Seismic site-response variability : from site-classification to soil non-linear behaviour

Julie Regnier

UNIVERSITÉ PARIS-EST

Thèse présentée pour obtenir le grade de

Docteur de l'Université Paris-Est

Spécialité: Sciences de l'Univers et Environnement

par

Julie Regnier

Ecole Doctorale : SCIENCES, INGÉNIERIE ET ENVIRONNEMENT

Variabilité de la réponse sismique: de la classification des sites au comportement non-linéaire des sols.

Seismic site-response variability: From site-classification to soil non-linear behaviour.

Thèse soutenue le 21 mai 2013 devant le jury composé de:

Philippe Gueguen	*Rapporteur*
Fäh Donat	*Rapporteur*
Françoise Courboulex	*Examinateur*
Alain Pecker	*Examinateur*
Fernando Lopez-Caballero	*Examinateur*
Jean-Francois Semblat	*Directeur de thèse*
Luis-Fabian Bonilla	*Directeur de thèse*
Etienne Bertrand	*Conseiller d'étude*

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE
CETE Méditerranée

IFSTTAR

CETE Méditérranée, Laboratoire de Nice, 56 Bd de Stalingrad 06359 Nice.
IFSTTAR, 14-20 Boulevard Newton Cité Descartes, Champs sur Marne F-77447 Marne la Vallée Cedex 2.

Résumé

La configuration géologique des couches de sol proches de la surface peut modifier fortement le mouvement sismique en surface (effets de site lithologiques). Ces effets de sites peuvent être évalués numériquement ou empiriquement. Ce travail s'inscrit dans la thématique de l'évaluation des effets de site lithologiques et de la variabilité de la réponse sismique des sites. Il se base principalement sur l'analyse d'enregistrements de séismes. Dans un premier temps, nous avons étudié la variabilité de la réponse sismique par classe de sites et pour des mouvements sismiques faibles (variabilité inter-sites), puis une partie de la variabilité inter-évènements en un site, due aux effets du comportement non-linéaire des sols. Enfin, nous avons inversé les courbes de réponse du site afin de préciser à quelles profondeurs le comportement du sol est non-linéaire.

La base de données accélérométriques KiK-net (Japon) est constituée de 688 forages instrumentés avec un capteur en surface et un autre en fond de puits. Cette configuration en réseau vertical permet de calculer en chacun des sites la courbe de réponse empirique du forage. Ce réseau a été choisi compte tenu du nombre important de sites instrumentés, du nombre de séismes enregistrés (plus de 46000 enregistrements ont été recueillis et analysés) et de l'existence de caractérisation géotechnique des sites (profils de vitesse de propagation des ondes de cisaillement et de compression). L'analyse de la variabilité de la réponse sismique par classe de site nous a permis de préciser les paramètres utiles afin d'améliorer l'évaluation des effets de site. En régime linéaire, il s'agit du gradient du profil de vitesse calculé jusqu'à 30 m (B_{30}) ou 100 m (B_{100}) et de la fréquence de résonance fondamentale (f_0). Ces paramètres pourraient être utilisés afin de réduire la variabilité inter-sites du mouvement sismique en surface (i.e. variabilité dans les lois d'atténuation).

L'analyse de la variabilité inter-évènement associée au comportement non-linéaire des sols nous a fourni les paramètres pertinents pour l'analyse des sites en régime non-linéaire : B_{30}, l'amplification maximale (A_{pred}) et la fréquence associée (f_{pred}). Cette analyse a également montré :

- que le PGA (Peak Ground Acceleration, accélération maximale du sol) est un paramètre pertinent pour l'étude des effets du comportement non-linéaire des sols sur la réponse des sites.
- que, quelque soit le site, le comportement non-linéaire des sols affecte la courbe de réponse du site à partir d'accélérations modérées (75 cm/s^2 en fond de puits).
- que la caractérisation non-linéaire d'un site, en vue de l'évaluation des effets de sites, pourrait être réduite à la caractérisation des couches de sol superficielles.

