This research is carried out in SINAPS@ project in collaboration with CASHIMA and SIGMA projects.

Référence: SINAPS@-2015-V1-A1-T3-1
Version : 01

Correction of surface records of their site effect before developing GMPE: an alternative approach to get reference incident ground motion (application to KiK-net data)

Fabrice HOLLENDER, Laetitia FOUNDOTOS, Aurore LAURENDEAU, Pierre-Yves BARD

With the collaboration of: Bruno HERNANDEZ, Stéphane DROUET, Olga KTENIDOU, Vincent PERRON, Jean LETORT

AUTHORS		REVIEWS		APPROVAL	
NOM	DATE & VISA	NOM	DATE & VISA	NOM	DATE & VISA
F. Hollender CEA/DEN/CAD/DPIE		M. Nicolas Leader SINAPS@ WP1		C. Berge-Thierry Coordinator SINAPS@	
L. Foundotos CEA/DEN/CAD/DPIE		F. Hollender co-advisor			
A. Laurendeau CEA/DAM/DASE/SLDG					

Submittted to the external review of scientific committee SIGMA

REVIEW		APPROVAL	
NOM	DATE &	NOM	DATE & VISA
Faccioli E		F. Hollender Leader WP3-SIGMA	
Scherbaum F.		G. Senfaute Coordinator SIGMA	

Executive Summary

The initial motivation of our work results from the observation that, within a site-specific seismic hazard study, the methods used to evaluate the site effect on the "host side" within the host-to-target adjustments were not the same than the ones used to evaluate the site effect on the "target side". We thus wanted to test an alternative concept in which we attempt to withdraw the site effect for each station of an accelerometric network from the data before deriving a GMPE, using methods as close as possible than the ones used on "target" sites. The KiK-net network (with couple of sensors –downhole and surface– for each sites + velocity profiles provided up to at least 100 m depth) offers the opportunity to test this approach with the possibility to check the correction process. Nevertheless, this work would be useless if we could not transpose it to other databases (with no downhole sensors or high quality Vs profile). So, in parallel, we wanted to develop a methodology that could be transposed to other databases.

In this work, we tested different approaches to evaluate the "transfer function" describing the site effect: 1D simulation, Generalized Inversion Technique (GIT), Standard Spectral Ratio (SSR) on records between surface and downhole sensors. These parameters will be first evaluated in terms of Fourier Spectra.

We showed that, in terms of transfer functions (expressed in Fourier Spectra) between the downhole and the surface sensors, the three methods are (statistically and in mean) in good agreement. However, the 1D simulation leads to a slight underestimation of site effect between 6 and 15 Hz and to a strong overestimation above 15 Hz. This is due to a "scaling" issue of the Q_S parameter. Indeed, we never get a Q_S profile and we have to deduce it from other parameters (V_S for example). In our work, we used the "standard" $Q_S = V_S/10$ scaling which seems no more relevant at high frequency.

We derived two GMPEs. The first one was based on surface data, corrected for site effect using 1D simulation. For this first GMPE, we needed as a first step to convert the 1D simulated transfer function expressed in Fourier spectra toward amplification factor, expressed in Response Spectra. The second one was based on downhole data "transposed" to surface using the Cadet et al. (2012) method to take into account free-surface phenomenon. The correction methodology proposed by Cadet et al. (2012) is already expressed in Response Spectra domain. This last GMPE was developed in order to check the first one. Globally, these two GMPEs are in good agreement, even at high frequency, that may suggest that the correction method, based on 1D simulation site effect estimation, is robust.

These two GMPEs are also compared to the one, derived from the surface data set, without station-by-station site effect correction, but with a standard host-to-target correction. Around 10 Hz, the difference between our new GMPEs and standard one is high, up to a 3 or 4 factor. This observation is essential and may have a high implication on seismic hazard assessment, if verified. We explain this high difference by the fact that we explicitly attempt to remove site effect from each station. As we demonstrated, the site effect is high at high frequency (for each class of stations) and this is fact is confirmed by 3 different methods (when analyzed in Fourier domain considering surface to downhole ratios). The host-to-target classical approaches probably do not fully take into account this feature.

We also discussed about the use of GIT in order to estimate site effect (within the perspective of applying our approach on database where we do not have down-hole sensors and where the information quality about velocity profiles is lower). The main difficulty of GIT is the fact that "site terms" are expressed in a relative form, we have thus to "post-process" them to provide and absolute site effect evaluation. Nevertheless, as far as we have some additional information in order to better parametrize the inversion, we think that GIT may provide robust results that may use for site effect remove from data before GMPE derivation.

FOREWORD

This work results from a collaboration between the SINAPS@ project and the Cashima* project. It is the continuation of the previous works initiated within the Cashima project (Laurendeau 2013, Sigma deliverable D3-36).

The database used in the present work (presented in chapter 3) is the same than the one use for these previous works.

For the present report, the work consisting in computing transfer function with 1 D simulation and SSR and the development of GMPE was performed by Aurore Laurendeau, funded by the SINAPS@ project. This results in the redaction of parts 4.1, 4.2, 4.3 and 5.

The work on Generalized Inversion Techniques, the comparison and synthesis between different methods, the “kappa” discussion was performed by Laetitia Foundotos, funded by the Cashima project, and Fabrice Hollender. This results in the redaction of parts 1, 4.4, 4.5, 6, and 7.

* According to the memorandum of understanding linking EdF, CEA, Areva and Enel for the development of the SIGMA program, the work performed by the Cashima program is the contribution of the CEA to the workpackage 3 (site effect) of the SIGMA program.

TABLE OF CONTENTS

1 INTRODUCTION

1.1 CONTEXT: THE "HOST TO TARGET ADJUSTMENT"

Advanced site-specific seismic hazard studies involve the determination of the amplification due to local geological features of the site of interest (called "target site"). This amplification characterization could be performed using numerical simulation (based on 1D, 2D or 3D physical properties models) and/or by empirical approached (e.g. "standard spectrum ratio" based on records of earthquakes at both "site" and "reference" locations). This amplification is a function of frequency, but could also be a function of the incidence and back-azimuth angles of considered seismic scenario, and also a function of the incidence ground motion level (in case of non-linearity consideration).

This amplification is usually expressed with respect to the bedrock underlying the basin (or more generally the soil layers) of a surface reference station where the same bedrock outcrops. The depth of this bedrock depends on the site (in the case of sites considered within the Sigma project, this depths goes from few tens of meters to approximately 1000 m for the Grenoble test site, which is a very deep site). At those depths, the shear wave velocity (V_S) is usually quite high (more than 1500 m/s, even if one can find of course some exception as the bedrock underlying the Pô plain sites investigated by Sigma Italian partners).

Thus, in order to derive site amplification to real seismic hazard estimation, we need to have "input ground motions" which are representative to these high V_S values. The available GMPEs are often not representative of such V_S values due to the lack of accelerometric stations located on very hard rock.

In order to overcome to this issue, the state-of-the-art practice is to apply the "host-to-target adjustments" (H2T). These corrections aim to take into account the site conditions characterizing in mean the accelerometric station sites used for the GMPEs derivation (called "host" side) in order to correct them and then in order to make them applicable for the "target" site.

Today, most of the "host to target" approaches use the couple of parameter "V_{S30}, κ_0", using a concept of correction proposed by Boore and Joyner (1997), then used in different forms, optimized and associated to different kinds of data processing by many others (Campbell 2003, Scherbaum et al. 2004, Cotton et al. 2006, Al Atik et al. 2010, etc.).

Basically:

- the V_{S30} parameter is used to perform an impedance correction, often with the quarter wavelength approach, using smooth generic V_S profiles down to 8 km depth (the real V_S profile of a given site is not taken into account). This correction does not take into account attenuation and cannot reproduce sharp contrast effect.
- The κ_0 parameter is used to take into account attenuation.

The V_{S30} correction, when used to go from "standard rock" conditions (typically V_{S30} = 800 m/s) toward a "hard rock" conditions (typically V_{S30} = 2000 m/s) leads to decrease ground motion levels on the whole frequency band, this decrease being larger at high frequency. On the contrary, the κ_0 correction, still going from "standard rock" conditions (typically κ_0 = 0.04 s) toward a "hard rock" conditions (typically κ_0 = 0.01 s) induces a drastic increase of the seismic level at high frequency. The combination of both correction leads to a quite moderate decrease of seismic level at low and intermediate frequencies and can leads to rather high increase at high frequency. Figure 1 and Figure 2 give some illustrations of such V_{S30}, κ_0 corrections.

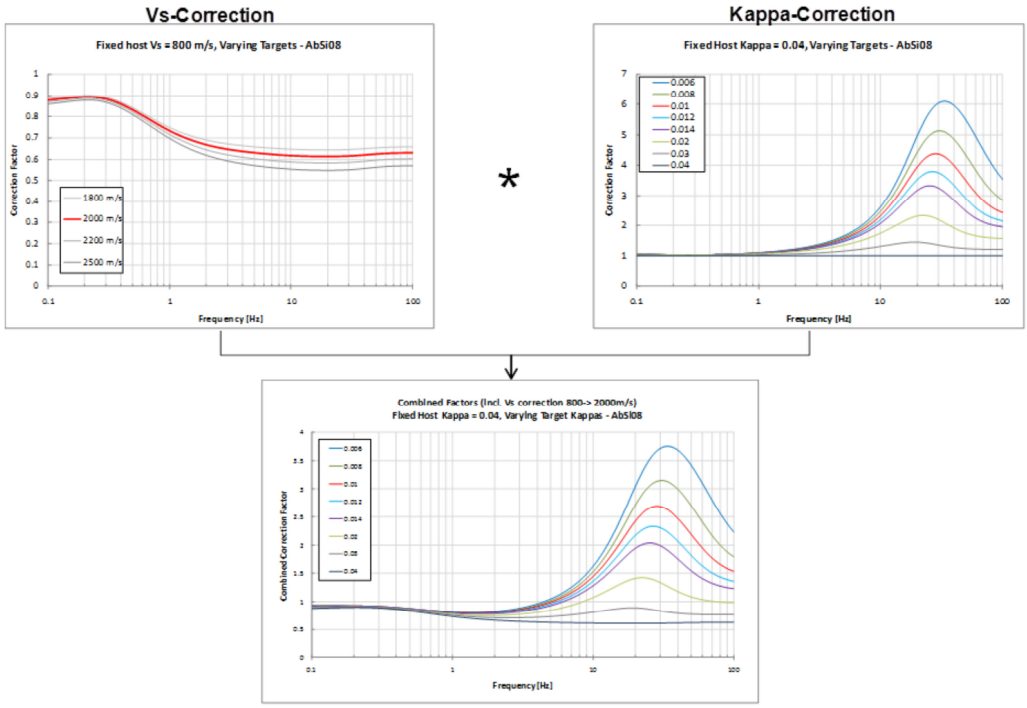

Figure 1: Hard rock to standard rock ratio: Top left: V_{S30} correction; top right: κ_0 correction; bottom: combination of V_{S30}, κ_0 corrections (from Biro and Renault 2012).

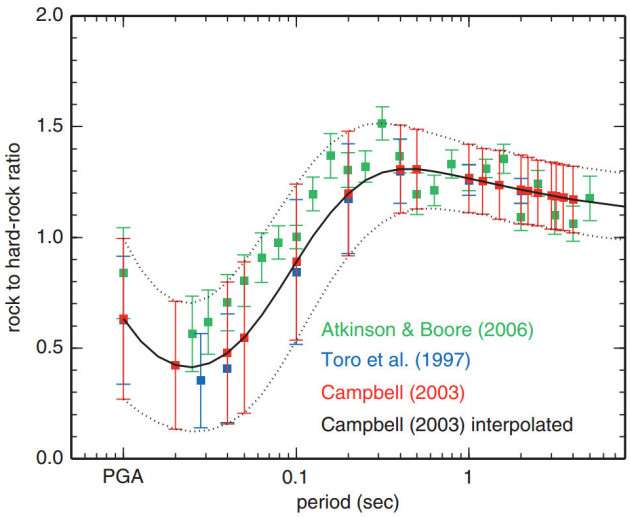

Figure 2: Standard rock to hard rock ratio (from Van Houtte et al. 2011).

This calls for several comments:

1. The methods used to compute site effect on the “target side” are different to the ones, more global, used to estimate site effects on the “host side”. If one or the other method class introduce a bias in the site effect evaluation, this could introduce a risk of “site effect double counting” or on the contrary, a site effect underestimation (this issue being also a matter of frequency range).
2. The investigated scale is not the same. On the “target side”, the site effect is evaluated between the bedrock and the surface, that is to say at a typically pluri-hectometer scale. Let’s call this site effect the “local site effect (LSE)”. On the contrary, on the “host side”, the “V_{S30}, κ_0” corrections aim to take into account (at least “theoretically”) features down to several kilometers (do not be confused by the

"30" number in the V_{S30} parameter name: here, V_{S30} value is used to select a given generic profile down to several kilometers). Since this correction is linked to a accelerometric station and since it aims to take into account all features that are not included in the source term nor in the regional attenuation term, let's call this site effect the "station site effect (SSE)" (this will be especially useful when we will refer to Generalized Inversion Technique results). Using the "local site effect" and "station site effect" concept, one can also introduce the "residual site effect (RSE)" as being the ratio between the SSE and the LSE for a given site (see Figure 3).

3. The κ_0 correction implicates correction even in the intermediate frequency range (below 10 Hz). For example, on Figure 1, one can see that the correction is already significant at 5 Hz, whereas the "predilection" frequency range of κ_0 measurement on signal is often above 10 Hz.

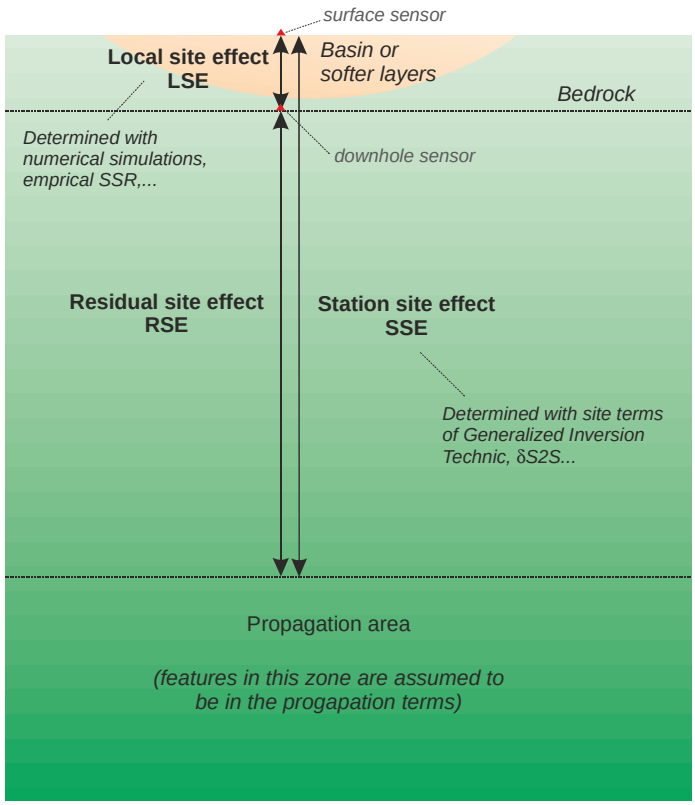

Figure 3: Notions of "station" site effect, "local" and "residual" site effects.

1.2 SPECIFIC ISSUES CONCERNING THE κ_0 DETERMINATION AND USE

The possible methods used to determine κ_0 are numerous and are summarized in Ktenidou et al. (2014). The "original" method was proposed by Anderson and Hough (1984). This is probably the method which is the most closely linked to the attenuation physical interpretation of κ: Within the Ktenidou et al., 2014, taxonomy, this method leads to the determination of κ_{0_AS}. However, from a practical point of view, this method is highly "time consuming" to implement on the whole station set of a given accelerometric network, especially if we wish to apply it taking care of the whole necessary precautions (for example: being sure that the frequency range where κ is measured is not affected by a site effect that could bias the slope to spectrum decay).

Thus, more "systematic" and easily applied approaches are often used. For example, the response spectra approach (Silva and Darragh, 1995) leads to determine κ_{0_RESP}. This method was used in

Laurendeau et al. 2013. Another one is based on the direct measurement on the site amplification transfer function, for example coming from Generalized Inversion Technique (Drouet et al., 2010) that leads to determine κ_{0_TF}.

In this kind of approaches, the determination of κ_0 could be bias by a local site effect as already demonstrated by Parolai and Bindi, 2004. In the case of κ_{0_TF}, we illustrated and example of such bias on a RAP Pyrenees station within a previous Sigma deliverable (Hollender et al. 2014) (see Figure 4).

Here, it is no more possible to assume that κ_0 is only linked to attenuation features. We prefer to consider it as proxy of high frequency content of signals associated to a given station. It is obvious that "hard rock stations" produce statistically more high frequency than "standard rock stations", but let's open the discussion considering that this feature could be produced by both:

- a "lack" of attenuation for harder sites with respect to softer site (original assumed phenomenology of κ_0),
- the fact that statistically, harder sites have higher frequency site effects than softer sites (due to thin weathered layers of few meters, even on most rock sites).

We often considered that "rock site" is synonymous of "site without site effect". Over the few last years, with an increasing effort on characterization of accelerometric network sites (e.g. Pileggi et al. 2011), we realized that this assumption is no more tenable, especially for high frequency concerns.

If our assumption is right, then this would have an impact on how to implement host-to-target adjustments (H2T). If the "lack-of-attenuation" feature of the κ_0 phenomenology has definitely to be considered for H2T, on the contrary, the "high-frequency-site-effect" feature has not to be "re-injected" on the target side since possible local site effect due to possible near surface softer layers is already taken into account in the local site effect evaluation of the target site.

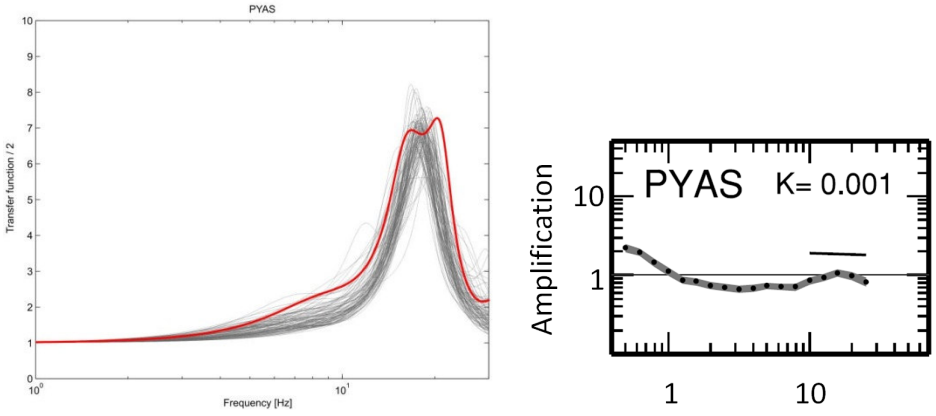

Figure 4: On the left: 1D transfer functions computed with "best estimated" Vs profile (red) and the one thousand profile sets deriving from the "acceptable misfit" approach (gray) for PYAS site. On the right: site transfer function for PYAS from generalized inversion of Drouet et al. (2010). Solid lines indicate the regression of the high frequency part of the transfer functions, which leads to the κ-values indicated on top of each frame. One can guess the high frequency site effect (bump between 10 and 20 Hz). This clearly affects the kappa estimation.

1.3 OBJECTIVES AND ORGANIZATION OF THE PRESENT WORK

The first objective of our work is to test an approach of GMPE adjustment that uses, on the "host side", site effect correction methods that are as close as possible as the ones usually used on the "target

side". Of course, it will not be possible to perform on each site of a given accelerometric network the same level of site effect estimation than the ones usually reached on target sites. So, we will use some simplifications and we will look for a "statistical validity" on the whole network instead of a "perfect correction" station by station.

This work will be performed on a subset of the KiK-net database, which has the advantages of 1/ providing V_S and V_P profiles, down to at least 100 m, based on in situ geotechnical measurements, and 2/ providing data from surface and downhole accelerometers that allows to better control site effect estimation from various methods.

This will allow us to evaluate site effect using 1/ a purely empirical method using Standard Spectral Ratio (SSR) between surface and downhole sensors, 2/ transfer function computed with 1D simulation using Thomson-Haskell method (Thomson,1950; Haskell, 1953).

The comparison between empirical SSR and computed transfer function will allow us to select station for which the 1D assumption is the more suitable.

Then, the basic idea will be to "substract" the local site effect (estimated with 1D simulation) from surface data before deriving the GMPE. The availability of downhole data will also allow us to apply the correction proposed by Cadet et al. (2012) in order to "convert it" to surface and then deriving another GMPE. The comparison of both GMPE will help us to comment the robustness of the obtained results.

Of course, the information provided by the KiK-net network (geotechnical profiles, downhole sensors…) are not available on all network and our work would be useless if it could not be apply to other database. So, our second objective is to develop a methodology that could also be applied other network, without downhole data and without such complete information set about velocity profiles. So, we here propose to apply the Generalized Inversion Technique (GIT) (Drouet et al. 2008, Drouet et al. 2010) on our data set (on both surface and downhole data). The ratio between surface and downhole site terms will also be compared to empirical SSR and ratio computed with 1D simulation in order to comment the robustness of results and to understand the difference between methods.

The site terms coming from GIT results should be representative of the "station site effect" (SSE: see above) whereas the transfer function computed with 1D simulations should be representative of the "local site effect" (LSE). This give the opportunity to evaluate the "residual site effect" (RSE) and then draw some comments and perspective on host to target adjustment approaches and discuss about the need of application of a "residual" correction to the proposed GMPE.

At a first glance, one can consider that our correction based on the LSE deconvolution is applied as a substitute of the V_{S30} correction and that the κ_0 correction will still have to be applied. This is not the case. Indeed, the attenuation is taken into account within the LSE estimation over the thickness of material between the downhole and the surface sensors. Moreover, as we suggested above, a part of κ_0 is also linked to high frequency site effect that we attempt to withdraw from data. So we implicitly perform at least a part of the κ_0 correction.

More generally, on the κ_0 issue, we will propose some figures (comparing κ_{0_TF} and κ_{0_resp} for example) and we generally will get poor correlation.

The overall task organization of the present work is illustrated on Figure 5.

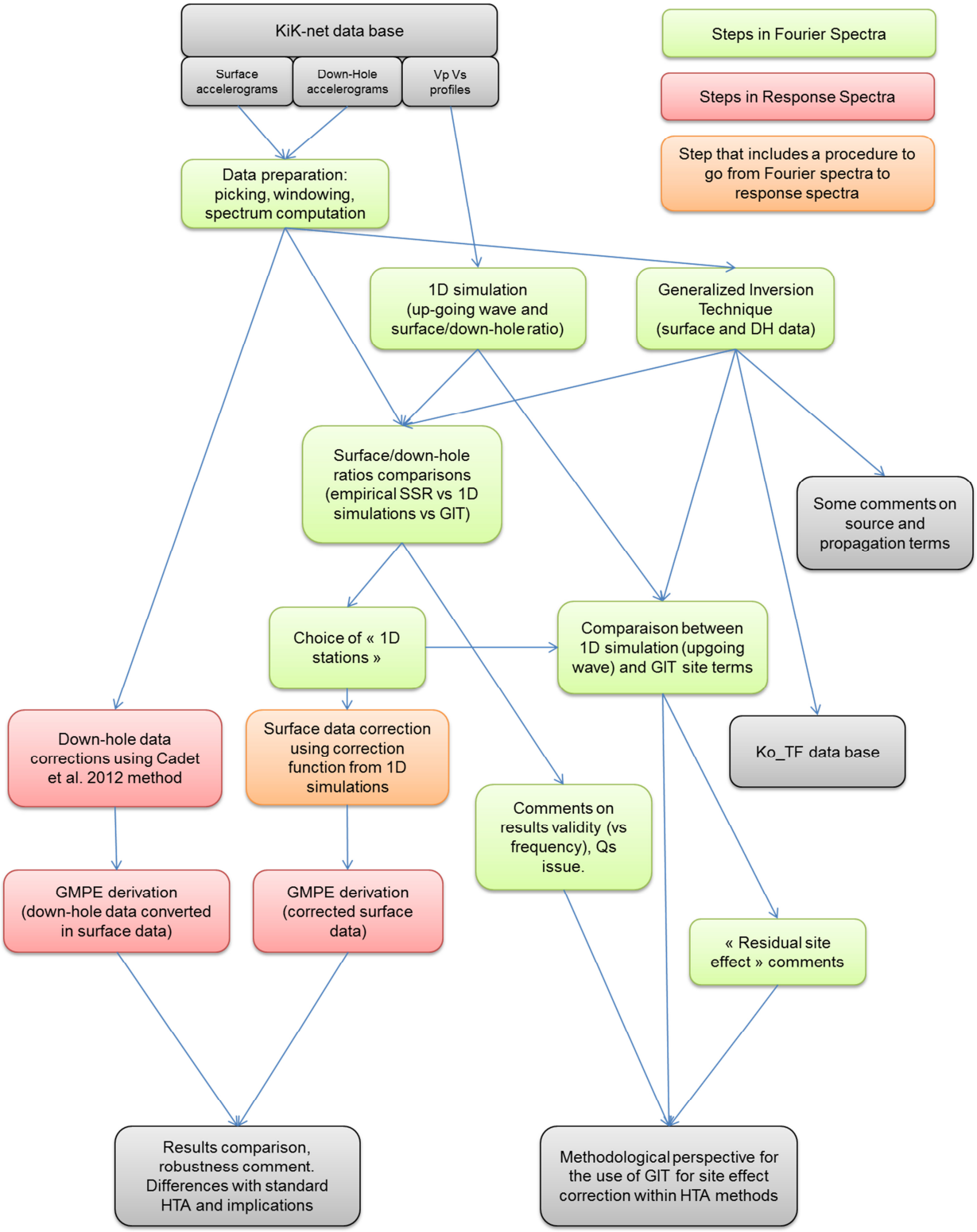

Figure 5: General task framework of the present work. Colors give the "spectral domain" in which the work is achieved.

2 NOTATIONS

Apeak: amplitude of the peak of the response spectra

BTF: the theoretical standard spectral ratio between TF and TF_dh

fdest: the frequency for which destructive interferences are maximal

fpeak: frequency of the peak of the response spectra

GIT: Generalized Inversion Technique

H2T: Host-to-target adjustment

HVSR: the empirical horizontal-to-vertical spectral ratios

LSE: local site effect

r: Pearson's correlation coefficient

RSE: residual site effect

SSE: station site effect

SSR: the empirical standard spectral ratio

std: standard deviation

TF: theoretical transfer function computed at the surface

TF_dh: theoretical transfer function computed at the sensor depth

3 THE KIK-NET DATASET

Japan is in an area of high seismicity, where a lot of quality digital data are recorded and made available to the scientific community. Indeed, after the destructive 1995 Kobe earthquake, Japanese scientists installed dense and uniform networks that cover the whole of Japan: the Hi-net (high-sensitivity), F-net (broadband), KiK-net and K-NET (strong-motion) networks [Okada et al., 2004]. The KiK-net network offers the advantage of combining pairs of sensors (one at the surface, and one installed at depth in a borehole). Each instrument is a three-component seismograph with a 24-bit analog-to-digital converter; the KiK-net network uses 200-Hz (until 27 January 2008) and 100-Hz (since 30 October 2007) sampling frequencies. The overall frequency response characteristics of the total system is flat, strictly speaking, from 0 to 15 Hz, after which the amplitude starts to decay. The response characteristics are approximately equal to those of a 3-pole Butterworth filter with a cut-off frequency of 30 Hz [Kinoshita, 1998; Fujiwara et al., 2004]. This filter restricts the analysis to frequencies below 30 Hz if the signal is not deconvolved by the transfer function [Oth et al., 2011], if allowance is made for the 3 dB drop.

In the present study, we have implemented a dataset based on the active Japanese shallow crustal accelerometric dataset build by Laurendeau et al. (2013). The different selection criteria are the following:

- Events between April 1999 and December 2009,
- Events described in the F-net catalog and for which Mw_{Fnet} is larger than 3.5,
- Shallow crustal events with a focal depth less than 25 km were selected and offshore events were excluded,
- Sites for which in surface $V_{S30} \geq 500$ m/s and at depth $V_{Shole} \geq 1000$ m/s,
- Sites with a complete velocity profile between the surface and the depth,
- Surface records with a predicted PGA > 2.5 gal using a magnitude-distance filter (Kanno et al., 2006),
- Following a visual inspection, faulty records, like S-wave triggers, and records from multi-events were eliminated or shortened,
- Records with at least 1 s of pre-event noise available,
- Events and sites with a minimum of three records (see Figure 6).

Our dataset finally consists of 2086 six-component records. The magnitude-distance distribution is shown on Figure 7. The magnitude range is 3.5 to 6.9 and the rupture distance range is 4 to 290 km. Figure 8 shows the locations of the 164 observation sites and the 272 earthquake epicenters. Figure 9 shows the V_S distribution of surface and depth records. In surface, there are no sites with V_{S30} larger than 1500 m/s. The downhole sites allow expanding the distribution until 3000 m/s. The distribution median is around 650 m/s in surface and around 1900 m/s at depth. Figure 10 shows the complete distribution. This distribution is approximately uniform for larger V_S than 800 m/s. In this case, the distribution median is around 1100 m/s.

