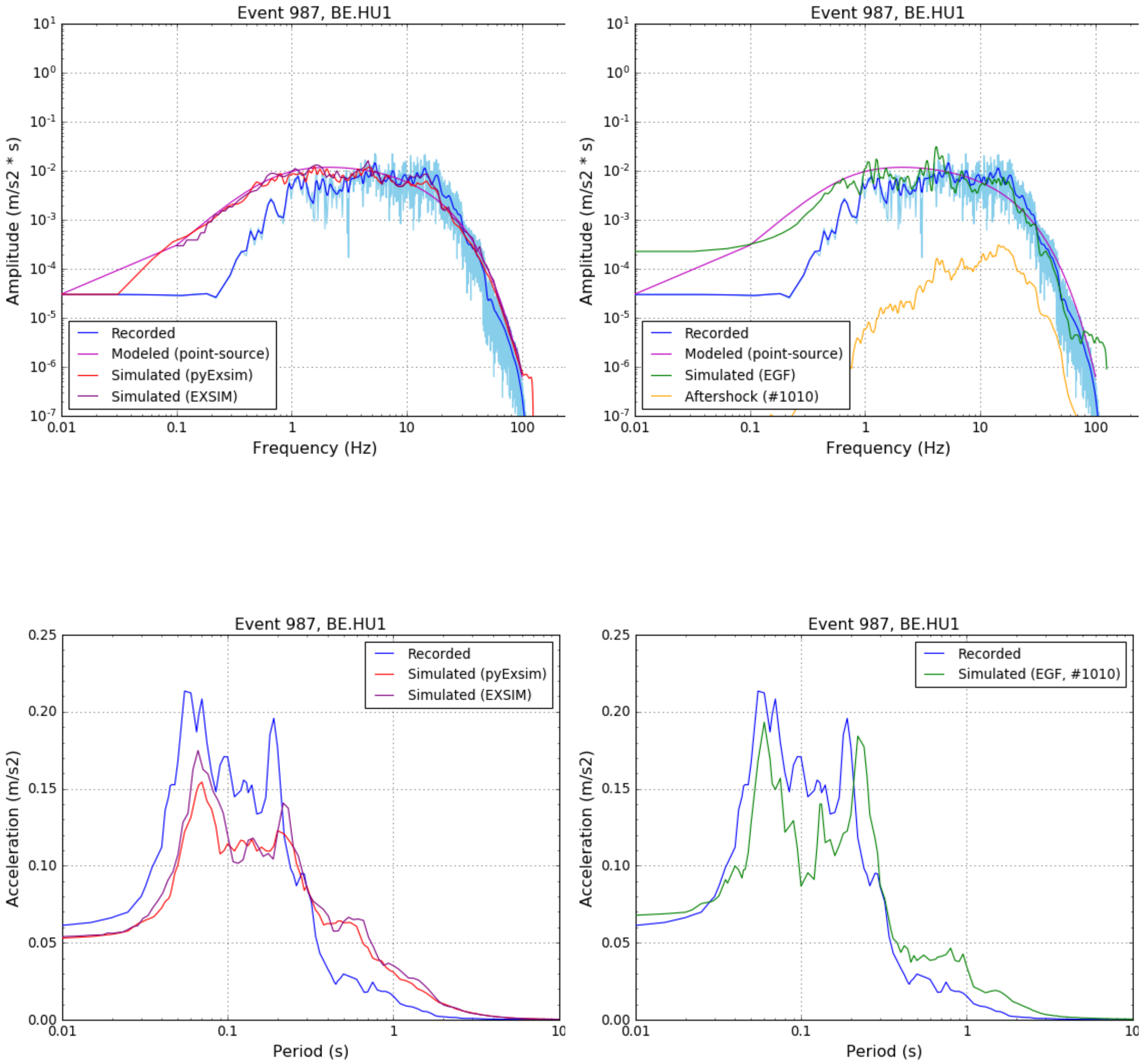
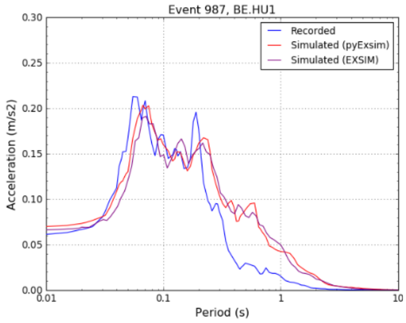


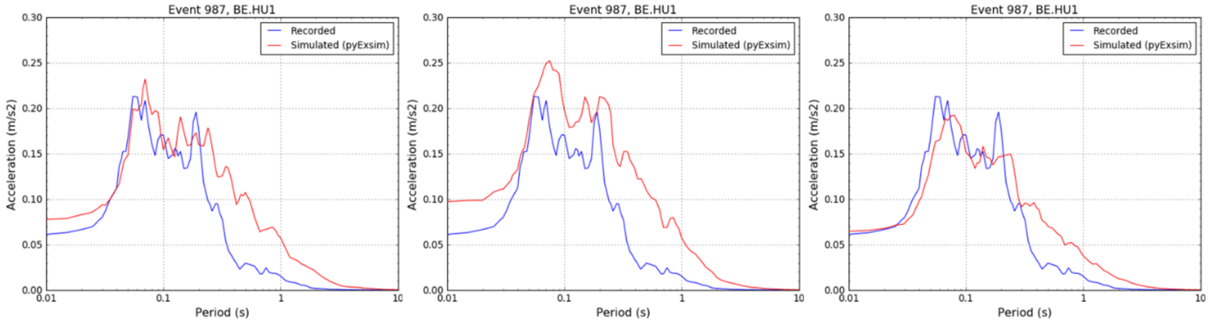
Figure 79 - Comparison of recorded and simulated Fourier amplitude spectra (top row) and response spectra (bottom row). Left column: simulated with python implementation of EXSIM (red) and original fortran program (purple); Right column: simulated with python implementation of EGF (green). Note that vertical scale is logarithmic for Fourier spectra and linear for response spectra

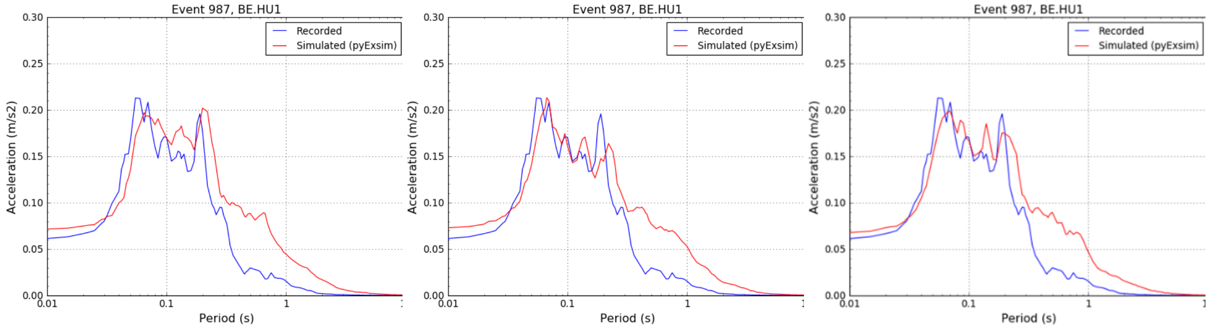
In the next step, a sensitivity analysis was conducted for both methods to better understand the impact of different source parameters (moment magnitude M_W, rupture velocity $Vrup$, corner frequency f_0 and corresponding stress drop $\Delta\sigma$), site parameters (average shear-wave velocity in the top 30 m V_{S30}, site transfer function) and modeling parameters (number of subfaults, max. pulsing fraction, slip distribution). The results in Figure 80 show that the EXSIM method is most sensitive to the correct estimation of mainshock corner frequency and site parameters.

For the EGF method, we analyzed the influence of aftershock properties (moment magnitude M_W, corner frequency f_0 and corresponding stress drop $\Delta\sigma$) and modeling parameters (number of subfaults, max. pulsing fraction, method for calculating HF scaling factor, phase randomization). Site effects cannot be varied as they are included in the aftershock record. The results are shown in

Figure 81 and indicate that the EGF method is most sensitive to the correct estimation of aftershock corner frequency and magnitude (the effects of which are accumulated due to summing over all subfaults). It is also clear that phase randomization improves the results.


Preferred model


$M_W = 5.4$ | $f_0 = 0.83\ Hz\ (\Delta\sigma = 109\ bar)$ | $Vrup = 2\ km/s$


Num. subfaults: 672 | Num. subfaults: 42 | Max. pulsing fraction: 0.25

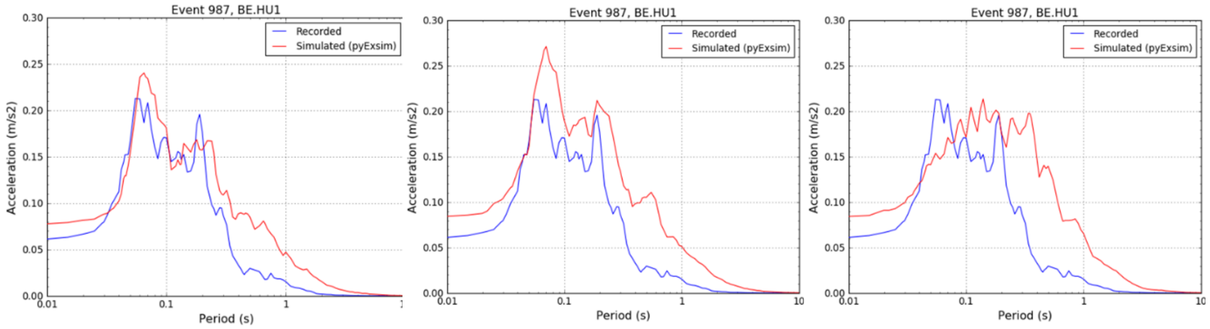
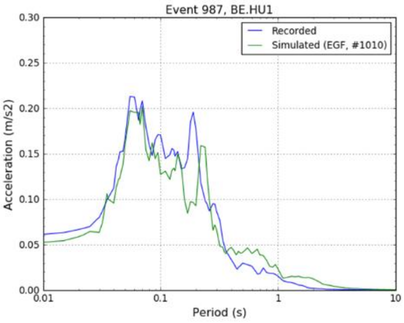
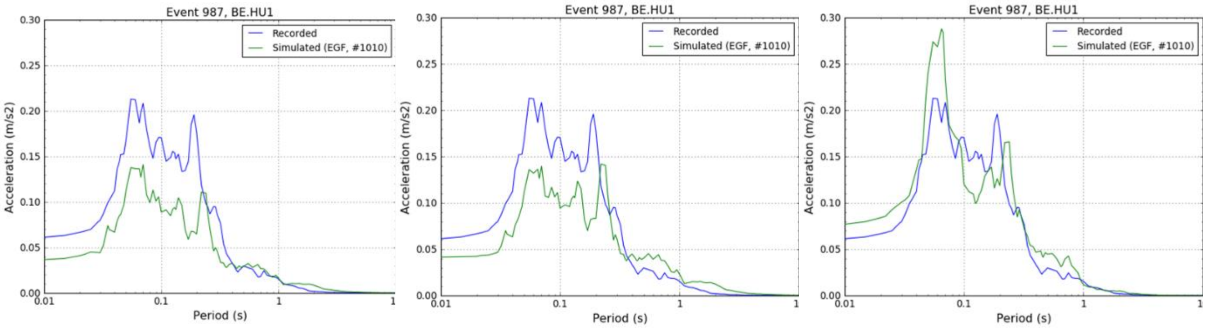