Cette dernière assertion peut avoir une influence importante pour la caractérisation non-linéaire des sites. Elle a été confirmée par l'inversion comparée des fonctions de transfert forage en régime linéaire et non-linéaire. Finalement, notre travail a également montré :

- que la profondeur à parti de laquelle le comportement non-linéaire des sols n'a plus d'influence sur la réponse du site dépend du site de l'intensité du mouvement sismique incident.
- que l'analyse de sensibilité de la colonne de sol ńà prioriż est un bon outil pour déterminer la résolution de l'inversion compte tenu des informations disponibles ainsi que les paramètres du sol contrôlant les pics d'amplification.
- que l'utilisation conjointe de l'analyse de sensibilité et de la comparaison des fonctions de transfert, sans étape d'inversion, pouvait être suffisante pour évaluer les profondeurs où le sol peut avoir un comportement non-linéaire important.

Abstract

Local geology can strongly affect seismic ground motion at the surface. These so-called site-effects can be evaluated either numerically by simulating the seismic wave propagation or empirically using earthquake recordings analyses or statistical correlations between site parameters and site effects. This thesis concerns the improvement of site effect evaluation and the analysis of the variability of the seismic site response. This work is mainly based on the analysis of earthquake recordings. First, we analysed the site response variability between sites for similar incident ground motion (weak motion) and then, the site response variability between events at one site caused by non-linear soil behaviour. Finally, we used an inversion method to find the depths where soil non-linear behaviour mostly occurs.

We analyzed various earthquake recordings from the KiK-net database in Japan (more than 46 000), which is composed of more than 688 surface/borehole instruments. The vertical array configuration allows the computation of the empirical site response in borehole condition. This database was chosen because of its large amount of instrumented sites located at sediments, its large amount of accelerometric data and the existence of characterisations of the shear and compressive wave velocity profiles down to the borehole depth.

The analysis of the seismic response variability per site classes indicated which parameters must be measured to improve the site-effect assessment. In the linear range, the parameters are the Vs profile gradient calculated down to 30 m (B_{30}) or 100 m depth (B_{100}) and the fundamental resonance frequency of the site (f_0). These additional parameters to Vs30 can be used to reduce the surface motion variability between-sites (such as in GMPEs). Besides, the analysis of inter-event site response variability caused by non-linear soil behaviour leaded to the relevant parameters for the analysis of site effects due to strong earthquakes: B_{30}, The maximal amplification (A_{pred}) and the associated frequency (f_{pred}). This analysis showed as well:

- that the PGA (Peak Ground Acceleration) is a relevant parameter for non-linear site effect assessment.
- that, whatever the site, non-linear soil behaviour affects the site response already for moderate solicitations (75 cm/s^2 at the down-hole station).
- that information on the non-linear soil behaviour of the superficial layers only is sufficient to fully assess the non-linear site response

This last conclusion may have a large impact for non-linear soil characterisation. It has been confirmed by inversion of linear and non-linear borehole site responses and comparison of the obtained Vs profiles. At the same time, the present work showed:

- that the depth from which non-linear soil behaviour has no influence on site response depends on the site and on the intensity of the seismic input motion.
- that careful attention needs to be taken when inverting data from vertical arrays. Sensitivity analyses are a powerful tool to evaluate the resolution of the inversion considering the available information and the soil parameters that are well solved during the inversion.
- That the combined use of sensitivity analysis with comparison of transfer function, without inversion, could be enough to assess the depth where the non-linear soil behaviour mostly takes place.

Remerciements

Dans l'Egypte ancienne, Theuth, ancien dieu créa les sciences et l'écriture; il alla présenter ses inventions au roi de l'Egypte en soumettant non seulement ses recherches mais leurs utilités et possibles effets. Le Roi écouta attentivement et, pour chacune de ses découvertes critiqua constructivement le discours de Theuth. En ce qui concerne l'écriture le roi tint ces propos : "Les discours écrits sont immuables, ils gardent le silence et ne peuvent se défendre; ils ne font pas distinctions entre les lecteurs, ne s'adaptent pas et ne répondent qu'une chose toujours la même". Le mythe de Theuth (Platon, 1933). Le discours n'est donc rien sans son père. Après cette lecture, si des questions et critiques vous submergent, ne blâmez pas le discours mais allez voir l'auteur.