Note that in the next parts, the datasets used will be not exactly the same depending on the application.

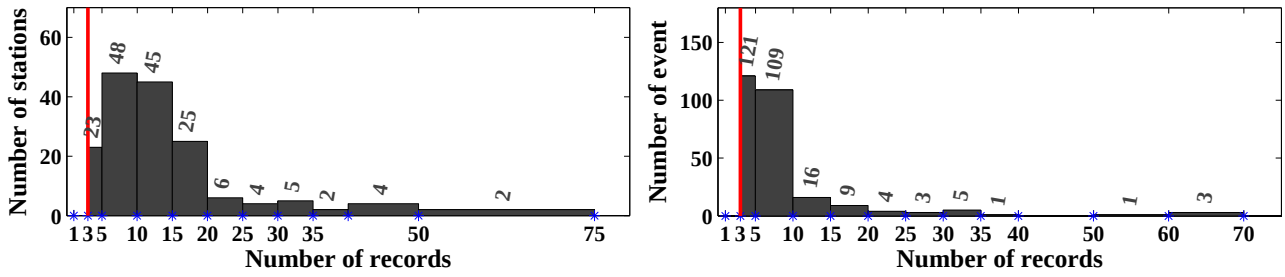

Figure 6: Number of records left) by station and right) by event.

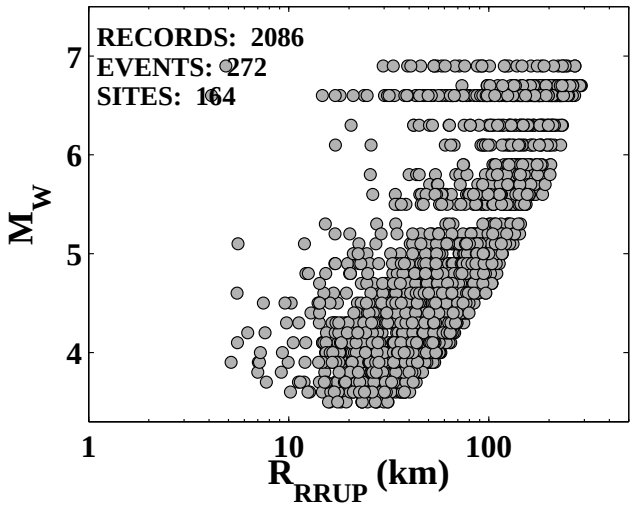

Figure 7: Distribution of the moment magnitude (M_W) and the rupture distance (R_{RUP}) of the selected KiK-net records.

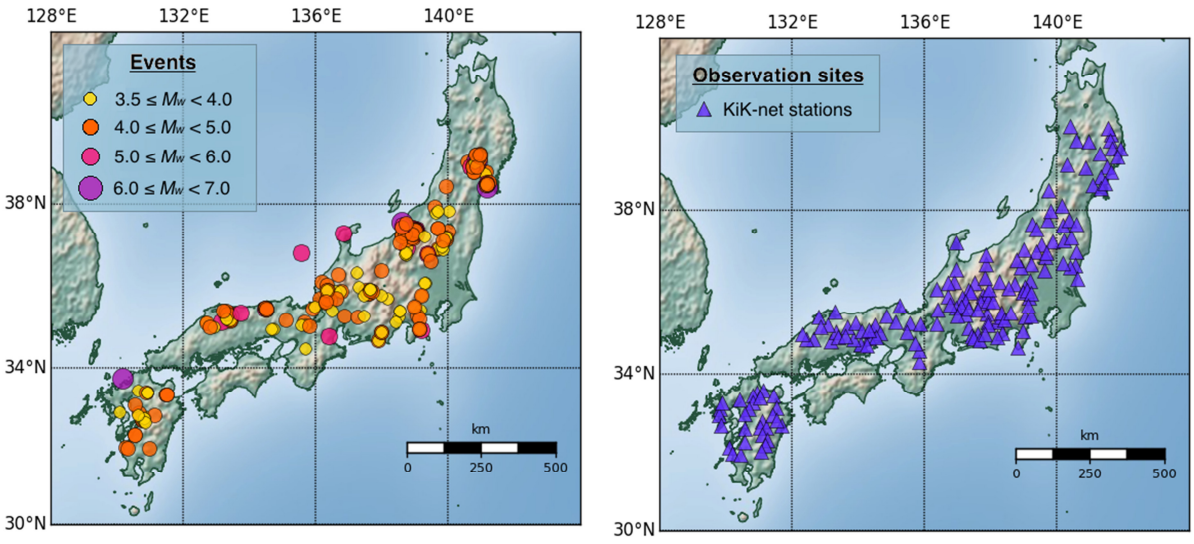

Figure 8: Events and recordings stations used in this study.

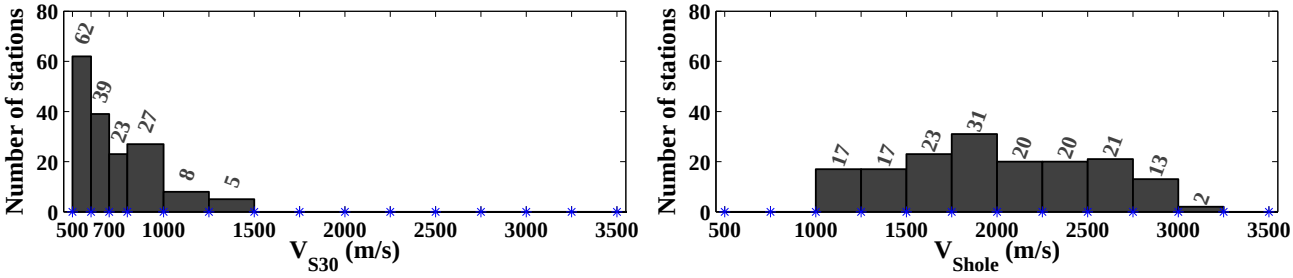

Figure 9: V_S distribution in terms of number of stations: right) in surface and left) in depth.

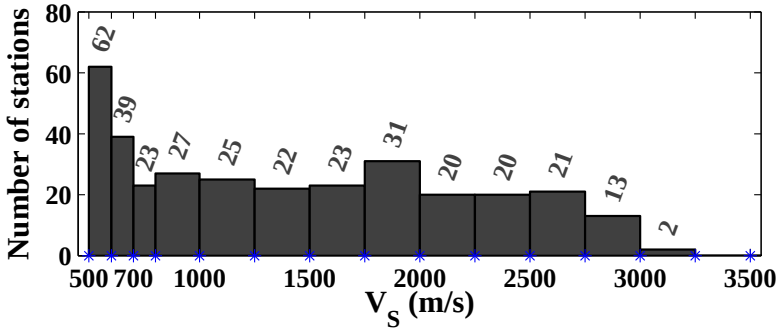

Figure 10: Complete V_S distribution in terms of number of station.

4 IMPLEMENTATION OF THE DIFFERENT APPROACHES TO ASSESS TRANSFER FUNCTIONS

4.1 SIGNAL PROCESSING

4.1.1 Time series windowing

4.1.1.1 S-wave window

The transfer functions are computed for the S-wave window to be compatible both with the model used in the generalized inversion and with the results in term of SH-waves from the Thompson-Haskell methodology. Firstly, the P- and S-wave arrivals are picking (noted T_P and T_S), as well as the end of the signal in the case of multishock events (noted T_{end}). In the way of Ktenidou et al. (2014), the theoretical S-wave window duration (D_{S_THEO}) is defined in respect of source and propagation terms. The chosen D_{S_THEO} definition is not exactly the same and it is the following:

[1]

$$D_{S\_THEO} = \frac{1}{f_c} + (T_S - T_P)$$

in which $(T_S\text{-}T_P)$ the time interval between the P- and S-wave arrivals and f_c is the source corner frequency defined as following:

[2]

$$f_c = 0.37\,\beta_S \left[\frac{16 \times \Delta\sigma \times 10^5}{7 \times M_0}\right]^{\frac{1}{3}}$$

in which β_S is the shear wave velocity (3500 m/s), $\Delta\sigma$ the stress drops (10 bars) and M_0 the moment magnitude defined such as:

[3]

$$M_0 = 10^{((M_W+6.03)\times 3/2)}$$

Figure 11 shows the effect of the source and the path on D_{S_THEO} definition. The maximum duration due to the source is around 17 s and due to the propagation is around 38 s. About the total duration, D_{S_THEO}, the minimum is 1.2 s and the maximum is 55 s.

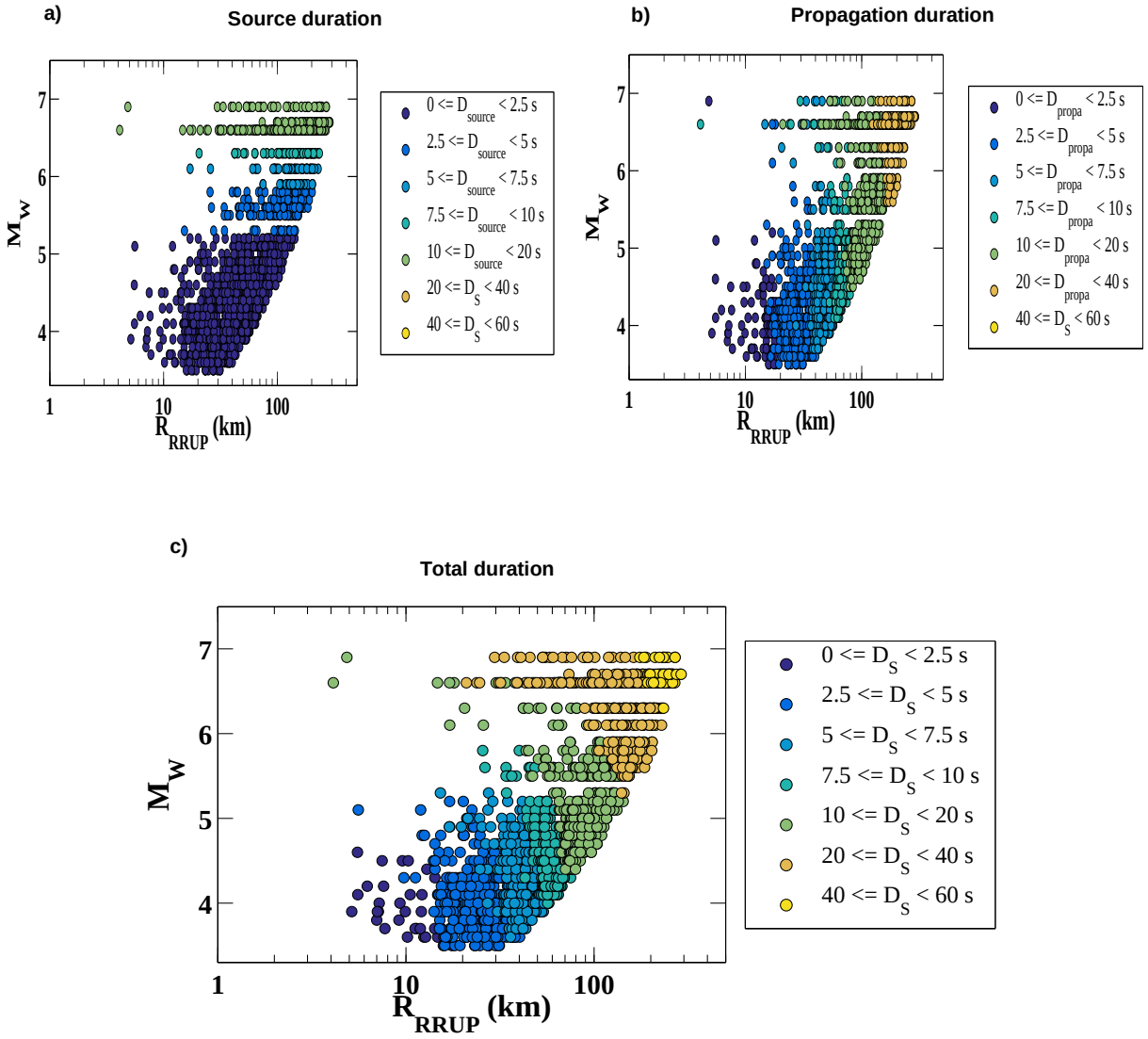

Figure 11: This figure shows the duration of the S-wave windows (c) in respect of (a) source (M_W effect) and (b) path (R_{RUP} effect) term duration.

It is necessary to fix a minimum duration for the S-wave window to have a good resolution at low frequency. For the generalized inversion, the maximum magnitude inversed is 5.0 giving a source corner frequency of 0.52 Hz with the previous parameters. We sought to go down to a minimum frequency of 0.3 Hz. To define the minimum duration of the S-wave windows, different duration range have been tested, as shown on Figure 6, and finally we have decided to fix a minimum duration of 10 s (D_{S_min} = *10 s*). In the dataset, 1074 records have a theoretical duration lower than 10 s. Finally, the S-wave windows are defined as:

[4]

$$D_S = min\left\{max\left\{D_{S\_min}, D_{S\_THEO}\right\}, D_{S\_AV}\right\}$$

in which D_{S_AV} is the duration available between T_S and T_{end}.

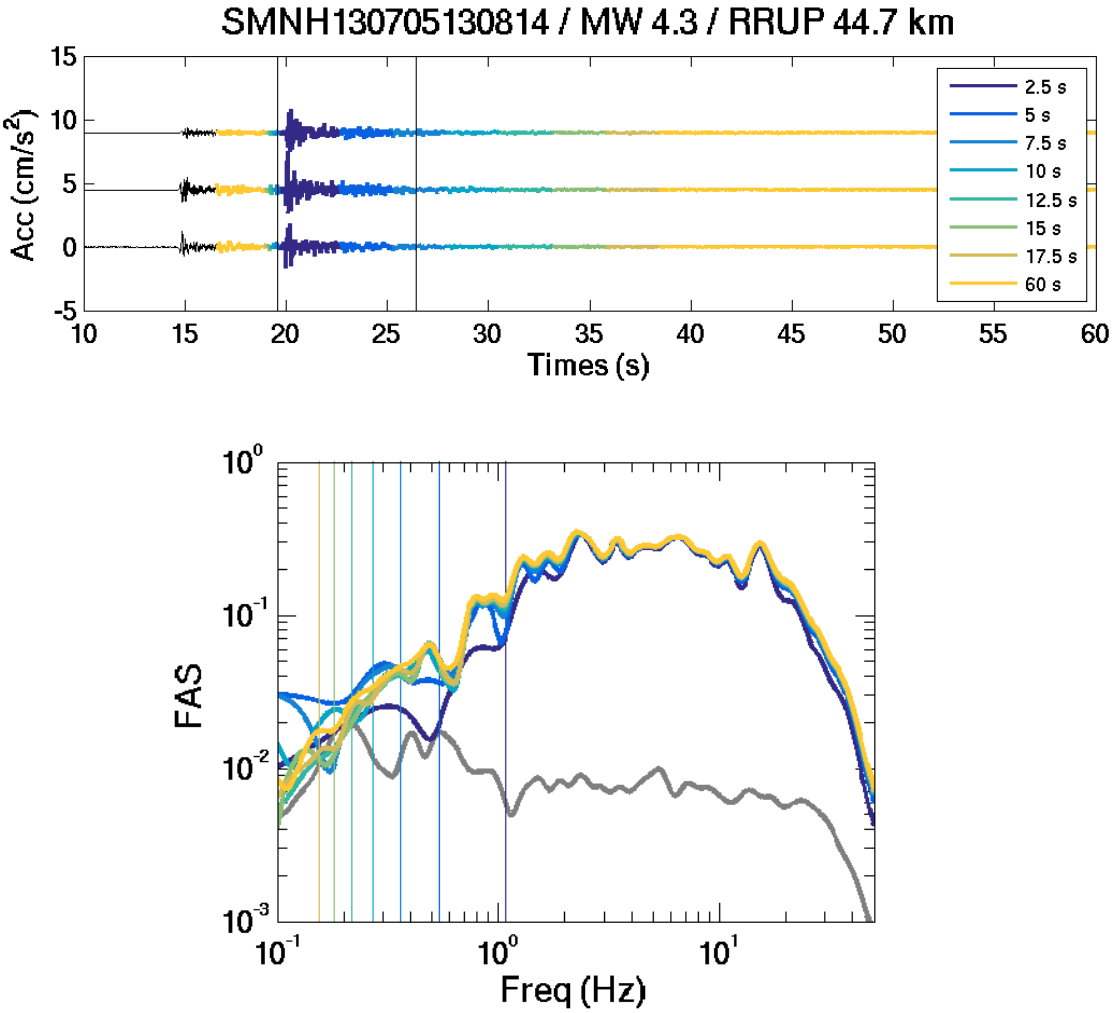

Figure 12: Illustration of the S-wave window duration impact on the computation of the Fourier acceleration spectrum (FAS). a) The times histories and the selected window are shown. The window is a little bit larger because the cosine taper is applied outside the limit interval of the defined S-wave window, i.e. [T_S T_S+D_S]. b) The corresponding FAS are shown. FAS in grey correspond to the noise spectrum. The vertical lines indicate the minimum frequency associated with the S-wave window and allowing to have at least 3 wavelengths (f_{min} = 3 / D_{S_min}). In this example, it is necessary to have at least 10 s of signal to have a stable spectrum at 0.3 Hz.

4.1.1.2 Noise window

To be consistent with the minimum S-wave window duration, we would like to have at least 10 s of noise. However, 10 s of noise is not available before the event for all records, as shown on Figure 13. The dataset contains 311 records without 10 s of pre-event noise. That's why we have tested different noise window definitions for which the energy is then compared. The idea is to select a noise window with a sufficient duration but also a window with a representative level of energy (without seismic signal included).

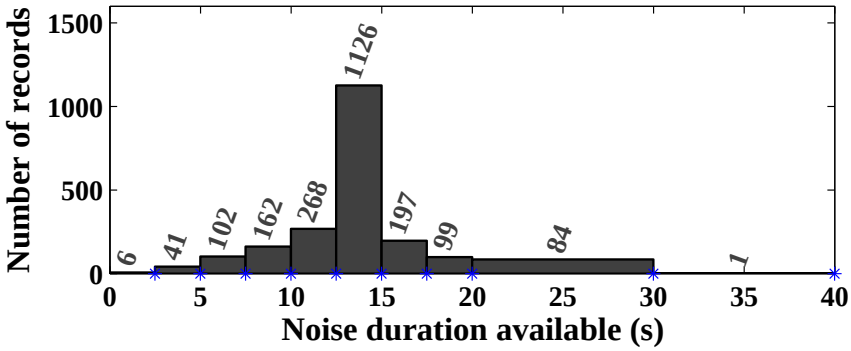

Figure 13: Histogram illustrating the number of records versus the pre-event noise duration available.

The time intervals of the 3 compared noise windows are defined as following:

[5] N1:

$$Int_{N1} = [0 \ ; \ T_P - 0.1s]$$

[6] N2:

$$Int_{N2} = \left[max \left\{ \begin{matrix} T_{end} - max \left\{ \begin{matrix} 10 \\ T_p - 0.1 \end{matrix} \right\} \\ T_S + D_S \end{matrix} \right\} \ ; \ T_{end} \right]$$

[7] N3:

$$Int_{N3} = \left[max \left\{ \begin{matrix} T_{end} - max \left\{ \begin{matrix} 10 \\ T_p - 0.1 \\ D_S \end{matrix} \right\} \\ T_S + D_S \end{matrix} \right\} \ ; \ T_{end} \right]$$

N1 is a pre-event noise window while N2 and N3 are post-event noise windows. N2 and N3 are different only when the S-wave window is larger than 10 s and larger than the pre-event noise window. When it is possible (at least 5 s of noise for N1 and at least 10 s for N2 and N3) the mean energy is estimated for the 3 noise windows. The mean energy (E) is defined as:

[8]

$$E = \frac{\sum_{i=fmin}^{Nf} \sqrt{FAS_{EW}(f)^2 + FAS_{NS}(f)^2 + FAS_Z(f)^2}}{Nf}$$

in which f_{min} is the minimum frequency defined according to the window length ($f_{min}=3/D_{min}$), *Nf* is the number of frequency point, *FAS* are the Fourier acceleration spectra computed for the 3 components.

The energy for the 3 noise windows is compared. Some factors are implemented in order to prefer N2 and N3 if N1 duration is lower than 10 s and to prefer N1 if N1 duration is larger than 10 s.

The duration of the 3 noise windows (N1, N2 and N3) are noted D_{N1}, D_{N2} and D_{N3} and their mean energy E_{N1}, E_{N2} and E_{N3}, respectively. Some factors are implemented in order to favor N2 and N3 if N1

duration is lower than 10 s and to favor N1 if D_{N1} is larger than 10 s. The complete scheme is presented on Figure 14. Table 1 presents the number of selected records for each noise window type.

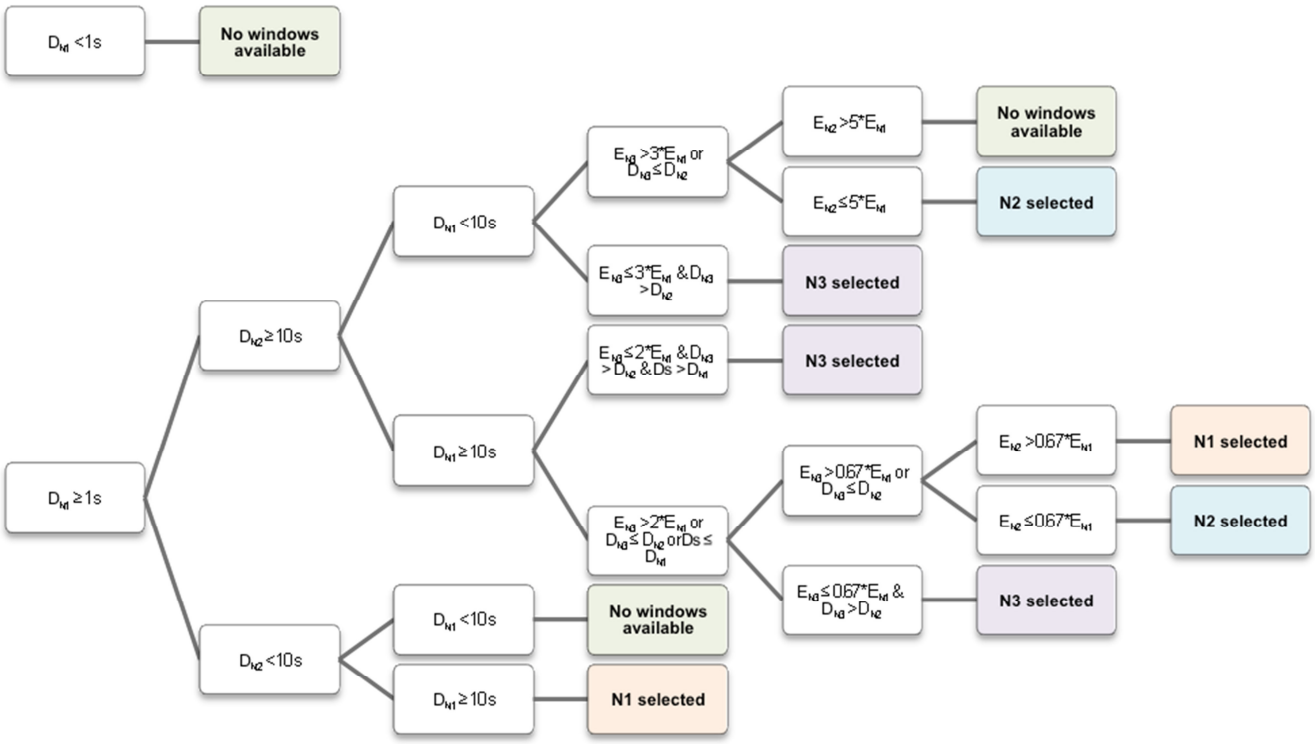

Figure 14: Scheme of the noise window selection methodology.

Table 1: This table indicates the number of records selected both at depth and in surface for each noise window definition.

	Depth	Surface
Pre-event noise	1567	1578
End-event noise (short)	57	96
End-event noise (long)	462	412

4.1.2 Processing

For each component in surface and at depth, the following processing is applied:

Step 1 - A first-order baseline operator is applied to the entire record, in order to have a zero-mean of the signal, and a simple baseline correction is applied by removing the linear trend.

Step 2 - A 5% cosine taper is applied on each side of the signal. In the case of the S-wave window, the taper is applied outside the defined time interval (D_S), in order not to bias the maximal amplitude.

Step 3 - The records with a 200-Hz sampling frequency are resampled to 100 Hz.

Step 4 - At the end of the records, zeros are added in order to have the same length for each records (here 8192 samples).

Step 5 - We computed the Fourier acceleration spectra (FAS) from each component and also from the 2D complex time-series of the two orthogonal horizontal time histories following Thompson et al. (2012).

Step 6 - FAS are smoothed according to the Konno and Ohmachi (1998) smoothing with b=30. Then, FAS are resampled for a 500 logarithmically spaced sample vector between 0.1 and 50 Hz.

Step 7 - The signal-to-noise ratio (SNR) is computed and the continuous frequency band with SNR>3 is noted. The minimum frequency is at least $f_{min}=3/D_{min}$, so f_{min} is at maximum 0.3 Hz when $D_{min}=10$ s, and f_{min} is lower for the strong events at large distances. Figure 15 gives an idea of the number of records with a good SNR on the whole frequency band at depth and in surface for the 2D complex series and the vertical component. At low frequency, SNR seems to be better for the horizontal components that for the vertical component. At least 70 % of the records could be used in the next parts to compute the transfer functions from the different methods.

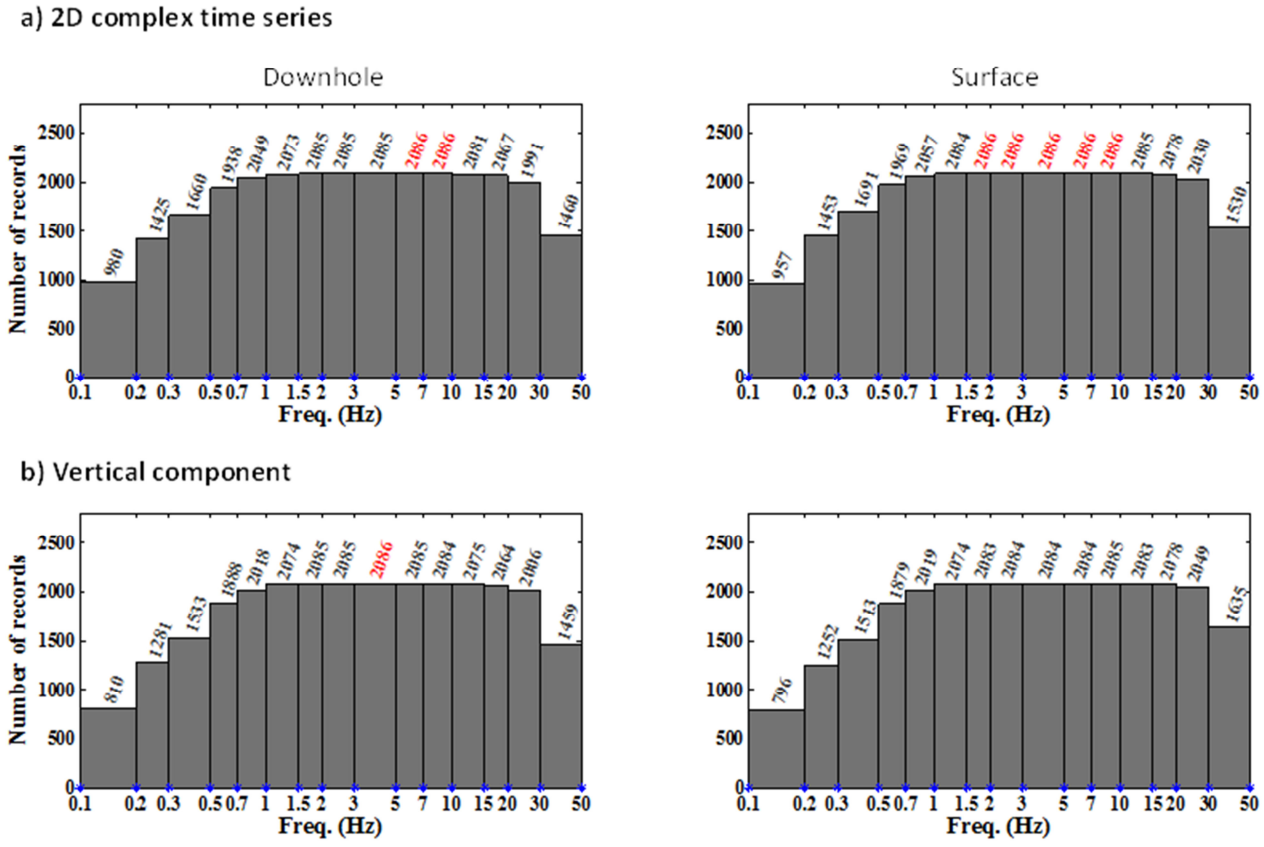

Figure 15: Number of records with a continuous signal-to-noise ratio larger than 3 for different frequency range at depth and in surface. These are the results for the 2D complex time-series (a) and for the vertical component (b). The record is counted in a class if its lower limit is greater than or equal to f_{min} and if its upper limit is less than or equal to f_{max}. In red, 100 % of the records have a good SNR for the frequency class.

