Development of a new approach to define reference ground motions applicable to existing strong-motion databases

Hussein Shible

▶ **To cite this version:**

Hussein Shible. Development of a new approach to define reference ground motions applicable to existing strong-motion databases. Earth Sciences. Université Grenoble Alpes [2020-..], 2021. English. NNT : 2021GRALU021 . tel-03467074

HAL Id: tel-03467074

https://theses.hal.science/tel-03467074v1

Submitted on 6 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE GRENOBLE ALPES

Spécialité: **Sciences de la Terre et de l'Univers et de l'Environnement (CESTUE)**

Arrêté ministériel : 25 mai 2016

Présentée par

Hussein SHIBLE

Thèse dirigée par **Philippe GUEGUEN,** directeur de recherche Université Gustave Eiffel, ISTerre, et
codirigée par **Fabrice HOLLENDER,** ingénieur-chercheur**,** Commissariat à l'énergie atomique et aux énergies alternatives (CEA)

préparée au sein du **Laboratoire Institut des Sciences de la Terre** dans **l'École Doctorale Terre, Univers, Environnement**

« Développement d'une nouvelle approche pour définir les mouvements sismiques de référence applicables aux bases de données existantes.»

« Development of a new approach to define reference ground motions applicable to existing strong-motion databases.»

Thèse soutenue publiquement le **13 Septembre 2021** devant le jury composé de :

Hiroshi KAWAZE
Professeur à Université de Kyoto, Japon, Rapporteur.
Adrian RODRIGUEZ-MAREK
Professeur à VirginiaTech, Etats-Unis, Rapporteur.
Fransesca PACOR
Ingénieur docteur à l'INGV, Italie, Examinatrice.
Pierre-Yves BARD
Ingénieur général des ponts des eaux et des forêts à ISTerre, Grenoble, invité.
Paola TRAVERSA
Ingénieur-chercheur, EDF, Aix-en-Provence, Co-encadrante invitée.
Fabrice HOLLENDER
Chercheur-Ingénieur, CEA, Cadarache, Codirecteur de thèse.
Philippe GUEGUEN
Directeur de recherche à ISTerre, Grenoble, Directeur de thèse.
Cécile CORNOU
Directrice de recherche IRD à lSTerre, Grenoble, Présidente du jury.

Résumé

La réalisation d'études de l'aléa sismiques site-spécifiques nécessite, d'une part, d'évaluer l'amplification locale due aux effets de site (méthodes numériques ou empiriques), et d'autre part, de disposer de mouvements "incidents" ou "de référence". Le mouvement de référence doit représenter les propriétés du substratum rocheux du site étudié, les caractéristiques des sources sismiques régionales et les paramètres d'atténuation crustale. Aujourd'hui, de tels mouvements ne sont toujours pas disponibles ou adaptés aux sites dont le substratum géologique est rigide (roche dure) à la surface. L'état actuel de la pratique pour surmonter ce problème est d'appliquer des corrections basées sur la physique pour tenir compte des différences spécifiques au site. Cependant, des études récentes ont montré que ces corrections sont à l'origine d'incertitudes importantes qui peuvent surestimer l'aléa sismique. De plus, des progrès significatifs ont été réalisés au cours des dernières années. Plusieurs méthodologies ont été proposées pour récupérer ces mouvements sismiques de référence, dont l'une est l'approche par déconvolution. L'approche de déconvolution repose sur la suppression des effets de site des enregistrements de surface pour obtenir le mouvement du substratum rocheux sous-jacent. Cette dernière approche a été validée sur la base de données Japonaise KiK-net bénéficiant d'un très haut niveau de caractérisation, ce qui permet d'obtenir des réponses théoriques du site. Malgré les progrès réalisés, la " transposition " de cette approche au contexte stable français et européen pose une question scientifique centrale : comment pallier le manque de caractérisation des stations accélérométriques européennes et l'absence d'enregistrements de fond des forages ? Dans ce travail de thèse, nous proposons les techniques d'inversion généralisées comme élément clé pour fournir des réponses empiriques aux sites. Les inversions généralisées réduisent les hypothèses sur les effets de site, présentes dans les estimations théoriques, et elles réduisent la nécessité de caractériser tous les sites à quelques sites seulement. Premièrement, une évaluation fiable des effets de site à partir d'inversions généralisées nécessite une condition de site de référence appropriée. Ici, nous proposons d'utiliser des estimations théoriques pour obtenir la réponse absolue du site d'une ou plusieurs stations, à imposer dans les inversions, en profitant des caractérisations de site existantes dans la base des données. De plus, nous considérons également une source d'incertitude dans les caractérisations des stations, qui est les hétérogénéités verticales à petite échelle dans la proche surface, et nous montrons qu'elle pourrait avoir un impact sur les estimations théoriques aux hautes fréquences. Deuxièmement, l'utilisation d'inversions généralisées nécessite l'exploration des incertitudes liées aux différentes possibilités d'application sur différentes configurations de jeux de données. Après avoir assuré une méthodologie robuste et fiable pour estimer les effets de site, nous complétons la validation de la déconvolution sur le réseau japonais KiK-net. Au final, nous dérivons une équation de prédiction des mouvements de référence du sol avec une méthodologie transposable au contexte français et européen. Le travail réalisé dans ce manuscrit servira de base à un futur travail qui sera mené sur les bases de données européennes.

Mots-clés: *Aléa sismique, Site-spécifique, Mouvements de référence, Approche de déconvolution, Inversions généralisées.*

Abstract

The realization of site-specific seismic hazard studies requires, on the one hand, to evaluate the local amplification due to site effects (numerical or empirical methods), and on the other hand, to have "incident" or "reference" movements. The reference motion should represent the bedrock properties of the studied site, the characteristics of the regional seismic sources, and the crustal attenuation parameters. Today, such motions are still not available or adapted for sites with rigid geological substratum (hard-rock) at the surface. The state-of-the-practice to overcome this issue is to apply physics-based corrections to account for site-specific differences. However, recent studies have shown that these corrections underly significant uncertainties that may overestimate the seismic hazard. Moreover, considerable progress has been made in the last few years. Several methodologies were proposed to retrieve these incidental seismic movements, one of which is the deconvolution approach. The deconvolution approach relies on removing site effects from surface recordings to obtain the underlying bedrock motion. The latter approach was validated on the Japanese KiK-net database benefiting from a very high level of characterization, which allows obtaining theoretical site responses. Despite the progress, the "transposition" of this approach to the stable French and European context raises a central scientific question: how to overcome the lack of characterization of European accelerometric stations and the absence of downhole recordings? In this PhD work, we propose generalized inversion techniques as a key element to deliver empirical site responses. Generalized inversions reduce the assumptions on site effects present in the theoretical estimations, and they reduce the necessity for site characterizations for all sites to only a few ones. First, reliable retrieval of site effects from generalized inversions requires a proper reference site condition. Here, we propose to use theoretical estimations to obtain the absolute site response of one or several stations, benefiting from existing site characterizations in the dataset. Also, we consider a source of uncertainty, which is the vertical small-scale heterogeneities in the near-surface, and we show that it could impact theoretical estimations in the high frequencies. Second, the use of generalized inversions requires exploring the related uncertainties due to different possibilities of application on different dataset configurations. After ensuring a robust and reliable estimate of site effects, we complete the validation of the deconvolution approach on the Japanese KiK-net network. In the end, we derive a reference ground-motion prediction equation with a methodology that can be transposable to the French and European context. The work presented in this document serves as a basis for future work in the French and European context.

Keywords: *Seismic hazard, Site-specific, Reference motion, Deconvolution approach, Generalized inversion techniques.*

Acknowledgments

A single-person effort can never achieve the success of research projects. All the factors that surround the work will definitely have a contribution. Before mentioning each, I should first thank the responsibles at both CEA and ISTerre for providing the opportunity and the continuity for such a project. I would also apologize to all the people that I could forget to mention in the following.

First of all, I want to thank Fabrice Hollender for his supervision of my PhD. Thanks a lot for your patience, Kindness, and support during the whole PhD, especially the last months, when my stresses were doubled due to successive Covid lockdowns. For sure, I will not remember all, but I would say simply thanks for everything. I know that words will never be enough. Thanks also to Paola Traversa for all the support and very helpful discussions. You will be one of the main reasons of success for this project. Thanks for everything. Thank you also, Philippe Gueguen, for always being by our side whenever we needed your help and for the continuous support. What I should not forget also, is that after doing my master's internship with Pierre-Yves Bard on a very close and related subject (to my PhD thesis), I learned lots of things, most of which I continued to use during my PhD. Thank you very much for everything.

Thanks to Adrian Rodriguez-Marek and Hiroshi Kawaze for accepting to be the reviewers of my PhD work. Thanks also to Cecile Cornou and Francesca Pacor for accepting to review my work and participate in the PhD defense jury.

I would thank all the people with whom I had contact during my PhD. Thanks to all the people I met at CEA, ISTerre and with whom I had discussions at different stages of my PhD, like Konstantinos Trevlopoulos, Alain Dujardin, Aurore Laurendeau, Vincent Perron, and Stephane Drouet. Thanks to the CEA team, especially Michèle and Franck, for their motivation and continuous support. Thanks to my colleague and friend Ioannis from the University of Thessaloniki. We had lots of exchanges on common subjects, and we had many interesting discussions. I would also thank the people who reviewed our work during the SIGMA-2 project meetings, Christophe Martin, Ian Main, Yousef Bozorgnia, Yoshi Fukushima.

During my PhD. I was lucky to meet several researchers, especially during the workshops we organized for the GITEC benchmark. Special thanks to Dino Bindi for all the nice, constructive, and interesting discussions, especially about generalized inversion techniques. I would say that I gained a lot from our discussions. I would also thank Adrien Oth, Ben Edwards, Hiroshi Kawase, Peter Klin, Raul Castro.

My endless thanks go to my family members. They did not contribute to this document's scientific content, but they did contribute to my support and internal serenity. Thank you very much for being by my side whenever I needed it. I dedicate all the successes I achieve in my life to my family.

Table of Contents

General Introduction

The earth is composed of several tectonic plates that are in continuous motion throughout geological ages. These movements lead to accumulated stresses at the earth's discontinuities, known as seismic faults. The stored energy keeps increasing over the years to reach a point where the frictional forces can no longer hold. At this point, the energy is released by rapid displacements across the entire fault and thus generates seismic waves in all directions. As the wavefronts propagate through the earth's crust, they undergo complex phenomena that are still not well constrained to this day. However, like other types of waves, seismic waves are expected to undergo attenuation along the followed crustal path, mainly in intrinsic and scattering forms. When the waves approach the earth's surface, they pass through less rigid and less dense materials due to both the decrease in the earth's pressure and the effects of geological weathering. The waves undergo additional modifications in the near-surface layers, known as site effects. These effects have been identified as amplification of amplitudes and elongation of durations of waves. Thus, a complete understanding of earthquake physics is conditioned by considering the different contributions of source, path, and site effects as schematized in Figure 0.1.

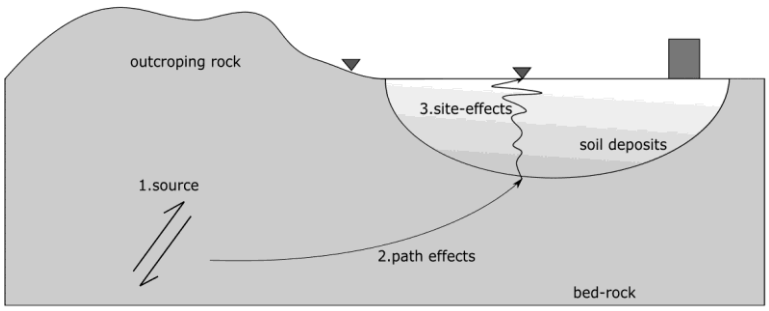

Figure 0.1: Illustrations of different contributions to seismic waves: source, path, and site effects.

Earthquakes are among the natural phenomena that cause the most damages and losses to human beings. Although people have witnessed earthquakes since early ages, the losses caused by earthquakes began to be greater with increased urban development. In recent decades, several devastating earthquakes were witnessed in different parts of the world, after which the human, economic and ecological consequences have been catastrophic. Some of the most remarkable events since the 1980s are the Michoacan earthquake (1985) in Mexico, followed by the earthquakes of Kobe (1995) in Japan and Haiti (2010) in Chile, and many others that resulted in great human and economic losses. Another dramatic event is the Tōhoku earthquake (2011) with magnitude 9, which caused a large tsunami on the northeast coast of Japan, resulting in the death of about 22,000 people and caused the Fukushima Daiichi nuclear power plant accident.

Since earthquake phenomena are unpredictable, decision-makers attempt to take action to mitigate seismic risks. These risks are generally seen as a combination of the seismic hazard in the studied region and the existing vulnerability and exposure. Here, engineering seismologists are generally asked to provide hazard evaluations that can be translated into engineering practices and recommendations. To this end, seismic hazard assessment (SHA) aims to provide ground motion intensities consistent with a low probability of being exceeded, which is an essential step in a) establishing and

revising building regulations and b) reassessing the safety of already constructed facilities and buildings.

Despite the scientific progress, many aspects of earthquake physics remain unclear. The prediction of the intensity and timing of an earthquake is not (yet) possible on this day. Thus, hazard evaluations are based on the assumption that past earthquakes will have the same spatial and temporal properties as future earthquakes. Once sufficient knowledge about past earthquakes is available, the hazard evaluation generally takes two forms, either (1) probabilistic (PSHA) or (2) deterministic (DHSA). First, PSHA has been widely used by experts and researchers over the last few decades for seismic safety decisions. The main output of a PSHA study is a hazard map for the desired ground motion intensity measure. The usual representation of a hazard map is in terms of a probability of exceedance of 10 percent for a specific intensity measure. These probabilities are based on a return period, which can vary with the overall objectives between 475 years for regular buildings and 10,000 years for critical facilities. Second, the DSHA study is generally performed by assuming a set of controlling seismic scenarios (hypothetical or realistic). The accuracy of DSHA output generally depends on uncertainties associated with the used parameters, unlike the probabilistic approach that combines sources of uncertainties in the form of logic trees to have a global view of embedded uncertainties.

Whatever the chosen approach (deterministic or probabilistic), the main element for hazard evaluations is the distribution of ground shaking amplitudes (e.g., response spectra) and other intensity measures, assessed by ground motion prediction equations (GMPEs). They are also called attenuation relationships or ground-motion models. GMPEs result from multivariate nonlinear mathematical regressions in which ground motion amplitudes are fitted to predefined functional forms. These functions generally take into account source, path attenuation, and site effects in different ways. To provide a simple model for engineering applications, researchers parametrize the different contributions with simple terms as variables. For example, earthquake magnitudes are used to account for the source effects, distance dependence to account for attenuation, and other geotechnical/geophysical parameters to account for site effects. Since GMPEs are defined over a specific variables range, they are considered applicable to future scenarios falling within this range, referred to as the domain of validity of the GMPE. Many discussions and efforts are being made to overcome this drawback in applications and extrapolate beyond validity domains, primarily by improving the understanding of earthquake physics and incorporating it into predictions. As not all earthquake-related phenomena are well constrained, the latter extrapolation steps face many difficulties in the application. Indeed, the work presented in the current text addresses some of the questions on site-related validity domains of GMPEs, where the details will be more clarified in the following paragraphs.

In the beginning, the spatial distribution of ground motion was assumed to be the same as the distribution at a single site when sampled over time. Such assumption is known in the literature as the ergodic assumption (Anderson and Brune, 1999). An ergodic assumption is unavoidable when no sufficient recordings are available, and it thus allowed combining recordings from regions with different characteristics in a GMPE determination. However, the availability of well-recorded ground motion has allowed researchers to begin searching for more accurate and region-specific ground motion

predictions, primarily by relaxing the ergodic assumption. This step in SHA requires integrating site- and path-specific effects and is thus expected to reduce uncertainties associated with GMPEs (Atik et al., 2010). As this last step still seems to be advanced today regarding path effects, partially non-ergodic GMPEs are addressed through site-specific ground motion predictions (Rodriguez- Marek et al., 2013; Kotha et al., 2016). In the context of critical facilities, precise estimations of the seismic hazard are essential, and site-specific SHA is a major concern.

Moreover, site-specific ground motion prediction generally attempts to account for the influence of local site conditions on ground motion. For this aim, two main approaches are followed, (1) a generic approach with simple site proxies and (2) an approach where full site-specific effects are considered. The generic approach consists of simplifying the practice by parameterization or categorization of site conditions in GMPEs. For example, the averaged S-wave velocity in the first 30 meters of soil (denoted by V_{S30}) is used to account for local impedance effects either directly or through using it to classify into site categories. Also, the high-frequency slope (κ) of Fourier spectra is introduced to account for shallow attenuation effects. The κ parameter was first introduced by Anderson and Hough (1984) as an observation on Fourier spectra, and then it became interpreted as a site-specific property (Boore, 2003). Also, other terms were explored, such as fundamental site resonance frequencies (f_0) or depth to bedrock (h_{800}). Though the parameterization approach seems relatively simple and practical for hazard evaluations, the site responses seem to be oversimplified, where important features are not taken into account (e.g., local resonance effects). On the other hand, a fully site-specific approach precisely evaluates site responses with ground response analyses (GRAs). Mainly, GRAs are performed with theoretical approaches based on reflectivity models in 1D conditions or numerical simulations of wave propagation in multidimensional site conditions, which provide site transfer functions (TFs) that reflect the site effects. Also, GRA can be done with empirical estimations that use available recorded data to estimate site amplifications with respect to a reference component. Here, the reference components can be either the horizontal components on a near outcropping bedrock (SSR) or vertical components (HVSR). Besides, generalized inversion techniques (GIT) is another tool of GRAs that benefits from the presence of several recordings at each station in a network to decompose the Fourier amplitude spectra into an independent source, path, and site effects.

Fully site-specific SHA is generally performed by estimating the hazard on reference rock sites and integrating the evaluated site-specific effects (in either generic or site-specific methods). In fact, reference site conditions are generally considered for sites with negligible local site effects. Thus, it is commonly considered that stiff hard-rock sites are good candidates of reference sites, which are the closest possible to seismic bedrock conditions (i.e., V_{S30} > 1500 m/s). Nevertheless, most existing strong-motion databases lack a sufficient number of accelerometric stations on hard-rock sites, making the reference conditions outside of the domain of validity of most GMPEs. In view of addressing this difficulty, the main approach proposed was to derive GMPEs on available data (host sites) and apply site-specific corrections to account for differences with reference sites (target sites). The latter approach is called hereafter Host-to-target-adjustments, denoted “HTTA”. Mainly, standard-rock sites (i.e., V_{S30} around 800 m/s) are the used host sites since they are the most rigid sites that lie

within the validity range of most GMPEs. These corrections account for two main points:

- V_S-correction accounts for impedance effects between host and target sites. This correction uses deep velocity profiles to estimate the site amplification by the quarter-wavelength (QWL) approximation (Boore, 2003). The adjustment factor is then obtained by the ratios of QWL-amplifications between the standard-rock and reference hard-rock sites.
- κ-correction accounts for the high-frequency attenuation effects through the κ parameter. It is generally considered that soft-rock sites have high κ values (~0.030 ms), while rigid hard-rock sites have significantly lower κ values (0.001-0.010 ms). This leads to predict a lack of attenuation for hard-rock sites compared to soft rock sites. Consequently, soft-rock motion should be amplified, mainly at high frequencies, to obtain hard-rock motion. Though the latter seems counterintuitive, it was adopted in several applications.

When applying V_S-κ corrections, several sources of uncertainties are worth to be discussed. First, deep "measured" V_S profiles used in the V_S correction are not present in most cases for both host and target sites. Instead, an interpolation using V_{S30} was proposed based on the few profiles proposed by Boore and Joyner (1997) for western U.S. regions (Cotton et al., 2006). In addition, V_{S30} values, which became an element to select deep V_S profiles, are inferred in many cases from geological and topographical correlations, which are subjected to huge uncertainties. Also, the QWL approach neglects all possible local resonance effects in site responses. On the other hand, both κ and V_{S30} values are not always estimated for all the host sites in the databases, and neither for target ones. Besides, the estimation of κ requires a sufficient number of well-recorded events to have reliable site-specific estimations. Indeed, κ estimations are subjected to huge uncertainties, as shown in several publications (e.g., Ktenidou et al., 2015). Moreover, the adjustment factors, defined in the Fourier domain, are not applied in a unique way. As GMPEs were derived in the response spectral domain, the HTTA application required the back and forth conversions to the Fourier with the help of the random vibration theory RVT and its inverse IRVT (Bora et al., 2015, 2016). Though they were accepted in the engineering seismology community, these processes include additional uncertainties and non-uniqueness of the solution.

In the aim to reduce the uncertainty sources in the classical HTTA approach, several alternatives were evoked in the last few years in order to avoid completely V_{S30}-κ corrections. One of these proposed alternatives is based on the correction of recordings to obtain hard-rock motion. Usually, reference bedrock corresponds to very rigid structures, i.e., V_{S30}>1500 m/s, which are close to deep-enough downhole sensor conditions (for instance, this case is available in the KiK-net network). However, downhole motion is affected by downgoing waves that induce destructive resonance frequencies and cannot fully represent outcropping rock motion. Cadet et al. (2012) proposed correcting downhole recordings from borehole effects by removing downgoing wave effects from recordings to obtain what represents outcropping rock motion. This procedure was thus tested in the work of Laurendeau et al. (2018) on the Japanese KiK-net data to obtain virtual hard-rock motion from borehole sensors. The

corrected dataset, named after DHcor, was then used to derive a GMPE with an extended validity range of V_{S30}.

However, downhole sensors are not usually present elsewhere in most existing strong-motion databases. Thus, Laurendeau et al. (2018) explored a different approach based on the correction of surface recordings. The correction was done by deconvolving site-effects estimated by theoretical 1D site responses. Theoretical site responses were possible because of the presence of characterized profiles for most of the KiK-net sites. Hence, the corrected dataset was called SURFcor and was used directly for an additional GMPE derivation with an extended V_{S30} range. The possibility of deriving GMPEs for both DHcor and SURFcor on KiK-net was advantageous from a methodological perspective.

Despite the advantages present in the deconvolution approach of surface data by Laurendeau et al. (2018) and the consistency with the results of DHcor, the deconvolution approach presented several limitations. First, they benefited from the presence of sensor pairs at surface and borehole levels in the Japanese KiK-net array. Besides, for site response estimations, V_S-profiles for almost all sites were essential. These put the approach's applicability to other datasets under question, especially that most existing strong-motion databases lack the information present on KiK-net. Consequently, we believe that the deconvolution approach needs additional work to ensure its applicability and transposability to other datasets, more precisely in the European/French context.

Within the PhD framework, we propose generalized inversion techniques (GIT) to overcome these challenges. First, site amplification functions obtained by GIT can be considered as empirical estimates of site effects as they result only from the Fourier spectra decomposition. Therefore, the 1D assumption is relaxed, and the methodology can be extended to all sites of interest. Besides, empirical estimations of site responses include all unconstrained controlling phenomena, unlike the theoretical estimates, where the level of input information highly affects their accuracy.

The main scientific questions that have been addressed in this PhD can be summarized in the following points:

- To understand the high-frequency discrepancy between 1DSH theoretical and empirical site response estimates highlighted by Laurendeau et al. (2018). Thus, we were interested in identifying, with a physical interpretation, possible sources of uncertainties that could bias 1DSH site responses.
- As we aim to use empirical estimates of generalized inversions, it is crucial to understand the proposed methods and highlight the best way to use them. Also, it is interesting to explore the epistemic uncertainty associated with GIT site terms in the framework of a precise site-response prediction.
- In Generalized inversions, the reference choice is an important step to ensure proper source and site separations. The current practice in GIT aims to fix the reference site response to unity even if it is not exactly the real case. Consequently, the obtained site responses at all other stations from inversions are considered relative to the reference site. An additional objective within the PhD is to propose fixing one or several reference sites for which we know, with a level of confidence, the absolute site response, which can be a challenging

step. Hence, we propose looking for stations for which we are sure that the 1DSH site response is close enough to reality and can be considered a reference in GIT inversions.

- Extend the deconvolution approach using GIT site responses and conduct comparisons with the approach of Laurendeau et al. (2018) for methodological aspects.
- Keep a transposable sense of the work. In other words, we avoided all network-specific applications that cannot be transposed elsewhere, especially in the French/European context.

Thus, I present hereafter in this document the elaborated work as illustrated in Figure 0.2, over different chapters as follows:

- **Chapter 1** presents a state-of-the-art of the main elements elaborated in the work. I also present the basic definitions in seismic hazard assessment and the importance of accounting for site responses, current practice, and proposals. GMPE developments are presented as they are the primary tool in SHA. Afterward, methods of site response estimations, which are addressed in the work, are briefly illustrated.
- In **chapter 2**, I discuss a possible source of uncertainty in the prediction of theoretical site responses. Then, I shed the lights on the small-scale heterogeneities that are observed in high-resolution geophysical measurements. To account for this, I present the adapted geostatistical model to account for these uncertainties. In the end, I show the impact of these heterogeneities on the high-frequency response of a subset of KiK-net sites in a representative statistical way. The work performed in this chapter seems important whenever theoretical site responses are addressed.
- Before going ahead with the GIT application on KiK-net, I present in **chapter 3** a methodological benchmark on generalized inversion techniques (GITEC, Generalized Inversion TEchnique Comparisons). The main objectives of this part were (1) to understand the different existing GIT methodologies to solve the inversion problem; (2) to explore the level of inter-methods variabilities that could result between different approaches. Several researchers from different countries were invited to apply their inversion schemes to this small project, where the different GIT schemes were applied on different dataset configurations over two phases. In the text, I show a summary of the main results and comparisons.
- I present in **chapter 4** the application of GIT to the Japanese KiK-net data. In this part, I present the dataset used with the main data selection criteria considered. I clearly illustrate the different steps followed, starting with regionalization of the Japanese network till the site response estimations. Following the recommendations of chapter 3, I include both a parametric and a non-parametric inversion scheme. In the end, I perform an additional interpretation on site responses, focusing on the “absolute” site response aspects and the identification of sites with dominant 1D effects.

- **Chapter 5** includes the application of the deconvolution approach to obtain reference motion GMPE. The work in this chapter uses the site response estimations presented in the preceding chapter. To keep track of results of different works in literature (especially that of Laurendeau et al., 2018) and to achieve our objectives (transposable methodology), we consider a new "1D-subset" of KiK-net sites in addition to a mixed subset that does not include 1D restrictions. The 1D subset served a comparison with the methodology and results of Laurendeau et al. (2018), while the non-restricted dataset allowed us to overview the results in a more generalized case (similar to what we would encounter in future applications where sites with dominant 1D effects cannot be easily identified).
- In the end, discussions and conclusions are illustrated based on the findings of the previous chapters. Also, perspectives are presented for future applications of the deconvolution approach in the French/European context.

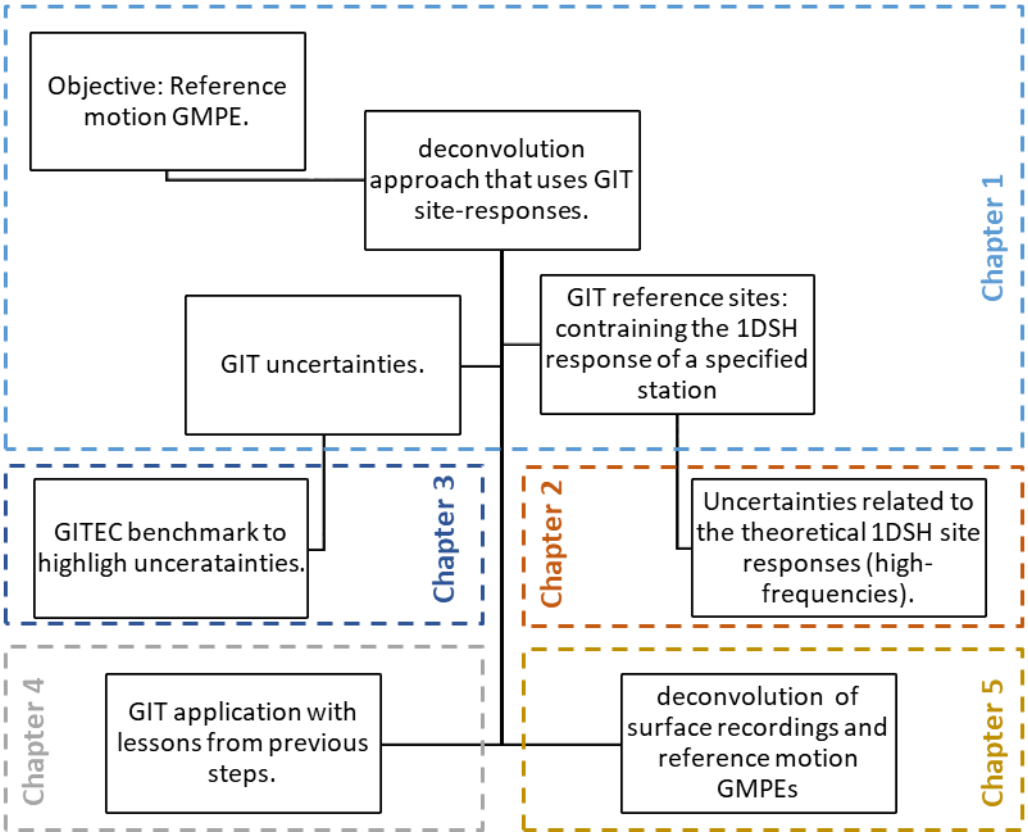

Figure 0.2: Workflow of the thesis over the different chapters.

Chapter 1: Reference motion and site effects in SHA

Precise seismic hazard evaluations are essential for site-specific engineering applications, such as the design of critical facilities. Remarkably, the seismic safety of nuclear facilities requires periodic site-specific hazard reassessments that permit to consider the advances of methods and knowledge in engineering seismology. The latter can lead to significant modifications of expected hazard levels, which imposes verifications of facilities and equipment to cope with the new hazard levels. As a result, if the concerned facilities are sensible to the new hazard levels, they are usually shut down or reinforced.

Until today, an absolute prediction of an earthquake occurrence and ground shaking level is not possible in a purely theoretical way due to several complexities of the earth's structure. Instead, researchers perform regressions on recorded data in the past to predict ground shaking levels of future earthquakes. The latter is only possible under the assumption that earthquake spatial and temporal properties are invariant with time. The equations used in the regressions are given the generic term "ground motion prediction equations" or "GMPEs". They equally provide a statistical estimate of the expected ground motion level and a standard deviation representing the uncertainty for a given earthquake scenario. Thus, GMPEs do not literally predict future earthquakes but instead evaluate average scenario-specific intensities based on previous observations.

In the context of site-specific SHA, site effects are a compulsory part to consider. However, different approaches are proposed to integrate site effects into SHA studies, which can be very simple or detailed. Despite the latter, the common principle is to assess ground motion on reference site conditions with the help of GMPEs and then integrate site-specific parameters to account for local site responses (Renault, 2014; Rodriguez-Marek et al., 2014; Ameri et al., 2017) and illustrated in Figure 1.1. Thus, the identification of reference motion is a critical question in such studies. Since site-responses are defined as the effects undergone by ground motion between the underlying seismic bedrock and the surface level, the reference motion is generally meant to be at the bedrock level, where no site effects are yet acquired. It is common to consider that reference ground motion, for a specific site, exists on a near outcropping bedrock, which is not always easy to find. However, the definition of reference motion has always been a sensitive issue since the early 1990s (Steidl et al., 1996).

In the following sections, the main aspects of site-specific SHA are briefly detailed to illustrate the background of the PhD work. Therefore, we will elaborate on the basics of GMPEs related to our work, namely accounting for site effects in GMPEs. Since reference conditions, the main target of the present work, represent "complete absence of site effects", we propose defining reference ground motion by removing the associated site responses from recordings in a given region. This step is named the "deconvolution approach", and it requires a proper estimation of site effects for all sites in the dataset. For this aim, we will address the state-of-the-practice to define reference motion, and we will consequently detail the most common methods (and mainly those addressed in this text) for precise site effects estimations.

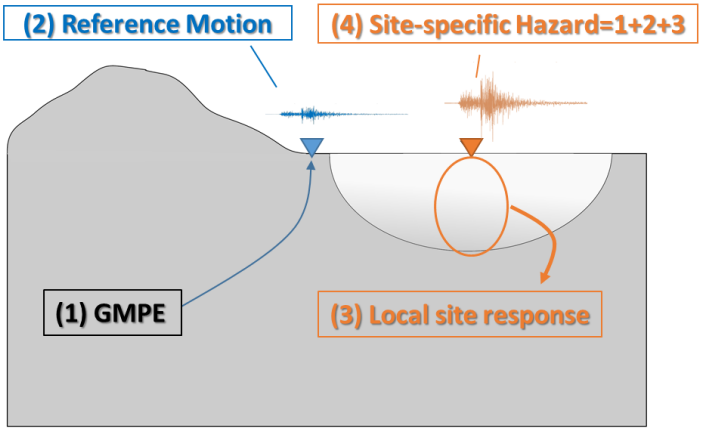

Figure 1.1: The concept of site-specific seismic hazard assessment: (1) GMPEs are used to estimate the reference motion(2). Then, integration of local site responses (3) is necessary to obtain site-specific hazard estimates (4).

1.1 Ground-motion prediction equations (GMPE)

GMPEs are models mainly resulting from regressions on empirical and sometimes simulated data. In either case, the regressions use simplified functional forms to model the recorded data robustly. These models ought to describe the relationship between a desired intensity measure Y that describes the ground motion and a group of variables related to the seismic scenario and stands for the different contributions (source, path, and site), as in Eq(1.1).

$$\ln(Y) = f_{source}(M) + f_{path}(R) + f_{site}(V_{S30}, etc\ \ldots) + \delta \qquad (1.1)$$

Here, f_{source} is a function that mainly scales ground motion prediction with magnitude (M), f_{path} is the scaling with source-to-site distance (R), and f_{site} is the part accounting for site-effects that can be a function of site-specific information as V_{S30} or other parameters. This relation is usually established by using multivariate mixed-effects nonlinear regressions, which results in a residual term δ.

1.1.1 Intensity measures (IMs)

Time histories of earthquakes are not directly used in GMPE developments. Instead, they are treated and used to estimate IMs, which are parameters that best describe the ground motion impact on buildings. In other words, they are often used to quantify ground motion severity. The common IMs are:

- Peak-based, which are measures of absolute maximum values of ground motion's time history characteristics. A large portion of GMPEs is derived from horizontal peak ground accelerations (PGA), while fewer ones are derived for peak ground velocities and displacements (PGV and PGD, respectively). PGA is known for its close relation to dynamic forces, for which structures are exposed. Also, it can be considered close to the response of very stiff structures. PGV and PGD could be more related to lower frequency responses (higher proper periods) and have several uses. For example, in addition to PGA, they were used to constrain the elastic response spectra for design (Newmark and Hall, 1982). Also, PGV/PGA ratios were examined in some studies for their correlation with inelastic structural deformations (Zhu et al., 1988; Sucuoğlu et

al., 1998) or relation to nonlinear soil behavior (Chandra et al., 2016; Guéguen, 2016).

- Duration-based, which results from the integration of functions related to ground motion time histories over the entire duration. Some of these measures are the cumulative absolute velocity (CAV), Arias Intensity (I_A), root-mean-square acceleration, and characteristic intensity (I_C). Since these duration-based IMs are not exploited within this text, further related details are omitted.
- Frequency-response-based, which are computed from the response of single degree of freedom oscillators (SDOF). The structure response is commonly approximated by that of an SDOF, characterized by a specific resonance frequency. Since the natural frequency can vary depending on the structure characteristics, a frequency-dependent acceleration response spectrum is usually derived. The most common is compute pseudo-spectral acceleration response (PSA) for a 5% damped oscillator. Figure 1.2 simplifies the concept of a response spectrum given an input earthquake signal (borrowed from Kramer, 1996). The response spectrum is principally constructed by calculating maximum oscillator response at several resonance frequencies, and they are combined in the form of a complete spectrum.

The vast majority of GMPEs are derived for PGA and PSA, as detailed in the reports of Douglas (2020). This is mainly due to the fact that most earthquake-resistant designs are based on PGA and PSA (Bommer and Alarcon, 2006). In the nuclear industry particularly, the prediction of response spectral accelerations over a specific frequency band grabs more attention than PGA. These facilities are sensitive to high-frequency motions between 10 Hz and 20 Hz, which can be higher than a PGA value. The latter leads to limited use of PGA in such precise hazard analysis.

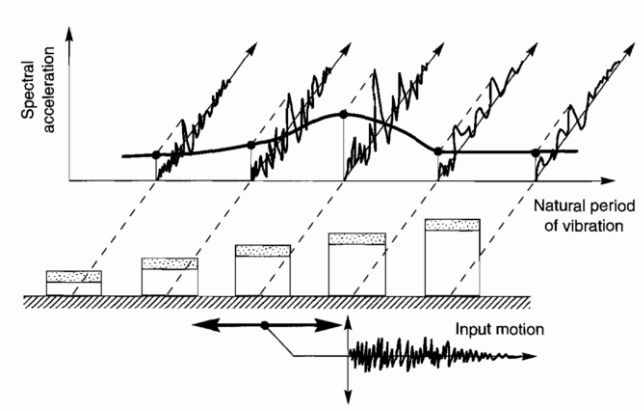

Figure 1.2: The figure describes how the response acceleration spectrum is obtained from recorded signals as input motion for SDOF systems. The maximal acceleration obtained at different natural frequencies (periods) of the SDOF oscillator constitutes the response spectrum (figure adopted from Kramer, 1996).

1.1.2 How GMPEs are derived

A vital step before deriving a GMPE is to construct a ground motion database with enough metadata and well-defined IMs. First, a network of seismic stations should have been installed and recording for several years, covering as many earthquakes as possible. Figure 1.3 illustrates the concept of strong-motion database construction. Each earthquake "i" recorded is usually characterized by a series of source-related

parameters, e.g., epicenter, depth, magnitudes, etc. On the other side, each recording station *"j"* is usually characterized by site-specific parameters such as localization, installation type, and a basic idea about the sub-surface soil structure. Seismic signals recorded at each station are treated and used to construct strong-motion databases used for GMPE regressions

Before Brillinger and Preisler (1985), regressions on recorded data were performed by the fixed effects models as described in Eq(1.2). In these models, δ_{ij} is simply the error of the regression model for earthquake "i" at site "j", which is normally distributed. Also, they are called fixed effects models because their main concern is the regression coefficients and no particular analysis is performed for residuals δ_{ij}.

$$\ln(Y_{ij}) = f_{source}(M_i) + f_{path}(R_{ij}) + f_{site}(V_{S30\,j}, etc\ldots) + \delta_{ij} \quad (1.2)$$

Brillinger and Preisler (1985) first presented an algorithm to include random effects, in addition to fixed effects. Random effects are proposed to account for dependencies in data that we do not consider in the functional form used in regressions. The essence of the method lay in the decomposition of residuals (δ_{ij}) into between-event and within-event residuals to better understand event-related terms. Then, this modeling type was revised by Abrahamson and Youngs (1992) and delivered a more stable algorithm that was afterward used in most of GMPE derivations in the last few decades. Recently, Bates et al. (2015) published an R-package dedicated to nonlinear mixed-effects regressions (nlmer) that greatly simplified the regression task for GMPE developers. The importance of residuals decomposition in GMPEs will be highlighted afterward.

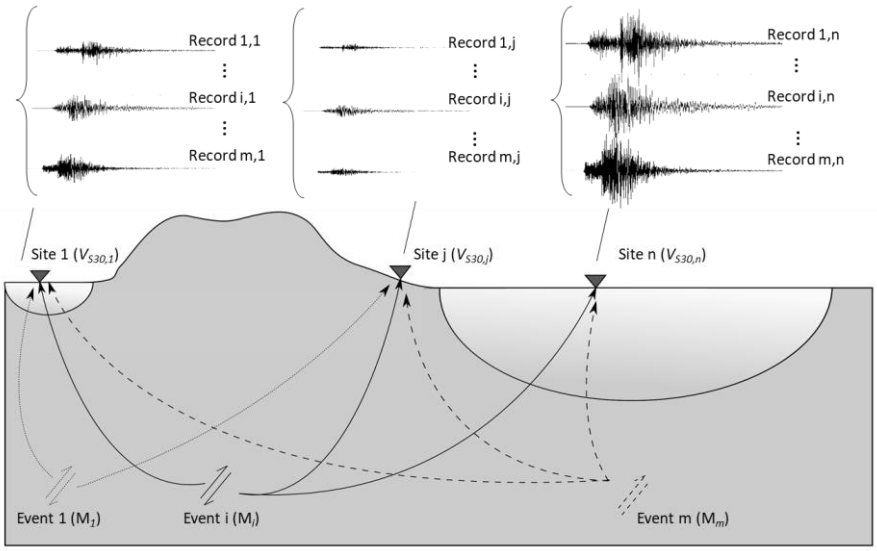

Figure 1.3: An illustration of a strong-motion dataset construction, which is aimed to develop GMPEs, using each event i at each site j.

First, GMPEs started in the late 1960s with elementary functional forms (as reported in Douglas, 2020). The limited amount and quality of recordings available were not enough to observe detailed aspects of ground motion considered in hazard analyses nowadays. After several devastating earthquakes in different parts of the world, seismic hazard analyses received increased attention. Thus, deployment of seismic networks and construction of strong-motion databases massively started in the regions exposed to seismic hazard such as in Japan (KiK-net, K-net, F-net, etc.), United States

(NGA East-1 and-2), and Europe (RESORCE then ESM datasets). The worldwide actions lead over three decades to enlarge the databases and include vast amounts of records that allowed addressing additional complexities in GMPEs.

In a GMPE functional form, dependence on the source (f_{source}) is usually described by the magnitude M. Additional source details and complexities are considered if they are available and well-constrained such as faulting style, hanging-wall effects, depth to rupture surface, and many others. However, in this text, we will avoid more detailing about source aspects and limit the study to magnitude dependence.

Despite the uncertainties related to the definition of M (local magnitudes, moment magnitudes, etc.), a clear dependence of IMs on M is always found and described in different ways in GMPE derivations. When datasets are limited to low-to-moderate seismicity context, a linear and simple form is usually considered (e.g., Berge-Thierry et al., 2003). Then, the linear dependence is extended to a quadratic polynomial in moderate seismicity contexts (e.g., Bindi et al., 2014). Afterward, a hinge magnitude (M_h) started to be introduced when recordings of large events appeared in databases (Rodriguez-Marek et al., 2011; Kotha et al., 2016), beyond which scaling returns to a linear form. Other studies with a wide range of magnitudes preferred to include a piece-wise linear relation with multiple frequency-dependent magnitude hinges (M_{h1} and M_{h2}) as in Boore et al. (2014) and Kotha et al. (2018).

The path terms (f_{path}) are inspired by physical models that are used for crustal wave propagation. First, anelastic attenuation is usually included by an exponential decay with distance r while geometrical spreading is inversely proportional to r. Distance decay in GMPE derivations was found to be magnitude dependent, as Cotton et al. (2008) first suggested, and it was considered in various GMPE derivations afterward. Another aspect of distance scaling is the near-source saturation effects of ground motion explored in several studies (e.g., Atkinson et al., 2016). Near-source effects are usually introduced in f_{path} in the form of effective depths at short distances.

Early functional forms of GMPEs were not prepared to account for site effects in an f_{site} term. Full ergodic assumptions were assumed by default for application purposes. Advances in the last decades showed significant importance to take site-specific ground-motion variations. As each site has its geological structure, ground motion characteristics are expected to vary differently from one site to another. First, site effects were considered similar for sites of the same category (i.e., soil or rock conditions, as in (Berge-Thierry et al., 2003; Zhao et al., 2006, 2016). Alternatively, site effects are also considered through site proxies, and mainly by the time-averaged S-wave velocity in the first 30 meters below the soil surface, i.e., the V_{S30}. However, accounting for site effects using V_{S30} underwent different forms, either directly using the value itself or indirectly using classified sites into V_{S30}-based soil categories.

1.1.3 Comments on site effects in GMPEs

Since we are interested in the site-specific ground motion predictions in this text, it is important to recall some of the main aspects of site conditions considered in GMPEs. Site effects can be accounted for by two main approaches: a generic and site-specific one. Generic methods provide average estimates of site effects and accept lower accuracy than site-specific methods, intending to provide a quick and simple practice for engineering applications. These methods mainly rely on simple site proxies such

as V_{S30}, which are used either directly (e.g., Rodriguez-Marek et al., 2011; Bindi et al., 2017) or indirectly by classifying soil conditions into different categories (e.g., Zhao et al., 2016). However, a site-specific method delivers accurate site responses resulting from site-specific ground response analyses, which require additional costs and efforts.

In the context of SHA for critical facilities, an increased level of attention is generally recommended for site effects. Recommendations for nuclear facilities usually consider categorization or parametrization of site-effects as an exaggerated simplification. Instead, a detailed site-by-site to account for amplification is strongly desired. Some of the reasons standing behind these recommendations are mentioned hereafter:

- V_{S30} is usually obtained by geophysical surveys investigating soil properties profiles up to at least 30 m depths. Since the surveys are usually expensive in time and money, a low portion of well-characterized sites is present in the strong-motion databases. Instead, geological and topographical correlations with V_{S30} are performed (based on existing measured profiles) to complete this missing site information. However, these correlations inject additional uncertainties and bias in the amplification estimations.
- The proportion of measured V_{S30} in sites is relatively lower than inferred sites, as in several databases.
- Many sites can have the same V_{S30} (or same site category, usually based on V_{S30}) but can reveal very different site responses. V_{S30} describes the first 30 meters of soil profiles, while deeper soil layers can play an essential role in site effects. Thus, there is a non-unique site-effect for the same V_{S30}. In another way, V_{S30} parameterizations lead to an average response without considering local resonance effects.
- A detailed site-response assessment may require investigations for lateral and vertical spatial variations of soil properties at every single site. A few hundred meters between two stations can hold significant changes in site responses.
- Since the work of Boore and Joyner (1997), an additional site proxy to V_{S30} started to be used, which is the high-frequency attenuation term, κ. This term was initially introduced by Anderson and Hough, (1984) to describe the unexplained high-frequency attenuation observed on Fourier spectra. However, clear physical bases of such parameterization remained absent until today, with correspondingly high levels of uncertainties in its measurements (Ktenidou et al., 2015).

1.1.4 GMPE residuals

Despite accounting for different contributions in GMPEs, ground motion representation remains very simple compared to the complexities of the physical processes involved in ground motion generation and propagation. Thus, the departure of data from the GMPE average predicted is considered to be in a random manner. Ground-motion residuals are defined as the difference between observations and predictions to capture the unexplained dispersions from this average as in Eq(1.3).

$$\delta = \ln(Y_{obs}) - \ln(Y_{pred}) \quad (1.3)$$

Ground-motion residuals are generally assumed to follow a normal distribution with zero mean and a total standard deviation (σ_T). This allows seismologists to express ground-motion predictions into an explained (fixed-effects) and an unexplained (σ_T) component that represents the related variability. In the context of hazard analyses, it is customary to distinguish between epistemic uncertainty, which results from incomplete knowledge of physical processes, and aleatory uncertainty due to the random nature of the processes. However, it is sometimes challenging to separate epistemic from aleatory uncertainties regarding the models used in regressions, where explainable processes not accounted for in a GMPE can appear as random variations. To simplify the separation, one can consider as epistemic all the variabilities that can be reduced by improved knowledge and data. At the same time, all the rest falls by default into aleatory variability.

Random effects are introduced in regressions to extract more information from residuals, in addition to fixed effects, permitting the identification of ≪repeated residuals≫ for a specific occurrence. In particular, mixed-effects (fixed and random) regressions were initially introduced by Brillinger and Preisler (1985) and revised afterward by Abrahamson and Youngs (1992) to extract event-related residuals treated as random effects. These residuals were called between-event residuals (δB_e) to capture what we do not account for in regressions. Based on this, what remains from residuals after deduction of δB_e is called within-event residuals ($\delta W = \delta - \delta B_e$). Each of the latter variability components is considered to follow zero-mean normal distributions with standard deviations τ for between-event terms and ϕ for within-event terms.

The variability associated with GMPEs (i.e. σ_T) appear to be important since it describes the uncertainty on the median predictions, and they can have a powerful influence on final hazard estimates (especially those of PSHA). In fact, median and standard deviations of GMPEs are obtained using a broad range of earthquakes and sites, and they are applied to analyze the hazard at a single source-site combination. This practice was always applied under the ergodic assumption (Anderson and Brune, 1999). However, the increased amount of well-recorded earthquakes allowed researchers to relax such an assumption, even if partially. It has been shown in several studies (e.g., Lin et al. 2010; Chen and Tsai 2002; Atkinson 2006; Morikawa et al. 2008; Anderson and Uchiyama 2010) that removing the ergodic assumption leads to lower variability of ground-motion. However, removing the ergodic assumption requires the definition of the site- and path-specific ground-motion models, which need enough data to constrain them.

Under a full ergodic assumption, all the variability observed is automatically classified as aleatory variability. Thus, the key to reducing the aleatory σ_T is to identify repeatable effects at single sites and events that can be transferred from aleatory variability into epistemic uncertainty. For this aim, researchers started to break down ground-motion residuals into several components to improve seismic hazard analyses (e.g., Walling 2009; Atik et al., 2010). The first necessary step was to decompose residuals into smaller variability parts and attribute to each one a specific acronym. The within-event residuals of GMPEs include systematic site- or path-specific effects that can be identified. Similarly, between-event residuals contain systematic source-specific effects. Removing these systematic effects from ground-motion residuals is considered

as the key to remove ergodic assumption, reduce the aleatory σ_T and improve hazard estimates. At the end, all repeatable effects that can be identified can thus be transferred into epistemic uncertainties, and all that remains unexplained will be thus considered as a random process.

The level to which δ can be decomposed depends on the assumption and the level of the event- and site-specific information available. Though the present ways to further decompose δB_e and δW_{es} into smaller variability components can vary from one application to another. These decompositions mainly break down the ergodic hypotheses about events, regions, path, and site levels. With the help of mixed-effects regressions, between and within-event components can be decomposed as follows:

- Between-event (δB_e): Earthquakes of similar parameters can produce different ground motions (e.g., Joyner and Boore, 1981) probably because of distinct physical characteristics that could not be accounted for in GMPEs. For a well-recorded event, such unaccounted effects can be viewed as event-specific non-ergodic effects with random effects and highlighted by δBe. If more information is available about earthquake seismogenic zones, GMPE developers might be interested in location-to-location variabilities $\delta L2L \sim N(0, \tau_L)$ of earthquakes. The latter component is a location-specific one that follows that identifies repeatable effects more accurately than δB_e.
- Within-event (δW_{es}): Excluding the source, two main explainable components can be generally identified: site-to-site and path-to-path components. First, site-to-site variability represents the systemic deviation of the observed amplification from the predicted one (the result of site-proxies) and is generally denoted $\delta S2S_s \sim N(0, \phi_{S2S})$. Second, path-to-path variabilities denoted $\delta P2P_{es} \sim N(0, \phi_{P2P})$, represent the average shift of event- and site-corrected ground-motion from the median site-specific model prediction. In other words, $\delta P2P_{es}$ signifies how many specific characteristics of the travel path can lead to systematic deviations from the median predictions. All the remaining unaccounted effects are shifted to a residual term $\delta W_{0,\mathrm{es}} \sim N(0, \phi_0)$ accounting for record-to-record variabilities. In some studies, as Rodriguez-Marek et al. (2011), ϕ_0 is also denoted as ϕ_{ss} (“ss” standing for single station). It is worth mentioning that it is common to merge the path-to-path component into the residual term $\delta W_{0,\mathrm{es}}$ whenever it is not addressed.

Since we are working in the context of site-specific hazard evaluation in this text, we will need to tackle later on the within-event residual components of ground motion. More precisely, we will use the site-to-site residuals due to their close relation to site-specific amplification factors.

1.2 Ground-motion prediction for reference rock conditions

After displaying the primary tool in SHA studies, the GMPEs, we address in this section the first key component in site-specific hazard estimates that is ground-motion prediction at reference conditions. Ground-motion is said to be at reference conditions if it has not yet gained site effects while approaching the surface. At the bedrock level below a site, site effects are generally minimal. We usually consider that sites are at reference conditions if their properties are reasonably close to seismic bedrock

properties, i.e., hard-rock material with V_S exceeding 1.5 km/s. Therefore, researchers have always addressed hard-rock sites to identify recorded reference motion.

GMPEs are derived on datasets with a specific range of site conditions, commonly called “host sites”. As shown in Figure 1.4, existing strong-motion databases are most often representative of stiff-soil to soft-rock site conditions (i.e., V_{S30} 300-800 m/s). It is clear that reference sites are not well constrained (even not at all sometimes) in most GMPEs, and hard-rock sites lie on the periphery, or even outside, of validity domains. Thus, ground-motion prediction at such “target sites” is not straightforward.

The proposed practice to overcome the reference motion issue has been first to derive ground motion on available host sites, then adjust the GMPEs to get a ground motion estimate under reference hard-rock conditions. They often refer to this kind of approach as “Host-to-target-adjustments” (HTTA). Several projects of safety reassessment of critical facilities around the world have considered HTTA. For instance, seismic hazard for Swiss nuclear power plants was first estimated within the PEGASOS and PRP projects and then refined by considering the HTTA approaches (Biro and Renault, 2012). Another application, which used these adjustments, is the “Thyspunt” Nuclear Siting Project (Rodriguez-Marek et al., 2014) and the Hanford site in the U.S. (Coppersmith et al., 2014).

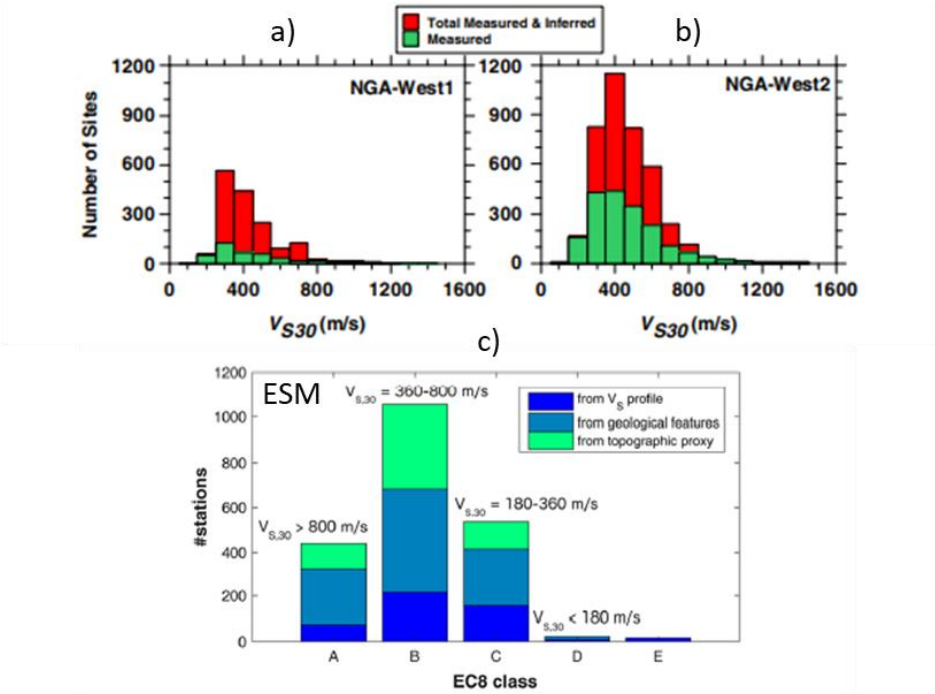

Figure 1.4: Site distributions in V_{S30} for American NGA (a and b) and European ESM datasets (c), indicating measured and inferred values. Figures are adapted from Bozorgnia et al., (2014) and Lanzano et al., (2019).

1.2.1 The essence of Host-to-Target-Adjustments (HTTA)

The principle of such adjustments is to consider possible differences between host and target sites using physics-based models. The latter requires sufficient knowledge of the underlying physics in addition to a well-defined methodology (or strategy) of application. Since the source and path effects are weakly constrained, they are sometimes neglected, and the focus is on adjustments accounting for site-effect differences. The site differences considered in these corrections are the impedance contrasts based on deep V_S-profiles (denoted by V_S-correction) and the high-frequency

attenuation differences (denoted as κ-correction). Figure 1.5 illustrates the principle of correction between host and target sites. After defining the adjustments, they are applied in the Fourier domain with IRVT and RVT conversions from and to the response spectral domain.

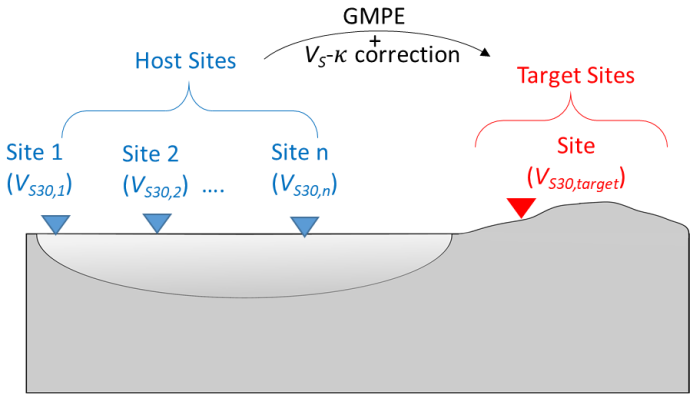

Figure 1.5: the essence of the Host-to-target adjustments. GMPEs derived on host sites are applied for target sites after physics-based V_S-κ corrections.

- V_S-correction is applied in a way to account for differences in impedance contrasts. This is done mainly by modeling site amplification by the QWL approach (initially proposed by Joyner and Boore, 1981). The QWL amplification is computed from impedance ratios (i.e., proportional to $V_{S,\,surface}/V_{S,\,bedrock}$) and results in a simple monotonic function with frequencies. The method uses very deep V_S-profiles that go to several kilometers of depth (initially proposed in Boore and Joyner, 1997, then updated in Boore, 2003). It is based on measured deep V_S-profiles (Figure 1.6). However, these deep V_S-profiles are not always available for target sites or even host sites, making the estimation difficult. Hence, Boore and Joyner (1997) used the V_S-profiles through interpolations, based on V_{S30} only, to obtain a deep generic profile that allows amplification estimation. However, several attempts to improve this practice were made by defining region-specific reference rock profiles (Poggi et al., 2011, 2013). Finally, the ratios between generic-rock and hard-rock amplifications (with the QWL-estimations mentioned above) constitute the V_S-correction.

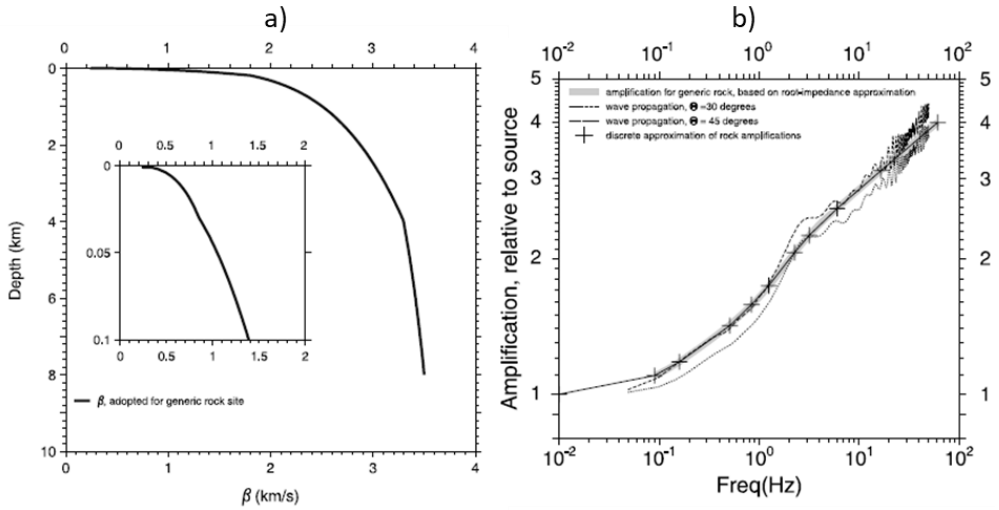

Figure 1.6: a) S-wave velocity profiles commonly adopted for generic rock sites in V_S-correction of HTTA. b) The amplification approximation by the quarter-wavelength method (in grey) based on the generic

V_S-profile. The amplification is compared to other empirical and theoretical estimates. Both figures are adopted from Boore (2003).

- In addition to impedance effects, high-frequency attenuation effects are accounted for with the κ-correction factors. The "κ" term was initially introduced by Anderson and Hough (1984) to describe the physically unexplained high-frequency slope. The model, which is usually used to parametrize these high-frequency effects of the Fourier amplitude spectrum of an earthquake, is $\exp(-\pi\kappa f)$. This attenuation κ model is usually added to QWL amplification to describe both impedance and attenuation effects (Figure 1.7). Based on this, if the attenuation features on target sites are lower than host sites (i.e. $\kappa_{target} < \kappa_{host}$), then κ-correction amplifies the high-frequency content of target sites. This is considered the case for hard-rock sites (usually target sites), which often have κ values much lower than soil to soft-rock sites. As a consequence of attenuation correction, hard-rock motion presents an amplification compared to standard-rock sites.

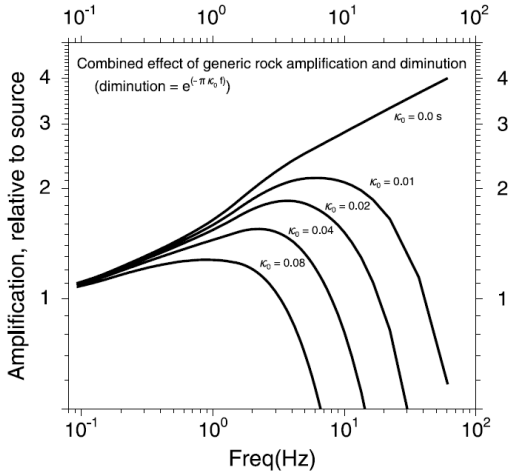

Figure 1.7: Combined effects of site amplification and attenuation effects described in Boore (2003).

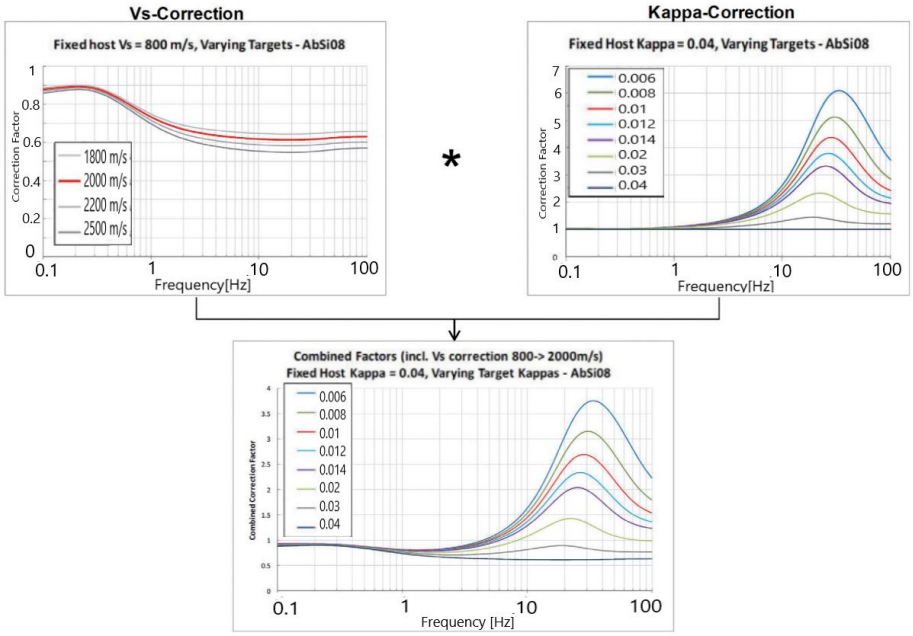

Figure 1.8: Correction factors proposed to go from 800 m/s to 2000 m/s site conditions in Biro and Renault's work (2012). The GMPE used to derive adjustment factors was that of Abrahamson and Silva (2008). (modified from Biro and Renault 2012)

To sum up, host-to-target corrections consider two correction factors related to V_S and κ parameters. An example application is present in Biro and Renault (2012). The general workflow in such analyses starts with V_{S30} and κ estimations on host sites then on target sites. Once these parameters are defined, the correction factors are derived separately (for V_S and κ) and applied, as shown in Figure 1.8.

1.2.2 Limitations of HTTA

Though HTTA methods look like an appealing approach, mainly due to the generalized and straightforward adjustment factors, the approach embeds several approximations and sources of uncertainty worth the discussion. These uncertainties appear to influence the final hazard estimates (Biro and Renault 2012). First, the resulting adjustment factors generally predict an amplification of hard rock over soft rock motion, which seems counterintuitive, as shown in Figure 1.8. Moreover, this method must have V_{S30} and κ estimations, which are accompanied by significant uncertainties. In some cases where measures are not available, the values are assumed subjectively, especially for reference hard-rock sites (i.e., a standard range of κ between 2-10 ms). Eventually, we can group sources of uncertainty in the HTTA adjustments into three main points, detailed afterward.

- The physical origins of κ have been a controversial issue, mainly due to unclearness, possible misinterpretations, and estimation biases. As initially introduced by Anderson and Hough (1984), κ was defined as the slope measured on the high-frequencies of the Fourier amplitude spectrum. Nevertheless, researchers started to address the possible decomposition of contributions to κ into a source, path, and site term (Ktenidou et al., 2015; Perron et al., 2017). The latter is still unclear, and different hypotheses about κ-origins are made each time in the practice.
- Apart, κ was introduced in HTTA as a site-related factor. As mentioned earlier, the defined κ-model led to amplify ground-motion on stiffer hard-rock compared to less stiff generic-rock sites. Although it is a statistical fact that hard-rock motion does exhibit high-frequency contents, careful attention is necessary for on-site installation conditions. A possible explanation for the lack of attenuation on hard-rock sites at high-frequency can be possible due to the local amplification of thin shallow layers, which causes high-frequency resonances. For instance, Hollender et al. (2018) carried the investigations for a set of French stations (RAP, French permanent Accelerometric network) that included several hard-rock sites. Hollender et al. (2018) then performed amplification estimations over the selected stations and showed that shallow weathered rock lying over very hard-rock sites stand behind the high-frequency resonances. Besides, other phenomena can interfere in the high-frequency aspects, such as small-scale local heterogeneities inducing scattering effects (Ktenidou et al., 2015; Cabas et al., 2017; Shible et al., 2018). On the other side, issues such as depth and type of installation could interfere in modifying the frequency content of recorded data, as evoked by Hollender et al. (2020). Above all, κ definition and measurements need carefulness to avoid bias with other physical phenomena.
- The application of HTTA requires reliable estimations of both V_{S30} and κ parameters on host sites as on target sites. Generally, existing databases allow

enough measures neither for V_{S30} nor κ. First, geological/topographic correlations usually help to overcome the lack of V_{S30} measurements in the datasets. These correlations are known today to incur significant uncertainties, as detailed earlier. Second, we should not forget that the V_S-correction initially uses site-specific deep V_S profiles, not simply V_{S30}. Being a parameter that describes only the first 30 meters below the surface, V_{S30} seems an exaggerated tool to select several-kilometer-depth profiles. As for κ estimations, the methods proposed in the context of HTTA are various and generally lead to significant variabilities even if records are present (Ktenidou et al., 2014). These estimations require enough amount of good quality records on both sides. However, κ-estimations are in many cases missing for target sites and even for host sites. To overcome this, several works tried to establish a kind V_{S30}-κ correlations to infer missing values of κ. These correlations were based on available measurements from host sites and proposed to provide κ-values whenever missing (e.g., Chandler et al., 2006; Drouet et al., 2010; Edwards et al., 2011; Houtte et al., 2011; Cabas et al., 2017). Figure 1.9 shows some V_{S30}-κ correlations from different works based on performed measurements. In the end, such correlations show huge uncertainties and large dispersions that doubt their reliability.