C'est après quelques années de réflexions, au cours desquelles j'ai réalisé mes premiers travaux de recherche, que j'ai débuté ce travail. Je commencerai mes remerciements, en saluant les personnes qui m'ont encouragée dans cette voie. Je pense tout particulièrement à Fabian Bonilla auprès de qui j'ai réalisé mon stage de Master, à Jean-François Semblat qui a accepté la direction de cette thèse et j'ajoute mes collègues de l'ERA Risque Sismique et du laboratoire de Nice à cette époque, Anne-Marie Duval (sans qui cette thèse n'aurait pas vu le jour), Etienne Bertrand, Michel Pernoud et Jean-Philippe Devic qui ont permis la réalisation de ce projet. J'insiste et je remercie profondément, Jean-François Semblat, Fabian Bonilla et Etienne Bertrand, d'avoir assuré la co-direction de cette thèse et de m'avoir conseillée judicieusement dans les labyrinthes que peuvent parfois prendre les réflexions de recherche. Un dédicace spéciale à mes collègues au quotidien, Nathalie Glinsky, Diégo Mercerat et Philippe Langlaude, pour leur écoute et leurs conseils. Je remercie également, l'ensemble de mes collègues du laboratoire de Nice, de leur bonne humeur et de leurs aides. Merci aussi à la direction du CETE Méditerranée ainsi qu'aux collègues de l'IFSTTAR qui m'ont accueilli les bras ouverts à Paris.

De nombreuses personnes m'ont conseillée et m'ont transmis certaines de leurs connaissances. Je remercie Fernando Lopez-Caballero et Céline Gélis pour leur collaboration pendant le projet RAP (Réseau Accélérométrique Permanent) sur la variabilité des paramètres du sol et l'influence sur la réponse des sites. Je remercie Héloïse Cadet, avec qui j'ai étroitement collaboré sur les corrélations canoniques. Je remercie Aurore Laurendeau, Pierre-Yves Bard et Fabrice Cotton avec qui j'ai travaillé sur la base de données japonaise KiK-net. De ce travail de thèse sont nées de nombreuses collaborations qui vont se poursuivre dans des projets de recherche, je remercie la "Prenolin team" de la confiance qu'elle me porte (Alain Pecker, Pierre-Yves Bard, Fabrice Hollender, Fabian Bonilla et Etienne Bertrand).

Amis des quatre coins de la France et d'ailleurs, je vous remercie d'être ce que vous êtes, tout simplement. Je terminerai par remercier ma famille si attentive et dévouée. Leurs points de vue se sont révélés bien souvent de bonnes directions à suivre.