4.2 TRANSFER FUNCTIONS FROM THE EMPIRICAL METHODS (SSR)

A classical way to estimate site effects consists in using spectral ratios of ground motion recordings. The horizontal-to-vertical spectral ratios (HVSR) allow only highlighting f_0 (under some assumptions) that characterize the resonance frequency in presence of strong contrasts (e.g., Lermo and Chavez-Garcia, 1993). The standard spectral ratios (SSR, initially proposed by Borcherdt 1970) allow quantifying the transfer function amplitude (Lebrun et al., 2001). The spectral ratios are computed relative to a site reference chosen either on the surface or on the downhole but near the site, in order that the characteristics of the incident wave field are similar. The choice of the reference site is a sensitive issue. In this section, the empirical SSR and HVSR are estimated for the KiK-net sites.

4.2.1 The dataset

For each record, the signal-to-noise ratio (SNR) is estimated for the 2D complex time-series of the S-wave window. The record is selected if SNR is larger than 3 for a continuous frequency band between 0.5 and 15 Hz both for the vertical component and the 2D complex time series. In addition, the geometrical mean of the three empirical spectral ratios is estimated if there are at least 3 records for the site. Finally, the empirical spectral ratios are computed for 152 sites from 1488 records (see Figure 16).

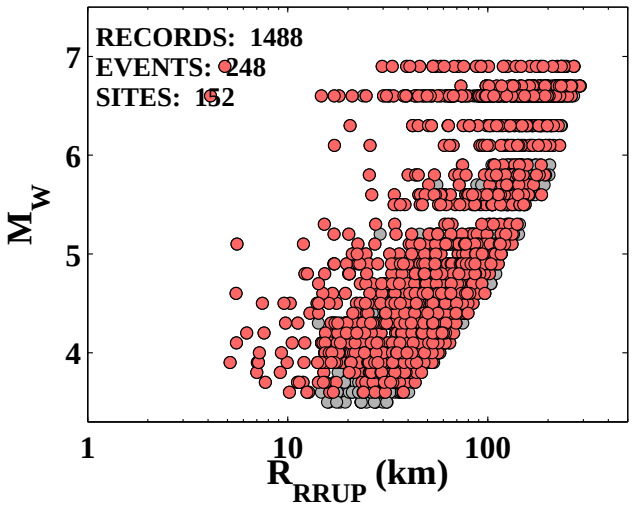

Figure 16: Distribution of the moment magnitude (M_W) and the rupture distance (R_{RUP}) of the KiK-net records selected to compute the empirical spectral ratios (red dots). These records have a continuous signal-to-noise ratio larger than 3 between 0.5 and 15 Hz (in this case the signal is the S-wave window). Besides, the empirical spectral ratios are computed for each site with at least 3 good records. The grey dots correspond to the initial data not selected.

4.2.2 The empirical spectral ratios for a particular site

The three empirical ratios (SSR, HVSR_surf and HVSR_dh) are estimated both in average and also with respect of the peak ground acceleration (PGA) and of the azimuth. This representation allows to highlight the nonlinear and geometrical effects, respectively. In this section, some examples are described.

Figure 17 presents the example of a site with relatively flat transfer functions, such as what is expected for a reference motion. However, the station MYGH06 is characterized by a relatively low V_{S30}, around 600 m/s. The only peak observed on SSR is caused by destructive interferences at depth, because a desamplification is also observed on HVSR_dh. With a reference at the surface, this peak would not be present, and SSR would be different.

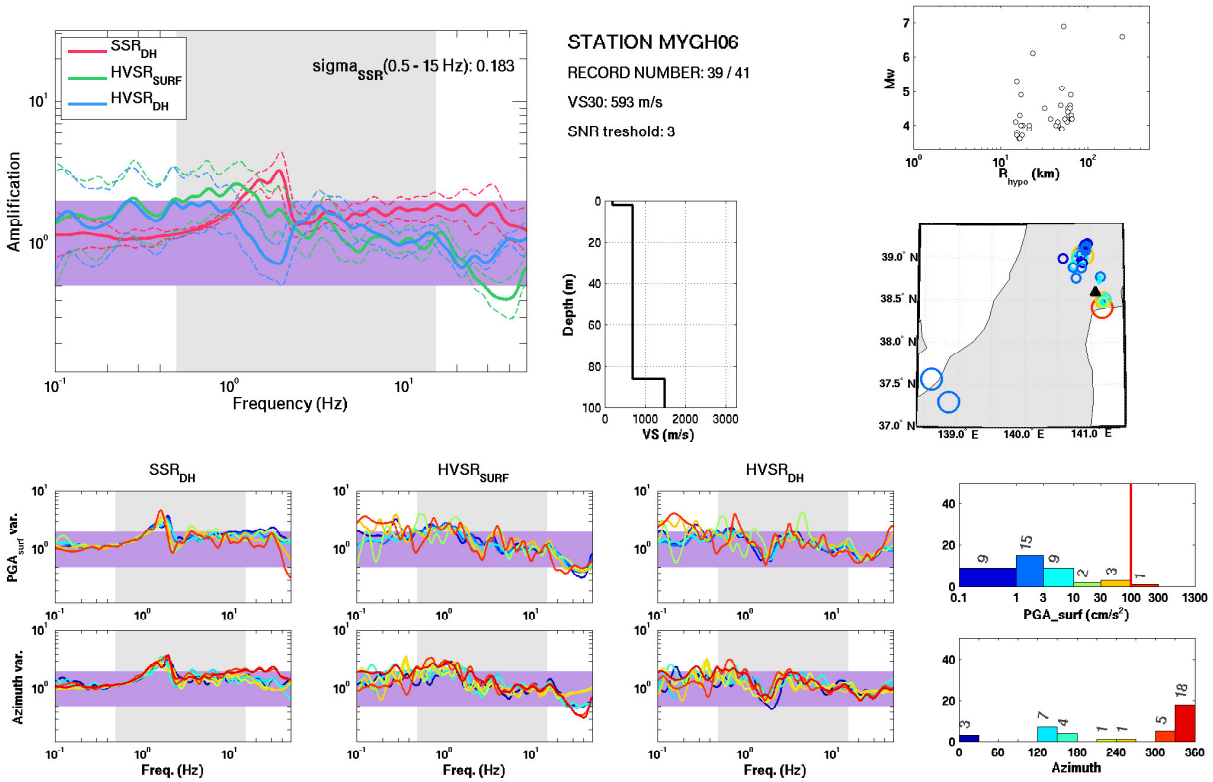

Figure 17: Example of a site with relatively flat transfer functions and with a relatively low V_{S30} (MYGH06). On the top left, the mean empirical spectral ratios and the mean ± 1 standard deviation are shown. The number of records used is indicated at the right of this subplot and their magnitude-distance distribution on the left. On the centre, the available shear wave velocity profile ($V_S(z)$) is given. On the left at the middle, the location of the station is shown by a black triangle and the location of the different sources by coloured circles. The size of these circles depending on the magnitude and the colours corresponds to the observed PGA at the surface represented in a histogram below. On the bottom, the mean empirical spectral ratios are computed with respect to ranges of surface PGA (top line) and azimuth (bottom line). On the right, the histograms give the colour associated with a bin and the number of records used to compute the mean spectral ratios. In the histogram about PGA, the vertical red line corresponds to a PGA of 0.1 g, the limit chosen by Thompson et al. (2012) for the nonlinearity. The purple patch corresponds to amplification levels between 0.5 and 2 and allows to highlight the amplification. The grey patch corresponds to the frequency range for which the signal-to-noise-ratio (0.5 -15 Hz) is estimated and is higher than 3.

On the contrary, Figure 18 displays an example of a site showing a large amplification, while this site, TCGH14, is characterized by a relatively high V_{S30}, around 850 m/s. This amplification around 10 Hz reaches 30.

Figure 19 presents the example of IWTH25 site which is characterized by variable transfer functions. This station has recorded numerous ground motions with the advantage to have a good distribution in terms of PGA and azimuth. Above 5 Hz, SSR and HVSR_surf are different especially for the mean spectral ratios computed for PGA larger than 0.3 g. Besides, SSR displayed for different azimuth range present some amplification between 1 and 5 Hz in the case of motion recorded on the north. The other azimuthal motions give a SSR more flat at these frequencies. The large variability observed at this station could be explained both by nonlinearity and geometrical effects. In addition, the broadband amplification observed is representative of a basin.

The station NIGH07 is another example of a site presenting nonlinearity (see Figure 20). However, for this station, the mean spectral ratios, especially SSR, evolve gradually with the PGA. The frequencies of the amplification peak decrease as well as the amplitude when the PGA increases. In this example, the limit of 0.1 g chosen by Thompson et al. (2012) for the nonlinearity does not seem a priori appropriate.

Figure 21 presents the spectral ratios for the NGNH35 site, which is a site having a good correlation coefficient between empirical and theoretical spectral ratios and it is used as an example in the section about theoretical spectral ratios.

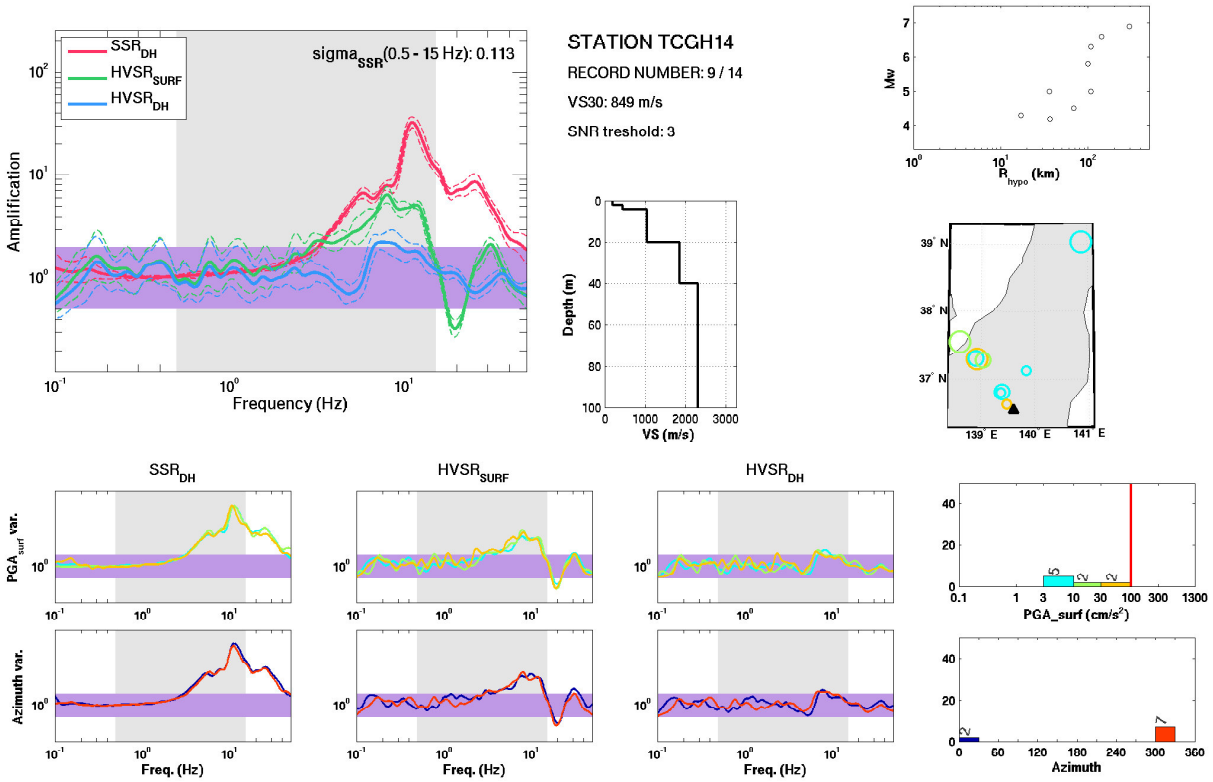

Figure 18: Example of a site showing a large amplification with a relatively high V_{S30} (TCGH14).

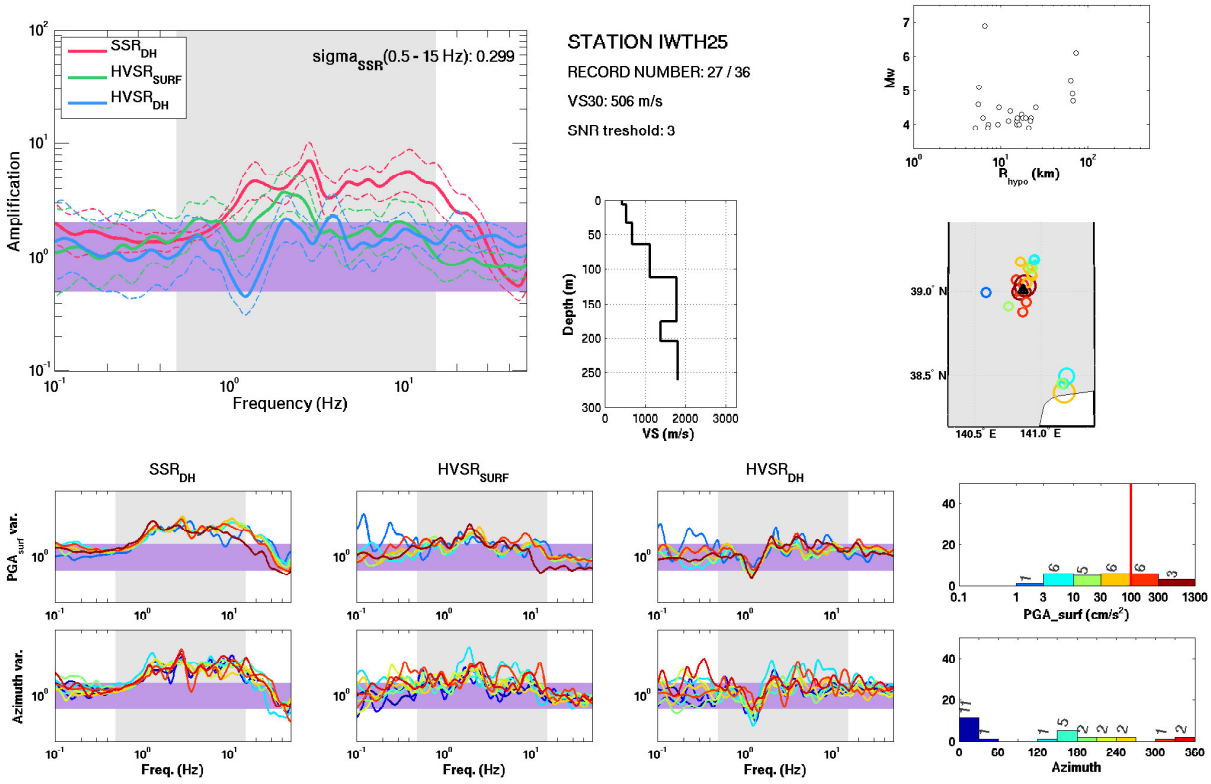

Figure 19: Example of a site with a variable transfer function (IWTH25). This large variability could be explain by geometrical effects

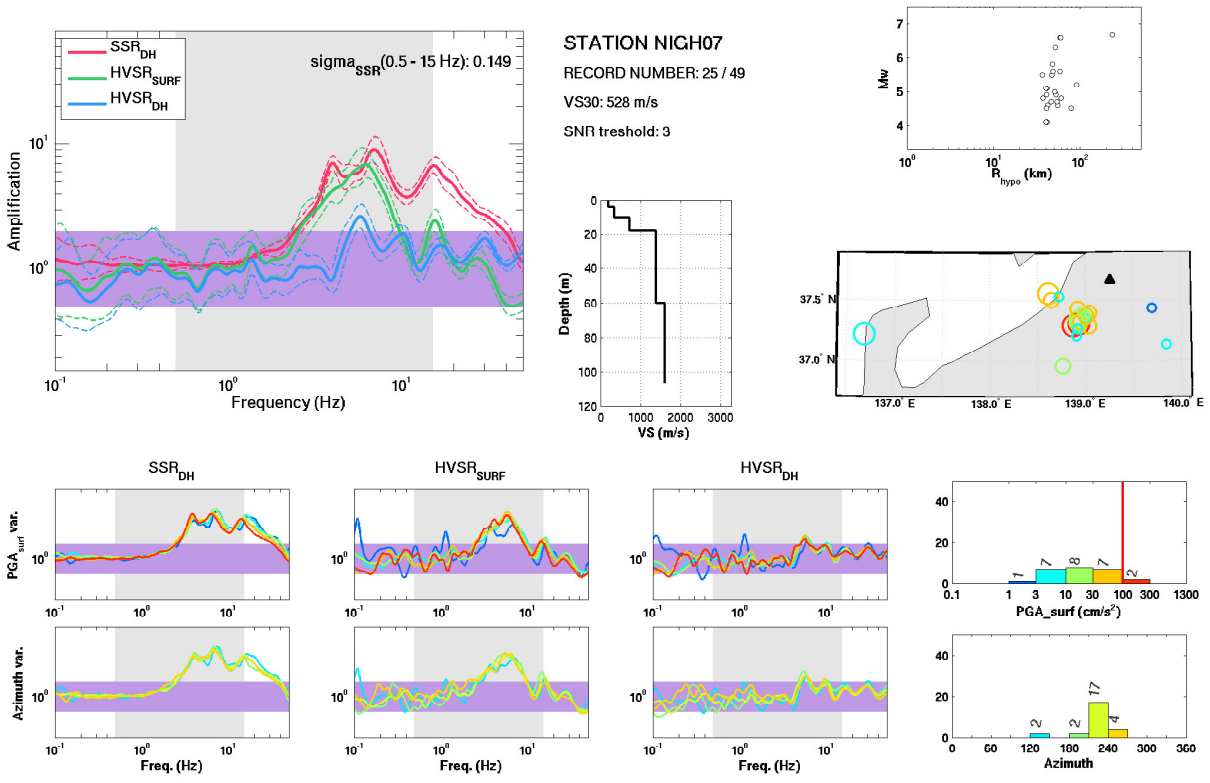

Figure 20: Example of a site with a transfer function showing nonlinear effect (NIGH07).

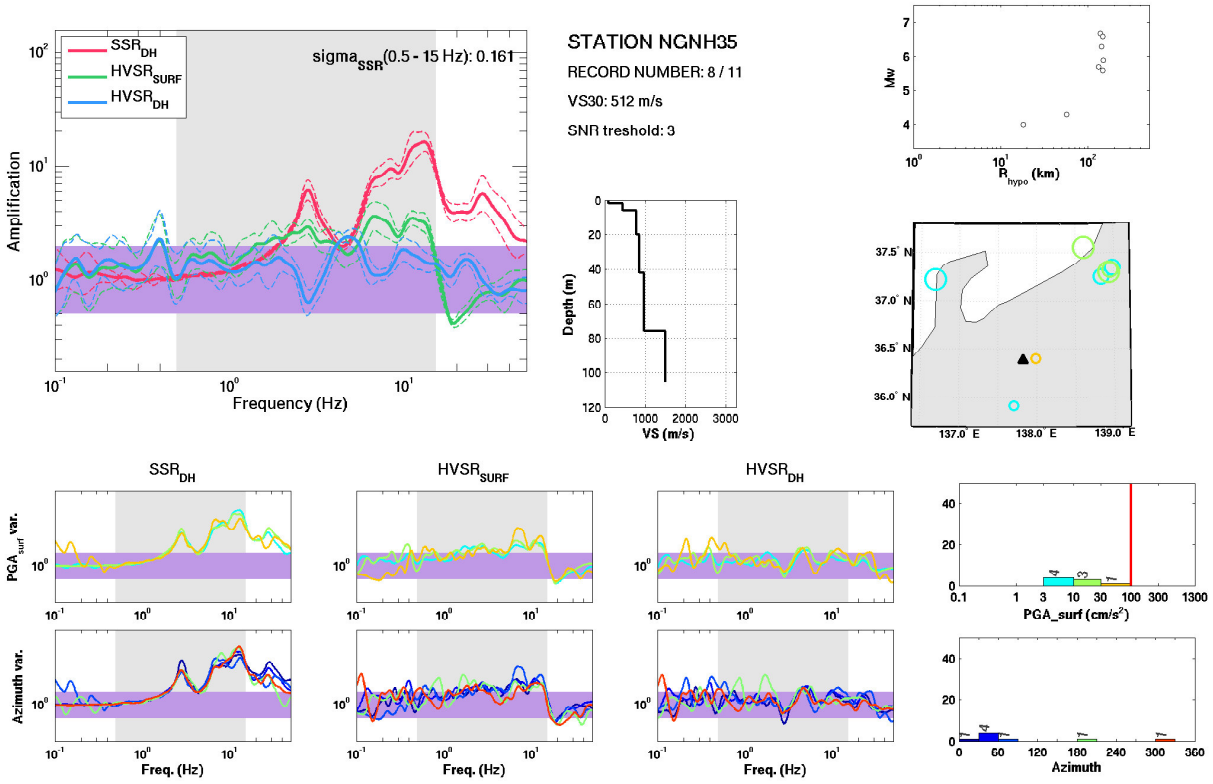

Figure 21: Example of a site with a very good agreement between empirical and theoretical spectral ratios (NGNH35).

4.2.3 The mean empirical spectral ratios

For the sites with $V_{S30} \geq 500$ m/s, the mean empirical standard spectral ratio presents a large amplification around 6 at 10 Hz (see Figure 10). The variability is larger around this frequency. Some sites even have amplification larger than 30 (8 sites). This amplification around 10 Hz can be also observed on HVSR_surf but with lower amplitude. HVSR_dh has an approximately flat mean ratio and with an amplitude around 1.

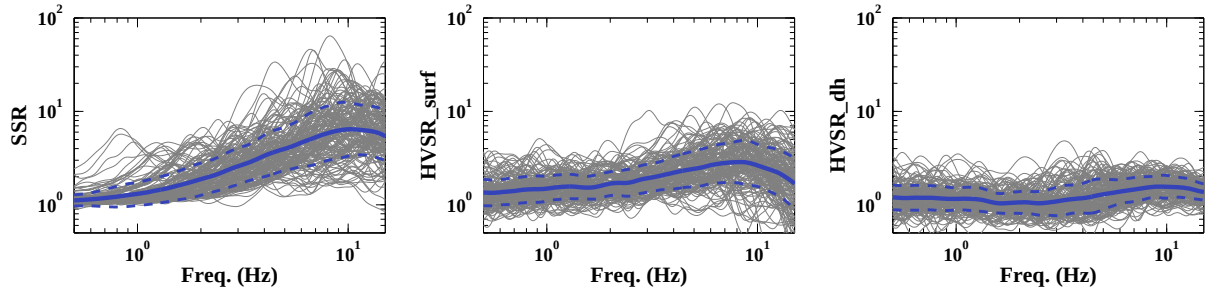

Figure 22: Empirical spectral ratios of the 152 selected sites (in grey). The mean spectral ratios are represented by the blue solid lines and the mean more or less one standard deviation by the blue dashed lines.

Figure 23 shows the mean empirical spectral ratios for different V_{S30} range. The peak of the mean spectral ratios is larger frequency when V_{S30} increases (for SSR and HVSR_surf). Thus, the amplitude tends to be larger at high frequency for the highest V_{S30}. In fact, there are two competing mechanisms at work: the amplification due to low-velocity shallow layers for which the frequency of the peak increase with V_{S30}, and the site attenuation effect, κ_0, which are lower when V_{S30} increases.

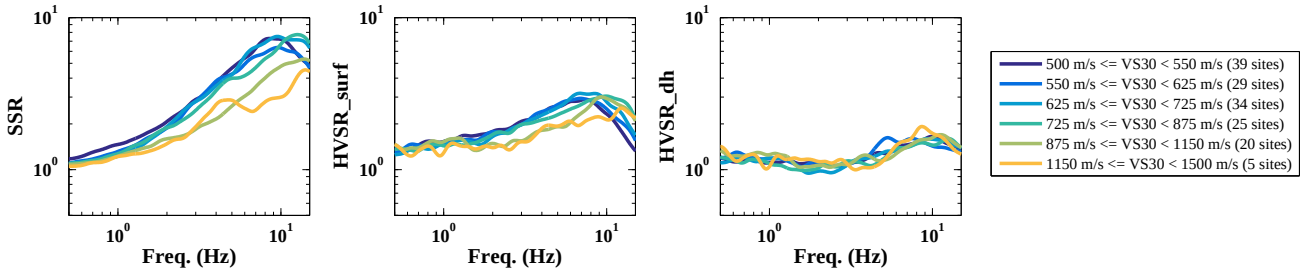

Figure 23: Mean empirical spectral ratios for different V_{S30} range.

4.3 TRANSFER FUNCTIONS FROM 1D SIMULATION

In this section, the same notation than Cadet et al. (2012) is used.

4.3.1 Computation of the theoretical transfer functions

The theoretical transfer functions (*TF*) is computed from a 1D reflectivity model (Kennett 1974) reproducing the response of horizontally stratified layers excited by a vertically incident SH plane wave (original software written by J.-C. Gariel and P.-Y. Bard and used previously in a large number of investigations: e.g., Bard and Gariel 1986; Theodulidis and Bard 1995; Cadet et al. 2012). In addition to the shear-wave velocity profile ($V_S(z)$) derived from downhole measurements, the profile of P-waves velocity ($V_P(z)$), unit masses ($\rho(z)$) and damping ($Q_P(z)$ and $Q_S(z)$) are also needed to compute the transfer functions. The relationships of Brocher (2005) are used to get $V_P(z)$ and $\rho(z)$ from $V_S(z)$:

[9]
$$V_P(z) = 0.9409 + 2.0947\,V_S(z) - 0.8206\,V_S(z)^2 + 0.2683\,V_S(z)^3 - 0.0251 V_S(z)^4$$

[10]
$$\rho(z) = 1.6612 - 0.4721\,V_P(z) + 0.0671\,V_P(z)^2 - 0.0043\,V_P(z)^3 + 0.000106\,V_P(z)^4$$

in which $V_S(z)$ and $V_P(z)$ are in km/s and $\rho(z)$ in g/cm^3.

The quality factors are deduced from $V_S(z)$ according these relationships:

[11]
$$Q_P(z) = max \begin{Bmatrix} V_P(z)/20 \\ V_S(z)/5 \end{Bmatrix}$$

[12]
$$Q_s(z) = V_S(z)/X_Q$$

in which X_Q is a variable, classically chosen equal to 10 (e.g., Cadet et al. 2012).

The quality factors are assumed to be independent of frequency. The TFs are computed for 2048 frequency samples regularly spaced from 0.1 to 50 Hz. We compute the transfer function *TF* at the surface, and also the one at the sensor depth *TF_dh*. The Konno and Ohmachi (1998) smoothing is applied with a b coefficient of 10, therefore lower than in the case of real data but theoretical TFs require a stronger smoothing. Besides, the “theoretical standard spectral ratio” is estimated and it is called *BTF* (*BTF=TF/TF_dh*).

Figure 24 shows the impact of different input parameters for BTF computation. The variations of the incidence angle have little effect on BTF, while the variations of Q_S definition have stronger impact, especially at high frequency for the harmonics and also the smoothing strongly decreases the amplitudes.

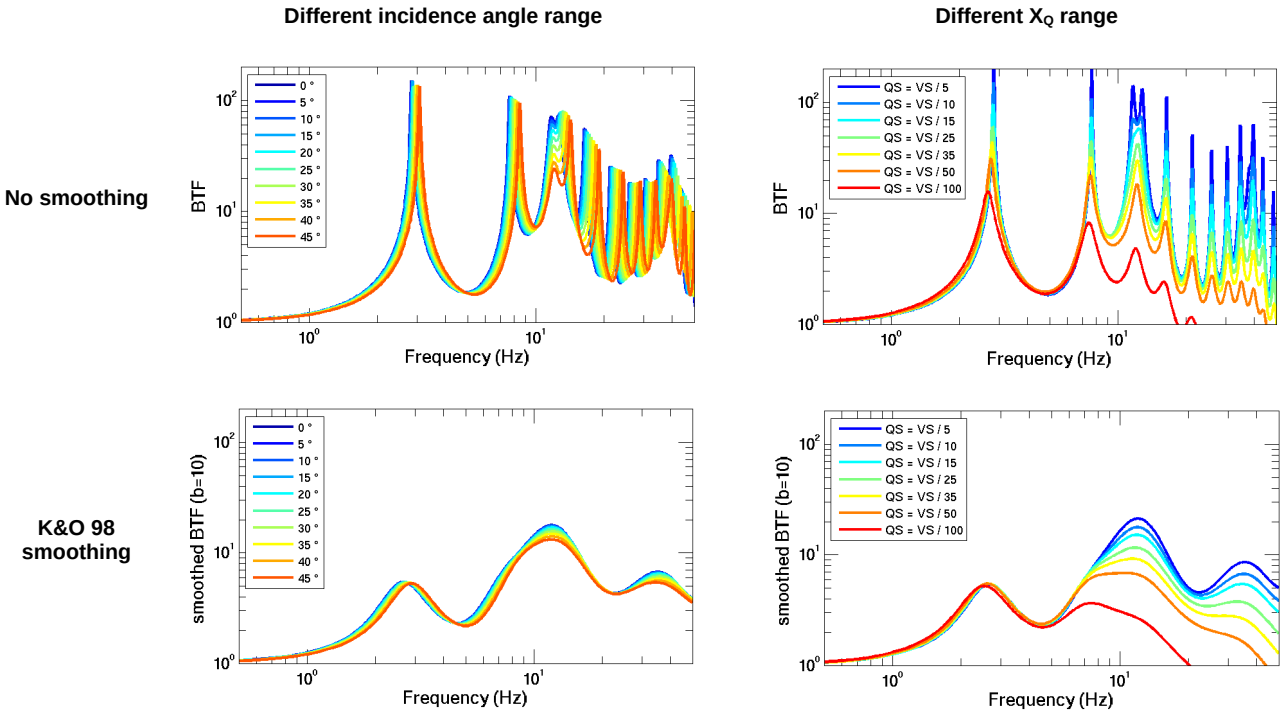

Figure 24: Illustration of the effect of different input parameters (incident angle, Q_S definition and smoothing) on the theoretical spectral ratios, BTF, in the example of the station NGNH35 (see Figure 21).