- Furthermore, Ktenidou and Abrahamson (2016) extended the analyses on hard-rock sites in the eastern U.S. by comparing classical adjustments with empirical factors. Ktenidou and Abrahamson (2016) considered a dataset that contains few hard-rock sites, which allowed them to analyze empirical hard-rock motion. Their results of comparison between soft-rock and hard-rock motion showed that hard-rock could have higher κ-values (~0.02 s) than those usually assumed (~0.001 s).

To conclude, both V_{S30} and κ parameters and their existing correlations still suffer from uncertainties that hinder obtaining robust estimations of adjustment factors. If we put aside the estimations, there remain several questions surrounding the application of V_{S30}-κ corrections. Being introduced in the Fourier domain, κ was not directly compatible with the usual output of GMPEs (i.e., response spectra). For this aim, conversions to and from the Fourier domain were adopted in practice (Atik et al., 2014; Bora et al., 2015), mainly through the random vibration theory (RVT) and its inverse (IRVT). These back and forth conversions introduce additional uncertainties since they are highly nonlinear and non-unique, especially at high frequencies (Bora et al., 2016).

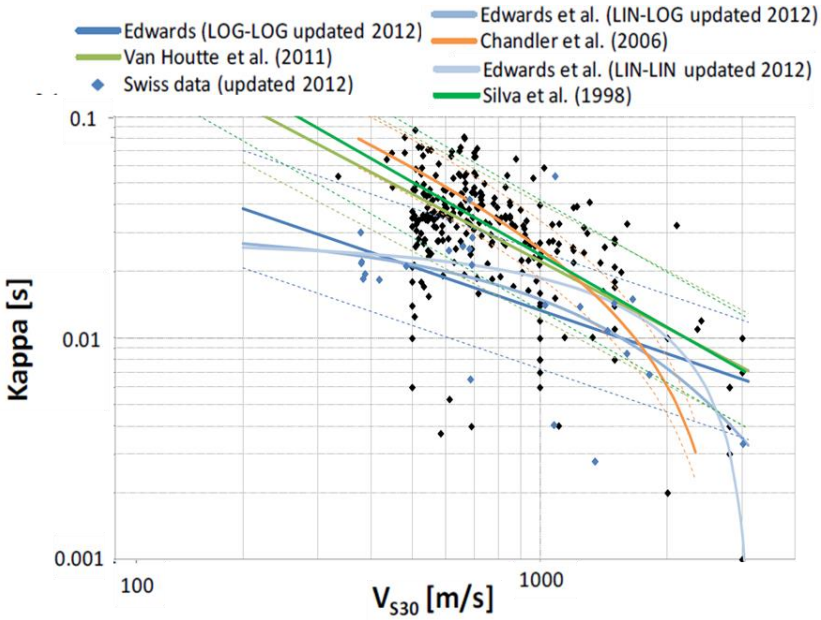

Figure 1.9: Established correlations between V_{S30} and κ from different publications. The Figure is adopted from Biro and Renault (2012) and slightly modified.

1.2.3 Recent alternatives and advances concerning scaling ratios:

This part sheds light on the recent alternatives evoked in recent years to overcome the limitations of the classical HTTA. Viewing the numerous sources of uncertainties in HTTA, alternatives should be investigated to improve and gradually replace the current practice. Some of these alternatives are briefly presented hereafter, highlighting the one addressed in the thesis work. We believe that presenting a global overview of the problem always helps localize this thesis's work better.

1.2.3.1 Fourier domain GMPEs

It was proposed to derive GMPE directly in the Fourier domain to reduce the uncertainty related to back and forth conversions from the response spectral domain (Bora et al., 2015, 2016). Despite changing the domain, Fourier GMPEs were still derived in the same way as for oscillator frequencies. Also, the functional forms for Fourier GMPEs remained very similar to those of "standard" GMPEs. The main advantages can be illustrated as follows.

- Fourier GMPEs lead to partially reduce the sources of uncertainties, but not totally. Fourier GMPEs can be directly adjusted (i.e., with HTTA) in the Fourier domain. However, conversion is still needed towards the response spectral domain (needed in engineering applications). Currently, the latter is only possible with the RVT approach, which requires a duration model (Bora et al., 2015). Sources of uncertainties due to conversions are reduced by avoiding the IRVT conversion, but the RVT and the associated variability were still used.
- The second advantage is that HTTA is directly applicable in the physical Fourier domain. This practice shows more consistency than the standard one since the correction factors are derived in the Fourier domain.

However, the main issues related to adjustment factors, detailed earlier, are not wholly avoided with Fourier GMPEs. Notably, the corrections of impedance effects remain very simplified, based on V_S profiles from the U.S. (i.e., Boore and Joyner, 1997) and

simplified site effects (QWL approximation). Also, the estimation of site-specific parameters is a must, and thus missing measurements lead to the use of the dispersed V_{S30}-κ correlations.

1.2.3.2 Empirical evaluations of scaling ratios

It is always interesting to predict reference motion with empirical methods and utterly independent approaches of the HTTA. First, empirical evaluations help validate, or not, the hypotheses made in the HTTA. The challenging part in empirical estimations is the need for enough good-quality data. Second, the consistency or discrepancies between approaches can be highlighted if entirely independent alternatives are considered. Hence, independent empirical evaluations appear as an interesting step.

An example of empirical correction factors is done in Ktenidou and Abrahamson (2016). They considered two datasets mostly of hard-rock sites of eastern United States. Although most of their sites were characterized by inferred V_{S30} values, they tried to compute scaling ratios between soft-rock and hard-rock sites empirically. Their empirical evaluation of scaling ratios proposed two models (1 and 2) corresponding for the scaling from 760 m/s to 2400 m/s and 2000 m/s. Even though these datasets of hard-rock sites were relatively small, they showed significant differences with the classical HTTA. The empirical scaling ratios showed values much closer to unity, particularly in high-frequencies, than those proposed from the standard adjustments by Houtte et al. (2011), as shown in Figure 1.10.

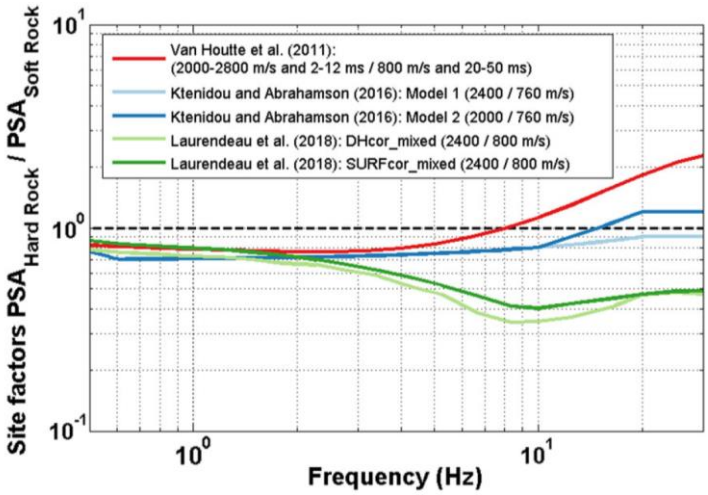

Figure 1.10: standard-rock to hard-rock scaling ratios comparison from different recent publications that proposed alternative approaches to the standard HTTA practice (Bard et al., 2020)

As stated in the work of Ktenidou and Abrahamson (2016), more effort should be made to obtain reliable empirical scaling factors. For instance, much more data is needed from hard-rock sites. As well, there is a need to characterize hard-rock sites better.

1.2.3.3 GMPEs based on corrected recordings

Another alternative proposed to avoid V_{S30}-κ corrections is based on the correction of recordings to account for hard-rock motion. Usually, reference bedrock corresponds to very rigid structures, i.e., V_{S30}>1500 m/s. Such conditions are usually present at the level of boreholes, which is a few hundreds of meters in depth. Cotton et al. (2008) and Rodriguez-Marek et al. (2011) first pointed out the possibility of using borehole records as hard-rock motion. However, borehole data was not much used since it is affected

by downgoing waves that induce destructive resonance frequencies and cannot fully represent outcropping hard-rock motion. Looking to benefit from downhole motion, Cadet et al. (2012) proposed a method to correct downhole recordings from downgoing wave effects. This correction procedure aimed to obtain an acceptable representative of outcropping rock motion.

The procedure of Cadet et al. (2012) was later used in Laurendeau et al. (2018) on Japanese KiK-net data to define reference hard-rock motion from downhole-corrected data, denoted DHcor. This correction allowed to have “virtual” hard-rock sites that can extend the validity range of the GMPE to hard-rock conditions. Then, DHcor was used to derive a GMPE that allowed the evaluation of the soft-rock to hard-rock scaling ratios.

However, the presence of downhole sensors is not usually the case elsewhere in most existing strong-motion databases. Thus, Laurendeau et al. (2018) explored another approach based on the correction of surface recordings. The latter was done by deconvolving (or removing) site-effects estimated by theoretical 1D site responses from surface recodings. Thus, theoretical site-responses were possible because most KiK-net sites are well-characterized with V_S profiles that reach the bedrock. The surface-corrected dataset was called SURFcor and was used directly for GMPE derivation. In the end, the possibility of applying both DHcor and SURFcor on KiK-net was advantageous from a methodological point of view. Both surface and borehole data correction procedures are illustrated in Figure 1.11.

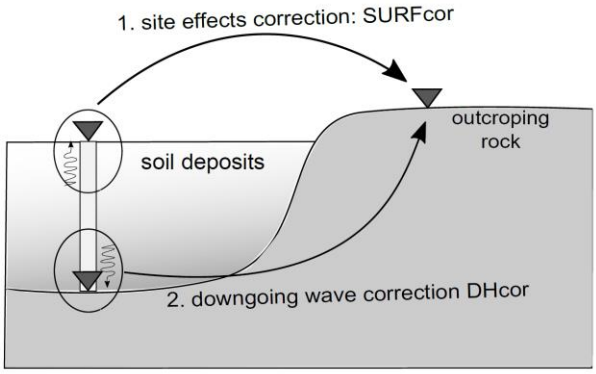

Figure 1.11: Schemes that represent the explored correction approaches of Laurendeau et al. (2018).

As presented in Figure 1.10, Laurendeau et al. (2018) evaluated the scaling ratios between standard-rock and hard-rock motion. Their newly derived ratios were also compared to those of the standard V_S-κ correction of Houtte et al., (2011) and those of Ktenidou and Abrahamson (2016). The main differences that grab attention were at high frequencies with a factor that reaches four (around 10 Hz) between corrected-dataset approaches and the standard one. Also, both models of Ktenidou and Abrahamson (2016) showed discrepancies, with the standard V_S-κ correction, that can reach a factor of 2 beyond 20 Hz. These comparisons highlight the need to revise and improve how the impedance contrasts and attenuation differences are accounted for in the HTTA.

Two main points can summarize the advantages of both data correction approach. First, it results in an extended validity range of GMPEs that permit ground-motion prediction for hard-rock conditions. The standard V_S-κ adjustment was wholly avoided,

providing a completely independent approach with several sources of uncertainties abandoned (related to the V_S-κ adjustments).

Despite the advantages presented, the data correction approach showed several limitations. First, they benefited from the presence of surface and borehole sensors of the Japanese KiK-net arrays. Also, for site-response estimations, V_S-profiles for almost all sites were essential. This fact made the applicability of the approach face several limitations. As the deconvolution approach is the foremost concern addressed in the thesis work, more details are discussed later in this chapter.

1.2.3.4 Site-specific residuals

Kotha et al. (2017) presented an appealing empirical approach to calibrate GMPEs, namely with site-specific residuals ($\delta S2S$). The essence of the method lies in the mixed-effects regressions, where both fixed and random effects can be determined. With the help of such regressions, repeated residuals over a specific site or event could be identified. The latter could be used as a calibrating term for the GMPE considered, allowing the consideration of site effects in an empirical and frequency-dependent way (without any parametrizations).

Thus, the approach of $\delta S2S$ seems interesting since it is empirical, reduces assumptions and approximations, and avoids site parameterizations. The main challenge of such an approach is having enough recordings and exploiting a wide range of possible magnitudes and distances. Even though the non-parametric form of such site-specific residuals could be more accurate, its representativeness of site-effects is limited to the used dataset and the corresponding magnitude and distance ranges.

Another essential point to be mentioned about site-specific residuals ($\delta S2S$) in GMPEs is that they can be considered relative site terms, only applicable in the derived GMPE. In the way it results from GMPEs, the average of $\delta S2S$ at each frequency is null. This fact could be interpreted as referring all site terms to a virtual reference condition, as stated in Kotha et al. (2017), represented by the median site-effect in the dataset. This fact also indicates that these terms, as they exactly result, are relative and can be used only with the same derived GMPE. Finally, it is possible to use these $\delta S2S$ terms to obtain amplification terms which would be of essential benefit in the context of a site-specific SHA. To this aim, a proper reference station(s) should be chosen as proposed in Kotha et al. (2018), for instance, and presented in Figure 1.12.

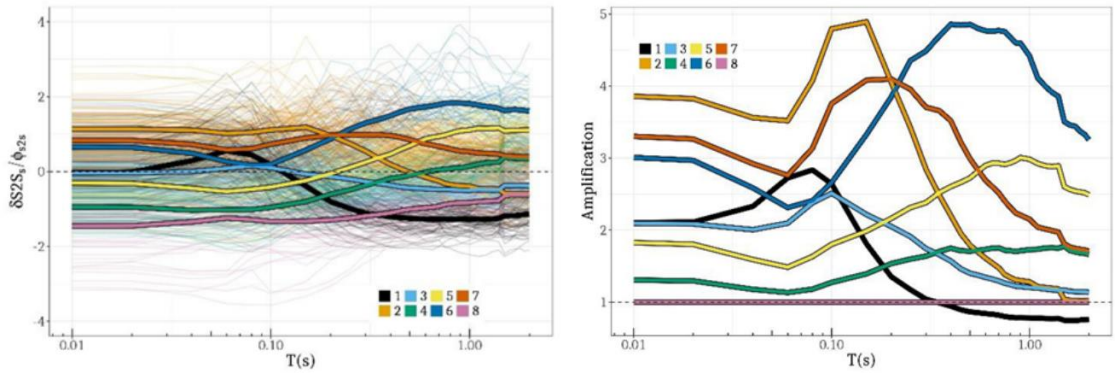

Figure 1.12: a) Site-to-site residuals ($\delta S2S$) obtained by Kotha et al., (2018) on KiK-net data. b) Normalized site-to-site residuals by the chosen reference condition (group 8) to represent amplification terms.

1.3 Site response estimations

Although seismic waves travel tens of kilometers from seismic faults towards the ground surface, ground motion undergoes significant modifications near the surface in the last hundreds of meters, i.e., the site-effects. Increased attention is devoted to understanding the effects of the near-surface geology on ground motion. Near-surface geology is highly variable where many phenomena interfere, and a large number of techniques have been developed for their analysis. For instance, Kawase (2003), Faccioli and Pessina (2003), and Singh et al. (2003) discussed in detail this topic. Accurate hazard estimates require a sufficient understanding of these effects. Thus, this section discusses the main physical aspects and the precise methods of estimations addressed in our work. For additional details, one can refer to (Kramer, 1996; Anderson, 2007).

1.3.1 Physical background

The limits of site effects, or the near-surface term, are generally meant to include all acquired effects by seismic waves above the “seismic bedrock”. It is known that seismic compressional and shear waves increase with depth. The bedrock generally has a shear wave velocity at 3.0 km/s or higher (e.g., Takemura et al., 1995). Also, the major release of seismic energy occurs just below the bedrock level.

Generally, the S-wave velocity increases with depth due to the increased pressure. In rock materials, microcracks and joints are closed gradually with increasing pressure along with depth. Since rock velocity increases with pressure increase (Christensen,1989; Jaeger and Cook, 1976), the influence of the surface can thus reach a few kilometers depth. Thus, surface rock outcrops may show site effects and cannot be suitable for reference motion definition (e.g., Castro et al., 1990; Humphrey and Anderson, 1992; Steidl et al., 1996). For soil deposits, the compression due to gravitational loads will cause both density (ρ) and elastic moduli (μ) to increase with depth. The shear wave velocity, $\beta = \sqrt{\mu/\rho}$, also increases with depth since μ increases more rapidly than ρ.

Inversely speaking, the propagating waves pass through the material with low β as they approach the surface. The decrease of the velocity is accompanied by the amplification of seismic waves by the principle of energy conservation. Aki and Richards (1980) defined the amplification gain for S waves acquired by waves propagating between depth Z (with ρ_Z and β_Z) and the surface (with $\rho_0$and β_0) as $G_A = \sqrt{\rho_Z\,\beta_Z/\rho_0\,\beta_0}$.

The direction of propagation of seismic waves changes as wave velocity changes from one layer to another. Snell’s law for refraction is the fundamental expression that determines how the seismic ray will refract. A seismic ray that passes from layer 1 of velocity β_1 to layer 2 of velocity β_2 with an angle of incidence j_1 will refract with an angle j_2, which can be determined from the general statement of snell’s law in Eq(1.4).

$$\frac{\sin(j_1)}{\beta_1} = \frac{\sin(j_2)}{\beta_2} \tag{1.4}$$

Let us consider a ray path with a large significant incident angle that passes through a horizontal stack of layers. The gradual decrease of layer velocities mimics the real case

near the earth's surface. By the time the seismic ray reaches the ground surface, multiple refractions would have bent the ray nearly towards the vertical direction. This explains why it is common in engineering practice to consider a vertical incidence of ground motion.

The physics at an interface between two layers is much more complicated (Chapman, 2002). For a vertically propagating plane SH-wave, there is a downgoing reflected ray besides the upgoing refracted ray, as shown in Figure 1.13. The transmitted waves can thus be described by a transmission coefficient which is the ratio between the incident and transmitted amplitudes, which is the ratio in Eq(1.5). Also, the reflection coefficient is defined as the ratio between the incident and reflected amplitudes in Eq(1.6).

$$T_{21} = \frac{A_T}{A_0} = \frac{2\rho_2\beta_2}{\rho_2\beta_2 + \rho_1\beta_1} \tag{1.5}$$

$$R_{22} = \frac{A_T}{A_0} = \frac{\rho_2\beta_2 - \rho_1\beta_1}{\rho_2\beta_2 + \rho_1\beta_1} \tag{1.6}$$

On the other hand, when the waves pass from a softer to a more rigid material, $\beta_1 \gg \beta_2$, the transmission coefficient approaches zero, and the reflection coefficient approaches unity. This means that the energy is almost totally reflected in the low-velocity material. When the waves reach the surface and get reflected, they are trapped near the surface where the low-velocity material. The phenomena of trapping the energy near the surface induce the surface waves (Rayleigh and Love).

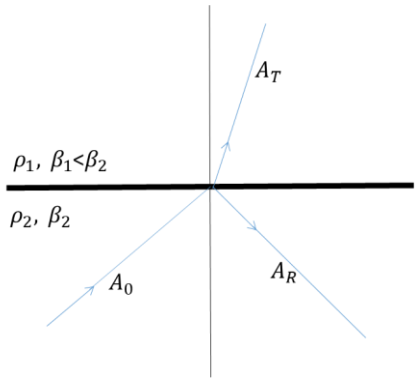

Figure 1.13:a ray that passes from higher to lower impedance layer.

As shown in Eq(1.5), the upper limit of the amplification factor at a single interface is 2. After multiple interfaces, the amplification factor can significantly exceed 2 and lead to significant damage to buildings at the surface (e.g., Hough et al., 1990). Thus, energy trapped in the low-velocity zone near the surface can induce resonances at specific frequencies. If we assume a vertical direction of SH-wave propagation in a homogenous half-space, the resonance occurs when the wavelength λ is four times the layer thickness, that is, $h_0 = \lambda/4$, where h_0 is the total layer thickness. Noting that $\beta = f\lambda$ for a monochromatic shear wave of velocity β, the resonance can occur at a sequence of frequencies f_k given by Eq(1.7).

$$f_k = \frac{(2k-1)\beta}{4h_0}, with\ k = 1,2,3 \ldots \quad (1.7)$$

Another aspect of site effects is the sub-surface topographic effects. Sub-surface topography effects are mainly present in the case of sedimentary basins. In addition to the amplification of vertically propagating waves in the sedimentary basin, seismic rays can hit oblique interfaces. Consequently, waves can follow multiple paths and converge towards a specific point or diverge from it. This leads to an increase in the ground motion duration and amplitude of local vibrations (Bard and Bouchon, 1980a, 1980b; Kawase and Aki, 1989). The most well-known example of basin response causing severe ground motions is the Mexico City basin, which experienced in 1985 an earthquake of M_w 8.1. The 1985 earthquake lasted around 300 s, severely stressed and damaged high-rise buildings, and led to a death toll of 10 000 people.

On the other hand, surface topography can also affect seismic waves, such as mountains and hilltops. It is expected that topographic effects are visible when the wavelengths are comparable to the geologic feature (Sanchez-Sesma and Campillo, 1991). However, the effect of topography on strong motion is not yet well understood due to the difficulty in numerical modeling of the effects and the necessity to have complex subsurface geologic structure (Bard and Tucker, 1985, Geli et al., 1988, Assimaki and Gazetas, 2004).

In addition, there are several aspects of near-surface effects that are not mentioned in this section. In addition to sedimentary soil deposits, sedimentary basins, and surface topographic effects, unconsolidated soil deposits, for instance, can present a non-linear behavior when subjected to strong vibrations. Perhaps, exceeding the elasticity limits of the hook's law will lead to permanent deformations. Eventually, these aspects will not be taken into account in the current work.

1.3.2 One-dimensional theoretical estimates

Since the 1920s, seismologists and earthquake engineers have worked a lot towards developing quantitative methods to predict the influence of local soil conditions. The existing numerical approaches differ predominantly in the level of complexities they address. These techniques can be grouped according to the problem's dimensionality, i.e., one-, two-, and three-dimensional analyses. Although multi-dimensional problems appear to be a generalization of the 1D case, they are much more complicated because of the required level of information and increased computational costs. Thus, 1D analyses of site responses were have been more used for site-specific studies because of their simplicity and closeness to observed site responses in many cases.

Basic one-dimensional computations assume simple geometries, typically infinitely-horizontal stratified layers with linear soil behavior. They also assume that local site response is predominantly caused by the SH-waves (horizontally polarized shear waves) vertically propagating from the seismic bedrock. If we assume the case of a uniform damped soil of thickness H, damping ξ, and velocity β and lying on visco-elastic bedrock, the derived complex transfer function is represented in Eq(1.8) by the ratio of incident to reflected amplitudes u and u', as illustrated in Kramer (1996). The computation of site response is based on the determination of the transfer function, relying on the principle of superposition at the free surface. However, realistic

conditions generally differ from mathematical idealizations of a perfectly homogeneous, flat-layered media for which the transfer function becomes much more complicated to be estimated.

$$TF(f) = \frac{u}{u'} = \frac{1}{\cos[\ 2\pi f H/\beta(1 + i\xi)\ ]} \tag{1.8}$$

The response of multiple layers to an incoming plane wave was first solved by Thomson (1950) through reflectivity modeling methods. Then, Haskell (1953, 1960, 1962) modified the method to simulate surface wave propagation. These methods represent wave propagation in the frequency-wavenumber domain, and they mainly deal with propagator matrix computations (Kennett, 1974; Kind, 1976; Kennett, 1983; Müller, 1985). The solution in a linear medium is a common problem (e.g., Aki and Richards, 1980, 2002) that computers can quickly evaluate. Linear 1D reflectivity models for incident SH plane waves have been used in several investigations (Bard and Gariel 1986; Theodulidis and Bard 1995; Theodulidis et al. 1996, Cadet et al., 2012, Laurendeau et al., 2018).

Computer software allows considering several known complexities at once, such as transmitted and reflected energies, oblique incidence, multiple paths, and local resonances. For instance, the computer program SHAKE (Schnabel et al., 1972) is also one of many widely used in engineering practice that extends the solution to non-linear soil response of vertically incident S-waves by an equivalent-linear approach. Other methods and programs exist, but they will not be further detailed.

Recently, several researchers evoked issues related to sources of uncertainties in the predicted 1D response. If the methods estimating site responses are trusted, sources of uncertainties can come from input information, namely the input velocity profiles. For instance, Parolai et al. (2015) explored the effects of introducing unmeasured vertical heterogeneities on the site response. Also, several movements have improved 1D site response predictions by integrating statistical approaches to account for measurement uncertainties with empirical constraints (Griffiths et al., 2016, Teague et al., 2018; Passeri et al., 2020).

In conclusion, one-dimensional ground response analysis has been considered adequate to represent the wave propagation conditions (Kramer1996; Stewart et al. 2008) in cases where topographic, and basin effects are minor. The required information for 1D methods is thus a velocity profile for P and S waves, density, and damping properties (or quality factors) up to the bedrock level. In this work, we will use mainly the site responsible for vertically propagating SH-waves in a 1D context.

1.3.3 Empirical estimates

Empirical evaluation of site effects mainly uses instrumental recordings of earthquakes. As defined in Field and Jacob (1995), the main challenge in these methods is to separate site response from source and path effects. Several methods have been proposed to achieve this goal. The essence of all empirical estimates lies in the definition of reference motion recorded on a site.

1.3.3.1 Standard Spectral ratios (SSR)

The spectral ratio technique was first introduced by Borcherdt (1970) and is still widely used. Figure 1.14 illustrates the principle of this method, which consists of comparing the records on nearby sites (separated by an epicentral distance Δd) through spectral ratios. Mainly, the site response at a given station, which is commonly described by an amplification function, is estimated by the spectral ratio of smoothed Fourier amplitude spectrum at the station U_s to that at a nearby outcropping bedrock U_r. Each Fourier spectrum is considered to hold contributions from source $G(f)$, path $A(f)$, and local site conditions $S(f)$. Thus, ensuring a reference choice in the nearby will result in ratios independent of source and path effects. SSR for seismic noise also exists (Irikura and Kawanaka, 1980), but no related details will be delivered in this text.

Eq(1.9) represents the typical definition of an SSR ratio at a site s with respect to a reference r. $U_{i,s}$ and $U_{i,r}$ stand for the Fourier amplitude spectra of the signal recorded on the site and the reference, respectively, for a specific event i. Usually before division, the spectra are smoothed with the Konno and Ohmachi (1998) procedure. For a statistically representative amplification, SSR is computed as the median (i.e., geometrical mean) for all ratios results from events i going from 1 to n (n events)

$$\mathrm{SSR_{s/r}}(f) = \prod_{i=1}^{n} \sqrt[n]{\frac{U_{i,s}}{U_{i,r}}} \tag{1.9}$$

For a reliable estimate of the amplification by an SSR ratio, three conditions must be fulfilled. First, the outcropping bedrock should be practically near the site under study to ensure that differences are only due to site conditions, where path differences are negligible, and source differences do not exist because it is the same earthquake for both records (Yu and Haines 2003; Drouet 2006; Cadet et al. 2012). A general limit of distance aperture between sites is considered to be one-fifth of the epicentral distance. Second, the reference site should not be affected by any site effects. This is guaranteed if the chosen reference is located on an unweathered horizontal bedrock. These two conditions appear to be essential for this technique (Steidl et al. 1996). A sufficient amount of recordings to a representative statistical number of recordings with a good signal-to-noise ratio (>3) on both sites.

$$SSR_{s/r} = \frac{U_s}{U_r} = \frac{[G(f).A(f).S(f)]_s}{[G(f).A(f).S(f)]_r} = \frac{[S(f)]_s}{[S(f)]_r} \; if \; \Delta d \ll R_{epi}$$

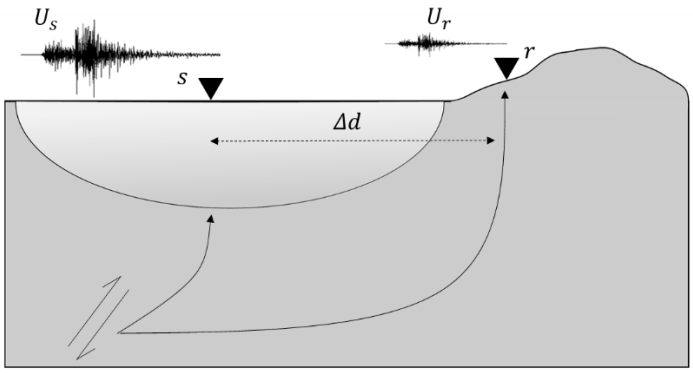

Figure 1.14: Illustration of the principle of standard spectral ratios techniques.

Many studies have shown the reliability of such methods in evaluating site responses (e.g., King and Tucker 1984; Aki 1993; Field 1996; Faccioli and Vanini 2003; Bindi et al. 2009; Thompson et al. 2012). Besides, SSR is regularly considered as a reference method to validate empirical results from other approaches (Rovelli et al. 1991; Malagnini et al. 1993; Field and Jacob 1995; Horike et al., 2001; Sato et al., 2012; Cultrera et al. 2014) or to validate results from numerical modeling (Jongmans and Campillo 1993; Cornou and Bard 2003; Triantafyllidis et al. 2006; Maufroy et al. 2015; Maufroy et al. 2017).

In some cases, the reference site for SSR techniques is the downhole array (Figure 1.14). Despite the downgoing wave effects, many studies have addressed surface to borehole spectral ratio as an empirical estimation of the site effects between the surface and the borehole. For the case of KiK-net arrays, SSR is usually performed with the downhole sites as a reference, named as SSR_{dh} (Cadet et al., 2012; Thompson et al., 2012; Laurendeau et al., 2018; Pilz and Cotton, 2019), we denote here as SSR_{dh}. This step facilitated the task to have site-by-site empirical amplification for all KiK-net sites to control the estimates from other numerical or empirical approaches, though it might be biased with downgoing waves.). Despite the downgoing wave effects, many studies have addressed surface to borehole spectral ratio as an empirical estimation of the site effects between the surface and the borehole. For the case of KiK-net arrays, SSR is usually performed with the downhole sites as a reference, named as SSR_{dh} (Cadet et al., 2012; Thompson et al., 2012; Laurendeau et al., 2018; Pilz and Cotton, 2019), we denote here as SSR_{dh}. This step facilitated the task to have site-by-site empirical amplification for all KiK-net sites to control the estimates from other numerical or empirical approaches, though it might be biased with downgoing waves.

1.3.3.2 Horizontal to vertical spectral ratios (HVSR)

After introducing the SSR techniques, Nakamura (1989) proposed to compute the spectral ratio between the horizontal and the vertical components of a recording as in Eq(1.10). $U_{i,H}$ and $U_{i,V}$ represent the fourier amplitude spectrum of the horizontal (combining two orthogonal directions) and the vertical components, respectively. Similar to SSR techniques, HVSR results from the geometrical mean of ratios at each recorded event i.

$$\mathrm{HVSR}(f) = \prod_{i=1}^{n} \sqrt[n]{\frac{U_{i,H}}{U_{i,V}}} \tag{1.10}$$

Nakamura's (1989, 1996) proposal was initially to use ambient noise recordings. This method is attractive because of its simplicity and economy. Then, the method was applied to the S wave portion of earthquake recordings by Lermo and Chávez-García (1993) in Mexico. Afterward, HVSR gained attention since it proposed simplifying the reference choice task using the vertical component instead of the reference horizontal components (see Chavez-Garcia et al., 1996, 1997; Lachet et al., 1996; Riepl et al., 1998; Theodulidis and Bard, 1994; Theodulidis et al., 1996; Bonilla et al., 1997; Yamazaki and Ansary, 1997; Zaré et al., 1999; Haghshenas et al., 2008). Mainly, the HVSR shapes exhibited excellent experimental stability and highlighted the fundamental resonance frequency. HVSR has become a standard tool to assess for

the site fundamental resonance frequency, f_0 (Bard, 2008; Haghshenas et al., 2008; Griffiths et al., 2016a). Also, it was shown to be well correlated with surface geology and much less sensitive to source and path effects.

After being initially proposed for amplification estimation by Nakamura (1989), several studies disagreed later and related it to the fact that vertical components might exhibit some amplification effects (Bindi et al., 2009; Ameri et al., 2011). While clear peaks in the HVSR often indicate a strong impedance contrast in the structure beneath, interpretations of flat HVSR ratios are complicated and misleading.

1.3.3.3 Generalized inversion techniques (GIT)

The spectral ratios method introduced by Borcherdt (1970) was afterward generalized by Andrews (1986). The latter proposed to look for source and site effects simultaneously in large datasets of recordings through an inverse problem. In a dataset of n_e events and n_s sites, generalized inversion schemes are based on the principle that the S-wave Fourier amplitude spectrum can be separated into three main components as Eq(1.11).

$$FAS_{ij}(f)=E_j(f).A_{ij}(r,f).S_i(f) \tag{1.11}$$

FAS_{ij} is the Fourier amplitude spectrum recorded at site i (from 1 to n_s) for event j (from 1 to n_e). $E_j(f)$ is the source function, $A_{ij}(r,f)$ is the path contribution for event j recorded at site i, and $S_i(f)$ is the site response term. Each of these terms is assumed independent from the other for a given recorded spectrum. Applying the logarithm on Eq(1.11), we get Eq(1.12), which constitutes a linear system of equations of the form $Ax = b$, mainly solved in the least-squares sense at each frequency (see appendix for more details).

$$\log_{10}(FAS_{ij}) = \log_{10}(E_j) + \log_{10}(A_{ij}) + \log_{10}(S_i) \tag{1.12}$$

Initially, Andrews (1986) aimed to estimate source parameters, especially the seismic moments (M_0) and the corner frequency (f_c), after removing path and site contributions. In his definition, the main task was to separate source and site effects, while attenuation was simply corrected by a geometrical spreading factor ($1/r$). In this case, the linear system presented a single unresolved degree of freedom, which is the non-uniqueness of the solution due to two unknowns (source and site). The possible solution is to identify reference site(s) and constrain its site response to unity (amplification-free) to resolve this issue.

Castro et al. (1990) introduced a more generalized scheme that reduces the assumptions on attenuation terms. They principally accounted for a frequency- and distance-dependent attenuation ($A(r,f)$) that includes both anelastic attenuation and geometrical spreading. In this case, the inversion scheme becomes fully non-parametric. Consequently, the unresolved degrees of freedom are doubled. The latter required two steps that are (1) constraining reference site(s) to unity and (2) constraining the attenuation function to start from 1 at a given reference distance ($A_{ij}(R_{ref}, f) = 1$).

On the other hand, several researchers applied the inversion scheme differently, for which they assumed *a priori* models for the source and attenuation terms (Boatwright et al., 1991, Drouet et al., 2008, 2010; Edwards et al., 2008, 2013). The system of equations becomes highly non-linear since several parameters are included (seismic moment, corner frequency, quality factors, etc.). Thus, a multi-step inversion is performed to estimate the parameters successively (Edwards et al., 2008), or the system is solved in a linearized way, such as the gauss-newton method for linearized systems used in Drouet et al. (2008). The parametric inversion scheme is generally more complicated than in the general inversion approach. Thus, every author has his own implementation.

Although the parameterization of source and path effects removes the need for a reference site, there still exists an unconstrained frequency-independent degree of freedom in this procedure. This unresolved degree is present in the trade-off between the low-frequency spectrum level (seismic moment) and the mean level of site response factors at all sites. In principle, this scaling factor can be constrained by *a priori* information, for example, seismic moment estimations or the site response estimated from the geological structure at the site. However, the unavoidable uncertainties on these parameters may increase the value of the scaling factor.

1.4 Scope of the work

The current work focuses on improving ground motion prediction on reference rock. Thus, the main concern will be reference motion and site effects without detailed source and path effect analyses. Among the approaches investigated by Laurendeau et al. (2018) on the KiK-net dataset, the deconvolution approach seems an interesting step towards more precise reference ground-motion prediction.

However, we think that the deconvolution approach needed additional work to ensure its applicability and transportability to other datasets of interest, more precisely as European or French datasets in the low-to-moderate seismicity context. The main limitations of the deconvolution performed by Laurendeau et al. (2018) are:

- V_S-profiles for KiK-net data is an exceptional characteristic not present elsewhere. Most of the existing databases do not have many measured V_S profiles. In fact, the most assessed site parameter in databases is the V_{S30} proxy only, which is also not always measured (i.e., sometimes inferred from geological correlations). It is known that V_{S30} is not sufficient for the estimation of a full frequency-dependent amplification. Thus, it is challenging to estimate theoretical 1D-site responses for all stations in other datasets.
- The presence of downhole sites permitted first to identify sites with dominant 1D effects. The 1D site identification was possible with the procedure introduced by Thompson et al. (2012) used further in several works (Laurendeau et al., 2018; Pilz and Cotton, 2019). The procedure is mainly based on the correlation of surface-to-borehole transfer functions between theoretical and empirical approaches. In the absence of borehole stations, the surface-to-borehole transfer function estimations are unattainable.
- The validity of the 1D assumption on site response is not generally guaranteed on all dataset sites. For example, Thompson et al. (2012) considered a subset

of 100 KiK-net stations in an attempt to identify sites with dominant 1D effects. The latter showed that only 16 percent of stations are ideal for 1D modeling. Laurendeau et al. (2018) further applied a similar procedure to Thompson et al. (2012) and identified around 97 “1D sites” out of 138 (around two-thirds). On the same track, Pilz and Cotton (2019) identified only around 55 percent of 1D sites (of a total of 354 sites) while Zhu et al. (2020) proposed modifications on the 1D effect identification procedure based on absolute site term comparisons, which lead to highlight only 76% of 1D sites out of 90 sites. Despite the different ratios, it seems that statistics conducted in different works lead to conclude that a significant percentage of sites are possibly not in 1D conditions and that the 1D assumption cannot be generalized as representative of the existing real site response in a precise site-specific study. Nevertheless, theoretical estimates can help to understand some specific aspects of site responses at a considered station.

- While conducting a statistical comparison between mean theoretical, empirical, and approximative QWL site responses on the KiK-net subset, Laurendeau et al. (2018) found discrepancies at high frequencies. The main issue was that the theoretical 1DSH estimate (BTF) over-predicted empirical site responses (SSR_dh), on average, as shown in Figure 1.15. The latter mainly highlights a lack of attenuation in the theoretical responses to be furthermore investigated.

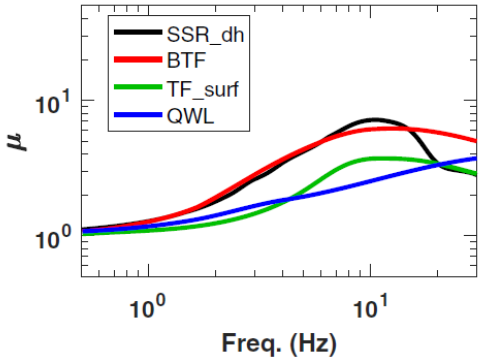

Figure 1.15: Comparison of the mean spectral amplification from different methods. (Laurendeau et al., 2018)

The main element that can allow overcoming these challenges is a robust estimation of site effects with less required site-specific information and reduced assumptions as much as possible. After illustrating the main approaches for site effect estimations in the previous sections, we believe that GIT can provide promising results of site effects estimations necessary to achieve the study's goals. GIT delivers empirical site terms directly from data by separations of source, attenuation, and site contributions. For the latter separation to be reliable, an adequate reference site needs to be fixed. In view of addressing reference sites in GIT, we propose to use 1DSH site responses. Once GIT site responses are estimated, their removal from surface recordings is feasible in the Fourier domain. Finally, GMPE derivations can take place on corrected surface recordings.

In this text, we continue working on the KiK-net dataset to keep track of previous results on the same network. Also, the work presented in the following chapters constitutes articles proposals.

Chapter 2: Vertical heterogeneities in Velocity profiles to account for uncertainties: application to Fourier site-responses in KiK-net arrays.

Context

The work in this chapter is motivated by two points. First, Laurendeau et al. (2018) highlighted a discrepancy between empirical and 1DSH transfer functions at high-frequency. Second, we aim to apply generalized inversions techniques to estimate site responses empirically. These techniques exhibit a degree of freedom generally resolved by imposing a constraint on one or several site responses, commonly known as the reference condition. Since we aim to use the 1DSH response to constrain the reference site(s), they should be carefully estimated. In a separate conference paper (Shible et al., 2018), we showed that small-scale heterogeneities could contribute to uncertainty in high-frequencies, commonly neglected. After updating the dataset used in both works, we consider addressing the related branch of the uncertainty. We also add new discussion elements in the chapter, which constitutes an article proposal.

Abstract

A vital step in seismic hazard analyses is to estimate site responses. For this aim, one of the most used techniques is one-dimensional theoretical modeling, namely the site response to a vertically propagating SH-wave in a horizontal stack of layers. These one-dimensional analyses need as input a minimum knowledge about the site structure, i.e., S-wave velocity (V_S) profiles from geophysical measurements.

If we assume that the considered site has almost a 1D structure, then sources of uncertainties in the 1DSH site response will mainly originate from the input V_S profiles. Several works recently started to address sources of uncertainties in one-dimensional site response predictions. We shed the lights in this work on the aleatory aspect of these uncertainties, which are mainly the unaccounted vertical heterogeneities in low-resolution V_S profiles. After performing a statistical validation from a set of measured profiles, we apply a geostatistical model to account for small-scale heterogeneities in the 1D site response of a subset of KiK-net sites. In the end, we show that small-scale heterogeneities can induce additional attenuation in the high-frequency site response.

Keywords: *Site responses, 1DSH, V_S profiles, heterogeneities.*

2.1 Introduction

Site-specific seismic hazard assessments have become an essential task in earthquake engineering that requires a careful prediction of ground responses to earthquakes. One of the tools mainly used to estimate site-response is the one-dimensional analysis that assumes infinitely horizontally stratified soil layers. These 1D analyses are feasible if we have geophysical measurements that deliver a minimum knowledge about the S-wave velocity (V_S) structure below the considered site. However, 1D analyses can lead to inexact estimations since they rely on assumptions much simpler than reality. Thus, it is generally preferred to validate these 1D analyses by empirical observations. Empirical estimations of site responses, especially the standard spectral ratio (SSR, Borcherdt, 1970), usually require the definition of a good reference site that represents an outcropping bedrock nearby the studied site (Steidl et al., 1996; Cadet et al., 2010; Hashash et al., 2014).

However, reference site identification is sometimes difficult and is replaced by downhole arrays (Cadet et al., 2012; Laurendeau et al., 2018). If the downhole sensors are deep enough (i.e., reaching the bedrock), they may represent the limit at which no site effects are present before. Also, they avoid additional possible uncertainties associated with nearby outcropping rocks considered as reference. These downhole arrays are used to investigate several questions, such as the effectiveness of linear, equivalent linear, and non-linear 1D modeling. For example, if 1D analyses do not match empirical observations, then two reasons are possible. The first is that we could hypothesize that 1D conditions are broken down, and multi-dimensional effects are dominant. The second possibility is that the discrepancies can originate from uncertainties related to site investigations, e.g., obtaining oversimplified site profiles that bias the 1D estimation. As a first step, quantifying uncertainties and their impact on predicted site responses seems more straightforward to consider than investigating the appropriateness of the 1D assumptions in the whole geological context.

In seismic hazard applications, primarily the site-specific, sources of uncertainties in site responses must be appropriately considered, as those that are due to site characterizations. Inherent uncertainties by V_S profiles are essential since they lead to uncertainties on the predicted 1D site responses. Epistemic uncertainty and aleatory variability are the main contributors to uncertainties on V_S profiles necessary for 1D simulations. Epistemic uncertainties can be due to inadequate modeling or insufficient knowledge of the physical processes undergoing. Thus, they could be improved with better knowledge. However, aleatory variability cannot be reduced and reflects the random nature of soil properties. Though the concepts are clear, it is challenging to separate practically between epistemic uncertainties and aleatory variabilities (Griffiths et al., 2016a; Teague and Cox, 2016). Epistemic uncertainty could be investigated by considering different kinds of V_S measurements as in the InterPacific project (Garofalo et al., 2016). Aleatory variability is generally taken into account through randomization models of V_S profiles (Toro, 1995; Parolai et al., 2015; Pilz and Fäh, 2017; Parolai, 2018; Teague et al., 2018; Passeri et al., 2020). However, such randomization models can lead to unrealistic estimations and should be constrained by site signatures (Griffiths et al., 2016a; Teague et al., 2018).

Randomization is done by assuming statistical distributions of some parameters as V_S and layer thicknesses. However, the calibration of key parameters to generate site-specific randomized profiles remains challenging. For this aim, invasive (borehole) and non-invasive (surface) measurements can be used if available. First, non-invasive measurements indirectly provide V_S profiles through several steps. They generally use active or passive sources to measure ground vibration over a large area (radius of tens or hundreds of meters). Then, these measurements are treated to construct an experimental dispersion curve. Indeed, the V_S profile that fits dispersion curves is non-unique, and both epistemic and aleatory uncertainties influence the resulting profiles. Invasive measurements include a direct estimation of V_S and can be influenced by epistemic uncertainties (Garofalo et al., 2016). However, invasive measurements can be used to infer for aleatory variability of V_S as they provide high-resolution profiles. Several invasive measurements of the site can help account for aleatory uncertainty, but it is a rare case. A single invasive measurement can help to quantify vertical soil heterogeneities, often addressed as spatial variability.

This text shows an interest in investigating the impact of vertical spatial variability on 1D-ground response analyses. Therefore, we propose a geostatistical model to account for vertical heterogeneities in soil properties based on the comparisons between resolutions of invasive and non-invasive measurements. The model we propose was inspired by previous studies (Pilz and Fäh, 2017; Parolai, 2018), highlighting the impact of accounting for vertical scattering effects in numerical simulations. The work to validate the model is also inspired by Hashash et al. (2014), which studied the distribution of V_S fluctuations from a defined mean intending to define a reference V_S value. We also consider the proposals of Teague et al. (2018) to bound the resulting uncertainty on site responses with the available site signature. With the presence of downhole arrays on KiK-net sites, we revised the validity of 1D assumptions for some stations using the criterion of Thompson et al. (2012). The work finally leads to a clear and straightforward procedure to account for vertical heterogeneities in 1DSH site response simulations. In the end, we show that including vertical V_S heterogeneities can lead to additional unaccounted attenuation in the high-frequency.

2.2 Data and observations

2.2.1 Observations

One of the invasive geophysical measurements is PS Suspension logging (PSSL), which measures the compressional P- and shearing S-wave downhole velocities. These measurements are usually made in a single borehole filled with a fluid in which sensors are displaced with a constant spacing. These measurements often allow obtaining a high-resolution P- & S-wave profile since it provides a measurement at each chosen depth in the borehole, which is not the case in surface wave methods. Thus, non-coarse (unsmooth) profiles of PSSL can reflect some aspects of the vertical small-scale heterogeneities below the surface. Although downhole methods (DH) are considered invasive, they usually lead to coarse profiles since depth sampling is low. On the other hand, multi-channel analysis of surface waves (MASW) is an example of a non-invasive measurement that leads to coarse profiles. MASW uses seismic surface waves generated by different sources at the surface. Overall, different types of measurements can help assess epistemic uncertainties, as in Garofalo et al. (2016).

In this study, we are interested in the impact of aleatory uncertainty on measured profiles that could be observed by comparing high and low-resolution profiles to study vertical heterogeneities. V_S profiles were collected from published work of the InterPacific Project (Garofalo et al., 2016) and the ongoing V_S profile database in the U.S. (Ahdi et al., 2018). In the InterPacific project, V_S profiles were measured by different methods, including DH and PSSL, to assess the reliability, resolution, and variability of different geophysical methods. In Ahdi's database, the data was collected from different sources, often presenting one or more measurements of V_S profiles (DH, PSSL, or DH and PSSL at the same time) along with geotechnical logs and laboratory tests whenever available. We selected from this database sites that had at least measured V_S profiles with PSSL.

Figure 2.1 shows measured V_S profiles of some sites from PSSL and DH methods, down to depths between 40 and 110 m. Most of the PSSL measured profiles have a depth sampling of 0.5 m. These profiles will form the basis to justify our proposed

geostatistical model afterward, though the model will be integrated into the estimation site responses for a group of KiK-net sites. We also note that PSSL profiles correspond to continuous linear V_S as a function of depth, while V_S profiles should be discretized into sublayers of constant V_S when performing 1D numerical simulations.

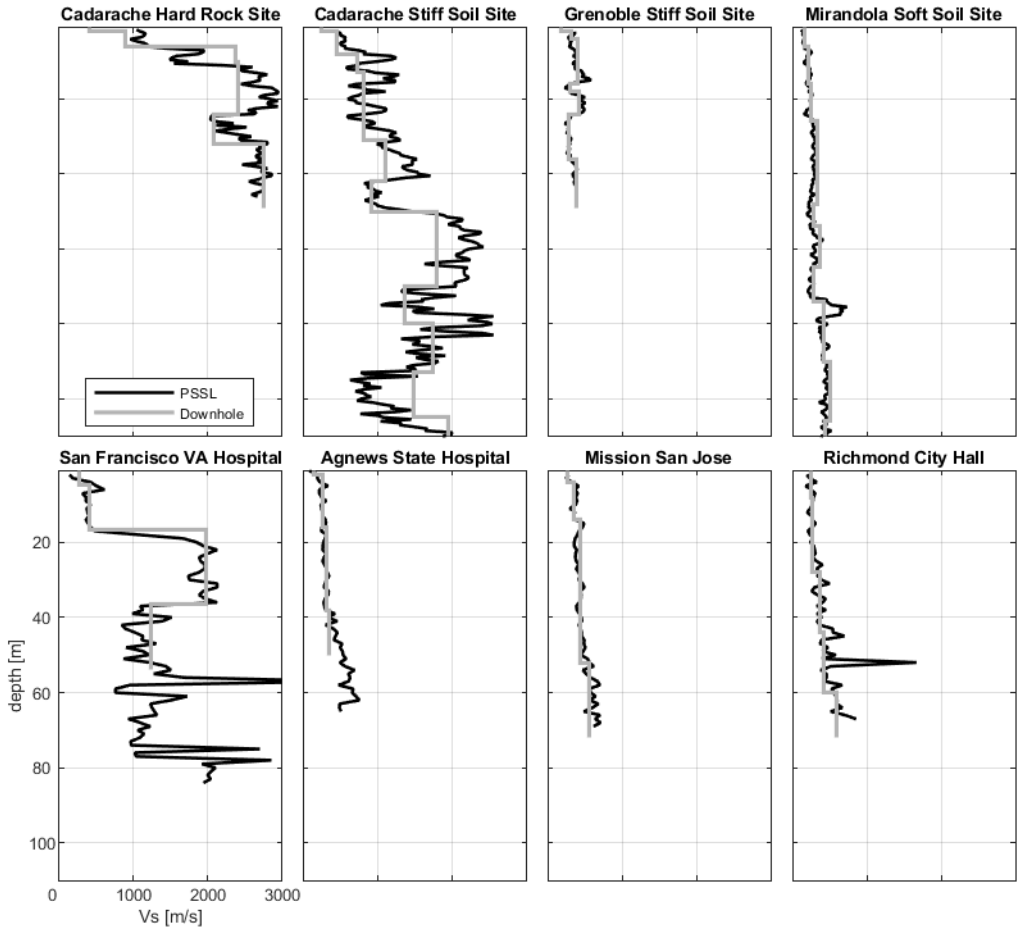

Figure 2.1: The figure shows V_S profiles measured by PSSL high-resolution and Downhole low-resolution profiles. The sites in the first row are from measurements performed within the InterPacific project. The second 4 sites are from the U.S. profiles database (Ahdi et al., 2018).

2.2.2 Measurements of V_S perturbations and Geostatistical model validation

The data we considered should provide information on the aleatory variability of V_S profiles at the within-profile scale. To avoid confusion, it is worth mentioning that V_S profile in a specific borehole may not hold an aleatory nature (i.e., well-determined values), but we treat the small-scale heterogeneities as aleatory to be able to reproduce them with statistical models. In this way, we account for the between-profile variability resulting from several boreholes on the same site (i.e., boreholes few meters apart).

In fact, within-profile variability could be captured if at least two different measurements are available. One of these measurements must have high resolution, i.e., a V_S measurement over short depth intervals. As a consequence, variability can be assessed by comparison to available low-resolution profiles. Though we had only 19 sites with measured PSSL profiles, it is interesting to investigate the distribution of vertical heterogeneities.

We are interested in studying the distribution of heterogeneities, which are first observed as differences between DH and PSSL profiles. We consider that these differences, which resemble vertical heterogeneities, can potentially impact site response analyses. Generally, one can use DH profiles as baseline profiles whenever available, but an alternative should be replaced if a smooth measurement was not available. In the absence of DH measurements, the alternative we propose here is to average V_S profiles over different layers and consider the average profile as a baseline. However, this averaging method does not lead to unique profiles and can be done in several ways, namely the definition of the sub-layers. Here, we adopt three possible definitions of the average (Figure 2.2a):

- 2 m moving average.
- 5 m moving average.
- Averaging after visual inspection of strong impedance contrasts.

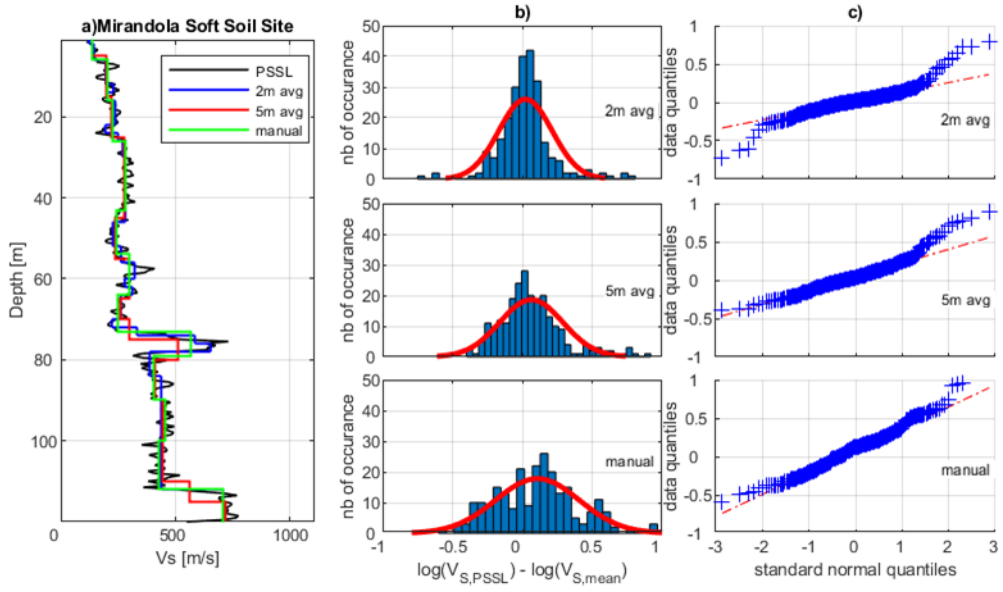

Figure 2.2: a) Different mean definitions for a PSSL measurement: 2 and 5 m averaging in addition to averaging over manually selected layers. b) Distribution of heterogeneities estimated with different baselines. c) Quantile-Quantile plot to verify the distribution of heterogeneities estimated with the different baselines.

Such PSSL measurements are not always possible, and we need to find a way to generate randomly perturbated profiles to consider the heterogeneities. We denote hereafter the vertical heterogeneities as the difference in V_S between PSSL ($V_{S,PSSL}$) and the baseline ($V_{S,baseline}$) defined by X in Eq(2.1).

$$X = \ln\left(\frac{V_{S,PSSL}}{V_{S,baseline}}\right) \tag{2.1}$$

Then, we consider that X is a random variable with a statistical distribution to be explored. Since we cannot confirm *a priori* the distribution, it is convenient to look for normal distribution validity. Figure 2.2b shows the distribution of X measurements using the different defined baselines. Histograms show that the X-distribution is close to a normal distribution, with zero mean and different standard deviations, σ_{scat}. The observations of X appears to be more clustered at the center than that expected for a normal distribution. Then, we perform an additional check by establishing the *Q-Q* plots

in Figure 2.2c, which shows data quantiles versus standard normal quantiles. Whatever the baseline considered, data shows deviation at the tails from the one-to-one line for quantiles beyond 2σ. At the tails, the data have more expected values than a normal distribution.

These measurements are applied to all selected sites, and the results are summarized in Table 2.1. For most profiles, the mean is closer to one for the 2 m moving average baseline. However, other baselines result in quite different values, probably due to different layerings that fit less well the local means. The question of the most reliable baseline profile remains a subjective choice and should be furthermore investigated. We also dropped off the discussions related to correlation lengths of vertical heterogeneities since their measurements are highly affected by the predefined baseline.

Table 2.1: Measurements of the mean and standard deviation of heterogeneities from the InterPacific and Ahdi Database profiles.

Site name	μ_1	σ_1	μ_2	σ_2	μ_3	σ_3
Cadarache Hard Rock Site	1.047	0.154	1.111	0.215	1.256	0.245
Cadarache Stiff Soil Site	1.030	0.272	1.058	0.338	1.242	0.503
Grenoble Stiff Soil Site	1.009	0.216	1.029	0.265	0.985	0.409
Mirandola Soft Soil Site	1.023	0.190	1.067	0.225	1.115	0.300
Agnews State Hospital	1.070	0.273	1.090	0.206	1.231	0.138
Belmont2-Story Building	0.959	0.476	0.853	0.616	0.796	0.650
Capitola Fire Station	1.024	0.308	1.032	0.280	1.074	0.367
Castaic Area High School	1.036	0.149	1.100	0.231	1.312	0.287
Halls Valley	1.046	0.393	1.108	0.360	1.485	0.441
Lexington Dam	1.023	0.177	1.126	0.185	1.149	0.181
Martinez VA Hospital	1.057	0.150	1.170	0.225	1.424	0.346
Menlo Park VA Hospital	1.034	0.378	1.051	0.359	1.161	0.385
Mission San Jose	1.011	0.136	1.051	0.131	1.159	0.137
Pacific Heights Fire Station	1.050	0.266	0.940	0.199	0.982	0.259
Patterson Pass Road	1.083	0.138	1.130	0.136	1.203	0.119
Piedmont Junior High School	1.128	0.470	1.549	0.401	1.423	0.528
Point Bonita	1.061	0.487	1.134	0.449	1.247	0.469
Richmond City Hall	1.061	0.297	1.088	0.262	1.184	0.259
San Francisco VA Hospital	1.057	0.446	1.191	0.468	1.555	0.853
Averaged values per baseline	1.040	0,280	1,100	0,290	1,210	0,360

In the end, we consider a normal distribution to take into account vertical heterogeneities in 1DSH simulations. The values estimated in Table 2.1 illustrate a possible range of sigma values and help calibrate the simulations. However, these estimations might be performed to specific sites of interest if measurements were available. Otherwise, a proper choice of distribution parameters should be considered in the application. We conducted these measurements in our study to confirm the possible range of values and distributions, and we do not intend to generalize the values obtained to other applications.

2.2.3 The geostatistical model proposed

Spatially heterogeneous soil properties were found to impact ground motions at KiK-net sites (Thompson et al., 2009). Several studies addressed the validity of 1D conditions at KiK-net sites (Thompson et al., 2012; Laurendeau et al., 2018; Pilz and Cotton, 2019; Zhu et al., 2020) and found that they can be dominant in around 30 percent of sites. Kaklamanos et al. (2020) investigated different physical hypotheses to modify existing V_S profiles to improve 1D site response analyses. One of the most commonly used methods to account for the aleatory nature of soil properties is the method of Toro (1995) which randomizes soil profiles. Toro's (1995) model assumes a Poisson's distribution of depths to define sublayers and lognormal distribution of V_S to generate randomized values around a predefined median (i.e., baseline).

The Toro model has shown that it could lead to increased variability (Griffiths et al., 2016a, 2016b; Teague and Cox, 2016) and can lead to unrealistic dispersion curves. Others proposed to constrain the realization of the Toro model by site signatures to avoid unrealistic variability (Teague et al., 2018; Passeri et al., 2019). Passeri et al. (2020) proposed a new geostatistical model based on travel times of S-waves along with measured profiles and a sub-layering model. However, for both Toro (1995) and Passeri (2020) models, specific distributions are assumed not only for perturbated V_S but also for layer thicknesses and inter-layer correlations. The latter issue requires validation for each of the distributions assumed. In a broad application of accounting for uncertainties in V_S profiles, hereafter, we propose a simpler model based on the direct observations of V_S profiles.

Profiles delivered by PSSL methods appear to have small-scale perturbations. These perturbations can have significant consequences on the high-frequency amplification and appear to follow a random process, and could only be quantified and applied statistically. We consider a model that generates a similar scale of heterogeneities close enough to observations from PSSL. We mimic these observations through thin sub-layers with variable velocities as addressed recently (Parolai et al., 2015; Parolai, 2018). Parolai (2018) showed that considering a random perturbation in the travel time for the first 100 meters would lead to a decay in the high-frequency spectra. Another similar analysis was done by Pilz and Fäh (2017) based on the contribution of scattering to near-surface attenuation after statistical observations of the seismic wavefield at sites of the Swiss seismic networks. They showed that attenuation properties show a clear dependency on the local shallow subsoil conditions with differences in the structural heterogeneity of the shallow subsoil layers producing different scattering regimes.

The perturbation of the velocity profile is expected to mimic the effects of scattering by introducing in a similar way thin sub-layers down to a maximum depth h_{scat}, with the following assumptions:

- Thin sublayers of thickness dh over the whole considered depth range on which we apply a perturbation in the travel time.
- The S-wave travel time is perturbed by a factor X in the velocity profile for each sub-layer according to Eq(2.2). Here, t_{new} is the perturbed travel time, V_s is the original unperturbed *S*-wave velocity profile, dh is the sublayer thickness, and X is the random perturbation introduced.

$$t_{new} = \frac{dh}{X.V_s} \quad (2.2)$$

- The perturbation factor X is chosen to follow a lognormal distribution, after the observations presented in this text, with a median value μ_{scat}= 1 and a standard deviation σ_{scat} that can take values between 0.2 and 0.5. For a case study, σ_{scat} is expected to be tuned over site-specific measurements, but we propose to set it to common values (e.g., 0.3) in case of a lack of information.
- The only constraint is that the total travel time in each sublayer of the original profile should not highly change (with a tolerance of 5 percent). It is also important to note the quality factor Q is determined in our 1D simulation using the scaling relation Q=V_S/10 as in several studies (Cadet et al., 2012; Laurendeau et al., 2018), which reflects the velocity heterogeneities in the quality factor.

All randomization models can result in high uncertainties if they are not constrained either by lower and upper bounds of perturbations or by empirical site observations. Several studies have addressed the possibility of matching randomized profiles to empirical site signatures, such as dispersion curves and fundamental f_0, to constrain the uncertainty obtained and avoid additional uncertainties. These site signatures are expensive to obtain or simply unavailable for many sites, as the case for KiK-net data (no dispersion curves are available). However, we benefit from the possibility to estimate HVSR on KiK-net arrays to estimate f_0, which can be used to constrain the geostatistical model realizations.

2.3 KiK-net data, empirical estimations, and site signature

The usual way to assess the dominance of 1D-effects in site effects is to compare observed and theoretical transfer functions through the classification procedure initially proposed by Thompson et al. (2012). The comparison is based on the estimation of Pearson's correlation coefficient *r* between empirical and 1DSH transfer functions. Its development was facilitated with KiK-net data due to surface and downhole recordings that allow direct estimation of the standard spectral ratios with downhole sites as reference (SSR_{dh}). In this text, we consider KiK-net data to investigate the impact of aleatory uncertainty on the 1D site-response estimations and, consequently, the 1D classification procedure. Previous works have addressed this question showing that a low percentage (~30%) of KiK-net sites show dominant 1D effects (Thompson et al., 2012; Laurendeau et al., 2018). We will try to find if improvements are observed when accounting for uncertainties introduced by vertical V_S profile perturbations.

In KiK-net arrays, downgoing wave effects can be substantial and affect the SSR_{dh} estimation. Destructive interferences can occur in borehole sensor levels at different frequencies f_{dest}, which can cause an amplification peak in SSR_{dh} ratio before (or after) f_0. That is why these empirical estimations can be biased, and comparisons should be performed among frequencies beyond f_{dest}. Usually, HVSR is estimated, highlighting site fundamental frequency f_0 (Lermo and Chávez-García, 1993). However, identifying a clear f_0 is conditioned by the presence of a strong impedance contrast in the layers beneath the site.

SSR_{dh} and HVSR are estimated for KiK-net sites over frequencies between 0.1 and 50 Hz using recordings between 1997 and 2017 (Bahrampouri and Rodriguez-Marek, 2019). A frequency band of 0.5-20 Hz is fixed for the signal-to-noise ratio test (SNR>3). Figure 2.3 shows transfer functions for four KiK-net sites using empirical (SSRdh and HVSR) and 1DSH estimations through the borehole transfer function, BTF. SSR_{dh} and BTF are the comparable estimates of amplification in these figures since both are surface to downhole ratios. HVSR represents the amplification with respect to a vertical component that might exhibit some amplification. However, HVSR is helpful to identify f_0 unlike the SSR_{dh}, which can show peaks before f_0 due to destructive interferences in the borehole. For example, AOMH11 has its f_0 at 6.5 Hz identified with HVSR while its first peak corresponds SSR_{dh} to f_{dest}. We also note that generally empirical estimations show consistent peak frequencies concerning f_0, while it is not the case for 1DSH transfer functions, as for EMH03 and AOMH11. The horizontal shift of the fundamental resonances implies a low *r*-value. Thus, the site will be classified as far from having dominant 1D effects. On the other side, high *r*-values are obtained when f_0 peaks are aligned, potentially proposing a 1D classification of the site as the case for IWTH27 and AICH20.

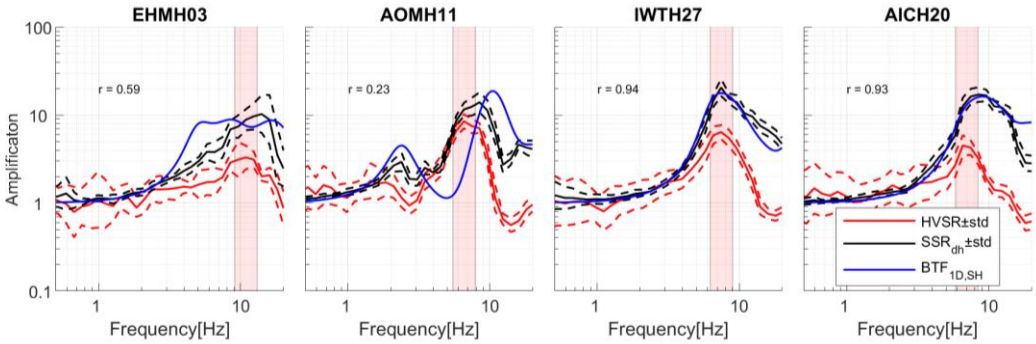

Figure 2.3: Amplification factors as estimated from standard spectral ratios SSR_{dh}, horizontal to vertical spectral ratios HVSR, and theoretical borehole transfer functions BTF_{1DSH}. The red-shaded area aims to highlight the fundamental frequency f_0.

After thorough comparisons of site effects, we can see that many of the low *r*-valued sites correspond to shifted resonance peaks. Hereafter, we question whether uncertainties in the VS profiles could stand behind these shifts and explain or improve the classification. Several studies applied this 1D site identification on the KiK-net sites (Thompson et al., 2012; Laurendeau et al., 2018; Pilz and Cotton, 2019), and it could be interesting to revise the classifications taking into account such V_S uncertainties. The fundamental peak shift can also potentially impact the mean amplification in high-frequency.

Aleatory uncertainty that we aim to test in this text will be through introducing vertical heterogeneities in V_S profiles. The geostatistical model proposed should keep track of site footprints such as the site-specific dispersion curves and f_0. Teague et al. (2018) addressed this issue through an application to Garner Valley downhole array. For the KiK-net arrays, we do not have available dispersion curves, but it is possible to estimate f_0 from HVSR. Thus, we use the model developed in Shible et al. (2018) and improve it by selecting profiles that match the available site signature. We select 56 profiles with different *r*-values and having clear f_0 peaks to ease the discussions.