Résumé étendu

CONTEXTE

Le mouvement sismique ressenti en surface résulte de la combinaison des effets liés à la source sismique, au milieu de propagation et à l'amplification locale des ondes sismiques par la géologie de surface (effets de site lithologiques) ou la topographie (effets de site topographiques). Les effets de site lithologiques sont la conséquence du piégeage des ondes sismiques dans les couches de sol mou situées près de la surface et reposant sur un substratum rocheux plus rigide (Bard & Bouchon, 1985; Kramer, 1996). Localement, au cours d'un séisme, une grande partie de la variabilité du mouvement sismique (d'un site à un autre) est directement imputable aux effets de site (e.g. Riepl et al., 1998). L'observation de données sismologiques a mis en avant les amplifications très importantes résultant de ces effets de site. L'un des exemples les plus parlant est le séisme de Mexico en 1985 pour lequel des accélérations du sol ont dépassé les accélérations enregistrées sur la zone épicentrale à plus de 350 km. C'est pourquoi il est nécessaire de prendre en compte ces effets de site dans l'évaluation de l'aléa sismique. De plus, dans les relations empiriques de prédiction du mouvement sismique (GMPEs [2]), la variabilité est définie par l'écart-type (σ_{GMPE}) a une influence très importante lors de l'évaluation probabiliste de l'aléa sismique. Cette variabilité est divisée en deux termes: variabilité inter-évènements et inter-sites. La première définit la variabilité du mouvement sismique en surface d'un évènement à un autre (ayant un couple magnitude/distance épicentrale équivalent). La seconde concerne la variabilité du mouvement sismique en surface entre les sites pour un même événement. Cette variabilité est intimement liée aux effets de site (Al Atik et al., 2010; Rodriguez-Marek et al., 2011). Ainsi, une évaluation robuste des effets de site est nécessaire pour la gestion et la prévention du risque sismique ou lors de l'établissement des équations prédictives du mouvement sismique tenant compte des conditions de site (e.g., Cotton et al., 2006; Douglas, 2006; Douglas et al., 2009).

Les effets de site peuvent être évalués de différentes manières: empiriquement, via l'analyse d'enregistrements de séismes (englobant des méthodes spécifique à un site ou à une classe de

[2] GMPEs: Ground Motion Prediction Equations

site) ou numériquement à l'aide de modèles numériques intégrant des comportements simples ou plus complexes du sol ainsi qu'une géométrie de site 1D, 2D ou 3D. L'utilisation des méthodes empiriques est limitée à une étude ponctuelle et dans des pays à faible sismicité, aux enregistrements de séismes faibles mais ne nécessite pas de connaissance préalable des paramètres du site. L'observation du mouvement sismique et plus précisément de la réponse des sites a montré qu'en un site donné, même pour des séismes ayant un couple magnitude/distance équivalent, la courbe de réponse des sites (courbe d'amplification dans le domaine fréquentiel), bien qu'ayant une allure similaire, est variable d'un séisme à un autre. Ainsi, la réponse sismique empirique des sites doit être mesurée en utilisant un nombre significatif d'enregistrements. En outre, la courbe d'amplification peut-être très modifiée lorsque l'on considère des séismes ayant induit au sein des matériaux des déformations fortes et donc un comportement non-linéaire.

Les méthodes numériques, selon leur complexité, requièrent une connaissance précise de la géométrie et des paramètres géotechniques du sol. Toutefois, la description géotechnique détaillée du site est bien souvent limitée ou évaluée avec une incertitude importante (liée à des erreurs de mesures et à la variabilité spatiale des paramètres du sol).

Pour pallier à ces limitations, des méthodes intermédiaires, basées sur des corrélations entre les paramètres du sol et du séisme et la réponse du site empirique sont attrayantes. Ainsi, dans les réglementations parasismiques européennes (EC8) et américaines (NERPH), les effets de site sont pris en compte via le paramètre Vs_{30} (Vitesse moyenne de propagation des ondes de cisaillement sur les 30 premiers mètres). Ces méthodes intermédiaires sont applicables sur une large échelle (commune, départements...). Certains auteurs ont même proposé des corrélations entre la pente du relief et Vs_{30} ou directement les effets de site à des échelles régionales (Wald & Allen, 2007; Allen & Wald, 2009). Néanmoins, ces méthodes ne sont pas suffisamment site-spécifiques et peuvent résulter en une évaluation très grossière des effets de site.

Cette thèse s'inscrit globalement dans la thématique de l'amélioration de l'évaluation des effets de site. Afin de réaliser une évaluation pertinente de ces effets, il est nécessaire de calculer non seulement l'amplification moyenne du site mais aussi la variabilité associée. L'objectif principal de ce travail est de comprendre et de quantifier la variabilité de la réponse des sites en suivant la même décomposition que proposée dans les GMPEs, i.e. inter-sites et inter-évènements. Les évènements sismiques faibles à moyens, bien qu'induisant dans les matériaux et les structures des déformations faibles, ont un taux de récurrence très élevé et doivent par conséquent être pris en compte. D'un autre coté, les évènements sismiques forts, bien que rares, doivent également être considérés compte tenu de leur potentiel destructeur. C'est pourquoi, dans cette étude nous avons considéré à la fois le comportement linéaire (dans le cas de séisme faibles) et non-linéaire (dans le cas de séismes forts) des sols.