Random slip distribution | $V_{S30} = 800 m/s$ | No site transfer function

Figure 80 – Sensitivity analysis of response spectra modeled with EXSIM for the Roermond earthquake at station BE.HU1 (red curves). Blue curves correspond to recorded response spectrum. Top: original modeling with $M_W = 5.3$, $f_0 = 0.73\ Hz$ (equivalent to $\Delta\sigma = 74\ bar$), uniform slip = 0.33 m, num. subfaults = 168, max. pulsing fraction = 0.5, $V_{rup} = 3\ km/s$, $V_{S30} = 1240\ m/s$, and estimated site transfer function

Preferred model



$M_{W\_as} + 0.1$ $f_{0\_as} = 9\ Hz$ Num. subfaults: 42



Max. pulsing fraction: 0.25 HF factor: formula No phase randomization

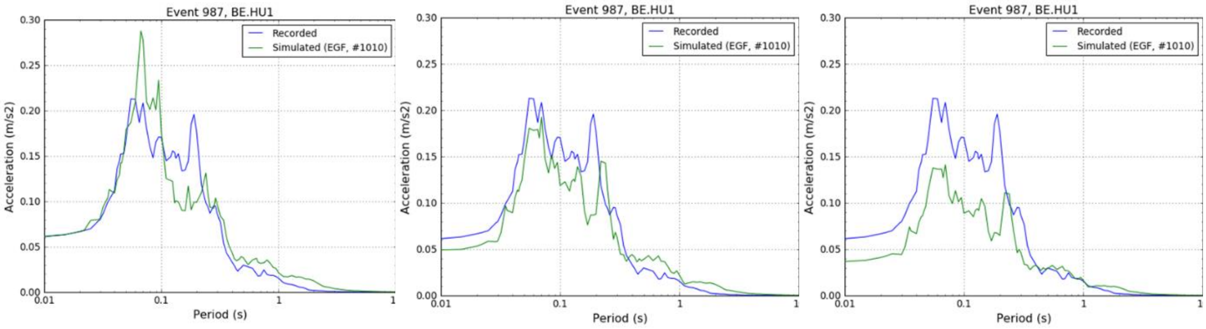
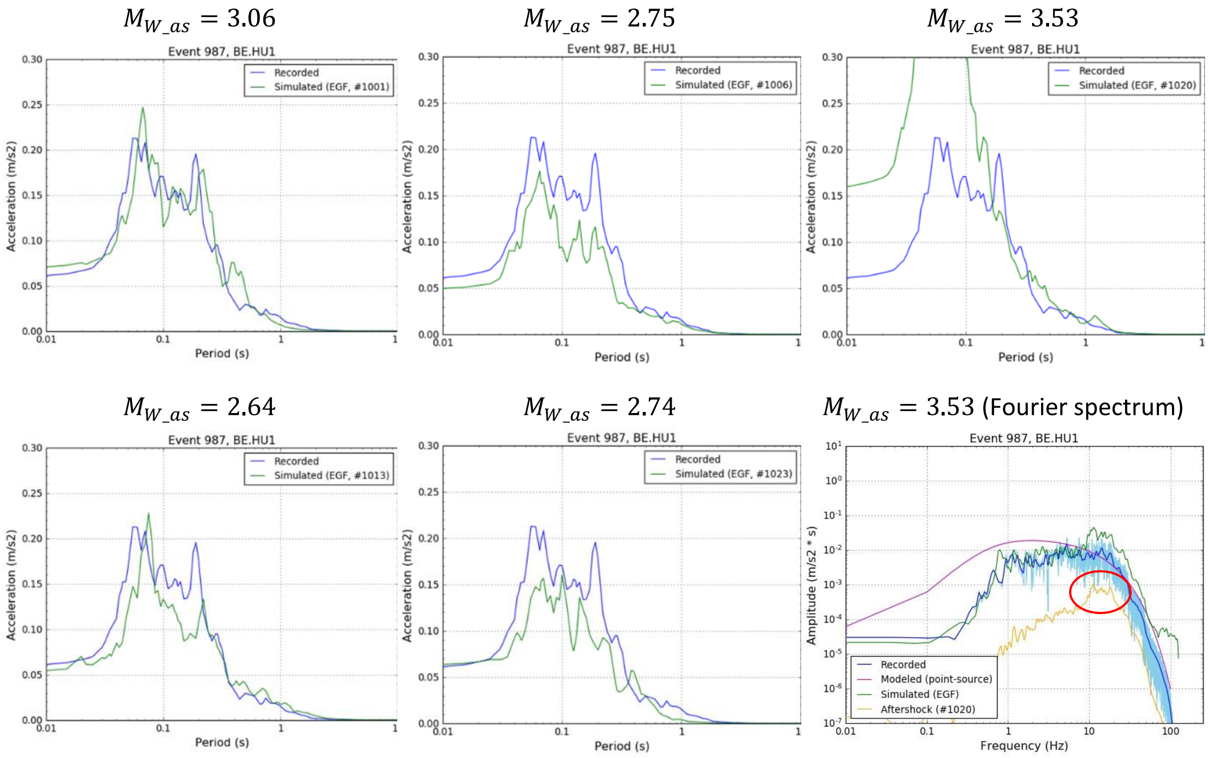


Figure 81 – Sensitivity analysis of response spectra modeled with EGF for the Roermond earthquake at station BE.HU1 (green curves). Blue curves correspond to recorded response spectrum. Top: original modeling with $M_{W\_as} = 2.98$, $f_0 = 7\ Hz$, uniform slip = 0.33 m, num. subfaults = 270, max. pulsing fraction = 0.5, $V_{rup} = 3\ km/s$, calculation of HF scaling factor by integration, and phase randomization

We also compared the EGF results based on different aftershocks with varying magnitudes. The results (Figure 82) are relatively similar, except for the case in the top right panel. This particular aftershock is a remote aftershock which occurred on a different fault than the mainshock rupture. The Fourier amplitude spectrum of this aftershock also has a different shape, as shown in the bottom right panel of Figure 82).

Finally, we applied both methods to stations in the epicentral region (pink triangles in Figure 83). As these stations were deployed after the mainshock occurred, recorded data are not available. Figure 84 shows that, except for one case, the two methods yield very different results. This is most likely due to the radiation factor we use in EXSIM. At these short distances, it is more appropriate to use variable radiation factors based on the focal mechanism rather than the average radiation factor. However, uncertainties on the focal mechanism may still lead to relatively large differences. The EGF method may be more reliable in this case, provided that the focal mechanism of the aftershock is similar to that of the mainshock. The PGA values obtained with the EGF method (left end of green curves in Figure 84) range between 1 and ~3.5 m/s², which is similar to the range obtained by Gariel et al. (1994).



$M_{W\_as} = 3.06$ $M_{W\_as} = 2.75$ $M_{W\_as} = 3.53$

$M_{W\_as} = 2.64$ $M_{W\_as} = 2.74$ $M_{W\_as} = 3.53$ (Fourier spectrum)

Figure 82 – Response spectra modeled with EGF for the Roermond earthquake at station BE.HU1 (green curves). Blue curves correspond to recorded response spectrum. Bottom right: acceleration response spectra corresponding to the case in the top right panel

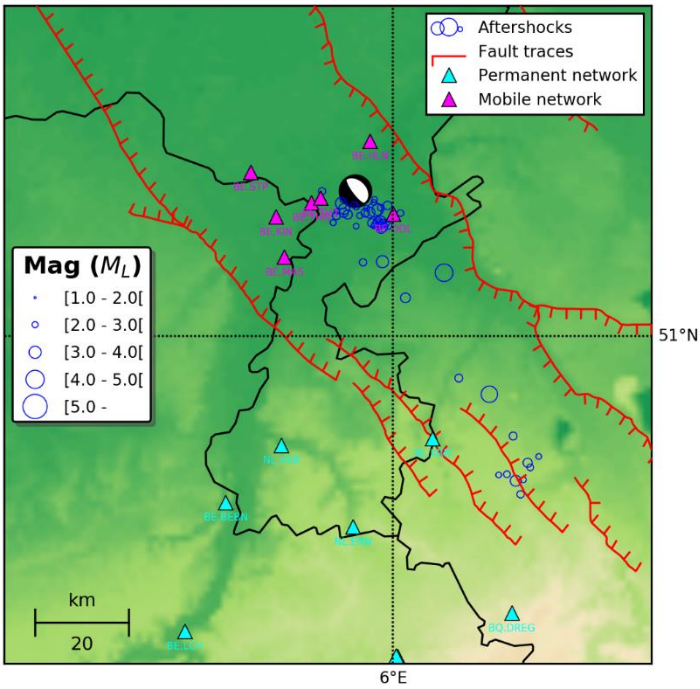


Figure 83 – Map of the epicentral area, showing location of aftershocks, seismic stations and faults

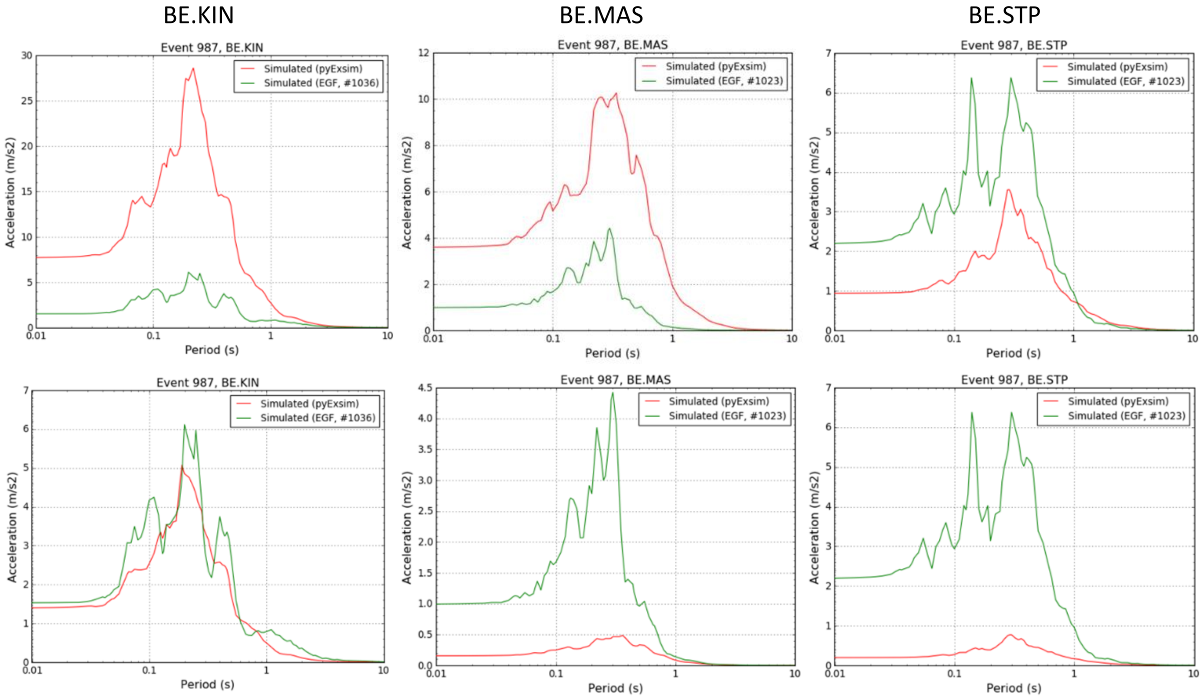


Figure 84 – Response spectra modeled with EXSIM (red curves) and EGF (green curves) for the Roermond earthquake at three stations close to the epicenter. Top row: using average S-wave radiation factor (0.55); Bottom row: using variable radiation factor based on focal mechanism and station azimuth. Note that this only affects EXSIM spectra, while EGF spectra are the same in both cases (note different vertical scales)

These preliminary results indicate that both EGF and EXSIM are very promising methods to simulate ground motion for past large earthquakes in stations where the mainshock was not well recorded, but good-quality aftershock records are available, as well as for hypothetical large earthquakes at sites with well-constrained site characteristics (V_{S30}, κ, site amplification, …).