2.4 Tuning the model before application:

The geostatistical model we propose to generate random heterogeneities in V_S profiles does not result in a single and unique profile. Thus, to account for uncertainties, we considered generating 40 realizations per site and taking the median of all or some of the resulting site responses for statistical validity. We consider that 40 realizations are enough, in a first step, to explore the impact of heterogeneities. Figure 2.4 shows an application for 3 KiK-net sites by setting perturbations up to 30 m, σ_{scat} to 0.3 and inter-layer spacing dh to 0.5 m. This choice of parameter values is just first step to test our proposed model. However, this selection will be more discussed and justified afterward. It is evident in the obtained responses that there is an important variability in intermediate to high frequencies, mainly resulting from shifting the resonance frequencies to lower or higher values. We believe that an additional constraining condition by site signature should be applied to keep matching realizations only. The last step is expected to achieve a reliable improvement of site response prediction.

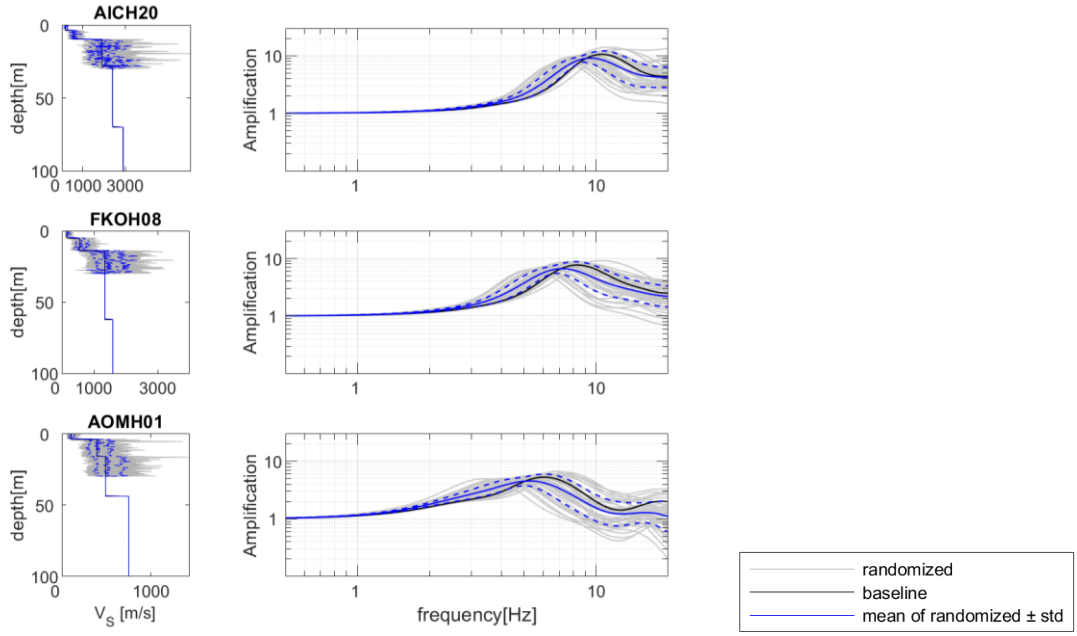

Figure 2.4: For each site, 40 realizations of the randomized models are performed, and the corresponding 1DSH responses are shown. The median and logarithmic standard deviations are plotted as well.

The chosen geostatistical model has a physical background that conceives some interpretations and consequences. The small-scale heterogeneities to be introduced tend to alter high frequencies mainly beyond 5 Hz (short wavelengths). However, this should be verified for each site since different realizations of the same distribution can highly differ and affect the mean travel times (mean V_S) between any two depths that control lower frequencies. The impact on frequencies depends finally on the existing geological structure and impedance contrasts for the considered site. Therefore, to detail more, we conduct a sensitivity study on the key physical parameters of this model to generate random profiles.

2.4.1 Quarter wavelength approximation

First, the depth of introduced heterogeneities h_{scat} and vertical perturbations should be introduced to account sufficiently for vertical heterogeneities. This depth is strongly related to the frequency band of interest, and it is important to quantify this relation

even in an approximate manner. To find the depths that should control frequencies within the frequency band at 0.5-20 Hz, we propose the quarter wavelength approximation (QWL). QWL was initially proposed by (Joyner and Boore, 1981) and subsequently optimized by Boore (2003) to display a depth-frequency relation. This approximation establishes a correspondence between the average velocity up to a depth z (V_{SZ}) of superficial soil layers up to depth z and the quarter-wavelength frequency by $f_{QWL}(z) = V_{sz}(z)/4z_{QWL}$. Velocity averaging is done through the travel times as in Eq(2.3). Then after applying the quarter-wavelength condition, we get Eq(2.4).

$$\frac{z}{\overline{V_{sz}}} = \int_0^z \frac{du}{V_s(u)} \tag{2.3}$$

$$\int_0^z \frac{du}{V_s(u)} = \frac{1}{4f_{QWL}(z)} \tag{2.4}$$

We set the f_{QWL} value to 5 and 10 Hz separately, and we solve the last equation to find depths needed to validate the equality. This procedure was applied at each site using the provided velocity profiles in KiK-net. However, considering lower frequencies (< 5 Hz) in the analyses will be unnecessary since the small-scale perturbations should, in principle, not affect low frequencies.

The results displayed in Figure 2.5 indicate that depths between 0 and 50 m control intermediate frequencies around 5 Hz, while depths between 0 and 20 m control high frequencies around 10 Hz. Thus, the controlling depths are shallower for higher frequencies than lower frequencies. Since we are interested in the high-frequency part, we consider it sufficient, in the first step, to introduce heterogeneities down to 50 m depth at most. In this way, we ensure that the study is limited to depths that can affect high-frequency and avoid unnecessary heterogeneities beyond a specific depth.

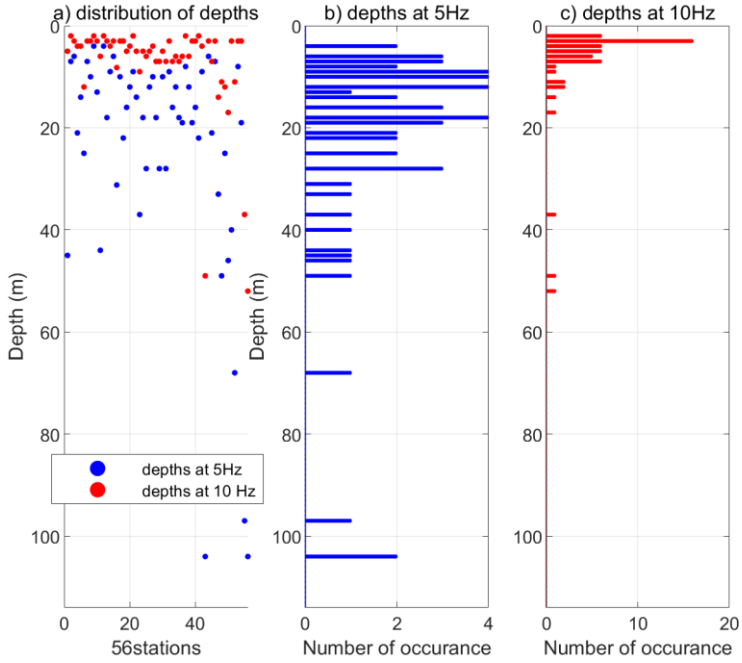

Figure 2.5: Quarter wavelength (QWL) method applied at frequencies 5 Hz and 10 Hz, showing the QWL depths on a scatter plot(a), and the number of occurrence for QWL depths is shown for all the 56 sites (b and c).

2.4.2 Sensitivity analyses of key parameters

A sensitivity study was needed to show the impact of each of the key parameters: h_{scat}, dh and σ_{scat}. First, varying h_{scat} can show its impact on the predicted transfer functions. In fact, vertical heterogeneities are expected to be more important in superficial layers that are often subjected to alteration and weathering than deep and rigid layers. Figure 2.6 shows the effect of variation of h_{scat} between 20, 30, 50, and 70 m on three stations, EHMH03, AOMH11, and IWTH27. The test on the three stations resulted in an invariant mean and variability of the predicted site responses for all frequencies. This result is expected as our proposed geostatistical model is applied by varying V_S over thin sub-layers, which should alter only short wavelengths. Thus, as dh is kept constant, limited modifications are observed with varied h_{scat}. It could also be interesting to test for other combinations of dh and h_{scat}. Finally, for dh = 0.5 m, h_{scat} can be set to 30 m.

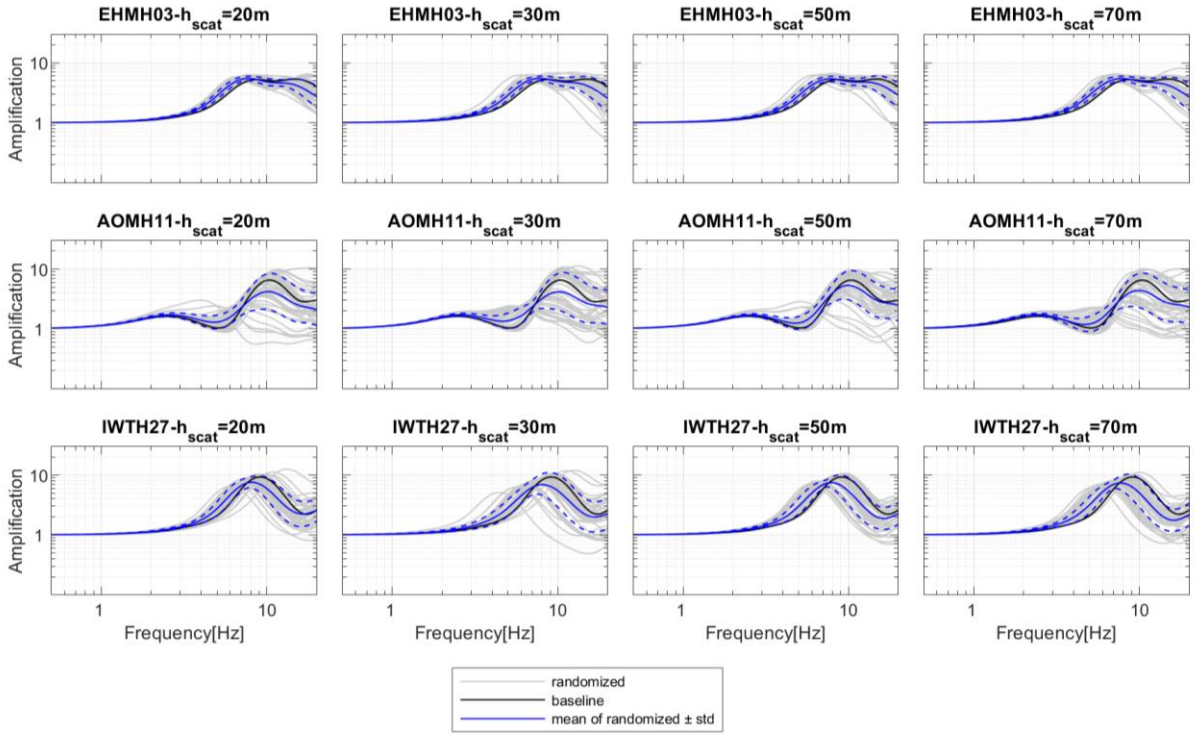

Figure 2.6: The same concept as Figure 2.4 but with varying h_{scat} to show the impact on the mean and variability of 1DSH transfer functions.

Inter-layer spacing, dh, remains a subjective decision to discretize the profile, which is also the case in high-resolution geophysical measurements. The most convenient is to consider inter-layer spacing at 0.5 m, as in most measurements. However, increasing dh can lead to some numerical limitations, especially when it exceeds some layer thicknesses. In the latter case, changes for layers depths in the baseline model can be considered with models that consider depth-dependent layer thickness distributions, as in the randomization model of Toro (1995). Indeed, calibrations from real measurements with 2 m or higher spacing, as performed earlier in this text, could appear somehow problematic in statistical testing of the model due to unclear layer-by-layer perturbations. In Figure 2.7, we show the results of varying the dh for the three KiK-net stations. As dh increases (for instance, dh=2 or 3 m), predicted mean site-responses show slight changes in mean values as well as limited variabilities. The results can change from one site to another but always shows limited effects for an increased inter-layer spacing. Hence, we consider the dh=0.5 for two reasons; the first is that maximal effects are observed for lower values of dh, and the second is that 0.5 m is very commonly used in PSSL methods.

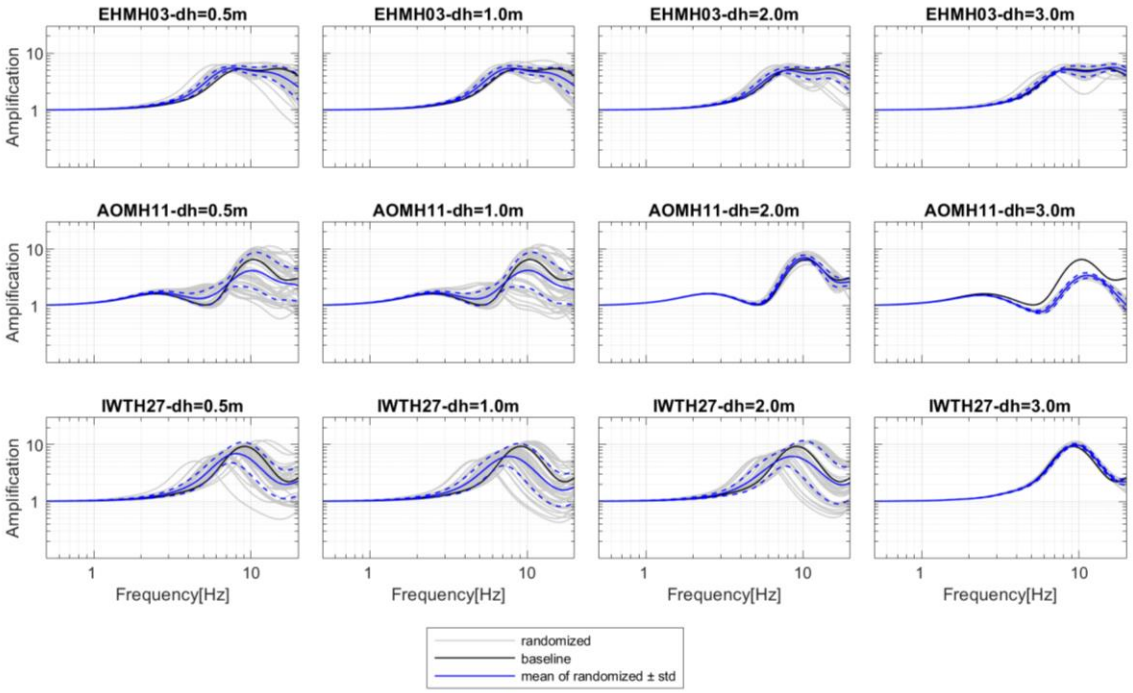

Figure 2.7: the same concept as Figure 2.6 but with varying dh.

We have seen previously that perturbations severity, ensured by σ_{scat}, can be site-dependent. Knowing that we need robust estimations of σ_{scat} for the considered KiK-net sites, we perform an additional sensitivity analysis on this parameter to highlight its effects. Physically, increasing the heterogeneities level could be interpreted as increasing scattering effects leading to additional attenuation effects in the transfer function. Figure 2.8 shows the results of testing different σ_{scat} values. Increasing σ_{scat} leads to a more attenuated site-response as well as increased variability. In the end, we consider that σ_{scat}=0.3, which is the median of measured values of Table 2.1.

It appears from the performed tests that mean site-responses and variability changes are controlled mainly by the shift of site fundamental frequency f_0. This shift is often observed towards the lower frequencies, leading from one side to reduce the amplification beyond 10 Hz and from another side to a smoothing of the mean site-response.

It is also interesting to see such an effect for all the sites selected. So we compute the ratio between predicted initial and randomized site responses as shown in Figure 2.9a. Rigid sites associated with high V_{S30} are shown to have limited changes from initial site responses, while soft-soil sites are highly affected. As for the variability of the predicted randomized responses, rigid sites (V_{S30}>1200 m/s) show limited variability, also suggesting a limited effect of randomized profiles, unlike soft- and stiff-soil sites (V_{S30}<800 m/s). This observation proposes that ground motion on rock sites, which is often considered a reference, is less susceptible to vertical small-scale heterogeneities.

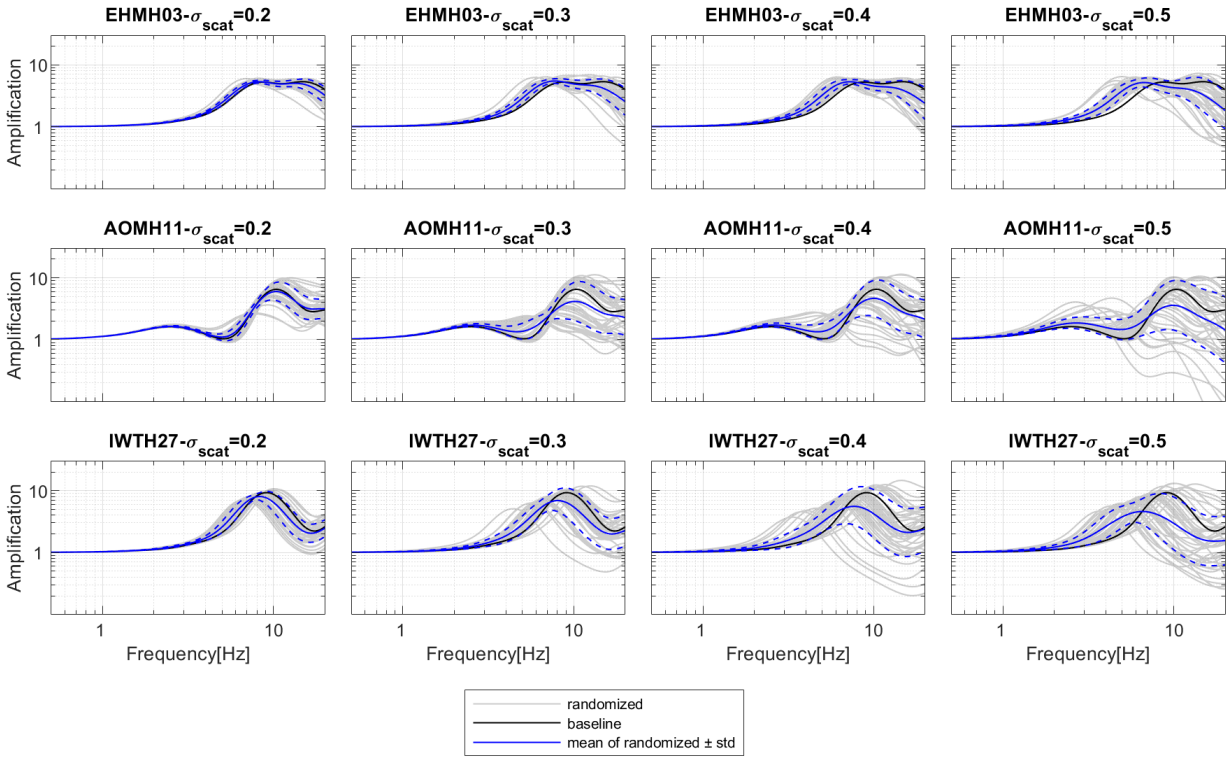

Figure 2.8: the same concept of Figure 2.6 but with varying σ_{scat}.

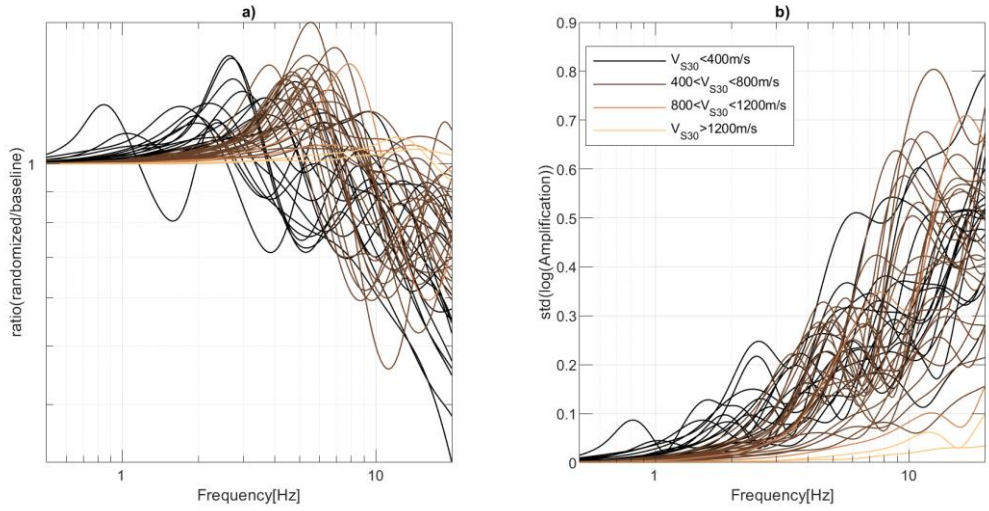

Figure 2.9: a) ratios between new and initial site-responses b) the variability of new median site-responses represented by logarithmic standard deviations categorized into different V_{S30} classes.

2.5 Discussions:

In what follows, we conduct direct comparisons between empirical and theoretical site-responses modified with the geostatistical model over the selected KiK-net sites assuming h_{scat} =30 m, dh = 0.5 m and σ_{scat} = 0.3. We introduce the proposed improvements to match available empirical site signature, f_0. As a reminder, this application mainly highlights modifications on the predicted 1DSH site-response, especially in the high-frequency range.

2.5.1 An additional constraint to select randomized profiles

To assess if any improvement can be delivered to 1DSH site responses, we can have two possibilities of interpretations. First, if the fundamental peaks are slightly shifted, then they can be recovered by the proposed geostatistical model. Second, if all the generated realizations do not explain the differences with empirical observations, then it is highly possible that the sites are highly multi-dimensional. In this case, no improvement could be attained, whatever the modification is.

As seen earlier, variability in the site resonance peaks could be observed if vertical heterogeneities are introduced. However, it is essential to constrain such variability with empirical observations for each considered site. We consider that f_0 from HVSR should be attained by one or several randomized profiles. Therefore, we propose to select site responses that match empirical f_0. Recent studies also started addressing HVSR as it could be used in the reliable estimate of empirical transfer functions for KiK-net sites after a correction procedure (Kawase et al., 2019; Zhu et al., 2020). However, in this text, we use HVSR only to identify true f_0 and not for amplification estimation. We also do not claim that the model will lead to perfect matching peaks with HVSR on all KiK-net sites. In the cases where it is impossible, the explanation that holds is that the site will be far enough from 1D conditions to have matching f_0.

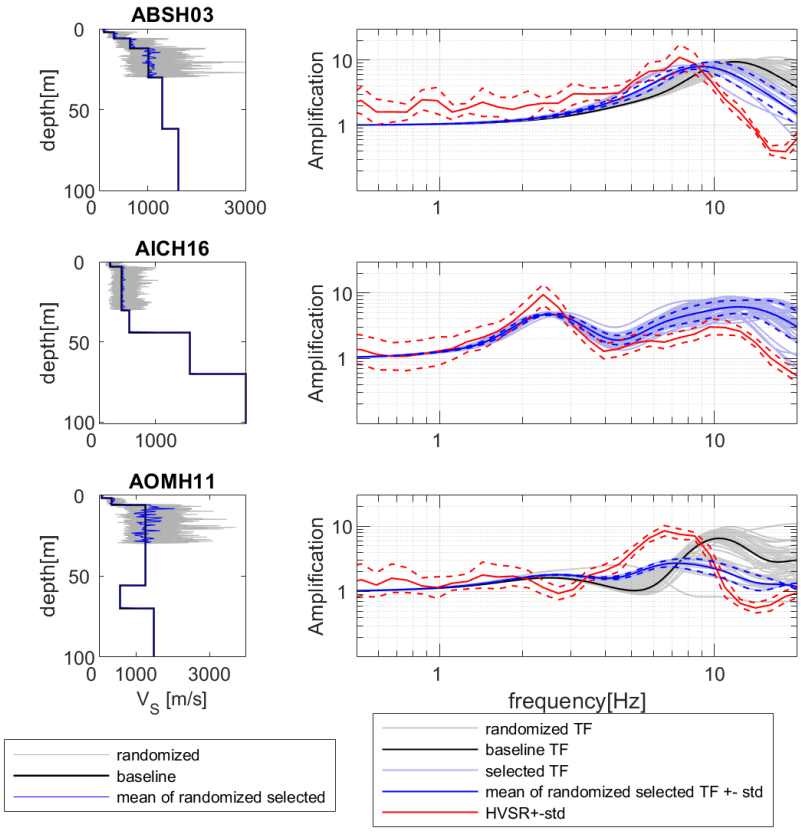

Figure 2.10: Baseline and randomized V_S profiles and the corresponding transfer functions. After comparing with HVSR, best matching realizations are selected, and their median and standard deviation are shown.

Figure 2.10 shows site responses of randomized profiles compared to HVSR. The constraint we apply is done first through a comparison of empirical and predicted f_0.

Then we select only the profiles that match with empirical f_0. For the considered sites, perfect alignment of fundamental peaks between HVSR and TF_{1DSH} is not always possible, but we tend to decrease the shift. For example, the AOMH11 station shows a possibility to optimize the shift of peaks in frequency values but leads to much lower amplification than the baseline. For such a case, if improvements to frequency shifts are obtained upon an unexpected drop in amplification, then the site is most probably dominated by multi-dimensional effects.

Several studies addressed the possibility of accounting for profile uncertainties, which proposed the concept of site signature to constrain such uncertainty and derive robust site effects (Griffiths et al., 2016a, 2016b; Teague et al., 2018). As these site signatures are purely empirical and derived from data, we should have a reliable estimate of them (i.e., dispersion curves and f_0). It is also important for development purposes to have the empirical amplification for some sites to compare the resulting 1D site responses. However, for KiK-net sites, not all needed details are available for this purpose. For example, experimental dispersion curves are not present, and the only fo could be identified from HVSR. Unlike the case of GVDA of Teague et al. (2018), where true amplification was available, downhole sites of KiK-net were used as a reference in spectral ratios computations. Downhole sensors are subjected to destructive interferences from down-going waves, which could bias SSRdh from the true amplification. However, we assume in this text that comparisons beyond the fundamental destructive frequency should hold and can be sufficient to identify 1D sites.

Figure 2.11 shows the comparison between SSR_{dh} and different 1DSH estimations (i.e., baseline case profile, perturbated profiles, selected matching ones, and the corresponding median). Generally, the fundamental peak observed on HVSR is also observed on SSR_{dh}, and thus optimizing the 1DSH peaks with HVSR is expected to optimize the resemblance between BTF and SSR_{dh}. We thus revised the *r* coefficient estimation (r_{before}) after introducing heterogeneities (r_{after}). Many sites show a significant improvement in the *r*-value estimation (as MYGH07), increasing the number of sites that could be classified as 1D sites (as ABSH03 and AICH16). However, some sites show increments in r_{after} values, and it remains challenging to match empirical responses and be classified as 1D sites (AOMH11).

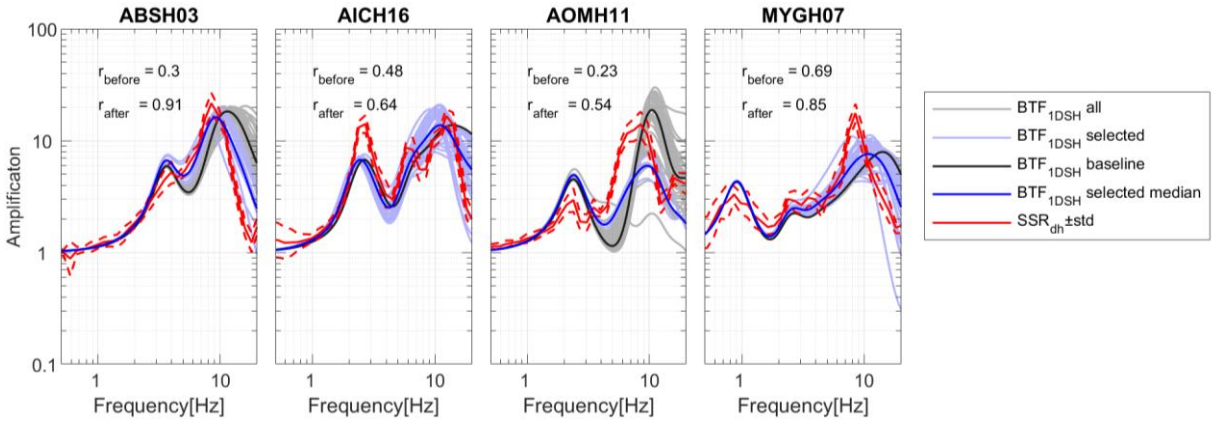

Figure 2.11: Site amplification as estimated by surface to borehole ratios from empirical (SSR_{dh}) and 1DSH (BTF) approaches. Randomized and baseline site responses are shown as well as the amplification of selected realizations. Initial r coefficient and revised ones are also displayed.

2.5.2 Statistical impact for selected sites

We conduct the previous steps on the 56 selected sites to derive some statistical conclusions. It appears that this procedure reduces the discrepancies between theoretical and empirical estimations to some extent. As shown in Figure 2.12a and b, a significant improvement is observed from r_{before} to r_{after}, and several sites enter into the list of 1D sites. The applied procedure clearly tends to increase the initial *r*-values. This observation is due to the improvement of peak alignment, as shown earlier in this study.

We then are interested to see the median impact of such a geostatistical model on site-response prediction. We compute the ratios between the median selected randomized and the baseline case, as in Figure 2.13a. We can see that most of these sites modified the response by shifting high and intermediate frequency peaks into lower ones. That is why we observe a median drop that reaches 0.55 at 30 Hz. In Figure 2.13b, we limit the analyses to r_{after}>0.6, which defines the modified list of 1D sites. The mean BTF shows a significant drop at high-frequency, which appears closer to the mean SSR_{dh}. These results could be one of the explanations to the issue raised in Laurendeau et al. (2018) when they found intermediate and high-frequency discrepancies when comparing mean empirical and theoretical 1DSH site-response estimates.

Finally, the comparisons of Figure 2.13b show that median BTF still shows a slight difference from the median SSR_{dh}. The latter could be possibly improved if we consider site-specific values for model realizations, i.e., μ_{scat} and σ_{scat} values that can change from one site to another. Besides, medians over a broader site list could lead to more representative statistics. However, the work done using the selected 56 KiK-net stations is enough to conclude a non-negligible effect of small-scale heterogeneities in the subsurface on the high-frequency ground motion.

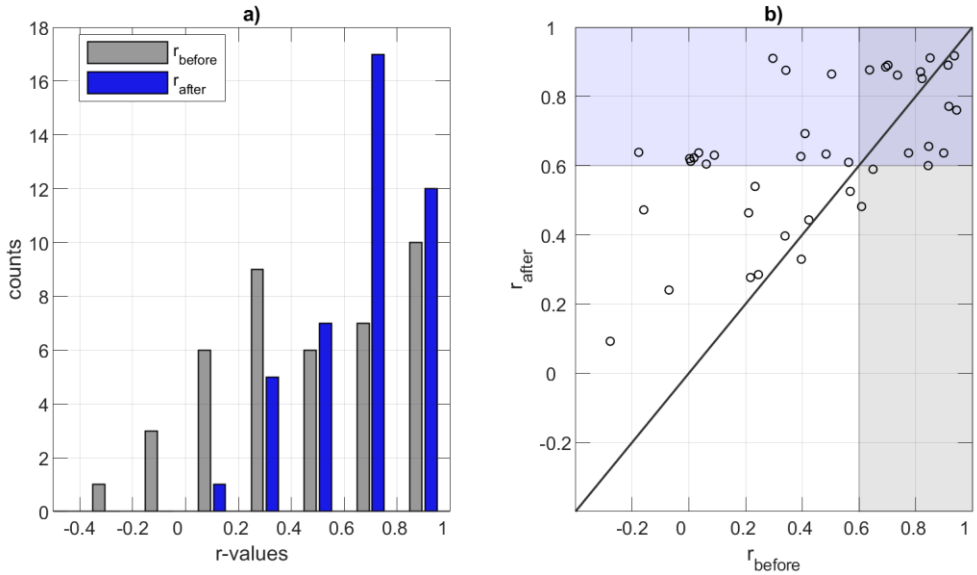

Figure 2.12: Comparing the initial and revised r values for the selected sites: a) distribution of values between 0.4 to 1. b) one-to-one plot of r-values. The patch corresponds to beyond which the site is considered 1D.

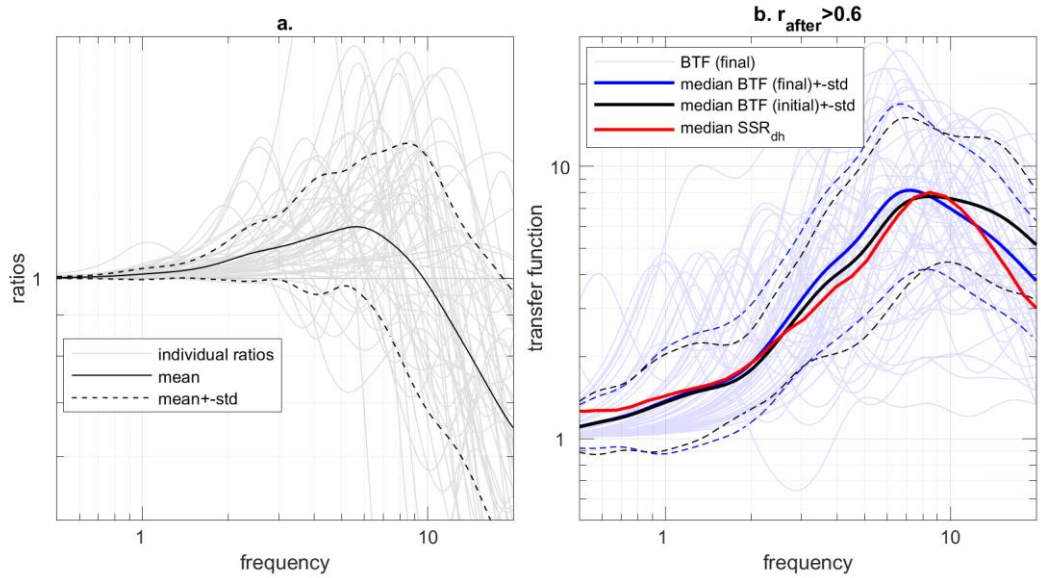

Figure 2.13: a) Ratios between new and initial site responses for all sites, their median, and the standard deviation. b) mean theoretical and empirical surface-to-borehole transfer functions (BTF and SSR_{dh}). BTFs in light blue are functions for each site after limiting the comparisons to $r_{after}>0.6$. (initial BTF is the baseline site response without heterogeneities, final BTF is the site response with heterogeneities).

2.6 Conclusions

A new application of modeling spatial variability was presented in this text to account for uncertainties. We were interested in assessing the final impact of heterogeneities on the 1D predicted site responses. For this aim, we focus on observations from direct (high-resolution) and indirect (low-resolution) site measurements to quantify vertical heterogeneities. Though the distributions of heterogeneities were tested on some specific sites, we were interested in assessing the impact of introducing such variability in 1DSH analyses at KiK-net sites. Therefore, we developed a simple model to generate randomized profiles accounting for vertical heterogeneities and thus mimicking, though in a limited way, scattering conditions. After sensitivity analyses, we also see that introducing vertical heterogeneities between 20 to 70 m has a similar effect. The main parameter appears to be the deviance from the median (σ_{scat}).

We consider that free constraints will lead to increased uncertainties, and thus limitation to site signature is necessary (Teague et al., 2018). However, at KiK-net, we have only fundamental peaks f_0 from HVSR that could be identified. We thus conduct a series of comparisons between selected KiK-net sites. We find out that constraining the realized estimations with matching f_0 improves 1D site response prediction. However, some sites appear to be far from 1D conditions, and that accounting for vertical spatial variability is not enough to be considered as 1D sites. In the end, we conclude that small-scale perturbations can affect the theoretical transfer functions, mainly by shifting the resonance peaks towards lower frequencies and leading to a lower amplification in the high frequencies.

Chapter 3: GITEC, A Generalized Inversion Techniques Benchmark

Context

GIT methods can be applied in different ways, either parametric or non-parametric. Since each method is applied differently, variabilities could exist between their results. This question appears important in the framework of the thesis. As stated earlier, the deconvolution approach relies on a reliable estimate of site responses, and we are aiming to estimate them with GIT. Thus, it is interesting to investigate the uncertainties associated with GIT site responses. Since GIT involves a source and attenuation effects separation, possible trade-offs and biases are possible. These issues were the main motivations for conducting the benchmark presented in this chapter, which is also an article draft in the final preparation stages.

Abstract

Generalized inversion techniques (GITs) have become popular for determining seismological parameters (e.g., source, attenuation, site responses), particularly in low-to-moderate seismicity regions. Indeed, GITs can potentially provide reliable site response estimates when a minimum number of recordings is available, as well as valuable information about source features (e.g., magnitudes, corner frequencies, stress drops) and regional attenuation characteristics. Since the early 1990s, significant advances have been made on GITs, where different approaches and alternative basic hypotheses were investigated, such as the application of 'nonparametric' and 'parametric' inversion schemes. In this context, several scientific questions have arisen that depend on the final scope of the GITs: What is the optimal inversion strategy for a given dataset configuration? What is the impact of the different assumptions and implementations on the reliability of the results? What is the dependence of the results on the chosen reference conditions? Is it possible to quantify the associated epistemic uncertainties? Here, we have considered and compared the different approaches of GITs to improve the understanding of each for use for different applications. A methodological benchmark that includes different GIT methods and dataset configurations is set up to fulfill the objective, using a simple synthetic dataset, a French regional sparse dataset, and an Italian national dense dataset. The benchmark is developed in two phases: (i) phase I: a free phase with no common constraints; and (ii) phase II: a constrained phase with unified reference conditions. Despite unifying the reference conditions in the different inversions, the variability was not reduced. Discrepancies are observed between different terms of GITs at different levels. Site responses appear to be the most robust estimates for GITs, compared to source and attenuation term parameters. The way that stress drops of earthquakes and quality factors for crustal attenuation are parameterized appears to lead to significant variability between different approaches. Finally, uncertainties are addressed by quantification of the inter-method variability for the different terms and parameters.

Keywords: *benchmark, generalized inversions techniques, parametric, non-parametric, inter-methods variability*

3.1 Introduction

Seismic waves initiate from faults and ruptures in the Earth's crust and propagate from the source to the Earth's surface through different paths to impact built structures and installations. The ground motions observed at the surface can be greatly affected by several factors, such as the rupture characteristics (i.e., source effects), the seismic

wave attenuation properties of the Earth materials between the source and a specific site (i.e., propagation effects), and the amplification of ground-motion amplitudes induced by certain lithological and topographical configurations in the near-surface subsoil (i.e., site effects). After several destructive earthquakes during the last few decades (e.g., Mexico 1985, Kobe 1995, Haiti 2010, Tōhoku 2011), accurate evaluation of strong ground-motion factors has become a necessary step for realistic predictions of the ground motion that can be expected for future strong earthquakes.

Seismic hazard assessment is carried out to predict the intensity of ground shaking either in a given region or for a specific site. In seismic hazard assessment, the ground-motion assessment is generally carried out using ground-motion models, known as ground-motion prediction equations (GMPEs), and they are determined through regression studies on empirically recorded data. Recently, probabilistic hazard studies have been progressively refined to improve site-specific hazard estimates by relaxing the ergodic assumption (Rodriguez-Marek et al., 2013; Kotha et al., 2016).

Indeed, recent probabilistic seismic hazard assessments have progressively moved from the use of fully ergodic ground-motion models toward nonergodic models (e.g., Landwehr et al., 2016). Site-specific seismic hazard assessment requires knowledge of the amplification of the ground motion at the site considered. Most of the current GMPEs base the evaluation of site responses on the V_{S30} proxy (time-averaged S-wave velocity in the first 30 m of the profile) and the high-frequency attenuation parameter, κ. The V_{S30} proxy can reflect some site properties, but it is not enough to describe the frequency-dependent amplification of ground motion induced by the soil properties beneath the site considered. In addition, several sources of uncertainties are associated with κ measurements and applications to take into account local attenuation effects. As a result, there is increased attention toward improving the practice through more precise site-effect estimations.

Within this scope, Bard et al. (2020) reviewed the main practices to obtain site-specific ground-motion predictions and provided several recommendations. One of the main recommendations was to investigate site effects with generalized inversion techniques (GITs), mainly because these represent an alternative tool for evaluating empirical site responses in the Fourier domain. However, the latter use of GITs does not hide that, along with site terms, they can deliver other source and attenuation parameters, such as stress drop and quality factors, which are of the main elements used in forward stochastic modeling (Boore, 2003). Indeed, GITs were first introduced by Andrews (1986), based on the assumption that the Fourier spectrum of a recorded seismic signal can be separated into three main components: source, path, the site factors. Then, generalized inversions were used in numerous studies that focused on crustal attenuation (e.g., Parolai et al., 2004; Bindi et al., 2006), source parameters (Oth et al., 2017), and site-response estimations (Edwards and Fäh, 2017; Kawase et al., 2019).

In the context of ground-response analysis, for example, GITs provide empirical estimations of site responses, as an empirical method to overcome some of the weaknesses of other empirical approaches. For example, the standard spectral ratio was initially introduced by Borcherdt (1970) as the ratio of the horizontal components of the ground motion recorded at a specific site to those of a reference site. Thus, the

estimation of amplification by standard spectral ratios requires identifying the reference motion, which is generally considered the nearest outcropping bedrock. However, in some cases, near outcropping bedrock without large regional variations is not easy to find. On the other hand, horizontal-to-vertical spectral ratios (HVSRs) were initially introduced by Nakamura (1989) to estimate site amplification, and they have been addressed in several investigations (Field and Jacob, 1995; Bonilla et al., 1997; Parolai et al., 2004). However, it has been shown that HVSRs can underestimate the site amplification due to amplification of the vertical components of the ground motion (Haghshenas et al., 2008; Bindi et al., 2009; Ameri et al., 2011). More recently, there have been several attempts to improve estimations of HVSRs (Kawase et al., 2019b; Zhu et al., 2020). As a result, generalized inversions are greatly appreciated as an additional empirical method, whereby several limitations can also be avoided.

Apart from empirical estimations, numerical site transfer functions are also widely used to estimate site responses, especially those in one-dimensional (1D) assumptions (Bard and Gariel, 1986; Bardet et al., 2000; Kottke and Rathje, 2008). Numerical simulations generally use site-specific input information from geophysical, geological, and geotechnical investigations (e.g., S-wave velocity profiles, unit masses, layer thicknesses). V_S information used for numerical simulations is usually assumed as a set of homogenous V_S layers, which neglects all possible lateral or vertical small-scale heterogeneities. Such variations were recently addressed in several studies in terms of their potential impact on final site-response estimates (Shible et al., 2018; Teague et al., 2018). Hence, the accuracy of numerical simulations appears to be highly dependent on the accuracy of the information provided. From this perspective, GITs can also have an essential role in validating numerical site responses by providing empirical estimates that result simply from the decomposition of the Fourier spectra, with no prior dimensionality assumptions or site information uncertainty.

In general, results of generalized inversions of ground motion, including source and attenuation parameters, have been used for several purposes. Bora et al. (2015) used GITs to develop a Fourier amplitude spectrum GMPE that can be adjusted to account for source, path, and site conditions for the regions of interest. The adjusting parameters were stress drops ($\Delta\sigma$), a quality factor (Q), and the high-frequency attenuation parameter κ. Perron et al. (2017) focused on the estimation of κ for sites in the Provence area of southern France, and they indirectly used GIT results of Drouet et al. (2010) to compare with the Q values obtained from the path term of κ. Based on the GIT studies of Drouet et al. (2010) in France, Drouet and Cotton (2015) also performed stochastic simulations of ground motion, and they determined the fully stochastic GMPEs for southern France based on these synthetic data.

Overall, there are two main ways to perform generalized inversions. First, there are the ‘nonparametric’ approaches (Castro et al., 1990; Parolai et al., 2004; Bindi et al., 2006; Pacor et al., 2016) that describe a linear model with some constraints for the unresolved degrees of freedom of the problem. Then, the ‘parametric’ approaches can become highly nonlinear and require *a-priori* models for the source and attenuation terms (Hartzell, 1992; Salazar et al., 2007; Drouet et al., 2008; Edwards et al., 2008). These GIT approaches have become widely used, and several studies have developed different schemes with different assumptions and constraints through the years.

The choice of the most appropriate GIT scheme for an application might depend on the dataset geometry and characteristics. The most appropriate assumptions can vary for different dataset configurations and the aims of a study. However, these choices have not been deeply explored in previous studies. In addition, the resulting 'inter-method' uncertainty for the estimated physical parameters has not been addressed previously in the framework of detailed parameter estimation. These issues were found motivating enough for us to set up an international benchmark to compare the different present implementations of GITs, to investigate their advantages and drawbacks and the relative performance of one inversion technique with respect to others, as well as to explore the uncertainty for the estimated seismological parameters from different implementations of the inversion procedure.

GITEC benchmark (Generalized Inversions TEchniques Comparisons) was organized based on the idea of performing several inversions with multiple GIT schemes on different datasets. The benchmark addresses the possible differences between different approaches on the same dataset and between different datasets. The main scope of GITEC is to compare and improve the knowledge of the performances of different generalized spectral inversion methods and the underlying hypotheses they use to estimate ground-motion parameters.

This GITEC 'methodological benchmark' is expected to serve several goals:

- First, the presence of different implementations of inversion schemes (i.e., parametric, nonparametric) makes it important to investigate the pros and cons of each approach, depending on the dataset characteristics considered (e.g., dataset geometry, configuration, and others).
- Exploration of the dependence of the results on the assumptions and reference conditions or the strategies followed to solve the trade-offs.
- Addressing the estimation of epistemic uncertainties (inter-methods) on the inverted terms and parameters.

The work presented here summarizes the results and the main conclusions reached with the GITEC benchmark. Inversions were run by the six teams that participated in the benchmark. Initially, inversions were performed on a simple synthetic dataset as a sort of 'sanity check' for all inversion schemes and codes. The inversions were then performed on two real datasets available from previous studies, with some slight modifications and updates. The impact of the choice of the reference condition was explored by performing two phases of inversions (I, II), where the reference conditions were changed from phase I to phase II. In the first run, the teams' reference conditions were freely set while unified in the second run. The impact of the dataset size on the inversion performance was also explored in two steps. First, the two datasets considered were chosen to have different numbers of records. Then, in a separate test, the number of records in one of the datasets was decimated before running the inversions.

We do not pretend to present here the 'best practice' for GIT applications, but instead, we aim to provide a comparison of the different existing methods to address the question of the uncertainties in GITs. In the end, analysis of the results from different approaches and inversion schemes allowed characterization of the inter-method variability. We also highlight the relative robustness of the seismological terms and the relative suitability of the different approaches as functions of the dataset configuration.

3.2 About Generalized Inversion Techniques

3.2.1 Overview

Generalized inversion schemes are based on the principle that the far-field S-wave Fourier amplitude spectrum can be separated into three main components, as indicated in Eq (3.1). FAS_{ij} is the Fourier amplitude spectrum recorded at site i for event j, $E_j(f)$ is the source function, $A_{ij}(f)$ is the path contribution for event j recorded at site i, and $S_i(f)$ is the site-response term. Each of these terms is assumed independent from each other for a given recorded spectrum.

$$FAS_{ij}(f)=E_j(f).A_{ij}(r,f).S_i(f) \quad (3.1)$$

Applying a logarithm to Equation (3.1), we get the linear equation that provides a linear system of the form $A.x = b$, where b is the data vector, x is the solution of the system, and A describes the system matrix (Andrews, 1986; Castro et al., 1990). In the way that this system is defined, it has two undetermined degrees of freedom. The system can then be solved if two additional constraints are applied to one or several site responses. In GITs, this is often called the reference site condition.

The inversion scheme can be implemented following either a nonparametric or a parametric approach. The difference between these two approaches lies in the assumptions made *a priori*. In a nonparametric inversion, no models are assumed *a priori*, and the inversion is carried out at each frequency point to separate the three contributions described in Equation (3.1). Nonparametric inversion provides the source spectra $E(f)$ for each earthquake, the average attenuation curves $A(r,f)$ as functions of the distance at each frequency, and the site amplification $S(f)$ as a function of frequency for all of the sites. On the other hand, analytic models are assumed to account for the source and attenuation terms in the parametric approach, while the site term is kept nonparametric. Parametric inversions provide the direct output of estimates of the parameters for the models assumed. When performing nonparametric inversions, these parameters are retrieved by performing post-inversion parameterization for each term.

To run parametric inversion schemes, parametric models need to be assumed *a priori*. The model of Brune (1970) and Eshelby (1957) is usually used for the source term, as described in Eqs (3.2) and (3.3). M_0 is the seismic moment, f_c is the corner frequency of the earthquake, $\Delta\sigma$ is the stress drops, $R_{\theta\phi}$ is the source radiation pattern, which is assumed to be constant ($R_{\theta\varphi}$= 0.55 for S-waves; Boore and Boatwright 1984), ρ is the density in the crust, β is the S-wave velocity of the medium near the source, and v_s is the average S-wave velocity along the path. Generally, we assume that $\beta = v_s = 3.5\ km/s$ and $\rho = 2800\ kg/m^3$, as in several studies (e.g., Drouet et al., 2010, Bindi et

al., 2017). With M_0 obtained from inversions, the moment magnitude M_W can be deduced using the relation of Hanks and Kanamori, (1979), as in Eq(3.4).

$$E(f) = \left(\frac{2R_{\theta\phi}}{4\pi\rho\beta^3}\right)\left(\frac{(2\pi f)^2 . M_0}{1 + (\frac{f}{f_c})^2}\right) \tag{3.2}$$

$$\Delta\sigma = \frac{7}{16} M_0 \left(\frac{f_c}{0.37 v_s}\right)^3 \tag{3.3}$$

$$M_W = \frac{\log_{10}(M_0) - 9.1}{1.5} \tag{3.4}$$

The attenuation of seismic waves is acquired by waves propagating through the crust, so it is usually constrained to unity at the source. In the nonparametric approach, the attenuation is derived directly over the available distance bins in the data. When we have no data at short distances, the attenuation is fixed to 1 at a distance defined as the reference distance (R_{ref}). This assumption leads to shifted sources at the chosen R_{ref}. Thus, nonparametric source terms need to be rescaled by R_{ref} for correct interpretation. The path term assumed accounts for both anelastic attenuation and the geometrical spreading decay effects, as given by Eq(3.5). The frequency-dependent quality factor is described by Q_0 and the exponent α while γ is the coefficient of the geometrical spreading. It is worth mentioning that in a parametric approach R_{ref} is implicitly set to 1.

$$A(r, f) = \exp\left(-\frac{\pi(r - R_{ref})f}{Q_0 f^{\alpha} v_s}\right) . \left(\frac{R_{ref}}{r}\right)^{\gamma} \tag{3.5}$$

Here, an additional attenuation representation is proposed for the quality factor, Q. A more generalized evaluation of Q, which is mainly assumed as $Q_0 f^{\alpha}$ in Equation (3.5), can be provided for a given distance and frequency. We chose to evaluate the generalized Q at 50 km, where attenuation is not yet undergoing additional complexities, and where we can consider that γ = 1. Also, we chose the frequency to be in the middle of the frequency range considered in the dataset. Indeed, other choices are also possible for detailed interpretations, but we considered only one (r, f) couple, with the general notation $Q_r(f)$. The main consequences of such a parameterization are:

- Having an alternative to the actual Q-model parameters can show better how much the approaches are consistent in their attenuation.
- That the $Q_r(f)$ term reveals more information attenuation than the Q_0 value.

The $Q_r(f)$ term can be improved for better interpretation of the attenuation properties. However, it appears a good start to exclude possible trade-offs (i.e., between Q_0, α, and γ) from comparisons.

3.2.2 Methodologies and inversion schemes involved in the benchmark

The GITEC benchmark includes several inversion schemes, as listed in Table 3.1, and these can be classified into three main categories:

- Full nonparametric inversion schemes. These inversions can be carried out as:
    1. One-step inversion, including simultaneous inversion of all of the terms in Equation (6). The scheme developed by (Bindi et al., 2009; Oth et al., 2011) was run. In the following, these inversions have the reference #01N1.
    2. Two-step inversion, where the attenuation is solved first, then the source and site are solved from the attenuation-corrected spectra. Two 2-step nonparametric inversion schemes were run:
        a. the scheme developed by Klin et al. (2018). In the following, these inversions have the reference #02N2.
        b. the scheme developed by Castro et al. (1990). In the following, these inversions have the reference #03N2.
- Full parametric inversion schemes, divided into the following different inversion schemes:
    1. The scheme implemented by Drouet et al. (2010). In the following, these inversions have the reference #04P.
    2. The scheme implemented by Drouet et al. (2010), further re-developed and modified by Grendas et al., (2018). In the following, these inversions have the reference #06P.
    3. The scheme implemented by Edwards et al. (2008). In the following, these inversions have the reference #07P.

Semi-parametric (or partially nonparametric) inversion schemes, where only the attenuation model is parameterized with a Q model and a geometrical spreading factor, while the source and site terms are kept nonparametric. This scheme was implemented by Nakano et al. (2015). In the following, these inversions have the reference #05SP.

Table 3.1: The inversion schemes of the GITEC benchmark.

#	Inversion Method	Teams	Affiliation	Scheme	Code
1	Bindi et al., (2017) Oth et al., (2011)	Adrien Oth and Dino Bindi	ECGS, Luxembourg GFZ Potsdam, Germany	Non-parametric (1-step inversion)	#01N1
2	Klin et al., (2018)	Peter Klin	OGS, Trieste, Italy	Non-parametric (2-step inversion)	#02N2
3	Castro et al.,(1990)	Raul Castro	Ciscese, Mexico	Non-parametric (2-step inversion)	#03N2
4	Drouet et al., (2010)	Hussein Shible	CEA-Cadarache, France	parametric	#04P
5	Nakano et al., (2015)	Hiroshi Kawaze	Kyoto University, Japan	Semi-parametric	#05SP
6	Grendas et al., (2018)	Ioannis Grendas	ITSAK, Greece	parametric	#06P
7	Edwards et al., (2008)	Ben Edwards	University of Liverpool, England	parametric	#07P

3.3 The datasets considered

To investigate the performances of the different GIT schemes with respect to different datasets, we considered:

- A synthetic dataset generated as the forward problem solution according to Eq(3.1), (Figure 3.1.a-c). The dataset configuration is based on the geometry of the Existing Swiss network, with events in the magnitude range of 3-5.5 (Figure 3.1.j-i). Generated signals are issued from 50 stations and 100 events. Hypocentral distances are in the range of 0 to 200 km.
- A sparse regional dataset: the RAP dataset in the French Alps region, from Drouet et al. (2008, 2010). The final dataset (Figure 3.1.d-e) consists of 72 earthquakes in the Alps area (recorded between 1998 and 2006) with hypocentral distances reach up to 250 km. Hypocentral distances of recorded events come from the French national network agency (RéNaSS) and local magnitudes from RéNaSS and LDG (another French national agency). Focal depths are ranging between few kilometers and 10 km. We would also note that other existing and updated databases (e.g., Traversa et al., 2020) were not considered in the aim to test for GIT performance on data-poor and less constrained datasets.

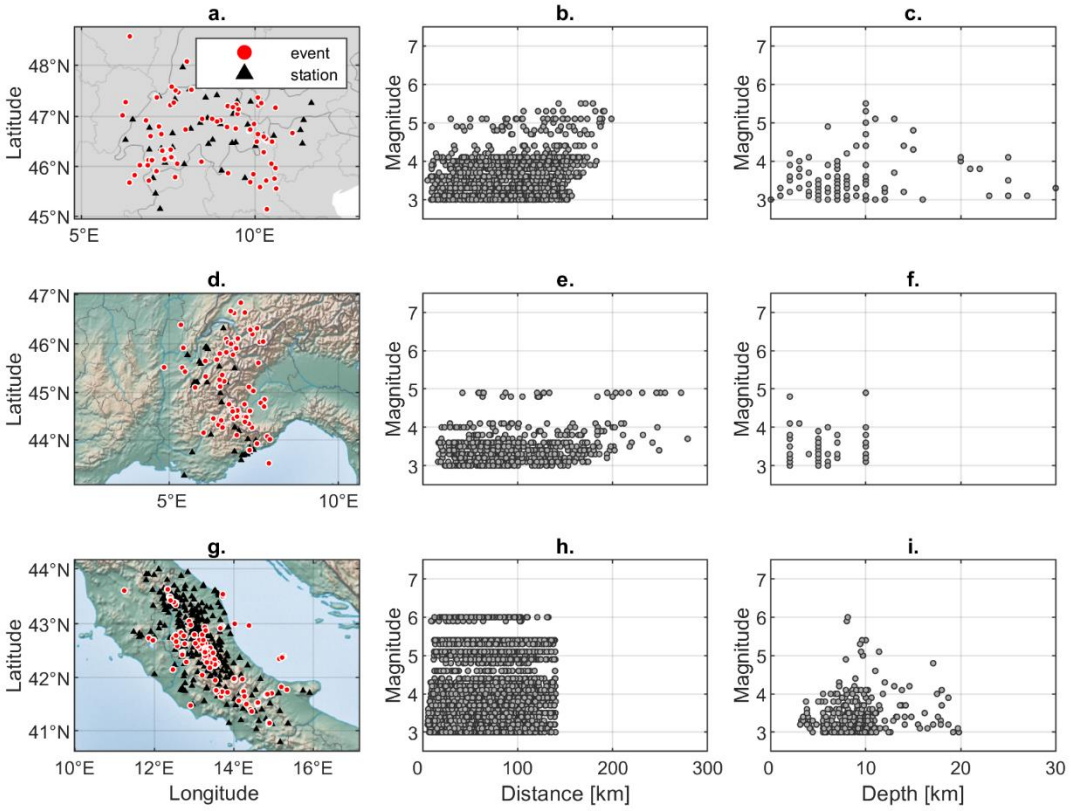

Figure 3.1 (a, d, g) Maps showing the synthetic dataset and the earthquakes and stations for France and central Italy. (b, e, h) Magnitudes and hypocentral distances for the respective datasets. (c, f, i) Depth distribution of the respective sets of events, with magnitudes. Note: Local magnitudes M_L were considered for the France and Italy data.

- A very dense regional dataset from Central Italy, data come from (Bindi et al., 2017) and (Pacor et al., 2016) with some updates and extensions. The

considered dataset (Figure 3.1.g-i) consists of 231 earthquakes recorded by 309 stations which include the 2009 L'Aquila sequence (M_w= 6.1) and spans the period between July 2008 to January 2017. In this dataset the local magnitudes vary in the range 3.0-6.1 mainly concentrated within 3.0-4.5 range, and the hypocentral distances reach up to 140 km. The depth of earthquakes is mainly distributed within 5 to 10 km range.

As the primary objective of GITEC was to perform consistent comparisons of results provided by the different approaches and schemes, it was essential to provide synthetic datasets that serve as a control for inversion results and the associated variability. The advantage of using a synthetic dataset is that the inversion is performed over data obtained from the forward problem, so the underlying seismological parameters are known *a priori*.

In the initial steps of the benchmark, a Japanese dataset was also considered. However, we focus only on the previous three datasets here.

3.4 Generalized inversions on synthetic and real data

3.4.1 Sanity check using a synthetic dataset

Performing a sanity check with a simple synthetic dataset helped to avoid all kinds of ambiguities in the results that might make the interpretations difficult. The synthetic data were generated using the Brune model for the sources, with a homogenous attenuation model for the earth crust with a quality factor Q = 600 (frequency-independent). The reference (nonamplifying) site was also provided as input to the teams.

The results provided by the different inversion schemes led to the correct parameter values, as highlighted by the $M_{w,GIT}$-M_L plots in Figure 3.2. Comparing the inverted corner frequencies within the range [0.3-10 Hz] with the inverted M_w showed that the stress drops obtained varied within the range of 35 bars to 60 bars, which is considered acceptable knowing that the correct values of the stress drops used in the generation of the synthetic dataset were 50 bars.

The attenuation values obtained by the different teams are considered acceptable compared to the values used in the synthetics (Table 3.2), except for inversion #03N3, which showed a significant difference from the true values. This difference gave a first indication that we might face strong trade-offs when dealing with attenuation parameters. On the other hand, the inverted site terms were consistent with the true site amplification. Finally, this test on such a simple dataset appeared beneficial and allowed some code adjustments that were helpful in the continuation of the exercise.

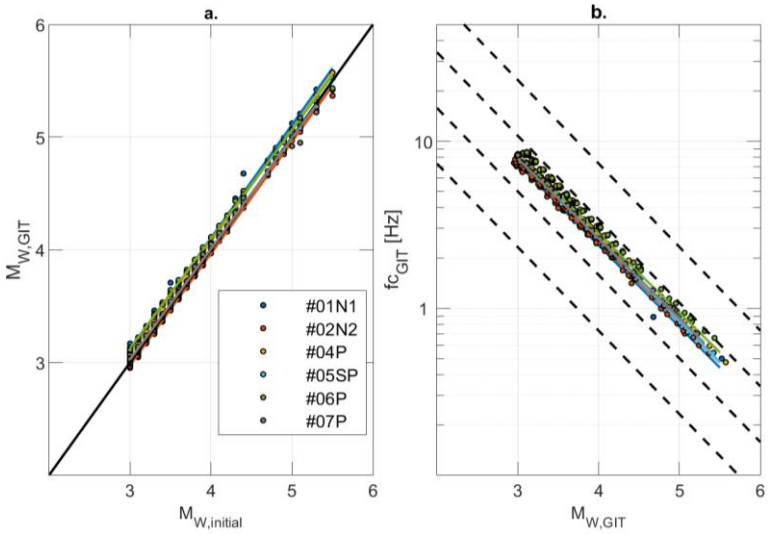

Figure 3.2: Synthetic dataset for source results. (a) The 1:1 plot of the M_w obtained from the inversions ($M_{w,GIT}$) with respect to the input magnitudes (M_L). (b) Corner frequencies versus inverted M_w showing the stress drop distributions that resulted from the inversions. Dashed lines show stress drop levels of 1, 10, and 100 bars.

Table 3.2: Mean estimations of the attenuation parameters from the different inversions.

inversion	$\mathbf{Q_o}$	$\boldsymbol{\alpha}$	$\boldsymbol{\gamma}$
Synthetic	600,00	0,00	1,00
#01N1	533,00	0,06	1,02
#02N2	517.64	0.07	0,99
#04P	555,87	0,03	0,99
#05SP	539,08	0,04	1,00
#06P	504,62	0,06	0,97
#07P	514,00	0,05	0,98

3.4.2 The two phases and the reference conditions

First, generalized inversions on the real datasets were carried out without any *a-priori* common constraints or assumptions; i.e., each team performed the inversions independently. This starting phase (phase I) was important in exploring the different choices to solve the inverse problem and eventually exploring the resulting variability. Indeed, constraints can vary depending on the scheme followed. However, the common constraint types between the exploited approaches were mainly the choice of R_{ref} and the reference site condition.

The reference distances and the sites for attenuation and amplification were chosen independently in phase I, as reported in Table 3.3. For R_{ref}, its definition was essential in the nonparametric approach, while it was indirectly set to 1 in the parametric and semiparametric approaches. The reference site choices were based either on trial inversions to identify flat responses or previous studies from the literature that used similar datasets. The aim behind phase I was to observe the global variability of the results due to both different choices and inversion implementations.

In phase II of the inversions on the same real datasets, reference conditions were unified as much as possible. R_{ref} was set to 20 km and 10 km in the nonparametric inversions for each of the France and Italy datasets. The correct choice of reference site generally requires *a-priori* knowledge about the site conditions. Therefore, the

inversions were constrained by selecting reference sites from those that were well characterized (i.e., with measured velocity profiles available).

Table 3.3: Reference sites and distances considered in the France and Italy datasets for phase I, the free phase.

Inversion	R_{ref}***(km)***	***Reference sites***
France		
#01N1	20	OGLE, ISOL
#02N2	16	CALF, ISOL, NBOR, OGAN, OGCH, OGFB, OGGM, OGLE, OGMU, SAOF, STET
#03N2	1	BELV, CALF, NBOR, OGAG, OGAN, OGCH, OGDI, OGGM, OGMO, OGMU, OGSI, OGTB, SAOF
#04P	1	CALF, ISOL, NBOR, OGAN, OGCH, OGFB, OGGM, OGLE, OGMU, SAOF, STET
#06P	1	CALF, ESCA, ISOL, MENA, NBOR, OGAG, OGAN, OGCA, OGCH, OGDI, OGGM, OGLE, OGMA, OGMB, OGMO, OGMU, OGSI ,OGTI ,SAOF ,STET
#07P	1	none
Central Italy		
#01N1	20	MNS, RM06, RM08
#02N2	20	CLN, LSS
#04P	1	CLN, LSS
#05SP	1	CLN, LSS
#06P	1	All sites with 'rock' classification following EC8 classes provided with the dataset, based on V_{S30} inferred or measured values.
#07P	1	none

First, the selected sites were associated with rock geological conditions, i.e., V_{S30} >1000 m/s, with at least 20 records. For the French stations, the V_{S30} was obtained from the characterization campaign data of Hollender et al. (2018). Then, 1DSH numerical simulations were carried out to estimate the theoretical transfer functions using the V_S profiles provided. The 1D reflectivity model (Kennett, 1974) was used to derive the responses of the horizontally stratified layers excited by a vertically incident SH plane wave (original software written by Gariel and Bard, and used previously in a large number of investigations: e.g., Bard and Gariel, 1986; Cadet et al., 2012; Laurendeau et al., 2018). We aimed to select the stations with the flattest site responses from the 1DSH simulations as the reference sites in the inversions. This procedure led to the choice of stations OGCH and LSS as the reference stations for the France and Italy datasets, respectively. The theoretical site responses that were fixed in the inversions are shown in Figure 3.3.

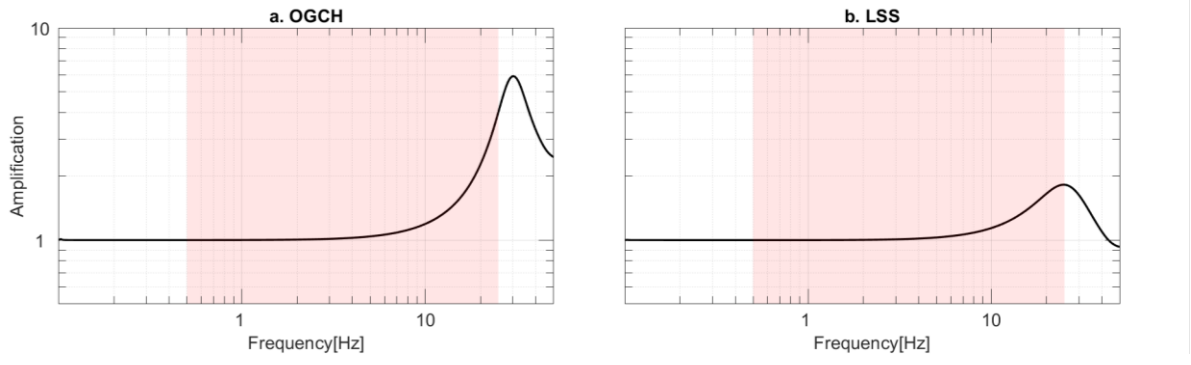

Figure 3.3: Amplification functions for each of the chosen reference stations (i.e., OGCH, LSS), as estimated from the 1DSH numerical simulation based on the available V_S profiles. The bandwidth used is shaded in red.

Inter-method comparisons were performed later and were not affected by the different reference choices. This step was ensured by fixing the same reference conditions by all of the participating teams in phase II. In this section, detailed analyses are conducted before the evaluation of the inter-method variability. Individual source-by-source or site-by-site comparisons are avoided in this text. Only the global statistical analyses are addressed here for simplicity and clarity.