La thèse est organisée en trois parties. Dans la première partie, les notions fondamentales sur

la propagation des ondes sismiques sont définies et la base de données accélérométriques utilisée au cours de ce travail est décrite en détail. Dans la seconde partie, les résultats de l'étude statistique sur la variabilité de la réponse sismique empirique des sites sont présentés et, dans la dernière partie, l'inversion de données de forage est réalisée afin de caractériser l'évolution des propriétés élastiques des matériaux lors de séismes forts.

Dans la suite de ce résumé étendu nous allons présenter chacune des parties en détaillant plus précisément les parties 2 et 3 qui constituent les principaux résultats de la thèse.

PREMIÈRE PARTIE: PRÉSENTATION DES NOTIONS FONDAMENTALES ET DE LA BASE DE DONNÉES ACCÉLÉROMÉTRIQUES

Le premier chapitre de la première partie consiste en une description de la propagation des ondes sismiques dans un milieu stratifié, unidimensionnel et idéalisé (homogène, isotrope et visco-élastique) sous faibles déformations afin d'introduire le calcul de la réponse sismique des sites 1-D linéaire. La formule analytique de la fonction de transfert forage (rapport du champ d'onde sismique en surface par le champ d'onde sismique total à la station en fond de puits) pour le cas d'un site mono-couche a été définie et analysée. Si la station en fond de puits est située à l'interface entre les sédiments et le substratum sismique, la fonction de transfert forage est indépendante des propriétés du substratum. Par ailleurs, des sites ayant un rapport Vs (des sédiments) sur l'épaisseur des sédiments similaires, conduisent à des fonctions de transfert forage identiques. Ces deux remarques ont des implications très importantes concernant l'inversion des fonctions de transfert forage empiriques.

Le logiciel utilisé dans ce travail pour le calcul de la réponse numérique sismique des sites linéaire 1-D a été amélioré afin de minimiser l'écart entre les observations et les simulations. La particularité et l'inconvénient majeur des enregistrements de séismes en forage réside dans le fait que la station en fond de puits n'enregistre pas seulement les ondes incidentes mais aussi les ondes descendantes issues des réflexions multiples sur les différentes interfaces entre les couches de sol et la surface libre. Si la station en fond de puits se trouve juste au niveau de l'interface entre les sédiments et le substratum (dans le cas d'une configuration de site très simple) les fréquences de résonance correspondent à celles de la fonction de transfert du site classique (site sur référence en surface) mais si elle est située juste au-dessus ou en dessous de cette interface des pics de pseudo-résonance, provenant de l'interférence destructive entre les ondes incidentes et descendantes, seront présents dans la fonction de transfert forage et masqueront les fréquences de résonance du site. Par ailleurs, la comparaison entre les observations et les calculs numérique a montré que l'amplitude de ces pics n'est pas correctement calculée. Ainsi, au logiciel initial nous avons ajouté les mêmes traitements que ceux réalisés sur les données empiriques (même lissage) et même pas d'échantillonnage en fréquence ainsi que l'implémentation de l'atténuation dépendant de la fréquence. L'étude de sensibilité de la réponse sismique des sites aux paramètres d'entrée

sur des cas synthétiques a montré que les fréquences des pics de résonance étaient principalement contrôlés par la vitesse et l'épaisseur des couches sédimentaires et que les amplitudes associées dépendaient fortement du paramètre de lissage et de l'atténuation, laissant peu d'influence à la densité des matériaux.