4.3. Additional results

Thanks to the BELSHAKE database, it has been possible to recalibrate the local magnitude scale used in Belgium. This is a very important side result, which was not part of the project itself.

4.3.1. RECALIBRATION OF THE BELGIAN LOCAL MAGNITUDE (M_L) SCALE

The current scale that is routinely used to determine M_L of local earthquakes recorded by the Belgian seismic network has been in use since its definition by Camelbeeck (1985). Since that time, the network has been further improved, resulting in a lower detection threshold. In addition, geothermal energy plants have come into operation, triggering many small earthquakes that are often only recorded in nearby stations. These evolutions have brought to light two drawbacks of the current M_L scale:

- The distance metric of the scale is epicentral distance, which does not take into account focal depth. As a result, earthquakes generating similar amplitudes at similar distances have the same magnitude, regardless of their focal depth;
- The scale has not been calibrated for distances smaller than 10 km, which makes it difficult to determine M_L for smaller earthquakes that are only recorded in nearby stations.

Thanks to the quality-checked records, integrated instrument response information and comprehensive metadata, the digital waveforms in the BELSHAKE database provide the perfect opportunity to recalibrate the Belgian M_L scale. We do this as a function of hypocentral distance, while paying particular attention to short distances (few tens of km) and ensuring a good agreement with the original scale at larger distances.

First, we determined maximum peak-to-peak displacement amplitudes on the vertical component in the S window of all records in the database using an automated procedure. We found that the M_L values computed from these amplitudes are reliable and that both the amplitudes and M_L values are more consistent than those in the ROB catalog. An added bonus is that more amplitude data could be collected this way than there are in the ROB catalog, despite the smaller number of earthquakes in the BELSHAKE database.

In the second step of the analysis, we fitted different relations to the magnitude-corrected station amplitudes ($\log_{10} A_{station} + M_{L\_ref} - M_{L\_event}$, with $M_{L\_ref} = 3.0$), as a function of hypocentral rather than epicentral distance. To reduce the influence of short-distance stations, only events with $\geq 2/3$ of contributing stations at distances beyond 30 km were included. Some results are shown in Figure 85. The top left panel shows the selected magnitude-corrected station amplitudes as a function of epicentral distance, along with the original scale. The total standard deviation of this relation with respect to the dataset is 0.231. The overall shape of the dataset represents regional attenuation (geometrical spreading and anelastic attenuation), which we want to capture by fitting, while the vertical variability reflects random variations and site effects. The top right panel shows the fit for a common relation with linear and logarithmic distance terms, similar to the original scale but including a reference distance, with respect to hypocentral distance:

$$\log_{10} A_0 = c_1 + c_2 \log_{10}\left(\frac{R_h}{R_{ref}}\right) + c_3\left(R_h - R_{ref}\right)$$

This relation has a standard deviation that is almost identical (0.230) to that of the original scale with respect to epicentral distance. Other relations were tested, but these did not improve the fit. However, if we change to a logarithmic distance scale (bottom left panel), apparently straight sections with different slopes can be discerned. This shape is very similar to piecewise continuous geometrical spreading models with a hinged trilinear shape in log-log space, such as the one used to model horizontal S-wave Fourier spectra in NE America (Atkinson & Boore, 1995). We thus tried to fit a piecewise continuous powerlaw model to our dataset by performing a (bounded) minimization to find knickpoint distances and exponents. The total standard deviation of this powerlaw fit is reduced to 0.220. However, it diverges from the original relation at large distances (> 400 km), mainly because it does not take into account anelastic attenuation. If we include regional anelastic attenuation in the form of a gradient of the high-frequency attenuation parameter kappa (see also §3.2.1.2), we obtain the fit in the bottom right panel of Figure 85, with an almost similar total standard deviation of 0.219, but more importantly better matching the original relation at larger distance.

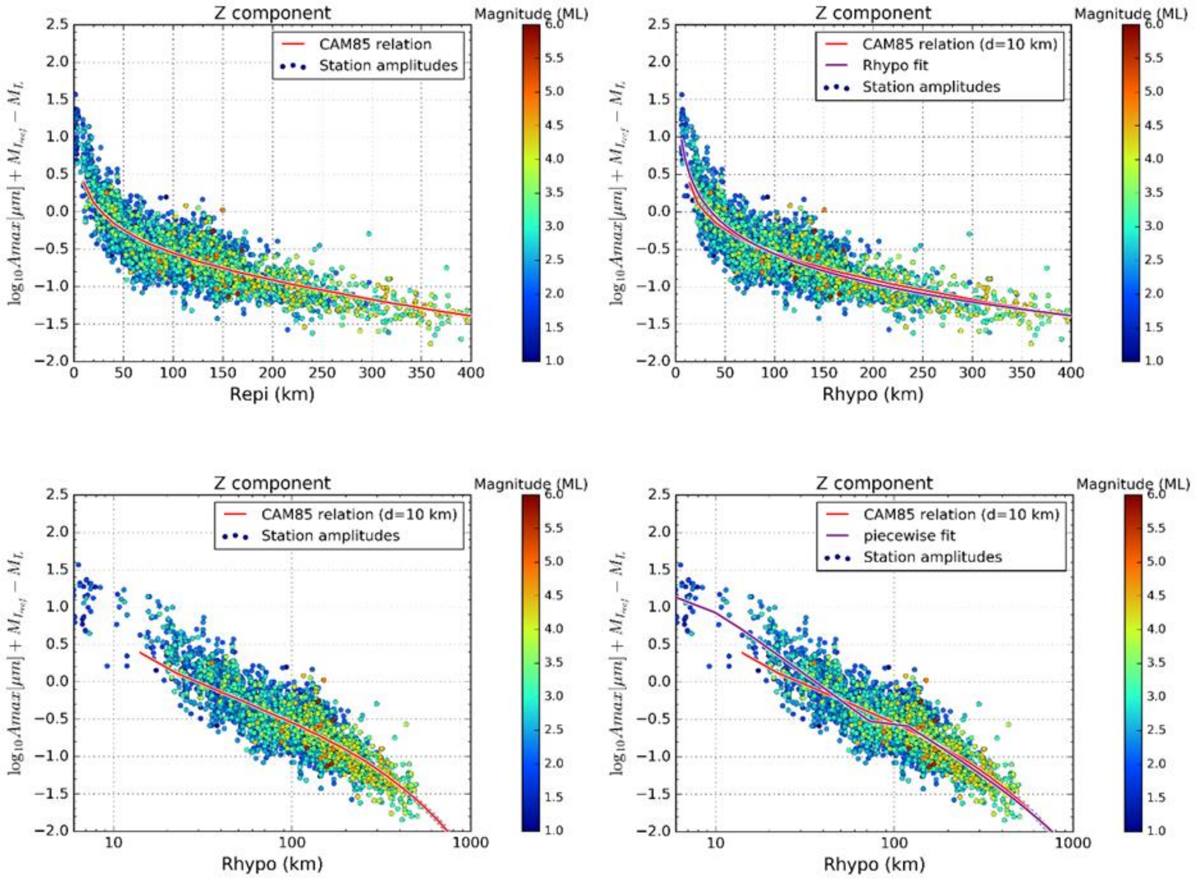


Figure 85 - Selected magnitude-corrected station amplitudes (measured on the Z component) versus distance, along with different fits. Top left: original M_L scale versus linear epicentral distance; Top right: “Rhypo” fit versus linear hypocentral distance; Bottom left: original M_L scale (adjusted to a focal depth of 10 km) versus logarithmic hypocentral distance; Bottom right: piecewise continuous powerlaw + kappa gradient fit versus logarithmic hypocentral distance. Amplitudes are color-coded according to event M_L values

We also determined amplitudes in a similar fashion for the horizontal components, combining the two horizontal components for each station in a single value by taking the geometric mean (GM). A similar fit (piecewise continuous powerlaw + kappa gradient) yields a total standard deviation of 0.258 (left panel in Figure 10), which is significantly higher than that obtained for the Z component. Although the use of horizontal components thus does not provide any advantage over the Z component for M_L determination, it does allow comparing with existing models of geometric spreading and anelastic attenuation of S waves, which are generally available for the horizontal component only. Both the powerlaw exponents and the kappa gradient thus obtained appear to be higher than expected. It is important to realize, however, that these models, e.g. the spreading model for NE America mentioned earlier, are Fourier-domain models. Time-domain and Fourier-domain amplitudes do not decay following the same powerlaw exponents, but should be related through the duration (T) according to Parseval’s theorem:

$$A_{time} = T^{-0.5} A_{fourier}$$

In the BELSHAKE database, we have computed durations following different metrics, a.o. D5-95 which is the duration corresponding to the buildup from 5% to 95% of the total Arias Intensity (see

§3.1.7). The right panel in Figure 86 shows the results for the horizontal (GM component) S wave. Similar to the magnitude-corrected amplitudes (the shape of which is dominated by geometrical spreading), a log-log plot of these durations suggests a piecewise continuous powerlaw with a flat midsection and knickpoints at more or less the same distances. If we integrate the duration model directly into the amplitude fitting procedure, the fits (not shown here) are not improved visually, but the kappa gradient is reduced from 0.46 ms/km to 0.18 ms/km, which is in better agreement with the results obtained in our dedicated study of this parameter (see §3.2.1.2). The fitted geometrical spreading model is shown in Figure 87. The powerlaw exponents obtained when integrating the duration model in the fit are also slightly lower than without, but still higher (in absolute terms) than those in the NE American model, indicating stronger attenuation through geometrical spreading in our region. Finally, our model also includes an additional knickpoint at ~10 km. It is not so well constrained by the data, and it is not clear whether it is related to geometrical spreading or to stronger attenuation due to shallow depth.