3.4.3 Results for the sparse regional dataset, the French Alps

Comparisons of the results from the different inversion schemes are provided here for the France dataset. As shown in Figure 3.4, the variability in the attenuation appeared low when limiting the comparison within each type of approach (i.e., nonparametric, parametric). However, more significant discrepancies were observed in the attenuation curves at large distances when comparing each type of approach. The wide range of distances in the dataset probably increased the complexity for the attenuation model for distances >100 km. These complexities might be better captured with nonparametric attenuation functions that are derived directly from the data. For example, between 50 km and 200 km, the low-frequency attenuation showed discrepancies up to a factor of 4 between the parametric and nonparametric approaches. At high frequencies, these discrepancies can increase to a factor of around 10 for large distance ranges.

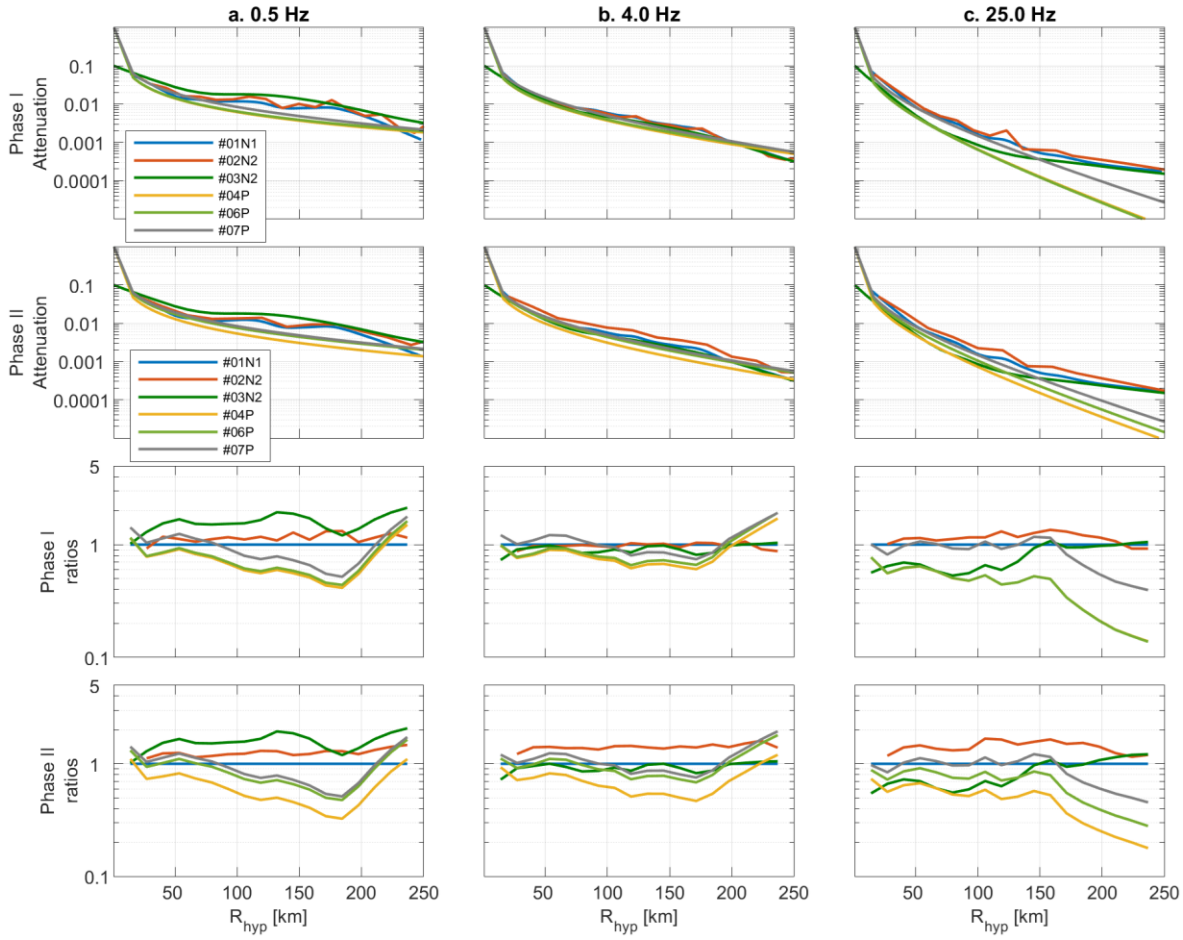

Figure 3.4: France dataset results from phases I and II. Top: Attenuation curves from the different approaches for the three frequencies (i.e., 0.5, 4, 25 Hz), plotted as functions of the hypocentral distance. Bottom: Ratios of the attenuation curves with respect to #01N1.

Table 3.4 summarizes the attenuation parameters obtained by the inversion schemes. The parameters provided by the different participating teams show some variability in both inversion phases: the quality factor Q_0 in the range 55-350, frequency dependence α in the range 0.3-1.1 and geometrical spreading γ in the range 0.7-1.2. For some inversions, attenuation was not affected by reference choices as in #03N2 and #07P, mainly since the reference choice doesn't affect the attenuation terms in these inversions. For very similar results of non-parametric attenuation, like that of #01N1 and #02N2 in the first phase, the values for Q_0 seems different which highlights a strong trade-off between these parameters. The other representation of Q, i.e. $Q_r(f)$ estimated at 50 km and 4 Hz ($Q_{50km}(4Hz)$), shows more consistency within the same approach (either parametric or non-parametric). The latter representation reflects more what is observed with the non-parametric attenuation curves in Figure 3.4. Also, these estimations were not much affected between the two phases of inversions.

Table 3.4: Summary of the mean estimations of the attenuation parameters from the different inversions for phases I and II.

Inversion	Q_0		$Q_{50km}(4Hz)$		α		γ	
	Phase I	**Phase II**	**Phase I**	**Phase II**	**Phase I**	**Phase II**	**Phase I**	**Phase II**
#01N1	277.64	80.82	38.73	38.55	0.51	1.02	1.26	0.76
#02N2	55.28	68.73	40.16	41.92	1.12	1.12	0.58	0.87
#03N2	76.75	76.75	45.97	45.97	1.03	1.03	0.70	0.70
#04P	239.68	341.73	118.68	110.12	0.44	0.31	1.08	1.07
#06P	236.68	389.33	121.43	136.99	0.45	0.27	1.01	1.07
#07P	203.00	203.00	151.32	151.32	0.52	0.52	0.97	0.97

Figure 3.5 shows the source term results from the two phases of the inversions. Acceleration source spectra are compared between the nonparametric approaches with the fitted Brune model from the parametric inversions. The discrepancies reach a factor of 10 for the mean source spectra (for all of the events) obtained from each inversion scheme. In addition, variable slopes are clearly seen between the inversions at high frequencies, with variation across negative and positive values.

Figure 3.6 shows the mean site responses from the different inversions of phases I and II. Differences can be clearly seen for both phases, with a factor of around 3 at the lower frequencies that increases to 10 times or higher at high frequencies. Thus, unifying the reference conditions provides more comparable mean amplifications (especially for the nonparametric approach) at the lower frequencies, but there was no decrease in the variability at the higher frequencies.

If we compare both sources and sites, it can be seen that the approach that results in the higher high-frequency slope on the sources corresponds to the lower slope on the sites. After conducting phase I only, these slope observations were attributed to the reference site's different choices, which was one of the main motivations for carrying

out phase II of the iterations. However, the high-frequency variability remained significant and was not reduced in phase II. Nevertheless, these data show that there is generally an impact of the choice of the reference condition on the source spectra obtained, so this choice represents an important step in the inversions.

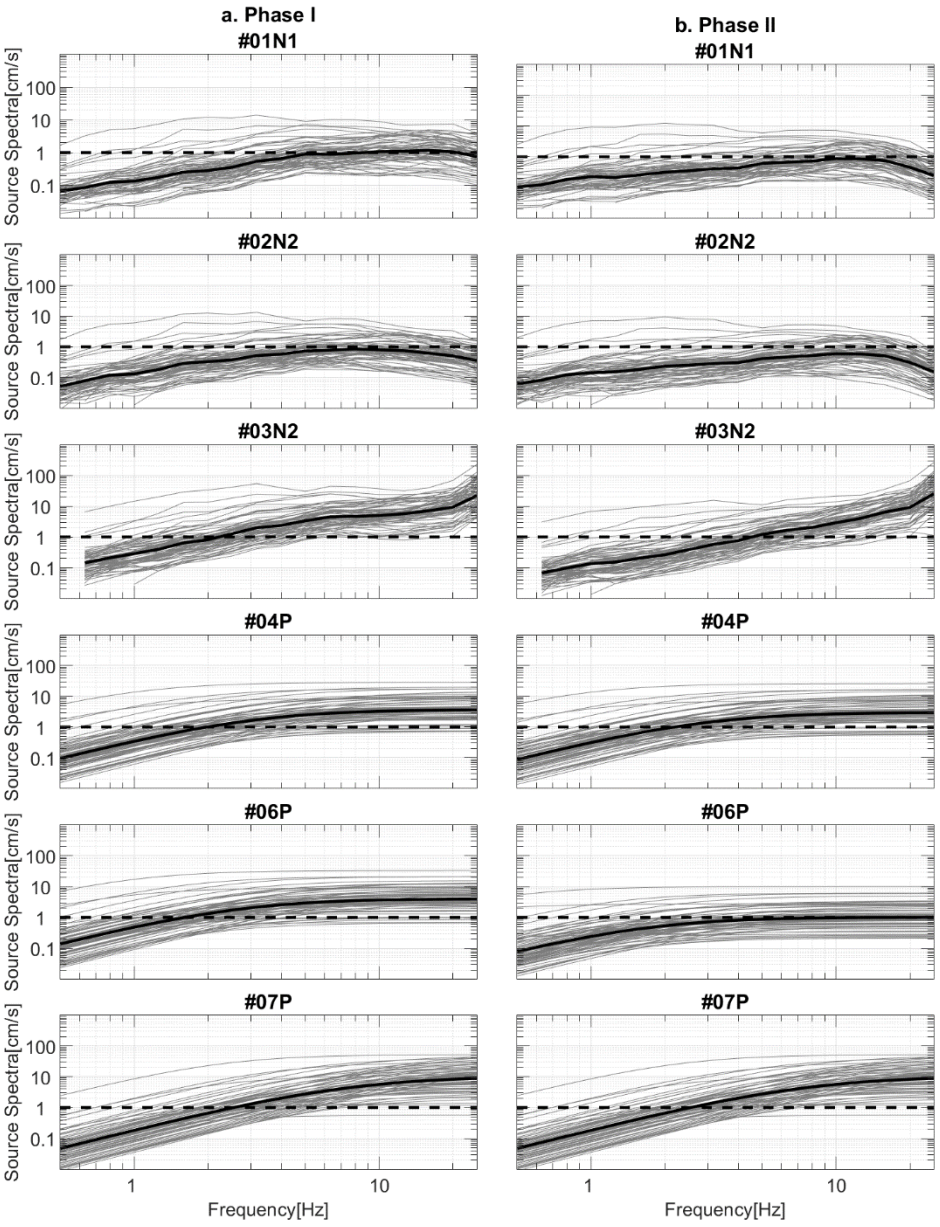

Figure 3.5: France dataset results from phases I and II. Overall comparisons of the source spectra obtained from the different inversions (as indicated), and showing the means for all of the spectra (bold black).

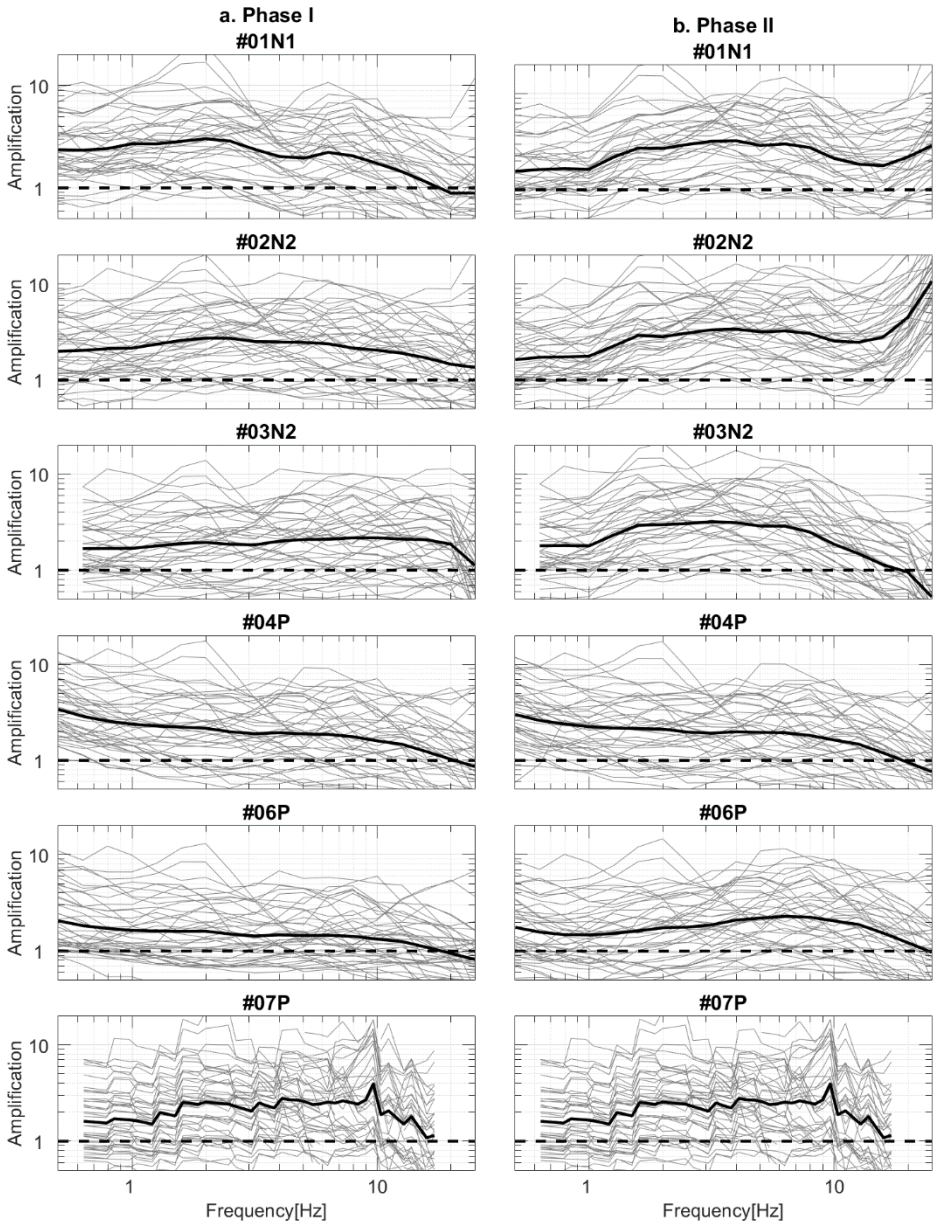

Figure 3.6: France dataset results from phases I and II. Overall comparisons of the site amplifications obtained from the different inversions (as indicated) and showing the means for all of the sites (bold black).

Next, we explored the impact of these differences seen for the nonparametric curves on estimating the physical parameters. A post-inversion parameterization was performed to estimate M_w, f_c, and $\Delta\sigma$. To conduct consistent comparisons with the parametric inversions, the same models assumed in these schemes were used to post-fit the nonparametric source spectra (Eq(3.2)). Figure 3.7 (a) shows the M_w obtained from the inversions *versus* the M_L of the original dataset. Limited dispersion can be seen among the results for both phases I and II, which is not the case for the other source parameters.

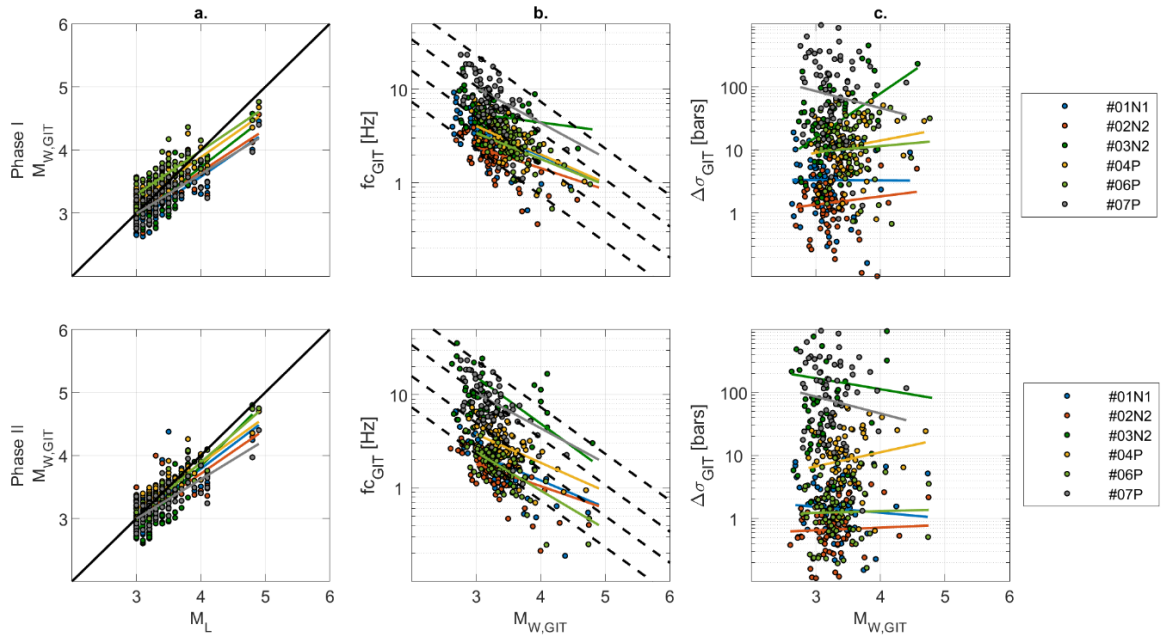

Figure 3.7: France dataset source results. (a) Plot (1:1) of the M_w from the inversions ($M_{w,GIT}$) with respect to M_L values provided. (b) Distribution of the corner frequencies f_c versus $M_{w,GIT}$ estimated from the inversions. Dashed lines, stress drop lines of 1, 10, and 100 bars. (c) Stress drop distributions from the GIT approaches versus the $M_{w,GIT}$.

Figure 3.7b, c shows the $f_{c,GIT}$-$M_{w,GIT}$ and $\Delta\sigma_{GIT}$-$M_{w,GIT}$ distributions, respectively. Significant discrepancies can be seen for the mean $\Delta\sigma$ between phases I and II. This reflects important variability, despite the unifying of the reference sites in phase II. The estimated $\Delta\sigma$ provided by approaches #01N1, #02N2, and #06P lie in the range of 0.1 bars to 10 bars, while the approaches #04P and #07P result in $\Delta\sigma$ values in the range of 1 bar to 1000 bars, with means of around 10 bars and 100 bars, respectively. These data highlight that the inter-method variability for these source parameters is significant even when the reference site conditions are fixed (i.e., in phase II).

3.4.4 Results for a dense regional dataset, the Central Italy dataset

Following the France dataset inversions, the results from the Italy dataset are presented in this section. At first glance, the path terms appear to show good agreement between the different approaches. However, when the ratios were computed with respect to one of the specific teams (#01N1), the inter-method differences reached a factor of 4, which was relatively large (Figure 3.8). In phase I, the nonparametric approaches appeared to have close estimations at all frequencies and distances (#01N1, #02N2), while the parametric approaches showed relatively variable curves. In phase II, the attenuation curves from all of the approaches were variable despite the unified reference station. Thus, significant variability affects these attenuation terms even though the distance range was limited to 140 km.

The data presented in Figure 3.9 show the source spectra obtained from the different schemes. For phase I, the data showed similar shapes of the source spectra, with slight differences in amplitude at high frequencies. However, for phase II, the high-frequency slopes appeared variable, with a tendency to decay beyond 10 Hz for the nonparametric approach of source spectra. Within each type of inversion, consistent data were obtained. With the parametric source spectra being fixed to the Brune model,

the high-frequency slopes for the source terms cannot exist (i.e., #07P, #04P, #06P inversion schemes).

Table 3.5 shows the results for the attenuation parameterizations, where greater variability is seen for the Q_o values as well as for α and γ. For instance, the Q_o values range from 45 to 500, along with variations in α and γ, which shows the significant trade-offs among these parameters. This might be due to trade-offs induced by the simplicity of the attenuation model used, which excludes any possible slope changes with distance (i.e., geometrical spreading effects) or by neglecting the possible regional variations. For the $Q_{50km}(4Hz)$ representation, the values are more consistent (between different inversions) than seen for the Q_o values, with also little changes between phases I and II. Although the variability is reduced with the $Q_{50km}(4Hz)$ representation, it is still present and reflects the discrepancies observed in Figure 3.8.

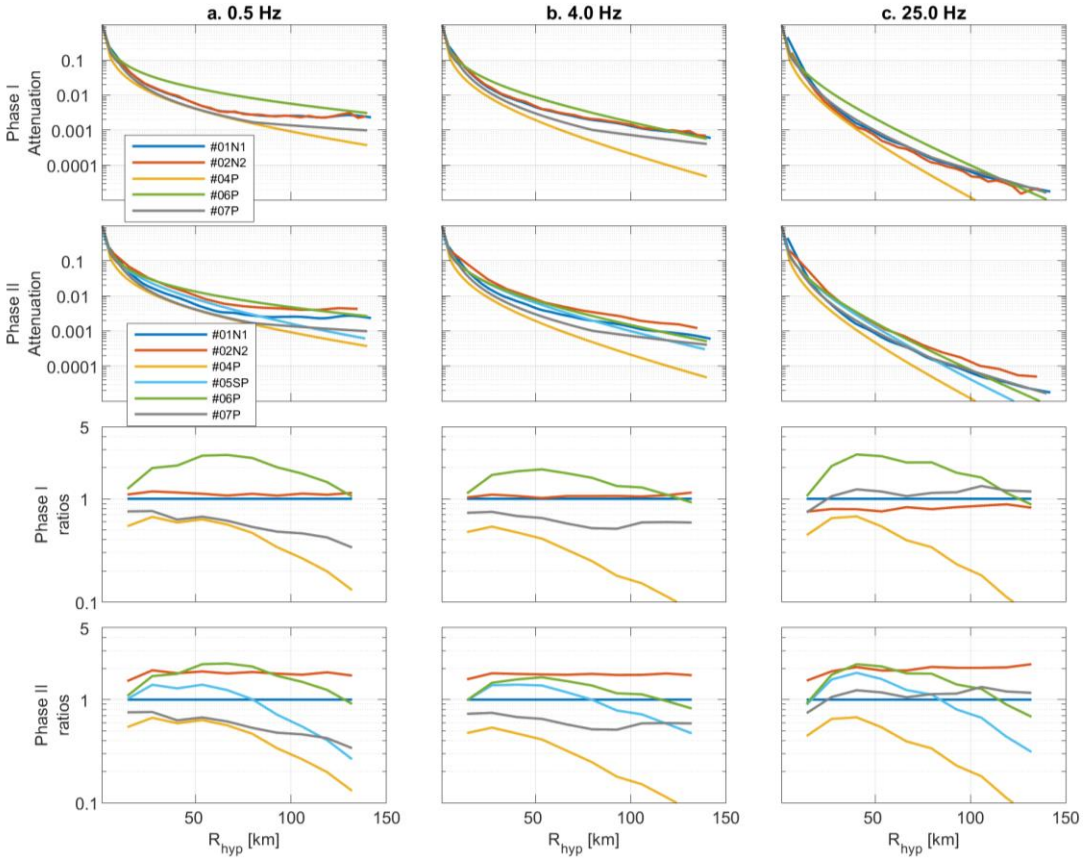

Figure 3.8: Central Italy dataset results from phases I and II. Top: Attenuation curves from the different approaches for the three frequencies (i.e., 0.5, 4, 25 Hz), plotted as functions of the hypocentral distance. Bottom: Ratios of the attenuation curves with respect to #01N1.

Table 3.5: This table summarizes the mean estimations of the attenuation parameters from different inversions

Inversion	Q_o		$Q_{50km}(4Hz)$		α		γ	
	PhaseI	PhaseII	PhaseI	PhaseII	PhaseI	PhaseII	PhaseI	PhaseII
#01N1	513.84	358.04	35.81	35.81	0.22	0.40	2.16	2.23
#02N2	150.86	90.95	40.18	49.10	0.64	0.70	1.88	1.37
#04P	56.13	56.13	51.50	51.50	0.59	0.59	1.26	1.26
#05SP	-	45.60	-	82.23	-	0.74	-	1.01
#06P	88.78	98.98	98.87	90.38	0.49	0.46	0.97	1.03

#07P	122.87	122.87	60.07	60.07	0.54	0.54	1.61	1.61

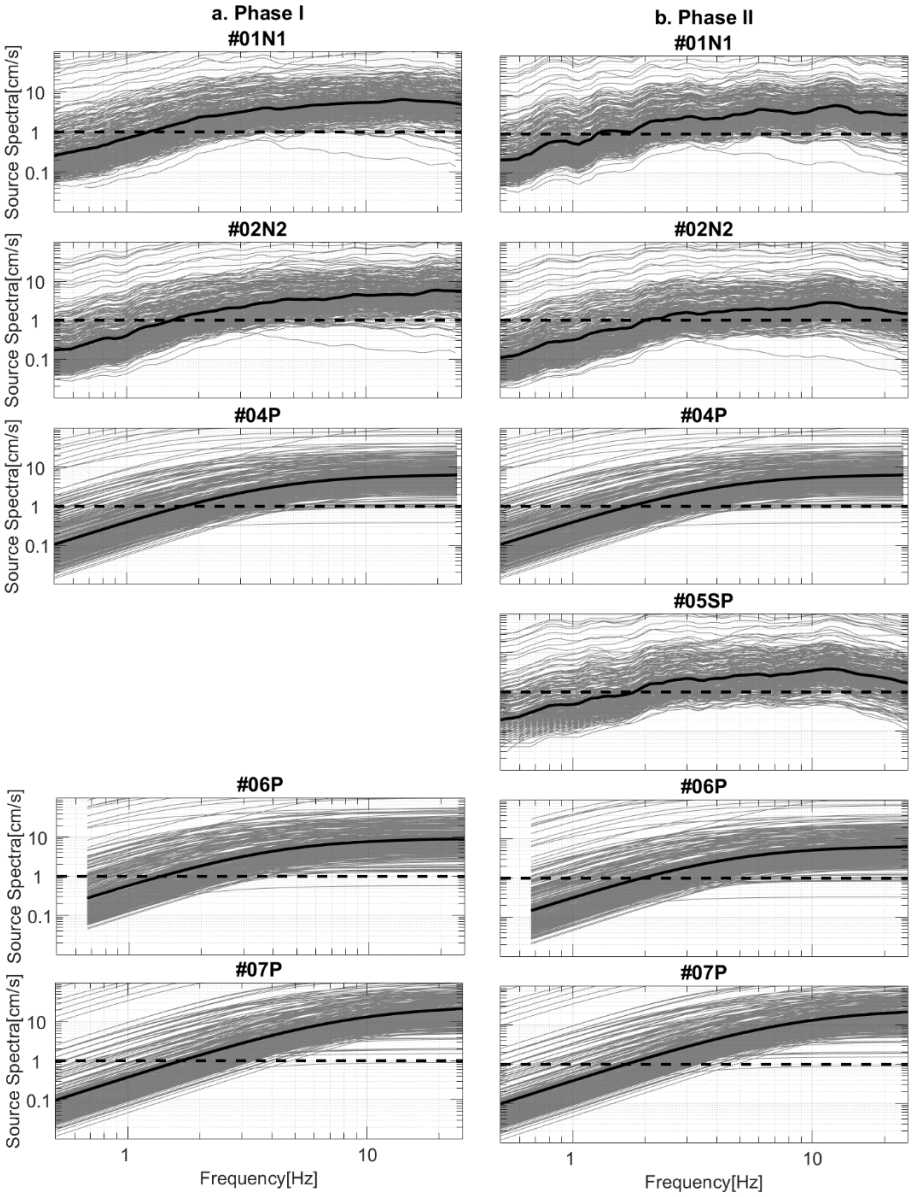

Figure 3.9: Italy dataset results for phases I and II. Overall comparisons of the source spectra obtained from the different inversions (as indicated), and showing the means for all of the spectra (bold black).

Figure 3.10 shows the overall views of the site term estimations provided by the different teams. The main observations for the Italian site amplifications are the differences in the high frequencies. Phases I and II show that the inversions can give different responses to the same reference site at high frequencies. For example, both methods showed slope changes, as for #01N1 and #02N2, while other inversions, such as #04P, #06P, and #07P, were not significantly affected. This appears to be partially counterbalanced by the larger high-frequency decay in the source spectra obtained by these inversion schemes, highlighting a potential trade-off between these two terms.

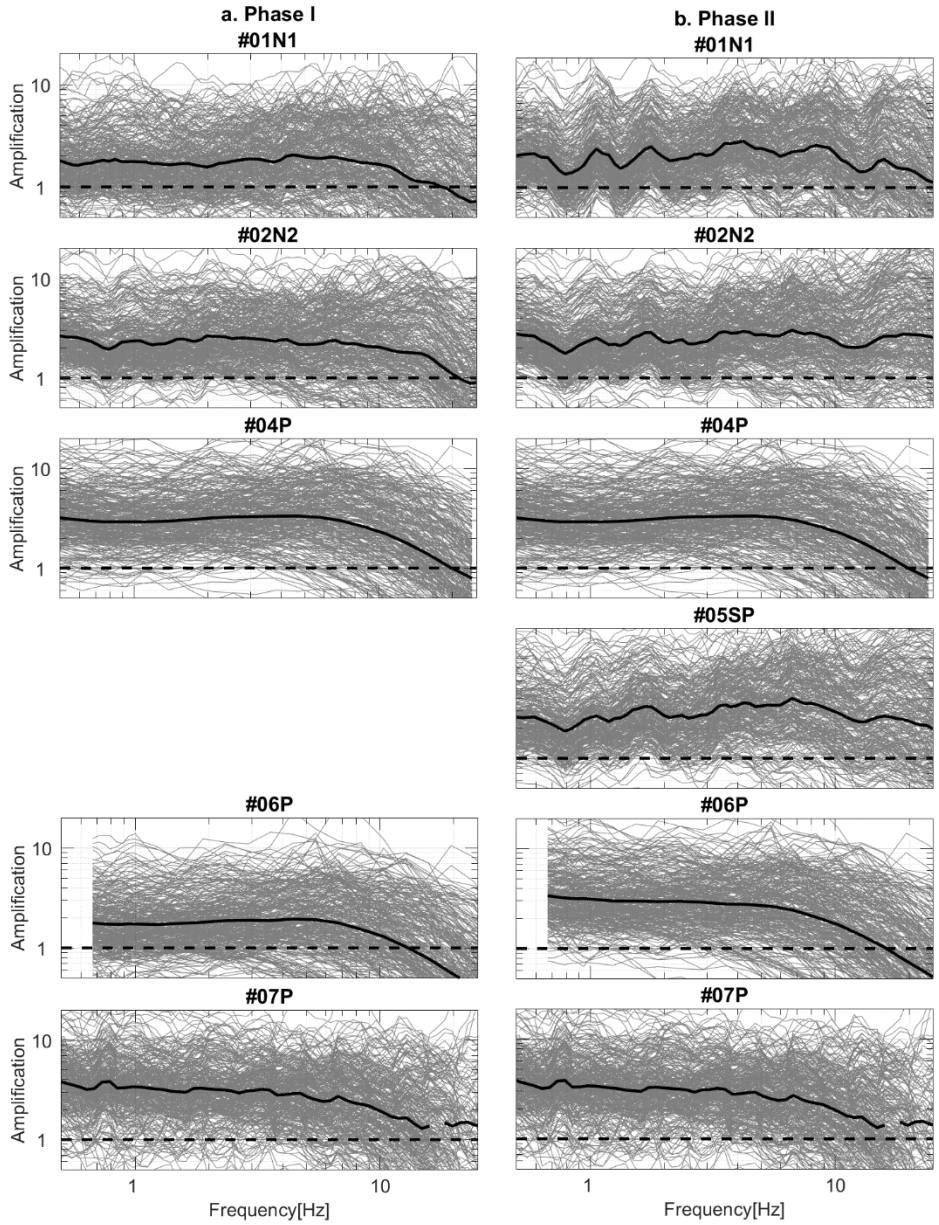

Figure 3.10: Central Italy dataset results from phases I and II. Overall comparisons of the site amplifications obtained from the different inversions (as indicated), and showing the means for all of the sites (bold black).

Although the semiparametric approach (#05SP) was applied in phase II only, it showed a decay on the source spectra. This suggests that high-frequency decay of the source spectra cannot be avoided by estimating the attenuation parameters first in a two-step inversion to avoid the source-site trade-offs.

Comparisons of source parameters obtained directly from the parametric inversions or post-inversion fitting of the nonparametric schemes are shown in Figure 3.11. The M_w-M_L plots show some consistency, where the GIT data overestimate low M_L and underestimate high M_L. This consistency was maintained between phases I and II. The stress drops $\Delta\sigma$ again showed significant variability in both phases I and II, along with the corner frequencies, $f_{c,GIT}$. Generally, the parametric schemes resulted in higher $\Delta\sigma$ than the nonparametric schemes, although within-type variability was also significant.

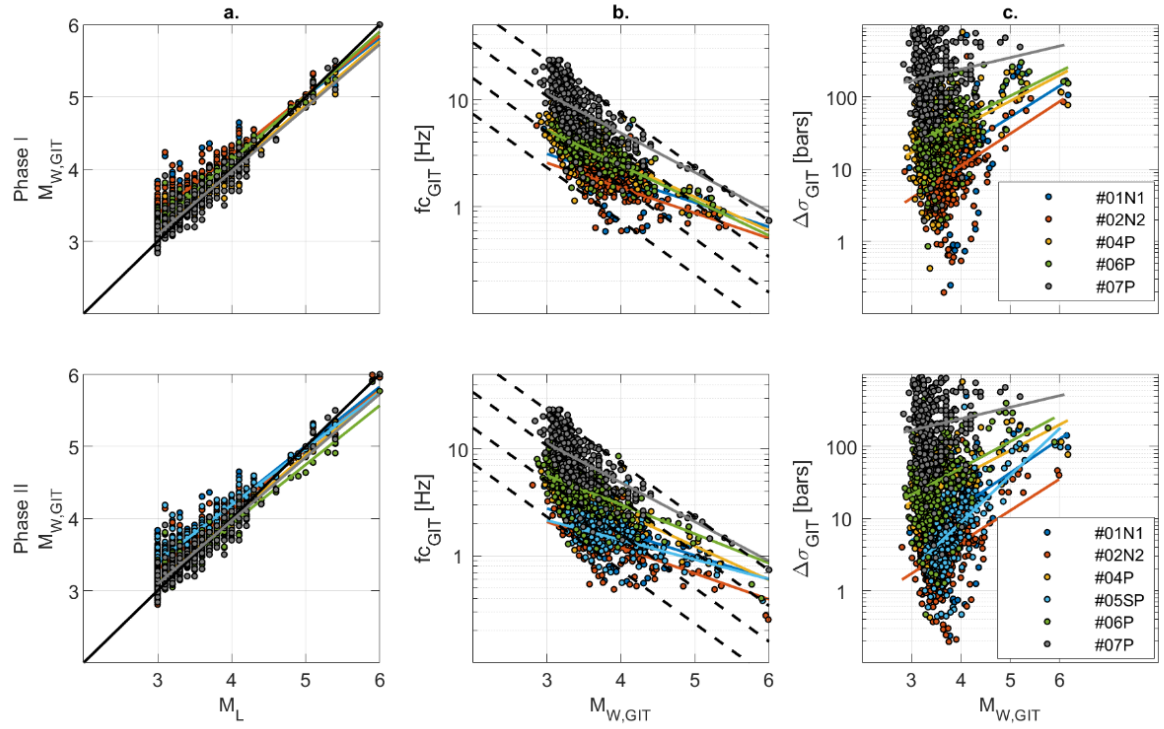

Figure 3.11: Central Italy dataset source results. (a) Plot (1:1) of the M_w obtained from the inversions ($M_{w,GIT}$) with respect to the M_L values provided. (b) Distribution of the corner frequencies f_c versus $M_{w,GIT}$ estimated from the inversions. Dashed lines, stress drop lines of 1, 10, 100, and 1000 bars. (c) Stress drops distributions from the GIT approaches versus the $M_{w,GIT}$.

3.5 Uncertainties associated with GIT results

One of the main objectives of the GITEC benchmark was not only to investigate but also to quantify the overall uncertainties associated with the GIT data. Generally speaking, uncertainties can result from either uncertainty in the estimates of the seismological parameters provided by a given inversion scheme (i.e., intra-method uncertainties) or epistemic uncertainties related to the choice of the inversion scheme (i.e., inter-method uncertainties) and to the characteristics of the dataset on which the GIT is based. The intra-method uncertainties can be described by the standard deviations associated with the results of the selected inversion procedure. However, the inter-method uncertainties can be captured by exploring the variability of the data provided by the different inversion approaches for the different dataset configurations. Here, we explored the inter-method variability, focusing only on phase II of the iterations for simplicity. Finally, a representative short summary of results is provided for both phases.

3.5.1 Strategy for inter-method uncertainties characterization

Given that the parametric and nonparametric inversion approaches do not provide the same results as the direct output (i.e., parametric inversions provide direct estimates of the seismological parameters, while nonparametric inversions provide spectral curves of the source, attenuation, and site terms), the comparisons of the variability of the results are best carried out in two steps.

The inter-method uncertainties were first quantified for the nonparametric spectra of the source, attenuation, and site terms. As the parametric GIT provides only parameters, the source spectra and attenuation curves were reconstructed from their models. The second step was to quantify the variability for the estimations of the physical parameters. For the nonparametric approaches, post-inversion parameterizations were considered (using the hypotheses and models consistent with those adopted by the parametric inversion schemes and previously presented in the text). In the end, the full (i.e., parametric and nonparametric) inter-method variability can be analyzed for each dataset considered. To determine a sort of minimum level variability, the quantification was performed over the synthetic dataset. Then the variabilities associated with the real dataset inversions are presented.

It is generally known that when uncertainties are addressed, large numbers of observations or estimations are needed to obtain meaningful statistics. However, for this benchmark, only a total of six inversion schemes were applied to each dataset. On this basis, other forms of variability indicators are defined and adopted. First, the standard deviations are calculated in the natural logarithmic scale, *std*($\log_{10}$), for the nonparametric spectra, termed spectral variability. Then, the parametric variability is presented by the coefficient of variation (*COV*) on the estimated parameters. As an alternative, (max-min)/2 is introduced instead of standard deviations and the ratio of (max-min)/(max+min) for *COV*, where max and min are the maximum and minimum estimations of the variable considered. This is indicated as the *maxmin* estimation in both cases.

3.5.2 Quantification of uncertainties

3.5.2.1 Spectral variability (std and maxmin of $\log_{10}$)

For each event, the nonparametric source spectra, attenuation, and site (from the different inversions) are used to estimate both the standard deviations and *maxmin* (in $\log_{10}$). Then, the estimations are presented for each source and site as functions of the frequency, while they are presented as functions of distance and frequency for the attenuation curves, as shown in Figure 3.12. The color scale used in Figure 12 corresponds to the values of the variability estimate (i.e. *std*($\log_{10}$) or (max-min)/(max+min)).

If the panels of Figure 3.12 are compared horizontally, it can be seen that the site responses are less prone to variability for both representations (i.e., *std*, *maxmin*). Also, the synthetic data show an expected minimal variability between the methods, with a tendency for lower variability for the site terms. However, the source spectra of the synthetic data show minimal variability of 0.2 (*std*) and 0.3 (*maxmin*), while the variability of the site responses is around 0.1 (for *std* and *maxmin*). Despite the differences in the mean site responses, the variability of the site responses for the France and Italy datasets is much more limited than for the source terms. The variability of the site terms starts to increase significantly beyond 10 Hz (around 0.4).

If the panels in Figure 3.12 are compared vertically, the same term variability for the different datasets can be considered. For example, the sources for the France dataset show high variability that starts at 0.3 std at low frequencies and exceeds 1 std at high

frequencies. The Italy dataset shows more consistency in terms of lower variability for the source spectra, which does not exceed 0.2 std before 8 Hz, and reaches 0.5 std afterward. The attenuation terms show similar variabilities that increase beyond 100 km in both of the real datasets to reach 0.4 std or 0.5 std. As for the site responses, similar robustness is achieved up to 10 Hz for both datasets.

The reduced variability in central Italy is probably due to the good geometry and density of the network, which allowed the recording of each event tens or even hundreds of times. After comparisons between the results from the France and Italy datasets, it appears that the redundancy in the datasets (in terms of the number of recordings, the data sampling for the distance and magnitude ranges) is essential and helpful to reduce the variability associated with generalized inversions. In other words, the more data there are from the region under study, the lower the variability.

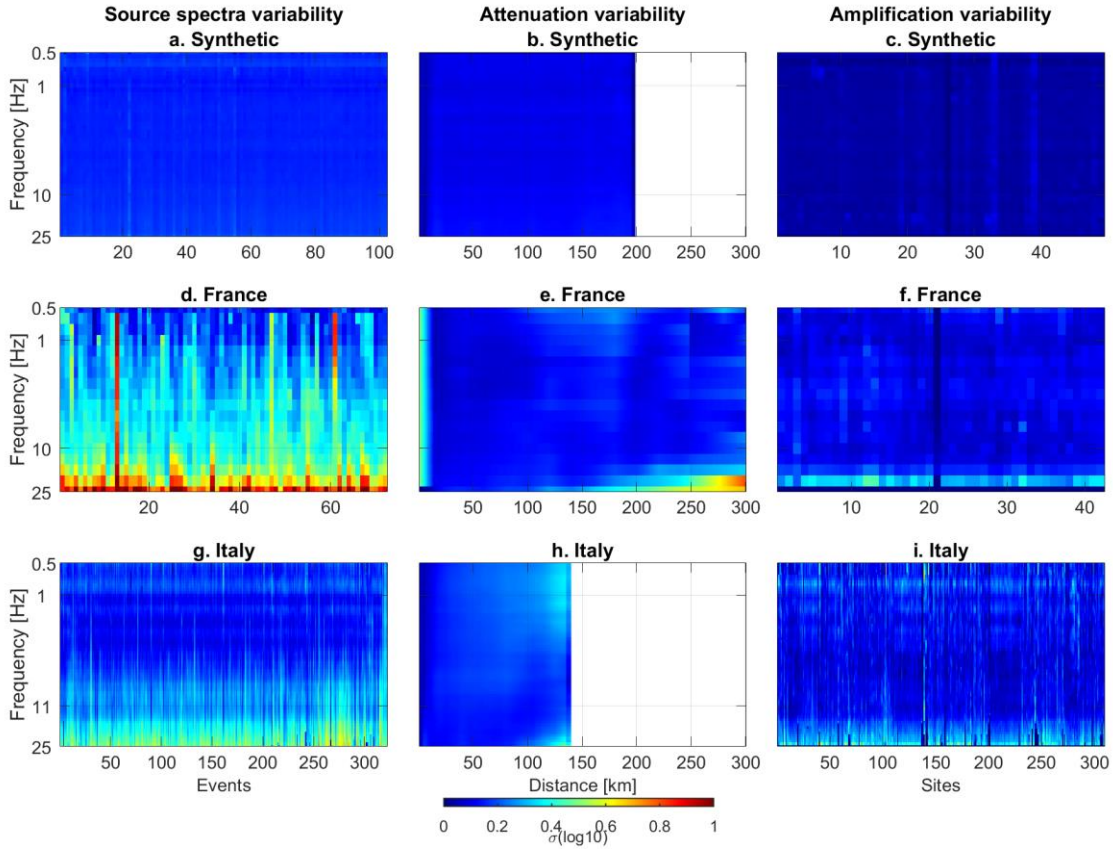

Figure 3.12: Synthetic, France, and Italy datasets. Estimations from the constrained phase (i.e., phase II) of the inter-methods variabilities (as indicated), as standard deviations in log_{10} at each frequency, for the source spectra and the amplifications obtained after the inversions. The single color scale is given at the bottom.

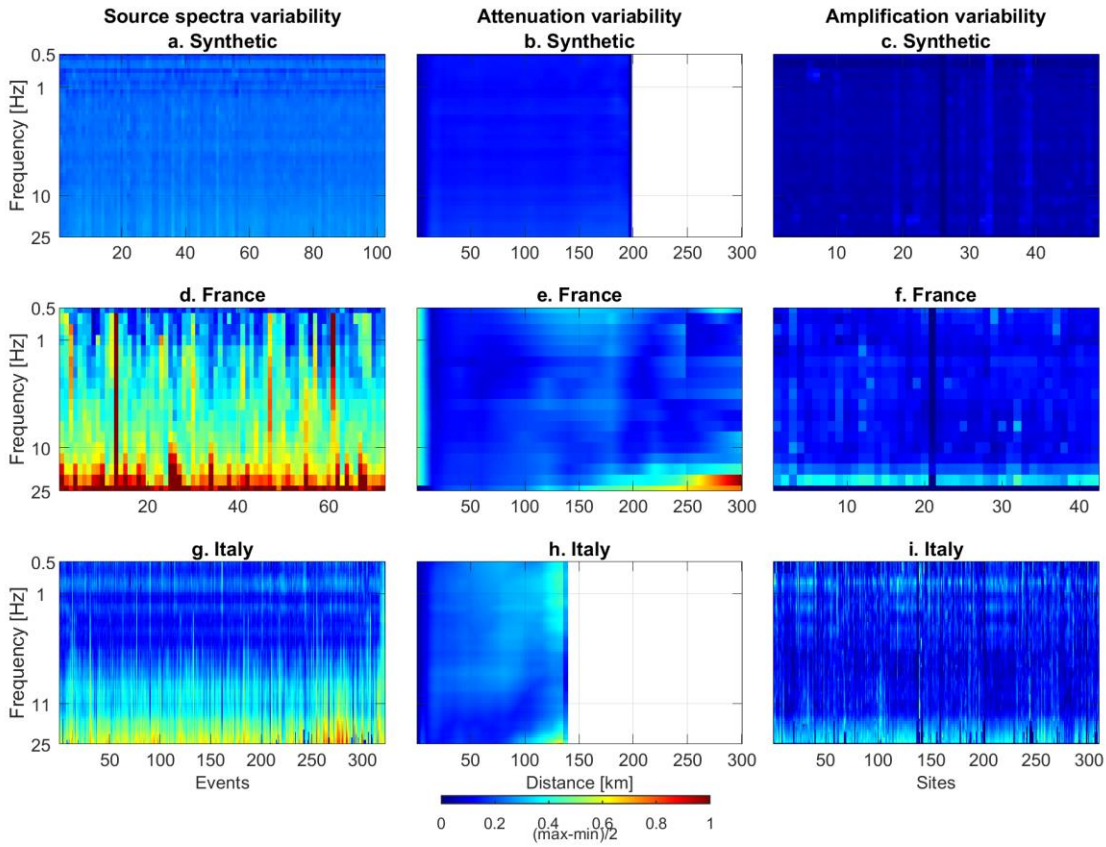

Figure 3.12: continued

3.5.2.2 Variability in source parameters (COV and maxmin):

Figure 3.13 and Figure 3.14 show the estimations of the variability of the source parameters using *COV* and *maxmin*. For each event, the variability from the six inversion types is assessed with *COV* and *maxmin*. In both cases, the distributions of these estimations are plotted over constant bins. In each bin, the average number of recordings is reported on the right axis (N_{rec}).

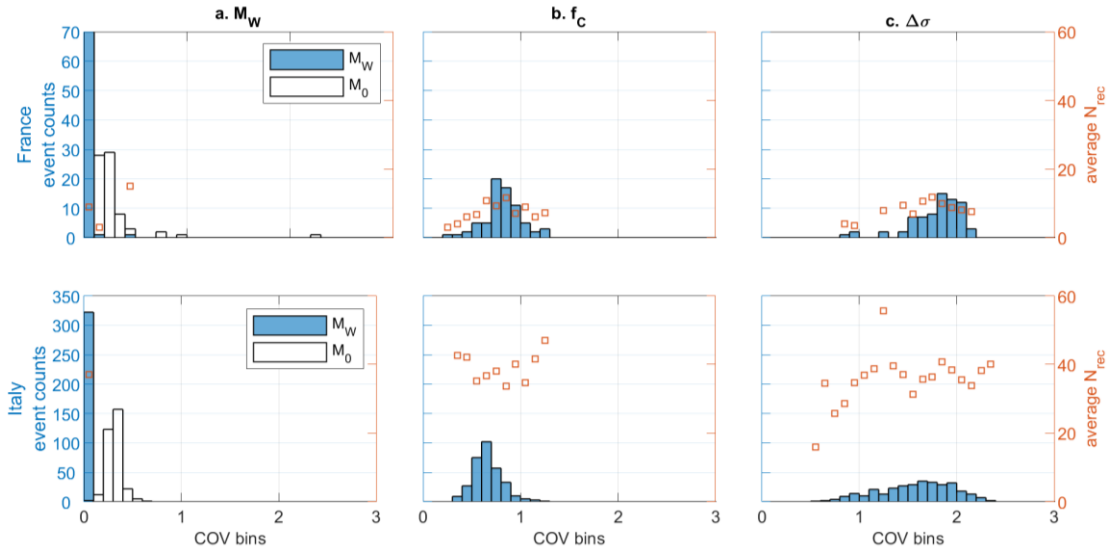

Figure 3.13: France and Italy datasets. Estimated variability for each of the source physical parameters M_w, f_c, and $\Delta\sigma$. Variability estimations are grouped into different bins of 0.1 std width. For each bin, the mean numbers of recording sites (Nrec) are also shown.

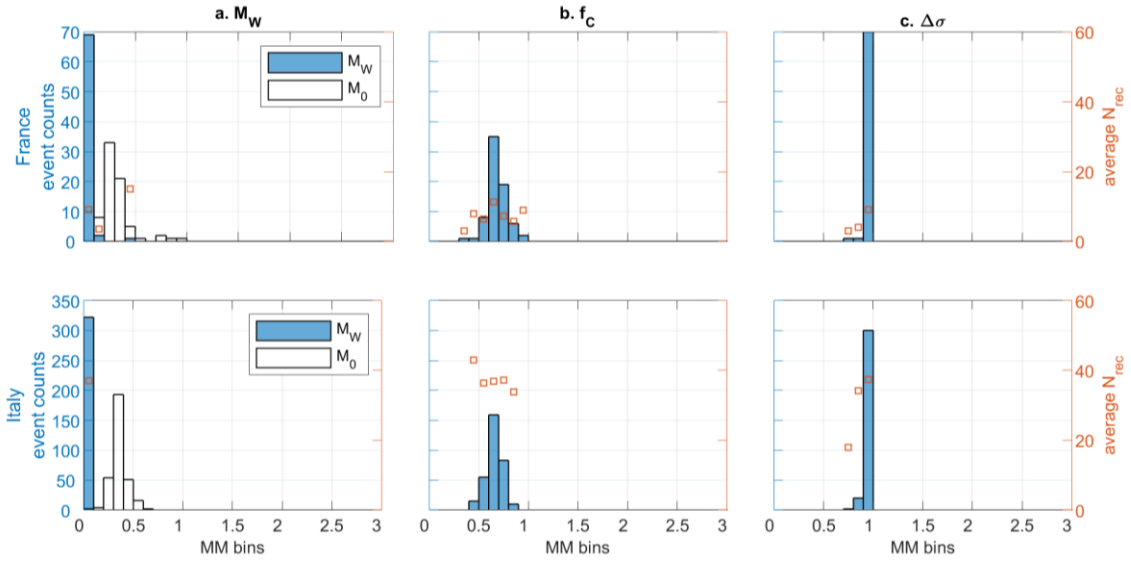

Figure 3.14: France and Italy datasets. Estimated maxmin for each of the source physical parameters M_w, f_c, and $\Delta\sigma$. Maxmin estimations are grouped into different bins of 0.1 widths. For each bin, the mean numbers of recording sites (Nrec) are also shown.

The statistics carried out for the France and Italy datasets show very low inter-method variability for the magnitude M_w, as compared to that for the f_c and $\Delta\sigma$ parameters. Indeed, the variability in terms of seismic moment M_0, which is directly used in the inversions, is higher than for that of M_w. This shows that the deduced M_w has lower variability, as it benefits from the relationship given in Eq(3.4). On the other hand, for the comparison of the two datasets, f_c and $\Delta\sigma$ show very small drops in their variability for central Italy (mainly in Figure 3.13). The *maxmin* representation shown in Figure 3.14 appears to limit the variability estimations more, where fewer bins are obtained than for Figure 3.13.

Any specific event in the France dataset was recorded on average by 10 sites, while for the Italy dataset, the events were recorded on average by 40 sites. This might represent evidence that data redundancy (i.e., to have events covered by as many sites as possible) is important for GIT. It is also interesting to note that the N_{rec} over the different bins for the stress is lower for the Italy dataset. When N_{rec} decreases, the variability is higher. However, it remained limited to *COV* of 1, while this extended beyond 1 for the France dataset.

3.5.2.3 Summary and comparisons of estimated variability

For a representative summary of the variability, we computed the generalized values for each dataset. For the nonparametric curves, the three frequencies of 0.5, 4, and 25 Hz were picked to compute the means of the variability per dataset for all of the sources and sites and at all distances within the common range for attenuation (i.e., [1-140 km]). Figure 3.15 and Figure 3.16 show the values plotted directly instead of in a tabulated form.

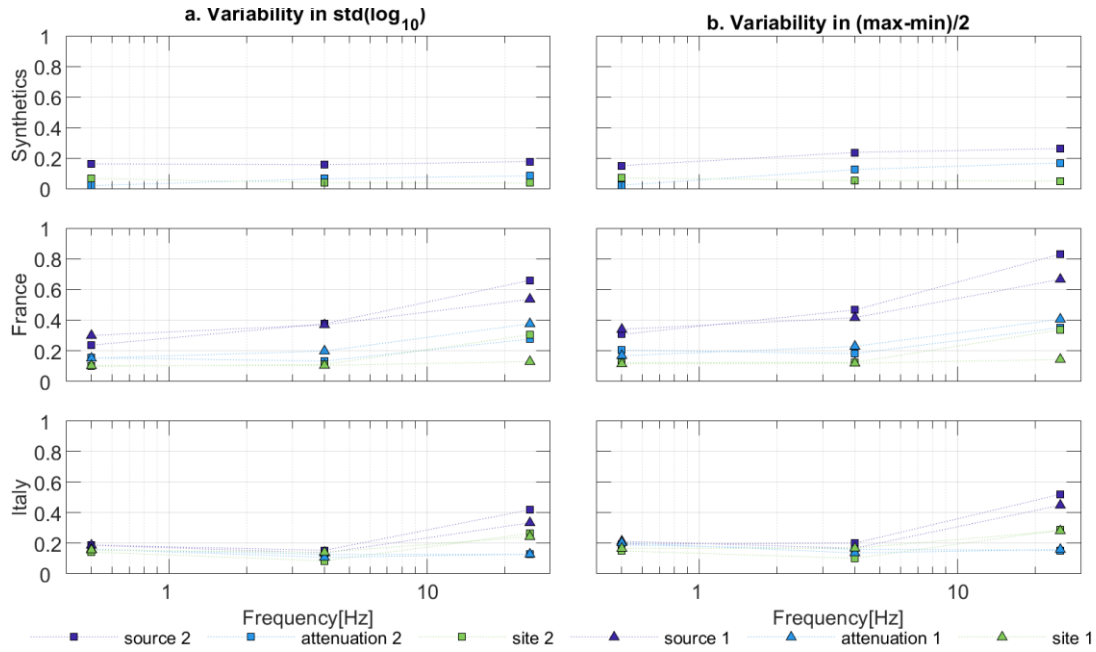

Figure 3.15: Phase I and II mean spectral variability estimations for all of the events, for attenuation at distances between [0-140 km], and for all of the sites. Three frequency points are considered.

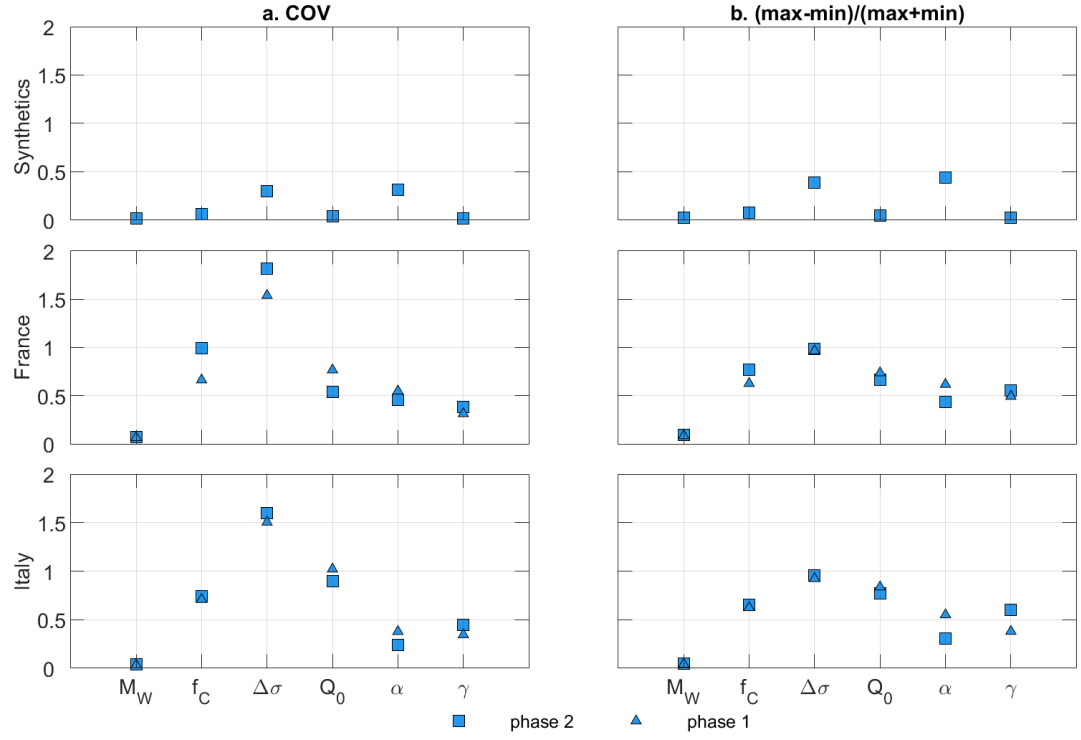

Figure 3.16: Phases I and II mean parameter variabilities for all of the events, with attenuation delivered by the different approaches.

Figure 3.15 shows the mean spectral variability for both indicators (i.e., *std* and *maxmin*). What can be seen from these comparisons is the high-frequency variability of the sources and sites. However, the highest variability is seen for the source and attenuation terms. The site terms represent the most robust terms of these estimations. Evidently, unifying the reference conditions in phase II did not decrease the variability observed.

The mean *COV* was also evaluated for all of the events, along with the mean attenuation parameters for each dataset (Figure 3.16). Although the *maxmin* estimations were generally lower than *COV*, the two GIT parameters that are most prone to high variability are $\Delta\sigma$ and Q_0. This might serve as a reference for the uncertainty on these parameters using GIT approaches.

3.5.3 Possible origins of variability: regional Variations

It is important to start to understand the origins of the variability observed. Therefore, we plotted the regional distributions of the spectral variability of the sources and sites in Figure 3.12. These figures were constructed from the variability estimates are presented in Figure 3.17 and Figure 3.18. At each source (or station), the color scale of the marker corresponds to the level of variability ($std(\log_{10})$), while the size represents the number of recordings (nb_{rec}). In this way, any regional changes in the variability can be captured.

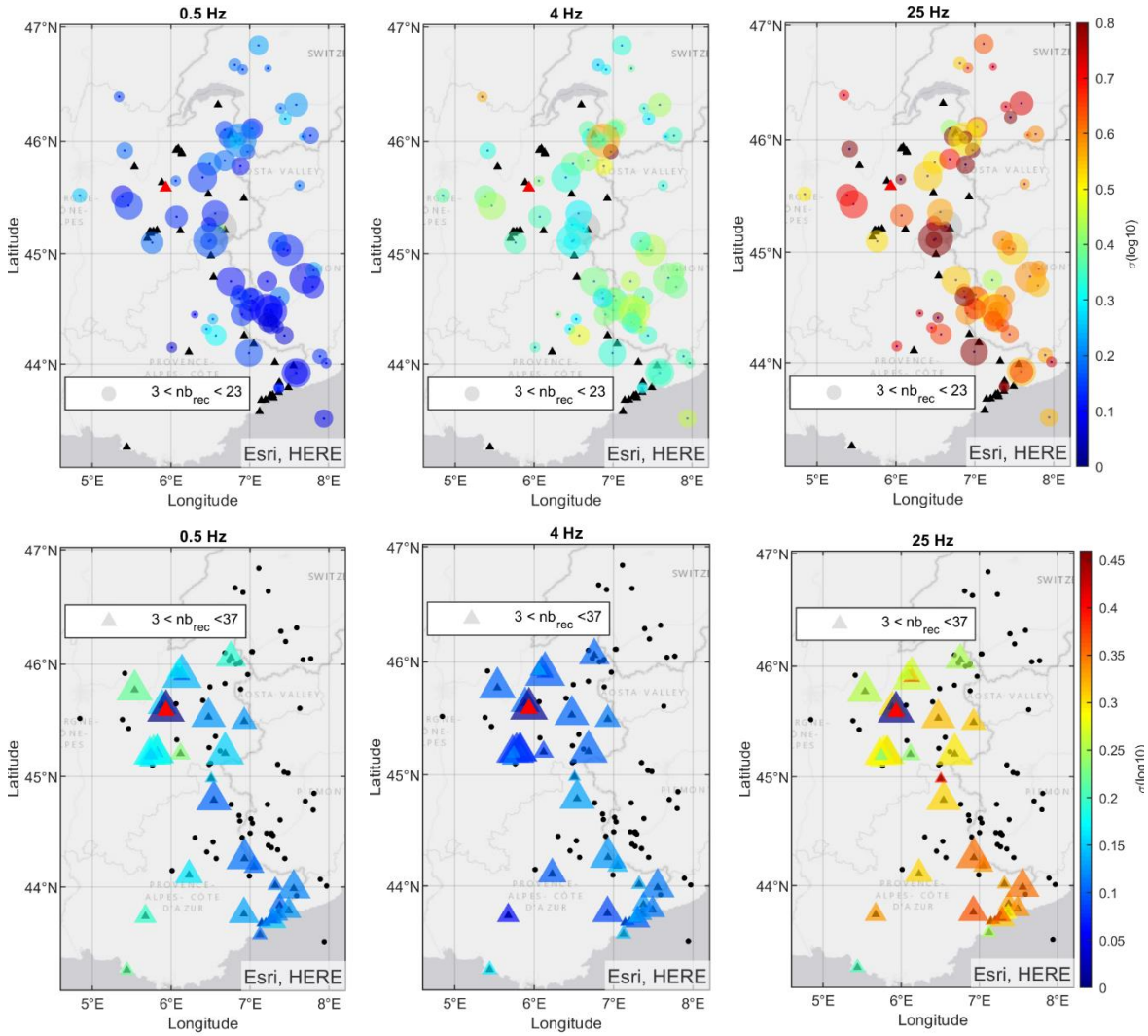

Figure 3.17: France dataset. Maps of the variabilities for the sources and sites in terms of std of $\log_{10}$ at three different frequency points. Triangles and circles, stations and earthquakes, respectively; symbol size, proportional to number of recordings; symbol color, inter-method variability at the given site or source.

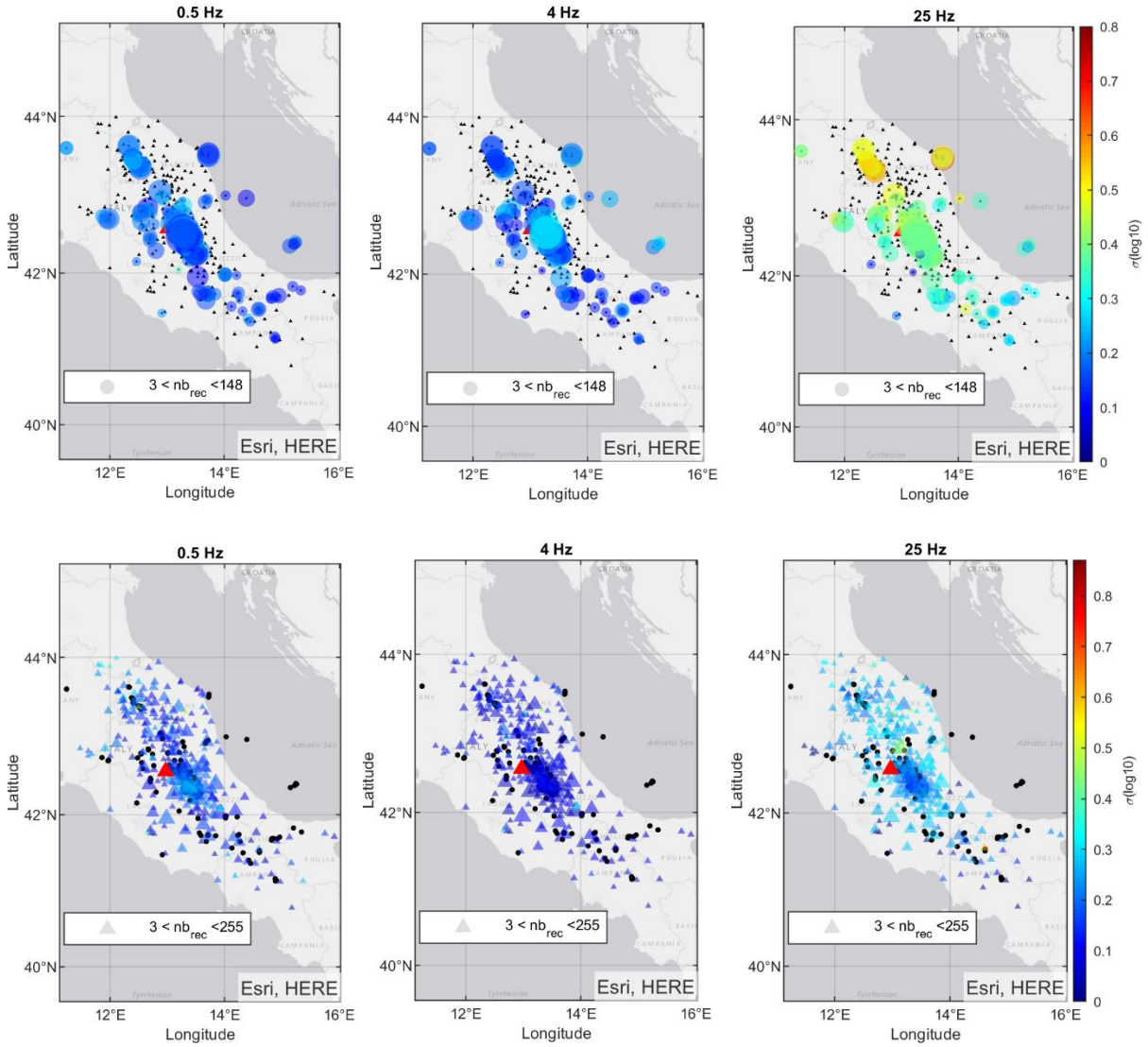

Figure 3.18: Italy dataset. Maps of the variabilities for the sources and sites in terms of std of log_{10} at three different frequency points. Triangles and circles, stations and earthquakes, respectively; symbol size, proportional to number of recordings; symbol color, inter-method variability at the given site or source.

For the French Alps data (Figure 3.17), the source spectra variability showed limited regional dependence, while the site variability showed regional changes at high frequencies in particular. At 4 Hz and 25 Hz, the sites in the southeast of France showed higher variability, although this is not clear at the lower frequency of 0.5 Hz. Going from the north to the south of the Alps, there was increasing variability.

Also, for the central Italy case (Figure 3.18), there were regional changes in the spectral variability, again at high frequencies in particular. The center of the dataset region (i.e., near the reference station marked in red) contrasts with the northern parts at all frequencies. For example, at high frequency, increasing variability is observed from the center to the northern regions.