4.3.2 The mean theoretical spectral ratios

The mean theoretical standard spectral ratio (*BTF*) presents a broadband amplification between 8 and 30 Hz with amplitude around 5 (see Figure 25). The variability is larger around this frequency range. The amplification difference between low frequency and high frequency is even more pronounced on TF. BTF includes the destructive interferences can be observed on TF_dh. For our dataset, the destructive frequencies are between 0.5 and 7 Hz.

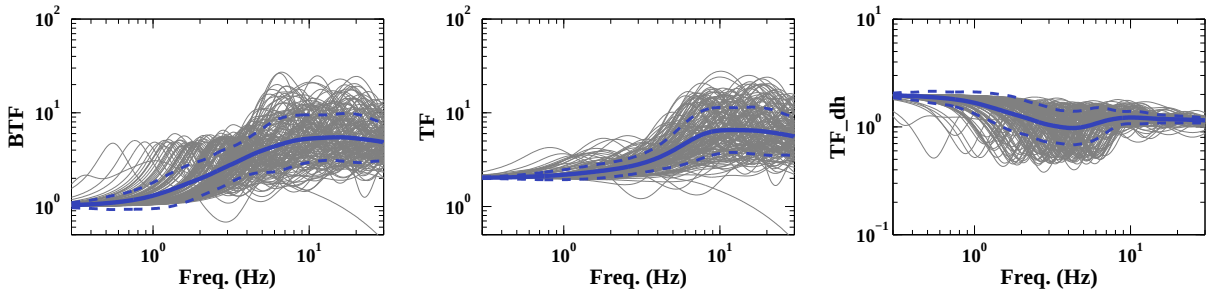

Figure 25: Theoretical transfer functions of the 152 selected sites (in grey). The mean spectral ratios are represented by the blue solid lines and the mean more or less one standard deviation by the blue dashed lines.

Figure 26 shows the mean theoretical transfer functions for different V_{S30} range and for two Q_S definitions. For *BTF* and *TF*, the frequency peaks increase more regularly with V_{S30} than for the empirical

spectral ratios (see Figure 23). Besides, the amplitude is higher going toward the high frequency when V_{S30} increases. This behaviour is similar to that is observed on the host-to-target adjustment curves, which are defined as a function of V_{S30} *and* κ_0 from stochastic simulation (Boore 2003). For example, the theoretical adjustment factor (soft-rock-to-rock ratio), showed in the study of Laurendeau et al. (2013) for a soft-rock with a V_{S30} of 550 m/s and an rock with a V_{S30} of 1100 m/s and computed in the same way than the one of Van Houtte et al. (2011), presents between 10 and 40 Hz a larger amplification for the rock than for the soft-rock. Around 25 Hz, the theoretical soft-rock-to-rock ratio is around 0.7. In the case of the Figure 18, the ratio between the first curve of TF (Vs around 525 m/s) and fifth curve of TF (VS around 1012 m/s) give 0.85 for X_Q=10 and 0.63 for X_Q=30.

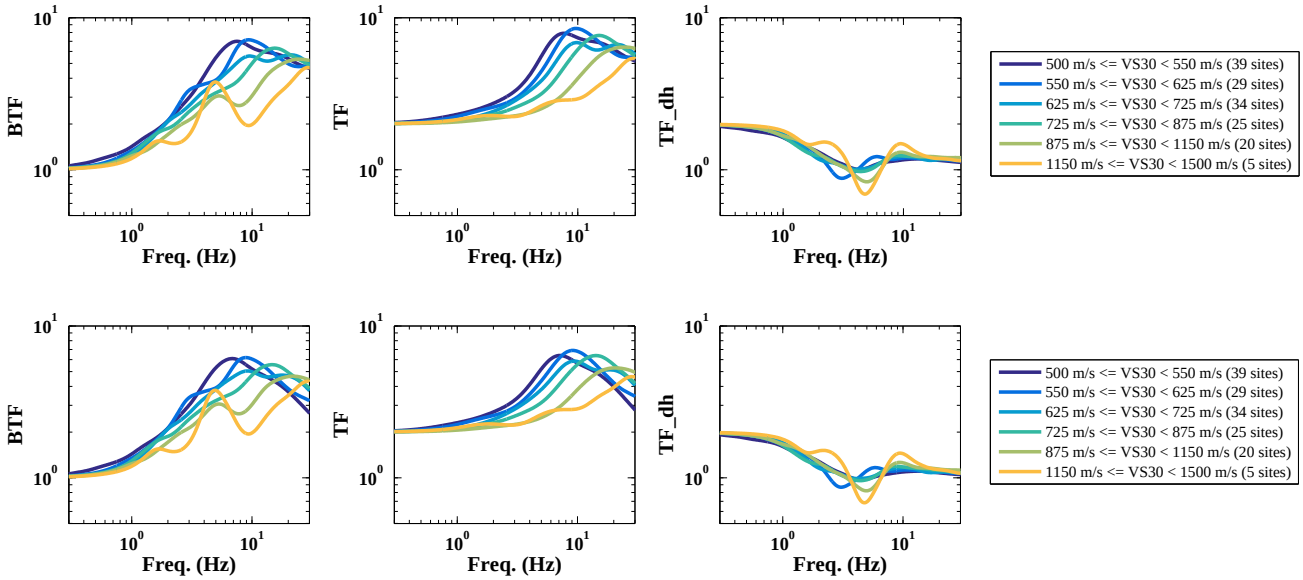

Figure 26: Mean theoretical transfer functions for different V_{S30} range and for two definitions of Q_S: a) X_Q=10 and b) X_Q=30.

4.4 TRANSFER FUNCTIONS FROM THE GENERALIZED INVERSION TECHNIQUE

4.4.1 The Generalized Inversion Technique (GIT)

A widely used method to simultaneously determine source, path and site terms from a large dataset of multiple earthquakes recorded at multiple stations is the Generalized Inversion Technique (GIT) introduced by Andrews (1986). In this part of the study, in order to compute transfer functions for each KiK-net site we apply the parametric inversion methodology developed by Drouet *et al.* (2008, 2010) to S-wave acceleration Fourier spectra of our dataset.

The far-field S-wave acceleration spectrum $A_{ijk}(r_{ij}, f_k)$ can be written as the product of source, propagation and site response terms:

[13]
$$A_{ijk}(r_{ij}, f_k) = \Omega_i(f_k) \times D_{ij}(r_{ij}, f_k) \times S_j(f_k)$$

where r_{ij} is the hypocentral distance from earthquake i to station j and f_k the frequency.

The source term is described using the Brune's source model (Brune 1970):

[14]
$$\Omega_i(f_k) \sim \frac{(2\pi f_k)^2 M_{0i}}{1 + \left(\frac{f_k}{f_{ci}}\right)^2}$$

where M_{0i} is the seismic moment and f_{ci} the corner frequency of event i.

The attenuation term, accounting for path effects, involves an-elastic decay and geometrical spreading:

[15]
$$D_{ij}(r_{ij}, f_k) = exp\left(-\frac{\pi r_{ij} f_k}{Q(f_k) v_S}\right) \times \frac{1}{r_{ij}^{\gamma}}$$

where v_S is the average S-wave velocity along the path, $Q(f_k) = Q_0 \times f_k^{\alpha}$ is the frequency-dependent quality factor and γ is the coefficient of geometrical spreading.

The site response term $S_j(f_k)$ corresponds to the site effect at the station j. This term is equal to unity for all frequencies in the absence of site effect (*i.e.* "rock" site conditions).

Equation [13] can be linearized by taking the base 10 logarithm of the Fourier spectra:

[16]
$$y_{ijk} = m_{0i} - \log_{10}\left(\frac{(2\pi f_k)^2}{1 + \left(\frac{f_k}{f_{ci}}\right)^2}\right) - \gamma \log_{10}(r_{ij}) - \frac{\pi r_{ij} f_k^{1-\alpha}}{\log_e(10) Q_0 v_S} + s_{jk}$$

where:

[17]
$$y_{ijk} = log_{10}\left(A_{ijk}(r_{ij}, f_k)\right)$$

[18]
$$m_{0i} = log_{10}\left(M_{0i} \times \frac{2R_{\theta\phi}}{4\pi\rho\beta^3}\right)$$

[19]
$$s_{jk} = log_{10}\left(S_j(f_k)\right)$$

with $R_{\theta\phi}$ the source radiation pattern, assumed to be constant ($R_{\theta\phi} = 0.55$ for S-waves, Boore & Boatwrigth 1984), ρ the density, β the S-wave velocity of the medium at the source and v_S the S-wave velocity along the path (we assume $\beta = v_S = 3.5\ km/s$ and $\rho = 2800\ kg/m^3$, as in Oth *et al.* (2010) for crustal earthquakes in Japan).

Equation [16] describes a classical linear system of the form $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{b}$ is the data vector containing the logarithmic spectral amplitudes, $\mathbf{x}$ is the vector containing the model parameters and $\mathbf{A}$ is the system matrix related them. This system is solved here by using an iterative Gauss-Newton inversion scheme.

4.4.2 The dataset

In order to prepare data for the generalized inversion, all the records of the dataset are processed as described in part 4.1 to compute S-wave acceleration Fourier spectra. However, KiK-net stations are characterized by frequency response of instruments corresponding to a Butterworth filter with a cut-off frequency equal to 30 Hz. In order to consider frequencies above this limit, a full correction of signals by deconvolution of this instrument response is necessary. Figure 27 shows an example of the correction we applied to data.

Moreover, the simultaneous analysis of small and large earthquakes imposed the use of a common frequency range with a good signal-to-noise ratio. Because of this constraint, we only keep data with signal amplitude greater than three times the noise amplitude and the minimum frequency is set to 0.3 Hz. Thus, for the generalized inversion the spectral amplitudes will be analyzed at 40 frequency points equally spaced on a logarithmic scale over the range 0.3-30 Hz.

We only keep events recorded by at least three stations and stations that recorded at least three events and to satisfy the far-field approximation, we imposed a minimum distance to rupture of 15 km. After the application of these selection criteria, the final dataset is composed of 1950 six-component records from 255 earthquakes recorded at 162 sites. Figure 28 shows the magnitude versus distance distribution of the data.

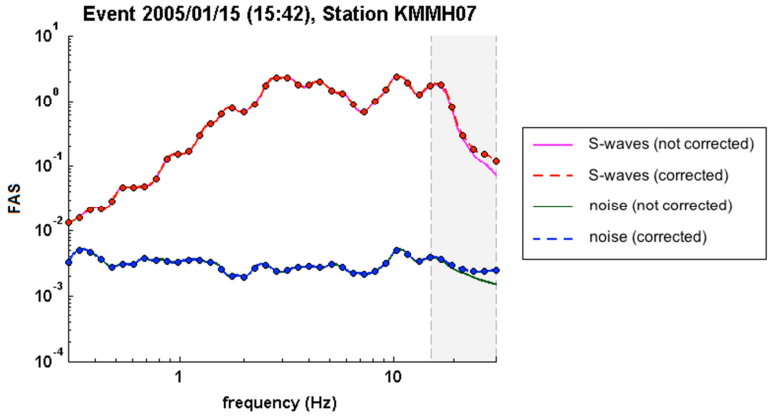

Figure 27 : Examples of S-waves and noise Fourier acceleration spectra (FAS) corrected by deconvolution of the instrumental response. For the inversion, the corrected spectral amplitudes are analyzed between 0.3 and 30 Hz at 40 frequency points (dots) equally spaced on a logarithmic scale.

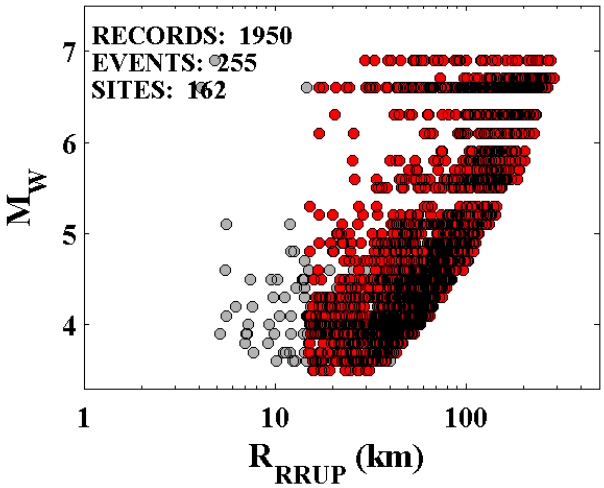

Figure 28: Distribution of the moment magnitude (M_w) and the rupture distance (R_{RUP}) of the KiK-net records selected for the generalized inversion (in red). The grey dots correspond to the initial data not selected.

4.4.3 Input parameters and reference conditions

In the inversion process, a system of equations must be solved where the unknowns are the seismic moment m_{0i} and the corner frequency f_{ci} for each event i, the site term s_{jk} for each station j and each frequency k and the attenuation parameters Q_0, α and γ.

In order to build the starting model, we give as input the three attenuation parameters. We set $\gamma = 1$ and the two others parameters are found in the literature. In Oth *et al.* (2011b), for the study of attenuation characteristics, the Japan is divided into five polygons following the attenuation tomography results of Pei *et al.* (2009). The Q models derived for the different polygons from Oth *et al.* (2011b) for crustal events are summarized in Table 2. As we want to use the same quality factor for the whole dataset, we take the mean value over the four polygons corresponding to crustal earthquakes: $Q_0 = 81$ and $\alpha = 0.71$.

Table 2: Q models derived byOth et al. 2011 for the Japanese territory divided into different polygons from dataset considering crustal events.

	Q_0	α
polygon 1 crustal	91	0.64
polygon 2 crustal	127	0.61
polygon 3 crustal	55	0.77
polygon 4 crustal	51	0.82
mean	81	0.71

Moreover, a reference conditions needs to be applied, either on the seismic moments or on site effects, which are the two constant parameters that control the amplitude of the spectrum (Andrews 1986; Field & Jacob 1995). As in Oth *et al.* (2010) and Drouet *et al.* (2011), we choose to impose the moment magnitudes of the largest events ($M_w \geq 5.5$ from F-Net catalog).

[20]

$$m_{0_{reference\ event}} = log_{10}\left[M_{0_{reference\ event}} \times \left(\frac{2R_{\theta\phi}}{4\pi\beta^3}\right)\right]$$

The standard deviation we impose on this reference condition is small enough to ensure that the moment magnitudes for the largest events will remain fixed during the inversion, allowing a robust estimation of the corner frequencies of these events. Indeed, the lower frequency limit (0.3 Hz) for the inversion, is likely above the corner frequency of the largest events and inverting for both moment magnitude and corner frequency for those events would results in a large bias. This reference condition thus allows overcoming this problem.

As in Drouet *et al.* (2008, 2010), we also impose that the average of the logarithms of the site effects at each frequency k over a list of j reference stations is null:

[21]

$$\sum_{j\ in\ list\ of\ reference\ stations} \log_{10}(S_{jk}) = 0$$

To define the list of reference stations, we use in a first inversion step all the surface stations within the reference list. From the obtained results (Figure 29), the stations showing the lowest and the flattest amplitudes are identified as rock sites and kept in the final list of reference stations. In our study, the stations that are kept are MYZH09 and NGNH22, which have respectively v_{S30} values equal to 973 m/s and 939 m/s estimated from S-waves velocity profiles. The final inversion is then performed for surface and borehole stations using these two surface stations as reference. Additionally, in order to compute "normalized" site effects, we normalize all the input spectra to a common generic amplification spectrum computed from a generic rock velocity profile with depth (Boore & Joyner 1997; Cotton *et al.* 2006) with a v_{S30} of 950 m/s as reference rock site in our case.

At this step, it is important to note that the site terms of a GIT have to be consider as relative transfer function: relative to the transfer function "postulated" for reference stations (here, a generic rock velocity profile v_{S30} = 950 m/s), with a forcing of the site effect of the chosen reference stations to be equal to one in mean (see relation [21]).

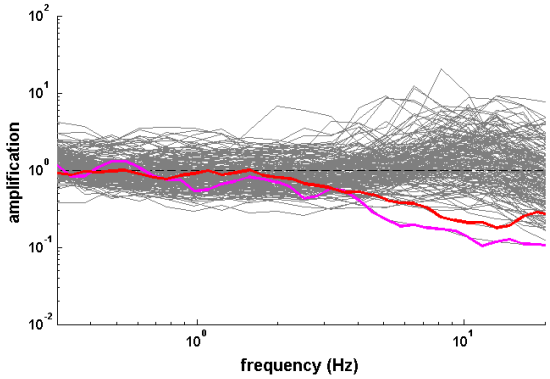

Figure 29: Site transfer function (grey lines) obtained for all surface stations in a first step of inversion. MYZH09 (in magenta) and NGNH22 (in red) are kept as reference stations.

4.4.4 Inverted attenuation coefficients and source parameters

Although the focus of this study is the inverted site transfer functions, the generalized inversion process has simultaneously determined source, path and site terms. We first present the results of inverted attenuation coefficients and source parameters, in order to ensure that this part of the spectral inversion has been well constrained.

The inverted attenuation parameters are: $Q_0 = 346$, $\alpha = 0.31$ and $\gamma = 1.1$. These parameters can be compared with the Q_0 and α values in Table 2 from Oth *et al.* (2011b). In Figure 30, the different attenuation models $\exp\left(-\frac{\pi f R_{ij}}{Q_0 f^{\alpha} v_S}\right)/R_{ij}^{\gamma}$, are plotted against distance for four different frequencies: 1 Hz, 5 Hz, 10 Hz and 30 Hz. The amplitudes are scales to get amplitude equal to 1 at 40 km. We can see that at short distances the amplitudes corresponding to the inverted attenuation model underestimate the models of Oth *et al.* (2011b), whereas at distances larger than 80 km the inverted model overestimates the others, except for $f = 30\ Hz$ for which all the models are almost the same. This is an important point on which we will discuss later.

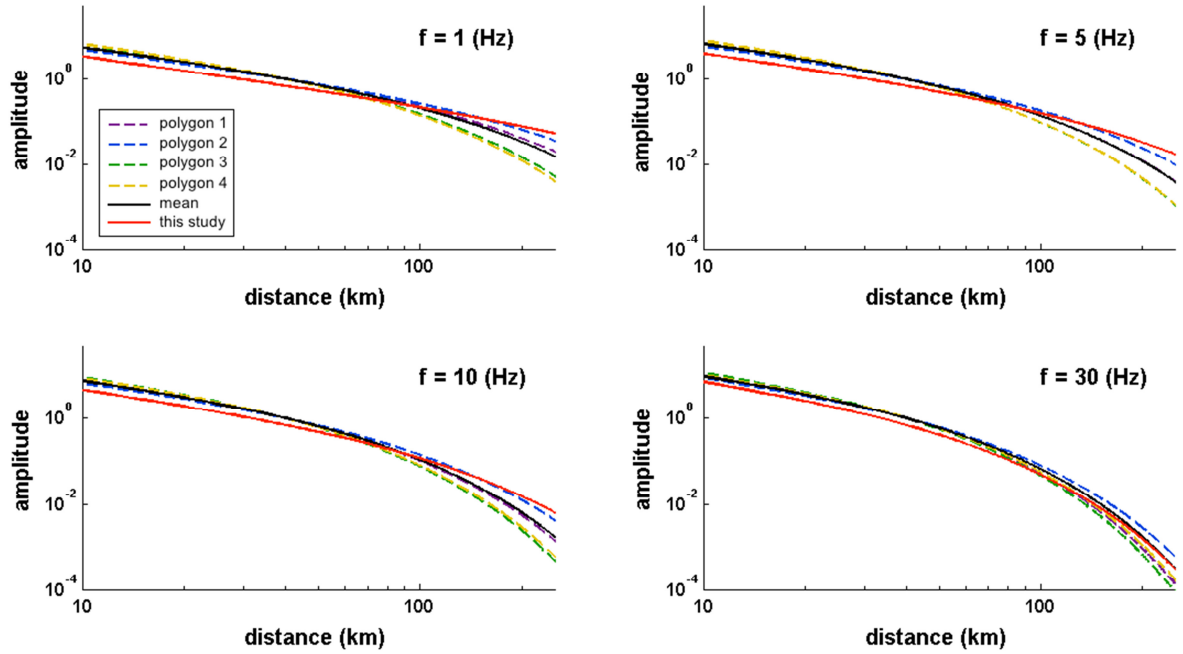

Figure 30: Attenuation models from Q_0 and α values in Table 2 and inverted attenuation parameters, for four different frequencies (1 Hz, 5 Hz, 10 Hz and 30 Hz) as a function of distance. Dotted lines correspond to attenuation models from Oth et al. (2011b) for polygons 1 to 4, black continuous line is the starting model for the inversion corresponding to the mean over polygons 1-4 and red continuous line is the inverted model.

Inverted source parameters such as moment magnitudes and corner frequencies can also be analyzed. For each earthquake of the database, moment magnitudes M_w are determined from the inverted seismic moments M_0 using the Hanks & Kanamori (1979) relationship:

[22]

$$M_w = \frac{log_{10}(M_0) - 9.1}{1.5}$$

These moment magnitudes are compared to moment magnitudes given by the F-Net catalog (Figure 31.a). We remind that we have imposed the moment magnitudes of events with $M_w \geq 5.5$. The inverted moment magnitudes show a good agreement with the F-Net moment magnitudes in the whole magnitude range, closed to the one-to-one relationship. On Figure 31.b the inverted corner frequencies f_c are represented as a function of moment magnitudes, with the theoretical relationship for three different constant stress drop values: 0.1, 1 and 10 MPa (1, 10 and 100 bars). We can see that values almost lie between 1 and 10 MPa and corner frequencies for large magnitude events are fairly well estimated even for frequencies below the lower frequency limit (0.3 Hz) of the inversion.

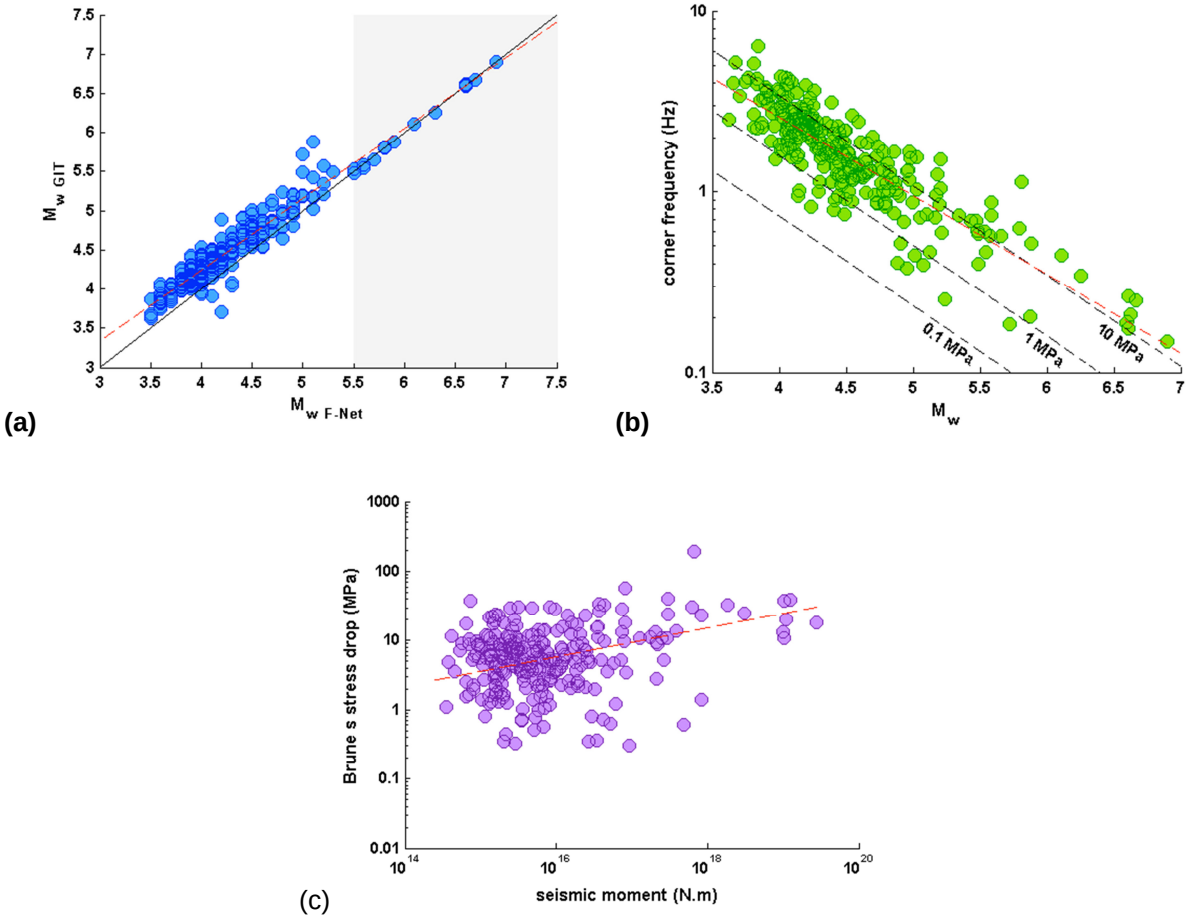

Figure 31: Inverted source parameters. (a) Moment magnitudes ($\boldsymbol{M_{w\,GIT}}$) as a function of moment magnitudes ($\boldsymbol{M_{w\,F-NET}}$) given by the F-Net catalog. The regression is denoted by the red dashed line and the black continuous line corresponds to the one-to-one relationship. The shaded grey area shows the limit of imposed magnitude ($\boldsymbol{M_w \geq 5.5}$) as reference condition; (b) Corner frequencies $\boldsymbol{f_c}$ as a function of moment magnitudes $\boldsymbol{M_w}$. The linear regression is denoted by the red dashed line and lines of constant stress drop of 0.1, 1 and 10 MPa are indicated; (c) Brune's stress drops $\boldsymbol{\Delta\sigma}$ as a function of seismic moments $\boldsymbol{M_0}$. The linear regression is denoted by the red dashed line.

Moreover, from the inverted seismic moments and corner frequencies, we also compute Brune's stress drops using the Brune (1970) relationship:

[23]

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

Brune's stress drops as a function of seismic moments are represented in Figure 31.c. The linear regression shows an increase of stress drops with seismic moments. Some authors also observe an increase of static stress drop with magnitude (*e.g.* Edwards *et al.* 2008; Drouet *et al.* 2010, 2011). However the scaling of stress drop with magnitude is still an open debate (*e.g*. Ide *et al.* 2003; Allmann & Shearer 2009; Baltay *et al.* 2011). Here, the scarcity of large events in the database does not allow giving a clear conclusion about this controversial issue.

4.4.5 Inverted site transfer functions

We determine the site transfer functions for all surface and borehole stations. These site transfer functions give the level of amplification at different frequencies with respect to the response of a common generic rock site with $v_{S30} = 950\,(m/s)$ as the reference. As shown by previous studies, the site amplifications are the most stable parameters coming out of such generalized inversion, because they are less sensitive to the trade-off between parameters (Field & Jacob 1995; Drouet *et al.* 2008). We also compute the ratios between surface and borehole transfer functions for each station. Figure 32.a shows example of the station OKYH02 ($v_{S30} = 1050\,m/s$) which presents a relatively flat transfer function with low amplitudes close to one and can thus be considered as a rock reference station. On the contrary, Figure 32.b shows an example of a site (AICH19, $v_{S30} = 587\,m/s$) characterized by site-effects with amplifications greater than 10 in a broad frequency range around 10 Hz.