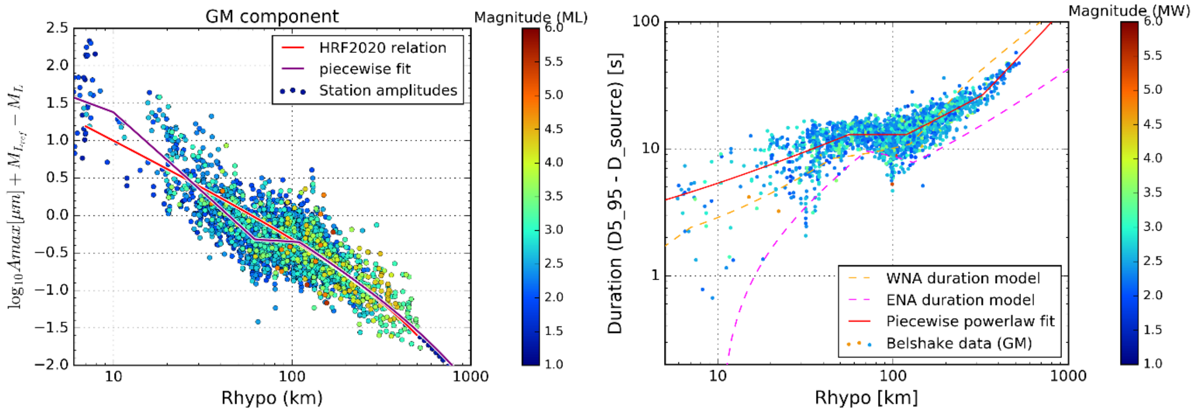


Figure 86 - Left: Selected magnitude-corrected station amplitudes (GM component) versus hypocentral distance, along with piecewise continuous powerlaw + kappa gradient fit. Right: path durations (D5-95 corrected for source duration) for horizontal S waves recorded at rock stations in the BELSHAKE database, along with piecewise continuous powerlaw fit and 2 models used in the USA

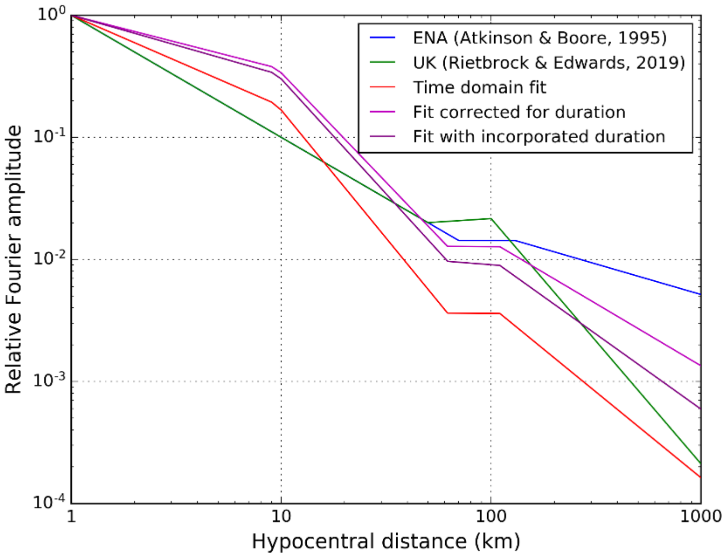


Figure 87 - Comparison of geometrical spreading model for eastern North America (Atkinson & Boore, 1995) and UK (Rietbrock & Edwards, 2019) with fitted spreading models for horizontal S-waves obtained in this study; note that all models are Fourier models, except for time-domain fit

Finally, we implemented the recalibration parameters in our M_L calculation code, and applied the recalibration based on the Z component to two sets of earthquakes: the BELSHAKE database (which contains only few earthquakes with $M_L < 2$) and earthquakes with $M_L \leq 2.0$ in the ROB catalog. The results are compared to the M_L values calculated with the original scale in Figure 88. For the first set, the agreement between both M_L calculations is very good, with zero bias and very low variability (standard deviation < 0.04). This is both expected and intended, as most earthquakes in BELSHAKE have been recorded by stations over a wide range of distances, and the recalibrated M_L scale should give the same results as the original one in this case. For the second set, the variability is clearly higher and there is a tendency for increasing negative bias (recalibrated M_L < original M_L) with lower M_L, which is expected as the contribution of faraway stations gradually drops. This demonstrates that we have successfully recalibrated the Belgian M_L scale, taking into account focal depth. Furthermore, the new scale can be used down to shorter distances (> 6 km) and has improved residuals compared to the original scale.

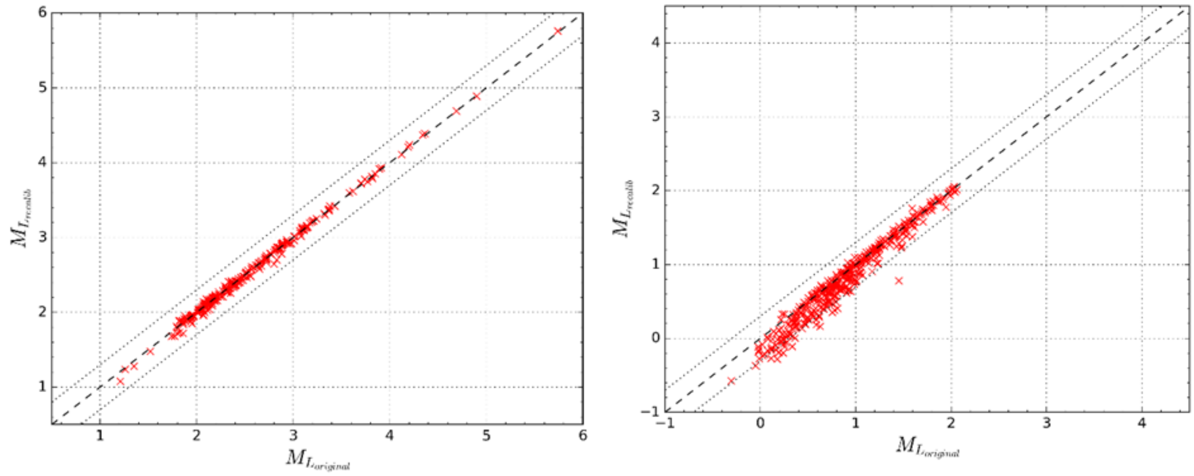


Figure 88 - Comparison of recalibrated M_L values with M_L values calculated with the original scale. Left: for earthquakes in the BELSHAKE database (M_L mostly ≥ 2.0); Right: for earthquakes with $M_L \leq 2.0$ in the ROB catalog

A paper describing these results and referencing the BELSHAKE project has been submitted. Furthermore, the geometrical spreading and duration models for horizontal S waves that we derived in this study are important for many other seismological applications as well as for the BELSHAKE project itself, e.g. for determining moment magnitudes (M_W) from displacement spectra, modeling of Fourier spectra and stochastic simulation of earthquake ground motion.

4.4. Recommendations

To further extend or build on the results of the BELSHAKE project, we have identified the following points of interest:

- First, it is important that the BELSHAKE database be kept alive, updated with new earthquake records, and possibly extended with lower-magnitude earthquakes. The BELSHAKE database has evolved into a valuable addition to ROB’s seismological databases, which thanks to its quality-controlled records and comprehensive metadata can be used for many seismological and seismic-engineering applications.

- Although the BELSHAKE database could benefit from the characterization of several stations of the Belgian seismic network in the EPOS-BE project "Upgrade of Belgian federal research infrastructure for EPOS" (Contract nr. FSIRI/33/EP1), funded by the Belspo, many stations, particularly discontinued ones, still lack quantitative characterization. It is recommended to continue efforts to determine V_{S30} profiles in these stations, which is important to link ground motion to site properties.
- In relation to site characterization, it would also be interesting to compare our κ_0 results with estimates based on alternative approaches, such as ambient noise analysis (e.g., Moratto et al., 2025).
- It could be very useful to apply the Generalized Inversion Technique (GIT, e.g. Cataldi et al., 2022; Morasca et al., 2025; Oth et al., 2017) to the ground-motion data in the BELSHAKE database. This relatively new method takes advantage of redundance of information (each event is recorded by multiple stations, each station has recorded multiple events) to invert simultaneously for source, path and site properties. GIT can either be non-parametric (no initial assumptions are made) or non-parametric (allowing the use of bounds on parameters that are already known). This would likely provide new insight into the frequency dependence of anelastic attenuation and lead to better constraints on earthquake stress drops, which would improve the accuracy and precision of ground-motion simulations.
- One of the objectives of the BELSHAKE project was to either develop a stochastic ground-motion model or to calibrate an existing model for the attenuation characteristics in Belgium. We obtained more concrete results with the second line of investigation, in which we implemented a relatively simple method for adjusting GMMs, identified 2 or 3 models that are suitable for adjustment, and updated our preferred model; in addition, thanks to data exchange other research groups could apply a more sophisticated adjustment method, that resulted in the adjustment of another GMM that can be applied in Belgium. We also made developments related to the first line of investigation, stochastic simulation. We obtained promising results, demonstrating that this method can be used to approximate ground motion produced by the 1992 Roermond earthquake, the largest event in the BELSHAKE database. However, there was not enough time to develop a stochastic ground-motion model during the project. If GIT proves successful and provides additional constraints, it would be very useful to pursue this approach. The best starting point would be the Rietbrock & Edwards (2019) GMM, which we identified as best matching the BELSHAKE data, and which also incorporates a segmented geometrical spreading function, similar to the geometrical spreading that we derived from the BELSHAKE database.
- Another ground-motion modelling approach that could be interesting to investigate is the hybrid stochastic-empirical method (HEM, Douglas et al., 2024), in which the ratios between stochastic models for host and target regions are used to adjust an empirical model for the host region to make it applicable to the target region. Considering that our analyses have shown that low-attenuation versions of NGA-West (i.e. California) GMMs show a relatively good match with the BELSHAKE database, this approach could result in a model for Belgium that has better empirical support for higher magnitudes.

5. DISSEMINATION AND VALORISATION

5.1. Open Datasets

We have published three datasets related to the BELSHAKE database on the Zenodo data repository:

- The ground-motion data in the BELSHAKE database have been published as a flatfile (plain text table containing the main intensity measures and metadata for each record): https://doi.org/10.5281/zenodo.15241694. A first version was released with restricted access in 2024-02; a second version with several improvements was released as open-access in 2025-04. On 3 Feb 2026, this dataset has been downloaded 275 times.
- Local magnitude (M_L) and moment magnitude (M_W) solutions for earthquakes in the BELSHAKE database recorded by stations of the Belgian seismic network are published as a separate dataset: https://doi.org/10.5281/zenodo.17514313. On 3 Feb 2026, this dataset has been downloaded 47 times.
- A third dataset contains individual kappa values for all records in the BELSHAKE database, and derived site-specific κ_0 and kappa gradient values, as well as regional kappa gradients, along with associated uncertainties: https://zenodo.org/records/17857380. On 3 Feb 2026, this dataset has been downloaded 38 times.

5.2. Participation to international effort to investigate ground motions caused by the 1992 Roermond earthquake

In January 2024, we were contacted by the ground-motion modelling team of the NAM (Nederlandse Aardolie Maatschappij), who wish to conduct a comparative study of ground motions due to the induced earthquakes in the Groningen gas field (Netherlands) and those caused by the 1992 Roermond earthquake. The latter is the largest instrumentally recorded earthquake in our region and a key event in the BELSHAKE database. Unfortunately however, most records at short to intermediate distance (< 100 km) were clipped and can therefore not be used to derive intensity measures. The epicentral shaking intensity of this important earthquake is therefore largely unknown. In the 30 years that have passed, more advanced methods have become available and computational capabilities have increased significantly. As a result, it may be possible to retrieve reliable information from records that were previously considered unusable, as well as to simulate epicentral ground motions.