We looked for recent studies of regional variations to identify any correlations within the regions considered. Indeed, regional variations are usually addressed in terms of

crustal attenuation, where regionalized maps of attenuation characteristics are established using attenuation tomography studies, such as those of Mayor et al. (2016) and (2018), in the European context. Indeed, these previous studies mainly used the coda waves to estimate quality factors from tens of thousands of recordings, and they displayed the results on maps that showed the lateral variations of the frequency-dependent quality factors.

After detailed comparisons of these maps, several aspects can be mentioned. First, within the France dataset, the variability changes at high frequencies tended to match the attenuation variations in the maps of Mayor et al. (2016), who focused on regions that included the southeast of France. Although the source variability was not apparent, the site variability changed from higher values at higher attenuation in the south of the dataset region to lower values with lower attenuation in the north. On the other hand, the Italy dataset shows that there was low variability with fewer attenuation zones in the center than in the northern parts at high frequencies. As a result, most of the higher variability (either for source or site) is seen in the higher attenuation zones.

In summary, in high attenuation regions, the properties observed additional variability compared to those of low attenuation. A direct interpretation of these correlations might be that regional variations that were not accounted for in the GIT inversions and models led to increased variability. The trade-offs allow this variability to be transmitted to either the source or the site or both terms.

3.6 Exploring the impact of the dataset size

After illustrating the results, we were interested in investigating the impact of a lack of data on the inversion results. For this, a ‘decimated’ subset of the France dataset was used, as follows:

- the number of sites was reduced from 42 to 14;
- the number of events was reduced from 72 to 24;
- The recordings per station/event (>3) was reduced, with a reduction in the total recordings from 645 to 107;
- The dataset referred to the OGCH hard-rock site as the chosen reference site.

With all of the other features remaining the same as for the initial GITEC France dataset, the inversion schemes of teams #01N1 (nonparametric) and #04P (parametric) were applied to both the original and this decimated dataset.

Examples of the results for two events are illustrated in Figure 3.19. These two inverted source spectra show that more significant changes affected the nonparametric GIT spectra when the data were reduced than the parametric GIT spectra. The ratios between the new and initial inverted spectra provide a clear illustration of the factors of the changes that occurred. In a second step, the mean ratio was computed (over all of the frequencies), and the events were grouped into bins following their mean ratio.

Applying the same procedure to the inverted site amplifications, Figure 19 also shows the distributions of the mean ratios obtained over the events and sites considered for both of the GIT methods. This ratio for events inverted with #01N1 covers a range of 0

to 5, while that for #04P remains limited. On the contrary, the site responses for the decimated dataset appear to be less affected for both methods.

These changes in the inverted spectra will definitely affect the physical parameter estimates, especially as it was not possible to post-fit some of the source spectra for the nonparametric method to a Brune model. This indicates the failure of the inversion to retrieve the event spectra when there are very few records available. Following these observations, it appears that the nonparametric GIT approach results in less reliable source spectra when there are reduced amounts of data compared to the parametric GIT approaches.

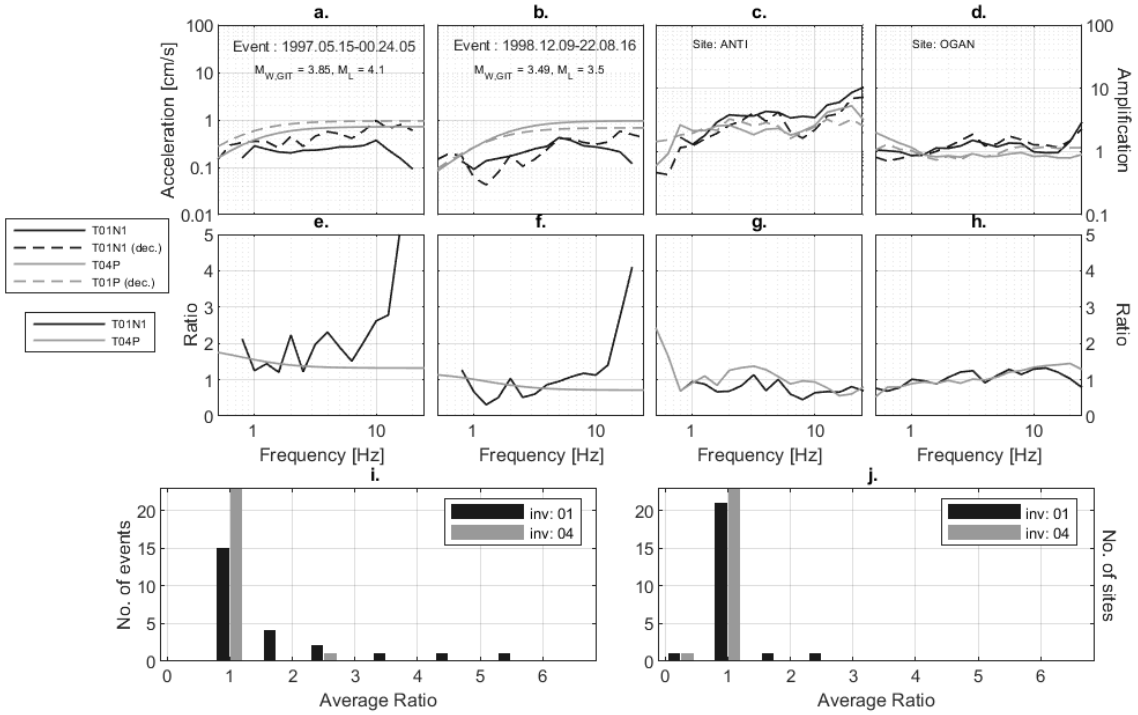

Figure 3.19: Decimated France dataset. Top: Examples of the sources for the two events (a, b) and sites (c, d) as considered in two inversions: once with the whole France dataset (solid lines) and once with the decimated dataset (dashed lines). Ratios of the above curves are also shown (e, f, g, and h). Bottom: Distributions of the mean ratios defined in the text for all of the events and the sites, as obtained after each of the inversions performed.

3.7 Summary and conclusions

This benchmark analysis was carried out to compare the methodologies of GITs across three datasets, one as synthetic data and two as real datasets used in previous studies. In addition, a simple synthetic dataset was considered to define the convergence of all of the methods to the correct solution. The inversions carried out served most of the initial benchmark objectives. The chosen datasets for the benchmark had different configurations and characteristics that can be attributed to some of the differences observed in the results of the inversions.

Two phases of inversions (I, II) were carried on the real datasets. First, attenuation terms showed variability with a factor of between 2 to 4 that changed with distance and frequency for both of the datasets. The resulting variability was mainly related to the source spectra when robust site terms were obtained. However, a strong trade-off appears to add variability of 0.4 std beyond 10 Hz at high frequencies.

The main observations and conclusions can be summarized as:

- The attenuation characteristics appear to be more complicated than the need for a simple homogenous quality factor. Significant attenuation changes with distance and regional variations of attenuations are behind the variability observed for the attenuation curves here. Also, such simple Q models can result in very unstable estimations of the attenuation parameters.
- The site terms appear to show the lowest variability among the different results, where site references can have important roles in terms of the average site responses in the dataset considered.
- The source terms appear to be subject to high variability. Data redundancy appears to be important to decrease this variability (i.e., sufficient numbers of recordings per event/site).
- The M_w source parameter shows less variability among the different approaches than the corner frequency (f_c) and the stress drops $(\Delta\sigma)$. This might be a result of the way that the parameterization uses M_0 directly, instead of M_w. The relation used to deduce M_w from M_0 ($M_w \propto log_{10}(M_0)$) leads to reduced variability in the results for M_w.
- The comparisons after data decimation suggest that the nonparametric scheme tends to deliver less reliable source terms when the amount of data is reduced while the site response remains more stable. Thus, keeping the dataset updated with recordings and including new stations can provide essential benefits for nonparametric GIT approaches.

Based on the different approaches used (i.e., parametric, nonparametric), each appears to have its own pros and cons relative to the desired application. With relatively few data in a given dataset, applying a nonparametric approach appears to be counter-indicated for both source and attenuation. For example, the attenuation model derived for such an approach can only be defined using the source-site distance points in the dataset. In contrast, with a parametric approach, the model assumed from the beginning is calibrated through inversion iterations to fit the recorded data characteristics.

Nonparametric GIT is accompanied by an assumption made on the source and attenuation parameters defined by the reference distance 'R_{ref}'. This assumption implies that all sources are shifted to the R_{ref}, whereby the attenuation starts from unity at this distance. This can be a very approximate assumption when there are little or no data in the short-distance range (e.g., <50 km).

On the other hand, nonparametric approaches appear to provide more information about earthquake sources and regional attenuation than parametric approaches because the unmodeled source spectra and attenuation curves can be visualized with fewer *a-priori* assumptions. Therefore, an important aspect is to reduce the *a-priori* assumptions as much as possible to improve the adequacy of the models.

For the preference between the GIT methods, several points can be mentioned. Where there is a sufficient amount of data, the nonparametric approach represents a reliable

tool. As nonparametric inversions estimate source and attenuation models based on the data, then the richer the data, the more robust the models are inverted. On the other hand, GIT on small datasets would be better carried out according to parametric inversion schemes under the conditions that provide the necessary constraints for the inversions from *a-priori* knowledge.

Following the several aspects indicated above relating to the reliability of GIT at high frequencies, several questions can be posed. First, it can be directly proposed that reducing the assumptions on sources (i.e., nonparametric spectra, instead of the Brune model) might be the reason behind the high-frequency discrepancies. In other words, the source spectra can follow a multiple corner frequency form. For example, the study of Bindi et al. (2019) that used data from central Italy proposed a source model that was anchored by a parameter they called 'κ_{source}'. . This confirmation for the presence of high-frequency slope for sources can only be accepted if the inversion schemes are shown to be stable and free of high-frequency trade-offs. Thus, the high-frequency performance of the inversion schemes remains under question. On the other hand, if the susceptibility of the nonparametric inversions to trade-offs is accepted, then these differences at high frequency can be misleading. Thus, the parametric models (especially the Brune model) might be preferred instead of accepting that these trade-offs occur. However, the final answer to this question about the high-frequency performance of GIT methods remains relatively unclear and needs further investigations and testing.

To conclude, the preference of either a parametric or a nonparametric approach does not appear straightforward in general. The preference might be case-dependent, based on the dataset of interest (e.g., the amount of data available, the geometry of the network, etc..). However, as best practice, we propose that it is best to proceed carefully with the two approaches in parallel, with continuous comparisons of the results from the different approaches before the direct use in other related applications.

3.8 Perspectives

At the end of the GITEC workshops, several open questions were defined. These remaining questions constitute the several perspectives for this work:

- Additional comparisons to understand the results and the origins of the variability.
- As correct attenuation parameterization has a primary role, a more complicated synthetic dataset was discussed. Such a synthetic dataset can take into account regional variations and the depth-dependence of the attenuation.
- What the minimum/maximum magnitude limits are below which (or above which) the inversions can resolve the corner frequencies appears to be an interesting question.
- High-frequency discrepancies should be further investigated to understand the observations better and build up reliable conclusions about the source and site of high-frequency attenuation.
- Additional aspects of dataset configurations are interesting to explore (e.g., the impact of the seismological network geometry with respect to earthquake

location and *vice versa*). This might help to improve the efficiency of future inversions by providing optimized datasets.

It is interesting to address the possibility of capturing more complicated seismic wave phenomena using GIT approaches, such as extending the assumptions on the point-source models generally used in GIT and for stochastic ground-motion simulations.

Chapter 4: Site response estimations on KiK-net sites

Context

Chapters 4 and 5 constitute the draft of two companion article proposals under preparation for submission to a scientific journal. The work in these chapters is an application of the deconvolution methodology with GIT site responses. This work on KiK-net was a necessary step before deriving GIT deconvolved GMPEs on other datasets. We aimed in this chapter to reduce the need for V_S profiles since they are not available for all sites in other networks. Nevertheless, V_S profiles for some sites will still be necessary and can be helpful to solve the GIT-related reference station question.

Abstract

Site-specific seismic hazard assessment involves the prior knowledge of 1) the input ground motion at the local bedrock and 2) the site-response characteristics. The deconvolution approach, which we aim to apply to obtain reference motion, requires site-response estimations in the Fourier domain. In this chapter, we propose to complement theoretical 1DSH site responses used in Laurendeau et al. (2018) with GIT empirical estimations (generalized inversion techniques). First, this step reduces the assumptions on site effects, and second, it reduces the necessity for site-specific V_S profiles. Consequently, reliable site response estimations are no longer limited to the presence of site information at each site. Also, GIT estimations can be considered as absolute global terms. The notions of absolute and global scales are detailed in this chapter to understand their significance. After estimating site responses for all KiK-net stations with different methods, i.e., empirical and 1D theoretical, we conduct comparisons and conclude a new list of 1D sites. The results obtained will be the primary input elements of the next chapter.

Keywords: *site responses, generalized inversions, 1D conditions, KiK-net.*

4.1 Introduction

A key component in site-specific seismic hazard assessment (SHA) is the determination of ground motion at the reference seismic bedrock, which most often coincides with very hard-rock materials (e.g., Rodriguez-Marek et al., 2014; Aristizábal et al., 2018). The definition of this reference motion allows us to estimate the site-specific amplification, which is essential in practical applications of SHA. This amplification is usually a function of frequency and can also be a function of the incident and back-azimuth angles for a given seismic scenario. The amplification can also depend on the ground-motion intensity if we consider non-linearity effects. Several methods are present to characterize this site-specific amplification, such as the numerical simulations using the S-wave velocity models, if available, in addition to the empirical methods like the standard spectral ratios (SSR) with a reference local outcropping bedrock.

The reference hard-rock sites usually have high values of the S-wave velocity V_S. However, ground motion prediction equations (GMPEs), used to assess regional seismic hazard, are not representative of these high V_S values due to a lack of accelerometric stations on very hard-rock sites (e.g., Laurendeau et al., 2013). To solve this issue, the state-of-the-art technique is to use the “host-to-target” adjustments

(HTTA), which rely on a site correction procedure between the host sites (i.e., those sites delivering records used in the GMPE derivation) and the target site (i.e., the studied specific site). These corrections could be applied in two different ways. First, there is the correction in terms of response spectrum (SA) using the point source stochastic model (PSSM, e.g., Campbell 2003; Scherbaum et al., 2004; Cotton et al., 2006; Van Houtte et al., 2011). Second, the correction can also occur in the Fourier domain by transforming response spectra to Fourier amplitude spectra (FAS) using the inverse random vibration theory (IRVT, Al Atik et al., 2014).

The HTTA corrections use a couple of parameters, mainly the "V_{S30}" and "κ_0". The parameter V_{S30} is the average shear wave velocity up to 30 meters, which has been used as a proxy to describe the sites, especially in GMPEs (e.g., Boore and Atkinson, 2008) and site classification in building codes (e.g., Eurocode 8). Besides, the term κ_0 is related to the high-frequency attenuation properties and has focused the interest of the engineering community to complement the V_{S30} proxy, at least for rock sites (Ktenidou et al., 2014). κ_0 has been introduced first by Anderson and Hough (1984) to characterize the attenuation of seismic waves in the first few hundreds of meters or kilometers beneath the site.

Since the standard practice is largely commented on for the lack of its accuracy (as detailed in the introduction and chapter one of this document), several alternatives are proposed. Within the thesis framework, we aim to transpose the deconvolution methodology of Laurendeau et al. (2018) to other datasets, mainly in European regions. The work of Laurendeau et al. (2018) benefited from several aspects present in the KiK-net network, which are not present elsewhere. Almost all KiKnet sites are characterized by velocity profiles, which are deep enough to reach the bedrock levels (i.e., few hundreds of meters). Also, two sensors exist at each station, one at the surface and another at the borehole level. These two factors first allowed Laurendeau et al. (2018) to estimates surface-to-borehole transfer functions using the velocity profiles. Then, a validation by empirical estimations was possible through the empirical surface to borehole spectral ratios. After identifying sites with dominant 1D effects, 1DSH theoretical transfer functions were used to remove site effects from surface recordings. After this correction, corrected surface recordings should, in principle, represent reference motion that is amplification-free. In the end, the applicability of the deconvolution method still needs improvements since most of the other existing strong-motion networks lack the same features as KiK-net arrays.

The limitations that face the deconvolution method of Laurendeau et al. (2018) should be broken down to ensure transposability to other networks. These limitations mainly concern the site effects estimations, the main element of the deconvolution approach. Laurendeau et al. (2018) needed to constrain the analyses to "1D sites" to use the 1DSH site responses in the deconvolution procedure. In fact, the limitation to 1D site conditions is a challenging step and not easy to achieve everywhere. First, the absence of velocity for each site prevents a theoretical estimation of all the site responses. Second, as downhole sensors are not available for other datasets, the essential elements to establish 1D-site classifications can be no more available. The third point is that the deconvolved reference motion was associated with downhole site velocities (V_{SDH}) whenever needed in GMPE regressions. Thus, we need to address the V_S conditions associated with the deconvolved reference motion on other networks and

other methods of site response estimation. In view of addressing these questions, we carry on our investigations on the KiK-net data before proceeding to European databases.

In this chapter, our objective is to extend the methodology of Laurendeau et al. (2018) on KiK-net with the help of GIT as an alternative for reliable empirical estimation of site responses. First of all, we considered updating the used KiK-net database to include recordings from 1998 up to 2017. Due to the enormous amount of data available, we avoid the manual approach of data treatment used in Laurendeau et al. (2018) (e.g., record-by-record manual picking of P- and S-wave arrivals, extensive visualizations of the recordings). Instead, we adopt an automatic picking procedure for P- and S-waves inspired by the picking tool developed in Baillard et al. (2014). As accuracy questions may arise, we apply several checks on recordings to reject or correct the faulty pickings. Afterward, we estimate the site responses through the possible empirical approaches (Standard spectral ratios and Horizontal-to-vertical spectral ratios) and 1DSH theoretical calculations that use the available site information for KiK-net stations. Afterward, we perform a detailed and regionalized generalized inversion (GIT) for entire Japan. Even though our interest in GIT lay in its resulting site terms, we needed to cope with the source and attenuation terms present in the definition of the inverse problem. Then, we choose simultaneously non-parametric and parametric inversion schemes to follow the recommendations of the GITEC benchmark presented in Chapter three. The results of source and attenuation terms are briefly detailed in our text, and we focus on the site terms. The use of GIT allows reducing the necessity of V_S profiles of all KiK-net sites to only a few sites.

Further on, we investigate the effect of the reference station choice in GIT on the absolute site responses (the surface transfer functions as obtained in GIT). In the end, we conduct comparisons between theoretical 1D and GIT approaches to identify a new list of sites with dominant 1D effects. These sites will be considered acceptable for the 1DSH site response estimation. The latter appears essential to keep track of the results of Laurendeau et al. (2018) and compare them with the new results of GIT site responses. Finally, all the estimated site responses here form the main elements of the following chapter.

4.2 KiK-net dataset adopted

The KiK-net network is one of several monitoring networks established throughout Japan after the Kobe earthquake in 1995, whose M_w was 6.9 (Okada et al., 2004). Since December 2011, the KiK-net network comprises 692 stations having two sensors at each site, one at the surface and another deep in the borehole. Depths of boreholes in the KiK-net network vary between 100 and 250 meters except for some sites where depths can reach a few kilometers. Sampling rates of KiK-net sensors vary between 100 and 200 Hz. This network provides V_S profiles measured by PS-logging measurements for around 655 stations.

In this study, we consider the dataset provided by Dawood et al. (2016), which comprises events from 1997 to 2011, updated later on by Bahrampouri and Rodriguez-Marek (2019) by events up to 2017. To address our objectives, we had to exploit time histories and not only the online provided spectral amplitudes. Event metadata from the F-net catalog and station metadata from the KiK-net site were provided in this

flatfile and consequently considered in our study. In addition, we selected recordings of events that were classified as shallow crustal.

4.2.1 Automated Onset detection

We adopt an automated procedure for a rapid and efficient P- and S-waves onset time detection (T_p and T_s), inspired by the algorithm of Baillard et al. (2014). The automatic picking we considered relies on calculating the Kurtosis value, defined as a fourth-order moment around the mean of the distribution. Kurtosis is a parameter that indicates whether the observations of a random variable follow a Gaussian distribution (often low-valued Kurtosis, e.g., near unity) or not (high values). Amplitudes of ambient vibrations generally follow a Gaussian distribution before it is disturbed by the arrival of seismic waves (a non-Gaussian field) in the seismic trace. This disturbance creates a rapid increase in the Kurtosis value, allowing the detection of onset times.

We calculate the Kurtosis over sliding windows on the seismic trace, which allows detecting any significant change in the record (i.e., P- or S-wave arrival). Figure 4.1 shows an example calculation of the Kurtosis noted as the first characteristic function CF_1. After a series of transformations passing by CF_2 and CF_3 functions (transformed characteristic functions 2 and 3, respectively), we can obtain a variable that shows two global minima that should correspond to P- and S-waves. A proper picking of the minima is ensured by searching within specified intervals for each onset. Indeed, it is prevalent to use theoretical arrivals of P-and S-waves to fix search intervals, which result from a 1D V_S-model of the crust. Thus, we use hereafter the model provided by the JMA agency for Japan.

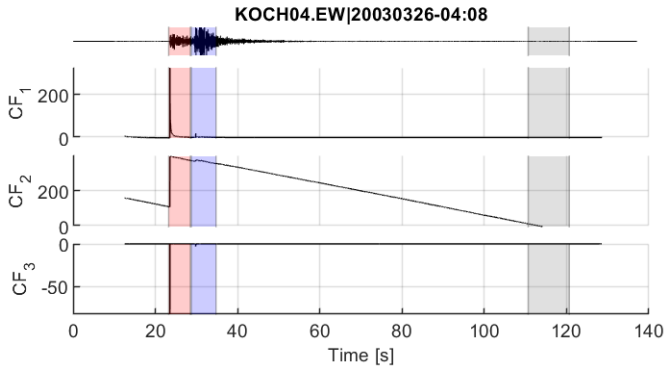

Figure 4.1: The characteristic functions obtained from the Kurtosis statistical parameter. CF1 and CF2 are the transformed functions to reach CF3, on which we pick the two minima that correspond to the P and S onset arrivals.

First of all, one could limit the automatic search for T_p to a 10 seconds interval, centered on the theoretical pick, $T_{p,theo}$. However, KiK-net earthquakes have the time reported in catalogs to the nearest minutes, which creates uncertainty of 60 seconds on the value of $T_{p,theo}$. Thus, we integrate the "Pphase-Picker" tool delivered by Kalkan (2016) along with our approach for improved accuracy of T_p picking. After series of trials, the Pphase-Picker seems to perform robustly for the KiK-net recordings. Eventually, we select T_p values resulting from the Pphase-Picker and ensure that it lies near the $T_{p,theo}$ (even if within 60 seconds). Any kind of impossibility to pick T_p within the latter conditions, the signal is flagged out for further visual and manual treatment.

Once T_p is picked, we define a search interval for T_s as shown in Eq(4.1). Here, $T^*_{s,theo}$ stands for pseudo theoretical S-wave arrival and is based on the theoretical arrival time differences, as $T^*_{s,theo} = Tp + (T_{p,theo} - T_{s,theo})$. The term α is arbitrarily chosen as 30% in our study after series of trials, and D_5 and D_{95} is the signal duration at 5% and 95% defined by the cumulative absolute velocity (CAV). We do this procedure for each component separately, for surface and borehole stations, and then we maintain the mean value of T_s from different components after excluding outliers.

$$T_s = [a_1\ ;\ a_2];\ where \begin{cases} a_1 = \max(T^*_{s,theo} - \alpha D_{95}\ , D_5) \\ a_2 = T^*_{s,theo} + \alpha . D_{95} \end{cases} \quad (4.1)$$

A series of checks are applied to ensure the accuracy of auto picking and thus to flag any possible errors that could occur. Each of the following cases will be flagged:

- Any significant deviation (>120 s) from the theoretical arrival.
- When the amplitude ratio of a window (10 s) after T_p to another window of equal length before shows values <1.
- When the maximum amplitudes do not lie in the S-wave window defined following Perron et al. (2018).
- All other kinds of impossibilities to detect T_p and T_s

Furthermore, this automated picking tool was tested on records that were manually picked to verify its accuracy. After picking T_p and T_s and visualization of flagged records and rejection of others, we finally define P- and S-wave and noise windows following Perron et al. (2018). In the end, Fourier amplitude spectra (FAS) of each window are computed, and the signal-to-noise ratio (SNR) of S-wave windows is evaluated over 33 frequency points between 0.5 and 20 Hz.

4.2.2 Selection of Ground Motion Recordings

Within the scope of the work, we set out some criteria to choose data from the enormous KiK-net dataset. For example, the data contains events that can be classified into different categories depending on the location with respect to faults and plate boundaries and its depths (Oth et al., 2011; Nakano et al., 2015). To maintain the focus on our objectives, we select data according to the following criteria (Figure 4.2):

- Active shallow crustal earthquakes (ACR shallow) following the classification of (García et al., 2012) as provided in the flatfile of (Bahrampouri and Rodriguez-Marek, 2019).
- Sites with available V_S-profiles, provided by the Japanese National Research Institute for Earth Science and Disaster Resilience (NIED).
- Focal Depths less than 30 km, where depths were rechecked by the Fnet-Catalog estimations based on moment-tensor solutions.
- Hypocentral distances (R_{hypo}) not exceeding 300 km.

For the application of GIT, we perform the data processing and preparation as follows. Both downhole and surface recordings are selected for generalized inversions with the aim to obtain surface and borehole site transfer functions simultaneously. Also, we retain only the frequencies of a recording's FAS that pass the SNR test (> 3) only if

they are at least 60% of the frequency points. In this way, we keep as many frequency points as possible. On the other hand, for GMPE derivations in the next chapter, we ensure that the entire frequency band 0.5 to 20 Hz should pass the SNR test. The current selection and treatment results in a subset of 912 events and 595 sites, forming 16335 recordings.

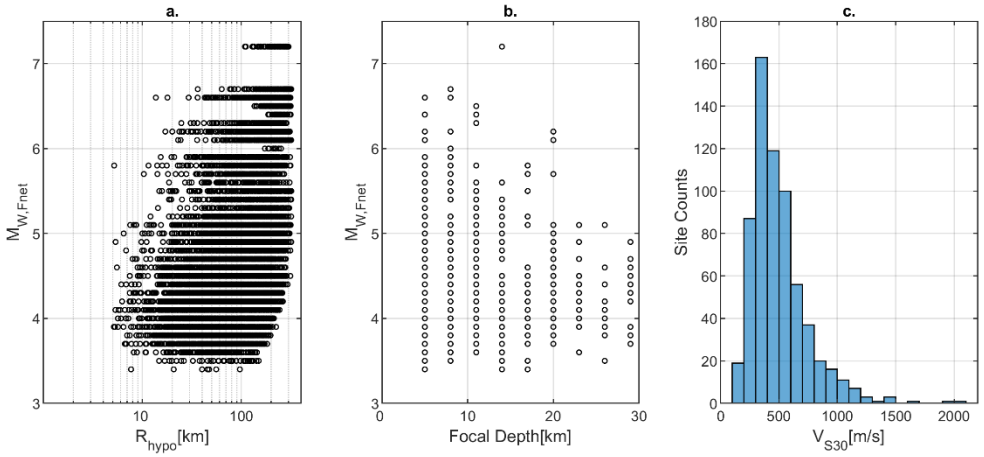

Figure 4.2: Distribution of the dataset (16335 recordings) ready for GIT analyses. a) Magnitude-Distance plot. b) Magnitude-depth plot c) V_{S30} distribution

4.2.3 Attenuation regionalization of Japan

KiK-net network covers the whole Japan, which lies on different tectonic plates and is crossed by volcanic fronts. That is why regional and lateral variations between different parts of Japan can be significant. Here, we propose to divide Japan into seven different attenuation zones (Figure 4.3) based on several studies of attenuation tomography. For example, Pei et al. (2009) investigated lateral variations of Q-values of S-wave in the upper crust and found (at 1 Hz) very low values along the volcanic fronts forming a boundary in central Japan. Eastern coasts showed high Q-values. Wang et al. (2017) and Wang and Zhao (2019) applied a similar but updated process to investigate for Q_s, Q_p and Q_p/Q_s. These studies delivered clearer maps, reached consistent conclusions regarding the distribution of Q-values, and confirmed the low Q-value boundaries throughout Japan. Based on this, we define polygons 1, 3, 5 and 6 for eastern coasts and 2, 4 and 7 as mainly separating into 2 parts from the island of Hokkaido in the north through central Japan towards Kyushu island in the south. We take into account lateral Q variations between the eastern and western coasts. The main reason behind subdivision of eastern coast into 3 polygons (3, 5 and 6) is the very low Q-values between the subduction trench and eastern coast found in the mentioned studies.

We set these polygons in a way to keep a minimum amount of data inside each polygon. The latter subdivision of Japan into polygons, to determine regionalized attenuation, remains a strong assumption. However, it can be sufficient to exploit and understand the lateral variability of seismic attenuation properties and exclude a strong possible bias in site-response estimation in the GIT application. To avoid additional complexity in the problem, we keep records that traverse two regions at most, and we remove records that traverse more than three regions if all regions traversed by the ray do not contain more than 60% of the ray path.

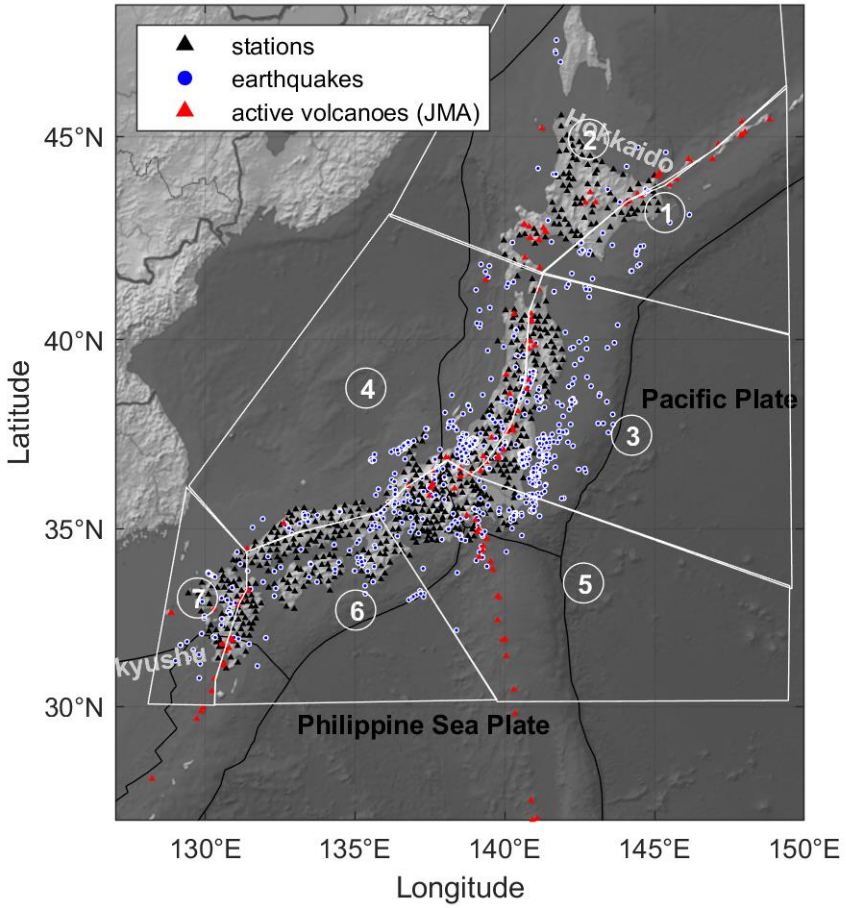

Figure 4.3: Japanese national map showing the distribution of stations and events in our KiK-net dataset. We also show the attenuation regions subdivided into polygons following several tomographic studies (Pei et al., 2009; Wang and Zhao, 2019).

4.3 Site response estimations

Generally, site responses are estimated through empirical approaches (directly from recordings) or theoretically through numerical modeling using measured site information.

4.3.1 Empirical

The central concept in empirical estimations is that the signal recorded at the surface results from a convolution between source, path, and attenuation effects. Then, defining a reference site plays a crucial role in the amplification estimation. Borcherdt (1970) first proposed to estimate the amplification for a site through the spectral ratios of recorded horizontal components to those of a reference site, which is often a nearby outcropping rock. Since rock motion is assumed to be amplification-free, spectral ratios are supposed to isolate site amplification from source and attenuation effects. However, identifying a reference site is not straightforward (Steidl et al., 1996) and remained subjected to limitations, especially in the presence of strong topographic effects. For the case of KiK-net arrays, SSR is usually performed with the downhole sites as reference (Cadet et al., 2012; Thompson et al., 2012; Laurendeau et al., 2018; Pilz and Cotton, 2019), which we denote here as SSR_{dh}. This step facilitates the task to have site-by-site empirical amplification for all KiK-net sites, though it might be biased with down-going waves (Cadet et al., 2012).

Then, horizontal-to-vertical spectral ratios (HVSR) gained attention since it proposed simplifying the estimations using the vertical component on the site instead of the reference horizontal components as first introduced by Nakamura (1989). Then, Lermo and Chávez-García (1993) extended the application of HVSR from microtremors to the strong S-wave part of earthquake signals. This method has shown helpfulness to reveal some site features in responding to earthquake motion with some underlying assumptions. Though initially proposed for amplification estimation by Nakamura (1989), several studies disagreed later and related it to the fact that vertical components might exhibit some amplification (Bindi et al., 2009; Ameri et al., 2011). HVSR has become a standard tool to assess the site fundamental resonance frequency, f_0 (Bard, 2008; Haghshenas et al., 2008; Griffiths et al., 2016a). While clear peaks in the ratios often indicate a strong impedance contrast in the geological structure beneath, interpretations of flat HVSR ratios can be complex and misleading.

For SSR_{dh} and HVSR computations, we consider the frequency band defined earlier (0.5-20 Hz). Fourier amplitude spectrum (FAS) of each component is computed and smoothed using Konno and Ohmachi's (1998) procedure with a smoothing parameter b=30. At each site, horizontal components are combined into a single component using the geometric mean, then SSRdh and HVSR ratios are computed for each event with recordings satisfying the SNR criteria over the considered bandwidth.

Figure 4.4a shows examples of empirical HVSR for four KiK-net sites, where clear peaks are observed at high or intermediate frequencies in addition to double peaks in some cases (AICH16). Clear, wide, unclear, or multiple peaks are generally encountered when observing the HVSR. Also, individual inspections of some sites can show strange wide amplification at low frequencies with a clear peak in intermediate to high frequencies. In Figure 4.4b, we show estimations of the SSR_{dh} for the same sites, which confirms fundamental peaks observed by HVSR curves despite the presence of the same peaks due to destructive interferences.

4.3.2 Theoretical

In the first level comes the theoretical approximation of QWL, introduced by Boore (2003), to account for a frequency-dependent site amplification factor (Poggi et al., 2012; Edwards et al., 2013). The generic practice aims to calculate the QWL amplification using the present V_S profile and add an attenuation parameter (i.e., κ_0) to account for cumulative damping effects. However, this approximation results in a steadily increasing amplification and thus misses the amplification peaks, especially in high frequencies leading to a rough estimation of amplification (Laurendeau et al., 2018). From our perspective, if V_S profiles are available, then 1D simulations remain much more accurate than QWL approximations, which we are trying to avoid.

In the absence of recordings, linear 1DSH numerical simulations can estimate the site transfer functions that represent the local amplification limited by the known V_S structure below the site. Here, we use the 1D reflectivity model (Kennett, 1974) to determine the response of horizontally stratified layers (using the software written by Bard and Gariel, 1986 and used later on in several investigations, e.g., Cadet et al., 2012; Laurendeau et al., 2018). The numerical simulations need an input definition of the density and damping factors, in addition to the V_S-profiles provided with the KiK-net network. We adopt for this aim the relationships defined by Brocher (2005). For the

quality factor, we assume the commonly used factor Q_S=V_S/10 (Cadet et al., 2012). As a result, we are capable of computing the surface transfer function (TF_{surf}) and the borehole one (TF_{dh}). We thus apply Konno-Omachi smoothing of coefficient b=30 (consistent with smoothing of real data) to avoid the sharp peaks obtained, especially in TF_{dh}. After smoothing, the borehole transfer function (BTF_{1DSH}) is obtained from the ratio of TF_{surf} and TF_{dh}.

Recent works tried to account for uncertainties in V_S-profiles and their impact on 1D site-response estimations (Griffiths et al., 2016a; Teague et al., 2018; Passeri et al., 2019). In addition to 1D assumptions, the considered V_S models are accompanied by aleatory uncertainties that remain somehow difficult to separate from epistemic ones. One of the methods proposed to account for aleatory uncertainties is Shible et al. (2018) (and Chapter two of this document) that introduces small-scale perturbations based on observations from high-resolution PSSL measurements. In fact, vertical heterogeneities introduced in this method account for shallow layer scattering conditions worth testing even in 1D conditions.

Therefore, in this text, we look to improve the 1DSH prediction by accounting for vertical heterogeneities proposed by Shible et al. (2018). We introduce vertical perturbations in the velocity profiles by assuming a lognormal distribution of the median μ_{scat}=1 and standard deviation σ_{scat}=0.3. Then, to have a representative accounting for heterogeneities, we believe that 30 realizations can be enough to explore the variability. Afterward, we apply a multi-step analysis to restrict the realizations only to reasonable generations of the profiles. The constraining consists of analyzing the HVSR curves and looking for clear peaks to identify the fundamental site frequency f_0. Second, whenever the peak is clear, we compare it to the peaks observed on TF_{1DSH}. Then, restriction criterion on the realizations consists of computing the median of all perturbed responses that show a close peak to those of HVSR, with dropping-off all that do not. Figure 4.4a and b show each of the TF and BTF compared to the response of their baseline profiles. The new 1DSH amplification show for the sites considered a shift of f_0 towards lower frequencies, which led to a high attenuation beyond 10 Hz. However, when no close peaks (to HVSR f_0) are obtained whatever the realization, we assume that the site is far enough from a 1D structure. Consequently, we consider by default the median of the 30 generated realizations. In this way, we ensure that we are improving the 1DSH prediction, using the available site signatures, only for those who can be strong candidates for a 1D classification.

The comparison between numerical and empirical approaches is indeed acceptable only if the site response is essentially one-dimensional. Thompson et al. (2012) proposed to select 1D sites with reliable velocity profiles by measuring the fit between SSR_{dh} and BTF_{1DSH} with Pearson's correlation coefficient (r_p). Pearson's linear coefficient measures the closeness of the spectral shapes of any two transfer functions. It was also used by Laurendeau et al. (2018) to select a subset of 1D KiK-net sites and by Pilz and Cotton (2019) in an extensive classification of sites into 1D or 2D/3D categories. We implemented a similar approach to perform the statistical comparison. In our work, we compute coefficient r_p over the frequency range $[\max(0.5\text{ Hz}, 0.5f_{dest}), \min(15\text{ Hz}, 7f_{dest})]$, where f_{dest} is the fundamental frequency of destructive interferences and was evaluated from the HVSR of borehole sensors. In Figure 4.4b we show the r_p coefficient before and after considering vertical

heterogeneities. Sites with dominant 1D effects are expected to have an increase in the r_p value while sites that have additional effects are not expected to have such an improvement.

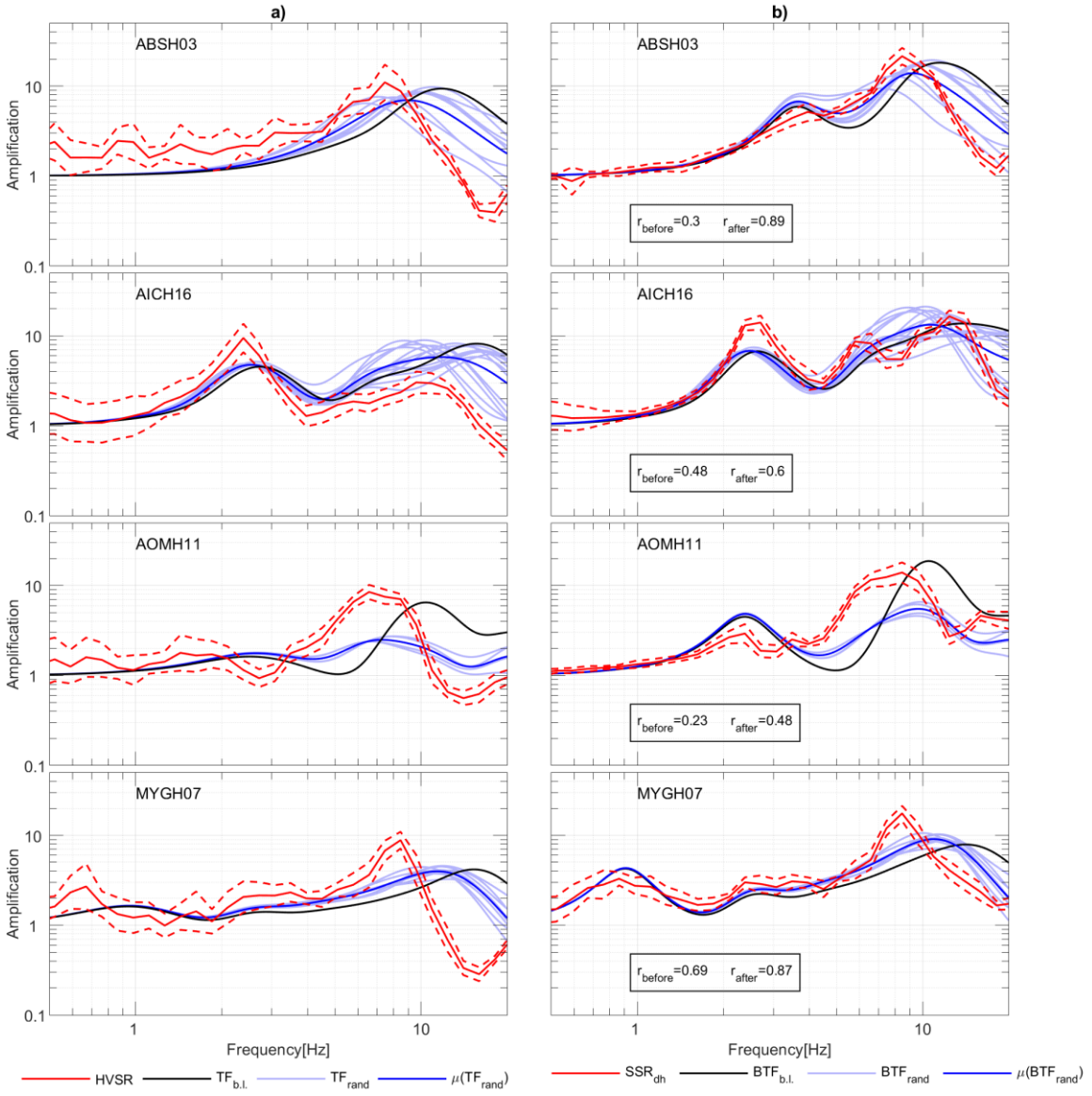

Figure 4.4: estimation of empirical and theoretical ratios for four KiK-nets a) HVSR is compared with baseline ($TF_{b.l.}$), 15 randomized (TF_{rand}) and median of randomized 1DSH site-responses b) comparisons of SSR_{dh} and BTF similar to (a). Pearson's correlation coefficient, used to verify the 1D assumption, is shown before and after including heterogeneities in the V_S-profiles.

4.3.3 Scales of the estimated functions

We show different site response estimations in a schematic representation (Figure 4.5) and associate each of the amplification terms with a scale of two levels, absolute and relative. By definition, site terms are relative if represented in terms of a ratio with a specified reference site term. These site terms become absolute terms if they are relative to an amplification-free site, i.e., representing the proper amplification. Though assumptions can vary from one estimation method to another, we believe that absolute terms of site effects are the ones that are essentially needed in SHA studies and mainly in our work. Indeed, absolute estimations of site responses are needed in precise applications, such as in the deconvolution approach (Laurendeau et al., 2018).

The 1DSH analyses estimate the absolute amplification at the surface (or any predefined level) using the available site information (V_S models). Though the 1DSH amplification can lead to absolute terms, it is limited to the considered site structure beneath the site, as shown in Figure 4.5. In this case, all site-specific information that is out of the explored V_S scale will not be included in the numerical simulations. We will define this scale in the text as the local scale, related to the amplification acquired by waves between the V_S profile depth and the surface. Thus, any additional or residual site effects can not appear in 1DSH estimations. To sum up, absolute site responses can represent local or global scales.

On the other hand, relative site responses represent the amplification compared to the motion at a chosen station, regardless of the absolute amplification exhibited by the site. This relative concept is mainly used on KiK-net arrays to compare and validate estimations from empirical and 1DSH estimations (SSR_{dh} and BTF_{1DSH}, respectively) as in several studies (Cadet et al., 2012; Thompson et al., 2012; Laurendeau et al., 2018).

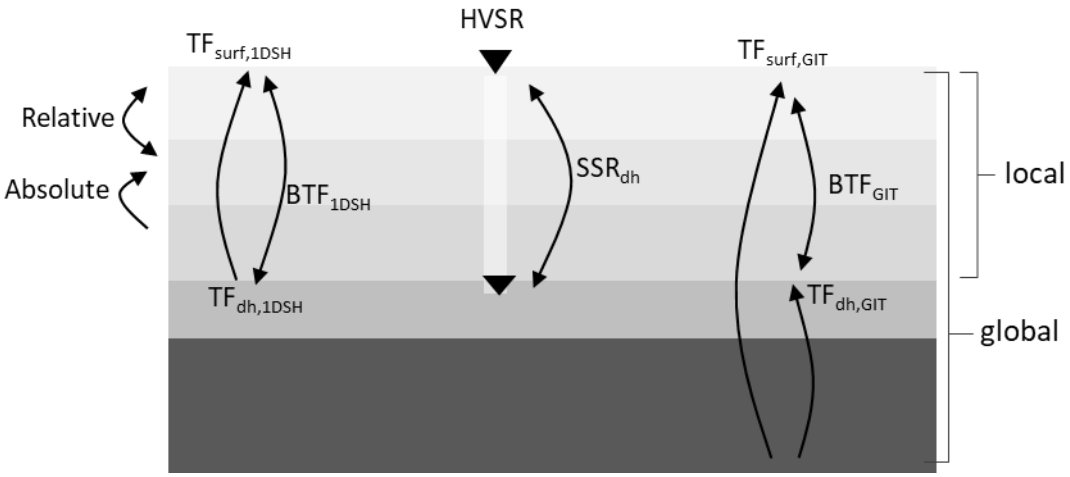

Figure 4.5: A schematic representation to differentiate between the different scales of the estimated transfer functions.

Eventually, we believe that discussions on absolute site amplification should pass from local to global conditions—the latter means including all the existing site effects, not only local ones. Empirical estimations of amplification serve as an essential tool to estimate the amplification without many assumptions, allowing the exploration of a global scale. Initially, HVSR (Nakamura, 1989) was intended to estimate the global amplification, but several studies consequently showed that vertical components could bias the amplification estimation (Haghshenas et al., 2008). The SSR techniques regarding a reference outcropping rock are the standard empirical approaches to estimate the correct amplification. An additional empirical estimation of site effects can be generalized inversion techniques (GIT), the main focus of the current study. GIT estimations could be considered empirical since they are based on the separation of the FAS into the source, path, and site contributions without prior strong assumptions.

GIT delivers non-parametric frequency-dependent site amplification, which permits to take site-effect analyses from local to global conditions, conditioned by the presence of enough recordings, with less effort to search for reference outcropping rocks. The traditional practice in GIT applications was to fix one or several sites as a reference, imposing their site-response to 1 (or median to 1) for all frequencies. Thus, the latter

practice limits analyses to relative site responses. On the contrary, we believe in going towards absolute global scales after fixing the correct response (with a certain level of confidence) of one or several stations as a reference in GIT inversions. Indeed, this step requires a knowledge of one or several site responses from other methods before inversions. Once correctly performed, the GIT application allows overcoming the lack of site characterizations, even if temporarily. However, the need for site-specific information for one or several stations can not be excluded whatever the studied dataset. This practice is expected to deliver absolute global site effects and avoid the impact of unaccounted features on the results.

4.4 Spectral inversion techniques application

4.4.1 Highlights on the inversion method

Generalized inversion schemes initiate from the principle of separation of the amplitude spectrum of ground motion as in Eq(4.2). FAS_{ij} is the observed Fourier amplitude spectrum at a given site for a given event, $E_j(f)$ is the source function for a given event j, $P_{ij}(f)$ is the path contribution upon an event j at a given site i, and $S_i(f)$ is the site response term for site i. Each of these terms are assumed independent of each other for a given spectrum.

$$FAS_{ij}(f)=E_j(f).P_{ij}(r,f).S_i(f) \quad (4.2)$$

Applying the logarithm on Eq(4.2), we get the linearized equation Eq(4.3). The linearized form provides a system of linear equations of the form $A.x = b$, where b is the data vector, x is the solution of the system, and A describes the system matrix. Following Andrews (1986) and Castro et al. (1990), two undetermined degrees of freedom need to be constrained on any two of the three terms of the equation. Often, the first is done by constraining the site response of one or several sites, while the second is related to the attenuation definition. The inversion scheme to solve the system can follow a non-parametric or a parametric approach.

$$\log_{10}(FAS_{ij}(f)) = \log_{10}(E_j(f)) + \log_{10}(A_{ij}(r,f)) + \log_{10}(S_i(f)) \quad (4.3)$$

The non-parametric approach (Castro et al., 1990; Bindi et al., 2009; Oth et al., 2011) provides the attenuation terms as a simple function of distance and frequency $A(r,f)$. Thus, the non-parametric attenuation is considered to be including all effects (geometrical spreading and intrinsic attenuation). In this case, the attenuation-related degree of freedom is resolved by assuming minimal effects near the source (i.e., equal 1). However, recordings at very short distances from earthquakes are not always guaranteed in common strong-motion databases (less than 5 km). Thus, attenuation is often constrained to one at the first distance bin R_0 where data is present (i.e., $A(r = R_0, \forall f) = 1$), also named as the reference distance. On the other hand, source terms are considered as shifted sources to R_0 though no prior assumptions are imposed on their spectral shapes. As a consequence of the attenuation constraint, inverted source spectra from non-parametric methods need to be corrected through the rescaling by R_0 to be interpreted correctly.

The parametric inversion schemes intend to use *a priori* models in the inversion equation. As a consequence, the system can become highly nonlinear. The same (or similar) models are also used in the post-treatment of the non-parametric results to

evaluate the physical parameters. For the source terms, it standard to use Brune's model (Brune, 1970), as in Eq(4.4).

$$E(f) = \left(\frac{2R_{\theta\phi}}{4\pi\rho R_0 \beta^3}\right)\left(\frac{(2\pi f)^2 . M_0}{1 + (\frac{f}{f_c})^2}\right) \quad (4.4)$$

Here, M_0 is the seismic moment and f_c the corner frequency of earthquake, $R_{\theta\phi}$ the source radiation pattern assumed to be constant ($R_{\theta\varphi}$= 0.55 for S-waves, Boore and Boatwright 1984). ρ is the density, β is the S-wave velocity of the medium at the source, and v_s is the average S-wave velocity along the path, assuming that $\beta = v_s$ =3.5 km/s and ρ =2800 kg/m^3 as taken in several studies (e.g., Drouet et al., 2010; Bindi et al., 2017). The values of the stress drops are also determined by Brune's model, as in Eq(4.5).

$$\Delta\sigma = \frac{7}{16} M_0 \left(\frac{f_c}{0.37 v_s}\right)^3 \quad (4.5)$$

The path terms, accounting for crustal attenuation effects, usually involve anelastic decay and geometrical spreading. Therefore, non-parametric attenuation functions can be fitted using the model in Eq(4.6).

$$A(r, f) = \exp\left(-\frac{\pi f (R - R_0)}{Q_0 f^{\alpha} v_s}\right) . G(R) \quad (4.6)$$

R_0 is the reference distance used, α is the frequency-dependence of the quality factor $Q = Q_0 f^{\alpha}$. $G(R)$ is the coefficient for the geometrical spreading and can be considered a multi-linear piece-wise function (e.g., Boore, 2003; Bindi et al., 2017) as in Eq(4.7). If we consider two break distances R_1 and R_2, there will be three geometrical spreading factors in each distance segment, i.e., γ_1, γ_2 and γ_3.

$$G(R) = \begin{cases} \left(\frac{R_0}{R}\right)^{\gamma 1} \; for \; R \leq R_1 \\ \left(\frac{R_0}{R_1}\right)^{\gamma 1} . \left(\frac{R_1}{R}\right)^{\gamma 2} for \; R_1 < R \leq \; R_2 \\ \left(\frac{R_0}{R_1}\right)^{\gamma 1} . \left(\frac{R_1}{R_2}\right)^{\gamma 2} . \left(\frac{R_2}{R}\right)^{\gamma 3} for \; R > \; R_2 \end{cases} \quad (4.7)$$

The parametric inversion scheme becomes highly nonlinear, and alternatives to the linear system solution should be used. For instance, Gauss-newton linearization of the parametric scheme was considered by Drouet et al. (2010) as an iterative solution. This approach tends to minimize the misfit between an initial model and the observed spectra. The Drouet et al. (2010) approach comprises defining *a priori* covariance matrices on each of the input parameters used to initialize the iterations. Thus, any known event or site information can be easily constrained in the inversions by decreasing the corresponding covariance values. This parametric approach was also considered in Grendas et al. (2018) and further optimized in Grendas et al. (2021) on a Greek database.

The choice of the approach type (non-parametric vs. parametric) was not straightforward after the results of the generalized inversions benchmark, GITEC. In

GITEC, a detailed comparison was conducted between the two GIT approaches and showed that parametric approaches with *a priori* models could sometimes underly strong assumptions. Ground motion can have additional complexities unaccounted for in the used models. The latter can have a substantial impact on site effects evaluation, the main focus of the study. Therefore, we decided to apply both approaches in this study.

The inversions are generally carried out, assuming that the dataset under study has no significant lateral variations of seismic attenuation properties. However, neglecting strong later contrasts can lead to strong biases in the inverted terms. If we want to consider regional differences, modifications to the methods should be injected. We do not intend to develop a new method of regionalized inversions at this stage, even though we are interested in accounting for regional effects. That is why we applied an approach similar to Oth et al. (2011) to regionalize attenuation effects. This approach relies mainly on predefining attenuation zones and deals with records based on their attributed regions.

4.4.2 Non-parametric attenuation:

At KiK-net sites, we had the advantage of obtaining attenuation functions either from the surface, borehole, or both. In principle, the only difference between surface and downhole arrays is the local site-effects acquired by seismic waves between the bedrock and ground surface. Consequently, attenuation effects should be the same whether they are obtained from the surface or downhole arrays since any additional local attenuation should be included in site effects. In the end, we decided to determine attenuation terms uniquely using the non-parametric approach to avoid “blind” model fitting of attenuation characteristics.

After the regionalization performed earlier, we consider surface stations to determine attenuation functions for each region. Note that trial inversions were performed in each region, and we observed stability of attenuation functions whatever the reference site in the region. Besides, attenuation was found to be robust and invariant whether obtained from the surface, borehole, or both, which presents an agreement with Oth et al. (2011). Since we apply a non-parametric approach for attenuation, data availability in different distance bins is essential for the estimation, especially in the short distances. Usually, R_0 could be chosen either as the first distance bin or at few distance bins beyond to have more data available. The regionalization performed earlier led to zones of different data distributions. Thus, R_0 was chosen for each region depending on the data recorded at short distances.

A one-step inversion was performed to determine attenuation terms. Figure 4.6 shows the results obtained for the whole of Japan and for each region. Visual inspection of attenuation functions shows that attenuation has similar effects for all frequencies before 80 km. Significant changes in decay appear after 80 km, mainly related to the geometrical spreading of seismic waves. The geometrical spreading controls the shape of the attenuation curves and appears to be frequency-dependent. Low frequencies seem to be more affected by the spreading decay with almost two main break distances around 80 and 130 km. Here, we define the breakpoints of geometrical spreading manually, and then the model fitting is done to obtain the physical attenuation parameters (e.g., Q-values). Table 4.1 and

Table 4.2 summarize the results of post-fitting, done in the least-square sense, with an attenuation model that considers both geometrical spreading and anelastic exponential decay (4.6).

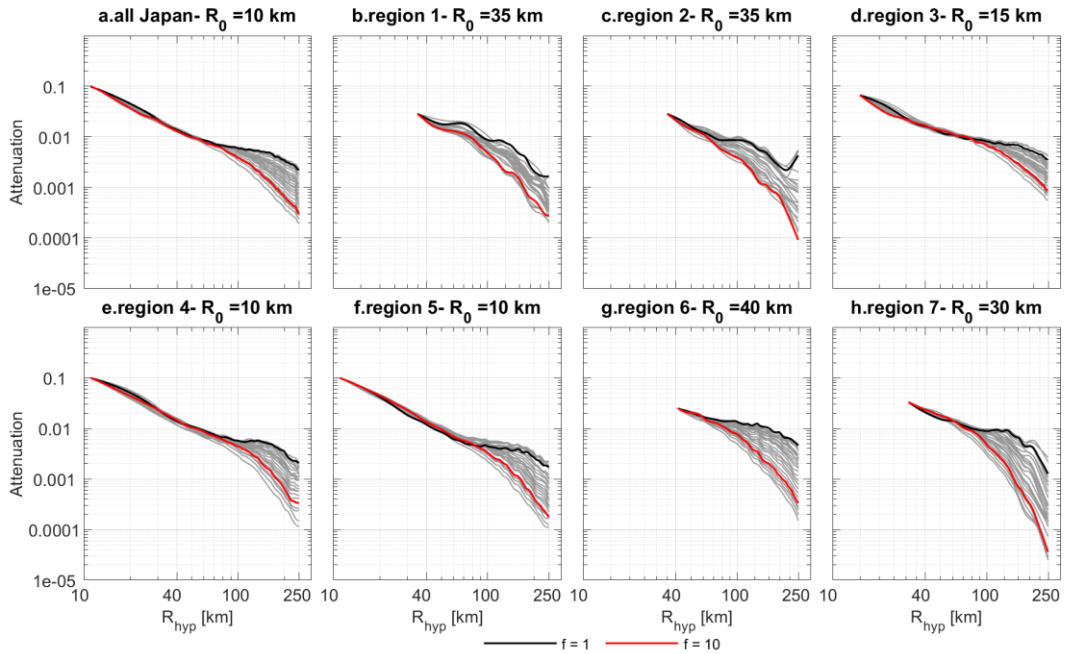

Figure 4.6: Non-parametric attenuation terms obtained from 1-step inversion for the entire of Japan and the different attenuation zones. The curves corresponding to 1 and 10 Hz are highlighted. No attenuation information could be delivered for distances smaller than the reference distances considered in each region.

Rapid analysis of the physical attenuation parameters could be done to validate the results. The break of the geometrical spreading decay is usually related to waves of long wavelengths that reach the Moho layer and bounces towards the surface. From the observed attenuation results, it is often seen that two main breakpoints can be identified with the commonly explored distance range 0-300 km. Theoretically, seismic waves start the decay in a pattern close to spherical spreading, which lets us expect a value of γ near unity. Afterward, the decay is expected to follow a cylindrical spreading leading to a decrease of γ value. However, additional complexities in the propagation medium and the Moho reflections could intervene in the spreading pattern. The best choice of the breakpoints to fit the model was identified by trial and error iterations. The results show that γ generally takes large values in first and third distance ranges, while lower values are observed in the intermediate distance between the two breakpoints.

Table 4.1: results of parametrization of the attenuation terms for the different regions obtaining Q_0 and α values.

region	Q_0	α
1	95.92±20.68	0.93±*0.24*
2	71.43±32.39	1.05±*0.35*
3	221.36±6.52	0.53±*0.09*
4	158.36±25.64	0.71±*0.21*
5	221.26±8.56	0.37±*0.10*
6	219.40±89.06	0.48±*0.34*
7	161.85±24.74	0.52±*0.22*

Table 4.2: Results of segmented parametrization for geometrical spreading factor by the piece-wise linear function. R_{max} and R_0 are the maximum distance range in the region and reference distance. respectively.

region	***Breakpoint(s) (km)***	R_0 ***(km)***	R_{max} ***(km)***	γ_1	γ_2	γ_3
1	70	35.00	180.00	1.44	0.18	-
2	80 – 150	35.00	180.00	1.04	0.33	1.47
3	60 – 160	15.00	180.00	1.20	0.07	0.78
4	80 – 140	5.00	180.00	1.00	0.09	0.76
5	70 – 150	5.00	180.00	1.08	0.00	0.53
6	80 – 140	40.00	180.00	0.60	0.66	0.97
7	70 – 130	30.00	180.00	0.83	0.77	2.60

As for anelastic attenuation factors, we ensured that model fitting is applied to a common distance range for all regions to have consistent and representative comparisons. Since crustal attenuation properties could examine lateral and vertical variations, including large distance ranges could reflect some information of deeper layers traversed by seismic rays. The latter reason stands behind fixing the same distance ranges (in different regions) to post-fit attenuation terms as Q_0 quality factor. As shown in the tables, Regions 1 and 2 show low Q_0 values. This result suggests high attenuation properties (below 100) and shows consistency with the recent tomographic studies, which we used to subdivide Japanese regions. Regions 3, 5, and 6, situated on one side of the volcanic arc, weak attenuation (high Q_0 values) while region 4 (situated to the opposite side of the arc) carries high mean attenuation. Region 7 (around the Kyushu region) shows an expected contrast with region 6. ~~However, if we had divided region 4 into two zones, we could have seen that northern parts carry higher attenuation than northern parts.~~ In the end, a perfect regionalization of Japan in GIT inversions remains complicated. Therefore, we consider that the current regionalization could be acceptable as it confirms more or less the results of attenuation from other studies.

4.5 Source and site separation with GIT

Before separating the site-response and source spectra, we rescale the non-parametric attenuation functions to the same reference distance (10 km) for all regions. Then, we correct surface and borehole spectra for the obtained attenuation effects to separate the source and site terms (Castro et al., 1990; Oth et al., 2011). In this way, we ensure, to some extent, that the source and site separation is unbiased by the lateral crust contrasts. This separation is also formulated in a system of equations, i.e., in Eq(4.8), which only accounts for source and site terms.

$$\log_{10}(FAS_{ij}(f)/A_{ij}(r,f)) = \log_{10}(E_j(f)) + \log_{10}(S_i(f)) \quad (4.8)$$

Also, we apply another separation using the parametric approach of Grendas et al. (2021) by fixing initial attenuation parameters to have minimal effects. Then we compare the results of source and site terms from the inversion approach.

This source-site separation is applied overall to the KiK-net network, first (1) for surface and borehole data simultaneously and second (2) for surface data only. The fact of including surface and borehole stations allows obtaining surface and borehole transfer functions ($TF_{surf,GIT}$ and $TF_{dh,GIT}$ respectively). Consequently, we can compute the GIT

borehole transfer function (BTF_{GIT}), defined as the ratio $TF_{surf,GIT}/TF_{dh,GIT}$, which is a relative term. On the other hand, inversions on surface data only could provide the absolute terms, $TF_{surf,GIT}$.

4.5.1 Reference site choice

Due to the unresolved degree of freedom, a reference condition on site responses is necessary for generalized inversions. The choice of this reference site is a crucial step for inversion results and should be taken carefully. The reference site question has been responded to differently throughout different studies. For example, several studies were imposing the average of all stations to unity to identify flat responses, then fixing the best flat and rock site response(s) to be amplification-free (Drouet et al., 2010; Oth et al., 2011; Bindi et al., 2017). In the absence of any site information, the latter strategy for reference choice is unavoidable, and it is the only solution to solve the inverse problem. However, rock sites could present some amplification at high frequencies caused by shallow weathered layers at the surface (Hollender et al., 2018). Besides, as outcropping hard-rock sites can be found in mountainous areas, additional factors could play an important in site response estimation, such as topographic effects. Thus, fixing an amplification-free site response for rock sites could lead to misestimations.

Furthermore, the study of Nakano et al. (2015) proposed to fix the response of a rock station for which 1D transfer function was estimated with the help of geophysical information available. If site information is already delivered and allows numerical simulations of surface transfer functions, it could be an essential step to obtain absolute site responses from GIT. Before proceeding with this step, several questions should be explored. First, the measured V_S profiles from different methods are subject to uncertainties (Teague et al., 2018; Passeri et al., 2019). Also, vertical 1DSH analyses could neglect some additional effects, such as multi-dimensional or topographic effects. Alternatively, if we decide to use empirical estimations such as the SSR method, a good reference site should be chosen, which is not easy for the KiK-net dataset.

We believe that constraining inversions with absolute correct reference conditions is crucial in GIT to ensure proper separation of source, attenuation, and site terms. Any improper fixing of a site response could lead to a significant bias on source terms. Also, we should always accept that we will not completely eliminate the uncertainties on absolute site terms. This study proposes two strategies to choose the reference site inversions, which could be transposed to other datasets.

The first reference choice is based on identifying a hard-rock site in the dataset, the YMGH06 station, characterized by a V_{S30} value of 2200 m/s. The spectral form of the amplification of YMGH06 is inspected through trial inversions. For example, an inversion is done with each of the parametric and non-parametric inversions after setting the average response of all sites equal to one (i.e., Ref_{ALL} condition). This inversion generally allows examining site responses with a minimum impact by the reference choice. Figure 4.7a shows the responses of the station YMGH06 from different approaches, empirical (GIT and HVSR) and theoretical 1DSH. Note here that GIT_n and GIT_p stands for nonparametric and parametric respectively. Though 1DSH response predicted a flat response, all empirical approaches (HVSR and GIT) proposed a flat response for lower frequencies that drops for higher frequencies,

mainly beyond 7-8 Hz. This discrepancy could be related to unaccounted effects in the 1DSH simulations. Therefore, we kept the same spectral form obtained from GIT, rescaled it to have low frequencies at unity, and used it as a constraint for YMGH06 in the final inversions. We denoted this reference choice by Ref1. One motivation for this choice is that hard-rock sites are less susceptible to variabilities and uncertainties that could be more expressed in soft soil sites. Besides, any discrepancy between theoretical and empirical approaches could be identified and taken into account as we proceeded (without detailed analyses of the origins of these aspects).

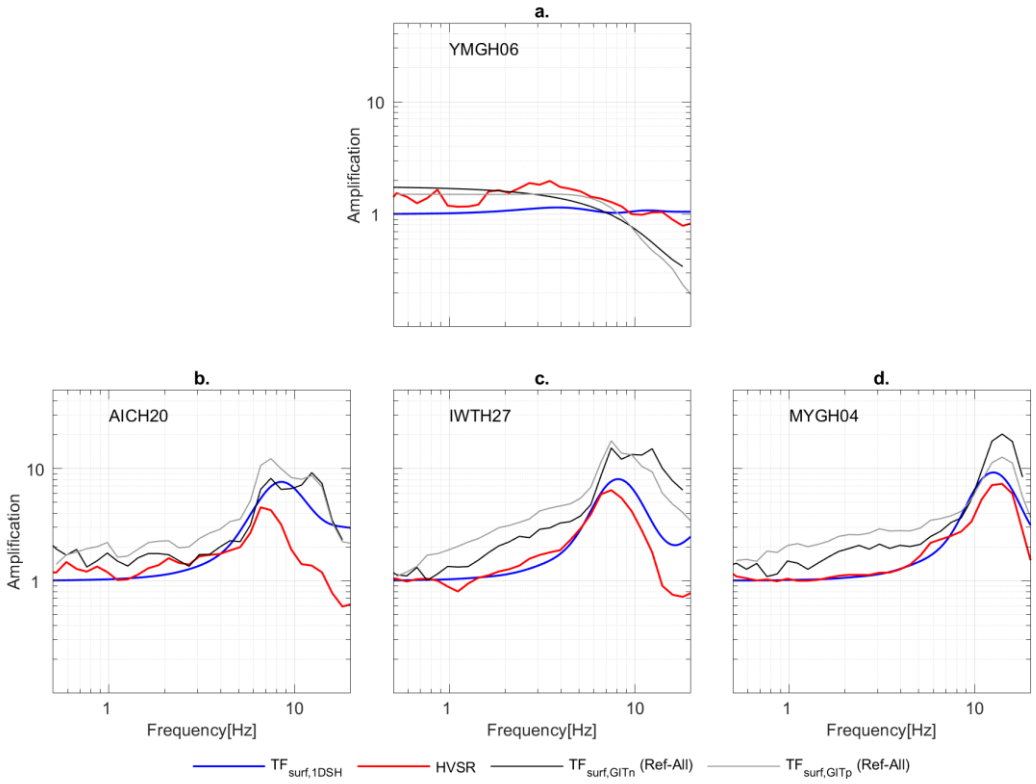

Figure 4.7: Comparison of site-responses obtained from a trial inversion with fixing the average of all site-responses to unity. The sites shown are the hard-rock YMGH06 used as Ref1 and the three stations AICH20, IWTH27, and MYGH04 used as Ref2.