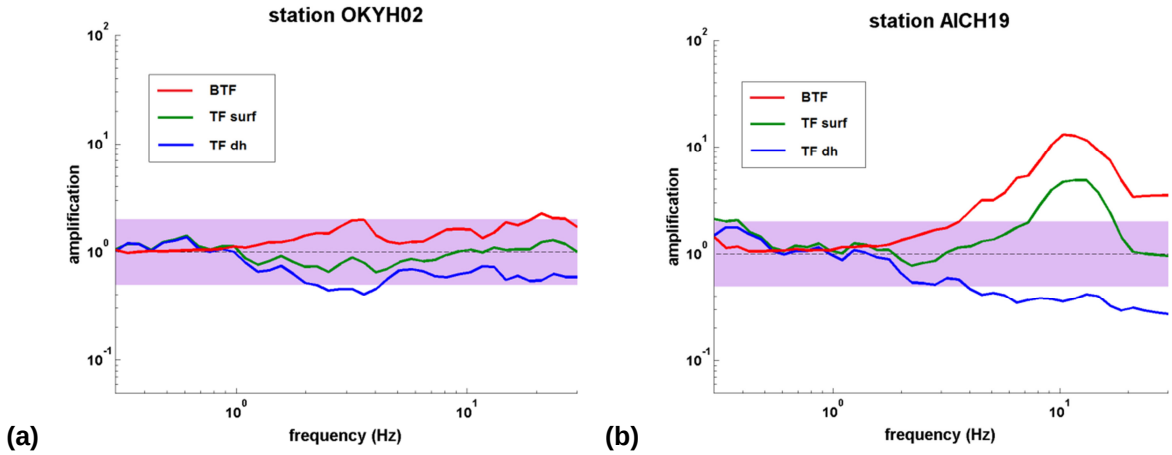

Figure 32. Examples of site transfer functions obtained from generalized inversion at surface (green lines) and borehole (blue lines) stations for (a) OKYH02 and (b) ISKH05. Ratios between surface and borehole transfer functions are represented with red lines. The purple patch corresponds to amplification levels between 0.5 and 2 and allows highlighting the amplification.

4.4.6 The mean spectral ratios from inverted transfer functions

Figure 33 shows the mean spectral ratios obtained from the inverted site transfer functions. We can observe a large amplification around 10 Hz associated with a large variability. This amplification is also observed on the mean transfer functions at the surface. For the mean borehole transfer functions, we obtain much lower amplitudes.

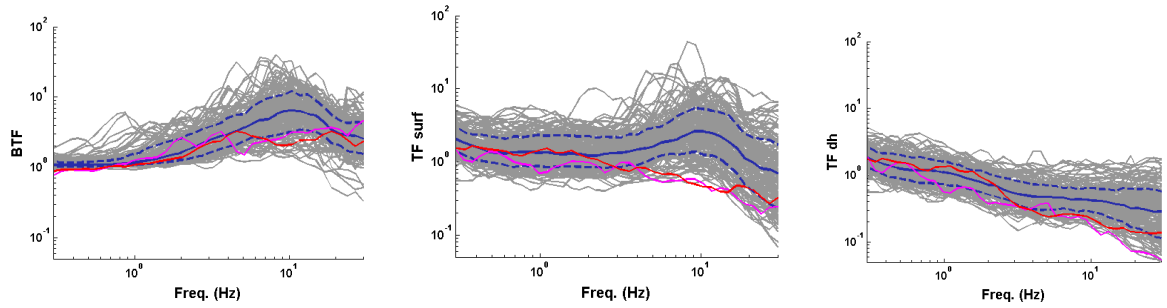

Figure 33: Transfer functions obtained at the 162 sites used for the generalized inversion (in grey). The mean spectral ratios are represented by the blue solid lines and the mean more or less one standard deviation by the blue dashed lines.

When we represent the mean site transfer functions for different V_{S30} ranges (Figure 34), as obtained from empirical and theoretical results, we can observe that the peaks have decreasing amplitude with increasing V_{S30} values and the frequency of the peaks increases with increasing V_{S30} values.

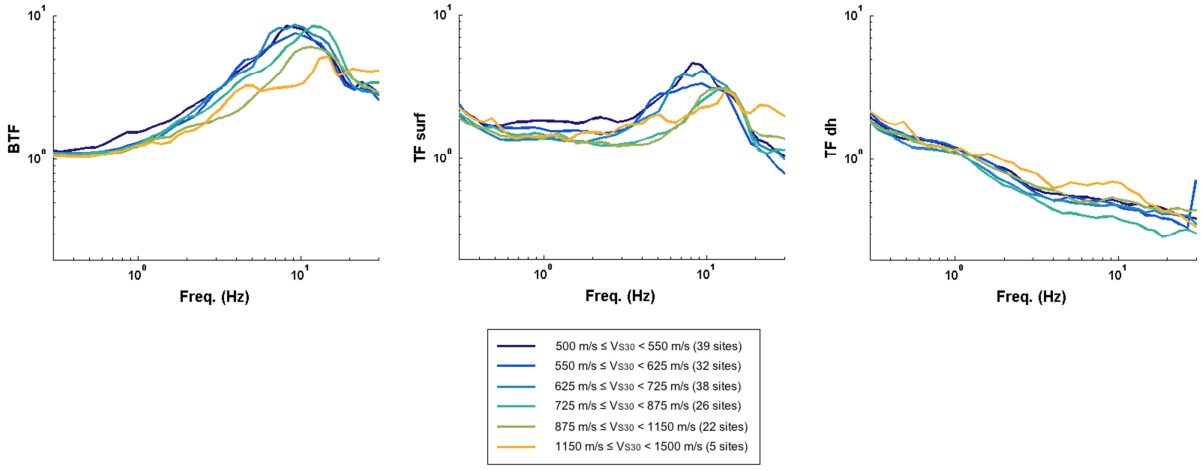

Figure 34: Mean inverted site transfer functions for different V_{S30} range.

4.5 COMPARISON OF THE DIFFERENT SITE TRANSFER FUNCTIONS

4.5.1 Comparison between empirical, theoretical and inverted site transfer functions

For each site we can compare the different transfer functions obtained from the empirical, theoretical and generalized inversion methods. Figure 35 and Figure 36 present examples of comparison for stations MYGH06 and NGNH35.

On this figures, only the green lines (ratio between surface and downhole data / simulations / site terms from GIT) are comparable between the three approaches. Indeed, the orange and purple lines for empirical are HVSR curves. Moreover, considering the same color lines for site terms from GIT, we remind that this information is relative to chosen reference stations and is not an absolute one (see 4.4.3). Moreover, site terms from GIT aim to characterize the “Station site effect”, whereas transfer functions from 1D simulation aim to characterize the “Local Site Effect” (see 1.1 and Figure 1).

The match between empirical and inversion ratios (green lines) are good in both examples. Considering the match with simulation, it is also quite good on the case of NGNH35 (with a slight shift in the fundamental frequency). For NIGH06, the ratio from 1D simulation produces much high amplification at high frequency. This will be an important issue later in our study.

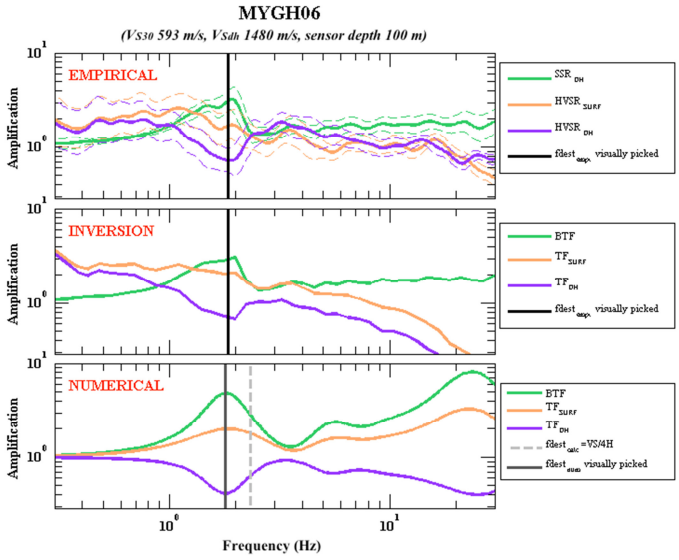

Figure 35: Comparison of the standard spectral ratios (green lines) and the site transfer functions obtained at surface (orange lines) and at depth (purple lines) stations from empirical, generalized inversion and theoretical methods for MYGH06 site.

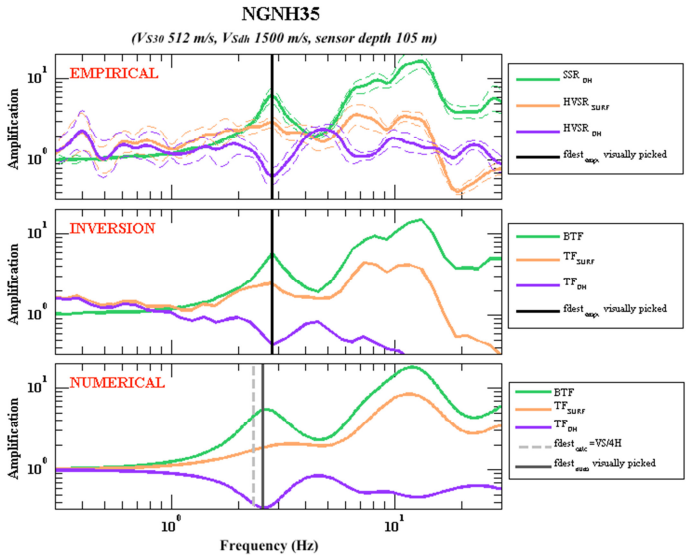

Figure 36: Comparison of the standard spectral ratios (green lines) and the site transfer functions obtained at surface (orange lines) and downhole (purple lines) stations from empirical, generalized inversion and numerical methods for NGNH35 site.

4.5.2 Selection of 1D sites

The discrepancy between empirical SSRs and ratios from 1D simulation could be explained by at least two issue:

1. The local site effect is not 1D but 2D or 3D, the 1D simulation is no more applicable,
2. The provided velocity profile of the station is not accurate.

Since we wish to test the possibility to use 1D velocity profile with our correction approach and since the KiK-net give us the possibility of checking the "1D assumption" through the use of empirical SSR (and implicitly the quality of the provided V_S profile), we will here continue our work only on station where the "1D assumption" remains acceptable.

In the way of Thompson et al. (2012), Pearson's correlation coefficient (r) is computed between [*max*(0.5 Hz, 0.5 x f_{dest}) *min*(15 Hz, 7 x f_{dest})], in which f_{dest} is the frequency for which destructive interferences are maximal.

Figure 37 displays this correlation coefficient with respect of V_{S30}. Thompson et al. (2012) define 1D sites as the one with r larger than 0.6. Thus, we have selected 108 sites with approximately 1D-behaviour.

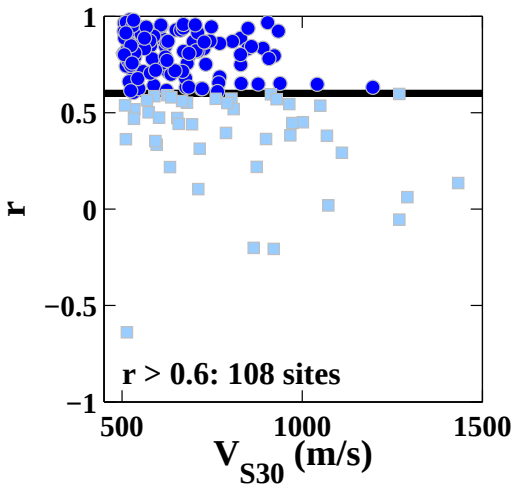

Figure 37: Pearson's correlation coefficient (r) versus V_{S30}.

4.5.3 Comparison of the mean spectral ratios obtained from the different approaches

After the selection of the 1D, we here compare the results from the three different approaches on the 108 selected "1D" sites, in terms of mean of surface to downhole ratios. The results are shown on Figure 38 with their associated standard deviation. Concerning the ratios deduced form 1D simulations, we tested here different "scaling" of the Q_S value from the V_S value (from V_S=Q_S/5 to V_S=Q_S/50).

We can see that global mean empirical and inverted spectral ratios are in very good agreement in the whole frequency range (even if one can observe that GIT provides a slightly lower mean than the empirical SSR method). Theoretical mean spectral ratios (from simulation) are in good agreement with the other methods for frequencies lower than 6 Hz. In the frequency range 6-15 Hz, numerical simulations

underestimate the amplification obtained from empirical and generalized inversion methods, the best fit is obtained for $Q_S=V_S/10$, for which, the underestimation remain moderate (around 10% at 10 Hz). Above 15 Hz, the situation is drastically different. Only the $Q_S=V_S/50$ seems to be acceptable in mean, whereas other scaling overestimate significantly other estimations. But this $Q_S=V_S/50$ produces also the worst standard deviation at this frequency.

Almost the same conclusions can be drawn from the comparison of the mean spectral ratios according to v_{S30} ranges (Figure 39). For v_{S30} values in the ranges 500-550 m/s and 550-625 m/s, a good agreement is obtained between methods, whereas for the ranges 625-725 m/s and 725-875 m/s the peak of amplification around 10 Hz is clearly underestimate by numerical simulations. The results obtained in the range 875-1500 m/s cannot be well interpreted because of the lack of data (only 9 sites).

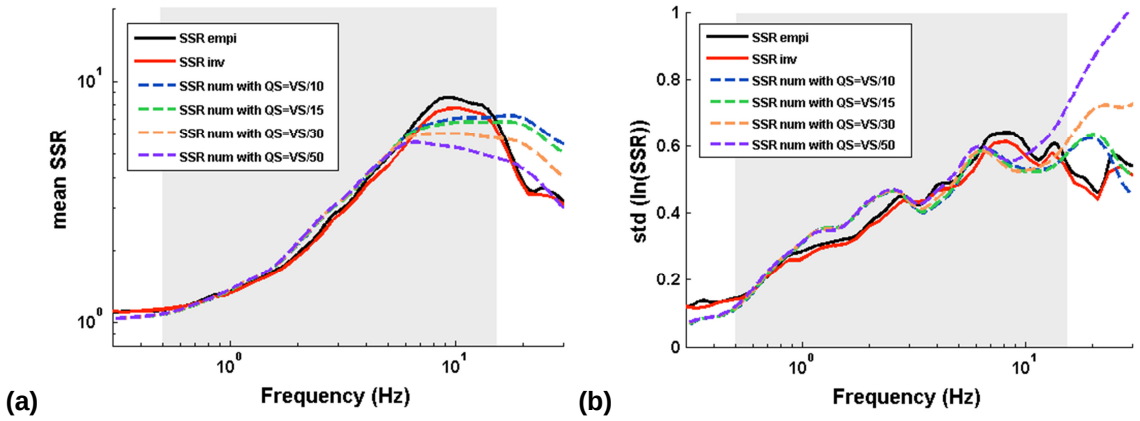

Figure 38: Comparison of (a) the global mean standard spectral ratios and (b) their associated standard deviations obtained from empirical (black line), generalized inversion (red line) and theoretical (dashed lines) approaches. The grey patches correspond to the frequency range (0.5-15 Hz) for which the different methods are comparable.

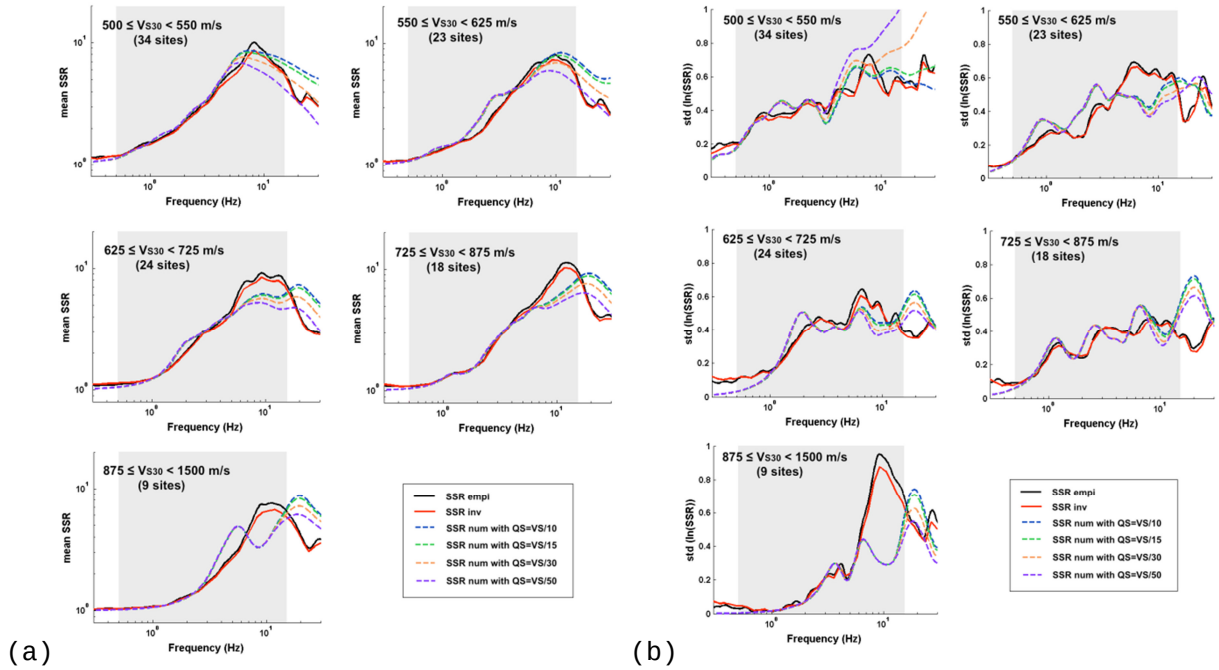

Figure 39: Comparison of (a) the mean spectral ratios according to v_{S30} ranges and (b) their associated standard deviations obtained from empirical (black line), generalized inversion (red line) and theoretical (dashed lines) approaches.

4.5.4 Partial conclusions

At the end of this section, one can formulate these partial conclusions:

In order the reproduce the site effect between the rock downhole sensor and the surface, the empirical approach (SSR) and the inversion approach (GIT) produced very similar an consistent results. This validates the robustness of GIT method for this purpose (the fact that site terms from GIT are relative ones is no more an issue here since we are dealing with ratios).

The “theoretical approach” (based on 1D simulation), when restricted to a choice of so called “1D stations” are also very similar to other method up to 6 Hz. Between 6 and 15 Hz, they provide a slightly underestimation of the site effect (in mean). Above 15 Hz, they could strongly overestimate the amplification.

The slight underestimation between 6 and 15Hz could be explained by the fact that even if we do perform a selection of “1D site”, this selection contained a given amount of “permissiveness” (in order to preserve a significant number of stations to continue the work) and that a part of 2D-3D site effect may be still remain in the database.

At high frequency, we think that the usual Q_S scaling is no more suitable. Assimaki et al. (2008) have defined by inversion the site parameters of some KiK-net sites, and especially $V_S(z)$ and $Q_S(z)$. For a given V_S, they found Q_S values lower than those defined with $Q_S = V_S/10$ for shallow layers. So the issue of attenuation have to be consider as a function of the depth and not only as a unique “scaling”. We will not solve this issue in the present work (this will a perspective for further work).

Thus, we will continue our work using the $Q_S=V_S/10$ keeping in mind this overestimation issue above 15 Hz.

5 GMPES DEVELOPMENT

In the context of the host-to-target issue, for the host region, the site is defined in terms of V_{S30} and for the target region, the site transfer function at different frequencies is considered. The current adjustment methods allow defining a rock motion by considering only corrections of GMPEs at least in terms of V_{S30} and if available in terms of κ_0 (κ_0 is often deduced from V_{S30}-κ_0 relationships).

The objective is to apply methods generally used in the target region to define rock motion from surface records corrected of the site effects or from downhole records corrected of the depth effects. From these datasets, simple ground motion prediction equations are developed and the results will be compared both with natural data and with classical adjustment method.

Different simple ground motion prediction equations are developed for different datasets to better understand the reference motion behavior.

The first datasets are based on real data without any corrections and it is called **DATA**. These models could be thus comparing with other GMPEs (e.g., Cotton et al. 2008; Rodriguez et al. 2011). The second type of datasets is based on downhole records corrected of the depth effects (correction of the hole in the downhole transfer function and consideration of the free surface effects) and it is called **DHcor**. The Japanese network offers the opportunity to have for each site two sensors, one at surface and one at depth. However, it is not the case for the most other networks, for which only surface data are recorded. This is why the third type of datasets is based on surface records corrected of the theoretical incident transfer function of each 1D-sites. This correction shall allow to get average motions similar to those observed with DHcor. This correction method is classically used in target region. This type of dataset is called **SURFcor**. The last model is obtained from the host-to-target methodology and it is called **H2T**. In this case, the GMPE obtained from surface records and defined for V_{S30}=800 m/s is adjusted to a hard rock characterized by a V_S around 2400 m/s with the adjustment factor of Van Houtte et al. (2011). This adjustment factor is defined for the κ_0 values from the V_{S30}-κ_0 relationships defined also by Van Houtte et al. (2011) from worldwide data. Thus, for V_{S30}=800 m/s, κ_0 is between 0.02 and 0.05 s and for V_{S30}=2400 m/s, κ_0 is between 0.002 and 0.012 s. This model could be compared with the others datasets only for V_{S30}=2400 m/s.

The different datasets are the following and are illustrated on Figure 40. The models 3, 5 and 7 will be presented only in Annexes (Figure 58 to Figure 60) in order to simplify the understanding in this section.

A. DATA:

1. *DATA_surf: Natural surface records*
2. *DATA_dh: Natural downhole records*
3. *DATA_mixed: Combination of natural surface and downhole records*

B. DHcor:

4. *DHcor: Corrected downhole records*
5. *DHcor_mixed: Combination of natural surface records and corrected downhole records*

C. SURFcor:

6. *SURFcor: Corrected surface records*
7. *SURFcor_mixed: Combination of natural surface records and corrected surface records*

D. H2T:

8. *Adjustment of the GMPE obtained from natural surface records with the adjustment factor of Van Houtte et al. (2011).*

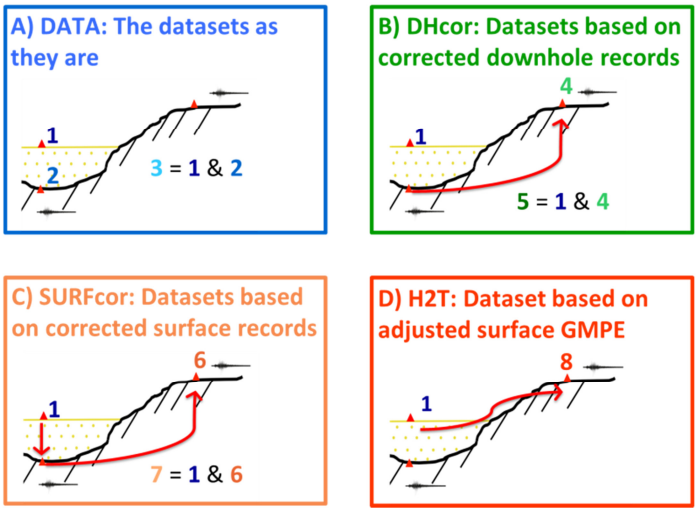

Figure 40: Schemes of the different datasets used to develop reference motion GMPEs. The coloured numbers will be used to characterize each dataset in the next figures of this section.

5.1 IMPLEMENTATION OF THE ROCK MOTION DATASETS

To build the reference motion datasets, the processing applied is similar to the one explained in the section 4.1. The only difference is that the signal-to-noise ratio is estimated from a signal window containing the whole signal, and not only the S-waves. A record is selected if SNR>3 for a continuous frequency band between 0.5 and 15 Hz for the two horizontal components. Then, the acceleration response spectra are computed as the geometrical mean of the two horizontal components.

In addition, only events with $M_W \geq 4.5$ are selected to develop new GMPEs. Actually, the goal of this last selection criterion is to avoid the magnitude-scaling problem (e.g., Cotton et al., 2008).

For each new dataset, DHcor and SURFcor, others processing are applied and they are described in the following subsections.

5.1.1 Datasets based on corrected downhole records (DHcor)

The downhole records have the advantage to represent the characteristics of sites with large V_S properties (see Figure 9). However, these records are affected by the downgoing waves and do not take into account the free surface effects. Cadet et al. (2012) have developed a correction factor to avoid these effects. This correction factor applies to the acceleration response spectrum in dimensionless frequency space defined according to the destructive frequency ($f_{adimensionless} = f / f_{dest}$). This correction factor can be applied to the downhole records, under the assumption that these sites at depth have a relatively 1D behavior. That's why the destructive frequency is picked directly on the empirical ratios (HVSR_dh and SSR) to avoid the biases that may be related to a shear-wave velocity profile not well defined (see Figure 41). Only if the destructive frequency is not clear on the empirical ratios, it is picked on the theoretical transfer function at depth (TF_dh). Finally, for 138 records, f_{dest} is from empirical ratios and for only 14 records is from theoretical one. In addition, one site with f_{dest} lower than 0.5 Hz is removed among them (GNMH05 with f_{dest_theo}=0.1 Hz).

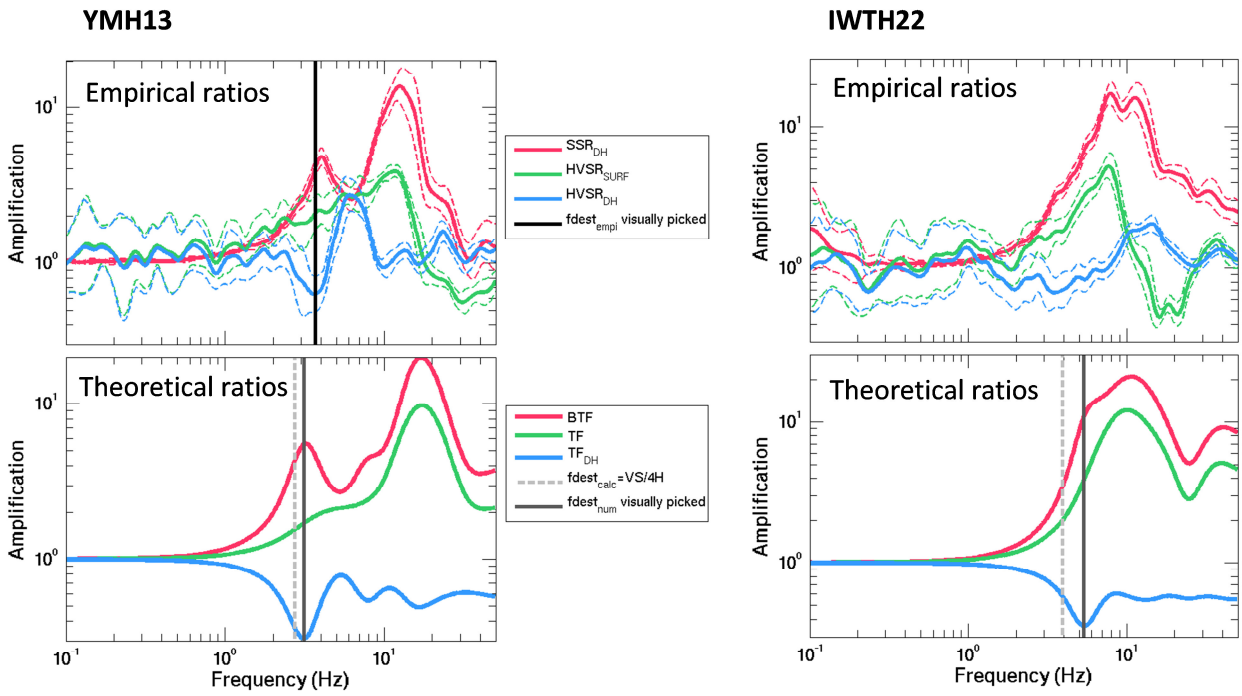

Figure 41: Illustration of the destructive frequency picking on the spectral ratios from 2 examples: for YMTH13, the destructive frequency is picked on the empirical HVSR_dh and for IWTH22; f_{dest} is picked on the theoretical ratio TH_dh. The theoretical destructive frequency is also reported (f_{dest_cal}= VS/4H, in dashed grey lines).

Globally, the destructive frequency f_{dest} picked on the empirical ratios is relatively similar than the one picked on the numerical ratios, as shown on Figure 42. Indeed, the linear regression is relatively similar to unitary line (X=Y) and the correlation coefficient is 0.84. However for some cases, we can observe a difference of several Hz. These differences could be explained especially by a not well defined velocity profile.

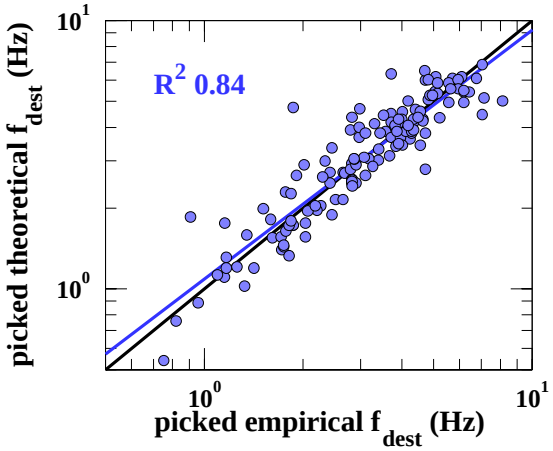

Figure 42: Comparison of f_{dest} picked on the empirical ratios versus the one picked on the theoretical ratio (137 sites).

Finally, 1040 records are selected to build a new dataset from corrected downhole records and to develop a GMPE (see Figure 43). For the natural data (**DATA**), the same dataset is used.