Thanks to the ongoing BELSHAKE project, we were in a position to contribute to this effort, and participate to a number of workshops together with scientists from NAM, KNMI (Royal Dutch Meteorological Institute, responsible for seismic monitoring in The Netherlands) and the University of Leeds. Vice versa, the resulting exchange of data and methods was also beneficial for the BELSHAKE project. In particular, this collaboration has allowed us to collect additional waveform data for the 1992 Roermond earthquake from other networks in Europe (Figure 89). In addition, we discovered waveform data for several aftershocks of the Roermond earthquake, recorded by a temporary network deployed by ROB in the epicentral area. These data had been overlooked before due to incorrect registration of these stations in the main ROB database. Furthermore, the collaboration provided the opportunity to implement methods to simulate ground motion from the Roermond earthquake based on extended fault ruptures (see §3.2.4 and §4.2.4). Finally, this collaboration has also resulted in co-authorship in a publication describing an improved method for reconstructing clipped waveform data, under the lead of Ben Edwards (Leeds University, UK). It is likely that other publications will follow in the coming years.

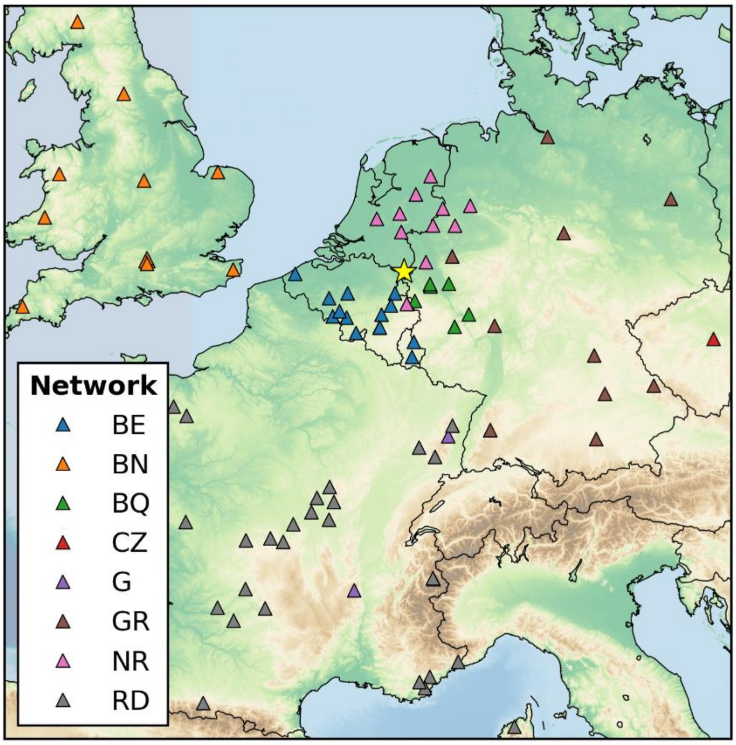


Figure 89 – Map showing seismic stations with waveform data of the 1992 $M_W = 5.3$ Roermond earthquake

5.3. Revision of the Belgian National Annexe to Eurocode 8

The results of the BELSHAKE project will have direct implications for the ongoing revision of the European building code EN-1998 (Eurocode 8), and more particularly the Belgian national annexe. This will no doubt be the most important valorisation of the project. The first version of the Belgian national annexe (Belgisch Instituut voor Normalisatie (BIN), 2002) is more than two decades old and was based on a primitive ground-motion model. Thanks to the BELSHAKE project, it will be possible for the first time to assess seismic hazard using GMMs that have been adjusted for the crustal attenuation characteristics in Belgium based on local data. Another important contribution concerns the moment magnitudes that we calculated for 294 earthquakes in the Belgian seismic catalog, whereas previously moment magnitudes had to be estimated from local magnitudes based on empirical relationships, which have larger uncertainties. These two main aspects will certainly allow computing more reliable seismic hazard maps for the country. The revision of the national annexe should be finalized in 2027.

5.4. Meeting presentations

5.4.1. PRESENTATIONS AT INTERNATIONAL CONFERENCES

- Onvani M., Vanneste K. & Almeida J., The BELSHAKE database of earthquake ground motion in Belgium, IUGG 28th General Assembly, Berlin (Germany), 13-20 July 2023 (oral);
- Vanneste K., Onvani M., Camelbeeck T., Lecocq T. & Rapagnani, G., Magnitude reassessment for digital records of the Belgian seismic network since 1985, IUGG 28th General Assembly, Berlin (Germany), 13-20 July 2023 (oral);

- Vanneste K. & Onvani M., The BELSHAKE database of earthquake ground motion in Belgium, British Seismology Meeting 2024, Reading (UK), 25-27 March 2024 (poster).
- Onvani M. & Vanneste K., The BELSHAKE database of earthquake ground motion in Belgium, 18th World Conference on Earthquake Engineering (WCEE2024), Milano (Italy), 30 June – 5 July 2024 (oral);
- Vanneste K. & Onvani M., Waveform processing in BELSHAKE, the Belgian ground-motion database, 18th World Conference on Earthquake Engineering (WCEE2024), Milano (Italy), 30 June – 5 July 2024 (oral);
- Vanneste K. & Onvani M., The BELSHAKE database of earthquake ground motion in Belgium, 102nd JLG / EFEHR Scientific Session, Luxembourg, 27-29 November 2024 (poster);
- Onvani M. & Vanneste K., Individual kappa calculation for BELSHAKE, site kappa and kappa gradient estimation across crustal domains in Belgium, 102nd JLG / EFEHR Scientific Session, Luxembourg, 27-29 November 2024 (poster);
- Vanneste K., Edwards B. & Ntinalexis M., Simulating ground motion from the 1992, M_W=5.3, Roermond (The Netherlands) earthquake using EGF and EXSIM, IAGA/IASPEI Joint Scientific Meeting, Lisbon (Portugal), 31 August – 5 September 2025:
- Vanneste K., Onvani M. & Van Noten K., The BELSHAKE database of earthquake ground motion in Belgium, IAEA International Workshop on Seismic Hazard Assessment of Nuclear Installation Sites in Low Seismicity Regions, Brussels, 4-7 November 2025.

5.4.2. PRESENTATIONS AT SEMINARS AND WORSKHOPS

- Vanneste K. & Onvani M., The BELSHAKE project, Belgian Eurocode 8 Mirror Committee meeting, Brussels, 13 May 2022 (oral);
- Onvani M. & Vanneste K., The BELSHAKE project, Rhine-Meuse Seismologists (RMS) meeting, Utrecht (NL), 22 June 2022 (oral);
- Onvani M., Vanneste K. & João Almeida, The BELSHAKE database of earthquake ground motion in Belgium, BELQUA Annual Scientific Workshop, Brussels, 7 March 2023 (oral);
- Vanneste K., Onvani M., Camelbeeck T. & Rapagnani G., Magnitude reassessment for digital records of the Belgian seismic network since 1985, Rhine-Meuse Seismologyists (RMS) meeting, Uccle, 28 March 2023 (oral);
- Vanneste K. & Onvai M., Progress made in the BELSHAKE project, Rhine-Meuse Seismologists (RMS) meeting, Bensberg (Germany), 24 October 2023 (oral);
- Vanneste K. & Onvani M., Current status of the BELSHAKE project, Eurocode 8 National Mirror Committee meeting, Brussels, 22 November 2023 (oral);
- Vanneste K. & Onvani M., The 1992 Roermond earthquake: a key event in the BELSHAKE ground-motion database, Workshop on ground motions of the 1992 Roermond earthquake, Schiphol (Netherlands), 13 March 2024 (oral);
- Vanneste K., Recovery of aftershock data of the 1992 Roermond earthquake recorded by the Belgian mobile network, Rhine-Meuse Seismologists (RMS) Meeting, Utrecht (Netherlands), 14 May 2024 (oral);
- Vanneste K. & Van Noten K., Progress made by ROB since 1st workshop, Workshop on ground motions of the 1992 Roermond earthquake, online, 19 June 2024 (oral);

- Vanneste K., Simulation of ground motion caused by the 1992 Roermond earthquake using EGF and EXSIM: A work in progress, Workshop on ground motions of the 1992 Roermond earthquake, online, 19 June 2024 (oral);
- Vanneste K., Simulating ground motion from the 1992, M_W=5.3, Roermond earthquake using EGF and EXSIM, Workshop on ground motions of the 1992 Roermond earthquake, Schiphol (Netherlands), 22 May 2025 (oral).

5.5. Participation to international workshops

- M. Onvani participated to a workshop on ground motion in France and surrounding areas, organized by EDF, Saint-Paul-lez-Durance (France), 19-20 October 2023 (oral presentation);
- K. Vanneste participated to a workshop on seismic hazard assessment of nuclear installation sites in low seismicity regions, organized by IAEA (International Atomic Energy Agency), Brussels, 4-7 November 2025.

6. PUBLICATIONS

6.1. Peer-reviewed publications

6.1.1. PUBLISHED

Vanneste, K., Onvani, M., Rapagnani, G., & Van Noten, K. (2026). The BELSHAKE Database of Earthquake Ground Motion in Belgium. Seismological Research Letters, https://doi.org/10.1785/0220250321

6.1.2. SUBMITTED

Onvani M. & Vanneste K., Analysis of the site-specific and regional components of kappa across crustal domains in Belgium. Submitted to Seismica, https://doi.org/10.21203/rs.3.rs-8484319/v1 (preprint).

6.2. Peer-reviewed publications mentioning the BELSHAKE project

6.2.1. PUBLISHED

Edwards, B., Ntinalexis, M., Vanneste, K., & Elk, J. V. (2025). Reconstruction of Clipped Time-History Records , with Application to Ground Motions from the ML 5.8 Roermond, Netherlands, 1992 Earthquake. Seismological Research Letters, p. 1–16, https://doi.org/10.1785/%25200220250139

6.2.2. SUBMITTED

Vanneste, K., & Camelbeeck, T., Recalibration of the local magnitude scale of the Belgian seismic network: Improved models of geometrical spreading and path duration. Submitted to Seismica, https://doi.org/10.21203/rs.3.rs-7850272/v1 (preprint).

6.3. Publications without peer review

Vanneste, K., & Onvani, M. (2024). Waveform processing in BELSHAKE, the Belgian ground-motion database. *18Th World Conference on Earthquake Engineering*, 12, https://www.researchgate.net/publication/382067932

7. ACKNOWLEDGEMENTS

We wish to thank all members, past and present, of the ROB seismology team who contributed to maintenance of the Belgian seismic network, routine phase picking and creation of the ROB seismic databases. In particular, we appreciated the help from Giovanni Rapagnani, who maintains the ROB waveform archive, station database and instrument response information, and who helped us deciphering response information for legacy data. We are grateful to the members of the follow-up committee, and in particular to Paola Traversa for inspiring discussions and sharing data and experiences from the RESIF/EPOS-France database.