The second choice of reference condition is made in a different way, which we denote Ref2. Here, we aimed to select sites that can be considered the best 1D sites with the help of Pearson's correlation coefficient, $r_p(\mathrm{SSR_{dh}}, \mathrm{BTF_{1DSH}})$. The last step was essential to have a valid 1D estimation. We then focused on stations with V_S profiles down to depths at which V_{Sdh} exceeds 2000 m/s. This condition ensures that the estimated 1D response is as "global" as possible. Then, trial inversions are performed by variating the reference conditions to identify the most stable inverted site terms. Among these sites, we selected three sites that showed a spectral form of amplification similar to that of 1DSH prediction (see Figure 4.7b,c, and d). To account for uncertainty for the fixed responses, limited flexibility was allowed to the constraint. Eventually, this second strategy is transposable to other datasets conditioned by the presence of site characterizations of only a few sites, not necessarily all.

4.5.2 Source results

Source terms were obtained with no prior assumptions from non-parametric GIT, while they are obtained directly by imposing Brune's model (Brune, 1970) for the parametric

GIT. However, we post-fit the non-parametric terms with the same source models to deliver consistent comparisons. Inversely, we reconstruct spectra for parametric schemes using Brune's model obtained parameters. In the end, we can compare parametric and non-parametric terms.

Figure 4.8 shows the results of non-parametric source spectra using reference choices Ref1 and Ref2. Mainly, the source spectra have a consistent overall average for all frequencies. Indeed, a drop in the high frequency is observed in the non-parametric inversion results, especially for the large magnitudes. Source spectra between the reference conditions chosen appear to be stable for the parametric approach. Simultaneously, slight changes in the spectral form are observed for the non-parametric inversions, especially in high frequencies. This ensures that the reference condition impacts the final source term estimations in the non-parametric scheme.

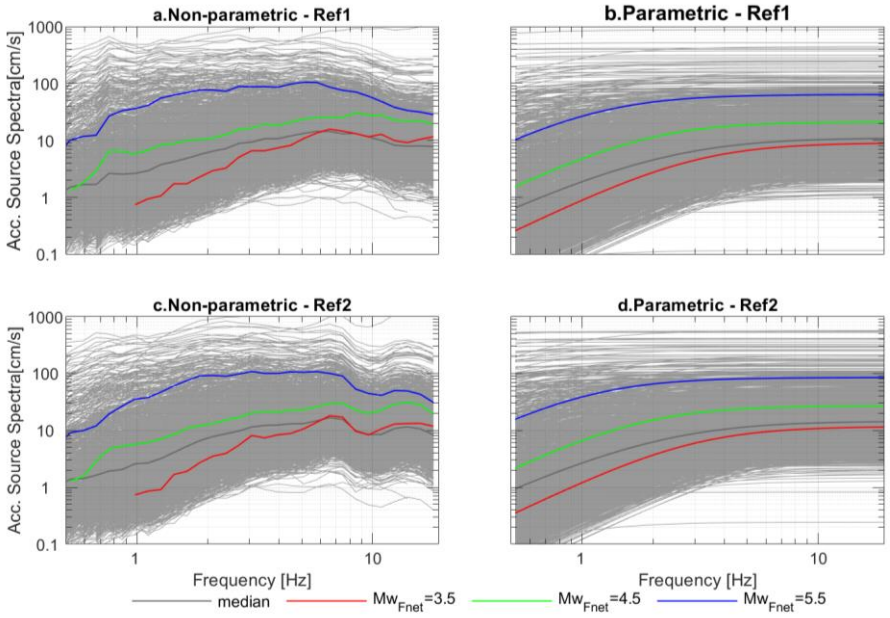

Figure 4.8: Non-parametric source spectra obtained from parametric and non-parametric GIT approaches. Source spectra of three events are highlighted. Besides, the results for the two reference conditions are displayed. The median source spectrum is plotted for each case.

Since the physical parameters could better reflect the differences, we compare those obtained from the two GIT schemes for Ref1 and Ref2. Figure 4.9a and b show the comparison of obtained corner frequencies f_c versus inverted magnitudes M_w. From comparisons between GITn and GITp results, we can see that some events, especially for higher magnitudes than 5, the non-parametric approach can not resolve for Brune's model. This issue is due to the fact that f_c does not lie anymore within the considered frequency bandwidth. The only solution to fix this problem is to enlarge the bandwidth, if possible, to capture higher and lower frequencies or fix input M_w of events in which we have confidence. Since the GITp defines *a priori* search intervals for M_w, the search range remains centered around predefined magnitudes. Figure 4.9c and d show the estimations of stress drops ($\Delta\sigma$) from both GIT_p and GIT_n. The average $\Delta\sigma$ obtained from different approaches remains consistent and ranges between 1 to 2 MPa (10 to 20 bars). These results of $\Delta\sigma$ are consistent with previous findings in several studies (Oth et al., 2011, 2015; Nakano et al., 2015).

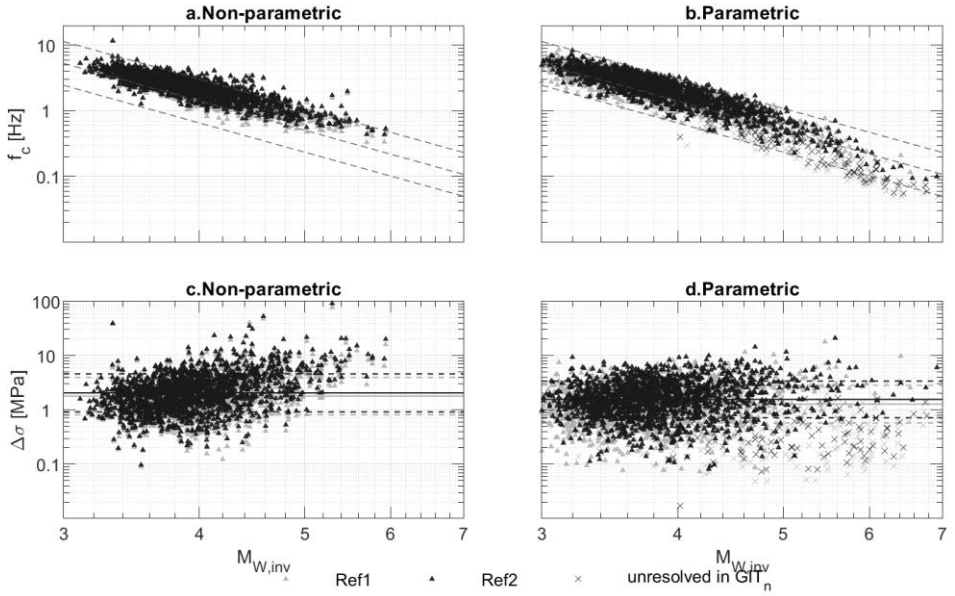

Figure 4.9: Source parameter estimated from parametric and non-parametric inversions for each reference condition. On subplots a and b, the dashed lines represent the stress drop constants for 0.1, 1, and 10 MPa. In c and d, the average stress drop is displayed for each reference condition case. Cross marks highlight unresolved source parameters in GIT_n without fixing input magnitudes.

4.5.3 Site results and some case studies

The fact of including surface and downhole data in the inversions allowed obtaining the relative transfer functions BTF_{GIT}. Figure 4.10 shows comparisons of BTF_{GIT}, BTF_{1DSH}, and SSR_{dh} at stations six different sites. These comparisons at different stations show the level of consistency between 1DSH and the empirical estimations. The sites, shown in Figure 4.10, have different r_p(BTF_{1DSH} and SSR_{dh}) values, from a poor to a strong fit. On the contrary, perfect consistency is observed when comparing empirical estimations (BTF_{GIT} and SSR_{dh}).

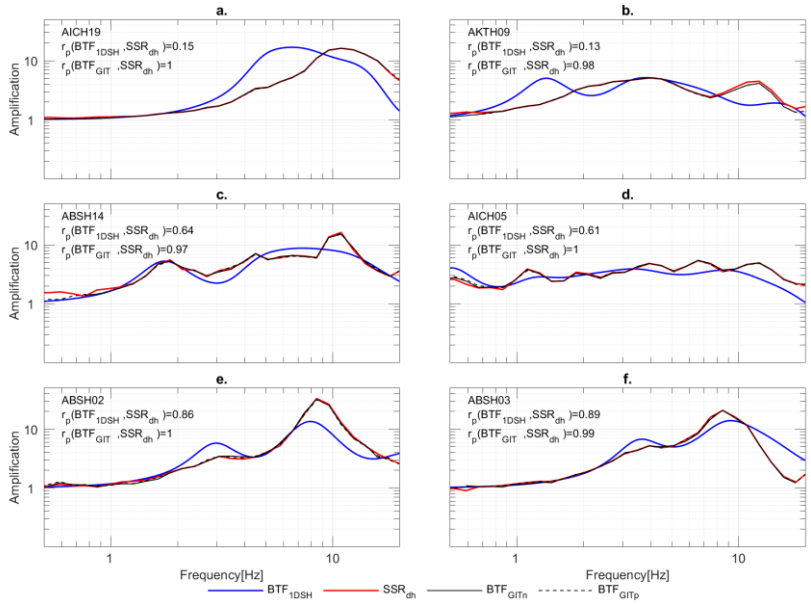

Figure 4.10: Comparisons of Standard-spectral ratios with borehole site as reference (SSR_{dh}) and Borehole transfer function (BTF_{1DSH}) obtained from 1DSH numerical modeling and Generalized inversions (BTF_{GIT}) at six stations. The goodness of the fit of empirical SSR_{dh} with each transfer function from GIT and 1D simulation is computed with Pearson's correlation coefficient r_p.

To have an overall view of the r_p estimation, we present the distribution of its values in Figure 4.11. These distributions show that 1DSH relative transfer functions can have both good and poor fits with empirical ones, while GIT relative terms mostly show high r_p. Applying the commonly used threshold ($r_p \geq 0.6$), we identified 300 sites that could be dominated by 1D effects and consequently classified as a 1D site. However, the 1D classification is based only on relative terms in this case.

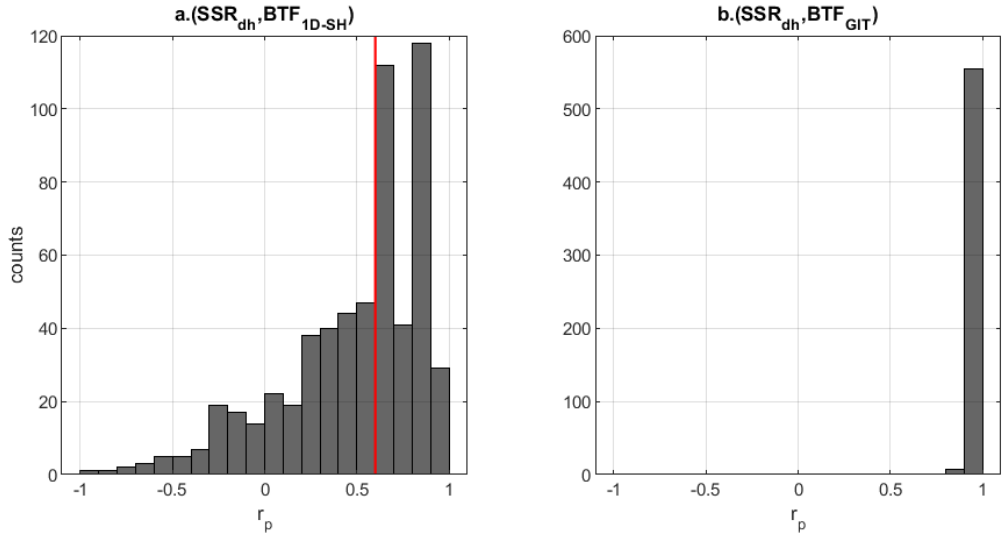

Figure 4.11: Pearson's correlation coefficient distribution between SSR_{dh} on one side and BTF_{1DSH-} and BTF_{GIT} on the other.

4.6 Discussions

4.6.1 Revising the 1D classification with absolute terms

After introducing the relative and absolute scales of transfer functions in a previous section, we discuss some consequences of considering absolute terms. The 1D site identification of Thompson et al. (2012) was first proposed on relative site terms of KiK-net because there was not yet a clear estimation method for absolute site terms. Here, we propose to estimate, with GIT, reliable absolute site terms by fixing the proper site reference condition. Consequently, we revised the 1D classification by comparing absolute terms of theoretical 1DSH and empirical GIT (i.e., $TF_{surf,1DSH,}$ and $TF_{surf,GIT}$).

For this aim, we now consider surface data only, with the conditions Ref1 and Ref2. The results of surface transfer functions are shown in Figure 4.12 for the same six sites presented earlier (in Figure 4.10). After observing that GIT_n and GIT_p lead to very similar values of r_p, we only display their average on the plots (i.e., $r_p(TF_{surf,1DSH},TF_{surf,GIT})$). We can notice from the comparison that there is an evolution of r_p values between absolute and relative terms. AICH19 and AKTH09 sites had low r_p from relative terms, while AICH19 shows an improvement despite the differences found. ABSH14 and AICH05 sites show a drop towards a poor fit when comparing absolute terms. Indeed, the two sites ABSH02 and ABSH03 show a good fit with both relative and absolute terms comparisons.

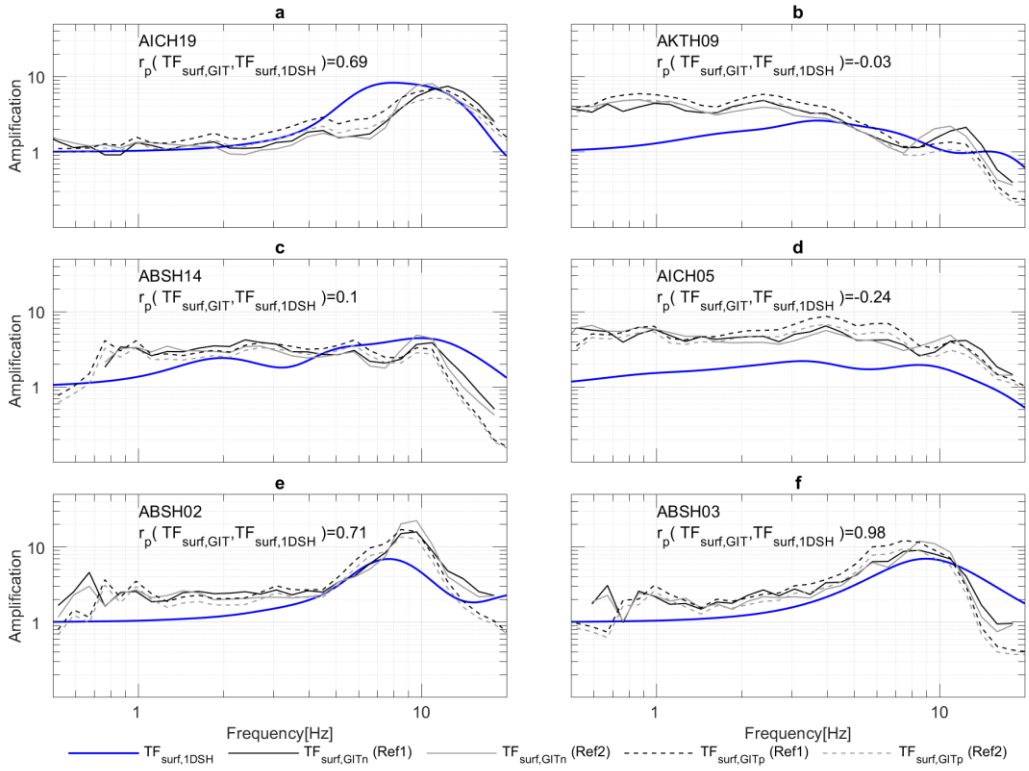

Figure 4.12: Absolute surface site-responses at six KiK-net sites obtained from 1DSH, GIT_n, and GIT_p approaches. The Pearson's correlation coefficient r_p is re-estimated for absolute terms.

We again conduct the statistics on r_p using the absolute terms in Figure 4.13a, where we show a comparison between r_p(BTF_{1DSH},BTF_{GIT}) and r_p($TF_{surf,1DSH}$,$TF_{surf,GIT}$). Around 110 sites were classified initially as 1D sites using relative terms appears not to be the case anymore when absolute terms. On the contrary, 97 sites that were not initially classified as 1D sites appear to have more comparable absolute 1DSH estimations with empirical responses. Alternatively, 190 sites showed a stable 1D classification when considering relative or absolute site responses. Figure 4.13b the distribution of new r_p values with respect to V_{S30} values. The latter shows that the very hard-rock sites (i.e., V_{S30}>1500 m/s) can not be classified as 1D in the "absolute" criteria.

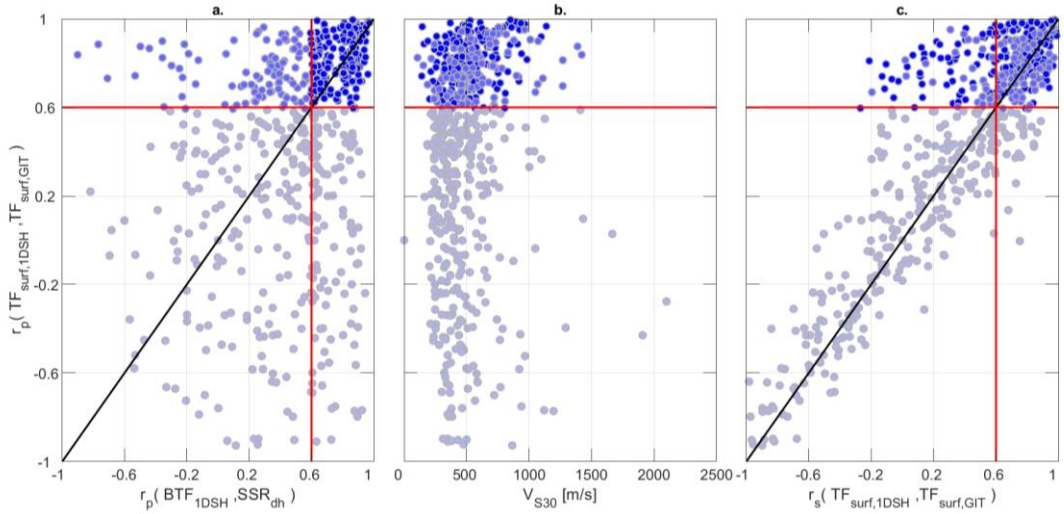

Figure 4.13: a) comparison of correlation coefficient obtained from absolute and relative site functions. b) distribution with V_{S30} of r_p. c) comparisons of r_p with Spearman's rank correlation coefficient r_s. The red lines represent the correlation threshold commonly used to classify one-dimensional site effects.

The observed changes in r_p values between relative ($r_{p,rel}$) and absolute ($r_{p,abs}$) estimations can be interpreted according to their values. However, these interpretations are conditioned by a high precision of V_S profiles.

- If $r_{p,rel}$ >0.6, then 1D effects are dominant up to the borehole level. In this case:
    - If $r_{p,abs}$ >0.6, then 1D effects are dominant on a global scale. Also, this can ensure that V_S used in the estimations were precise enough to capture real site effects. Consequently, all residual site effects between global and local scales are minimized.
    - If $r_{p,abs}$<0.6, it could result from V_S profiles that were not deep enough to capture global effects.
- If $r_{p,rel}$<0.6, then the site structure between surface and borehole are far from one-dimensional conditions. However, in our estimations, we could find $r_{p,abs}$ >0.6 for some sites with $r_{p,rel}$<0.6. This case could originate from several factors, i.e., oblique wave incidence, downhole transfer functions, and others. In the end, to avoid any ambiguity, we discarded all sites with $r_{p,rel}$<0.6.

An additional correlation coefficient could be added to the discussion, Spearman's rank correlation coefficient (r_s). It was also used by Zhu et al. (2020) in absolute site response analyses. Spearman's coefficient reflects how well a monotonic relation can describe the comparison between two transfer functions. Spearman's correlation between two variables is equal to Pearson's correlation between the rank values of those variables. Since r_p could be more sensitive to outliers than r_s, it is interesting to identify among stable 1D sites (between relative and absolute terms) those who also have a good r_s value. Figure 4.13c shows a comparison on absolute terms of r_p and r_s for all sites. If we consider that an acceptable r_s value should be greater than 0.6, we find around 203 sites acceptable for both r_s and r_p using absolute terms. If we constrain the 1D list using relative and absolute terms, we result in only 133 sites.

Finally, these 133 sites could be considered to be dominated potentially with 1D effects since they result from the intersection of correlation coefficients on absolute and relative terms, using r_p and r_s. A complete list of correlation coefficients estimated for the selected 1D sites is provided in an appendix.

4.6.2 Mean comparisons of relative and absolute site responses

Another essential part of the statistical analyses of site responses is the median comparison of site responses from different approaches. Figure 4.14 shows the medians of relative and absolute site responses for the final 1D sites. For the relative terms, we obtain consistent comparisons of medians as expected from site-by-site comparisons. Since 1D simulations probably include only local effects limited to the measured V_S profile, the relative term comparisons (surface to borehole) work fine. On both surface and borehole transfer functions, any additional effects from the (unaccounted) global scale will be systematically removed with ratios, leading to consistent medians. It is also worth mentioning that high-frequency discrepancies previously observed by Laurendeau et al. (2018) on median estimates are not observed here. This is mainly due to considering small-scale heterogeneities in the 1DSH site response computations.

For the absolute surface terms, the median from empirical approaches shows some differences with theoretical 1DSH. Slight differences of the medians are observed from GIT_n and GIT_p with conditions Ref1 and Ref2, limited to a factor of 30 percent at most. This level of GIT variability of site response was also indicated in the outcomes of the GITEC benchmark. However, we believe that the correct choice of reference station will play a key role in the robust estimation of absolute site responses, which is necessary to achieve our ultimate objective (i.e., robust estimation of site responses to use for the deconvolution approach).

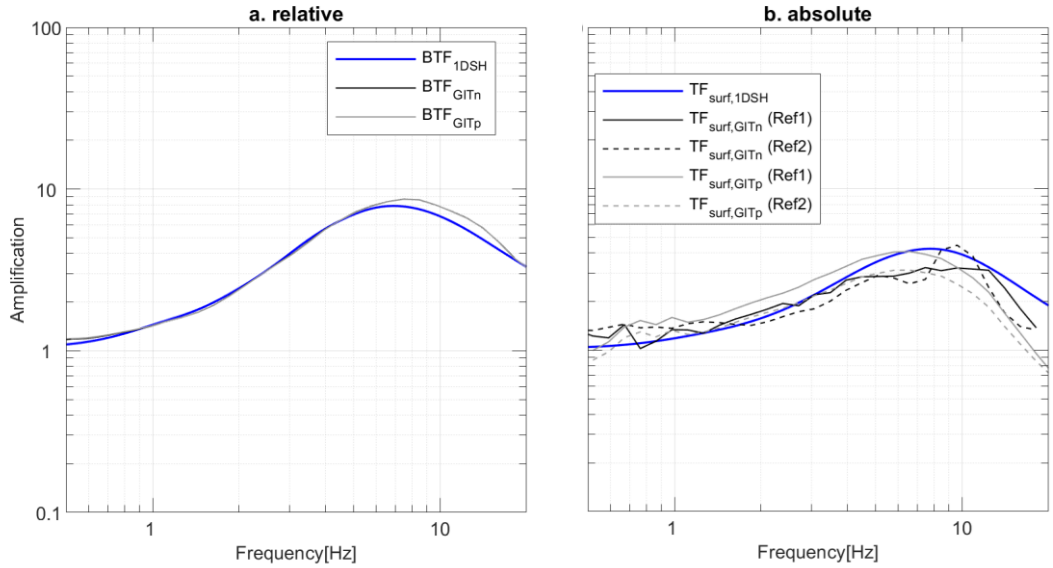

Figure 4.14: a) Average Relative term comparison (for identified 1D site list): mean site-responses of all sites considered BTF_{1DSH} and BTF_{GIT}. b) Comparisons of the mean surface transfer function obtained from GIT inversions TF_{surf} from 1DSH, GIT_n, and GIT_p (Ref1 and Ref2).

4.6.3 Site response estimations using GIT

This work was mainly conducted with the spirit of transposing it to other datasets, where we have less available site-related information. Therefore, the three main advantages of the current work could be illustrated as follows.

- First, the estimation of 1DSH site responses with the available V_S profiles underlies the one-dimensional assumption. This assumption is generally relaxed with empirical estimates of site responses, and more precisely, with GIT estimations. The only constraint needed for inversions is the appropriate reference condition. Eventually, the use of GIT leads us to estimate site responses even in the case of the presence of multi-dimensional effects.
- We do not have coupled sensors (surface-borehole) and V_S profiles down to borehole level in other common datasets. In the KiK-net network, these two features allow validating the use of the 1DSH approach by comparing it to empirical estimates. The use of GIT empirical estimates overcomes the absence (or the non-completeness) of these features in other networks.
- Despite having V_S profiles in the KiK-net data, additional sources of uncertainties control the final 1DSH response. These uncertainties could originate from unaccounted features near the soil surface, and thus, they can

be reduced when considering empirical methods, like GIT. This leads to avoid several questions and rely on empirical estimations of site responses.

4.7 Conclusions

The primary purpose of this study is to estimate site responses by applying generalized inversions. We aimed to extend the site effects determination beyond some limitations, first on the KiK-net network. The primary tool to do so is the generalized spectral inversion techniques (GIT), which can deliver all amplification functions for all sites, not only the 1D.

Since site responses are not the only parameters inverted in GIT, we must carefully perform inversions for reliable estimations of all associated terms. For this reason, we performed a regionalization of the Japanese KiK-net network with the help of several attenuation tomographic studies. We estimated attenuation characteristics for the different regions and found consistency with the recent studies. Then, we followed a regionalization approach similar to Oth et al. (2011), where we correct the dataset for attenuation effects. Afterward, we performed a source and site separation with the help of parametric and non-parametric inversions. The results led to consistent estimations for source parameters, which ensured a proper separation. The key reason for the success of this source-site separation was the proper choice of the reference condition. We tested two strategies to fix reference stations, needed by GIT inversion to solve the system, that lead to consistent estimations. An expected variability was observed in the order of 20 to 30 percent.

In the end, we proposed to revise the 1D site identification by applying the correlation coefficients directly on absolute terms (surface transfer functions) and not only on relative terms (surface-to-borehole ratios). This step highlighted only 133 out of 585 sites (~ 23%) with dominant one-dimensional effects.

Chapter 5: Ground Motion Model for Hard Rock Sites by Surface Recordings Correction

Context

This chapter is entirely dedicated to applying the deconvolution approach to surface recordings after estimating site responses. This chapter illustrates the main steps to reach reference motion GMPE. First, the KiK-net data is the same as that of the previous chapter, with some additional preparations and selections. Then, we make sure that the deconvolution approach works well on several recordings before application to all data. In the end, we derive reference motion that appears to no more include site effects. Also, we derive an additional reference motion GMPE that applies for all site conditions, non-restricted to 1D sites. In the end, deconvolved GMPE with GIT represents reference motion, in the sense that no site effects remain after deconvolution.

5.1 Introduction

Site-specific seismic hazard assessment relies on the definition of reference motion and integration of site effects into hazard predictions. Researchers always consider hard-rock motion as reference motion since hard-rock sites are the closest to bedrock conditions. However, the amount of recorded data on such rigid sites is insufficient to obtain ground motion prediction equations (GMPEs) representing such conditions.

To overcome the lack of data on hard-rock sites, the state-of-the-practice is to derive GMPEs for standard rock conditions present in most databases, then calibrate them with correction factors to account for hard-rock conditions. The last correction is commonly called “Host-to-target-adjustments”, HTTA. Several aspects, drawbacks, and uncertainties associated with HTTA applications were thoroughly detailed in previous chapters. Several alternatives to the HTTA were also detailed, mainly the deconvolution approach first applied by Laurendeau et al. (2018).

The deconvolution approach mainly consists of removing site effects from surface recordings to have amplification-free reference motion. Hence, the applicability of this approach is conditioned by the prior site effects estimates at each site. Laurendeau et al. (2018) first applied the deconvolution approach by using 1DSH site responses, and thus they needed to limit the analyses to 1D sites. In fact, in this text, we propose to apply the deconvolution approach with site responses from an empirical approach, namely from the generalized inversions techniques (GIT). However, we remind that continuing the work on KiK-net data is motivated by two main reasons. The first is to keep a comparison with the results of the approach of Laurendeau et al. (2018) from a methodological point of view. The second reason is that we benefited from KiK-net site characterizations to draw a clear methodology to constrain one or several reference sites in GIT inversions (the work detailed in the previous chapter).

The main objective of this chapter is to use the estimated site effects of the previous chapter in the deconvolution approach to obtain reference ground motion. Thus, we perform the deconvolution of surface recordings with theoretical site responses and GIT site responses. Since we aim to expand the application of deconvolution to other datasets on which we can not easily identify 1D sites, we find interest in constructing two datasets of surface recordings. The first one is composed of sites from the 1D list

(defined in chapter 4), and the second contains no restrictions on site conditions. Afterward, we derive a new GMPE suitable for hard-rock motion using the deconvolution approach. Residual analysis is conducted to assess GMPE model validity first and then the effectiveness of the deconvolution approach. Reference motion obtained from GIT-corrected datasets presents a significant drop in the total variability, mainly due to the drop of site-related variability. Comparisons in terms of rock to hard-rock ratios are presented at the end.

5.2 KiK-net dataset explored

In this study, we carry on with the same KiK-net subset prepared in the previous chapter. Thus all the data selection criteria are inherited (selection of magnitudes, distances, etc.). To remind, recordings of shallow crustal earthquakes were selected (following information from Dawood et al., 2016, and Bahrampouri and Rodriguez-Marek, 2019), automatically picked for P- and S-onset wave arrivals, and checked for Signal-to-noise ratio in Fourier domain (the details were already presented in the previous chapter). To fulfill the ultimate objective of the current study (to derive a hard-rock ground motion model), we apply some additional selection conditions to the data:

- We consider recordings with SNR >3 over [0.5-20Hz]. In consequence, no holes are allowed over the considered bandwidth.
- We keep sites with estimated responses by GIT over [0.5-20Hz].
- A minimum V_{SDH} was imposed to 1500 m/s so that all the deconvolved data correspond to hard-rock data.

After identifying a list of KiK-net sites where one-dimensional effects are dominant (using the Pearson's correlation coefficient between empirical and theoretical site-response estimations), we name these sites as 1D sites. First, the definition of a 1D site list allows having an applicable estimation of site effects from 1DSH simulations. Besides, this can allow us to capture further the impact of multi-dimensional effects on predicted ground motion and their variability.

After fixing the minimum number of recordings per site and per earthquake in each data subset to 3, Table 5.1 shows the statistics on the dataset prepared for the number of recordings, sites, and earthquakes. In a quick comparison to data prepared and used in Laurendeau et al. (2018), we have much more data recorded on 1D sites (7904) instead of 1000 recordings only. This increase in the number of recordings is due to expanding the dataset until 2017. Though the identified list of 1D sites contained 133 sites, the restrictions we imposed (mentioned earlier in this section) to develop the deconvolved GMPE lead to retaining only 99 1D sites.

Figure 5.1 shows data distribution for $SURF_{1D}$ and $SURF_{ALL}$. F-net magnitudes ($M_{w,Fnet}$) range between 3.2 up to 7.3. Joyner and Boore distance metric (R_{JB}) for GMPE regressions was chosen (R_{JB}) to capture near-source saturation effects, with a distance less than 300km. Finally, site conditions at the surface are 300< V_{S30}< 2200 m/s with a median around 500 m/s, and at downhole stations 1500<V_{SDH}<4000 m/s with a median around 2500 m/s.

Table 5.1 Statistics on the selected recordings from the KiK-net dataset for GMPE regressions.

	conditions	$N_{recording}$	N_{sites}	N_{events}	abbreviation
Surface data 0	None	31592	575	1740	-
Surface data 1	$V_{S,DH}$ >1500 m/s	18591	319	1485	$SURF_{All}$
Surface data 2	1D sites and $V_{S,DH}$ >1500 m/s	7904	99	1034	$SURF_{1D}$

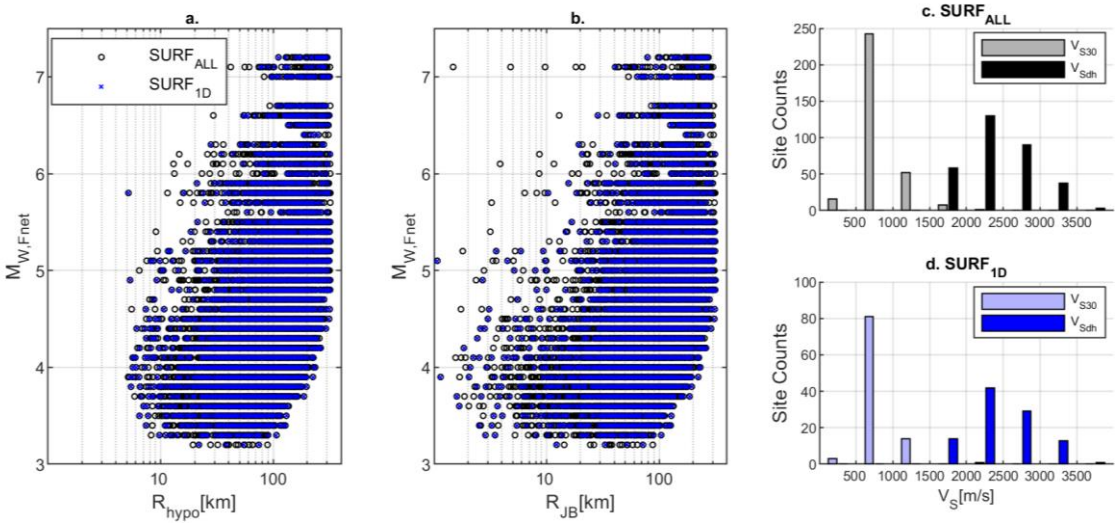

Figure 5.1: Data distribution for $M_{w,Fnet}$ versus hypocentral (a.), and JoynerBoore (b.) distance for each KiK-net subset. The distribution of sites V_{S30} and V_{SDH} is shown in c and d.

5.3 Deconvolution approach

The deconvolution procedure in the Fourier domain is aimed to remove site effects from surface recordings. This step should result in an amplification-free dataset conditioned by a robust estimation of site effects. The corrected surface dataset is expected then to represent reference rock conditions at the bedrock level. An essential element and a crucial step for this procedure is site amplification estimation. LA18 first presented this approach using 1DSH simulations, benefiting from site characterizations of KiK-net sites and limiting the analysis to identified 1D sites. In this work, we use several site amplification estimations intending to expand the methodology over the 1D limitation.

In the previous chapter, regionalized GIT was applied to the database using a parametric and a non-parametric approach with several choices of the reference stations (Ref1 and Ref2). Comparisons at the end of the chapter showed consistency between estimated site responses, with variability that does not exceed 20 percent (also following GITEC conclusions, Chapter 3). We selected one site-response estimation per inversion type (non-parametric or parametric) and per reference choice (REF1) to use these estimations.

Figure 5.2 shows the main steps of deconvolution of surface recordings using a 1DSH estimation at an arbitrarily chosen site, MYGH11, for several events of different $M_{w,Fnet}$. Surface time traces are transformed into the Fourier domain, where the deconvolution is performed. Amplifications estimated from either 1DSH, GIT_n, or GIT_p are removed from the FAS, and then site-corrected time histories are retrieved. In the end, pseudo acceleration response spectra (5%-damped) are calculated for the corrected data to obtain a new database. For this, we attribute the suffix “cor” to the

dataset name followed by the method of site-response estimation in this text. In summary, $SURF_{1D}$ undergoes corrections to obtain $SURF_{1D}cor_{1DSH}$, $SURF_{1D}cor_{GITn,}$ and $SURF_{1D}cor_{GITp}$, while from $SURF_{ALL}$ we could only obtain $SURF_{ALL}cor_{GITn}$ and $SURF_{ALL}cor_{GITp}$.

Since this concept of deconvolution is a newly proposed approach to estimate reference motion, we make a series of checks on final data distribution to avoid any bias. The first check was done on hundreds of deconvolved time histories and showed that removing site amplification from FAS leads to an expected decrease in signal amplitudes. An additional check was done in Figure 5.3, where we visually inspected the PSA distribution, magnitude scaling, and distance scaling compared to $SURF_{1D}$ data. Despite the amplitude drops at higher frequencies, no other significant changes impacting ground motion scaling were observed.

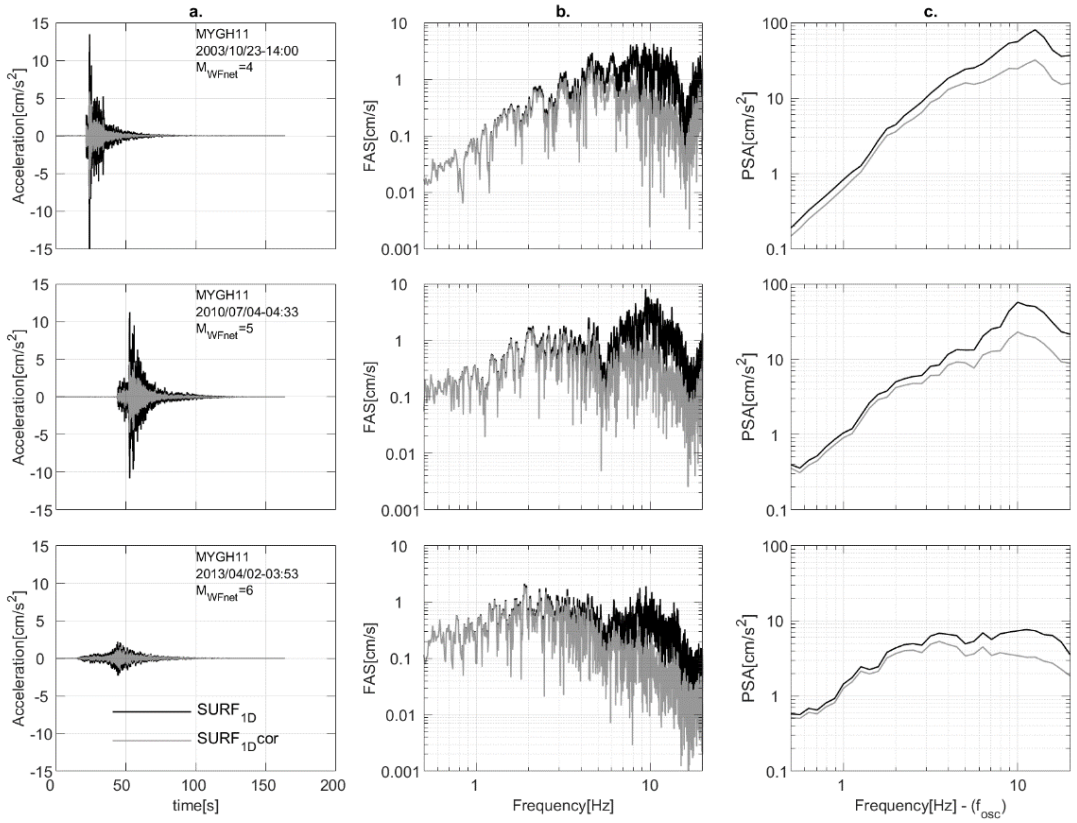

Figure 5.2: Deconvolution procedure main steps. Surface time traces (a) designated as $SURF_{1D}$ are transformed into the frequency domain (b). The correction with estimated site-response is performed in the Fourier domain before using the inverse Fourier transform to obtain corrected time trace (a) before computing the SDOF oscillator response (c). In this example, the site response was obtained by 1DSH simulation. The choice of one estimation (1DSH) is just to simplify the example illustration.

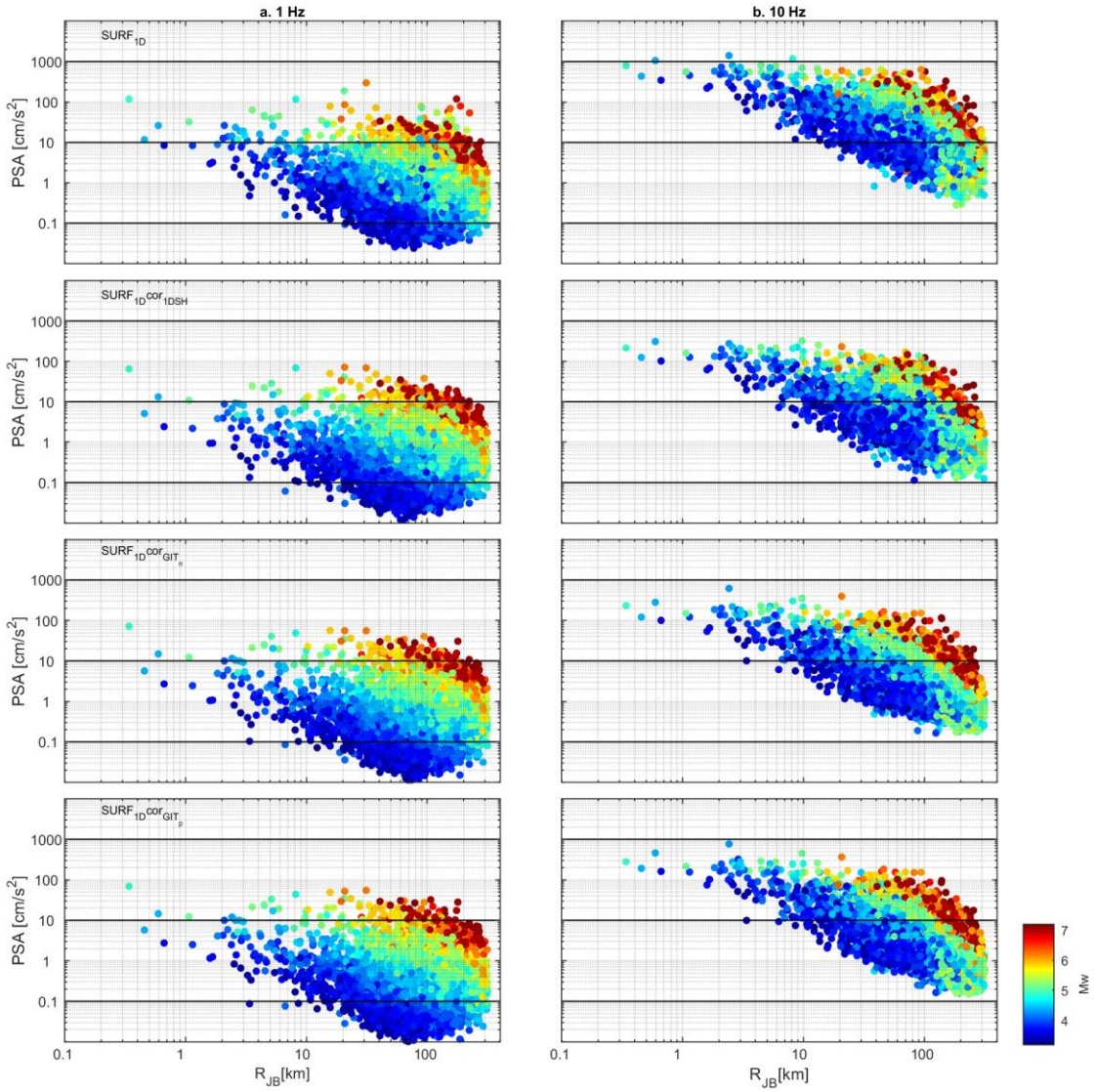

Figure 5.3: Data intensity distribution in terms of PSA to distance (R_{JB}) for each of the four defined 1D subsets, at two oscillator frequencies 1 and 10 Hz. the black lines are plotted to ease the visual comparison in PSA levels between different subsets. The color scale stands for the magnitude M_w.

5.4 GMPE functional form and first results

We apply mixed-effects regression provided within the "nlmer" package (Bates et al., 2015) to obtain a GMPE of the geometric mean of (5% damped) horizontal PSAs at 30 frequency points between 0.5 and 20 Hz (periods between 0.05 s and 2 s). We consider the general functional form as in Eq(5.1).

$$\ln\big(SA(f)\big) = F_m + F_d + F_s + \delta B_e + \delta S2S_s + \delta W_{o,es} \qquad (5.1)$$

The functions F_m, F_d and F_s are set to capture magnitude scaling, distance dependence, and site effects, respectively, in observed ground motion. In GMPE developments, residuals between model and data were first treated as random effects to obtain between-event residuals δB_e (Abrahamson and Youngs, 1992) aiming to improve model prediction. Later on, several motions were initiated to improve seismic hazard analyses towards non-ergodic assumption motivating the different decomposition of residuals related to path and site (Atik et al., 2010; Rodriguez-Marek et al., 2013; Kotha et al., 2016). In this text, we adopt decomposing the event-corrected

residuals into site-to-site (between-site) residuals, treated as random effects within the regressions, and a residual term $\delta W_{o,es}$ describing the record-to-record variability. The frequency-dependent random effects and residuals follow the normal distributions as follows:

- $\delta B_e = N(0,\tau)$ where τ estimates the between-event standard deviation.
- $\delta S2S_s = N(0,\varphi_{S2S})$ where φ_{S2S} is the between-site standard deviation.
- $\delta W_{o,es} = N(0,\varphi_{SS})$ where φ_{SS} is the single-station standard deviation.

The total aleatory variability associated with the GMPE is estimated as $\sigma = \sqrt{\tau^2 + {\varphi_{S2S}}^2 + {\varphi_{SS}}^2}$. Note that we considered the same notations as in Rodriguez-Marek et al. (2013) for variability terms.

After trials and testing several functional forms in regressions on KiK-net data, a multi-step regression was performed (as in the work of Kotha et al., 2018). The last step helped avoid strong trade-offs between regression coefficients, mainly for large magnitude and distance ranges and a vast amount of recordings.

5.4.1 F_d scaling

The first step of regressions was to capture distance scaling of ground motion of Eq(5.1). Observations Figure 5.3 shows PSA distribution at two frequencies (1 and 10 Hz) considering a magnitude-dependent decay of ground motion. Besides, a near-source saturation can be observed mainly at high frequency.

Near-source saturation effects have been differently represented in several GMPE derivations (Zhao et al., 2006; Rodriguez-Marek et al., 2013; Atkinson et al., 2016). A usual step to account for Near-source saturation effects is to add an effective depth parameter h to the distance measured on the surface (R_{JB} or R_{RUP}). In this study, we introduce a frequency-dependent h parameter without magnitude dependence. However, we include magnitude dependence for the spectral decay shape as defined in Eq(5.2).

$$F_d = \begin{cases} F_{d1}(R) = \left[b_1(f) + b_2(f).(M_w - M_{ref})\right].\ln\left(\sqrt{{R_{JB}}^2 + h(f)^2}\right) \quad (R \le R_b) \\ F_{d2}(R) = F_{d1}(R_b) + b_3(f).\ln\left(\frac{R_{JB}}{R_b}\right) + b_4(f).\left(R_{hypo} - R_b\right) \quad (R > R_b) \end{cases} \tag{5.2}$$

Since the distance scaling appears to change around 90 km (Figure 5.3), we set a breaking point in the distance scaling at 90 km. This step was also motivated by GIT results in the previous chapter, where geometrical spreading decay showed a slope change between 80 km and 120 km. Anelastic attenuation will be captured from PSA at distances larger than 90 km. Once the coefficients b_1, b_2, b_3, b_4, and h are obtained, they are fixed in the regressions steps afterward.

5.4.2 F_m scaling

PSAs are corrected for the distance scaling from the previous step before conducting the following regression. Though additional uncertainties can enter into discussions, different works took fault mechanisms into account through frequency-dependent terms for strike-slip, normal, or reverse faulting (Zhao et al., 2006; Bindi et al., 2010, 2014). However, recent analyses have indicated a weak dependence of ground motion

on faulting style (Kotha et al., 2018; Traversa et al., 2020) and avoided fault mechanism parametrizations.

The magnitude scaling displayed in Figure 5.3 tends to saturate at higher magnitudes. However, there is not enough data in the short distance for clear observations. A piece-wise linear or quadratic polynomial is usually defined for the magnitude scaling, with one or several magnitude breakpoints often called hinge magnitudes (M_h). For example, for the NGA-West database Campbell and Bozorgnia (2014) considered multiple frequency-independent hinges at 4.5, 5.5, and 6.5 for a wide range of moment magnitudes. Then, Boore et al. (2014) used a transition of M_h from 5.5 for frequencies >10 Hz to 6.2 for frequencies < 2.5 Hz. In this work, we define a quadratic magnitude scaling up to a hinge magnitude M_h, beyond which a linear dependence is imposed as often taken in GMPE regressions (Rodriguez-Marek et al., 2011; Kotha et al., 2018). However, M_h was preferred to be frequency-dependent following observation of KiK-net data, varying linearly between 6.2 at 0.5 Hz to 5.5 at 10 Hz, then it is kept constant at 5.5 above 10 Hz. The magnitude scaling functional form is presented in Eq(3), where a1, a2, a3, and a4 are the fixed-effects regression coefficients.

$$F_m = a_1(f) + \begin{cases} a_2(f)(M_w - M_h) + a_3(f)(M_w - M_h)^2 for\ M_w < M_h \\ a_4(f)(M_w - M_h) \qquad\qquad\qquad for\ M_w \geq M_h \end{cases} \tag{5.3}$$

5.4.3 F_s scaling

For site-specific hazard studies, site-to-site variability obtained with the help of random effects were discussed as elements of ground-motion site-calibration instead of V_{S30} proxy (Bora et al., 2015; Kotha et al., 2016, 2018). This approach would improve the practice and helps in physical explanations of the variability between sites, as weak correlations are found for high-frequency site-specific residuals V_{S30} and the attenuation parameter k_0 (Kotha et al., 2018; Bard et al., 2020).

Here, a reference motion GMPE is developed, with direct usability in site-specific studies. Therefore, we keep the V_{S30} parameterization to identify any possible dependency of the reference motion on V_{S30}. Keeping the V_{S30} parameterization also allows computing soft-rock to hard-rock scaling ratios usually evaluated in terms of this site rigidity proxy (Houtte et al., 2011; Ktenidou and Abrahamson, 2016; Laurendeau et al., 2018).

On the other side, it would be interesting to see if any correlation exists between site-specific residuals ($\delta S2S_s$) and V_S-values of the bedrock, which should be principally close to V_{SDH}. So, we define a usual linear dependence on V_{S30} with a reference velocity at 1000 m/s as in Eq(5.4).

$$F_s = c_1(f) \ln\left(\frac{V_S}{1000}\right) \ ;\ where\ V_S = V_{S30}\ or\ V_{SDH} \tag{5.4}$$

It is essential to understand that deconvolution delivers reference motion that corresponds to virtual hard-rock sites. However, it is crucial to know and define these virtual sites' characteristics in GMPEs, especially in site parametrizations. In this aim, the nature of the transfer function used in the deconvolution determines the corresponding V_S value.

- For instance, deconvolution of 1DSH transfer functions delivers ground motion representative of the borehole levels. Thus, setting V_S to V_{SDH} is reasonable in this case.
- For deconvolution with GIT terms, the resulting ground motion represents the reference site's conditions constrained in inversions. Hence, if GIT terms were relative to one site, then deconvolution leads to a dataset that represents a single site condition. However, we focused in the previous chapter on defining absolute site terms to constrain GIT inversions. This step leads to a deconvolved motion representative of the reference bedrock motion. In other words, the dataset is expected to have no more site effects. Despite that, we set $V_S = V_{SDH}$ in the following, and verifications will be made by inspecting the site coefficient and residuals.

5.4.4 Residuals check

Mixed-effects effects regressions allowed decomposing residuals, being treated as random effects, into between-event, -site, and -record variability, contrary to full fixed-effects regressions. With the help of these regressions, repeated event-specific and site-specific residual effects could be captured. Several investigations have shown that including the mixed-effects in each step of regressions prevents well-recorded events or sites from biasing the regression coefficients (Stafford, 2014; Kotha et al., 2017), though these intermediate-step residuals are not the final estimate to be considered. For each record, the fixed-effects are fixed in the functional form and deducted from PSA values at each frequency to obtain residuals split as random effects into δB_e, $\delta S2S_s$, and $\delta W_{o,es}$.

To apply this method to poorly described datasets (surface data only, less site characterization), we include non-1D sites in the analyses. The multi-step regressions were carried on $SURF_{1D}$ and $SURF_{ALL}$ datasets separately, where we expect higher site variability for $SURF_{ALL}$ data.

Figure 5.4 and Figure 5.5 show the residuals' variation with $M_{W,Fnet}$, V_{S30}, and R_{JB} at three different frequency values (low, intermediate, and high). Usually, random effects are the parameters to judge if the model used is adequate to capture magnitude and distance scaling. If any remarkable trend in the residuals is observed, the functional form can be updated, or a new predictor variable is introduced.

δB_e from both $SURF_{1D}$ and $SURF_{ALL}$ have comparable limits, showing higher residuals for high frequencies than low frequencies. The absence of a clear difference between both datasets is expected as the only difference is the 1D-site selection. Then, we divide the magnitude range into six bins equally spaced, and we display the median residual with one standard deviation marge. The absence of a clear trend in δB_e indicates an efficient regression analysis. $\delta W_{o,es}$ residuals were also similar between both datasets with no clear trends depending on R_{JB}. The median with one standard deviation marge over equally spaced distance bins does not show any trends.

The site residuals $\delta S2S_s$ constitute one of the essential parts of the discussion in this work. Discussions can also be interesting if site-parameterizations were dropped (i.e., F_S dependence on V_{S30}) and site effects are captured by $\delta S2S_s$. However, we insist in this text on including V_{S30}-proxy several reasons, detailed afterward. Despite this site

parameterization, some significant trends could be observed in residuals. Figure 5.4 and Figure 5.5 show that the used linear dependence on V_{S30} is not enough at low frequency (0.5 Hz) for the very few sites present at V_{S30}<200 m/s.

As the 1D selection had excluded most of the rock sites at V_{S30}>1000 m/s in $SURF_{1D}$, $SURF_{ALL}$ included all sites of V_{S30} beyond 1000m/s. For $SURF_{ALL}$, we can see that $\delta S2S_s$ includes additional attenuation effects at high frequencies, making the linear dependence on V_{S30} not enough. Here, we do not claim to enter into deep physical explanations that could stand behind these observations, but at least we can say that V_{S30} cannot be the only parameter to predict precise reference motion on hard-rock, at least for high frequencies.

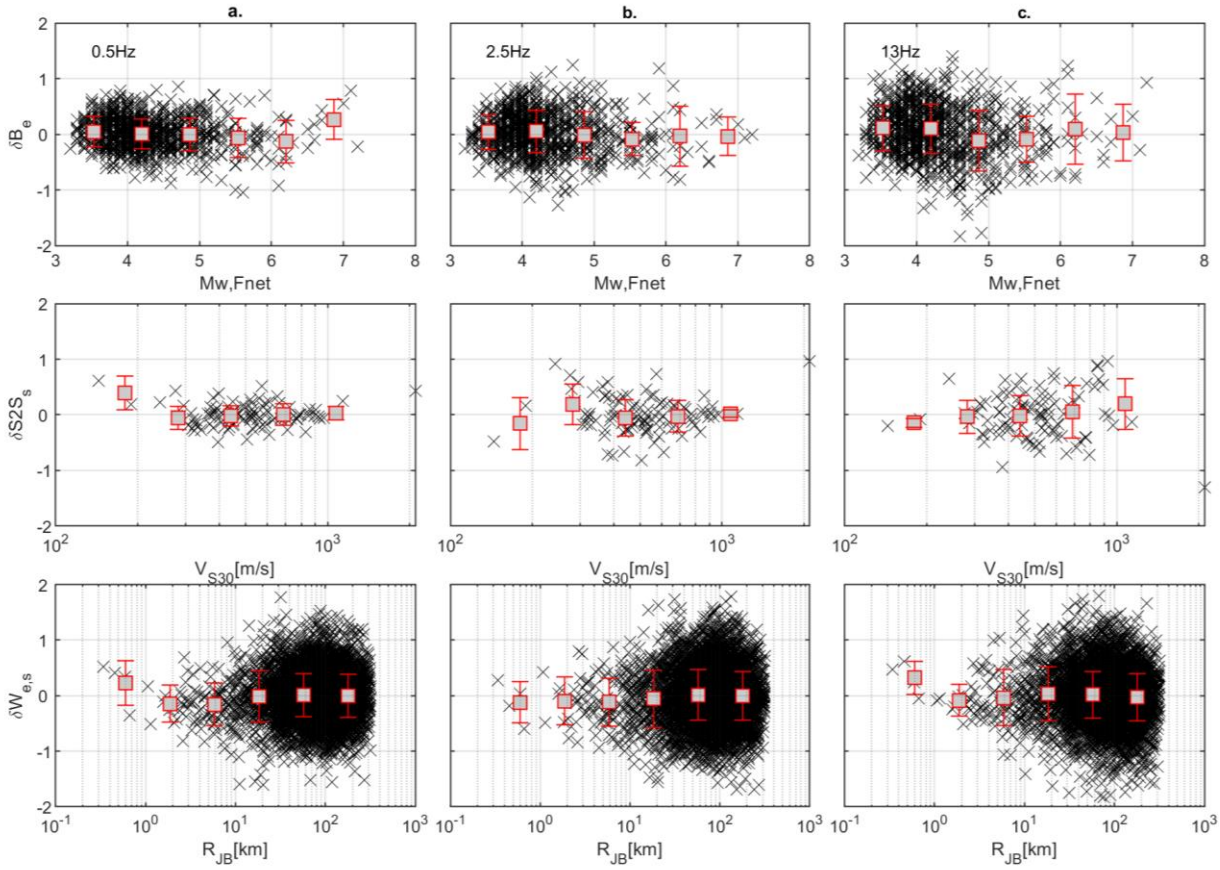

Figure 5.4: $SURF_{1D}$ residuals treated as random-effects at frequencies 0.5, 2.5, and 13 Hz. The upper row shows event residuals against $M_{W,Fnet}$. The middle row contains between site residuals against V_{S30}. The lower row contains record-to-record residuals against R_{JB}.

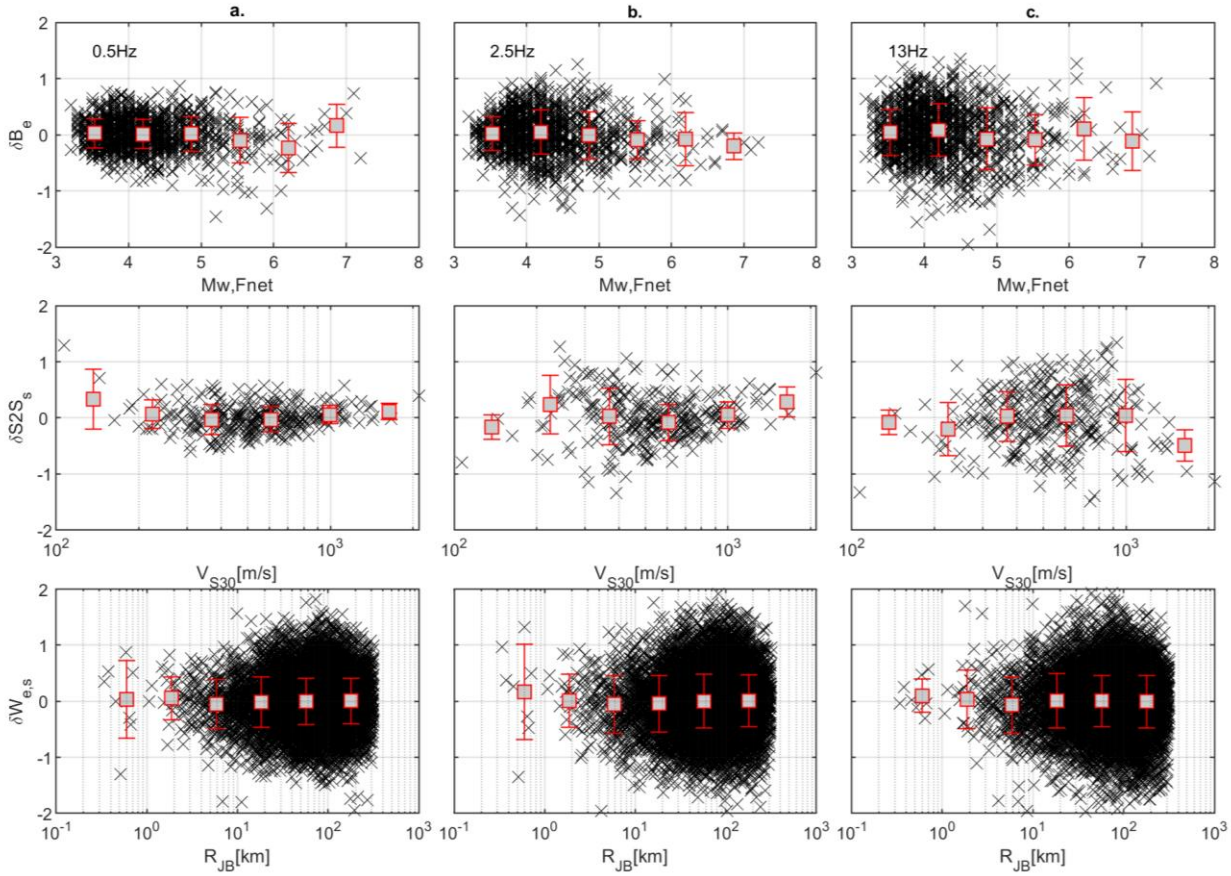

Figure 5.5: $SURF_{ALL}$ residuals treated as random-effects at frequencies 0.5, 2.5, and 13 Hz.

5.5 Results

5.5.1 Model scaling

After calibrating the regressions on surface data to best capture magnitude and distance scaling, we perform the same multi-step regressions on SURFcor datasets (i.e., $SURF_{1D}cor_{1DSH}$, $SURF_{1D}cor_{GITn}$, $SURF_{1D}cor_{GITp}$, $SURF_{ALL}cor_{GITn}$, and $SURF_{ALL}cor_{GITp}$. Since the major difference between different datasets defined is the site-effects, in Figure 5.6, we show the PSA values predicted versus the event- and site-corrected observations (i.e., PSA(f) - $\delta S2S_s$(f) - δB_e(f)) of one of the surface datasets, i.e., $SURF_{1D}$.

First, we can see that our model captures the distance scaling for all frequencies, especially at the breakpoint imposed at 90 km. We fixed in this distance-scaling model an exponential decay beyond 90 km (i.e., coefficient b_4) which fits well with observations. Besides, in the GIT results in the previous chapter, non-parametric attenuation terms showed a clear change in shape around 80 to 90 km that was modeled with geometrical spreading. This is in agreement with previous findings of Oth et al. (2011). Here, we consider a change in the geometrical spreading at R_b, but we introduce anelastic attenuation coefficients only after R_b as recently done in Kotha et al. (2018) for Japan. As for near-source saturation effects, the model presents slight differences in the PSA predictions compared to (Zhao et al., 2006; Kotha et al., 2017), mainly due to parameterization of effective depth h and its dependence on magnitudes.

Despite these slight differences, the GMPEs developed in this text serve well to our objectives.

Figure 5.7 shows the magnitude scaling of the derived GMPEs. Note that also, in these figures, PSA is corrected for site and event variabilities. It appears that for magnitudes < 6.2, the dependence (in the natural log) on $M_{w,Fnet}$ might better be described by a linear relation rather than quadratic since the concavity term a_3 obtained shows very low values for most frequencies. However, it is clear that we observe a change in the slope behind the hinge M_h. However, the main difference with other GMPEs derived for Japan lies in the near-source (short distance magnitude dependence). Unlike Kotha et al. (2018), the derived models do not show the same decrease of PSA for $M_{w,Fnet}$ >M_h for distances < 10 km. The latter may be due to the different accounting for near-source effects.

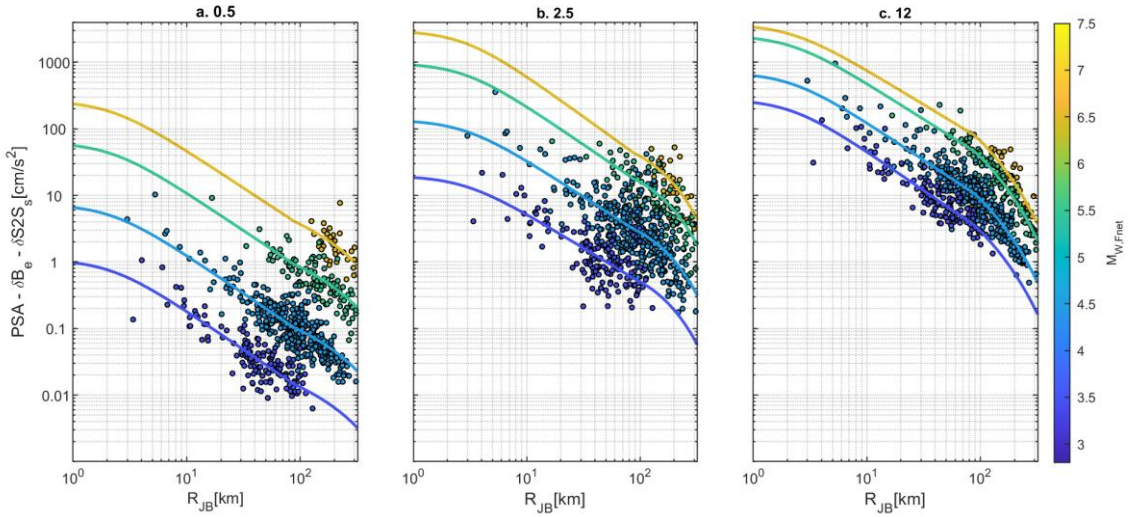

Figure 5.6: SURF$_{1D}$ GMPE. Distance scaling predicted PSA against the site- and event-corrected observed PSA at different magnitudes (3.5, 4.5, 5.5, and 6.5) at three frequencies 0.5, 2.5, and 12 Hz. The color scale stands for the predicted scenario magnitude $M_{w,Fnet}$.

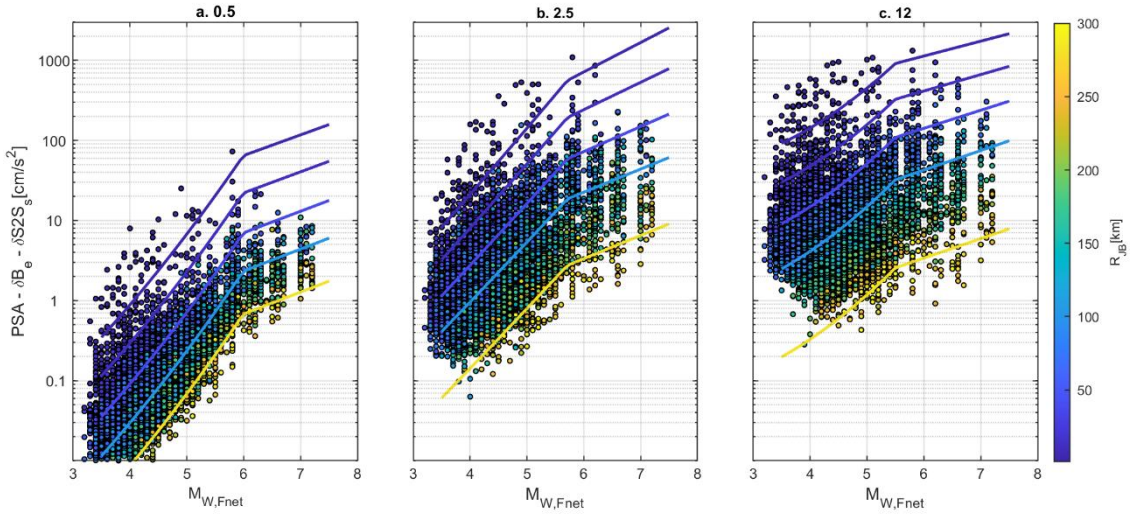

Figure 5.7: SURF$_{1D}$ GMPE. Magnitude scaling predicted PSA against the site- and event-corrected observed PSA at different distances (5, 15, 40, 110, 320 km) at three frequencies 0.5, 2.5, and 12 Hz.