Le second chapitre de cette partie est consacré à la présentation de la base de données accélérométriques utilisée dans cette étude. Nous avons choisi la base de données accélérométriques japonaise KiK-net. Elle est composée de 688 sites ayant des capteurs tridimensionnels en surface et en fond de puits, ce qui permet de calculer, en chacun des sites, les fonctions de transfert forage empiriques. De nombreux enregistrements sont disponibles allant de faibles accélérations à de très fortes. En chacun des sites où un profil de vitesse (S et P) était disponible, nous avons calculé la courbe de réponse du site forage pour plus de 46 000 enregistrements. Nous avons caractérisé la réponse linéaire en calculant la moyenne et la limite de confiance à 95 % de la courbe de réponse du site ($BFSR_{lin}^{95}$) ainsi que le rapport spectral des composantes horizontales par la composante verticale des enregistrements en surface (HV_{lin}^{95}) pour les séismes ayant un PGA à la station en fond de puits inférieur à 10 cm/s^2. Compte tenu de cette définition, les hypothèses suivantes sont la base des travaux réalisés sur la base de données:

- (1) $BFSR_{lin}^{95}$ est calculé en utilisant les enregistrements de séismes dont les PGA en fond de puits sont faibles; le comportement non-linéaire ne devrait pas affecter la courbe de réponse et cette évaluation empirique peut être comparée au calcul numérique linéaire.

- (2) $BFSR_{lin}^{95}$ est calculé en utilisant un nombre important d'enregistrements de séismes provenant de diverses sources sismiques. Ainsi, la variabilité de la courbe de réponse contient déjà la variabilité qui pourrait être causée par une configuration complexe du site. Ainsi, l'écart entre $BFSR_{lin}^{95}$ et BFSR calculé pour des séismes forts, peut-être associé au comportement non-linéaire des sols.

Finalement, nous avons défini les paramètres caractérisant le site, la courbe de réponse linéaire et le signal sismique incident, pertinents pour l'étude des effets de site. L'analyse de la distribution de ces paramètres sur l'ensemble de la base de données confirme que cette base est pertinente pour l'étude statistique des effets de sites et représentative d'une grande variabilité de configurations de sites. Par ailleurs, pour l'analyse du comportement non-linéaire des sols, plus de 50 sites ont enregistré au moins deux séismes dont le PGA en fond de puits est supérieur à 50 cm/s^2.

DEUXIÈME PARTIE: ANALYSE DE LA VARIABILITÉ DE LA RÉPONSE SISMIQUE DES SITES

La seconde partie est divisée en deux chapitres. Dans le premier chapitre, nous avons étudié la variabilité des profils de vitesse et de la courbe de réponse des sites linéaires par classe de sol (étude de la variabilité inter-sites). Dans le second chapitre, nous étudions la variabilité de la réponse sismique des sites en un site dû aux effets du comportement non-linéaire des sols (variabilité inter-évènements).

Vs_{30} est un paramètre standard permettant de classifier les sols selon leur réponse sismique. Plus Vs_{30} est important plus le matériau est rigide et moins d'amplification est attendue. Néanmoins, ce paramètre ne tient pas compte de la complexité du profil de vitesse des ondes de cisaillement en profondeur (au-delà de 30 m). Certains sites ayant des Vs_{30} similaires peuvent avoir des profils de vitesse très différents en profondeur et même sur les 30 premiers de sol. Dans ce chapitre nous proposons d'analyser l'influence du profil de vitesse en profondeur sur l'amplification du site linéaire. Hormis Vs_{30}, nous proposons un paramètre supplémentaire caractérisant le site, i.e. le gradient du profil de vitesse, calculé à partir du profil de Vs de la surface jusqu'à des profondeurs de 30 et 100 m (appelés B_{30} et B_{100}). Ce gradient caractérise l'évolution du profil de vitesse avec la profondeur: un gradient faible est relatif à des profils de vitesse ayant une évolution faible avec la profondeur alors qu'un gradient fort caractérise des sites ayant un contraste d'impédance marqué dans les couches de sol proches de la surface. En outre, la fréquence de résonance fondamentale du site (f_0) a également été utilisée. Des études antérieures (Cadet, Bard, Duval, et al., 2012; Luzi et al., 2011) ont montré qu'il s'agissait d'un paramètre pertinent pour l'évaluation des effets de site qui, de surcroit, est calculé à partir de la courbe de réponse du site et donc issue d'une source d'information indépendante du profil de vitesse.