REFERENCES

Abrahamson, N. A., Silva, W. J., & Kamai, R. (2014). Summary of the ASK14 ground motion relation for active crustal regions. *Earthquake Spectra*, *30*(3), 1025–1055. https://doi.org/10.1193/070913EQS198M

Akkar, S., Sandıkkaya, M. a., & Bommer, J. J. (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. https://doi.org/10.1007/s10518-013-9461-4

Akkar, S., Sandıkkaya, M. A., Şenyurt, M., Azari Sisi, A., Ay, B. Ö., Traversa, P., Douglas, J., Cotton, F., Luzi, L., Hernandez, B., & Godey, S. (2014). Reference database for seismic ground-motion in Europe (RESORCE). *Bulletin of Earthquake Engineering*, *12*(1), 311–339. https://doi.org/10.1007/s10518-013-9506-8

Al Atik, L., Abrahamson, N. A., Bommer, J. J., Scherbaum, F., Cotton, F., & Kuehn, N. (2010). The variability of ground-motion prediction models and its components. *Seismological Research Letters*, *81*(5), 794–801. https://doi.org/10.1785/gssrl.81.5.794

Al Atik, L., Kottke, A., Abrahamson, N., & Hollenback, J. (2014). Kappa (κ) Scaling of Ground-Motion Prediction Equations Using an Inverse Random Vibration Theory Approach. *Bulletin of the Seismological Society of America*, *104*(1), 336–346. https://doi.org/10.1785/0120120200

Ameri, G., Drouet, S., Traversa, P., Bindi, D., & Cotton, F. (2017). Toward an empirical ground motion prediction equation for France: Accounting for regional differences in the source stress parameter. *Bulletin of Earthquake Engineering*, *15*(11), 4681–4717. https://doi.org/10.1007/s10518-017-0171-1

Ancheta, T. D., Darragh, R. B., Stewart, J. P., Seyhan, E., Silva, W. J., Chiou, B. S. J., Wooddell, K. E., Graves, R. W., Kottke, A. R., Boore, D. M., Kishida, T., & Donahue, J. L. (2014). NGA-West2 database. *Earthquake Spectra*, *30*(3), 989–1005. https://doi.org/10.1193/070913EQS197M

Anderson, J. G., & Hough, S. E. (1984). A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies. *Bulletin of the Seismological Society of America*, *74*(5), 1969–1993.

Atkinson, G. M. (2008). Ground-Motion Prediction Equations for Eastern North America from a Referenced Empirical Approach: Implications for Epistemic Uncertainty. *Bulletin of the Seismological Society of America*, *98*(3), 1304–1318. https://doi.org/10.1785/0120070199

Atkinson, G. M. (2015). Ground-Motion Prediction Equation for Small-to-Moderate Events at Short Hypocentral Distances , with Application to Induced-Seismicity Hazards. *Bulletin of the Seismological Society of America*, *105*(2A), 981–992. https://doi.org/10.1785/0120140142

Atkinson, G. M., & Boore, D. M. (1995). Ground-motion relations for eastern North America. *Bulletin - Seismological Society of America*, *85*(1), 17–30. https://doi.org/10.1785/bssa0850010017

Atkinson, G. M., & Boore, D. M. (2006). Earthquake Ground-Motion Prediction Equations for Eastern North America. *Bulletin of the Seismological Society of America*, *96*(6), 2181–2205. https://doi.org/10.1785/0120050245

Baptie, B. (2017). *Earthquake Seismology 2015/2016* (p. 48). British Geological Survey.

Beauval, C., Tasan, H., Laurendeau, a., Delavaud, E., Cotton, F., Gueguen, P., & Kuehn, N. (2012). On the Testing of Ground-Motion Prediction Equations against Small-Magnitude Data. *Bulletin of the Seismological Society of America*, *102*(5), 1994–2007. https://doi.org/10.1785/0120110271

Belgisch Instituut voor Normalisatie (BIN). (2002). *Eurocode 8: Ontwerp en dimensionering van aardschokbestendige structuren—Deel 1-1: Algemene regels—Aardschokken en algemene regels voor structuren—Belgische toepassingsrichtlijn* (Issue NBN-ENV 1998-1-1: 2002 NAD-E/N/F, p. 17).

Biasi, G. P., & Smith, K. D. (2001). *Site Effects for Seismic Monitoring Stations in the Vicinity of Yucca Mountain, Nevada*. US DOE/University and Community College System of Nevada (UCCSN) Cooperative Agreement.

Bindi, D., Cotton, F., Kotha, S. R., Bosse, C., Stromeyer, D., & Grünthal, G. (2017). Application-driven ground motion prediction equation for seismic hazard assessments in non-cratonic moderate-seismicity areas. *Journal of Seismology*, *21*(5), 1201–1218. https://doi.org/10.1007/s10950-017-9661-5

Bindi, D., Massa, M., Luzi, L., Ameri, G., Pacor, F., Puglia, R., & Augliera, P. (2014). Pan-European ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods up to 3.0 s using the RESORCE dataset. *Bulletin of Earthquake Engineering*, *12*(1), 391–430. https://doi.org/10.1007/s10518-013-9525-5

Bommer, J. J., & Scherbaum, F. (2008). The use and misuse of logic trees in probabilistic seismic hazard analysis. *Earthquake Spectra*, *24*(4), 997–1009. https://doi.org/10.1193/1.2977755

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

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*(6), 3202–3216. https://doi.org/10.1785/0120090056

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. https://doi.org/10.1785/0120090400

Boore, D. M., & Atkinson, G. M. (2008). Ground-Motion Prediction Equations for the Average Horizontal Component of PGA, PGV, and 5%-Damped PSA at Spectral Periods between 0.01 s and 10.0 s. *Earthquake Spectra*, *24*(1), 99–138. https://doi.org/10.1193/1.2830434

Boore, D. M., Stewart, J. P., Seyhan, E., & Atkinson, G. M. (2014). NGA-West2 equations for predicting PGA, PGV, and 5% damped PSA for shallow crustal earthquakes. *Earthquake Spectra*, *30*(3), 1057–1085. https://doi.org/10.1193/070113EQS184M

Boore, D. M., & Thompson, E. M. (2014). Path durations for use in the stochastic-method simulation of ground motions. *Bulletin of the Seismological Society of America*, *104*(5), 2541–2552. https://doi.org/10.1785/0120140058

Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes. *Journal of Geophysical Research*, *75*(26), 4997–5009. https://doi.org/10.1029/JB075i026p04997

Bruyninx, C., Van Noten, K., Bamahry, F., Fabian, A., Zeckra, M., Legrand, J., Rapagnani, G., Frederick, B., Miglio, A., Jacques, D., Mesmaker, D., Moyaert, A., & Lecocq, T. (2023). *EPOS-BE - Upgrade of Belgian federal research infrastructure for EPOS* (p. 51). Royal Observatory of Belgium.

Bureau voor Normalisatie (NBN). (2011). *Eurocode 8—Ontwerp en berekening van aardbevingsbestendige constructies—Deel 1: Algemene regels, seismische belastingen en regels voor gebouwen—Nationale Bijlage* (Issue NBN EN 1998-1-ANB:2011(N)).

Buscetti, M., Traversa, P., Perron, V., Rischette, P., Hollender, F., & Shible, H. (2025). Epos-France Database of Earthquake Ground Motion for Mainland France: 1996 – 2021 Updated Version. *Seismological Research Letters*, *96*(2A), 1214–1226. https://doi.org/10.1785/0220240053

Camelbeeck, T. (1985). Some notes concerning the seismicity in Belgium. Magnitude scale. Detection capability of the Belgian seismic stations. In P. Melchior (Ed.), *Seismic activity in Western Europe, with particular consideration to the Liege earthquake of November 8, 1983* (Issue 144, pp. 99–108). Reidel.

Camelbeeck, T. (1993). Mécanisme au foyer des tremblements de terre et contraintes tectoniques: Le cas de la zone intraplaque belge. In *Série Géophysique: (N° hors-S*. Observatoire Royal de Belgique.

Camelbeeck, T., Van Noten, K., Lecocq, T., & Hendrickx, M. (2022). The damaging character of shallow 20th century earthquakes in the Hainaut coal area (Belgium). *Solid Earth*, *13*(3), 469–495. https://doi.org/10.5194/se-13-469-2022

Campbell, K. W. (2003). Prediction of strong ground motion using the hybrid empirical method and its use in the development of ground-motion (attenuation) relations in Eastern North America. *Bulletin of the Seismological Society of America*, *93*(3), 1012–1033.

Campbell, K. W., & Bozorgnia, Y. (2014). NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra. *Earthquake Spectra*, *30*(3), 1087–1114. https://doi.org/10.1193/062913EQS175M

Cataldi, L., Poggi, V., Costa, G., Parolai, S., & Edwards, B. (2022). Parametric spectral inversion of seismic source, path and site parameters: Application to northeast Italy. *Geophysical Journal International*, *232*(3), 1926–1943. https://doi.org/10.1093/gji/ggac431

Cauzzi, C., Faccioli, E., Vanini, M., & Bianchini, A. (2015). Updated predictive equations for broadband (0.01–10 s) horizontal response spectra and peak ground motions, based on a global dataset of digital acceleration records. *Bulletin of Earthquake Engineering*, *13*(6), 1587–1612. https://doi.org/10.1007/s10518-014-9685-y

CEN (Comité Européen de Normalisation). (2004). *Eurocode 8: Design of structures for earthquake resistance—Part 1: General rules, seismic actions and rules for buildings* (Issue EN 1998-1, p. 44). European Committee for Standardization.

Chiou, B., Darragh, R., Gregor, N., & Silva, W. (2008). NGA Project Strong-Motion Database. *Earthquake Spectra*, *24*(1), 23–44. https://doi.org/10.1193/1.2894831

Chiou, B. S. J., & Youngs, R. R. (2014). Update of the Chiou and Youngs NGA model for the average horizontal component of peak ground motion and response spectra. *Earthquake Spectra*, *30*(3), 1117–1153. https://doi.org/10.1193/072813EQS219M

Cobum, A. W., Spence, R. J. S., & Pomonis, A. (1992). Factors determining human casualty levels in earthquakes: Mortality prediction in building collapse. *Tenth World Conference on Earthquake Engineering*, 5989–5994.

Danciu, L., Giardini, D., Weatherill, G., Basili, R., Nandan, S., Rovida, A., Bard, P., Pagani, M., Reyes, C., Sesetyan, K., Villanova, S., & Cotton, F. (2024). *The 2020 European Seismic Hazard Model: Overview and Results*. (January), 1–36.