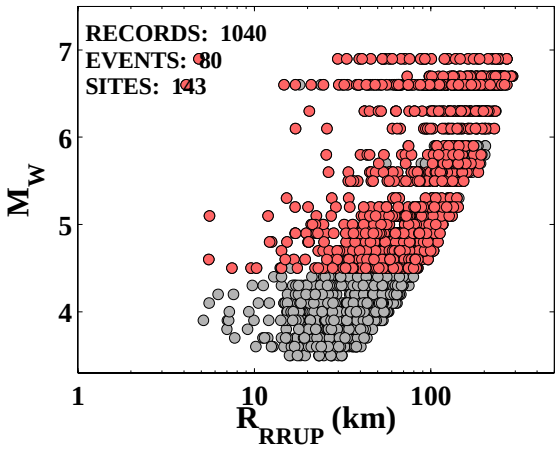

Figure 43: Distribution of the moment magnitude (M_W) and the rupture distance (R_{RUP}) of the KiK-net records selected to develop a GMPE from corrected downhole records (red dots). These records have a continuous signal-to-noise ratio larger than 3 between 0.5 and 15 Hz (in this case, the signal is the entire signal window). Only events with $M_W \geq 4.5$ and only sites with $f_{dest} \geq 0.5$ Hz are selected. The grey dots correspond to the initial data not selected.

5.1.2 Datasets based on corrected surface records (SURFcor)

In low-to-moderate seismicity region, in most cases, only surface data are available and few events are recorded. Therefore it is interesting to develop a reference motion dataset obtained after correction of the surface records. Indeed, the Japanese database with their double sensors offers the opportunity to compare the results from surface and downhole records. To correct this data, the theoretical amplification factor in terms of response spectra is computed for each 1D site with the methodology presented at the section 4.3. This theoretical amplification factor is defined as the ratio between an incident response spectra obtained from theoretical SH 1D simulation and an input response spectra used as input in the simulation. The objective is to correct the surface records of the incident theoretical response. This theoretical correction will be applied in terms of response spectra to the surface records.

For each 1D site, the theoretical amplification factors are computed as presented in the section 4.3.1 and with the following specific input parameters:

- 15 sample input motions, from RESORCE European databank, chosen for their wide range of frequency content (see Figure 44)
- Range of incidence angles (0 to 45°)
- $Q_S = V_S / X_Q$ with X_Q =[5, 10, 15, 25, 35, 50, 100]

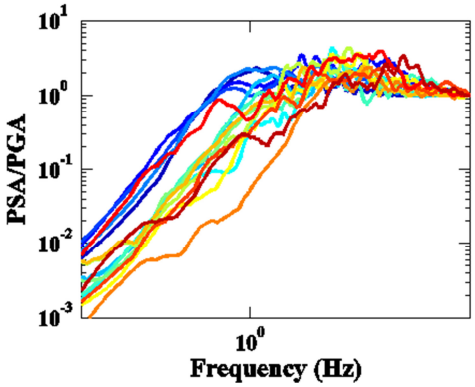

Figure 44: Illustration of 15 response spectra selected as input motions for theoretical computation. These spectra are normalized according to the PGA. The colours are defined according to the frequency of peak, from blue to red, i.e. from a low frequency peak to a frequency peak at high frequency. The 15 input motions come from RESORCE European databank and they are chosen for their wide range of frequency content.

Figure 45 illustrates the amplification factor computation for one example: the station NGNH35 which is a station with a high correlation coefficient between theoretical and empirical standard spectral ratios. HVSR_surf of this station is characterized by a first peak around 3 Hz with an amplitude of 3, and then by a double peaks at 6.6 Hz and 11.6 Hz with an amplitude of 3.5. On SSR, the amplitude of this last peak reaches 16. The amplification factor shows also a large amplification at this last frequency. However, the amplitude is clearly dependant of the frequency content of the input motions. In the case of this station with a site response at high frequency, the amplification factor is larger for the input motions having a peak at high frequency. Around this peak the amplification factor presents a large variability due to input motions. The standard deviations associated with the three different input parameters are presented on Figure 46. The variability due to different range of the incident angle is very low compared to the one due the input motions and the Q_S definitions. This variability increases around 3 Hz and stays larger for the high frequency band in the case of input motions, while in the case of the Q_S definitions, mainly the site amplification peak have a large variability. This is why in the following we have chosen not to vary the Q_S definition and to fix X_Q=10. Besides, this Q_S definition (Q_S=V_S/10) is largely used in simulation (Maufroy et al. 2015) but the formula is not necessarily adequate. Assimaki et al. (2008) have defined by inversion the site parameters of some KiK-net sites, and especially V_S(z) and Q_S(z). For a given V_S, they found Q_S values lower than those defined with Q_S = V_S / 10. This issue has not been addressed in this report but it is an important issue that could be addressed later.

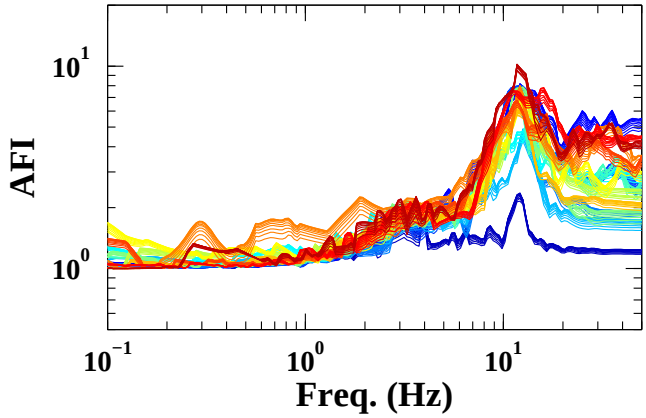

Figure 45: Illustration of the amplification factors obtained for the station NGNH35 for the 15 input motions and different range of incident angles. In this case, $Q_S=V_S/10$.

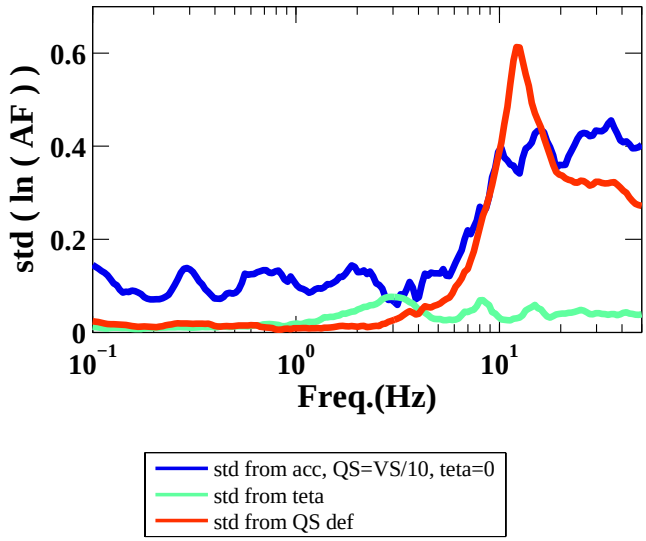

Figure 46: Illustration of the standard deviation associated with the amplification factor for the station NGNH35 due to three input parameters: in blue, due to input motions; in green, due to the range of incidence angles; and in red, due to the QS definition.

Figure 47 displays the mean and the standard deviation computed for the 108 1D-sites when the input motion and the incident angle vary. The variability is larger for the amplification factors having the larger amplitude (in red). In addition, the amplification factors with the peak at lower frequencies (in blue) have a larger variability up to 15 Hz.

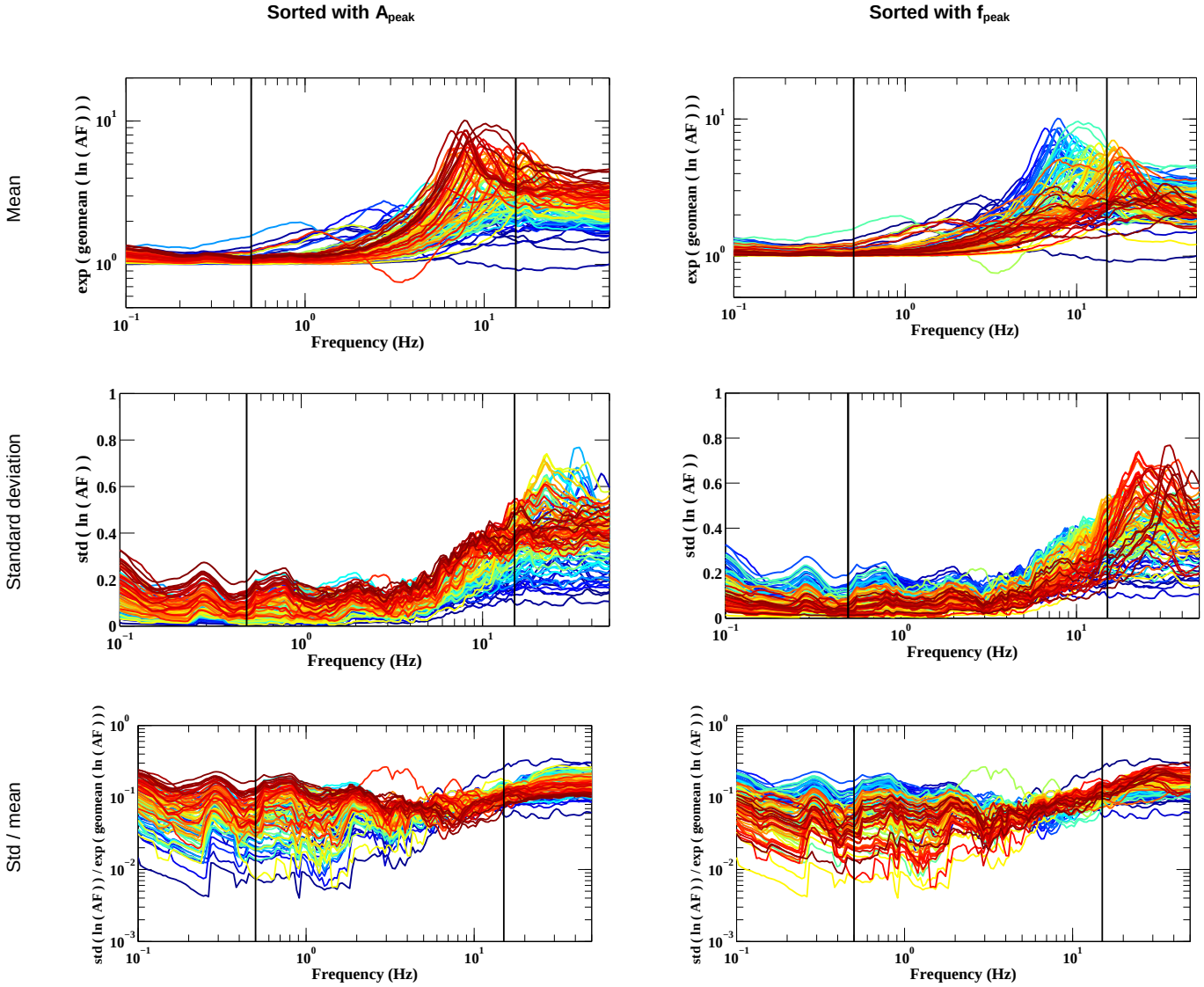

Figure 47: Analysis of the amplification factors computed for all the 1D sites (108) and sorted both according to the amplitude of the peak (A_{peak} on the left) and to the frequency of the peak (f_{peak} on the right). In this case, the mean is computed for the 15 input motions and the different range of incidence angles. The vertical lines represent the frequency limit for the empirical data.

Then, the records associated with 1D-sites, with a PGA lower than 0.1 g, and associated to a site and an event having recorded at least three records are selected. Thus, 704 records are selected (see Figure 48). Finally, each surface record is corrected of its site amplification factor in terms of response spectra. For these corrected records, we assume that the shear-wave velocity at depth is their corresponding velocity property.

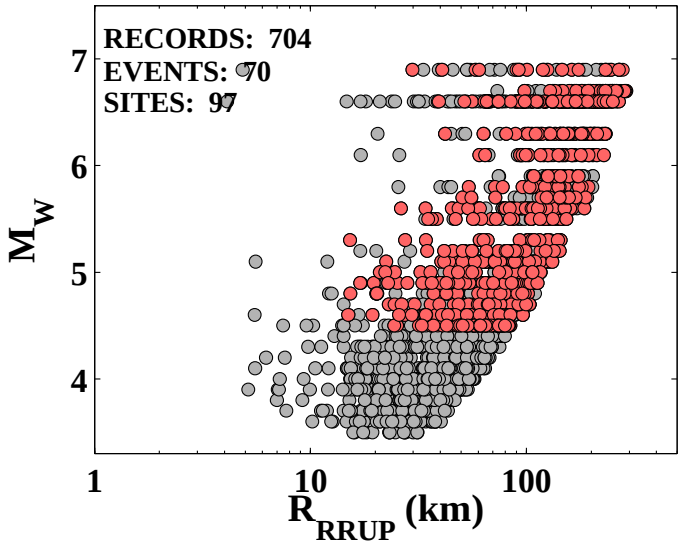

Figure 48: Distribution of the moment magnitude (M_W) and the rupture distance (R_{RUP}) of the KiK-net records selected to develop a GMPE from corrected surface records (red dots). These records have a continuous signal-to-noise ratio larger than 3 between 0.5 and 15 Hz (in this case, the signal is the entire signal window). Only events with $M_W \geq 4.5$ and only sites with 1D behaviour are selected. The grey dots correspond to the initial data not selected.

5.2 GMPEs

To analyze the behavior of the different datasets, we have chosen to develop ground motion prediction equations based on a simple functional form in order to get the coefficients directly. Indeed, the functional forms used in the literature are increasingly complex, requiring fixing some coefficients. These fixed coefficients are obtained beforehand either from theoretical simulation or empirically but from a well-defined subset. The regression algorithm of Abrahamson and Youngs (1992) is employed to develop GMPEs for the geometrical mean of the two horizontal components of SA in g. The following simple functional form is used:

[24] $$\ln(SA(T))_{es} = a_1(T) + a_2(T).M_W + a_3(T).M_W^2 + b_1(T).R_{RUP} - \ln(R_{RUP}) + c_1(T).\ln(V_S/1000) + \delta B_e(T) + \delta W_{es}(T)$$

where δB_e is the between-event and δW_{es} the within-event variability, associated with standard deviations τ and φ respectively, and $\sigma_{TOT}=(\tau^2+\varphi^2)^{0.5}$.

The regression coefficients obtained are given in Annexes (Table 3 to Table 9).

Figure 49 shows the evolution over the period of each coefficient and also of the variability terms for the different datasets. Between the datasets, the main differences are observed at short periods. b1, the coefficient of anelastic attenuation, is the more stable coefficient, but we observe some differences at short periods. SURFcor is the model with the more important attenuation and at the contrary DHcor is the lowest.

The site coefficient c1 is clearly negative at long periods. In this case, when V_S increases the amplification decreases. However, at high frequency, c1 is around zero and even positive for SURFcor and DATA_surf. In this case, when V_S increases, the amplification stays stable or even increases. For DATA_surf, c1 varies steeply between 0.09 and 0.15 s from -0.9 to 0.08. DATA_dh and DHcor records are relatively similar except than DHcor records are corrected of the depth effects. However, if c1 is relatively similar at short periods, it is different at long periods. Between 0.15 and 0.3 s, we observe the destructive

frequency correction effect, i.e. a hole for DATA_dh and at the contrary, a bump for DHcor. At long periods, DHcor is similar to SURFcor. These 2 methods should give similar results but if it is the case at long periods, at short periods we observe large differences.

Figure 50, showing predicted SA with respect of V_S, allows to better observe the differences between the models. DATA_surf predicts amplitudes larger than the other models for similar VS at periods lower than 0.15 s (6.67 Hz), and it is the contrary at long periods. DATA_prof is the model with the lower amplitudes because it is the only one model that does not include the free surface effect. At high frequency, especially at 12.5 Hz, the differences of the slope of DHcor and SURFcor are more visible. The red cross represents the amplitude obtained with H2T. This amplitude is clearly larger at periods lower than 0.1 s than the amplitude given by especially DHcor and DATA_surf.

The variability associated with the different datasets is mainly controlled by the intra-events variability because the inter-event variability is really stable between models. The intra-event variability is lower for the models based on downhole records: DHcor and DATA_dh. The intra-event variability is larger for the entire range of frequency in the case of DATA_surf. The intra-event variability is larger also around 0.1 s for SURFcor that is to say around the mean amplification peak.

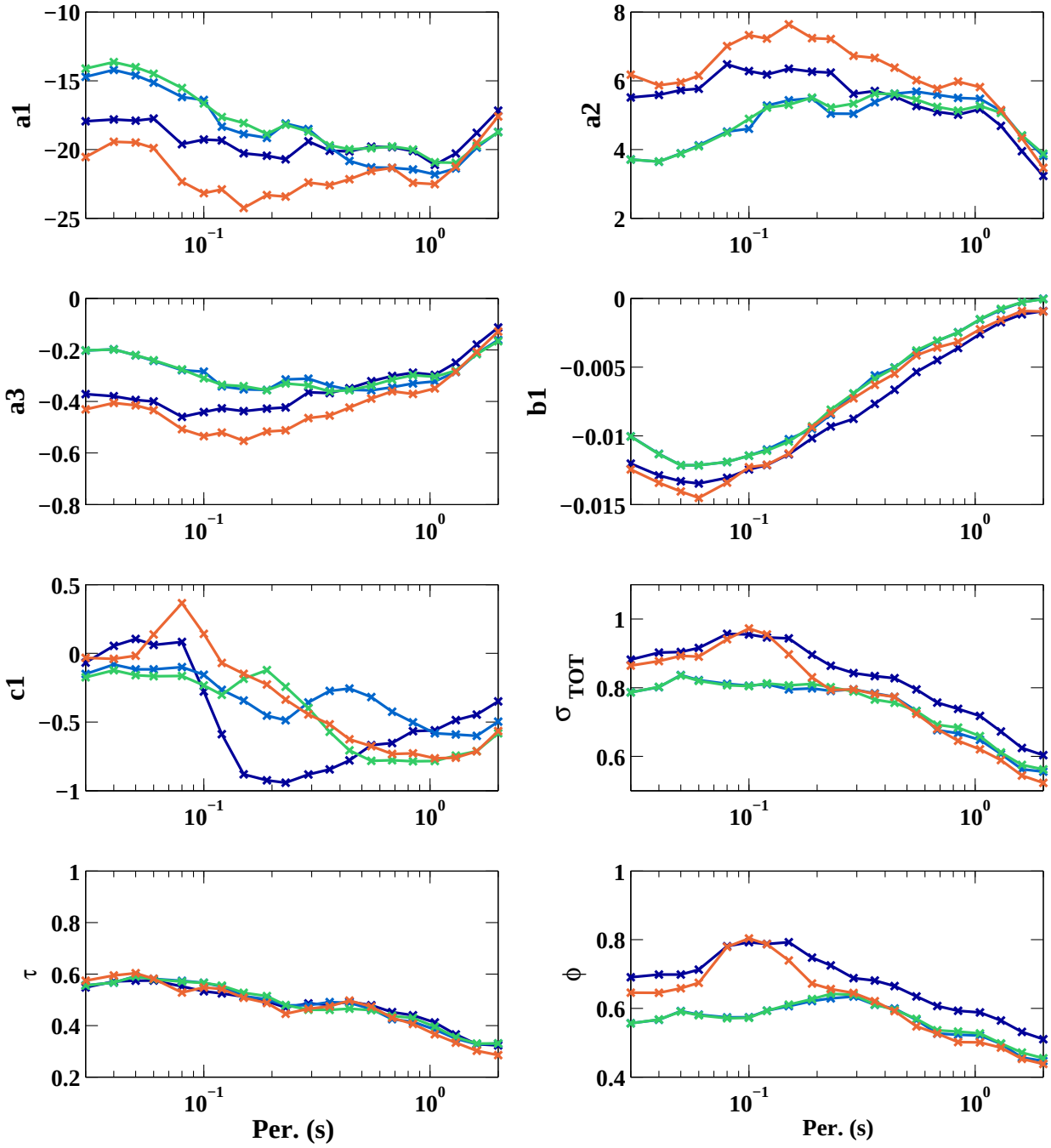

Figure 49: Regression coefficients obtained for DATA_surf (dark blue), DATA_dh (lght blue), DHcor (green) and SURFcor (brown) (the models are presented on Figure 40).

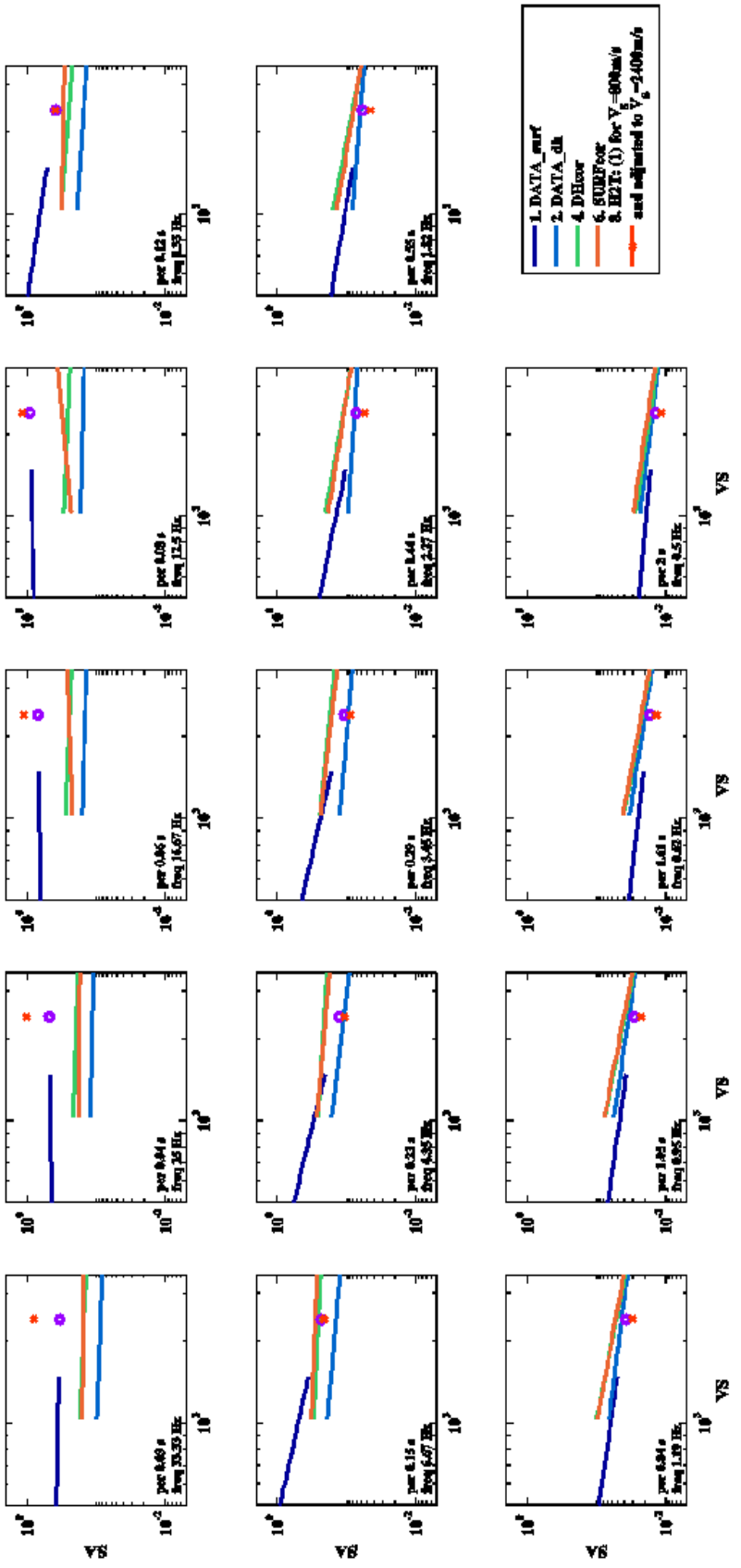

Figure 50: Predicted response spectra according to V_S for different periods and for the following scenario M_W 6.5 and R_{RUP} 20 km. The different models are presented on Figure 40. The purple point corresponds to the extrapolation of DATA_surf to 2400 m/s.

Figure 51 displays a comparison of DATA_surf, DATA_dh, DHcor, SURFcor and H2T models for three scenarios in terms of V_S and M_W 6.5 and R_{RUP} 20km. The first two scenarios correspond to the VS distribution medians of the surface and downhole records, respectively. The third scenario was defined so as to make a comparison with the host-to-target method, and especially in the case of the Van Houtte et al. (2011) adjustment factor is used.

For this specific scenario, DHcor and SURFcor predicted spectral acceleration medians are really close for each period. This supports the interest of using surface records corrected of their site transfer functions. However, as discussed previously (see Figure 49), the variability associated with these 2 models is larger at high frequency in the case of SURFcor. Around 10 Hz, the predicted spectral acceleration is 3 to 4 times lower in the case of DHcor and SURFcor than with H2T.

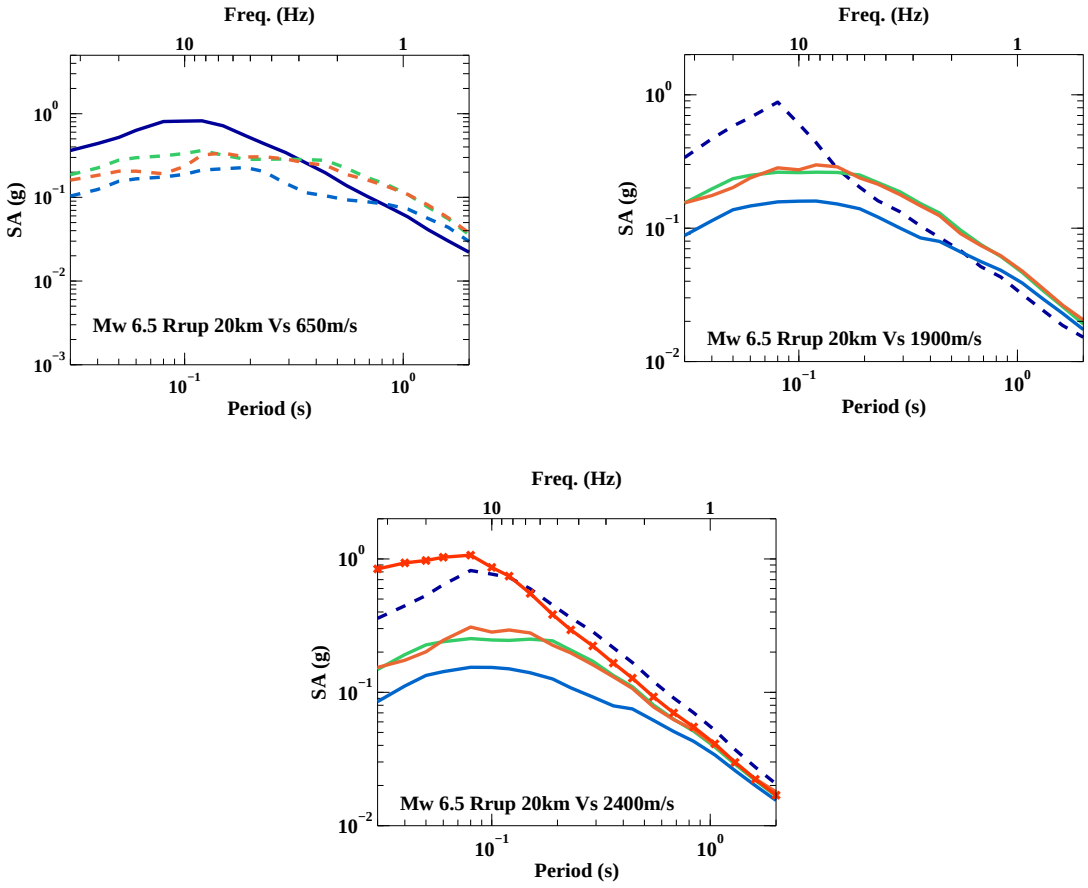

Figure 51: Comparison of the predicted SA obtained with the empirical models presented on Figure 40 for different V_S values and for a specific scenario M_W 6.5 and R_{RUP} 20 km. The models are presented with dashed lines when the V_S value is outside the range of V_S used to develop the model.

6 DISCUSSION

6.1 ESTIMATION OF THE "RESIDUAL SITE-EFFECT"

As discussed in introduction (1.1), site effects could be estimate through different approaches. A first group consists in the evaluation of the amplification between a given depth and surface (or between rock outcrop at surface and a given surface position within a basin). This group involves simulations, empirical approaches based on SSR evaluation. A second group aims to evaluate how a given site is different from the other ones. These approaches need a significant amount of earthquake records and distinguish features that are linked to the station, the regional attenuation and the sources. In the case of GIT, this corresponds to the "site terms". Another approach (not addressed in the present report) is the assessment of the "δS2S" terms within the GMPE derivation (Rodriguez-Marek et al. 2014).