5.5.2 Variability analysis between SURF$_{1D}$ GMPEs

As mentioned earlier, no significant differences in the results were observed for the event or path terms or residuals, even for the deconvolved datasets. This part

compares the residuals observed on 1D deconvolved datasets as seen in Figure 5.8, Figure 5.9, and Figure 5.10.

It is worth comparing the site residuals from different datasets. As corrected with 1DSH simulations, the $\delta S2S_s$ residuals in Figure 5.8 show a remaining variability that cannot be neglected, though no clear trends are observed in the $\delta S2S_s$ for low and intermediate frequencies. However, at high frequencies, a slight remaining decreasing trend in $\delta S2S_s$ residuals for V_{SDH}>3100 m/s (which correspond to the few sites of V_{S30}>1000 m/s shown in $SURF_{1D}$).

Then, if we exploit the $\delta S2S_s$ of $SURF_{1D}cor_{GITn}$ and $SURF_{1D}cor_{GITp}$ (Figure 5.9 and Figure 5.10), we find a significant drop in variability remaining after the deconvolution of site effects. As random-effects identify repeated residuals specific to sites, we suspect that estimations from the 1DSH approach, even on 1D sites, present a remaining misfit with empirical site responses. This could also be justified by the level of uncertainties accompanying theoretical response prediction and site information (Griffiths et al., 2016a; Teague et al., 2018).

Among the GIT-corrected datasets, site variability is much lower for GIT_n compared to GIT_p. This could be an indication of the performance of each GIT method in terms of site effects. Since the inversion in non-parametric approaches is less constrained than parametric inversions, especially for the source model, spectral forms of site-responses take more flexibility to fit the data. As we observe in this work, the GIT_n site-response removal results in much lower remaining residuals (as seen in the residual plots). We observe that site coefficient c1 (see tables in the appendix) is almost zero, suggesting a weak dependence of the GIT-corrected dataset on V_S (V_S used instead of V_{S30} for deconvolved recordings).

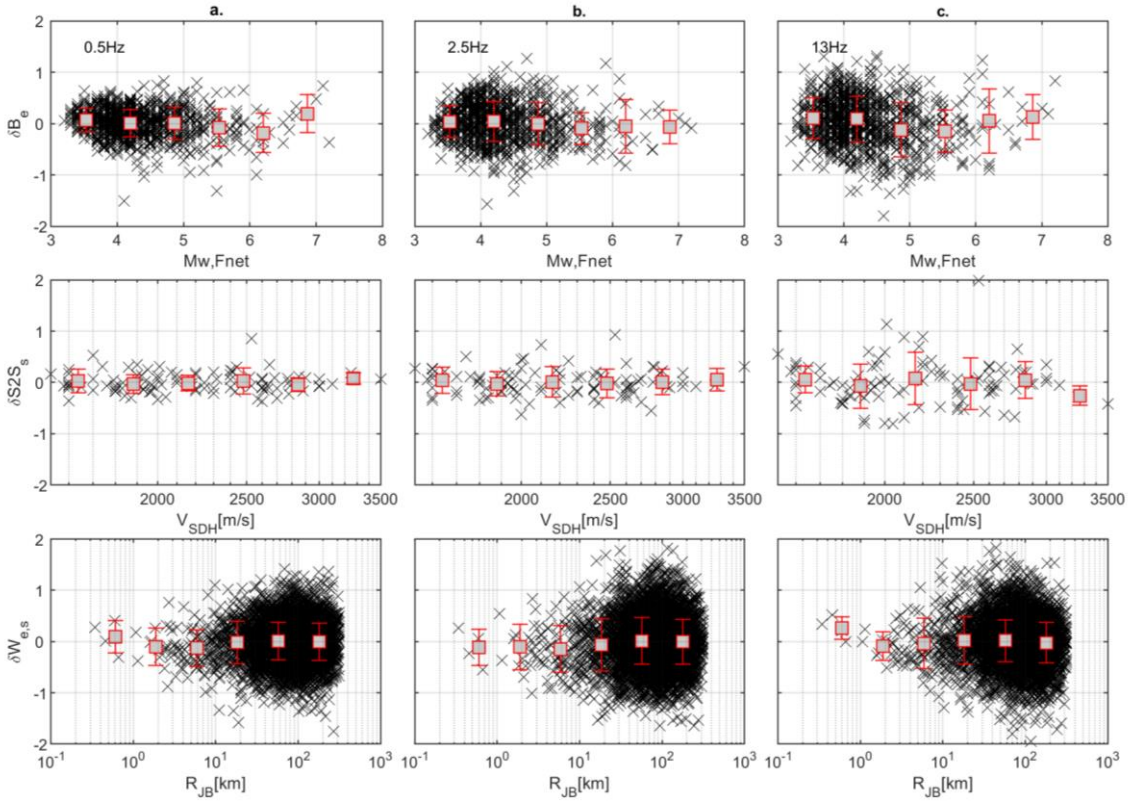

Figure 5.8: Between-event, -site, and –record residuals versus $M_{W,Fnet}$, V_{SDH}, and R_{JB}, respectively, for $SURF_{1D}cor_{1DSH}$.

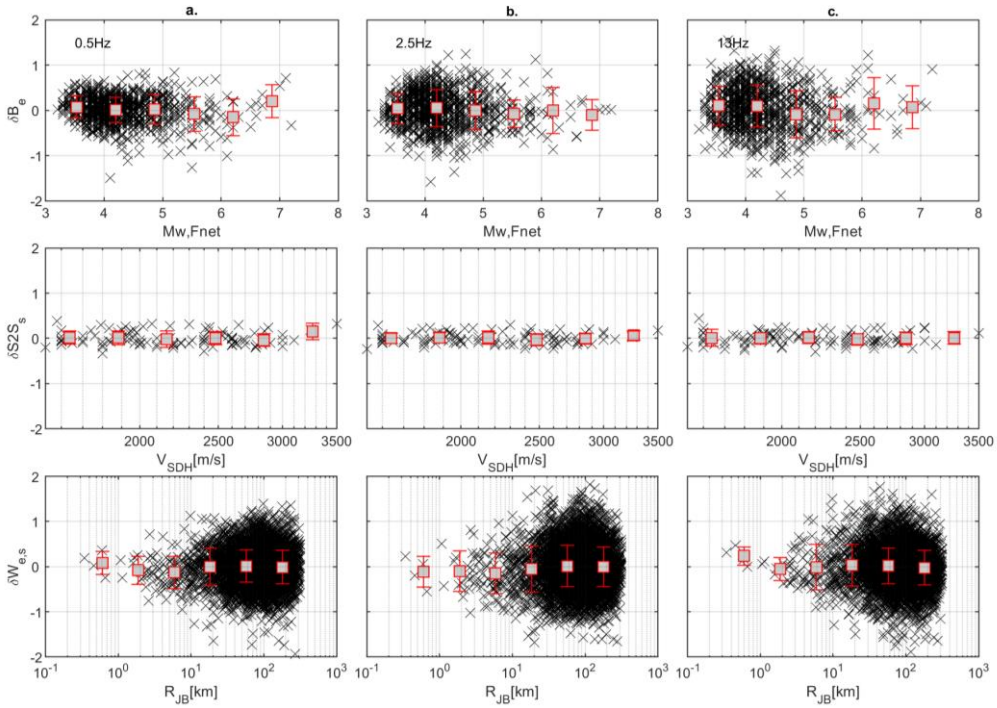

Figure 5.9: Between-event, -site, and –record residuals versus $M_{W,Fnet}$, V_{SDH}, and R_{JB}, respectively, for $SURF_{1D}cor_{GITn}$.

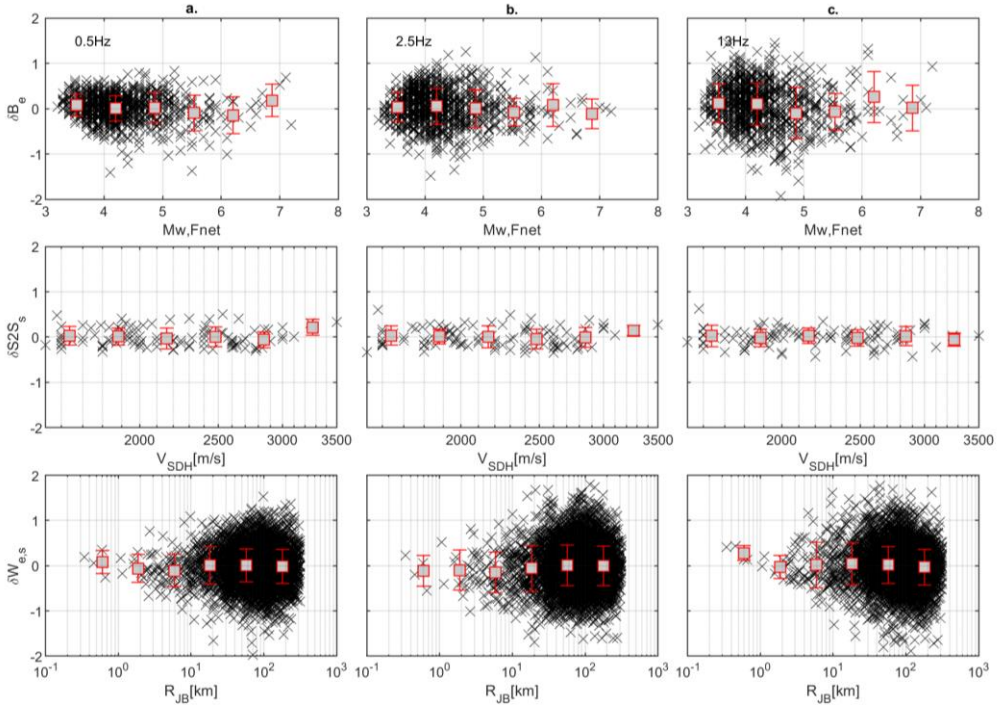

Figure 5.10: Between-event, -site, and –record residuals versus $M_{W,Fnet}$, V_{SDH}, and R_{JB}, respectively for $SURF_{1D}corGIT_p$.

Figure 5.11 shows another representation of the variability observed in terms of standard deviations for surface-1D datasets residuals. Globally, deconvolved GMPEs present a drop that reaches a maximum of 0.1 the total sigma values for all frequencies. Entering into details by decomposing the total sigma into τ, φ_{S2S} and φ_{SS} reveals the expected origins behind the drop. τ and φ_{SS} show invariance between different datasets, ensuring that the event and path residuals are kept far from significant modifications. The origin of reducing the sigma for deconvolved GMPEs is the site-effects removal. The best site-effect removal could be classified from the best

performing with GIT_n to the least performing with 1DSH estimations that show a higher remaining variability at high frequencies.

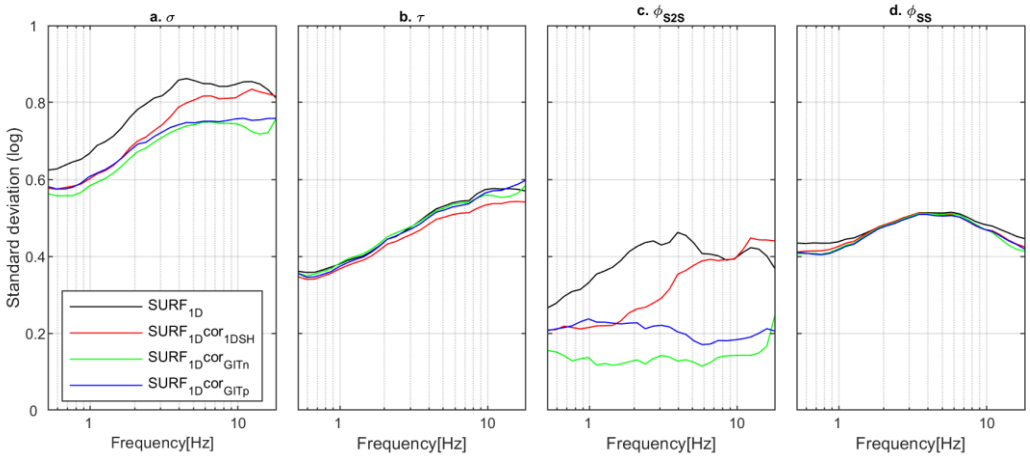

Figure 5.11: standard deviation of residuals for $SURF_{1D}$ and its "cor" datasets: a) total sigma b) between-event c) between-site d) single-station.

5.5.3 Variability analysis on $SURF_{ALL}$ GMPEs

The other interesting branch of discussion can be presented by comparing the performance of the deconvolution approach in removing site effects for all the sites not considered as 1D. For this aim, residuals of $SURF_{ALL}$ GMPEs are presented in Figure 5.12 Figure 5.13. Here, 1DSH can be applied since 1D conditions are not guaranteed. That is why GIT_n and GIT_p could only be used in deconvolution for $SURF_{ALL}$.

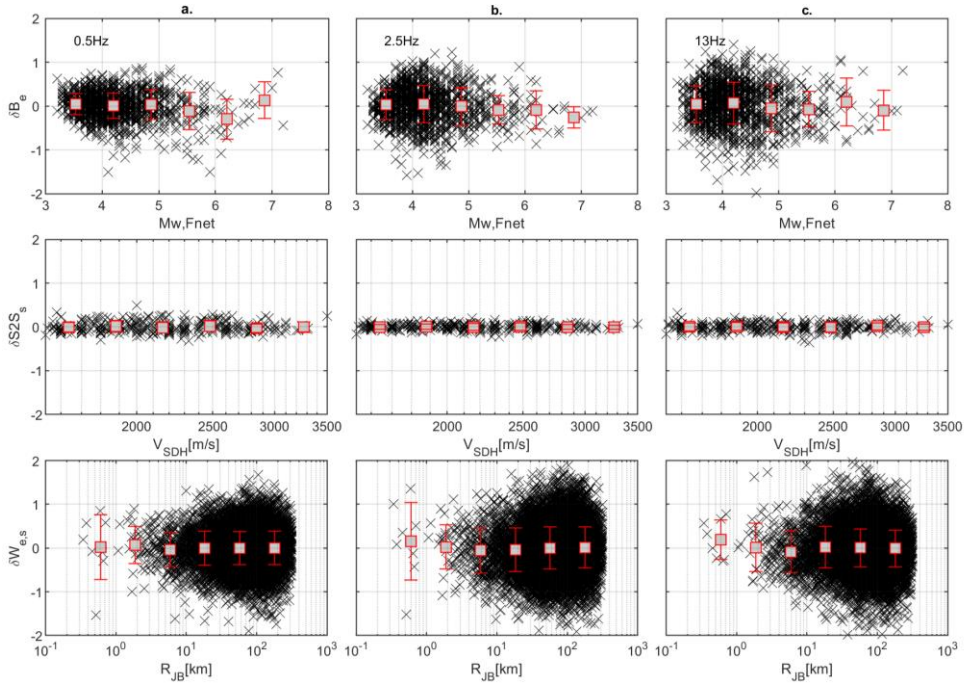

Figure 5.12: Between-event, -site, and –record residuals versus $M_{W,Fnet}$, V_{SDH}, and R_{JB}, respectively, for $SURF_{ALL}cor_{GITn}$.

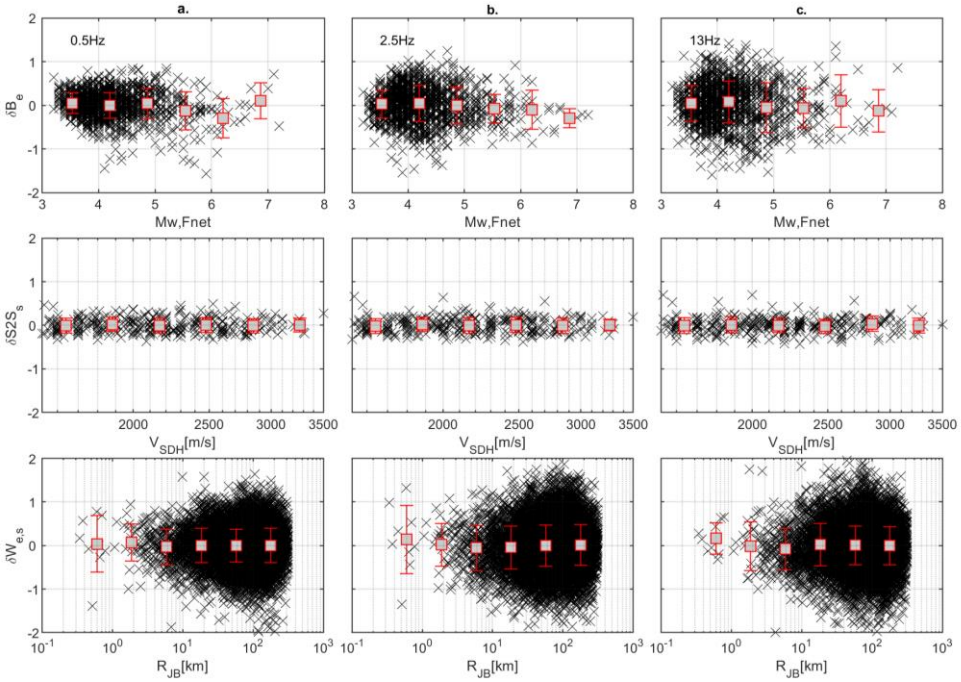

Figure 5.13: Between-event, -site, and –record residuals versus $M_{W,Fnet}$, V_{SDH}, and R_{JB}, respectively for $SURF_{ALL}cor_{GITp}$.

Compared to $SURF_{ALL}$ residuals, the drop of $\delta S2S_s$ is constantly observed and for all frequencies. Also, as a performance comparison, GIT_n seems to be slightly more effective in removing residuals than GIT_p. As presented in Figure 5.14, total sigma residuals of $SURF_{ALL}$, do not vary much from those of $SURF_{1D}$. However, the same level of drops is observed in standard deviations, always having the drop origin at the site-to-site variability φ_{S2S} only. These results and observations validate the possibility of a proper transposability of the deconvolution approach with the help of generalized inversion techniques, especially the non-parametric GIT.

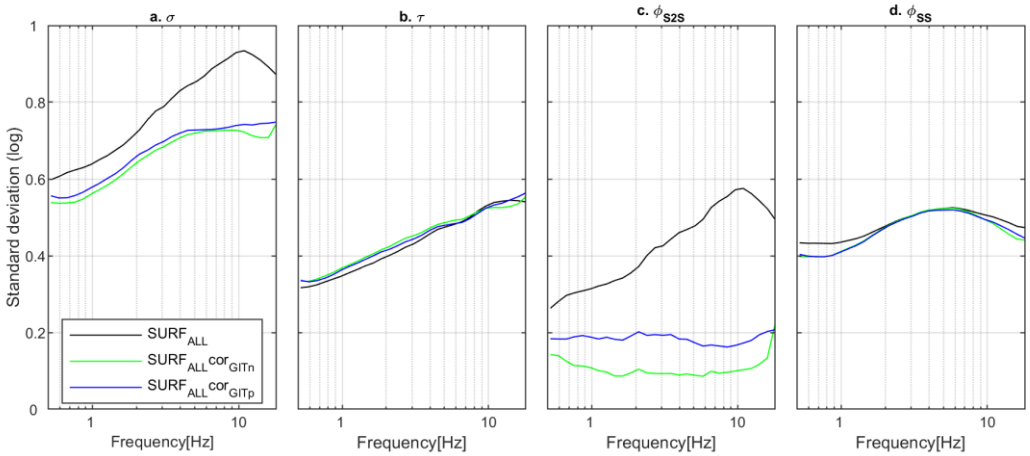

Figure 5.14: standard deviation of residuals for $SURF_{ALL}$ and its "cor" datasets: a) total sigma b) between-event c) between-site d) single-station.

5.6 Discussions

5.6.1 Results in terms of rock to hard rock scaling

As a direct implication of reference motion GMPE, the scaling factors that allow accounting for site effects could be revisited. These scaling factors were usually addressed to extend the GMPEs validity ranges for V_{S30} beyond 1500 m/s. Even

though V_{S30} might not be the only proxy to describe the site-effects fully, we hereafter look to re-establish these scaling factors between standard-rock sites of V_{S30}=800 m/s and very-hard-rock sites of V_{S30}≥1500 m/s. This factor was traditionally treated with a host-to-target adjustment taking into account V_{S30} and k_0 parameters (Houtte et al., 2011), and recent empirical investigations were carried (Ktenidou and Abrahamson, 2016; Laurendeau et al., 2018).

In this study, the derived GMPEs provide no residual $\delta S2S_s$ trends for 400 < V_{S30} < 800 m/s. We consider that any predictions of PSA (with our model) at sites of relevant V_{S30} are valid to obtain proper median predictions. The increased variability at high frequency in between-site residuals (φ_{S2S}) could be included in later discussion. Having a valid model for predictions at V_{S30} = 800 m/s and at V_{SDH} = 2000 m/s, we consider a scenario of M_w = 6.0 and R_{JB} = 20 km (Figure 5.15).

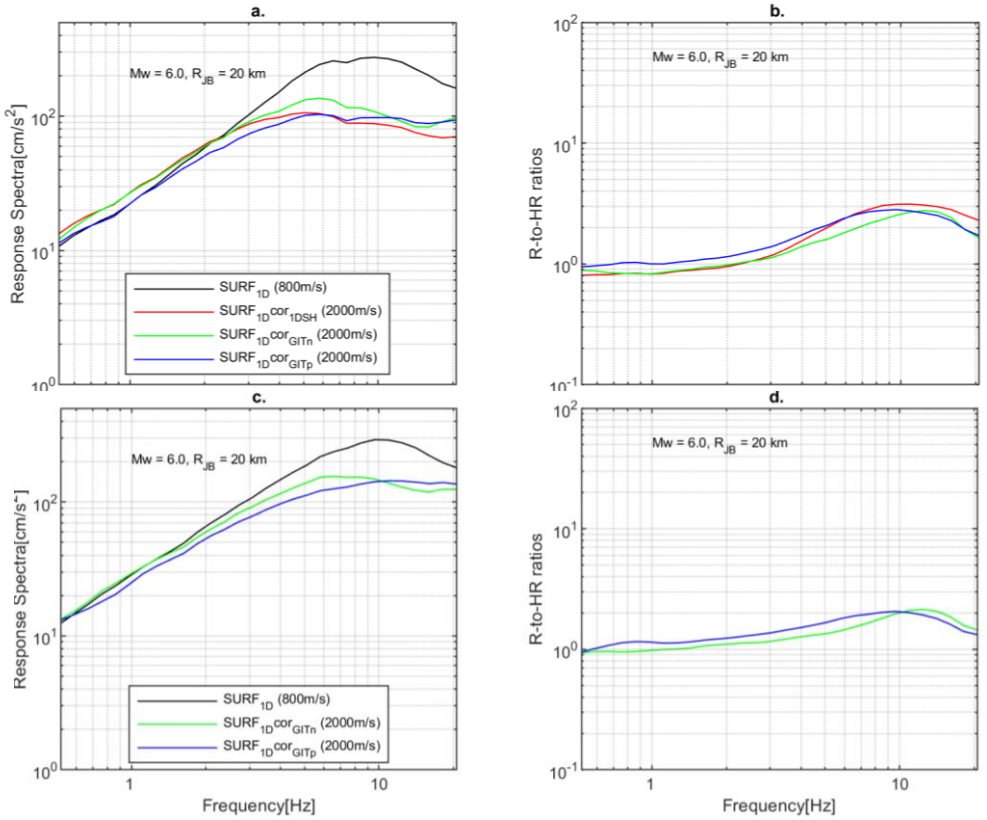

Figure 5.15: a,c) Pseudo response spectra predictions from $SURF_{1D}$ and $SURF_{ALL}$ GMPEs for rock and hard-rock conditions. b,d) Rock-to-Hard ratios obtained from $SURF_{1D}$ and $SURF_{ALL}$ GMPEs. The scenario chosen corresponds to M_W=6.0 and R_{JB}=20 km.

Despite the different levels of variability observed in the deconvolved datasets (i.e., SURFcor GMPEs), median predictions for response spectra (at several frequencies) are very close. However, in comparisons to the standard-rock predictions at the surface, both 1D and non-1D datasets present a factor of 2 to 3 at high frequencies, respectively, representing amplification of rock with respect to hard rock. The main reason behind the difference between scaling factors from $SURF_{1D}$ and $SURF_{ALL}$ is the median amplification for the sites we are considering. In the previous chapter, the median site responses for 1D and All_sites have shown differences in the observed high-frequency peak. This suggests that sites restricted to 1D conditions have higher average amplification than non-restricted sites. This could be possible since in a 1D

case, the site-effects prove most often amplification (i.e.,>1) while either amplification or de-amplification may accompany multi-dimensional effects (e.g., basin effects, topographic effects, etc…) according to geological complexities. However, considering all these complexities provide significant uncertainties. That is why it could be interesting in the next step to include all residual $\delta S2S_s$ to the rock-to-hard-rock scaling estimations.

Figure 5.16 shows the estimation of the scaling ratio for different site conditions, i.e. V_{S30} = 400, 600, 800 and 1000 m/s. These estimations follow a common shape that highly amplifies around 10 Hz. Though the relation between Fourier and response spectral domain is not linear, the average amplification observed in the Fourier domain controls in a clear way the median amplification observed in PSAs. This result is consistent between the different datasets and the deconvolution approaches.

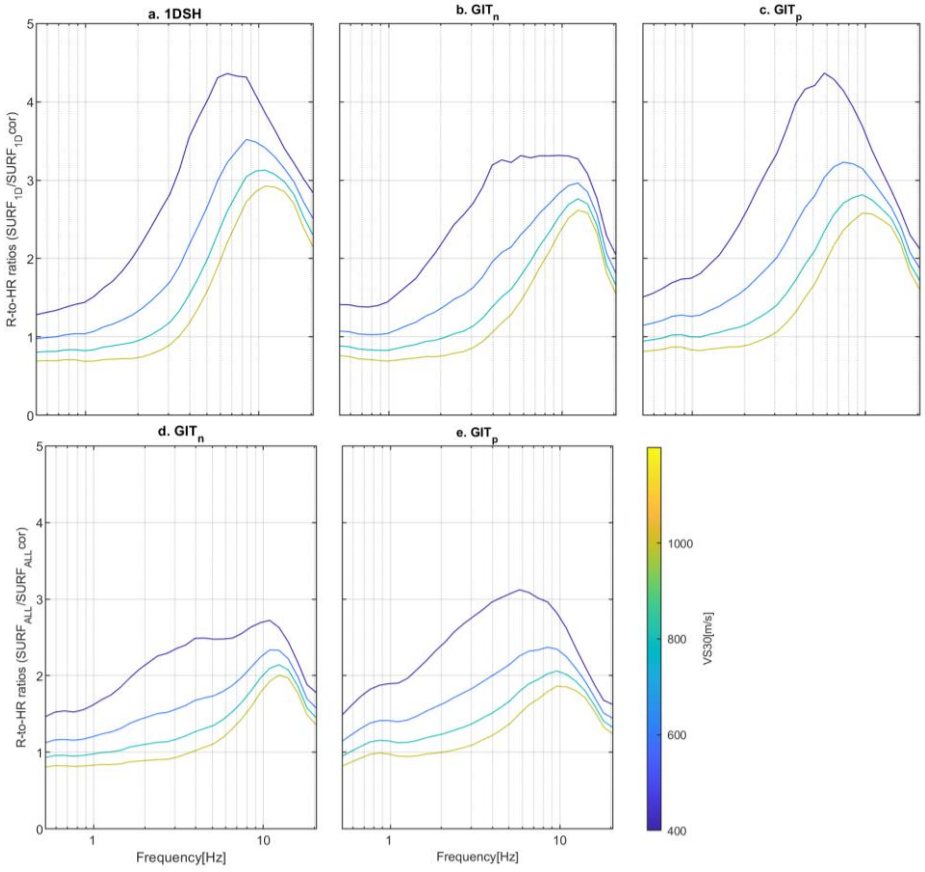

Figure 5.16: Rock to hard-rock scaling ratios obtained from GMPEs $SURF_{1D}$ and $SURF_{ALL}$ at several V_{S30} values from 400 to 1000 m/s.

5.6.2 Comparison of amplification factors from recent GMPEs

Figure 5.17a shows the amplification obtained in this study, a comparison of scaling factors from previous studies (17b), and the amplification factors from existing GMPEs (17c). The nearest factor to our results is Laurendeau et al. (2018), which applied site effect correction by 1DSH estimations with slightly different data configurations.

The comparisons of Figure 5.17a show that $SURF_{1D}$ leads to an amplification factor much greater than $SURF_{ALL}$, even though we expect 2D/3D effects can lead to much higher amplifications than in 1D conditions. This observation can be explained by the fact that we are comparing average responses. Site responses from $SURF_{ALL}$ have more significant variability, leading to smooth the final median value. Besides, we

identified 1D sites with the help of 1D theoretical estimates and V_S profiles. These sites essentially have a 1D geological structure and a reliable estimate of the V_S profiles deep enough to retrieve the global site amplification. Hence, a site that is in 1D conditions but has biased V_S profiles will be displaced from the 1D list towards the ALL list. Eventually, $SURF_{ALL}$ data might contain a mixture of 1D, 2D, and 3D conditions, and its median amplification (or scaling factor) does not necessarily indicate that 2D/3D conditions amplify ground motion less than 1D conditions.

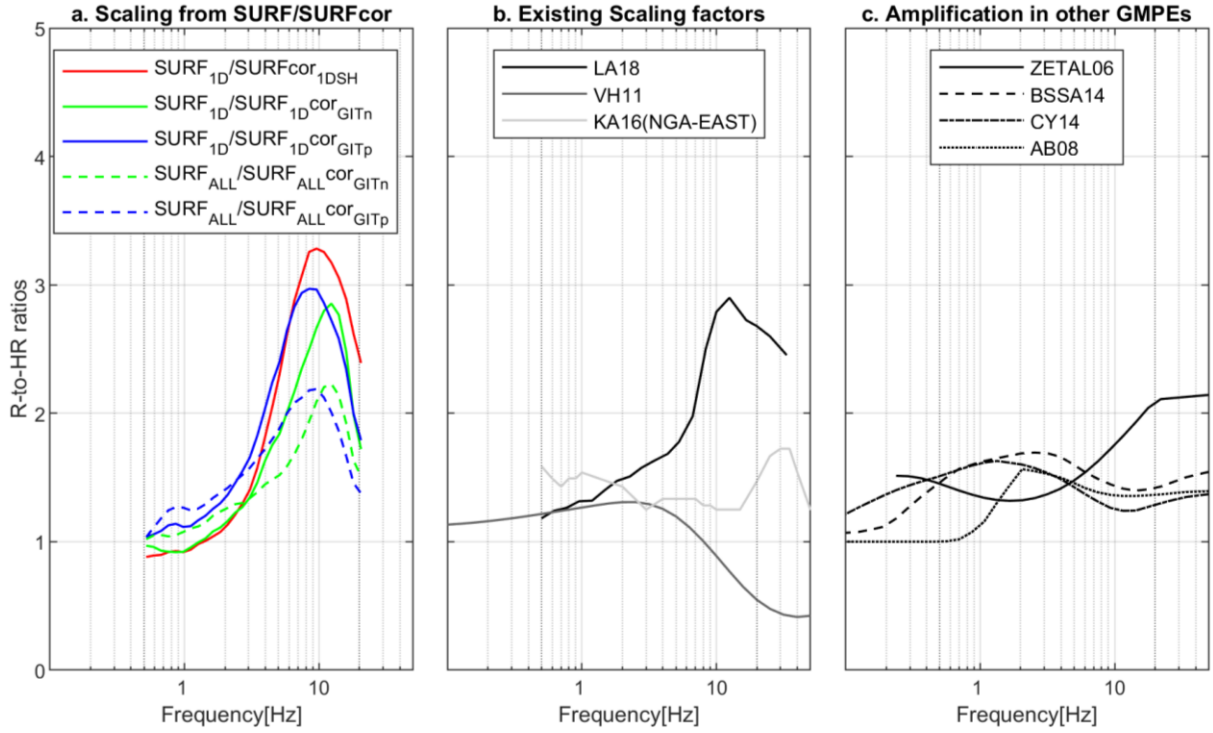

Figure 5.17: a) scaling factors obtained in this study for $SURF_{1D}$ and $SURF_{ALL}$. b) Recently proposed or observed ratios (Laurendeau et al., 2018; V.Houtte et al., 2011; Ktenidou and Abrahamson, 2016) c) amplifications terms from some existing GMPEs corresponding to ratios of predicted PSA at V_{S30}=800m/s to that at V_{S30}=1500m/s (Zhao et al., 2006; Boore et al., 2014; Chiou and Youngs, 2014; Abrahamson and Silva, 2008).

If we compare the ratios obtained with that proposed by Houtte et al. (2011) that uses the traditional Host-to-target adjustments (HTTA), we find a difference of factor 3 at high frequencies. HTTA scaling ratios overestimates ground motion for hard-rock sites compared to soft-rock sites. In fact, the mean correction factor of HTTA does not consider that hard-rock sites might have amplification in high frequencies. Instead, HTTA considers that there is a lack of attenuation at high frequencies. Thus, the assumed values of κ to adjust GMPEs for hard-rock sites can be deeply questioned in the mean (average) way it is applied for all sites (Bard et al., 2020). V_S profiles for host sites used for generic rock profiles selection (that are not limited only to V_{S30} but also to several kilometers depth) are poorly known and generally not constrained (Boore and Joyner 1997, Cotton et al. 2006). Additional doubts arise after the observed dispersion in the correlations of V_{S30}-κ, which also depends on how careful measurements are performed (Hollender, 2019; Hollender et al., 2020).

Ktenidou and Abrahamson (2016) investigated the reestimation of the rock-to-hard-rock ratios empirically on the NGA-East database. They compared purely empirical ratios with those obtained from the analytical models of HTTA approaches. For frequencies beyond 20 Hz, they proposed an empirical scaling from soft-rock to hard-

rock, unlike the standard HTTA. In fact, their new empirical rock-to-hard-rock ratios deviate in values from the traditional HTTA at high frequencies with a factor of 2. However, the scaling factor (as an absolute value) remains limited to 1.5, whereas our approach results in 2.5 to 3. In addition, the peak we observe with the deconvolution approach around 10 Hz in these ratios is observed at a higher frequency (30 Hz, which is not visible in our usable frequency-band 0.5-20 Hz). Since their dataset did not have many hard-rock sites (limited V_{S30} sampling for V_{S30}>1500 m/s), it might need more data on hard-rock sites for that the dataset to become more representative of hard-rock conditions.

Amplification factors from different GMPEs are displayed in Figure 5.17(c) as computed from the ratios between predicted PSA for generic rock (800 m/s) and hard-rock (1500 m/s). The latter hard-rock V_{S30} values were set to 1500 m/s in Figure 5.17(c) ratios since they represent the upper limit of most of the derived GMPEs. None of these amplification factors look comparable to the derived amplifications from deconvolved GMPEs in Figure 5.17(a). This discrepancy is due to the fact that hard-rock predictions of the other GMPEs are not reliable even at 1500 m/s because of lack of data.

To sum up, ground motion prediction is not well-defined for reference conditions in most GMPEs derived from empirical surface recordings. The lack of hard-rock recordings for GMPEs could be one challenge, but it is not the only one. The parametrization of site effects by site-specific proxies, which intends to simplify the practice, apparently increases the variability and leads to significant overestimations. Moreover, more uncertainties are injected in ground motion predictions when the proxies themselves are deduced from correlations derived in the same or different regions (whose measurements and interpretations could also be biased). This is the case, especially for the κ parameter assumed to be site-specific attenuation, where it could be biased with other phenomena (Hollender et al., 2020). Thus, the commonly assumed lack of attenuation for hard-rock sites could result simply from the amplification at high frequencies due to thin shallow layers (Hollender et al., 2018). Eventually, the robust predictions of reference motion should be performed independently from the existing host side GMPEs, where the deconvolved GMPE presents a promising approach.

5.7 Conclusions and perspectives

The purpose of this chapter was to use the results of generalized inversions in the deconvolution approach initially proposed in Laurendeau et al. (2018) after extending the site response estimations, in the previous chapter, to all sites in the dataset. The latter allows applying the deconvolution approach to other datasets where 1D sites can be easily identified.

The site coefficient derived from hard rock GMPE shows that removing the site effects from recordings probably weakens (or even vanishes) the dependence on V_{S30} (or V_{SDH}). Detailed comparisons of between-site residuals have shown that only the deconvolution approach decreases the site variability with different levels. Site-effects removal with 1DSH estimations leaves variability limits almost unchanged. However, GIT_n and GIT_p provide powerful corrections and minimized variabilities. It appears that for a perfect reference motion (a complete amplification-free motion), the ground motion amplitudes do not scale anymore with the corresponding value of V_S.

Results in terms of rock-to-hard-rock ratios are also presented, which meets with the conclusions of Ktenidou and Abrahamson (2016) regarding the traditional scaling ratios to account for hard-rock site conditions and the results Laurendeau et al., (2018). Again, these ratios propose that soft rocks have amplification in ground motion at high frequencies compared to hard-rock sites. This conclusion, again, opposes the scaling factors obtained by HTTA techniques, especially at high frequencies, and motivates to re-check the appropriateness of these scaling factors. The transportability of the approach was the primary concern throughout the studies in Chapters 4 and 5. The central perspective after this work is the derivation of a deconvolved reference motion GMPE for European and French regions.

5.8 Acknowledgments

Mahdi Bahrambouri and Adrian Rodriguez-Marek are thanked a lot for sharing their preparations of flatfiles and data treatment information of the KiK-net dataset. Vincent Bremaud is warmly thanked for helping with R-language. Thanks to the Japanese National Research Institute for Earth Science and Disaster Resilience (NIED) for providing the KiK-net data with all its metadata in open access online.

5.9 Supplements

Tables of regression coefficients are delivered for each GMPE in a separate appendix.

Conclusions and Perspectives

Overview of the work and general conclusions

The main objective of the work, illustrated through several chapters, is to define a reliable methodology to obtain reference ground motion in the context of site-specific hazard assessments. However, we were interested in a methodology that applies to other existing ground-motion databases, particularly in the French and European low-to-moderate seismicity areas.

The standard definition of site-specific seismic hazard assessment involves first predicting reference motion and then integrating site effects. However, the lack of empirical recordings on hard-rock sites has limited validity ranges of most GMPEs to sites with V_S values up to 1000 m/s. Thus, assessing reference motion, commonly associated to sites with $V_S \geq 1500$ m/s, is not possible from classical empirical GMPEs. Since the early 2000s, researchers have proposed to overcome the lack of hard-rock motion issues by performing corrections on ground-motion predictions for standard-rock sites ($V_S = 800$ m/s) to adjust them to hard-rock conditions. The corrections mainly consider site effect differences between recording host and target sites through factors, so-called host-to-target adjustments (HTTA). Eventually, the need to define reference motion in the engineering practice has led to applying HTTA despite the related uncertainties (see Chapter 1).

Laurendeau et al. (2018) initially proposed an alternative to the existing HTTA by deriving reference motion based on the correction of surface recordings of the host sites. This approach is also called the deconvolution approach, and it stands for the site response removal from surface recordings. This approach delivers a virtual database corresponding to reference rock conditions, directly used to derive reference GMPEs. Thus, site response estimations constitute the essential elements *a priori* required by the deconvolution approach. To this aim, Laurendeau et al. (2018) benefited from all available V_S profiles for the KiK-net sites to use 1DSH site responses to correct surface recordings. Laurendeau et al. (2018) carried out the work on a KiK-net subset to validate the methodology. However, the direct use of these deconvolved GMPEs in hazard assessment and operational studies is not acceptable for several reasons. First, the seismotectonic characteristics of Japan are not necessarily conserved for other regions such as Europe, for instance. Thus, there will be a need to define these deconvolved GMPEs at each region of interest.

To summarize, both the HTTA and deconvolution approaches perform corrections on recorded host data to obtain reference motion and can be considered alternative approaches to obtain ground motion at hard-rock sites.

The main comments about HTTA are summarized as follows:

- First of all, the HTTAs are defined in the Fourier domain, while ground motion predictions are usually for the response spectra. Hence, conversions are often made to Fourier spectra to apply the V_{S30} and κ corrections, then convert back to response spectra. This step is thus possible through the RVT and IRVT. Each of these processes leads to non-unique solutions and is accompanied by

additional sources of uncertainties, especially that the conversion with RVT back to response spectra needs an additional model for the duration.

- HTTAs are prone to several sources of uncertainties. Even though the ongoing research investigates developing GMPEs in the physical Fourier domain, which eliminates IRVT usage, several sources of uncertainties are still present. HTTA suffers variability in simple parametrizations to describe frequency-dependent phenomena, mainly the V_{S30} and κ. However, even if we accept using V_{S30} and κ corrections, they still face measurement issues in most strong-motion databases. These issues are mainly due to the lack of reliable estimations for either term at the target hard-rock sites.
- Each of the site proxies faces a lack of measurements and biases. In fact, measured V_{S30} and κ values are missing for a significant proportion of sites in common strong-motion databases. On the other side, the κ values face methodological biases, where measurements can be biased by site effects. For instance, Hollender et al. (2020) showed for a set of French stations that intermediate resonance peaks of standard-rock lead to positive κ values while high-frequency peaks observed at hard-rock lead to negative κ values. This high-frequency amplification of hard-rock sites is often due to soft shallow layers. Thus, lower κ values of hard-rock motion are likely not due to lack of attenuation but to high-frequency amplification. Eventually, the latter leads to over-predictions of the final hazard estimates at high-frequency.

The difference between HTTA and the deconvolution approach is that they apply the corrections with different assumptions, parameters, and related sources of uncertainties.

- The deconvolution approach reduces all the uncertainties in HTTA to only one, related to the estimation of the site-specific transfer functions. However, these uncertainties diminish when using empirical approaches of estimation. To this aim, GIT represents a helpful tool to estimate site responses instead of the 1DSH used in Laurendeau et al. (2018). Following the definition of GIT, its reliability is mainly controlled by a proper constraint on site responses to well separate the associated terms. Thus, ensuring a good reference site in GIT will control and constrain the whole correction procedure.
- Surface recordings correction keeps a symmetric approach in hazard estimation for both sides, the host and target. HTTA applies corrections to GMPEs derived for standard-rock conditions to obtain reference motion. Instead, the deconvolution of surface recordings is applied to each site before GMPE derivations. This step leads to virtual reference time histories that constitute a new database for GMPE regressions. Thus, target GMPEs are derived using target sites, and they are no longer deduced from host GMPEs. In this way, uncertainties related to correction parameters' measurements on host sites will not propagate to target sites, as the case for HTTA.

Defining a reference constraint in GIT appears essential. The reference is commonly related to an amplification-free site in most publications. Also, several authors used to constrain the average response of several stations to unity in the case of several reference candidates. These two possibilities are unavoidable if site-specific

information is unattainable. However, if a precise amplification is needed, the reference site must be constrained with a response closest to its true values. This step allows GIT to deliver absolute site responses.

One possibility of achieving absolute responses is using 1DSH theoretical site responses as a constraint for one or several stations in inversions. Hence, chapter 2 addressed one source of uncertainty related to 1DSH site responses, impacting high frequencies. This source of uncertainty is the unaccounted vertical small-scale heterogeneities in V_S profiles. We accounted for these heterogeneities by a simple geostatistical model, based on observations, and simple to apply in 1DSH simulations. We showed that including these small-scale perturbations in the velocity profiles can affect the 1D site classification criterion initially introduced by Thompson et al. (2012). Also, the work showed that these perturbations could also induce additional attenuation effects at high frequencies, which fits better with empirical estimates.

Second, the performance and reliability of the GIT method itself were addressed in the GITEC benchmark of chapter 3. In this benchmark, the existing schemes of generalized inversions were tested and compared on several datasets. The primary outcomes can be summarized in two main points. The first is that significant lateral variations appear to correlate with the high inter-method variabilities. Consequently, site response estimations are prone to bias, especially at high frequencies, if we wholly discard regional variations. The second outcome is that uncertainty on site responses from GIT methods appears limited, and they are the most robust elements of GIT results. These conclusions were considered in the rest of the work on KiK-net.

Afterward, we performed in chapter 4 an extensive analysis to estimate site effects on the KiK-net data using different methods. Empirical SSR, HVSR, and GIT were estimated from recorded data, while theoretical 1DSH site responses using the available velocity profiles. Here, GIT estimations were obtained after considering an attenuation regionalization of the Japanese regions. Then, we defined a list of 1D sites based on the comparison between empirical and theoretical 1D site responses. We also proposed to compare the absolute surface transfer functions and not only the relative surface-to-borehole ratios. In the end, a site is classified with 1D conditions if it simultaneously has a 1D geological structure and a reliable velocity profile. We

We concluded the work in chapter 5 by deriving the reference motion GMPE. Here, we defined two kinds of datasets, one for 1D site conditions and another for all sites in the dataset. The first (1D datasets) allowed us to compare the results with those of Laurendeau et al. (2018), while the second was interesting to explore the results in generalized conditions. One of the primary outcomes of this part was that the deconvolution approach, using either theoretical or empirical site responses, shows consistent estimations of rock-to-hard-rock scaling ratios. However, the only observed difference lies in the remaining variability level of 1DSH-corrected GMPEs. This observation indicates the robustness of GIT site terms in the corrections. Therefore, the final prediction values for hard-rock conditions, being overpredicted in the current state of practice, calls for increased attention towards reference motion prediction and accounting for site effects in hazard analyses.

It is worth mentioning that the deconvolved reference motion will represent a virtual hard-rock motion at specific V_S conditions. For example, depending on the used

method of site effects estimations, the V_S conditions of the deconvolved GMPE can be as follows:

- If GIT were used to obtain relative terms, then all the final hazard estimates will be relative to the reference site, i.e., V_S equalsV_{S30} of the surface reference station.
- If GIT with an absolute reference choice was considered (the main point introduced in chapter 4), its final predictions will be relative to the bedrock level where no site effects are present. In this case, one might consider that V_S equals V_S of the crust (i.e., 3500 m/s).
- If 1DSH numerical simulations were conducted, which is currently possible for KiK-net only, V_S equals V_{SDH}.

Overall, we should not forget that we need reference motion (i.e., amplification-free) while the standard practice targeted rigid hard-rock sites because of their closeness to reference bedrock conditions. Thus any unaccounted amplifications in the hard-rock motion will lead to under/overestimate the seismic hazard. For instance, since surface hard-rock sites are usually weathered in the first very few meters, the motion will be amplified at high frequencies and possibly deviate from an amplification-free motion. This issue is resolved by accounting for precise site-specific conditions, as proposed in the deconvolution approach.

Perspectives

The current work was done to validate the use of site responses from GIT in the deconvolution approach for French and European datasets, where borehole sensors or deep velocity profiles are not available everywhere. The methodology was tuned over the KiK-net network to minimize the necessity for site information with keeping reliable results. After showing that the deconvolution approach results in consistent GMPEs on KiK-net data, it appears that the methodology defined throughout the thesis is applicable in other regions.

We aim to derive Pan-European GMPEs for reference motion based on the lessons and outcomes of this work. Thus, the main steps in the future work can be summarized as follows:

- Identify one or several well-recording stations (i.e., number of events) in the studied regions with measured velocity profiles. Once these key stations are defined, their absolute theoretical site responses can help to constrain generalized inversions.
- Characterize new hard-rock stations which appear to be a potential candidate for reference constraint in GIT. This work has already started by conducting characterization campaigns in the last few years in France and Italy. This work was done under the supervision of Fabrice Hollender (CEA), where I participated several times in these campaigns.
- Investigate the impact of vertical heterogeneities on the 1DSH site responses using the geostatistical model.

- Identify strong lateral variations of the crust in the considered regions and consider them in the GIT inversions. This step should ensure a proper site and source separation even if the regionalization performed is not highly precise.
- After data correction, we will be interested in deriving GMPEs in both response spectral and Fourier domains. After the increased attention towards Fourier GMPEs in the seismological community, it seems essential to compare their final results with GMPEs in the response domain. This comparison will highlight possible discrepancies, especially that Fourier GMPE predictions are converted to the response domain with a duration model and the random vibration theory.

References

Abrahamson, N., and W. Silva, 2008, Summary of the Abrahamson & Silva NGA Ground-Motion Relations:, Earthquake Spectra, doi: 10.1193/1.2924360.

Abrahamson, N. A., and R. R. Youngs, 1992, A stable algorithm for regression analyses using the random effects model, Bulletin of the Seismological Society of America, 82, no. 1, 505–510.

Ahdi, S. K., S. Sadiq, O. Ilhan, Y. Bozorgnia, Y. M. A. Hashash, D. Y. Kwak, D. Park, A. Yong, and J. P. Stewart, 2018, Development of a United States Community Shear Wave Velocity Profile Database, 330–339, doi: 10.1061/9780784481462.032.

Ameri, G., F. Hollender, V. Perron, and C. Martin, 2017, Site-specific partially nonergodic PSHA for a hard-rock critical site in southern France: adjustment of ground motion prediction equations and sensitivity analysis, Bull Earthquake Eng, 15, no. 10, 4089–4111, doi: 10.1007/s10518-017-0118-6.

Ameri, G., A. Oth, M. Pilz, D. Bindi, S. Parolai, L. Luzi, M. Mucciarelli, and G. Cultrera, 2011, Separation of source and site effects by generalized inversion technique using the aftershock recordings of the 2009 L'Aquila earthquake, Bull Earthquake Eng, 9, no. 3, 717–739, doi: 10.1007/s10518-011-9248-4.

Anderson, J. G., 2007, Anderson, J. G. (2007). Physical processes that control strong ground motion, in Earthquake Seismology, Treatise on Geophysics, Vol. 4, Hiroo Kanamori (Editor), Elsevier, Amsterdam, 513–565., Elsevier, Amsterdam.

Anderson, J. G., and J. N. Brune, 1999, Probabilistic Seismic Hazard Analysis without the Ergodic Assumption, Seismological Research Letters, 70, no. 1, 19–28, doi: 10.1785/gssrl.70.1.19.

Anderson, J. G., and S. E. Hough, 1984, A model for the shape of the fourier amplitude spectrum of acceleration at high frequencies, Bulletin of the Seismological Society of America, 74, no. 5, 1969–1993.

Andrews, D. J., 1986, Objective Determination of Source Parameters and Similarity of Earthquakes of Different Size, in Earthquake Source Mechanics S. Das, J. Boatwright, and C. H. Scholz (Editors), American Geophysical Union, 259–267.

Aristizábal, C., P.-Y. Bard, C. Beauval, and J. C. Gómez, 2018, Integration of Site Effects into Probabilistic Seismic Hazard Assessment (PSHA): A Comparison between Two Fully Probabilistic Methods on the Euroseistest Site, 8, Geosciences, 8, no. 8, 285, doi: 10.3390/geosciences8080285.

Atik, L. A., N. Abrahamson, J. J. Bommer, F. Scherbaum, F. Cotton, and N. Kuehn, 2010, The Variability of Ground-Motion Prediction Models and Its Components, Seismological Research Letters, 81, no. 5, 794–801, doi: 10.1785/gssrl.81.5.794.

Atik, L. A., A. Kottke, N. Abrahamson, and J. Hollenback, 2014, Kappa (κ) Scaling of Ground-Motion Prediction Equations Using an Inverse Random Vibration Theory ApproachKappa (κ) Scaling of Ground-Motion Prediction Equations

Using an Inverse Random Vibration Theory Approach, Bulletin of the Seismological Society of America, 104, no. 1, 336–346, doi: 10.1785/0120120200.

Atkinson, G. M., E. Yenier, N. Sharma, and V. Convertito, 2016, Constraints on the Near-Distance Saturation of Ground-Motion Amplitudes for Small-to-Moderate Induced EarthquakesConstraints on the Near-Distance Saturation of Ground-Motion Amplitudes, Bulletin of the Seismological Society of America, 106, no. 5, 2104–2111, doi: 10.1785/0120160075.

Bahrampouri, M., and A. Rodriguez-Marek, 2019, Ground Motion Parameters for KiK-net Records: An Updated Database, Database, doi: 10.17603/ds2-e0ts-c070.

Baillard, C., W. C. Crawford, V. Ballu, C. Hibert, and A. Mangeney, 2014, An Automatic Kurtosis-Based P- and S-Phase Picker Designed for Local Seismic NetworksAn Automatic Kurtosis-Based P- and S-Phase Picker Designed for Local Seismic Networks, Bulletin of the Seismological Society of America, 104, no. 1, 394–409, doi: 10.1785/0120120347.

Bard, P.-Y., 2008, Foreword, Bull Earthquake Eng, 6, no. 1, 1–2, doi: 10.1007/s10518-008-9059-4.

Bard, P.-Y., S. S. Bora, F. Hollender, A. Laurendeau, and P. Traversa, 2020, Are the Standard V S -Kappa Host-to-Target Adjustments the Only Way to Get Consistent Hard-Rock Ground Motion Prediction?, Pure Appl. Geophys., 177, no. 5, 2049–2068, doi: 10.1007/s00024-019-02173-9.

Bard, P.-Y., and M. Bouchon, 1980a, The seismic response of sediment-filled valleys. Part 1. The case of incident SH waves, Bulletin of the Seismological Society of America, 70, no. 4, 1263–1286, doi: 10.1785/BSSA0700041263.

Bard, P.-Y., and M. Bouchon, 1980b, The seismic response of sediment-filled valleys. Part 2. The case of incident P and SV waves, Bulletin of the Seismological Society of America, 70, no. 5, 1921–1941, doi: 10.1785/BSSA0700051921.

Bard, P.-Y., and J.-C. Gariel, 1986, The seismic response of two-dimensional sedimentary deposits with large vertical velocity gradients, Bulletin of the Seismological Society of America, 76, no. 2, 343–366.

Bardet, J., K. Ichii, and C. Lin, 2000, A computer program for equivalent-linear earthquake site response analyses of layered soil deposits, University of Southern California, Department of Civil Engineering.

Bates, D., M. Mächler, B. Bolker, and S. Walker, 2015, Fitting Linear Mixed-Effects Models Using lme4, Journal of Statistical Software, 67, no. 1, 1–48, doi: 10.18637/jss.v067.i01.

Berge-Thierry, C., F. Cotton, O. Scotti, D.-A. Griot-Pommera, and Y. Fukushima, 2003, New Empirical Response Spectral Attenuation Laws for Moderate European Earthquakes, Journal of Earthquake Engineering, 7, no. 2, 193–222, doi: 10.1080/13632460309350446.

Bindi, D., L. Luzi, M. Massa, and F. Pacor, 2010, Horizontal and vertical ground motion prediction equations derived from the Italian Accelerometric Archive (ITACA), Bull Earthquake Eng, 8, no. 5, 1209–1230, doi: 10.1007/s10518-009-9130-9.

Bindi, D., M. Massa, L. Luzi, G. Ameri, F. Pacor, R. Puglia, and P. Augliera, 2014, Pan-European ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5 %-damped PSA at spectral periods up to 3.0 s using the RESORCE dataset, Bull Earthquake Eng, 12, no. 1, 391–430, doi: 10.1007/s10518-013-9525-5.

Bindi, D., F. Pacor, L. Luzi, M. Massa, and G. Ameri, 2009, The Mw 6.3, 2009 L'Aquila earthquake: source, path and site effects from spectral analysis of strong motion data, Geophys J Int, 179, no. 3, 1573–1579, doi: 10.1111/j.1365-246X.2009.04392.x.

Bindi, D., S. Parolai, H. Grosser, C. Milkereit, and S. Karakisa, 2006, Crustal Attenuation Characteristics in Northwestern Turkey in the Range from 1 to 10 Hz, Bulletin of the Seismological Society of America, 96, no. 1, 200–214, doi: 10.1785/0120050038.

Bindi, D., D. Spallarossa, and F. Pacor, 2017, Between-event and between-station variability observed in the Fourier and response spectra domains: comparison with seismological models, Geophys J Int, 210, no. 2, 1092–1104, doi: 10.1093/gji/ggx217.

Biro, Y., and P. Renault, 2012, Importance and Impact of Host-to-Target Conversions for Ground Motion Prediction Equations in PSHA, Proc. of the 15th World Conference on Earthquake Engineering., 10.

Bommer, J. J., and J. E. Alarcon, 2006, The Prediction and Use of Peak Ground Velocity, Journal of Earthquake Engineering, 10, no. 1, 1–31, doi: 10.1080/13632460609350586.

Bonilla, L. F., J. H. Steidl, G. T. Lindley, A. G. Tumarkin, and R. J. Archuleta, 1997, Site amplification in the San Fernando Valley, California: Variability of site-effect estimation using the S-wave, coda, and H/V methods, Bulletin of the Seismological Society of America, 87, no. 3, 710–730.

Boore, D. M., 2003, Simulation of Ground Motion Using the Stochastic Method, Pure appl. geophys., 160, no. 3, 635–676, doi: 10.1007/PL00012553.

Boore, D. M., and G. M. Atkinson, 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, no. 1, 99–138, doi: 10.1193/1.2830434.

Boore, D. M., and W. B. Joyner, 1997, Site amplifications for generic rock sites, Bulletin of the Seismological Society of America, 87, no. 2, 327–341.

Boore, D. M., J. P. Stewart, E. Seyhan, and G. M. Atkinson, 2014, NGA-West2 Equations for Predicting PGA, PGV, and 5% Damped PSA for Shallow Crustal Earthquakes:, Earthquake Spectra, doi: 10.1193/070113EQS184M.

Bora, S. S., F. Scherbaum, N. Kuehn, and P. Stafford, 2016, On the Relationship between Fourier and Response Spectra: Implications for the Adjustment of Empirical Ground-Motion Prediction Equations (GMPEs)On the Relationship between Fourier and Response Spectra, Bulletin of the Seismological Society of America, 106, no. 3, 1235–1253, doi: 10.1785/0120150129.

Bora, S. S., F. Scherbaum, N. Kuehn, P. Stafford, and B. Edwards, 2015, Development of a Response Spectral Ground-Motion Prediction Equation (GMPE) for Seismic-Hazard Analysis from Empirical Fourier Spectral and Duration ModelsDevelopment of Response Spectral GMPE for Seismic-Hazard Analysis, Bulletin of the Seismological Society of America, 105, no. 4, 2192–2218, doi: 10.1785/0120140297.

Borcherdt, R. D., 1970, Effects of local geology on ground motion near San Francisco Bay, Bulletin of the Seismological Society of America, 60, no. 1, 29–61.

Bozorgnia, Y. et al., 2014, NGA-West2 Research Project, Earthquake Spectra, 30, no. 3, 973–987, doi: 10.1193/072113EQS209M.

Brillinger, D. R., and H. K. Preisler, 1985, Further analysis of the Joyner-Boore attenuation data, Bulletin of the Seismological Society of America, 75, no. 2, 611–614.

Brocher, T. M., 2005, Empirical Relations between Elastic Wavespeeds and Density in the Earth's Crust, Bulletin of the Seismological Society of America, 95, no. 6, 2081–2092, doi: 10.1785/0120050077.

Brune, J. N., 1970, Tectonic stress and the spectra of seismic shear waves from earthquakes, Journal of Geophysical Research (1896-1977), 75, no. 26, 4997–5009, doi: 10.1029/JB075i026p04997.

Cabas, A., A. Rodriguez-Marek, and L. F. Bonilla, 2017, Estimation of Site-Specific Kappa (κ0)-Consistent Damping Values at KiK-Net Sites to Assess the Discrepancy between Laboratory-Based Damping Models and Observed Attenuation (of Seismic Waves) in the FieldEstimation of Site-Specific Kappa (κ0)-Consistent Damping Values at KiK-Net Sites, Bulletin of the Seismological Society of America, 107, no. 5, 2258–2271, doi: 10.1785/0120160370.

Cadet, H., P.-Y. Bard, and A. Rodriguez-Marek, 2010, Defining a Standard Rock Site: Propositions Based on the KiK-net DatabaseDefining a Standard Rock Site: Propositions Based on the KiK-net Database, Bulletin of the Seismological Society of America, 100, no. 1, 172–195, doi: 10.1785/0120090078.

Cadet, H., P.-Y. Bard, and A. Rodriguez-Marek, 2012, Site effect assessment using KiK-net data: Part 1. A simple correction procedure for surface/downhole spectral ratios, Bull Earthquake Eng, 10, no. 2, 421–448, doi: 10.1007/s10518-011-9283-1.

Campbell, K. W., and Y. Bozorgnia, 2014, NGA-West2 Ground Motion Model for the Average Horizontal Components of PGA, PGV, and 5% Damped Linear Acceleration Response Spectra, Earthquake Spectra, 30, no. 3, 1087–1115, doi: 10.1193/062913EQS175M.

Castro, R. R., J. G. Anderson, and S. K. Singh, 1990, Site response, attenuation and source spectra of S waves along the Guerrero, Mexico, subduction zone, Bulletin of the Seismological Society of America, 80, no. 6A, 1481–1503.

Chandler, A. M., N. T. K. Lam, and H. H. Tsang, 2006, Near-surface attenuation modelling based on rock shear-wave velocity profile, Soil Dynamics and Earthquake Engineering, 26, no. 11, 1004–1014, doi: 10.1016/j.soildyn.2006.02.010.

Chandra, J., P. Guéguen, and L. F. Bonilla, 2016, PGA-PGV/Vs considered as a stress–strain proxy for predicting nonlinear soil response, Soil Dynamics and Earthquake Engineering, 85, 146–160, doi: 10.1016/j.soildyn.2016.03.020.

Chiou, B. S.-J., and R. R. Youngs, 2014, Update of the Chiou and Youngs NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra:, Earthquake Spectra, doi: 10.1193/072813EQS219M.

Cotton, F., G. Pousse, F. Bonilla, and F. Scherbaum, 2008, On the Discrepancy of Recent European Ground-Motion Observations and Predictions from Empirical Models: Analysis of KiK-net Accelerometric Data and Point-Sources Stochastic SimulationsDiscrepancy of Recent European Ground-Motion Observations and Predictions from Empirical Models, Bulletin of the Seismological Society of America, 98, no. 5, 2244–2261, doi: 10.1785/0120060084.

Cotton, F., F. Scherbaum, J. J. Bommer, and H. Bungum, 2006, Criteria for Selecting and Adjusting Ground-Motion Models for Specific Target Regions: Application to Central Europe and Rock Sites, J Seismol, 10, no. 2, 137, doi: 10.1007/s10950-005-9006-7.

Dawood, H. M., A. Rodriguez-Marek, J. Bayless, C. Goulet, and E. Thompson, 2016, A Flatfile for the KiK-net Database Processed Using an Automated Protocol, Earthquake Spectra, 32, no. 2, 1281–1302, doi: 10.1193/071214EQS106.

Douglas, J., 2020, Ground motion prediction equations 1964_2020.

Drouet, S., S. Chevrot, F. Cotton, and A. Souriau, 2008, Simultaneous Inversion of Source Spectra, Attenuation Parameters, and Site Responses: Application to the Data of the French Accelerometric NetworkSimultaneous Inversion of Source Spectra, Attenuation Parameters, and Site Responses, Bulletin of the Seismological Society of America, 98, no. 1, 198–219, doi: 10.1785/0120060215.

Drouet, Cotton Fabrice, and Guéguen Philippe, 2010, v S30, κ, regional attenuation and Mw from accelerograms: application to magnitude 3–5 French earthquakes, Geophysical Journal International, 182, no. 2, 880–898, doi: 10.1111/j.1365-246X.2010.04626.x.

Edwards, B., and D. Fäh, 2017, Prediction of earthquake ground motion at rock sites in Japan: evaluation of empirical and stochastic approaches for the PEGASOS Refinement Project, Geophysical Journal International, 211, no. 2, 766–783, doi: 10.1093/gji/ggx328.

Edwards, B., C. Michel, V. Poggi, and D. Fäh, 2013, Determination of Site Amplification from Regional Seismicity: Application to the Swiss National Seismic Networks, Seismological Research Letters, 84, no. 4, 611–621, doi: 10.1785/0220120176.

Edwards, B., V. Poggi, and D. Fäh, 2011, A Predictive Equation for the Vertical-to-Horizontal Ratio of Ground Motion at Rock Sites Based on Shear-Wave Velocity Profiles from Japan and SwitzerlandA Predictive Equation for the Vertical-to-Horizontal Ratio of Ground Motion at Rock Sites, Bulletin of the Seismological Society of America, 101, no. 6, 2998–3019, doi: 10.1785/0120110023.

Edwards, B., A. Rietbrock, J. J. Bommer, and B. Baptie, 2008, The Acquisition of Source, Path, and Site Effects from Microearthquake Recordings Using Q Tomography: Application to the United KingdomAcquisition of Source, Path and Site Effects from Microearthquake Recordings Using Q Tomography, Bulletin of the Seismological Society of America, 98, no. 4, 1915–1935, doi: 10.1785/0120070127.

Eshelby, J. D., 1957, The determination of the elastic field of an ellipsoidal inclusion, and related problems, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 241, no. 1226, 376–396, doi: 10.1098/rspa.1957.0133.

Field, E. H., and K. H. Jacob, 1995, A comparison and test of various site-response estimation techniques, including three that are not reference-site dependent, Bulletin of the Seismological Society of America, 85, no. 4, 1127–1143.

García, D., D. J. Wald, and M. G. Hearne, 2012, A Global Earthquake Discrimination Scheme to Optimize Ground-Motion Prediction Equation SelectionA Global Earthquake Discrimination Scheme to Optimize Ground-Motion Prediction Equation Selection, Bulletin of the Seismological Society of America, 102, no. 1, 185–203, doi: 10.1785/0120110124.

Garofalo, F. et al., 2016, InterPACIFIC project: Comparison of invasive and non-invasive methods for seismic site characterization. Part I: Intra-comparison of surface wave methods, Soil Dynamics and Earthquake Engineering, 82, 222–240, doi: 10.1016/j.soildyn.2015.12.010.

Grendas, I., N. Theodoulidis, P. Hatzidimitriou, B. Margaris, and S. Drouet, 2018, Determination of source, path and site parameters based on non-linear inversion of accelerometric data in Greece, Bull Earthquake Eng, 16, no. 11, 5061–5094, doi: 10.1007/s10518-018-0379-8.

Grendas, I., N. Theodoulidis, F. Hollender, and P. Hatzidimitriou, 2021, A GIT algorithm for simultaneous estimation of seismic source, site response and regional-distance dependent attenuation parameters: application to synthetic and real data, J Seismol, 25, no. 2, 575–598, doi: 10.1007/s10950-020-09975-8.

Griffiths, S. C., B. R. Cox, E. M. Rathje, and D. P. Teague, 2016a, Mapping Dispersion Misfit and Uncertainty in Vs Profiles to Variability in Site Response Estimates, Journal of Geotechnical and Geoenvironmental Engineering, 142, no. 11, 04016062, doi: 10.1061/(ASCE)GT.1943-5606.0001553.

Griffiths, S. C., B. R. Cox, E. M. Rathje, and D. P. Teague, 2016b, Surface-Wave Dispersion Approach for Evaluating Statistical Models That Account for Shear-Wave Velocity Uncertainty, Journal of Geotechnical and Geoenvironmental Engineering, 142, no. 11, 04016061, doi: 10.1061/(ASCE)GT.1943-5606.0001552.

Guéguen, P., 2016, Predicting Nonlinear Site Response Using Spectral Acceleration Vs PGV/Vs30: A Case History Using the Volvi-Test Site, Pure Appl. Geophys., 173, no. 6, 2047–2063, doi: 10.1007/s00024-015-1224-5.