Dawood, H. M., Rodriguez-Marek, A., Bayless, J., Goulet, C., & Thompson, E. (2016). A Flatfile for the KiK-net Database Processed Using an Automated Protocol. *Earthquake Spectra*, *32*(2), 1281–1302. https://doi.org/10.1193/071214eqs106

Department of Geosciences, Bensberg Observatory, U. of C. (2016). *Bensberg Earthquake Network, International Federation of Digital Seismograph Networks*. https://doi.org/10.7914/SN/BQ

Derras, B., Bard, P. Y., & Cotton, F. (2014). Towards fully data driven ground-motion prediction models for Europe. *Bulletin of Earthquake Engineering*, *12*(1), 495–516. https://doi.org/10.1007/s10518-013-9481-0

Douglas, J. (2019). *Ground motion prediction equations 1964 2019* (p. 650). University of Strathclyde. http://www.gmpe.org.uk

Douglas, J., Aldama-Bustos, G., Tallett-Williams, S., Daví, M., & Tromans, I. J. (2024). Ground-motion models for earthquakes occurring in the United Kingdom. *Bulletin of Earthquake Engineering*, *22*(11), 5953–5954. https://doi.org/10.1007/s10518-024-01992-z

Douglas, J., & Aochi, H. (2008). A survey of techniques for predicting earthquake ground motions for engineering purposes. *Surveys in Geophysics*, *29*(3), 187–220. https://doi.org/10.1007/s10712-008-9046-y

Douglas, J., Edwards, B., Convertito, V., Sharma, N., Tramelli, A., Kraaijpoel, D., Cabrera, B. M., Maercklin, N., & Troise, C. (2013). Predicting ground motion from induced earthquakes in geothermal areas. *Bulletin of the Seismological Society of America*, *103*(3), 1875–1897. https://doi.org/10.1785/0120120197

Douglas, J., Gehl, P., Bonilla, L. F., & Gélis, C. (2010). A κ model for mainland France. *Pure and Applied Geophysics*, *167*(11), 1303–1315. https://doi.org/10.1007/s00024-010-0146-5

Drouet, S., Ameri, G., Le Dortz, K., Secanell, R., & Senfaute, G. (2020). A probabilistic seismic hazard map for the metropolitan France. *Bulletin of Earthquake Engineering*, *18*(5), 1865–1898. https://doi.org/10.1007/s10518-020-00790-7

Drouet, S., & Cotton, F. (2015). Regional stochastic GMPEs in low-seismicity areas: Scaling and aleatory variability analysis—Application to the French Alps. *Bulletin of the Seismological Society of America*, *105*(4), 1883–1902. https://doi.org/10.1785/0120140240

Drouet, S., Cotton, F., & Guéguen, P. (2010). VS30, κ, regional attenuation and MW from accelerograms: Application to magnitude 3-5 French earthquakes. *Geophysical Journal International*, *182*(2), 880–898. https://doi.org/10.1111/j.1365-246X.2010.04626.x

Edwards, B., & Fäh, D. (2013). A Stochastic Ground-Motion Model for Switzerland. *Bulletin of the Seismological Society of America*, *103*(1), 78–98. https://doi.org/10.1785/0120110331

El Bouch, A., Camelbeeck, T., & Martin, H. (2002). Atténuation des ondes sismiques en Belgique et dans les régions limitrophes à partir de la coda des tremblements de terre locaux. *Geologica Belgica*, *5*(1–2), 17–29.

Gariel, J. C., Horrent, C., Jongmans, D., & Camelbeeck, T. (1994). Strong ground motion computation of the 1992 Roermond earthquake, the Netherlands, from linear methods using locally recorded aftershocks. *Geologie En Mijnbouw*, *73*(2–4), 315–321.

Georges, P. (2025). *Developing a pan-European ground-motion model: Towards improved modeling of apparent anelastic attenuation in shallow seismicity* (p. 23).

Ghasemi, H., Allen, T., Sunketa, A., & Mckenna, C. (2026). AGMD: A Ground-Motion Database for the Australian Continental Crust. *Seismological Research Letters*.

Goraj, V. (2000). Etude de l'atténuation des ondes S dans la croûte en Belgique. In *Faculté des Sciences*. Université Catholique de Louvain.

Goulet, C. A., Kishida, T., Ancheta, T. D., Cramer, C. H., Darragh, R. B., Silva, W. J., Hashash, Y. M. A., Harmon, J., Parker, G. A., Stewart, J. P., & Youngs, R. R. (2021). PEER NGA-East database. *Earthquake Spectra*, *37*(1_suppl), 1331–1353. https://doi.org/10.1177/87552930211015695

Goulet, C. A., Kishida, T., Ancheta, T. D., Cramer, C. H., Darragh, R. B., Silva, W. J., Hashash, Y. M. A., Harmon, J., Stewart, J. P., Wooddell, K. E., & Youngs, R. R. (2014). *PEER NGA-East Database* (p. 98). Pacific Earthquake Engineering Research Center.

Grünthal, G., Stromeyer, D., Bosse, C., Cotton, F., & Bindi, D. (2018). The probabilistic seismic hazard assessment of Germany—Version 2016, considering the range of epistemic uncertainties and aleatory variability. *Bulletin of Earthquake Engineering*, *16*, 4339–4395. https://doi.org/10.1007/s10518-018-0315-y

Haak, H. W., van Bodegraven, J. A., Sleeman, R., Verbeiren, R., Ahorner, L., Meidow, H., Grünthal, G., Hoang-Trong, P., Musson, R. M. W., Henni, P., Schenková, Z., & Zimová, R. (1994). The macroseismic map of the 1992 Roermond earthquake, the Netherlands. *Geologie En Mijnbouw*, *73*(2–4), 265–270.

Heimann, S., Kriegerowski, M., Isken, M., Cesca, S., Daout, S., Grigoli, F., Juretzek, C., Megies, T., Nooshiri, N., Steinberg, A., Sudhaus, H., Vasyura-Bathke, H., Willey, T., & Dahm, T. (2017). *Pyrocko—An open-source seismology toolbox and library* [Computer software]. GFZ Data Services. https://doi.org/10.5880/GFZ.2.1.2017.001

Hinzen, K. (2019). Engineering Seismological Parameters of Two ML 4.3 Intraplate Earthquakes in the Northern Rhine Area, Germany. *Seismological Research Letters*, *90*(3), 1258–1267. https://doi.org/10.1785/0220180265

Hinzen, K.-G., Reamer, S. K., & Fleischer, C. (2020). Seismicity in the Northern Rhine Area (1995–2018). *Journal of Seismology*. https://doi.org/10.1007/s10950-020-09976-7

Hobiger, M., Bard, P. -Y., Cornou, C., & Le Bihan, N. (2009). Single station determination of Rayleigh wave ellipticity by using the random decrement technique (RayDec). *Geophysical Research Letters*, *36*(14). https://doi.org/10.1029/2009GL038863

Johnson, P. C. D. (2014). Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models. *Methods in Ecology and Evolution*, *5*(9), 944–946.

Kale, Ö., & Akkar, S. (2013). A new procedure for selecting and ranking ground-motion prediction equations (GMPEs): The Euclidean distance-based ranking (EDR) method. *Bulletin of the Seismological Society of America*, *103*(2 A), 1069–1084. https://doi.org/10.1785/0120120134

Kaneko, Y., & Shearer, P. M. (2014). Seismic source spectra and estimated stress drop derived from cohesive-zone models of circular subshear rupture. *Geophysical Journal International*, *197*(2), 1002–1015. https://doi.org/10.1093/gji/ggu030

Kenny, C. (2009). *Why do People Die in Earthquakes? The Costs, Benefits and Institutions of Disaster Risk Reduction in Developing Countries* (p. 40). The World Bank.

KNMI. (1993). *Netherlands Seismic and Acoustic Network. Royal Netherlands Meteorological Institute*. https://doi.org/10.21944/e970fd34-23b9-3411-b366-e4f72877d2c5

Kotha, S. R., & Traversa, P. (2024). A Bayesian update of Kotha et al. (2020) ground-motion model using Résif dataset. *Bulletin of Earthquake Engineering*, *22*(4), 2267–2293. https://doi.org/10.1007/s10518-023-01853-1

Kotha, S. R., Weatherill, G., Bindi, D., & Cotton, F. (2020). A regionally-adaptable ground-motion model for shallow crustal earthquakes in Europe. *Bulletin of Earthquake Engineering*, *18*(9), 4091–4125. https://doi.org/10.1007/s10518-020-00869-1

Krischer, L., Smith, J., Lei, W., Lefebvre, M., Ruan, Y., de Andrade, E. S., Podhorszki, N., Bozdağ, E., & Tromp, J. (2016). An Adaptable Seismic Data Format. *Geophysical Journal International*, *207*(2), 1003–1011. https://doi.org/10.1093/gji/ggw319

Ktenidou, O., Abrahamson, N. A., Darragh, R. B., & Silva, W. J. (2016). *A Methodology for the Estimation of Kappa (κ) from Large Datasets: Example Application to Rock Sites in the NGA-East Database and Implications on Design Motions*. PEER.

Ktenidou, O. J., Abrahamson, N. A., Drouet, S., & Cotton, F. (2015). Understanding the physics of kappa (κ): Insights from a downhole array. *Geophysical Journal International*, *203*(1), 678–691. https://doi.org/10.1093/gji/ggv315

Ktenidou, O. J., Darragh, R. B., Silva, W. J., Traversa, P., & Bozorgnia, Y. (2021). *Seismic Hazard Research: Kappa and high-frequency ground motion effects at hard-rock sites*. EPRI.

Ktenidou, O.-J., Cotton, F., Abrahamson, N. a., & Anderson, J. G. (2014). Taxonomy of κ: A Review of Definitions and Estimation Approaches Targeted to Applications. *Seismological Research Letters*, *85*(1), 135–146. https://doi.org/10.1785/0220130027

Luzi, L., Puglia, R., Russo, E., D’Amico, M., Felicetta, C., Pacor, F., Lanzano, G., Çeken, U., Clinton, J., Costa, G., Duni, L., Farzanegan, E., Gueguen, P., Ionescu, C., Kalogeras, I., Özener, H., Pesaresi, D., Sleeman, R., Strollo, A., & Zare, M. (2016). The Engineering Strong-Motion Database: A Platform to Access Pan-European Accelerometric Data. *Seismological Research Letters*, *87*(4), 987–997. https://doi.org/10.1785/0220150278

Manchuel, K., Traversa, P., Baumont, D., Cara, M., Nayman, E., & Durouchoux, C. (2018). The French seismic CATalogue (FCAT-17). *Bulletin of Earthquake Engineering*, *16*(6), 2227–2251. https://doi.org/10.1007/s10518-017-0236-1

Mayor, J., Traversa, P., Calvet, M., & Margerin, L. (2017). Tomography of crustal seismic attenuation in Metropolitan France: Implications for seismicity analysis. *Bulletin of Earthquake Engineering*, *16*(6), 2195–2210. https://doi.org/10.1007/s10518-017-0124-8

Megies, T., Beyreuther, M., Barsch, R., Krischer, L., & Wassermann, J. (2011). ObsPy—What can it do for data centers and observatories? *Annals of Geophysics*, *54*(1), 47–58. https://doi.org/10.4401/ag-4838

Morasca, P., D’Amico, M., Spallarossa, D., Bindi, D., Picozzi, M., Oth, A., & Pacor, F. (2025). GITpy: A Python Implementation of the Generalized Inversion Technique. *Seismological Research Letters*. https://doi.org/10.1785/0220250042

Moratto, L., Panebianco, S., Satriano, C., Stabile, T. A., & Priolo, E. (2025). Using ambient noise k0 estimation to improve microearthquake source parameters assessment in the High Agri Valley. *Geophysical Journal International*, *242*(2), 1–19. https://doi.org/10.1093/gji/ggaf191