From a theoretical point of view, these two groups do not investigate the site effect at the same scale. The first one addresses the site effect phenomenon at a typically hectometric scale (here, by reference to the depth at which the amplification evaluation is referring). For the second one, it is more difficult to know exactly the associate scale, but one can suggest intuitively that they could include features, linked to given station, that may affect ground motion in a deeper zone. In parallel, host-to-target adjustments usually refer to generic profiles expressed at a kilometric scale.

In order to distinguish this two "scales" of site effects, we proposed to talk about the "local site effect (LSE)" for the first group and to talk about "station site effect (SSE)" for the second one. In order to compare these two classes, we also introduce the concept of "residual site effect (RSE)" (see Figure 3). In the case of our study, we have the opportunity to discuss about this "residual site effect", with this corollary question: which is the part of site effect which is not taken into account locally, through the LSE?

The first step of this evaluation is the computation, for each station, of the transfer function using 1D simulation using the velocity profiles information provided by KiK-net station, using the scaling already discussed (4.3.1), applied on the subset of stations considered as "1D stations". This transfer functions are expressed as the ratio between the surface station with respect to an outcropping bedrock having the same characteristics (velocity, density, damping) of the deeper layer of the available profile (generally the velocities at the depth of the downhole sensor). This will be used to characterize the LSE and will be noted: LSE_{1D_SIM} (this term was also noted *TF* in previous part of this report).

The second step of this evaluation is to prepare site terms for GIT. It is worth here to remind two important features:

- Site GIT terms, as directly considered at the output of the inversion, are expressed as a relative amplification with respect to the mean amplification of few selected "reference sites" (here, MYZH09, NGNH22, see 4.4.3). Let's note this term: SSE_{GIT_REL} ("relative" station site effect).
- The mean amplification of this reference stations are forced to fit the amplification associated to a generic rock site of v_{S30} = 950 m/s using generic profile, as describe through equation [21]. Let's note GIT_{REF} this amplification postulated to be representative of the reference stations.

The "Residual Site effect" is hence expressed (in Fourier domain):

[25]

$$RSE = \frac{SSE_{GIT\_REL} \times GIT_{REF}}{LSE_{1D\_SIM}}$$

The result are given on Figure 52 (at this step, just consider the blue lines of this figure). The RSE computed for each station are the thin blue lines. The mean and associated standard deviation are shown with thick solid and thick dashed blue lines. We can see that the RSE is almost flat in the frequency range 0.5-12 Hz with amplitude equal to approximately 1.6. Above 12 Hz, the mean decreases down to 0.7, but we know that in this frequency range, le 1D simulation overestimate strongly the LSE_{1D_SIM} term due to Q_S scaling issue. We will not comment here more in details this RSE behavior above 12 Hz.

The 2 value of the RSE within the valid frequency range could, at a first glance, lead to the conclusion that there is still a amplification due to site effect “below” de LSE zone. But before formulating this conclusion, we should verify the possibility that a part of the “true” site terms “was not moving” to attenuation terms or source terms within the inversion process. Indeed, the assumption postulated to impose to the reference station site terms (equation [21]) is very strong and may impact the results of site terms as far we wish to examine it from an absolute point of view, and not only from a relative point of view.

In our opinion, the source parameters seem to have been well constrained by the inversion (cf. 4.4.4), since we imposed magnitude above M_W=5.5 and since the results obtained for other magnitudes seem to be in good agreement with M_W provided by F-net based catalogue. Conversely, the obtained attenuation terms are quite different from the ones evaluated by Oth et al. (2011b) (cf. Figure 30). Indeed, those attenuation parameters, especially the Q values are the most difficult parameters to invert. Here, we obtained output attenuation model that produces higher regional attenuation than the one proposed Oth et al. (2011b). In the propagation terms, this leads to ground motion amplitude lower by a factor approximately 1.6 in almost the whole frequency range for distances lower than 100 km.

If we re-inject the ratio computed for all frequencies between our obtained propagation terms and the ones deduced from the Oth et al. (2011b) (named COR_{ATT}: correction factor for attenuation term), then RSE computation becomes:

[26]

$$RSE = \frac{SSE_{GIT\_REL} \times GIT_{REF}}{LSE_{1D\_SIM}} \times COR_{ATT}$$

This correction leads to the red continuous line in Figure 52. The COR_{ATT} parameter being approximately equal to 0.6 , we obtain now a RSE almost equal to 1 between 0.5 and 12 Hz.

This could suggest that, at least within this frequency band, there is no “Residual Site Effect” and corollary, that the “Local Site Effect” equals the “Station Site Effect” (statistically and in mean and between 0.5 and 12 Hz), as far as the SSE is corrected to become an absolute measurement of site effect and not a relative one.

Of course, in order to formulate this conclusion, we had to perform a “post correction” of GIT site terms. We will have now to consolidate this conclusion by working on the inversion parametrization in order to get satisfactory regional attenuation terms and thus “clean” site effect terms.

Even without the COR_{ATT} correction, it is remarkable to see that the RSE amplification is almost flat at all frequencies between 0.5 and 12 Hz. This is very different, is terms of typology, from correction

associated to classical V_{S30} and κ_0 corrections that produce highly frequency dependent correction curves (see for example Figure 1 and Figure 2).

This could lead to this another suggestion: at least below 12 Hz, the correction based on 1D simulations, at the LSE scale (in the KiK-net example, from 100 or 200 m to surface in most cases) is sufficient the retrieve the whole station "specificity", and there is no need to apply a "residual" V_{S30} and/or κ_0 correction. We cannot conclude above 12 Hz until we fix the issue of Q_S "scaling".

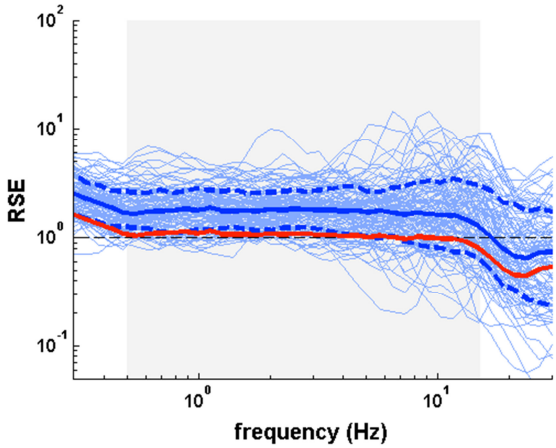

Figure 52: RSE evaluation for each 1D sites (light blue lines) with associated mean (blue continuous line) and standard deviation (blue dashed lines). The red continuous line corresponds to the mean spectral ratio corrected from the "COR_{ATT}" factor.

6.2 DISCUSSION ABOUT κ_0

Our work was not focused on the κ_0 determination or use. Nevertheless, our work provide the opportunity the draw a few figures. We will not comment in deep these figures but it seems to us interesting to show it for possible further discussions.

6.2.1 κ_0 from inverted transfer functions and comparison with other studies

From the site transfer functions obtained from the generalized inversion, we estimate values of the site parameter κ_0 (denoted as $\kappa_{0,TF}$ according to the nomenclature of Ktenidou *et al.* 2014). We compute κ_0 values by regression of the high frequency part of the transfer functions ($f \geq 10\ Hz$) in lin-log scale. Similar to Anderson & Hough (1984), we model the high frequency attenuation through $\exp(-\pi\kappa f)$, however in our case, the κ values are independent of distance since the inversion procedure has already accounted for this.

κ_0 values are computed for surface and downhole stations. For some of the surface stations, the presence of peaks due to high-frequency site-effects in the site transfer functions probably biases the results (Parolai and Bindi, 2004). For these specific sites, κ_0 values are computed above 20 Hz. For a few stations, we also obtain positive slopes and in which case κ_0 have negative values. However, for most of the stations, a good fit is obtained for the regression and we find mean κ_0 values equal to 0.018 s at the surface and 0.011 s for downhole, with relatively high standard deviation (0.011 s and 0.007 s respectively).

Figure 53 shows that κ_0 values at depth and at the surface are not well correlated. As expected, we obtain larger values at the surface, reflecting the influence on κ_0 of shallow layers between downhole and surface.

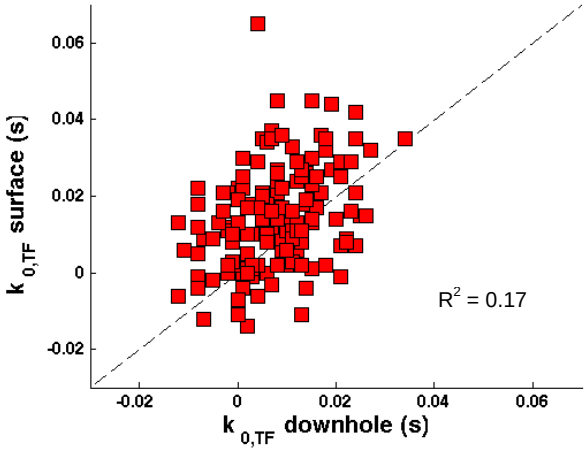

Figure 53: Comparison between surface and downhole κ_0 values. The black dashed line shows the one-to-one relationship.The correlation coefficient is R^2 = 0.17.

Our κ_0 values can be compared with the ones found for KiK-net stations in the literature. In the study of Van Houtte *et al.* (2011), κ_0 are estimated from the classical method of Anderson & Hough (1984) by the regression of the high-frequency decay of the Fourier spectra in lin-log scale ($\kappa_{0,AS}$ according to the nomenclature of Ktenidou *et al.* 2014).

They determined κ_0 values for 148 surface and borehole KiK-net stations of which 125 are in common with our study (Figure 54). Although there is a correlation between κ_0 values obtained from both methods which is better for downhole stations, our values are lower. It can be due to the fact that in the study of Van Houtte *et al.* (2011) they do not consider the effects at high frequencies of the instrumental filter applied to KiK-net stations which leads to an overestimation of the κ_0 values.

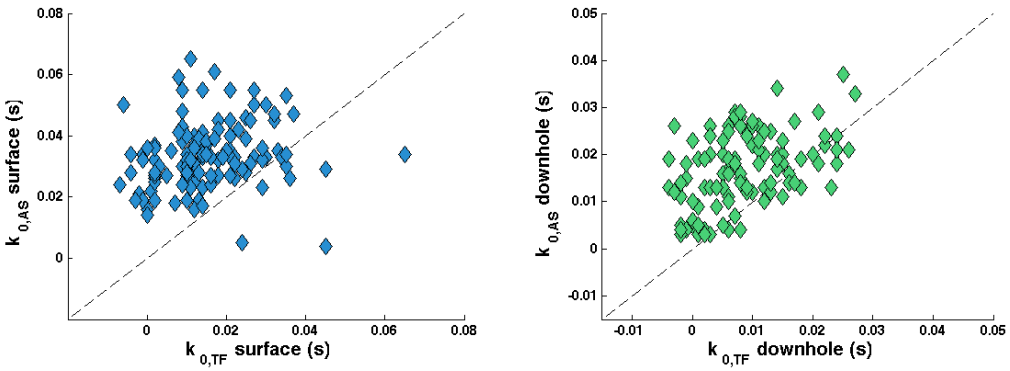

Figure 54: Comparison between $\kappa_{0,TF}$ determined in this study and $\kappa_{0,AS}$ values from Van Houtte et al. (2011) for surface (left) and borehole (right) KiK-net stations. The black dashed lines show the one-to-one relationship.

We can also compare our values with those obtained by Laurendeau *et al.* (2013) from the method of Silva & Darragh (1995) ($\kappa_{0,RESP}$ according to the nomenclature of Ktenidou *et al.* 2014). For the 21 surface sites we have in common with this study, we observe a correlation between the values obtained from the two methods, our values being slightly lower (Figure 55).

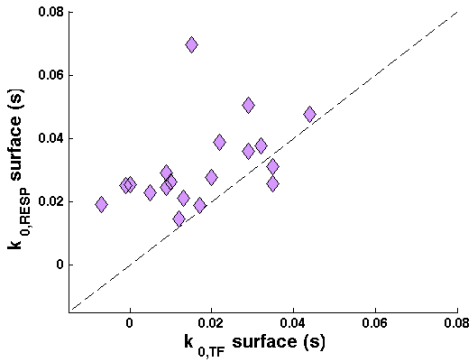

Figure 55: Comparison between $\boldsymbol{\kappa_0}$ determined in this study and $\boldsymbol{\kappa_0}$ values from Laurendeau et al. (2013) for surface KiK-net stations. The black dashed line show the one-to-one relationship.

6.2.2 Study of the $\boldsymbol{\kappa_0}$-$\boldsymbol{v_S}$ relationships

In order to study a potential relationship between the site parameters κ_0 and v_S, we represent κ_0 values obtained at surface stations as a function of v_{S30} and κ_0 values at downhole stations as a function of v_{Sdh} (Figure 56). The linear regression of κ_0 values at surface as a function of v_{S30} shows a flat relationship with no correlation, whereas the linear regression of κ_0 values downhole as a function of v_{Sdh} shows a decrease of κ_0 downhole with increasing v_{Sdh}values. Moreover our values are relatively widespread.

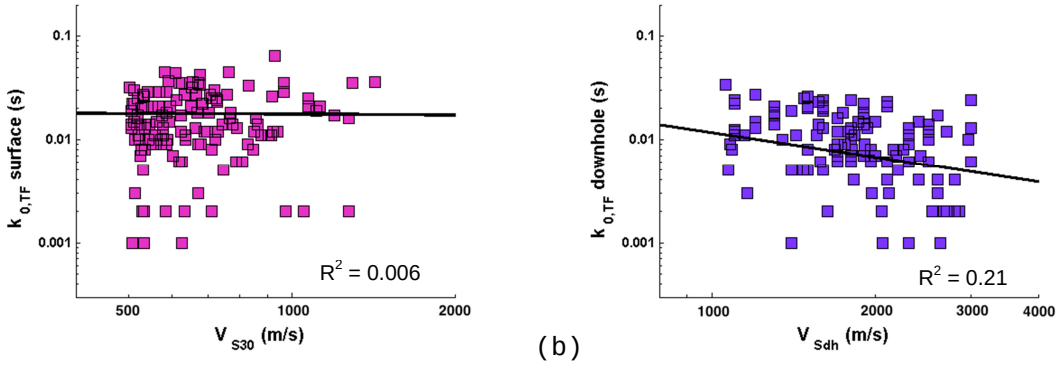

Figure 56: Left: surface $\boldsymbol{\kappa_0}$ as a function of $\boldsymbol{v_{S30}}$values. Right: downhole $\boldsymbol{\kappa_0}$ as a function of $\boldsymbol{v_{Sdh}}$. The solid black lines show the linear regression of the data.

When we compare our results with different κ_0-v_{S30} relationships obtained for different regions of the world from different methodologies and by different authors (*e.g.* Silva *et al.* 1999 (California), Chandler *et al.* 2006 (worldwide), Van Houtte *et al.* 2011 (Japan and NGA), Edwards *et al.* 2011 (Switzerland)), we cannot draw clear conclusions about the correlation between the site parameters κ_0 and v_{S30}.

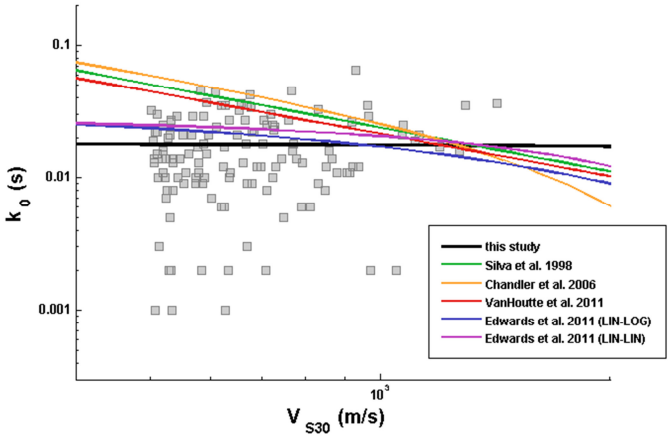

Figure 57: Comparison between κ_0 values as a function of v_{S30} obtained in this study (grey squares and black line for the regression) and different κ_0-v_{S30} relationships from the literature.

7 CONCLUSION

The initial motivation of our work results from the observation that, within a site-specific seismic hazard study, the methods used to evaluate the site effect on the "host side" within the host-to-target adjustments (H2T) were not the same than the ones used to evaluate the site effect on the "target side".

So, basically, we wanted to test an alternative concept in which we attempt to withdraw the site effect for each station of an accelerometric network from the data before deriving a GMPE, using methods as close as possible than the ones used on "target" sites. The KiK-net network (with couple of sensors – downhole and surface– for each site + velocity profiles provided up to at least 100 m depth) offered the opportunity to test this approach with the possibility to check the correction process.

Nevertheless, this work would be useless if we could not transpose it to other databases. Indeed, on most available databases, we do not have downhole sensors. Moreover, just a few stations have a reliable estimation of velocity profiles. We could except that in a near future, more and more stations should benefit from reliable velocity profiles (especially with surface-wave based methods that the Sigma program contributes to make reliable through the Interpacific sub-project). This may allow us to approximate the site effect through 1D transfer function simulation. However, we need a methodology that could be applied to a wider set of databases. In the short and medium term, 1D simulations will not be sufficient. That's why we also worked on Generalized Inversion Techniques (GIT) since it allows getting site effect information on virtually all stations, as far as we have a sufficient amount of records.

In this work, we thus test these different approaches to evaluate the "transfer function" describing the site effect:

- 1D simulation,
- GIT,
- Standard Spectral Ratio (SSR) on records between surface and downhole sensors.

In our work, SSR (that are only available on the KiK-net database) took the role of "validation" of the two other approaches.

We showed that, in terms of transfer functions between the downhole and the surface sensors, the three methods are (statistically and in mean) in good agreement. However, the 1D simulation leads to a slight underestimation of site effect between 6 and 15 Hz and to a strong overestimation above 15 Hz (Figure 38). This is probably due to a "scaling" issue of the Q_S parameter. Indeed, we never get a Q_S profile and we have to deduce it from other parameters (V_S for example). In our work, we used the "standard" $Q_S = V_S/10$ scaling which is no more relevant at high frequency. We still have to work on this issue, for example by developing a scaling approach that also takes into account the depth (as suggest by Assimaki et al. 2008).

We derived two GMPEs. The first, called $SURF_{COR}$ one was based on surface data, corrected for site effect using 1D simulation (on a subset on stations, for which the 1D assumption seems correct). The second one, called DH_{COR}, was based on downhole data "transposed" to surface using the Cadet et al. (2012) method to take into account free-surface phenomenon (and especially to correct the down-going wave destructive interference in downhole data). This last GMPE (that could not be used on other databases due to the lack of downhole sensors) was developed in order to check the first one. Globally, these two GMPE are in good agreement (Figure 51), even at high frequency, that may suggest that the GMPE method, based on 1D simulation site effect estimation, is robust.

These two GMPEs are also compared to the one, derived from the surface data set, without station-by-station site effect correction, but with a standard host-to-target correction (called H2T). At high frequency, the difference between $SURF_{COR}$ (or DH_{COR}) and H2T is high, up to a 3 or 4 factor. This observation is essential and may have a high implication on seismic hazard assessment, if verified. We explain this high difference by the fact that we explicitly attempt to withdraw site effect from each station for $SURF_{COR}$. As we demonstrated (for example on Figure 39), the site effect is high at high frequency (for each class of stations) and this is fact is confirmed by 3 different methods. The host-to-target classical approaches probably do not fully take into account this feature. We suggest that a part of this "high amount of high frequency" due to site effect is indeed considered as resulting from a lower attenuation instead of a local amplification in classical host-to-target adjustments. The fact that the DH_{COR} GMPE, derived from downhole measurements is consistent with the $SURF_{COR}$ strengthens this suggestion.

We also discussed about the use of GIT in order to estimate site effect (within the perspective of applying our approach on database where we do not have down-hole sensors and where the information quality about velocity profiles is lower). The main difficulty of GIT is the fact that the "site terms" are expressed in a relative form. We have thus to "post-process" them to provide and absolute site effect evaluation. Nevertheless, as far as we have some additional information in order to better parametrize the inversion, we think that GIT may provide robust results that may use for site effect "withdrawing" from data before GMPE derivation. We also suggested (but this conclusion is still weak and need to be confirm) that at least between 0.5 and 12 Hz, the whole site effect take place at the scale between the surface and the underlying bedrock. With the Kik-Net data we used, this was between the surface and 100 or 200 m depth area (since we select station where downhole sensors were within the bedrock). Thus, there is no "residual site effect" below this zone and our method does not need additional corrections.

The work to be done in the continuity of the present report is summarized here:

- better understand the Q_S scaling and propose a methodology in order to provide a better simulation for very high frequencies (> 15 Hz),
- perform new GIT on our dataset in order to get a better evaluation of regional attenuation, and thus, to get better absolute "site terms",
- deriving a corrected GMPE using the whole data site (not only the 1D station) in order to evaluate the bias induced by the lack of 1D site selection,
- deriving a corrected GMPE using the site effect evaluation from GIT.

8 REFERENCES

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

Al Atik, L., N. Abrahamson, F. Cotton, F. Scherbaum, J. Bommer, And N. Kuehn (2010). The variability of ground-motion prediction models and its components, Seismol. Res. Lett. 81, 794–801.

Allmann B.P. & Shearer P.M. (2009) Global variations of stress drop for moderate to large earthquakes, *J. Geophys. Res.* **114**, B01310, doi: 10.1029/1008JB005821.

Anderson J. & Hough S. (1984) A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies, Bull. Seismol. Soc. Am. 74, 1969-1993.

Anderson, J. G., and S. E. Hough (1984). A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies, Bull. Seismol. Soc. Am. 74, 1969–1993.

Andrews D.J. (1986) Objective determination of source parameters and similarity of earthquakes of different size, *in Earthquake Source Mechanics, pp. 259-267, eds. Das S., Boatwright J. & Scholz C.H., American geophysical Monograph 37*.

Aoi S., Kunugi T. & Fujiwara H. (2004) Strong-motion seismograph network operated by NIED: K-NET and KiK-net, *Journal of Japan Association for Earthquake Engineering* **4**, 65-74.

Assimaki, D., Li, W., Steidl, J. H., & Tsuda, K. (2008). Site amplification and attenuation via downhole array seismogram inversion: A comparative study of the 2003 Miyagi-Oki aftershock sequence. Bulletin of the Seismological Society of America, 98(1), 301-330.

Baltay A., Ide S., Prieto G. & Beroza G. (2011) Variability in earthquake stress drop and apparent stress, Geophys. Res. Lett. 38, L06303, doi: 10.1029/2011GL046698.

Bard P.-Y. and Gariel J.-C. (1986). The seismic response of two-dimensional sedimentary deposits with large vertical velocity gradients. Bulletin of the Seismological Society of America, 76(2):343–346.

Biro Y., and P. Renault (2012), Importance and Impact of Host-to-Target Conversions for Ground Motion Prediction Equations in PSHA, 15 WCEE, Lisboa 2012.

Boore D. & Boatwright J. (1984) Average body-wave radiation coefficients, Bull. Seismol. Soc. Am. 74, 1615-1621.

Boore D. & Joyner W. (1997) Site amplifications for generic rock sites, Bull. Seismol. Soc. Am. 87, 327-341.

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

Borcherdt RD (1970) Effect on local geology on ground motion near San Francisco bay. Bulletin of the Seismological Society of America, 60:29–61.

Brocher, T. M. (2005), Empirical relations between elastic wavespeeds and density in the earth's crust, Bulletin of the Seismological Society of America, 95 (6), 2081–2092.

Brune J.N. (1970) Tectonic stress and the spectra of seismic shear waves from earthquakes, J. Geophys. Res. 75, 4997-5009.

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, Bulletin of Earthquake Engineering, 10 (2), 421–448.

Campbell, K. W. (2003). Prediction of Strong Ground Motion Using the Hybrid Empirical Method and Its Use in the Development of Ground-Motion (Attenuation) Relations in Eastern North America. Bulletin of the Seismological Society of America 93:3, 1012–1033.

Campbell K.W. (2009) Estimates of shear-wave Q and k0 for unconsolidated and semiconsolidated sediments in Eastern North America, *Bull. Seismol. Soc. Am.* **99**, 2365-2392.

Chandler A.M., Lam N.T.K. & Tsang H.H. (2006) Near-surface attenuation modelling based on rock shear-wave velocity profile, *Soil Dyn. Earthq. Eng.* **26**, 1004-1014.

Cotton F., Scherbaum F., Bommer J. & Bungum H. (2006) Criteria for selecting and adjusting ground-motion models for specific target regions: application to central Europe and rock sites, J. Seismol. 10, 137-156.

Cotton, F., Pousse, G., Bonilla, F. and Scherbaum, F. (2008). On the discrepancy of recent European ground-motion observations and predictions from empirical models : Analysis of KiK-net accelerometric data and point-sources stochastic simulations. Bulletin of the Seismological Society of America, 98(5):2244.

Drouet S., Chevrot S., Cotton F. & Souriau A. (2008) Simultaneous inversion of source spectra, attenuation parameters, and site responses: application to the data of the French acccelerometric network, Bull. Seismol. Soc. Am. 98, doi: 10.1785/0120060215.

Drouet S., Cotton F. & Guéguen P. (2010) v_S30, κ, regional attenuation and M_w from accelerograms : application to magnitude 3-5 French earthquakes, Geophys. J. Int. 182, 880-898, doi: 10.1111/j.1365-246X.2010.04626.x.

Drouet S., Bouin M-P. & Cotton F. (2011) New moment magnitude scale, evidence of stress drop magnitude scaling and stochastic ground motion model for the French West Indies, Geophys. J. Int. 187, 1625-1644, doi: 10.1111/j.1365-246X.2011.05219.x.

Edwards B., Rietbrock A., Bommer J.J. & Baptie B. (2008) The acquisition of source, path and site effects from micro earthquake recordings using Q tomography: application to the United Kingdom, Bull. Seismol. Soc. Am. 98, doi: 10.1785/0120070127.

Edwards B., Faeh D. & Giardini D. (2011) Attenuation od seismic shear-wave energy in Switzerland, *Geophys. J. Int.* **185**, 967-984.

Field E.H. & Jacob K.H. (1995) A comparison and test of various site-response estimation techniques, including three that are not reference-site dependent, Bull. Seismol. Soc. Am. 85, 1127-1143.

Fujiwara, H., Aoi, S., Kunugi, T. and Adachi, S. (2004). Strong-motion observations of NIED: K-NET and KiK-net. In Proceedings of the COSMOS Workshop on Record Processing Guidelines.

Hanks T.C. & Kanamori H. (1979) A moment magnitude scale, *J. Geophys. Res.* **84**, 2348-2350.

Hanks T.C. (1982) f_{max}, *Bull. Seismol. Soc. Am.* **72**, 1867-1879.

Haskell, N.A. (1953). The dispersion of surfaces waves in multilayered media. Bulletin of the Seismological Society of America 43, 17-34.

Ide S., Beroza G.C., Prejean S.G. & Ellsworth W.L. (2003) Apparent break in earthquake scaling due to path and site effects on deep borehole recordings, *J. Geophys. Res.* **108**, B5-2271, doi: 10.1029/2001JB001617.

Kanno, T., Narita, A., Morikawa, N., Fujiwara, H. et Fukushima, Y. (2006). A new attenuation relation for strong ground motion in Japan based on recorded data. Bulletin of the Seismological Society of America, 96(3):879-897.

Kennett BLN (1974). Reflections, rays, and reverberations. Bulletin of the Seismological Society of America, 64(6):1685–1696.

Kinoshita, S. (1998). Kyoshin net (K-NET). Seismological Research Letters, 69(4):309-332.

Konno K., and Ohmachi T (1998). Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor. Bulletin of the Seismological Society of America, 88(1):228–241.

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, p. 135-146.

Ktenidou, O., T. Kishida, R. Darragh, W. Silva, and N. Abrahamson (2014), Fourier spectra and kappa0 (κ_0) estimates for rock stations in the nga-west2 project, in Proceedings of the 10th U.S. National Conference on Earthquake Engineering, Frontiers of Earthquake Engineering, Anchorage, Alaska.

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), Bulletin of the Seismological Society of America, 103 (6), 3131–3148.

Lebrun B., Hatzfeld D., and Bard P.-Y., (2001), A site effect study in urban area: experimental results in Grenoble (France), Pure and Applied Geophysics, 158:2543-2557.

Lermo J, and Chavez-Garcia FJ (1993). Site effect evaluation using spectral ratios with only one station. Bulletin of the Seismological Society of America, 83(5):1574–1594

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

Oth A., Bindi D., Parolai S. & Di Giacomo D. (2010) Earthquake scaling characteristics and the scale-(in)dependence of seismic energy-to-moment ratio: insights from KiK-net data in Japan, *Geophys. Res. Lett.* **37**, L19304, doi: 10.1029/2010GL044572.

Oth, A., Parolai, S. and Bindi, D. (2011a). Spectral analysis of K-NET and KiK-net data in Japan, Part I: Database compilation and peculiarities. Bulletin of the Seismological Society of America, 01(2):652-666.

Oth A., Bindi D., Parolai S. & Di Giacomo D. (2011b) Spectral analysis of K-NET and KiK-net data in Japan, Part II: On attenuation characteristics, source spectra, and site response of borehole and surface stations, *Bull. Seismol. Soc. Am.* **101**(2), 667-687, doi: 10.1785/0120100135.