Haghshenas, E., P.-Y. Bard, N. Theodulidis, and SESAME WP04 Team, 2008, Empirical evaluation of microtremor H/V spectral ratio, Bull Earthquake Eng, 6, no. 1, 75–108, doi: 10.1007/s10518-007-9058-x.

Hanks, T. C., and H. Kanamori, 1979, A moment magnitude scale, Journal of Geophysical Research: Solid Earth, 84, no. B5, 2348–2350, doi: https://doi.org/10.1029/JB084iB05p02348.

Hartzell, S. H., 1992, Site response estimation from earthquake data, Bulletin of the Seismological Society of America, 82, no. 6, 2308–2327.

Hashash, Y. M. A., A. R. Kottke, J. P. Stewart, K. W. Campbell, B. Kim, C. Moss, S. Nikolaou, E. M. Rathje, and W. J. Silva, 2014, Reference Rock Site Condition for Central and Eastern North AmericaReference Rock Site Condition for Central and Eastern North America, Bulletin of the Seismological Society of America, 104, no. 2, 684–701, doi: 10.1785/0120130132.

Hollender, F. et al., 2018, Characterization of site conditions (soil class, VS30, velocity profiles) for 33 stations from the French permanent accelerometric network (RAP) using surface-wave methods, Bull Earthquake Eng, 16, no. 6, 2337–2365, doi: 10.1007/s10518-017-0135-5.

Hollender, F., 2019, IMPROVEMENT OF PRACTICES FOR A MORE ROBUST SITE-SPECIFIC SEISMIC HAZARD ASSESSMENT.

Hollender, F., Z. Roumelioti, E. Maufroy, P. Traversa, and A. Mariscal, 2020, Can We Trust High-Frequency Content in Strong-Motion Database Signals? Impact of Housing, Coupling, and Installation Depth of Seismic Sensors, Seismological Research Letters, doi: 10.1785/0220190163.

Horike, M., B. Zhao, and H. Kawase, 2001, Comparison of Site Response Characteristics Inferred from Microtremors and Earthquake Shear Waves, Bulletin of the Seismological Society of America, 91, no. 6, 1526–1536, doi: 10.1785/0120000065.

Houtte, C. V., S. Drouet, and F. Cotton, 2011, Analysis of the Origins of κ (Kappa) to Compute Hard Rock to Rock Adjustment Factors for GMPEsAnalysis of the Origins of κ (Kappa) to Compute Hard Rock to Rock Adjustment Factors for GMPEs, Bulletin of the Seismological Society of America, 101, no. 6, 2926–2941, doi: 10.1785/0120100345.

Joyner, W. B., and D. M. Boore, 1981, Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 imperial valley,

California, earthquake, Bulletin of the Seismological Society of America, 71, no. 6, 2011–2038.

Kaklamanos, J., B. A. Bradley, A. N. Moolacattu, and B. M. Picard, 2020, Physical Hypotheses for Adjusting Coarse Profiles and Improving 1D Site-Response Estimation Assessed at 10 KiK-net Sites, Bulletin of the Seismological Society of America, doi: 10.1785/0120190263.

Kalkan, E., 2016, An Automatic P-Phase Arrival-Time PickerAn Automatic P-Phase Arrival-Time Picker, Bulletin of the Seismological Society of America, 106, no. 3, 971–986, doi: 10.1785/0120150111.

Kawase, H., 2003, Site Effects on Strong Ground Motions, in W. H. Lee, H. Kanamori, P. C. Jennings, and C. Kisslinger (Editors), Elsevier, Amsterdam, 1013–1030.

Kawase, H., and K. Aki, 1989, A study on the response of a soft basin for incident S, P, and Rayleigh waves with special reference to the long duration observed in Mexico City, Bulletin of the Seismological Society of America, 79, no. 5, 1361–1382, doi: 10.1785/BSSA0790051361.

Kawase, H., F. Nagashima, K. Nakano, and Y. Mori, 2019a, Direct evaluation of S-wave amplification factors from microtremor H/V ratios: Double empirical corrections to "Nakamura" method, Soil Dynamics and Earthquake Engineering, 126, 105067, doi: 10.1016/j.soildyn.2018.01.049.

Kawase, H., F. Nagashima, K. Nakano, and Y. Mori, 2019b, Direct evaluation of S-wave amplification factors from microtremor H/V ratios: Double empirical corrections to "Nakamura" method, Soil Dynamics and Earthquake Engineering, 126, 105067, doi: 10.1016/j.soildyn.2018.01.049.

Kennett, B. L. N., 1974, Reflections, rays, and reverberations, Bulletin of the Seismological Society of America, 64, no. 6, 1685–1696.

Klin, P., G. Laurenzano, and E. Priolo, 2018, GITANES: A MATLAB Package for Estimation of Site Spectral Amplification with the Generalized Inversion Technique, Seismological Research Letters, 89, no. 1, 182–190, doi: 10.1785/0220170080.

Kotha, S. R., D. Bindi, and F. Cotton, 2017, From Ergodic to Region- and Site-Specific Probabilistic Seismic Hazard Assessment: Method Development and Application at European and Middle Eastern Sites, Earthquake Spectra, 33, no. 4, 1433–1453, doi: 10.1193/081016eqs130m.

Kotha, S. R., D. Bindi, and F. Cotton, 2016, Partially non-ergodic region specific GMPE for Europe and Middle-East, Bull Earthquake Eng, 14, no. 4, 1245–1263, doi: 10.1007/s10518-016-9875-x.

Kotha, S. R., F. Cotton, and D. Bindi, 2018, A new approach to site classification: Mixed-effects Ground Motion Prediction Equation with spectral clustering of site amplification functions, Soil Dynamics and Earthquake Engineering, 110, 318–329, doi: 10.1016/j.soildyn.2018.01.051.

Kottke, A. R., and E. M. Rathje, 2008, "Technical manual for Strata." Report No.: 2008/10., Pacific Earthquake Engineering Research Center, University of California, Berkeley.

Kramer, S. L., 1996, Geotechnical Earthquake Engineering, Prentice-Hall civil engineering and engineering mechanics series, Upper Saddle River, New Jersey 07458, United States of America.

Ktenidou, O.-J., and N. A. Abrahamson, 2016, Empirical Estimation of High-Frequency Ground Motion on Hard Rock, Seismological Research Letters, 87, no. 6, 1465–1478, doi: 10.1785/0220160075.

Ktenidou, O.-J., N. A. Abrahamson, S. Drouet, and F. Cotton, 2015, Understanding the physics of kappa (κ): insights from a downhole array, Geophysical Journal International, 203, no. 1, 678–691, doi: 10.1093/gji/ggv315.

Ktenidou, O.-J., F. Cotton, N. A. Abrahamson, and J. G. Anderson, 2014, Taxonomy of κ: A Review of Definitions and Estimation Approaches Targeted to Applications, Seismological Research Letters, 85, no. 1, 135–146, doi: 10.1785/0220130027.

Landwehr, N., N. M. Kuehn, T. Scheffer, and N. Abrahamson, 2016, A Nonergodic Ground-Motion Model for California with Spatially Varying Coefficients, Bulletin of the Seismological Society of America, 106, no. 6, 2574–2583, doi: 10.1785/0120160118.

Lanzano, G., S. Sgobba, L. Luzi, R. Puglia, F. Pacor, C. Felicetta, M. D'Amico, F. Cotton, and D. Bindi, 2019, The pan-European Engineering Strong Motion (ESM) flatfile: compilation criteria and data statistics, Bull Earthquake Eng, 17, no. 2, 561–582, doi: 10.1007/s10518-018-0480-z.

Laurendeau, A., P.-Y. Bard, F. Hollender, V. Perron, L. Foundotos, O.-J. Ktenidou, and B. Hernandez, 2018, Derivation of consistent hard rock (1000 < VS < 3000 m/s) GMPEs from surface and down-hole recordings: analysis of KiK-net data, Bull Earthquake Eng, 16, no. 6, 2253–2284, doi: 10.1007/s10518-017-0142-6.

Laurendeau, A., F. Cotton, O.-J. Ktenidou, L.-F. Bonilla, and F. Hollender, 2013, Rock and Stiff-Soil Site Amplification: Dependency on VS30 and Kappa (κ0)Rock and Stiff-Soil Site Amplification: Dependency on VS30 and Kappa (κ0), Bulletin of the Seismological Society of America, 103, no. 6, 3131–3148, doi: 10.1785/0120130020.

Lermo, J., and F. J. Chávez-García, 1993, Site effect evaluation using spectral ratios with only one station, Bulletin of the Seismological Society of America, 83, no. 5, 1574–1594.

Mayor, J., M. Calvet, L. Margerin, O. Vanderhaeghe, and P. Traversa, 2016, Crustal structure of the Alps as seen by attenuation tomography, Earth and Planetary Science Letters, 439, 71–80, doi: 10.1016/j.epsl.2016.01.025.

Mayor, J., P. Traversa, M. Calvet, and L. Margerin, 2018, Tomography of crustal seismic attenuation in Metropolitan France: implications for seismicity analysis, Bull Earthquake Eng, 16, no. 6, 2195–2210, doi: 10.1007/s10518-017-0124-8.

Nakamura, Y., 1989, A METHOD FOR DYNAMIC CHARACTERISTICS ESTIMATION OF SUBSURFACE USING MICROTREMOR ON THE GROUND SURFACE, Railway Technical Research Institute, Quarterly Reports, 30, no. 1.

Nakano, K., S. Matsushima, and H. Kawase, 2015, Statistical Properties of Strong Ground Motions from the Generalized Spectral Inversion of Data Observed by K-NET, KiK-net, and the JMA Shindokei Network in JapanStatistical Properties of Strong Ground Motions from the Generalized Spectral Inversion of Data, Bulletin of the Seismological Society of America, 105, no. 5, 2662–2680, doi: 10.1785/0120140349.

Newmark, N. M., and W. J. Hall, 1982, Earthquake spectra and design, Earth System Dynamics.

Okada, Y., K. Kasahara, S. Hori, K. Obara, S. Sekiguchi, H. Fujiwara, and A. Yamamoto, 2004, Recent progress of seismic observation networks in Japan —Hi-net, F-net, K-NET and KiK-net—, Earth Planet Sp, 56, no. 8, xv–xxviii, doi: 10.1186/BF03353076.

Oth, A., D. Bindi, S. Parolai, and D. D. Giacomo, 2011, Spectral Analysis of K-NET and KiK-net Data in Japan, Part II: On Attenuation Characteristics, Source Spectra, and Site Response of Borehole and Surface StationsSpectral Analysis of K-NET and KiK-net Data in Japan, Part II, Bulletin of the Seismological Society of America, 101, no. 2, 667–687, doi: 10.1785/0120100135.

Oth, A., H. Miyake, and D. Bindi, 2017, On the relation of earthquake stress drop and ground motion variability, Journal of Geophysical Research: Solid Earth, 122, no. 7, 5474–5492, doi: https://doi.org/10.1002/2017JB014026.

Oth, A., H. Miyake, and D. Bindi, 2015, On the Relation of Earthquake Stress Drop and Ground Motion Variability, AGU Fall Meeting Abstracts, 51, S51A-2641.

Pacor, F., D. Spallarossa, A. Oth, L. Luzi, R. Puglia, L. Cantore, A. Mercuri, M. D'Amico, and D. Bindi, 2016, Spectral models for ground motion prediction in the L'Aquila region (central Italy): evidence for stress-drop dependence on magnitude and depth, Geophys J Int, 204, no. 2, 697–718, doi: 10.1093/gji/ggv448.

Parolai, S., 2018, κ 0 : Origin and Usability, Bulletin of the Seismological Society of America, doi: 10.1785/0120180135.

Parolai, S., D. Bindi, M. Baumbach, H. Grosser, C. Milkereit, S. Karakisa, and S. Zünbül, 2004, Comparison of Different Site Response Estimation Techniques Using Aftershocks of the 1999 Izmit Earthquake, Bulletin of the Seismological Society of America, 94, no. 3, 1096–1108, doi: 10.1785/0120030086.

Parolai, S., D. Bindi, and M. Pilz, 2015, k0: The role of Intrinsic and Scattering AttenuationShort Note, Bulletin of the Seismological Society of America, 105, no. 2A, 1049–1052, doi: 10.1785/0120140305.

Passeri, F., S. Foti, B. R. Cox, and A. Rodriguez-Marek, 2019, Influence of Epistemic Uncertainty in Shear Wave Velocity on Seismic Ground Response Analyses, Earthquake Spectra, 35, no. 2, 929–954, doi: 10.1193/011018EQS005M.

Passeri, F., S. Foti, and A. Rodriguez-Marek, 2020, A new geostatistical model for shear wave velocity profiles, Soil Dynamics and Earthquake Engineering, 136, 106247, doi: 10.1016/j.soildyn.2020.106247.

Pei, S., Z. Cui, Y. Sun, M. N. Toksöz, C. A. Rowe, X. Gao, J. Zhao, H. Liu, J. He, and F. D. Morgan, 2009, Structure of the Upper Crust in Japan from S-Wave Attenuation TomographyShort Note, Bulletin of the Seismological Society of America, 99, no. 1, 428–434, doi: 10.1785/0120080029.

Perron, V., F. Hollender, P.-Y. Bard, C. Gélis, C. Guyonnet-Benaize, B. Hernandez, and O.-J. Ktenidou, 2017, Robustness of Kappa (κ) Measurement in Low-to-Moderate Seismicity Areas: Insight from a Site-Specific Study in Provence, FranceRobustness of Kappa (κ) Measurement in Low-to-Moderate Seismicity Areas, Bulletin of the Seismological Society of America, 107, no. 5, 2272–2292, doi: 10.1785/0120160374.

Perron, V., A. Laurendeau, F. Hollender, P.-Y. Bard, C. Gélis, P. Traversa, and S. Drouet, 2018, Selecting time windows of seismic phases and noise for engineering seismology applications: a versatile methodology and algorithm, Bull Earthquake Eng, 16, no. 6, 2211–2225, doi: 10.1007/s10518-017-0131-9.

Pilz, M., and F. Cotton, 2019, Does the One-Dimensional Assumption Hold for Site Response Analysis? A Study of Seismic Site Responses and Implication for Ground Motion Assessment Using KiK-Net Strong-Motion Data:, Earthquake Spectra, doi: 10.1193/050718EQS113M.

Pilz, M., and D. Fäh, 2017, The contribution of scattering to near-surface attenuation, Journal of Seismology, 21, no. 4, 837–855, doi: 10.1007/s10950-017-9638-4.

Poggi, V., B. Edwards, and D. Faeh, 2012, The quarter-wavelength average velocity : a review of some past and recent application developments: </paper/The-quarter-wavelength-average-velocity-%3A-a-review-Poggi-Edwards/0302486e4c2173d4ce9f02203f076ccead78a1aa> (accessed October 2, 2020).

Renault, P., 2014, Approach and Challenges for the Seismic Hazard Assessment of Nuclear Power Plants: the Swiss experience, BGTA, doi: 10.4430/bgta0089.

Rodriguez-Marek, A., F. Cotton, N. A. Abrahamson, S. Akkar, L. A. Atik, B. Edwards, G. A. Montalva, and H. M. Dawood, 2013, A Model for Single-Station Standard Deviation Using Data from Various Tectonic RegionsA Model for Single-Station Standard Deviation Using Data from Various Tectonic Regions, Bulletin of the Seismological Society of America, 103, no. 6, 3149–3163, doi: 10.1785/0120130030.

Rodriguez-Marek, A., G. A. Montalva, F. Cotton, and F. Bonilla, 2011, Analysis of Single-Station Standard Deviation Using the KiK-net DataAnalysis of Single-Station Standard Deviation Using the KiK-net Data, Bulletin of the Seismological Society of America, 101, no. 3, 1242–1258, doi: 10.1785/0120100252.

Rodriguez-Marek, A., E. M. Rathje, J. J. Bommer, F. Scherbaum, and P. J. Stafford, 2014, Application of Single-Station Sigma and Site-Response Characterization

in a Probabilistic Seismic-Hazard Analysis for a New Nuclear Site, Bulletin of the Seismological Society of America, 104, no. 4, 1601–1619, doi: 10.1785/0120130196.

Salazar, W., V. Sardina, and J. de Cortina, 2007, A Hybrid Inversion Technique for the Evaluation of Source, Path, and Site Effects Employing S-Wave Spectra for Subduction and Upper-Crustal Earthquakes in El Salvador, Bulletin of the Seismological Society of America, 97, no. 1B, 208–221, doi: 10.1785/0120060076.

Sato, H., M. C. Fehler, and T. Maeda, 2012, Seismic Wave Propagation and Scattering in the Heterogeneous Earth : Second Edition, Springer Science & Business Media.

Shible, H., A. Laurendeau, P.-Y. Bard, and F. Hollender, 2018, Importance of local scattering in high frequency motion: lessons from Interpacific project sites, application to the KiK-net database and derivation of new hard-rock GMPE.

Stafford, P. J., 2014, Crossed and Nested Mixed-Effects Approaches for Enhanced Model Development and Removal of the Ergodic Assumption in Empirical Ground-Motion ModelsCrossed and Nested Mixed-Effects Approaches for Enhanced Model Development, Bulletin of the Seismological Society of America, 104, no. 2, 702–719, doi: 10.1785/0120130145.

Steidl, J. H., A. G. Tumarkin, and R. J. Archuleta, 1996, What is a reference site?, Bulletin of the Seismological Society of America, 86, no. 6, 1733–1748.

Sucuoğlu, H., S. Yücemen, A. Gezer, and A. Erberik, 1998, Statistical evaluation of the damage potential of earthquake ground motions, Structural Safety, 20, no. 4, 357–378, doi: 10.1016/S0167-4730(98)00018-6.

Takemura, M., M. Motosaka, and H. Yamanaka, 1995, Strong Motion Seismology in Japan, Journal of Physics of the Earth, 43, no. 3, 211–257, doi: 10.4294/jpe1952.43.211.

Tarantola, A., 2005, Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics.

Teague, D. P., and B. R. Cox, 2016, Site response implications associated with using non-unique Vs profiles from surface wave inversion in comparison with other commonly used methods of accounting for Vs uncertainty, Soil Dynamics and Earthquake Engineering, 91, 87–103, doi: 10.1016/j.soildyn.2016.07.028.

Teague, D. P., B. R. Cox, and E. M. Rathje, 2018, Measured vs. predicted site response at the Garner Valley Downhole Array considering shear wave velocity uncertainty from borehole and surface wave methods, Soil Dynamics and Earthquake Engineering, 113, 339–355, doi: 10.1016/j.soildyn.2018.05.031.

Thompson, E. M., L. G. Baise, R. E. Kayen, and B. B. Guzina, 2009, Impediments to Predicting Site Response: Seismic Property Estimation and Modeling SimplificationsImpediments to Predicting Site Response: Seismic Property Estimation and Modeling Simplifications, Bulletin of the Seismological Society of America, 99, no. 5, 2927–2949, doi: 10.1785/0120080224.

Thompson, E. M., L. G. Baise, Y. Tanaka, and R. E. Kayen, 2012, A taxonomy of site response complexity, Soil Dynamics and Earthquake Engineering, 41, 32–43, doi: 10.1016/j.soildyn.2012.04.005.

Thomson, W. T., 1950, Transmission of Elastic Waves through a Stratified Solid Medium, Journal of Applied Physics, 21, no. 2, 89–93, doi: 10.1063/1.1699629.

Toro, G. R., 1995, Probabilistic models of site velocity profiles for generic and site-specific ground-motion amplification studies., Technical Rep, 779574.

Traversa, P. et al., 2020, RESIF RAP and RLBP Dataset of Earthquake Ground Motion in Mainland France, Seismological Research Letters, 91, no. 4, 2409–2424, doi: 10.1785/0220190367.

Wang, Z., and D. Zhao, 2019, Updated attenuation tomography of Japan subduction zone, Geophys J Int, 219, no. 3, 1679–1697, doi: 10.1093/gji/ggz339.

Wang, Z., D. Zhao, X. Liu, C. Chen, and X. Li, 2017, P and S wave attenuation tomography of the Japan subduction zone, Geochemistry, Geophysics, Geosystems, 18, no. 4, 1688–1710, doi: 10.1002/2017GC006800.

Zhao, J. X. et al., 2006, Attenuation Relations of Strong Ground Motion in Japan Using Site Classification Based on Predominant Period, Bulletin of the Seismological Society of America, 96, no. 3, 898–913, doi: 10.1785/0120050122.

Zhao, J. X. et al., 2016, Ground-Motion Prediction Equations for Subduction Interface Earthquakes in Japan Using Site Class and Simple Geometric Attenuation FunctionsGMPEs for Subduction Interface Earthquakes in Japan, Bulletin of the Seismological Society of America, 106, no. 4, 1518–1534, doi: 10.1785/0120150034.

Zhu, C., M. Pilz, and F. Cotton, 2020, Evaluation of a novel application of earthquake HVSR in site-specific amplification estimation, Soil Dynamics and Earthquake Engineering, 139, 106301, doi: 10.1016/j.soildyn.2020.106301.

Zhu, T. J., W. K. Tso, and A. C. Heidebrecht, 1988, Effect of Peak Ground a/v Ratio on Structural Damage, Journal of Structural Engineering, 114, no. 5, 1019–1037, doi: 10.1061/(ASCE)0733-9445(1988)114:5(1019).

Appendix

A.1. Matrix formulations of Generalized Inversion Techniques

The basic equation of generalized inversion techniques (1.12) is used to construct the system of equations. However, depending on the input and output requirements, the system can have different definitions in the linear form ($Ax = b$). In each of the following, A is the system matrix, x is the vector of unknows, and b is the vector of observed Fourier spectra at a given frequency. Once they are well-defined, these systems are generally resolved in the least-squares sense.

- Attenuation can be retrieved in a first step by assuming that it is an implicit function containing all effects, with a smooth decrease from the starting distance $r = r_0$. This is based on the idea that amplitudes decrease with distance and that any undulations are related to other factors. The system can be represented at each frequency point, as follows:

$$\underbrace{\begin{pmatrix} & 1 & 0 & 0 & & & & 1 & 0 & 0 \\ & 0 & 1 & 0 & & & \cdots & 0 & 1 & 0 \\ & 0 & 0 & 1 & & & & 0 & 0 & 1 \\ & & \vdots & & & & & & \vdots & \\ w_1 & & 0 & & 0 & 0 & & & \cdot & \\ -w_2/2 & & w_2 & & w_2/2 & 0 & \cdots & & \vdots & \\ 0 & & -w_2/2 & & w_2 & w_2/2 & & & \cdot & \\ \cdot & & \cdot & & \cdot & \cdot & & & & \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} \log_{10}(A_1) \\ \log_{10}(A_2) \\ \vdots \\ \log_{10}(A_{N_d}) \\ \\ \log_{10}(\hat{S}_1) \\ \log_{10}(\hat{S}_2) \\ \vdots \\ \log_{10}(\hat{S}_{N_r}) \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} \log_{10}(FAS_1) \\ \log_{10}(FAS_2) \\ \vdots \\ \log_{10}(FAS_{N_r}) \end{pmatrix}}_{b}$$

  where A_m is attenuation function at each distance r_m, with m=1 to N_d, which is smoothed using the coefficients w_1 and w_2 (predefined constants). $\hat{S}_k$ is the attenuation-corrected Fourier spectra ($FAS(f)/A(r,f)$), with k =1 to N_r, which is expected to contain only source and site contributions. FAS_k is the observed Fourier spectra at each event-site couple, with k = 1 to N_r.

- Similarly, source and site contributions can be separated using Eq(4.8) as follows:

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & & 1 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 1 & 0 \\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} \log_{10}(G_1) \\ \log_{10}(G_2) \\ \vdots \\ \log_{10}(G_{N_e}) \\ \\ \log_{10}(S_1) \\ \log_{10}(S_2) \\ \vdots \\ \log_{10}(S_{N_s}) \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} \log_{10}(FAS_1) \\ \log_{10}(FAS_2) \\ \vdots \\ \log_{10}(FAS_{N_r}) \end{pmatrix}}_{b}$$

where G_i is the source function, S_j is the site function, with i and j go from 1 to N_e and N_s, respectively.

- However, the system can be formulated to be solved all at once, resulting in a single step inversion. This can be obtained by considering the following formulation:

$$\underbrace{\begin{pmatrix} \overbrace{\begin{matrix}1&0&0\\0&1&0\\0&0&1\end{matrix}}^{Source} & \cdots & \overbrace{\begin{matrix}1&0&0\\0&1&0\\0&0&1\end{matrix}}^{attenuation} & \cdots & \overbrace{\begin{matrix}1&0&0\\0&1&0\\0&0&1\end{matrix}}^{site} \\ \vdots & & \vdots & & \vdots \\ \cdot & & \begin{matrix}w_1 & 0 & 0 & 0\\ -w_2/2 & w_2 & w_2/2 & 0\\ 0 & -w_2/2 & w_2 & w_2/2\\ \cdot & \cdot & \cdot & \cdot\end{matrix} & \cdots & \cdot \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} \log_{10}(G_1) \\ \vdots \\ \log_{10}(G_{N_e}) \\ \log_{10}(A_1) \\ \vdots \\ \log_{10}(A_{N_d}) \\ \log_{10}(S_1) \\ \vdots \\ \log_{10}(S_{N_s}) \end{pmatrix}}_{x}$$

$$= \underbrace{\begin{pmatrix} \log_{10}(FAS_1) \\ \log_{10}(FAS_2) \\ \vdots \\ \log_{10}(FAS_{N_r}) \end{pmatrix}}_{b}$$

On the other hand, generalized inversions can rely on parametrized forms of each source and site contributions. This can lead to highly non-linear systems, with the impossibility of solving linearly in the least-squares methods. For example, authors attempting to apply parametrized inversions relied on the gaussian-newton approach to linearize and invert the obtained system (Drouet et al., 2008; Grendas et al., 2021). This method is based on defining a vector of model parameters m, whose dimension is equals to those of the system unknowns. Then, the iterative gauss-newton approach starts from an initial input m_{prior} and proposes a perturbation Δm on the model m_n at iteration $n+1$, as follows:

$$\Delta m = m_{n+1} - m_n = (G_n^t C_D^{-1} G_n + C_M^{-1})^{-1}\{C_n^t[g(m_n) - d_{obs}]\} + C_M^{-1}(m_n - m_{prior})$$

where G_n is the matrix of partial derivatives, C_D and C_M are model covariance matrices, $g(m)$ is the matrix of the forward model, and d_{obs} is the vector of observations. For more details, one can refer to Tarantola, (2005).

A.2. Final of list 1D sites of chapter 4

	Station	V_{S30} (m/s)	V_{Sdh} (m/s)	Number of recordings	r_p(relative)	r_p(absolute)	r_s(absolute)
1.	***ABSH02***	501	1430	6	0.86	0.71	0.63
2.	***ABSH03***	499	1630	6	0.89	0.98	0.92
3.	***AICH08***	449	1930	30	0.61	0.73	0.89
4.	***AICH15***	663	2600	49	0.87	0.94	0.87

5.	***AICH16***	364	2600	64	0.60	0.90	0.94
6.	***AICH17***	314	2200	59	0.66	0.67	0.76
7.	***AICH18***	499	2000	54	0.88	0.86	0.86
8.	***AICH20***	691	2900	42	0.69	0.93	0.80
9.	***AOMH12***	281	820	58	0.80	0.85	0.87
10.	***AOMH14***	361	770	19	0.60	0.67	0.65
11.	***AOMH17***	378	1450	95	0.82	0.74	0.77
12.	***EHMH10***	318	2880	24	0.76	0.77	0.79
13.	***EHMH11***	712	2570	21	0.68	0.79	0.75
14.	***FKIH02***	343	1840	19	0.86	0.83	0.79
15.	***FKIH07***	384	2180	49	0.82	0.71	0.77
16.	***FKOH01***	588	3100	45	0.84	0.85	0.89
17.	***FKOH05***	777	3000	28	0.63	0.94	0.63
18.	***FKOH08***	536	1600	112	0.67	0.69	0.84
19.	***FKSH07***	829	2600	92	0.88	0.97	0.80
20.	***FKSH09***	585	1960	395	0.67	0.69	0.94
21.	***FKSH10***	487	870	385	0.92	0.97	0.79
22.	***FKSH12***	449	2320	371	0.66	0.88	0.94
23.	***FKSH13***	562	1970	14	0.70	0.82	0.78
24.	***FKSH17***	544	1970	312	0.67	0.68	0.70
25.	***FKSH18***	307	2250	354	0.69	0.74	0.78
26.	***GIFH04***	380	2770	39	0.62	0.65	0.76
27.	***GIFH11***	904	2680	60	0.68	0.96	0.68
28.	***GIFH12***	667	1500	45	0.62	0.87	0.86
29.	***GIFH14***	627	2700	46	0.61	0.90	0.82
30.	***GIFH17***	429	1600	39	0.84	0.78	0.81
31.	***GIFH23***	588	1900	43	0.88	0.95	0.89
32.	***GIFH25***	469	2100	35	0.86	0.79	0.72
33.	***GNMH12***	407	1600	80	0.71	0.82	0.89
34.	***GNMH13***	323	2400	70	0.60	0.79	0.71
35.	***HRSH05***	371	2850	20	0.80	0.74	0.63
36.	***HRSH06***	279	1650	29	0.60	0.69	0.78
37.	***HRSH15***	579	2070	24	0.87	0.78	0.65
38.	***HRSH16***	455	2520	18	0.86	0.82	0.84
39.	***HRSH18***	402	2530	17	0.84	0.80	0.85
40.	***HYGH04***	472	2020	37	0.87	0.77	0.94
41.	***HYGH09***	364	1028	47	0.79	0.78	0.79
42.	***HYGH14***	698	1930	46	0.78	0.90	0.82
43.	***IBRH10***	144	2530	268	0.89	0.96	0.97
44.	***IBRH11***	242	2100	366	0.61	0.86	0.83
45.	***IBRH12***	486	1700	414	0.91	0.95	0.74
46.	***IBRH13***	335	3000	433	0.91	0.83	0.81
47.	***IBRH15***	450	1700	416	0.89	0.72	0.93

48.	***IBRH18***	559	2200	295	0.87	0.89	0.92
49.	***IBUH07***	259	1330	18	0.70	0.92	0.91
50.	***ISKH03***	311	870	69	0.81	0.68	0.83
51.	***ISKH07***	440	1080	34	0.67	0.70	0.84
52.	***IWTH01***	438	1300	81	0.89	0.82	0.91
53.	***IWTH02***	390	2300	109	0.86	0.92	0.96
54.	***IWTH05***	429	2600	162	0.64	0.79	0.97
55.	***IWTH07***	396	2040	115	0.89	0.90	0.85
56.	***IWTH09***	967	2610	124	0.68	0.98	0.68
57.	***IWTH18***	892	2630	156	0.97	0.99	0.83
58.	***IWTH19***	482	1270	154	0.86	0.95	0.66
59.	***IWTH21***	521	2460	135	0.93	0.94	0.88
60.	***IWTH22***	532	2780	173	0.88	0.94	0.98
61.	***IWTH23***	923	2200	192	0.74	0.95	0.83
62.	***IWTH27***	670	2790	256	0.80	0.92	0.99
63.	***KGWH04***	407	1760	11	0.82	0.79	0.86
64.	***KGWH05***	599	3500	11	0.64	0.73	0.71
65.	***KKWH15***	529	1300	6	0.85	0.91	0.67
66.	***KMMH05***	230	820	18	0.91	0.72	0.74
67.	***KNGH18***	388	1200	87	0.75	0.75	0.81
68.	***KNGH19***	731	1800	50	0.60	0.84	0.66
69.	***KNGH21***	792	2400	55	0.86	0.96	0.72
70.	***KOCH07***	385	1950	11	0.62	0.82	0.87
71.	***KOCH13***	530	2410	15	0.84	0.78	0.69
72.	***KYTH02***	518	1860	22	0.87	0.92	0.61
73.	***KYTH03***	637	1940	38	0.76	0.92	0.85
74.	***MIEH03***	435	2430	41	0.81	0.77	0.79
75.	***MIEH05***	590	1920	51	0.71	0.80	0.90
76.	***MIEH09***	608	2440	21	0.96	0.98	0.71
77.	***MYGH03***	934	2630	190	0.81	0.98	0.77
78.	***MYGH04***	850	2830	176	0.86	1.00	0.97
79.	***MYGH05***	305	690	59	0.77	0.63	0.76
80.	***MYGH08***	203	970	370	0.76	0.64	0.60
81.	***MYGH11***	859	2780	148	0.95	0.96	0.97
82.	***MYZH12***	319	1900	17	0.87	0.76	0.84
83.	***MYZH15***	446	2650	103	0.95	0.90	0.82
84.	***NARH03***	492	1832	42	0.62	0.83	0.74
85.	***NARH06***	371	1791	49	0.88	0.77	0.90
86.	***NARH07***	655	1900	11	0.82	0.95	0.61
87.	***NGNH09***	771	1800	15	0.78	0.96	0.90
88.	***NGNH13***	513	1160	60	0.81	0.91	0.74
89.	***NGNH17***	609	2700	36	0.86	0.92	0.87
90.	***NGNH20***	530	2400	52	0.84	0.90	0.82

91.	***NGNH23***	524	1100	80	0.65	0.86	0.86
92.	***NGNH30***	457	1000	68	0.85	0.82	0.61
93.	***NGSH01***	398	1700	32	0.61	0.62	0.84
94.	***NIGH07***	528	1600	101	0.64	0.99	0.95
95.	***NIGH11***	375	850	122	0.87	0.78	0.75
96.	***NIGH13***	461	910	58	0.65	0.67	0.70
97.	***NIGH16***	525	2250	52	0.60	0.68	0.71
98.	***OITH07***	276	1600	14	0.61	0.88	0.92
99.	***OKYH10***	554	2600	19	0.80	0.76	0.83
100.	***SAGH03***	434	1980	28	0.88	0.87	0.78
101.	***SAGH04***	724	3000	69	0.85	0.95	0.65
102.	***SIGH02***	569	1500	21	0.88	0.84	0.78
103.	***SITH06***	369	1100	128	0.82	0.67	0.67
104.	***SITH08***	521	1600	99	0.80	0.97	0.89
105.	***SITH11***	372	1800	110	0.90	0.75	0.85
106.	***SMNH03***	435	1300	29	0.61	0.76	0.71
107.	***SMNH06***	288	2200	15	0.72	0.66	0.61
108.	***SRCH05***	410	2010	5	0.91	0.78	0.78
109.	***SZOH30***	506	2000	47	0.69	0.95	0.71
110.	***SZOH33***	520	2100	19	0.94	0.86	0.65
111.	***SZOH34***	430	1100	42	0.90	0.78	0.66
112.	***SZOH35***	158	1470	30	0.60	0.89	0.85
113.	***SZOH37***	366	1400	27	0.85	0.93	0.86
114.	***SZOH43***	323	1080	9	0.87	0.77	0.76
115.	***TCGH07***	419	1870	154	0.71	0.89	0.89
116.	***TCGH11***	329	1600	177	0.93	0.92	0.82
117.	***TCGH13***	574	2010	365	0.68	0.75	0.92
118.	***TCGH14***	849	2300	249	0.86	0.88	0.76
119.	***TCGH15***	423	1170	19	0.95	0.83	0.66
120.	***TKCH05***	337	640	23	0.87	0.89	0.79
121.	***TKCH07***	140	530	9	0.82	0.74	0.86
122.	***TKSH04***	475	1880	16	0.77	0.66	0.64
123.	***TKSH05***	380	1740	29	0.72	0.68	0.62
124.	***TKYH12***	326	1800	83	0.80	0.86	0.81
125.	***TTRH03***	189	2000	29	0.91	0.90	0.76
126.	***TTRH04***	254	1350	14	0.86	0.78	0.66
127.	***WKYH01***	463	1580	34	0.66	0.68	0.85
128.	***WKYH02***	369	2180	14	0.97	0.98	0.86
129.	***WKYH10***	466	2300	26	0.72	0.83	0.63
130.	***YMGH10***	526	2400	10	0.68	0.64	0.71
131.	***YMGH12***	1138	3250	10	0.84	0.97	0.63
132.	***YMNH09***	768	1830	43	0.66	0.71	0.63
133.	***YMNH11***	295	1530	73	0.74	0.87	0.81

A.3. Tables of GMPE coefficients of Chapter 5

$SURF_{1D}$

freq	a_1	a_2	a_3	a_4	b_1	b_2	h	b_3	b_4	M_h	M_{ref}	R_{break}	c_1
0.520	5.621	2.365	0.117	0.572	-1.153	0.013	-2.174	-0.423	-0.004	5.993	4.400	90.000	-0.674
0.590	5.852	2.354	0.100	0.576	-1.159	0.001	-2.231	-0.340	-0.005	5.971	4.400	90.000	-0.689
0.670	5.961	2.279	0.078	0.562	-1.156	0.001	-2.262	-0.299	-0.005	5.949	4.400	90.000	-0.714
0.761	5.999	2.184	0.055	0.584	-1.140	0.006	-2.166	-0.253	-0.005	5.927	4.400	90.000	-0.723
0.864	6.035	2.081	0.030	0.593	-1.127	0.008	-2.091	-0.216	-0.006	5.906	4.400	90.000	-0.754
0.981	6.176	2.048	0.017	0.577	-1.120	0.003	-2.023	-0.248	-0.006	5.884	4.400	90.000	-0.804
1.113	6.348	2.042	0.009	0.683	-1.116	-0.006	-1.977	-0.216	-0.006	5.863	4.400	90.000	-0.856
1.264	6.501	1.976	-0.012	0.783	-1.117	-0.013	-1.979	-0.203	-0.006	5.841	4.400	90.000	-0.900
1.435	6.726	1.998	-0.019	0.901	-1.116	-0.032	-1.982	-0.107	-0.007	5.820	4.400	90.000	-0.951
1.628	6.931	2.015	-0.025	1.032	-1.115	-0.048	-1.991	-0.070	-0.008	5.798	4.400	90.000	-1.022
1.849	7.034	1.985	-0.023	1.007	-1.113	-0.046	-1.948	-0.047	-0.008	5.776	4.400	90.000	-1.110
2.099	7.211	2.028	-0.006	0.965	-1.108	-0.052	-2.116	-0.034	-0.009	5.755	4.400	90.000	-1.179
2.382	7.295	2.016	0.000	0.954	-1.091	-0.056	-2.404	0.133	-0.010	5.734	4.400	90.000	-1.225
2.705	7.523	2.052	0.014	0.960	-1.093	-0.066	-2.818	0.073	-0.010	5.713	4.400	90.000	-1.252
3.070	7.639	1.963	0.018	0.948	-1.099	-0.053	-2.618	0.001	-0.010	5.692	4.400	90.000	-1.248
3.486	7.779	1.906	0.020	0.938	-1.099	-0.046	-2.478	0.010	-0.011	5.671	4.400	90.000	-1.231
3.957	7.939	1.906	0.034	0.869	-1.093	-0.046	2.189	-0.099	-0.010	5.650	4.400	90.000	-1.207
4.492	8.159	1.939	0.068	0.821	-1.094	-0.046	1.834	-0.200	-0.010	5.629	4.400	90.000	-1.110
5.099	8.316	1.913	0.091	0.729	-1.096	-0.036	1.752	-0.239	-0.010	5.609	4.400	90.000	-1.003
5.789	8.486	1.875	0.098	0.722	-1.105	-0.037	1.847	-0.195	-0.011	5.588	4.400	90.000	-0.896
6.572	8.487	1.781	0.109	0.637	-1.089	-0.022	1.866	-0.308	-0.010	5.568	4.400	90.000	-0.744
7.460	8.380	1.595	0.114	0.460	-1.085	0.011	-1.814	-0.420	-0.010	5.547	4.400	90.000	-0.613
8.469	8.494	1.559	0.145	0.374	-1.112	0.031	-1.756	-0.624	-0.009	5.527	4.400	90.000	-0.502
9.614	8.504	1.456	0.157	0.374	-1.133	0.050	-1.913	-0.681	-0.009	5.506	4.400	90.000	-0.390
10.914	8.477	1.411	0.172	0.306	-1.131	0.060	-2.240	-0.909	-0.008	5.500	4.400	90.000	-0.300
12.390	8.487	1.429	0.180	0.348	-1.135	0.047	-2.175	-0.936	-0.008	5.500	4.400	90.000	-0.244
14.066	8.382	1.373	0.194	0.251	-1.144	0.062	-1.909	-1.002	-0.008	5.500	4.400	90.000	-0.192
15.968	8.355	1.366	0.199	0.243	-1.171	0.062	-1.856	-1.017	-0.008	5.500	4.400	90.000	-0.192
18.127	8.256	1.328	0.195	0.225	-1.199	0.070	-1.849	-1.007	-0.008	5.500	4.400	90.000	-0.253
	8.236	1.356	0.195	0.234	-1.222	0.069	-1.856	-1.040	-0.007	5.500	4.400	90.000	-0.306

SURF$_{1D}$cor$_{1DSH}$

freq	a_1	a_2	a_3	a_4	b_1	b_2	h	b_3	b_4	M_h	M_{ref}	R_{break}	c_1
0.520	6.307	2.347	-0.006	0.554	-1.168	0.026	2.804	-0.176	-0.005	5.993	4.400	90.000	-0.478
0.590	6.561	2.292	-0.029	0.571	-1.177	0.017	2.979	-0.078	-0.005	5.971	4.400	90.000	-0.521
0.670	6.715	2.217	-0.053	0.587	-1.170	0.010	3.022	-0.055	-0.005	5.949	4.400	90.000	-0.561
0.761	6.736	2.113	-0.078	0.624	-1.143	0.012	-2.778	-0.061	-0.006	5.927	4.400	90.000	-0.592
0.864	6.774	1.997	-0.100	0.632	-1.126	0.014	-2.564	-0.062	-0.006	5.906	4.400	90.000	-0.617
0.981	6.972	1.974	-0.105	0.633	-1.124	0.007	-2.477	-0.130	-0.006	5.884	4.400	90.000	-0.631
1.113	7.123	1.969	-0.103	0.748	-1.116	-0.002	-2.322	-0.104	-0.006	5.863	4.400	90.000	-0.624
1.264	7.219	1.889	-0.114	0.833	-1.110	-0.005	-2.162	-0.086	-0.006	5.841	4.400	90.000	-0.641
1.435	7.459	1.914	-0.115	0.992	-1.109	-0.028	-2.076	-0.007	-0.007	5.820	4.400	90.000	-0.657
1.628	7.624	1.942	-0.105	1.088	-1.107	-0.041	-1.985	0.031	-0.008	5.798	4.400	90.000	-0.629
1.849	7.7	1.916	-0.089	1.049	-1.104	-0.037	-1.842	-0.005	-0.008	5.776	4.400	90.000	-0.615
2.099	7.837	1.961	-0.056	0.984	-1.103	-0.038	-1.962	-0.006	-0.008	5.755	4.400	90.000	-0.602
2.382	7.874	1.949	-0.041	0.958	-1.088	-0.040	-2.202	0.129	-0.010	5.734	4.400	90.000	-0.606
2.705	8.08	2.011	-0.016	0.981	-1.093	-0.054	-2.548	0.063	-0.010	5.713	4.400	90.000	-0.627
3.070	8.202	1.979	-0.001	0.990	-1.107	-0.051	-2.425	0.040	-0.010	5.692	4.400	90.000	-0.654
3.486	8.204	1.923	0.006	0.969	-1.109	-0.044	-2.271	0.066	-0.011	5.671	4.400	90.000	-0.621
3.957	8.164	1.898	0.018	0.916	-1.101	-0.044	2.020	-0.049	-0.010	5.650	4.400	90.000	-0.565
4.492	8.189	1.970	0.059	0.876	-1.105	-0.049	1.787	-0.180	-0.010	5.629	4.400	90.000	-0.460
5.099	8.111	1.971	0.084	0.827	-1.111	-0.045	1.757	-0.186	-0.010	5.609	4.400	90.000	-0.316
5.789	8.005	1.933	0.095	0.825	-1.118	-0.047	1.841	-0.208	-0.010	5.588	4.400	90.000	-0.168
6.572	7.844	1.881	0.113	0.756	-1.105	-0.039	1.875	-0.347	-0.010	5.568	4.400	90.000	-0.102
7.460	7.618	1.716	0.125	0.599	-1.101	-0.008	-1.839	-0.440	-0.010	5.547	4.400	90.000	-0.094
8.469	7.641	1.673	0.156	0.499	-1.128	0.011	-1.780	-0.659	-0.009	5.527	4.400	90.000	-0.084
9.614	7.664	1.595	0.180	0.470	-1.152	0.029	-1.960	-0.822	-0.008	5.506	4.400	90.000	-0.144
10.914	7.591	1.549	0.197	0.414	-1.147	0.039	-2.265	-1.078	-0.007	5.500	4.400	90.000	-0.134
12.390	7.567	1.565	0.208	0.463	-1.150	0.027	-2.191	-1.129	-0.007	5.500	4.400	90.000	-0.092
14.066	7.462	1.520	0.230	0.379	-1.159	0.041	-1.920	-1.257	-0.006	5.500	4.400	90.000	-0.068
15.968	7.465	1.509	0.240	0.385	-1.190	0.041	-1.874	-1.367	-0.005	5.500	4.400	90.000	-0.020
18.127	7.55	1.495	0.237	0.390	-1.233	0.041	-1.889	-1.520	-0.004	5.500	4.400	90.000	-0.014
	7.653	1.545	0.247	0.441	-1.274	0.037	-1.958	-1.605	-0.004	5.500	4.400	90.000	0.032

SURF$_{1D}$cor$_{GITp}$

freq	a_1	a_2	a_3	a_4	b_1	b_2	h	b_3	b_4	M_h	M_{ref}	R_{break}	c_1
0.520	5.495	2.047	-0.017	0.198	-1.172	0.102	2.713	-0.010	-0.005	5.993	4.400	90.000	-0.050
0.590	5.736	2.019	-0.041	0.274	-1.184	0.085	2.914	0.054	-0.006	5.971	4.400	90.000	-0.009
0.670	5.892	1.990	-0.059	0.335	-1.179	0.068	2.941	0.066	-0.006	5.949	4.400	90.000	0.009
0.761	5.907	1.920	-0.078	0.387	-1.152	0.064	-2.650	0.029	-0.006	5.927	4.400	90.000	0.010
0.864	5.972	1.828	-0.098	0.426	-1.136	0.061	-2.454	0.030	-0.006	5.906	4.400	90.000	-0.008
0.981	6.234	1.845	-0.103	0.487	-1.133	0.043	-2.342	-0.061	-0.006	5.884	4.400	90.000	-0.032
1.113	6.453	1.868	-0.101	0.648	-1.126	0.026	-2.245	-0.067	-0.006	5.863	4.400	90.000	-0.046
1.264	6.582	1.812	-0.113	0.766	-1.120	0.016	-2.129	-0.065	-0.007	5.841	4.400	90.000	-0.060
1.435	6.83	1.858	-0.113	0.947	-1.119	-0.011	-2.049	0.035	-0.007	5.820	4.400	90.000	-0.067
1.628	7	1.902	-0.099	1.048	-1.117	-0.027	-2.000	0.046	-0.008	5.798	4.400	90.000	-0.042
1.849	7.113	1.910	-0.079	1.039	-1.116	-0.029	-1.860	0.000	-0.008	5.776	4.400	90.000	-0.037
2.099	7.267	1.963	-0.047	1.024	-1.117	-0.034	-1.983	0.013	-0.009	5.755	4.400	90.000	-0.028
2.382	7.305	1.949	-0.031	0.998	-1.104	-0.035	-2.203	0.164	-0.010	5.734	4.400	90.000	-0.036
2.705	7.51	2.011	-0.004	1.013	-1.110	-0.047	-2.564	0.084	-0.010	5.713	4.400	90.000	-0.045
3.070	7.651	1.992	0.014	1.019	-1.125	-0.046	-2.535	0.030	-0.010	5.692	4.400	90.000	-0.066
3.486	7.685	1.935	0.021	1.013	-1.127	-0.039	-2.398	0.058	-0.011	5.671	4.400	90.000	-0.065
3.957	7.743	1.934	0.039	0.970	-1.120	-0.042	2.140	-0.053	-0.010	5.650	4.400	90.000	-0.081
4.492	7.859	1.990	0.081	0.914	-1.125	-0.044	1.859	-0.175	-0.010	5.629	4.400	90.000	-0.086
5.099	7.93	1.997	0.106	0.877	-1.131	-0.043	1.828	-0.216	-0.010	5.609	4.400	90.000	-0.074
5.789	7.905	1.933	0.114	0.844	-1.135	-0.039	1.897	-0.245	-0.010	5.588	4.400	90.000	-0.034
6.572	7.782	1.834	0.128	0.758	-1.121	-0.023	1.915	-0.361	-0.010	5.568	4.400	90.000	-0.033
7.460	7.566	1.633	0.134	0.573	-1.114	0.013	-1.826	-0.468	-0.010	5.547	4.400	90.000	-0.027
8.469	7.654	1.612	0.168	0.478	-1.145	0.028	-1.783	-0.697	-0.009	5.527	4.400	90.000	-0.006
9.614	7.634	1.501	0.194	0.420	-1.170	0.055	-1.950	-0.800	-0.008	5.506	4.400	90.000	-0.027
10.914	7.576	1.438	0.213	0.358	-1.164	0.068	-2.214	-1.074	-0.007	5.500	4.400	90.000	-0.008
12.390	7.562	1.420	0.223	0.372	-1.167	0.060	2.129	-1.156	-0.007	5.500	4.400	90.000	0.040
14.066	7.448	1.330	0.247	0.240	-1.165	0.083	1.809	-1.273	-0.007	5.500	4.400	90.000	0.034
15.968	7.53	1.299	0.262	0.222	-1.204	0.085	1.817	-1.496	-0.005	5.500	4.400	90.000	0.057
18.127	7.713	1.291	0.261	0.268	-1.264	0.082	1.989	-1.692	-0.004	5.500	4.400	90.000	0.087
	7.759	1.352	0.271	0.290	-1.309	0.074	2.128	-1.797	-0.004	5.500	4.400	90.000	0.302

$SURF_{1D}cor_{GITn}$

freq	a_1	a_2	a_3	a_4	b_1	b_2	h	b_3	b_4	M_h	M_{ref}	R_{break}	c_1
0.520	5.719	2.159	-0.024	0.391	-1.179	0.058	2.835	-0.019	-0.006	5.993	4.400	90.000	0.058
0.590	5.995	2.128	-0.047	0.441	-1.190	0.044	3.062	0.026	-0.006	5.971	4.400	90.000	0.062
0.670	6.187	2.092	-0.069	0.502	-1.182	0.032	3.123	0.045	-0.006	5.949	4.400	90.000	0.061
0.761	6.207	1.998	-0.095	0.540	-1.148	0.034	-2.873	0.037	-0.006	5.927	4.400	90.000	0.046
0.864	6.269	1.898	-0.115	0.560	-1.133	0.034	-2.663	0.046	-0.006	5.906	4.400	90.000	0.028
0.981	6.474	1.909	-0.113	0.592	-1.131	0.022	-2.524	-0.065	-0.006	5.884	4.400	90.000	0.025
1.113	6.625	1.914	-0.109	0.718	-1.125	0.010	-2.352	-0.065	-0.006	5.863	4.400	90.000	0.030
1.264	6.711	1.837	-0.120	0.804	-1.115	0.006	-2.174	-0.068	-0.006	5.841	4.400	90.000	0.022
1.435	6.958	1.883	-0.118	0.980	-1.115	-0.020	-2.082	0.035	-0.007	5.820	4.400	90.000	0.010
1.628	7.123	1.928	-0.100	1.076	-1.113	-0.034	-2.008	0.051	-0.008	5.798	4.400	90.000	0.026
1.849	7.247	1.931	-0.079	1.062	-1.113	-0.035	-1.871	0.005	-0.008	5.776	4.400	90.000	0.017
2.099	7.411	1.982	-0.046	1.035	-1.113	-0.039	-1.991	0.015	-0.009	5.755	4.400	90.000	0.009
2.382	7.459	1.959	-0.030	1.000	-1.100	-0.037	-2.210	0.163	-0.010	5.734	4.400	90.000	-0.008
2.705	7.679	2.018	-0.003	1.009	-1.104	-0.049	-2.563	0.086	-0.010	5.713	4.400	90.000	-0.017
3.070	7.828	1.996	0.013	1.022	-1.118	-0.048	-2.523	0.037	-0.010	5.692	4.400	90.000	-0.038
3.486	7.881	1.950	0.022	1.018	-1.119	-0.042	-2.393	0.067	-0.011	5.671	4.400	90.000	-0.039
3.957	7.951	1.958	0.043	0.975	-1.115	-0.046	-2.158	-0.039	-0.010	5.650	4.400	90.000	-0.061
4.492	8.106	2.025	0.083	0.939	-1.117	-0.052	1.882	-0.154	-0.010	5.629	4.400	90.000	-0.084
5.099	8.194	2.031	0.108	0.903	-1.121	-0.050	1.826	-0.189	-0.010	5.609	4.400	90.000	-0.083
5.789	8.173	1.964	0.114	0.879	-1.123	-0.046	1.884	-0.214	-0.010	5.588	4.400	90.000	-0.049
6.572	8.072	1.883	0.128	0.808	-1.112	-0.034	1.923	-0.317	-0.010	5.568	4.400	90.000	-0.064
7.460	7.822	1.693	0.135	0.626	-1.106	0.002	-1.842	-0.421	-0.010	5.547	4.400	90.000	-0.064
8.469	7.849	1.676	0.169	0.530	-1.133	0.017	-1.777	-0.635	-0.009	5.527	4.400	90.000	-0.043
9.614	7.782	1.580	0.193	0.495	-1.163	0.040	-1.918	-0.712	-0.009	5.506	4.400	90.000	-0.062
10.914	7.662	1.526	0.213	0.428	-1.165	0.053	-2.133	-0.971	-0.007	5.500	4.400	90.000	-0.045
12.390	7.599	1.528	0.226	0.444	-1.173	0.045	-2.038	-1.052	-0.007	5.500	4.400	90.000	-0.001
14.066	7.472	1.448	0.250	0.321	-1.179	0.065	-1.794	-1.178	-0.006	5.500	4.400	90.000	0.020
15.968	7.492	1.405	0.267	0.307	-1.211	0.071	-1.793	-1.383	-0.005	5.500	4.400	90.000	0.097
18.127	7.541	1.349	0.253	0.309	-1.251	0.070	1.867	-1.648	-0.004	5.500	4.400	90.000	0.333
	7.663	1.422	0.273	0.307	-1.304	0.064	2.043	-1.772	-0.004	5.500	4.400	90.000	0.533

SURF$_{ALL}$

freq	a_1	a_2	a_3	a_4	b_1	b_2	h	b_3	b_4	M_h	M_{ref}	R_{break}	c_1
0.520	6.061	2.349	0.086	0.743	-1.256	0.021	4.171	-0.110	-0.006	5.993	4.400	90.000	-0.650
0.590	6.257	2.311	0.061	0.791	-1.256	0.010	4.200	-0.011	-0.006	5.971	4.400	90.000	-0.669
0.670	6.419	2.272	0.036	0.804	-1.248	0.000	4.179	0.073	-0.007	5.949	4.400	90.000	-0.683
0.761	6.550	2.220	0.013	0.874	-1.235	-0.004	4.130	0.167	-0.008	5.927	4.400	90.000	-0.683
0.864	6.655	2.130	-0.013	0.890	-1.229	-0.005	4.161	0.230	-0.008	5.906	4.400	90.000	-0.697
0.981	6.837	2.094	-0.034	0.952	-1.225	-0.017	4.163	0.234	-0.009	5.884	4.400	90.000	-0.723
1.113	6.975	2.054	-0.047	1.021	-1.219	-0.025	4.056	0.236	-0.009	5.863	4.400	90.000	-0.759
1.264	7.082	1.968	-0.066	1.065	-1.221	-0.026	4.070	0.298	-0.009	5.841	4.400	90.000	-0.799
1.435	7.153	1.898	-0.075	1.100	-1.216	-0.025	4.013	0.321	-0.010	5.820	4.400	90.000	-0.850
1.628	7.244	1.852	-0.076	1.111	-1.210	-0.025	3.936	0.327	-0.010	5.798	4.400	90.000	-0.883
1.849	7.423	1.866	-0.066	1.130	-1.210	-0.031	3.785	0.251	-0.010	5.776	4.400	90.000	-0.932
2.099	7.543	1.841	-0.055	1.099	-1.211	-0.028	3.837	0.244	-0.011	5.755	4.400	90.000	-0.967
2.382	7.661	1.809	-0.044	1.084	-1.215	-0.025	4.000	0.269	-0.011	5.734	4.400	90.000	-0.997
2.705	7.849	1.801	-0.032	1.071	-1.226	-0.029	4.284	0.245	-0.011	5.713	4.400	90.000	-1.011
3.070	7.913	1.748	-0.020	1.036	-1.222	-0.020	4.071	0.203	-0.012	5.692	4.400	90.000	-1.008
3.486	8.069	1.720	-0.009	1.061	-1.231	-0.018	4.023	0.121	-0.011	5.671	4.400	90.000	-0.983
3.957	8.229	1.743	0.018	0.990	-1.235	-0.019	3.910	-0.027	-0.011	5.650	4.400	90.000	-0.971
4.492	8.355	1.752	0.047	0.929	-1.235	-0.015	3.723	-0.180	-0.010	5.629	4.400	90.000	-0.927
5.099	8.471	1.753	0.068	0.925	-1.232	-0.017	3.765	-0.302	-0.010	5.609	4.400	90.000	-0.874
5.789	8.681	1.772	0.086	0.963	-1.245	-0.025	3.788	-0.326	-0.010	5.588	4.400	90.000	-0.796
6.572	8.759	1.725	0.095	0.954	-1.245	-0.022	3.770	-0.497	-0.010	5.568	4.400	90.000	-0.703
7.460	8.832	1.624	0.096	0.910	-1.262	-0.010	3.750	-0.641	-0.009	5.547	4.400	90.000	-0.626
8.469	8.955	1.641	0.129	0.855	-1.276	-0.004	3.678	-0.822	-0.009	5.527	4.400	90.000	-0.549
9.614	9.054	1.613	0.143	0.874	-1.295	-0.001	3.863	-0.969	-0.008	5.506	4.400	90.000	-0.451
10.914	9.070	1.570	0.158	0.801	-1.305	0.010	4.177	-1.174	-0.007	5.500	4.400	90.000	-0.378
12.390	9.097	1.538	0.166	0.749	-1.321	0.014	4.271	-1.237	-0.007	5.500	4.400	90.000	-0.294
14.066	9.039	1.521	0.182	0.692	-1.327	0.021	4.313	-1.330	-0.007	5.500	4.400	90.000	-0.234
15.968	8.976	1.487	0.189	0.618	-1.350	0.029	4.308	-1.337	-0.007	5.500	4.400	90.000	-0.222
18.127	8.900	1.475	0.190	0.622	-1.375	0.033	4.251	-1.317	-0.007	5.500	4.400	90.000	-0.255
	8.870	1.496	0.193	0.639	-1.399	0.033	4.332	-1.326	-0.006	5.500	4.400	90.000	-0.296

$SURF_{ALL}cor_{GITp}$

freq	a_1	a_2	a_3	a_4	b_1	b_2	h	b_3	b_4	M_h	M_{ref}	R_{break}	c_1
0.520	6.03	2.151	0.009	0.410	-1.270	0.091	-5.054	0.132	-0.006	5.993	4.400	90.000	-0.082
0.590	6.126	2.060	-0.021	0.508	-1.273	0.085	-5.047	0.208	-0.007	5.971	4.400	90.000	-0.038
0.670	6.261	2.026	-0.042	0.576	-1.267	0.067	-4.959	0.275	-0.007	5.949	4.400	90.000	-0.017
0.761	6.359	1.981	-0.059	0.670	-1.251	0.059	4.768	0.339	-0.008	5.927	4.400	90.000	-0.017
0.864	6.476	1.908	-0.079	0.706	-1.243	0.052	4.685	0.405	-0.009	5.906	4.400	90.000	-0.028
0.981	6.72	1.891	-0.094	0.804	-1.244	0.035	4.656	0.377	-0.009	5.884	4.400	90.000	-0.048
1.113	6.95	1.896	-0.104	0.950	-1.236	0.014	4.539	0.360	-0.009	5.863	4.400	90.000	-0.063
1.264	7.095	1.841	-0.116	1.026	-1.233	0.005	4.521	0.429	-0.010	5.841	4.400	90.000	-0.082
1.435	7.169	1.794	-0.119	1.100	-1.224	-0.001	4.300	0.461	-0.010	5.820	4.400	90.000	-0.079
1.628	7.241	1.769	-0.110	1.111	-1.219	-0.004	4.161	0.451	-0.011	5.798	4.400	90.000	-0.065
1.849	7.412	1.794	-0.094	1.173	-1.220	-0.014	3.916	0.345	-0.011	5.776	4.400	90.000	-0.059
2.099	7.519	1.790	-0.073	1.154	-1.220	-0.013	3.882	0.330	-0.011	5.755	4.400	90.000	-0.054
2.382	7.61	1.766	-0.055	1.141	-1.223	-0.011	3.973	0.341	-0.011	5.734	4.400	90.000	-0.055
2.705	7.789	1.776	-0.037	1.146	-1.237	-0.018	4.203	0.312	-0.012	5.713	4.400	90.000	-0.058
3.070	7.86	1.754	-0.016	1.115	-1.240	-0.014	4.080	0.263	-0.012	5.692	4.400	90.000	-0.057
3.486	7.962	1.727	-0.001	1.131	-1.249	-0.012	4.066	0.191	-0.012	5.671	4.400	90.000	-0.056
3.957	8.104	1.749	0.024	1.067	-1.255	-0.015	4.002	0.024	-0.011	5.650	4.400	90.000	-0.083
4.492	8.189	1.760	0.053	1.015	-1.258	-0.012	3.866	-0.139	-0.010	5.629	4.400	90.000	-0.091
5.099	8.237	1.768	0.078	1.010	-1.256	-0.014	3.941	-0.259	-0.010	5.609	4.400	90.000	-0.082
5.789	8.328	1.775	0.098	1.038	-1.267	-0.020	3.916	-0.342	-0.010	5.588	4.400	90.000	-0.062
6.572	8.335	1.721	0.109	1.021	-1.268	-0.017	3.921	-0.530	-0.009	5.568	4.400	90.000	-0.061
7.460	8.361	1.627	0.115	0.987	-1.288	-0.005	3.863	-0.705	-0.009	5.547	4.400	90.000	-0.058
8.469	8.45	1.635	0.149	0.914	-1.308	0.002	3.781	-0.934	-0.008	5.527	4.400	90.000	-0.051
9.614	8.542	1.613	0.172	0.897	-1.334	0.007	4.012	-1.129	-0.007	5.506	4.400	90.000	-0.061
10.914	8.576	1.557	0.184	0.834	-1.347	0.017	4.311	-1.398	-0.006	5.500	4.400	90.000	-0.056
12.390	8.628	1.512	0.193	0.782	-1.368	0.021	4.461	-1.530	-0.006	5.500	4.400	90.000	-0.037
14.066	8.594	1.475	0.215	0.735	-1.371	0.030	-4.415	-1.709	-0.005	5.500	4.400	90.000	-0.029
15.968	8.685	1.437	0.227	0.656	-1.414	0.037	-4.455	-1.876	-0.004	5.500	4.400	90.000	-0.004
18.127	8.837	1.458	0.236	0.691	-1.468	0.033	-4.566	-2.035	-0.003	5.500	4.400	90.000	0.049
	8.829	1.497	0.251	0.699	-1.509	0.031	4.884	-2.108	-0.003	5.500	4.400	90.000	0.218

SURF$_{ALL}$cor$_{GITn}$

freq	a_1	a_2	a_3	a_4	b_1	b_2	h	b_3	b_4	M_h	M_{ref}	R_{break}	c_1
0.520	6.145	2.264	0.014	0.544	-1.270	0.054	-5.032	0.162	-0.007	5.993	4.400	90.000	0.034
0.590	6.289	2.176	-0.018	0.609	-1.271	0.048	-5.069	0.225	-0.007	5.971	4.400	90.000	0.055
0.670	6.47	2.144	-0.043	0.676	-1.260	0.032	-5.037	0.292	-0.007	5.949	4.400	90.000	0.059
0.761	6.604	2.094	-0.065	0.772	-1.238	0.027	4.910	0.368	-0.008	5.927	4.400	90.000	0.041
0.864	6.716	2.014	-0.085	0.794	-1.230	0.024	4.861	0.431	-0.009	5.906	4.400	90.000	0.028
0.981	6.903	1.992	-0.093	0.868	-1.232	0.011	4.788	0.389	-0.009	5.884	4.400	90.000	0.026
1.113	7.028	1.963	-0.100	0.969	-1.222	-0.001	4.573	0.370	-0.009	5.863	4.400	90.000	0.032
1.264	7.132	1.891	-0.111	1.015	-1.216	-0.004	4.512	0.433	-0.010	5.841	4.400	90.000	0.021
1.435	7.193	1.838	-0.111	1.104	-1.206	-0.007	4.275	0.469	-0.010	5.820	4.400	90.000	0.020
1.628	7.255	1.811	-0.102	1.113	-1.200	-0.011	4.110	0.460	-0.011	5.798	4.400	90.000	0.030
1.849	7.437	1.833	-0.084	1.135	-1.201	-0.018	3.875	0.346	-0.010	5.776	4.400	90.000	0.024
2.099	7.554	1.831	-0.062	1.119	-1.200	-0.017	3.838	0.337	-0.011	5.755	4.400	90.000	0.018
2.382	7.654	1.808	-0.042	1.114	-1.201	-0.015	3.939	0.344	-0.011	5.734	4.400	90.000	0.010
2.705	7.848	1.815	-0.025	1.107	-1.211	-0.021	4.154	0.319	-0.012	5.713	4.400	90.000	0.005
3.070	7.919	1.791	-0.005	1.082	-1.212	-0.017	4.028	0.275	-0.012	5.692	4.400	90.000	0.006
3.486	8.024	1.761	0.009	1.107	-1.219	-0.015	4.000	0.198	-0.012	5.671	4.400	90.000	0.003
3.957	8.174	1.785	0.034	1.048	-1.228	-0.018	3.938	0.035	-0.011	5.650	4.400	90.000	-0.020
4.492	8.288	1.803	0.064	0.989	-1.231	-0.016	3.831	-0.121	-0.011	5.629	4.400	90.000	-0.036
5.099	8.348	1.803	0.089	0.991	-1.226	-0.017	3.882	-0.233	-0.010	5.609	4.400	90.000	-0.034
5.789	8.457	1.820	0.110	1.024	-1.234	-0.024	3.840	-0.300	-0.010	5.588	4.400	90.000	-0.027
6.572	8.458	1.772	0.123	1.015	-1.236	-0.020	3.831	-0.468	-0.010	5.568	4.400	90.000	-0.041
7.460	8.438	1.691	0.133	0.982	-1.256	-0.009	3.779	-0.622	-0.009	5.547	4.400	90.000	-0.038
8.469	8.47	1.696	0.166	0.912	-1.274	-0.002	3.683	-0.825	-0.008	5.527	4.400	90.000	-0.037
9.614	8.489	1.677	0.188	0.924	-1.303	0.002	3.889	-1.002	-0.008	5.506	4.400	90.000	-0.043
10.914	8.449	1.634	0.205	0.865	-1.320	0.012	4.146	-1.251	-0.006	5.500	4.400	90.000	-0.040
12.390	8.461	1.619	0.220	0.832	-1.347	0.013	4.304	-1.363	-0.006	5.500	4.400	90.000	-0.022
14.066	8.407	1.589	0.245	0.770	-1.358	0.023	4.277	-1.525	-0.005	5.500	4.400	90.000	0.003
15.968	8.435	1.542	0.254	0.697	-1.394	0.030	-4.315	-1.688	-0.004	5.500	4.400	90.000	0.079
18.127	8.399	1.514	0.242	0.710	-1.426	0.028	-4.291	-1.880	-0.003	5.500	4.400	90.000	0.341
	8.387	1.515	0.254	0.685	-1.473	0.035	-4.695	-1.942	-0.003	5.500	4.400	90.000	0.538