Mosca, I., Sargeant, S., Baptie, B., Musson, R. M. W., & Pharaoh, T. C. (2022). The 2020 national seismic hazard model for the United Kingdom. *Bulletin of Earthquake Engineering*, *20*(2), 633–675. https://doi.org/10.1007/s10518-021-01281-z

Munafò, I., Malagnini, L., & Chiaraluce, L. (2016). On the Relationship between M w and M L for Small Earthquakes. *Bulletin of the Seismological Society of America*, *106*(5), 2402–2408. https://doi.org/10.1785/0120160130

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R 2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Neefs, B., Van Noten, K., Vanneste, K., & Camelbeeck, T. (2025). The Belgian traditional macroseismic (BTM) database of the twentieth century. *Journal of Seismology*, *29*(5), 1005–1036. https://doi.org/10.1007/s10950-024-10266-9

Oth, A. (2013). On the characteristics of earthquake stress release variations in Japan. *Earth and Planetary Science Letters*, *377–378*, 132–141. https://doi.org/10.1016/j.epsl.2013.06.037

Oth, A., Miyake, H., & Bindi, D. (2017). On the relation of earthquake stress drop and ground motion variability. *Journal of Geophysical Research: Solid Earth*, 1–19. https://doi.org/10.1002/2017JB014026

Pacor, F., Paolucci, R., Luzi, L., Sabetta, F., Spinelli, A., Gorini, A., Nicoletti, M., Marcucci, S., Filippi, L., & Dolce, M. (2011). Overview of the Italian strong motion database ITACA 1.0. *Bulletin of Earthquake Engineering*, *9*(6), 1723–1739. https://doi.org/10.1007/s10518-011-9327-6

Pagani, M., Monelli, D., Weatherill, G., Danciu, L., Crowley, H., Silva, V., Henshaw, P., Butler, L., Nastasi, M., Panzeri, L., Simionato, M., & Vigano, D. (2014). OpenQuake Engine: An Open Hazard (and Risk) Software for the Global Earthquake Model. *Seismological Research Letters*, *85*(3), 692–702. https://doi.org/10.1785/0220130087

Paolucci, R., Pacor, F., Puglia, R., Ameri, G., Cauzzi, C., & Massa, M. (2011). Record processing in ITACA, the new Italian strong-motion database. In S. Akkar, P. Gülkan, & T. van Eck (Eds.), *Earthquake Data in Engineering Seismology: Predictive Models, Data Management and Networks* (pp. 99–113). Springer. https://doi.org/10.1007/978-94-007-0152-6_8

Perron, V., Hollender, F., Bard, P., Gélis, C., Guyonnet-Benaize, C., Hernandez, B., & Ktenidou, O. (2017). Robustness of Kappa (κ) Measurement in Low-to-Moderate Seismicity Areas: Insight from a Site-Specific Study in Provence, France. *Bulletin of the Seismological Society of America*, *107*(5), 2272–2292. https://doi.org/10.1785/0120160374

Perron, V., Laurendeau, A., Hollender, F., Bard, P.-Y., Gélis, C., Traversa, P., & Drouet, S. (2018). Selecting time windows of seismic phases and noise for engineering seismology applications: A versatile methodology and algorithm. *Bulletin of Earthquake Engineering*, *16*, 2211–2225. https://doi.org/10.1007/s10518-017-0131-9

Puglia, R., Russo, E., Pacor, F., Luzi, L., & Amico, M. D. (2014). A Software for the Processing of the Waveforms of the Italian Strong Motion Database. *Second European Conference on Earthquake Engineering and Seismology*, 1–3.

Ramadan, F., Lanzano, G., Pacor, F., Felicetta, C., Smerzini, C., & Traversa, P. (2024). Adjusting an active shallow crustal ground motion model to regions with scarce data: Application to France. *Bulletin of Earthquake Engineering*, *22*(8), 3727–3751. https://doi.org/10.1007/s10518-024-01890-4

Rietbrock, A., & Edwards, B. (2019). Update of the UK Stochastic Ground Motion Model. *SECED 2019 Conference*, 10.

Rietbrock, A., Strasser, F., & Edwards, B. (2013). A Stochastic Earthquake Ground-Motion Prediction Model for the United Kingdom. *Bulletin of the Seismological Society of America*, *103*(1), 57–77. https://doi.org/10.1785/0120110231

Sargeant, S. L., Stafford, P. J., Lawley, R., Weatherill, G., Weston, A.-J. S., Bommer, J. J., Burton, P. W., Free, M., Musson, R. M. W., Kuuyuor, T., & Rossetto, T. (2008). Observations from the

Folkestone, U.K., earthquake of 28 April 2007. *Seismological Research Letters*, *79*(5), 672–687.

Satriano, C. (2025). *SourceSpec – Earthquake source parameters from P- or S-wave displacement spectra (1.8)* [Computer software]. https://doi.org/10.5281/ZENODO.3688587

Satriano, C., Mejjia Uquiche, A. R., & Saurel, J. M. (2016). Spectral Estimation of Seismic Moment, Corner Frequency and Radiated Energy for Earthquakes in the Lesser Antilles. *American Geophysical Union, Fall Meeting*, S13A-2518. https://ui.adsabs.harvard.edu/abs/2016AGUFM.S13A2518S/abstract

Scherbaum, F., Cotton, F., & Smit, P. (2004). On the use of response spectral-reference data for the selection and ranking of ground-motion models for seismic-hazard analysis in regions of moderate seismicity: The case of rock motion. *Bulletin of the Seismological Society of America*, *94*(6), 2164–2185. https://doi.org/10.1785/0120030147

Scherbaum, F., Delavaud, E., & Riggelsen, C. (2009). Model selection in seismic hazard analysis: An information-theoretic perspective. *Bulletin of the Seismological Society of America*, *99*(6), 3234–3247. https://doi.org/10.1785/0120080347

Seabold, S., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. *9th Python in Science Conference*.

Stafford, P. J. (2014). Crossed and nested mixed-effects approaches for enhanced model development and removal of the ergodic assumption in empirical ground-motion models. *Bulletin of the Seismological Society of America*, *104*(2), 702–719. https://doi.org/10.1785/0120130145

Stafleu, J., Kraaijpoel, D., & Osinga, S. (2024). *Vs30 map of the Netherlands—Update of the proof-of-concept map of 2022* (p. 30). TNO.

Stanko, D., Markušić, S., Ivančić, I., Mario, G., & Gülerce, Z. (2017). Preliminary Estimation of Kappa Parameter in Croatia. *IOP Conference Series: Earth and Environmental Science*, *95*, 032014. https://doi.org/10.1088/1755-1315/95/3/032014

Toro, G. R. (2002). *Modification of the Toro et al. (1997) attenuation equations for large magnitudes and short distances* (Issue 1997, pp. 1–10).

Traversa, P., Maufroy, E., Hollender, F., Perron, V., Bremaud, V., Shible, H., Drouet, S., Guéguen, P., Langlais, M., Wolyniec, D., Péquegnat, C., & Douste-Bacque, I. (2020). RESIF RAP and RLBP Dataset of Earthquake Ground Motion in Mainland France. *Seismological Research Letters*, *91*(4), 2409–2424. https://doi.org/10.1785/0220190367

Van Houtte, C., Drouet, S., & Cotton, F. (2011). Analysis of the origins of κ (Kappa) to compute hard rock to rock adjustment factors for GMPEs. *Bulletin of the Seismological Society of America*.

Van Noten, K., Lecocq, T., Goffin, C., Meyvis, B., Molron, J., Debacker, T. N., & Devleeschouwer, X. (2022). Brussels' bedrock paleorelief from borehole-controlled power laws linking polarised H/V resonance frequencies and sediment thickness. *Journal of Seismology*, *26*, 35–55. https://doi.org/10.1007/s10950-021-10039-8

Van Noten, K., Lecocq, T., Sira, C., Hinzen, K. G., & Camelbeeck, T. (2017). Path and site effects deduced from merged transfrontier internet macroseismic data of two recent M4 earthquakes in northwest Europe using a grid cell approach. *Solid Earth*, *8*(2), 453–477. https://doi.org/10.5194/se-8-453-2017

Vanneste, K., & Camelbeeck, T. (n.d.). Recalibration of the local magnitude scale of the Belgian seismic network: Improved models of geometrical spreading and path duration. *Seismica?*

Vanneste, K., Neefs, B., & Camelbeeck, T. (2024). Testing the applicability of ground motion prediction equations for the Hainaut region (Belgium) using intensity data. *Bulletin of Earthquake Engineering*, (0123456789). https://doi.org/10.1007/s10518-024-01958-1

Vanneste, K., & Onvani, M. (2024). Waveform processing in BELSHAKE, the Belgian ground-motion database. *18Th World Conference on Earthquake Engineering*, 12. https://www.researchgate.net/publication/382067932

Vanneste, K., Vleminckx, B., Verbeeck, K., & Camelbeeck, T. (2014). Development of seismic hazard maps for Belgium. In C. Butenweg & D. Kaiser (Eds.), *DGEB Workshop "Seismic Hazard Harmonisation in Europe (SHARE)"* (pp. 61–68). DGEB.

Vanneste, K., Vleminckx, B., Verbeeck, K., & Camelbeeck, T. (2015). *Seismic Hazard Re-Evaluation of the Nuclear Power Plant Sites Doel and Tihange—Seismological Analyses and PSHA: Final Report* (p. 330). Royal Observatory of Belgium.

Verbeeck, K., Vanneste, K., & Camelbeeck, T. (2009). *Seismotectonic zones for probabilistic seismic-hazard assessment in Belgium* (p. 47). ONDRAF/NIRAS.

Villani, M., Polidoro, B., Mccully, R., Ader, T., Edwards, B., Rietbrock, A., Lubkowski, Z., Courtney, T. J., & Walsh, M. (2019). A Selection of GMPEs for the United Kingdom Based on Instrumental and Macroseismic Datasets. *Bulletin of the Seismological Society of America*, *109*(4), 1378–1400. https://doi.org/10.1785/0120180268

Weatherill, G., Kotha, S. R., & Cotton, F. (2020). A regionally-adaptable "scaled backbone" ground motion logic tree for shallow seismicity in Europe: Application to the 2020 European seismic hazard model. *Bulletin of Earthquake Engineering*, *18*(11), 5087–5117. https://doi.org/10.1007/s10518-020-00899-9

Yenier, E., & Atkinson, G. M. (2015). Regionally Adjustable Generic Ground-Motion Prediction Equation Based on Equivalent Point-Source Simulations: Application to Central and Eastern North America. *Bulletin of the Seismological Society of America*, *105*(4), 1989–2009. https://doi.org/10.1785/0120140332

Zhao, J. X., Zhang, J., Asano, A., Ohno, Y., Oouchi, T., Takahashi, T., Ogawa, H., Irikura, K., Thio, H. K., Somerville, P. G., & Fukushima, Y. (2006). Attenuation Relations of Strong Ground Motion in Japan Using Site Classification Based on Predominant Period. *Bulletin of the Seismological Society of America*, *96*(3), 898–913. https://doi.org/10.1785/0120050122