Les hypothèses majeures de l'analyse réalisée dans ce chapitre sont: (1) la courbe de réponse du site n'est pas affectée par les effets non-linéaires. (2) La réponse des sites n'est pas affectée significativement par des effets de configuration du site complexes (effets 2-D ou 3-D). Afin de satisfaire les hypothèses précédentes, nous calculons la courbe moyenne de la réponse du site linéaire en utilisant les enregistrements de séismes pour lesquels l'accélération maximale (PGA) à la station située en fond de puits soit inférieure à 10 cm/s^2. En outre, parmi les 668 sites pour lesquels un profil de Vs était disponible, 351 sites ont été sélectionnés compte tenu de la proximité de la réponse sismique du site empirique avec la réponse numérique 1-D linéaire.

Nous avons analysé la variabilité des profils de Vs et des courbes de réponse du site empirique linéaire selon le Vs_{30}, le gradient et la fréquence de résonance fondamentale. Nous avons constaté que le gradient calculé jusqu'à 30 m (B_{30}) permet de séparer, pour une gamme de Vs_{30} donné, les sites ayant des contrastes d'impédances forts près de la surface. Il permet également de mieux caractériser la réponse des sites, qu'en utilisant uniquement Vs_{30}. Le gradient calculé à partir du

profil de Vs jusqu'à des profondeurs plus importantes de 100 m (B_{100}) et la fréquence de résonance fondamentale (f_0) permettent de séparer les sites, ayant une gamme de Vs_{30} similaire, en trois types: sites sédimentaire profonds, sites moyens et sites ayant un fort contraste d'impédance près de la surface. Les courbes de réponse du site sont significativement différentes pour ces trois groupes de sites indiquant que le profil de vitesse en profondeur a une influence certaine sur la réponse du site. L'utilisation combinée de Vs_{30}, du gradient et de la fréquence de résonance fondamentale du site permet d'améliorer la classification des sites en mettant en avant l'influence du profil de vitesse en profondeur sur la courbe de réponse du site.

Dans le second chapitre de cette partie, nous analysons la variabilité de la réponse sismique des sites, d'un évènement à un autre, en un site. Nous nous intéressons aux effets du comportement non-linéaire des sols sur la réponse des sites. L'hypothèse majeure de ce chapitre est de considérer que la variabilité de la réponse du site imputable à la complexité géométrique du site et aux effets de sources sismiques est entièrement contenue dans la caractérisation linéaire de la réponse du site. Ainsi, les modifications, entre la courbe de réponse du site calculée à partir d'enregistrements de séismes forts par rapport à la caractérisation linéaire, sont causées par le comportement non-linéaire du sol. Afin de décrire les effets du comportement non-linéaire sur la réponse du site par événements, nous proposons le pourcentage de modifications de la courbe de réponse par rapport à la caractérisation linéaire (PNL_{ev}, Pourcentage de NonLinéarité) et le décalage de la fréquence des pics d'amplification (Sh_{ev}). Ces paramètres sont utilisés afin d'estimer la probabilité que la courbe de réponse non-linéaire soit significativement différente de la courbe de réponse linéaire. Nous avons constaté que, quel que soit le site considéré, cette probabilité est importante (> 20 %) même pour des niveaux d'accélération relativement faibles à la station en fond de puits (entre 30 et 75 cm/s^2). Ce qui indique que les effets du comportement non-linéaire des sols doivent être pris en compte même lors d'évaluations de la réponse sismique des sites pour des évènements modérés.