Parolai S. & Bindi D. (2004) Influence of soil-layer properties on k evaluation, *Bull.Seismol. Soc. Am.* **94**, 349-356.

Pei S., Cui Z., Sun Y., Toksoz M.N., Rowe C.A., Gao X., Zhao J., Liu H., He J. & Morgan F.D. (2009) Structure of the upper crust in Japan from S-wave attenuation tomography, Bull. Seismol. Soc. Am. 99, 428-434.

Pileggi D., Rossi D., Lunedei E. and Albarello D., (2011), Seismic characterization of rigid sites in the ITACA database by ambient vibration monitoring and geological surveys, Bull Earthquake Eng (2011) 9:1839–1854.

Rodriguez-Marek, A., Rathje, E. M., Bommer, J. J., Scherbaum, F., and Stafford, P. J. (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: 1601-1619.

Scherbaum, F., Schmedes, J. and Cotton F. (2004). On the Conversion of Source-to-Site Distance Measures for Extended Earthquake Source Models. Bulletin of the Seismological Society of America. 94:3, 1053-1069.

Silva W. & Darragh R. (1995) Engineering characterization of earthquake strong ground motion recorded at rock sites, *Electric Power Research Institute, Report TR-102261,* Palo Alto, California.

Silva W., Darragh R.B., Gregor N., Martin G., Abrahamson N. & Kircher C. (1998) Reassessment of site coefficients and near-fault factors for building code provisions, *Technical Report Program Element II: 98-HQ-GR-1010, Pacific Engineering and Analysis*, El Cerrito, California USA.

Silva,W. J., and R. Darragh (2012). Assessment of kappa for Vertical and Horizontal Motions at Rock Sites Using Spectral Shapes, Final Report no. EXT-TB-1089, Pegasos Refinement Project.

Steidl JH, Tumarkin AG and Archuleta RJ (1996). What is a reference site? Bulletin of the Seismological Society of America, 86(6): 1733–1748.

Theodulidis N. and Bard P.-Y. (1995). Horizontal to vertical spectral ratio and geological conditions—an analysis of strong ground motion data from Greece and Taiwan (smart-1). Soil Dynamics and Earthquake Engineering 14(3):177–197.

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.

Thomson, W.T. (1950). Transmission of elastic waves through a stratified solid. Journal of Applied Physics 21, 89-93.

Van Houtte, C., S. Drouet, and F. Cotton (2011), Analysis of the origins of κ (kappa) to compute hard rock to rock adjustment factors for gmpes, Bulletin of the Seismological Society of America, 101 (6), 2926–2941.

ANNEXES

Table 3: Regression coefficients based on DATAsurf.

Per (s)	a_1	a_2	a_3	b_1	c_1	tau	phi	sigma TOT
0.03	-17.951566	5.514416	-0.371507	-0.012018	-0.067190	0.548	0.691	0.882
0.04	-17.819776	5.587198	-0.379644	-0.012878	0.055032	0.571	0.698	0.902
0.05	-17.905037	5.725913	-0.394313	-0.013303	0.104841	0.574	0.698	0.904
0.06	-17.753037	5.766323	-0.399856	-0.013469	0.061061	0.575	0.713	0.916
0.08	-19.611574	6.478096	-0.459642	-0.013051	0.082796	0.552	0.781	0.957
0.10	-19.272902	6.281062	-0.441060	-0.012445	-0.277742	0.533	0.793	0.955
0.12	-19.335885	6.178943	-0.426939	-0.012134	-0.588696	0.525	0.787	0.946
0.15	-20.272738	6.351647	-0.437926	-0.011345	-0.880892	0.512	0.793	0.944
0.19	-20.454661	6.264484	-0.427683	-0.010180	-0.924174	0.494	0.748	0.896
0.23	-20.716935	6.238701	-0.423166	-0.009317	-0.941268	0.470	0.725	0.864
0.29	-19.406139	5.619114	-0.364499	-0.008765	-0.881612	0.486	0.688	0.842
0.36	-20.091474	5.700198	-0.367652	-0.007670	-0.844118	0.481	0.681	0.834
0.44	-20.130835	5.544209	-0.348887	-0.006646	-0.778945	0.493	0.665	0.828
0.55	-19.802570	5.265742	-0.321570	-0.005350	-0.668089	0.479	0.635	0.795
0.68	-19.834016	5.093812	-0.301458	-0.004494	-0.653181	0.452	0.607	0.757
0.84	-20.112142	5.014459	-0.288667	-0.003615	-0.566544	0.440	0.593	0.738
1.05	-21.114308	5.174928	-0.297239	-0.002598	-0.559995	0.412	0.588	0.718
1.30	-20.290978	4.686475	-0.249486	-0.001736	-0.485550	0.365	0.565	0.673
1.61	-18.780184	3.951829	-0.179487	-0.001156	-0.445989	0.328	0.532	0.624
2.00	-17.178281	3.231430	-0.112977	-0.000931	-0.349224	0.323	0.510	0.604

Table 4: Regression coefficients based on DATAprof.

Per (s)	a_1	a_2	a_3	b_1	c_1	tau	phi	sigma TOT
0.03	-14.710625	3.710741	-0.202259	-0.010035	-0.149826	0.556	0.557	0.787
0.04	-14.212067	3.643002	-0.198074	-0.011316	-0.081624	0.567	0.567	0.802
0.05	-14.601262	3.888712	-0.221248	-0.012142	-0.116182	0.592	0.592	0.837
0.06	-15.138420	4.118887	-0.242373	-0.012138	-0.116107	0.581	0.581	0.822
0.08	-16.180264	4.522478	-0.278498	-0.011897	-0.099873	0.573	0.574	0.811
0.10	-16.401070	4.602358	-0.284625	-0.011439	-0.156058	0.566	0.574	0.806
0.12	-18.341089	5.280081	-0.341499	-0.010994	-0.265399	0.552	0.593	0.810
0.15	-18.867702	5.431685	-0.352750	-0.010250	-0.342711	0.514	0.606	0.795
0.19	-19.157340	5.487858	-0.355150	-0.009504	-0.452770	0.501	0.621	0.798
0.23	-18.112198	5.041696	-0.314652	-0.008454	-0.486970	0.478	0.629	0.791
0.29	-18.516454	5.039596	-0.311891	-0.006943	-0.357103	0.479	0.634	0.795
0.36	-19.815705	5.373944	-0.338536	-0.005603	-0.273495	0.491	0.610	0.783
0.44	-20.834984	5.623010	-0.354691	-0.005023	-0.255623	0.488	0.600	0.773
0.55	-21.305097	5.685478	-0.357050	-0.003921	-0.318443	0.464	0.566	0.732
0.68	-21.331783	5.592631	-0.344839	-0.003114	-0.425302	0.424	0.526	0.676
0.84	-21.457980	5.500233	-0.330340	-0.002464	-0.502583	0.414	0.523	0.667
1.05	-21.817205	5.473264	-0.322090	-0.001544	-0.582228	0.385	0.521	0.648
1.30	-21.376457	5.126582	-0.285682	-0.000828	-0.590763	0.349	0.496	0.607
1.61	-19.865419	4.393908	-0.214949	-0.000288	-0.600344	0.328	0.457	0.562
2.00	-18.716830	3.813692	-0.161168	-0.000028	-0.497176	0.330	0.447	0.556

Table 5: Regression coefficients based on a mixed dataset: DATAsurf + DATAprof.

Per (s)	a_1	a_2	a_3	b_1	c_1	tau	phi	sigma TOT
0.03	-15.723481	4.432996	-0.271088	-0.011192	-1.107575	0.589	0.684	0.903
0.04	-15.378169	4.422642	-0.271911	-0.012280	-1.073050	0.619	0.695	0.931
0.05	-15.613759	4.607682	-0.290177	-0.012917	-1.089929	0.627	0.706	0.944
0.06	-15.759354	4.731433	-0.302308	-0.013008	-1.193937	0.625	0.720	0.953
0.08	-17.115075	5.263999	-0.348292	-0.012694	-1.313096	0.612	0.766	0.980
0.10	-17.297881	5.311637	-0.351507	-0.012107	-1.337639	0.589	0.753	0.956
0.12	-18.217233	5.578734	-0.371084	-0.011683	-1.340162	0.573	0.739	0.935
0.15	-19.068944	5.792264	-0.386744	-0.010869	-1.275957	0.547	0.728	0.911
0.19	-19.408845	5.799496	-0.384791	-0.009901	-1.161145	0.529	0.697	0.875
0.23	-19.090603	5.583710	-0.364046	-0.008933	-1.086479	0.503	0.685	0.850
0.29	-18.679930	5.292296	-0.335082	-0.007877	-0.942244	0.507	0.669	0.840
0.36	-19.702675	5.512316	-0.351096	-0.006641	-0.815081	0.509	0.653	0.828
0.44	-20.345391	5.590991	-0.352576	-0.005835	-0.681070	0.515	0.633	0.816
0.55	-20.462934	5.479643	-0.339760	-0.004638	-0.564667	0.497	0.597	0.777
0.68	-20.588970	5.364716	-0.325119	-0.003796	-0.473445	0.464	0.562	0.728
0.84	-20.857056	5.284345	-0.311922	-0.003032	-0.400062	0.454	0.552	0.715
1.05	-21.612599	5.366409	-0.313417	-0.002057	-0.349882	0.428	0.550	0.697
1.30	-21.022626	4.955958	-0.271999	-0.001258	-0.301717	0.386	0.528	0.654
1.61	-19.543277	4.227793	-0.202154	-0.000682	-0.267636	0.352	0.495	0.607
2.00	-18.128160	3.566306	-0.141136	-0.000419	-0.228689	0.348	0.477	0.591

Table 6: Regression coefficients based on DHcor.

Per (s)	a_1	a_2	a_3	b_1	c_1	tau	phi	sigma TOT
0.03	-14.113665	3.704214	-0.201753	-0.010046	-0.174127	0.556	0.556	0.786
0.04	-13.639009	3.647481	-0.198647	-0.011330	-0.123164	0.567	0.568	0.802
0.05	-14.009087	3.880121	-0.220574	-0.012137	-0.158976	0.591	0.591	0.835
0.06	-14.494232	4.090358	-0.239960	-0.012138	-0.166354	0.580	0.580	0.820
0.08	-15.532261	4.488765	-0.275567	-0.011897	-0.164682	0.571	0.571	0.807
0.10	-16.650717	4.889338	-0.310061	-0.011448	-0.232863	0.565	0.573	0.805
0.12	-17.643392	5.216036	-0.335742	-0.011082	-0.302293	0.555	0.593	0.812
0.15	-18.077811	5.301918	-0.340887	-0.010424	-0.185352	0.527	0.611	0.807
0.19	-18.873142	5.509471	-0.356551	-0.009309	-0.122373	0.514	0.627	0.811
0.23	-18.198278	5.221446	-0.330048	-0.008103	-0.243321	0.480	0.642	0.801
0.29	-18.701773	5.337684	-0.337929	-0.006914	-0.398491	0.461	0.640	0.789
0.36	-19.698751	5.624029	-0.361079	-0.005845	-0.571044	0.461	0.611	0.765
0.44	-19.978354	5.620888	-0.356172	-0.005060	-0.706588	0.465	0.596	0.756
0.55	-19.910826	5.452288	-0.338311	-0.003796	-0.782296	0.459	0.569	0.732
0.68	-19.779253	5.241334	-0.315227	-0.003074	-0.778058	0.438	0.536	0.692
0.84	-20.001027	5.132166	-0.298194	-0.002490	-0.786285	0.429	0.532	0.684
1.05	-20.935785	5.269402	-0.304187	-0.001514	-0.783819	0.395	0.527	0.659
1.30	-20.966709	5.064959	-0.280147	-0.000772	-0.744778	0.355	0.498	0.611
1.61	-19.737119	4.412556	-0.216419	-0.000281	-0.711256	0.330	0.471	0.575
2.00	-18.737647	3.875746	-0.166648	-0.000062	-0.582311	0.330	0.455	0.562

Table 7: Regression coefficients based on a mixed dataset: DHcor + DATAsurf.

Per (s)	a_1	a_2	a_3	b_1	c_1	tau	phi	sigma TOT
0.03	-15.72110	4.51391	-0.27818	-0.01115	-0.66730	0.587	0.643	0.870
0.04	-15.39099	4.50863	-0.27950	-0.01225	-0.64089	0.615	0.652	0.897
0.05	-15.61084	4.68525	-0.29697	-0.01288	-0.66831	0.624	0.664	0.911
0.06	-15.71368	4.79501	-0.30809	-0.01297	-0.78041	0.621	0.674	0.917
0.08	-17.07683	5.32533	-0.35366	-0.01266	-0.91538	0.607	0.717	0.940
0.10	-17.51451	5.46487	-0.36521	-0.01208	-0.95613	0.585	0.711	0.921
0.12	-18.13130	5.62245	-0.37502	-0.01169	-0.95569	0.573	0.706	0.909
0.15	-18.87494	5.79826	-0.38706	-0.01090	-0.84347	0.553	0.712	0.902
0.19	-19.42583	5.88272	-0.39193	-0.00973	-0.65928	0.533	0.692	0.874
0.23	-19.34633	5.75485	-0.37892	-0.00869	-0.57211	0.501	0.684	0.848
0.29	-19.11138	5.53645	-0.35638	-0.00782	-0.45135	0.501	0.660	0.829
0.36	-20.05911	5.73339	-0.37064	-0.00673	-0.38651	0.499	0.644	0.815
0.44	-20.35519	5.68004	-0.36103	-0.00584	-0.32720	0.508	0.631	0.810
0.55	-20.14595	5.43450	-0.33649	-0.00457	-0.29209	0.498	0.606	0.784
0.68	-20.09183	5.23798	-0.31432	-0.00377	-0.26633	0.475	0.578	0.748
0.84	-20.37765	5.15022	-0.30013	-0.00304	-0.23901	0.465	0.568	0.734
1.05	-21.36290	5.30444	-0.30790	-0.00204	-0.22270	0.435	0.564	0.713
1.30	-20.96537	4.95612	-0.27190	-0.00123	-0.20017	0.388	0.539	0.664
1.61	-19.58997	4.26087	-0.20494	-0.00067	-0.18841	0.353	0.510	0.620
2.00	-18.22620	3.61618	-0.14549	-0.00043	-0.16351	0.348	0.486	0.598

Table 8: Regression coefficients based on SURFcor.

Per (s)	a_1	a_2	a_3	b_1	c_1	tau	phi	sigma TOT
0.03	-20.548033	6.178139	-0.431012	-0.012443	-0.032397	0.574	0.646	0.864
0.04	-19.433998	5.866697	-0.405889	-0.013418	-0.040299	0.595	0.645	0.877
0.05	-19.495749	5.956209	-0.414966	-0.014049	-0.017635	0.603	0.658	0.893
0.06	-19.899394	6.151548	-0.433714	-0.014515	0.137168	0.582	0.674	0.891
0.08	-22.340819	7.008086	-0.507529	-0.013408	0.365145	0.528	0.779	0.941
0.10	-23.177435	7.327750	-0.534889	-0.012282	0.141910	0.548	0.803	0.973
0.12	-22.902112	7.225752	-0.520571	-0.012105	-0.069151	0.541	0.787	0.955
0.15	-24.252455	7.641910	-0.552566	-0.011292	-0.149562	0.508	0.739	0.897
0.19	-23.316835	7.238531	-0.516966	-0.009386	-0.226034	0.488	0.672	0.830
0.23	-23.420983	7.214199	-0.512108	-0.008352	-0.336613	0.445	0.656	0.793
0.29	-22.397855	6.727428	-0.464558	-0.007273	-0.444368	0.465	0.645	0.795
0.36	-22.594818	6.666578	-0.454463	-0.006287	-0.517069	0.475	0.621	0.782
0.44	-22.165077	6.381640	-0.423586	-0.005483	-0.625970	0.497	0.592	0.773
0.55	-21.561413	6.016098	-0.388842	-0.004143	-0.675378	0.473	0.548	0.724
0.68	-21.330093	5.769052	-0.360578	-0.003574	-0.732011	0.430	0.527	0.680
0.84	-22.422420	5.976850	-0.371127	-0.003162	-0.728137	0.406	0.502	0.645
1.05	-22.524295	5.813496	-0.349664	-0.002253	-0.764003	0.366	0.501	0.620
1.30	-21.251940	5.148351	-0.284742	-0.001548	-0.759880	0.334	0.486	0.589
1.61	-19.525375	4.328148	-0.207555	-0.000924	-0.712348	0.302	0.453	0.544
2.00	-17.596623	3.465464	-0.128734	-0.000949	-0.565888	0.285	0.438	0.523

Table 9: Regression coefficients based on a mixed dataset: SURFcor + DATAsurf.

Per (s)	a_1	a_2	a_3	b_1	c_1	tau	phi	sigma TOT
0.03	-19.746776	6.076082	-0.425007	-0.011990	-0.671765	0.616	0.647	0.893
0.04	-18.805976	5.841225	-0.406853	-0.012902	-0.727741	0.636	0.658	0.915
0.05	-18.722312	5.903814	-0.414204	-0.013500	-0.775854	0.644	0.670	0.929
0.06	-18.895381	6.065023	-0.429622	-0.013990	-0.819102	0.635	0.699	0.944
0.08	-21.106047	6.894836	-0.500522	-0.013201	-0.788053	0.612	0.789	0.998
0.10	-22.297238	7.290802	-0.534345	-0.012223	-0.815233	0.620	0.784	0.999
0.12	-22.649102	7.332881	-0.532669	-0.011756	-0.774065	0.598	0.760	0.967
0.15	-23.882296	7.649441	-0.555526	-0.010838	-0.679170	0.558	0.723	0.913
0.19	-23.068643	7.230250	-0.517367	-0.009179	-0.546107	0.533	0.667	0.853
0.23	-23.458235	7.254267	-0.515623	-0.008379	-0.464368	0.484	0.651	0.811
0.29	-22.546014	6.761395	-0.466472	-0.007587	-0.398132	0.501	0.637	0.810
0.36	-22.757252	6.670165	-0.453260	-0.006673	-0.342034	0.510	0.620	0.802
0.44	-22.538202	6.414632	-0.424627	-0.005829	-0.293159	0.527	0.598	0.797
0.55	-22.095266	6.079353	-0.392422	-0.004430	-0.249501	0.497	0.562	0.750
0.68	-22.024260	5.863972	-0.366762	-0.003822	-0.219334	0.454	0.548	0.712
0.84	-23.207936	6.096527	-0.379530	-0.003363	-0.185465	0.426	0.527	0.678
1.05	-23.415979	5.958665	-0.360349	-0.002409	-0.178246	0.386	0.531	0.656
1.30	-22.210079	5.318064	-0.297702	-0.001673	-0.172545	0.347	0.515	0.621
1.61	-20.431526	4.487600	-0.219897	-0.000995	-0.155284	0.309	0.481	0.572
2.00	-18.314402	3.590544	-0.138559	-0.000978	-0.116301	0.296	0.454	0.542

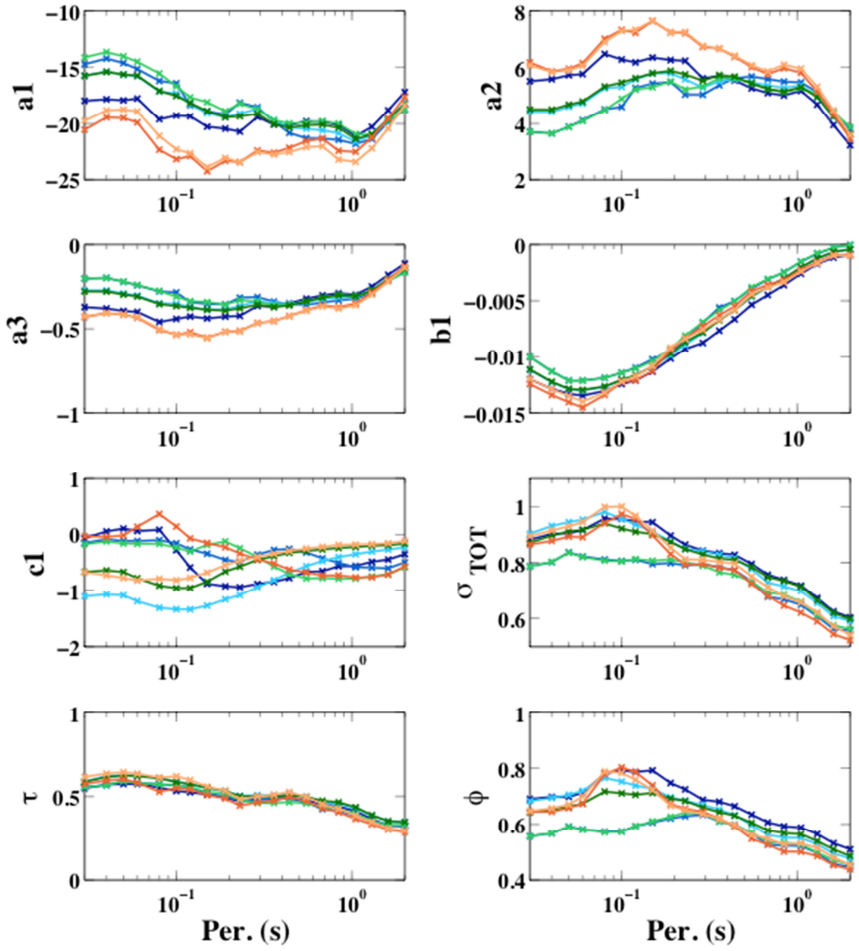

Figure 58: GMPEs coefficients.

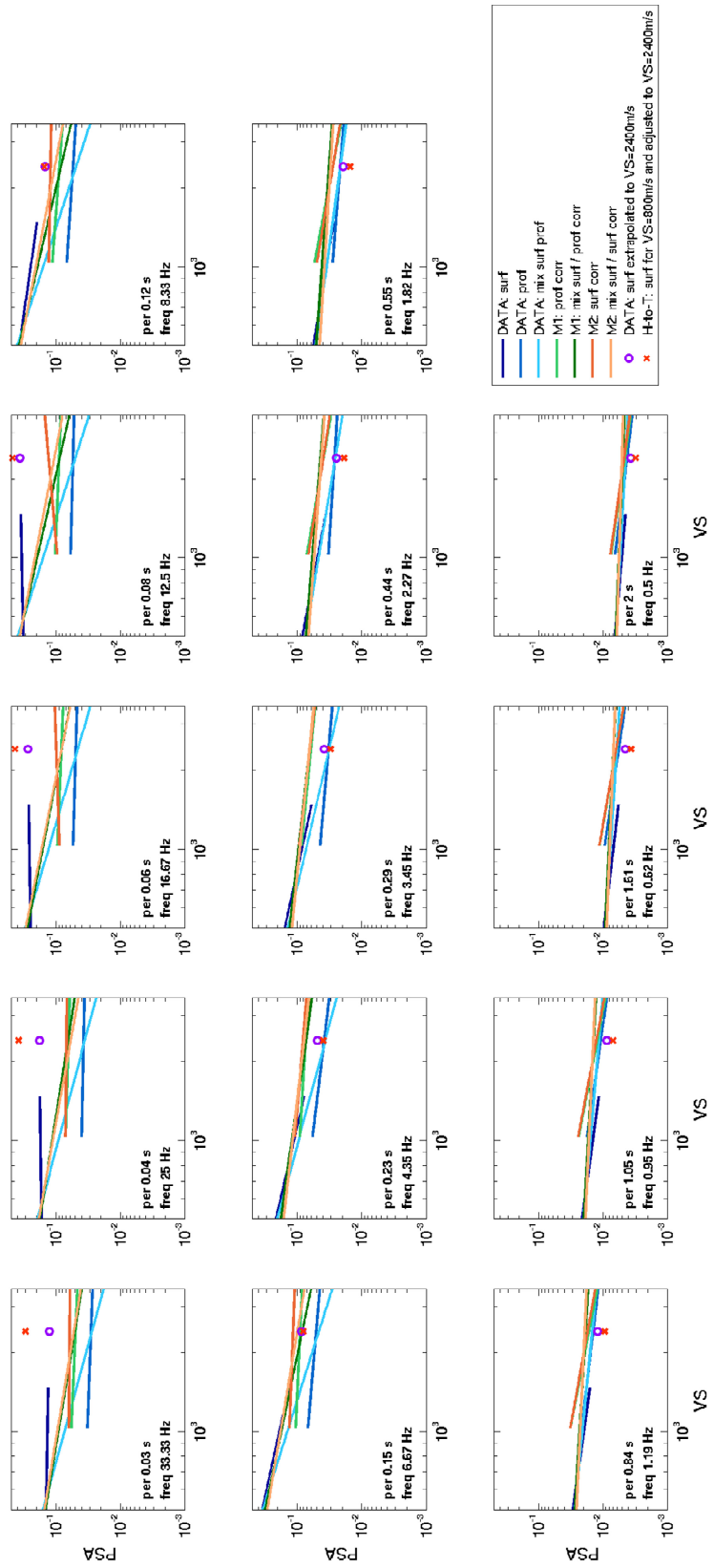

Figure 59: Predicted response spectra according to VS for different periods.

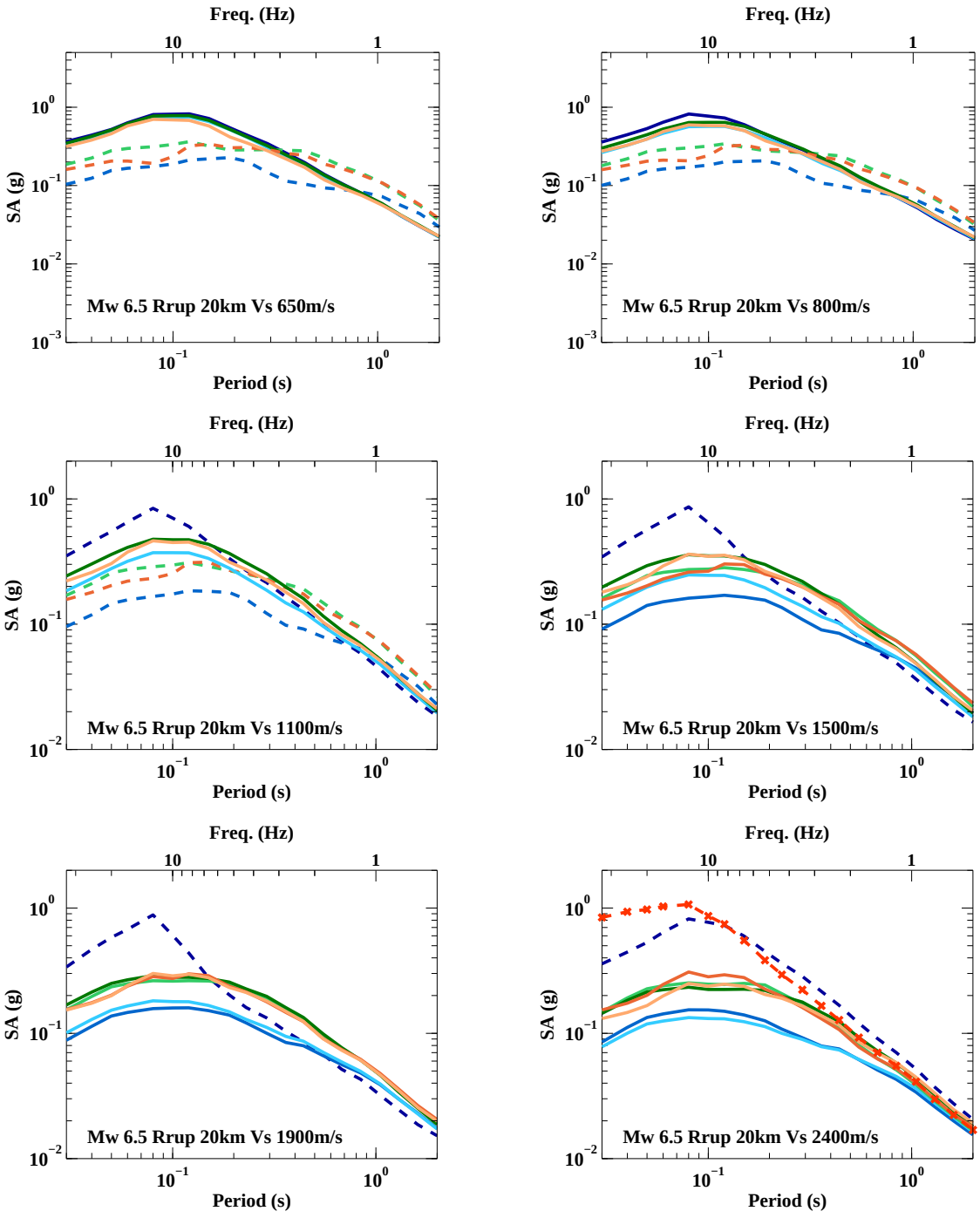
Freq. (Hz)
10
1
SA (g)
Period (s)
Mw 6.5 Rrup 20km Vs 650m/s
Mw 6.5 Rrup 20km Vs 800m/s
Mw 6.5 Rrup 20km Vs 1100m/s
Mw 6.5 Rrup 20km Vs 1500m/s
Mw 6.5 Rrup 20km Vs 1900m/s
Mw 6.5 Rrup 20km Vs 2400m/s

Figure 60