Par ailleurs, en 54 sites KIK-net ayant enregistré au moins deux séismes avec des accélérations maximales importantes (PGA à la station en fond de puits $>$ 50 cm/s^2), nous avons défini 4 paramètres supplémentaires site-spécifiques qui caractérisent les effets du comportement non-linéaire des sols sur la réponse du site: (1) un PGA seuil site-spécifique (PGA_{th}), PGA à partir duquel le PNL_{ev} est supérieur à 10%, (2) un PNL site-spécifique pour un PGA incident de 50 cm/s^2 (PNL_{site}), (3) un décalage de la fréquence site-spécifique pour un PGA incident de 50 cm/s^2 (Sh_{site}) et (4) une fréquence de la réponse du site à partir de laquelle on observe une dé-amplification de la courbe de réponse non-linéaire par rapport à la courbe de réponse linéaire (fNL). Sur la majorité des 54 sites ayant un comportement non-linéaire, nous avons observé que le comportement non-linéaire des sols augmentait le niveau de l'amplification pour des fréquences en-dessous de la fréquence fNL (compte tenu du décalage des fréquences des pics de résonance vers les basses fréquences). L'analyse des corrélations entre les paramètres non-linéaires site-spécifiques et les

paramètres caractérisant le site (Vs_{30}, B_{30}, f_0, f_{pred} et A_{pred}[3]) a montré que:

- (1) fNl se trouve entre la fréquence de résonance fondamentale du site et la fréquence prédominante lorsque ces dernières sont différentes,
- (2) les sites ayant des contrastes d'impédance prés de la surface (gradient important) et les sites ayant des effets de sites marqués (A_{pred} important) sont plus susceptibles d'avoir un comportement non-linéaire important.

Ces observations semblent indiquer que les couches de sol associées à la fréquence de résonance fondamentale i.e les couches de sol ayant le contraste d'impédance le plus en profondeur, ont un comportement moins non-linéaire que les couches de sol superficielles associées à la fréquence de résonance prédominante. Ainsi, il semblerait que le comportement non-linéaire du sol affectant la courbe de réponse du site se produise majoritairement dans les couches près de la surface. Par ailleurs, nous avons également montré que fNL et le décalage de la fréquence site-spécifique (Sh_{site}) peuvent être obtenus de façon similaire sur les fonctions récepteurs en surface. Cette dernière remarque montre qu'une partie des résultats obtenus dans cette étude peut être étendue à d'autres bases de données ayant uniquement un capteur en surface.

TROISIÈME PARTIE: INVERSION DE LA RÉPONSE SISMIQUE DES SITES

Dans le chapitre précédent, nous avons montré que l'effet non-linéaire semble se produire principalement dans les couches de sol proches de la surface. Cette conclusion peut avoir de très fortes implications quant à la caractérisation non-linéaire d'un site. L'étape suivante est de déterminer plus précisément les profondeurs auxquelles cet effet non-linéaire prend majoritairement place et à partir desquelles son influence sur le calcul de la réponse non-linéaire n'est plus significative. Dans cette partie nous inversons les courbes de réponse forage afin de trouver les propriétés élastiques équivalentes lors de séismes forts. Les objectifs principaux de l'inversion sont d'obtenir les propriétés élastiques (profils de Vs et de Q) à partir de l'inversion des courbes de réponse linéaire et de comparer les résultats avec ceux obtenus lors de l'inversion de courbe de réponse non-linéaire. L'inversion est réalisée à l'aide d'un algorithme hybride entre une méthode de recuit simulé (Kirkpatrick et al., 1983; Cerny, 1985) et une méthode de Nelder-Mead (Nelder & Mead, 1965), connue sous le nom de "downhill simplex method". Le logiciel utilisé a été développé par Liu Pengcheng (Liu et al., 1995). L'intérêt de combiner ces deux approches est d'associer l'efficacité d'une méthode de recherche directe (Nelder-Mead) à la faculté de trouver le minimum global d'une fonction coût complexe d'une méthode stochastique (recuit simulé).

Nous avons sélectionné 4 sites KiK-net avec différents profils de vitesse pour appliquer notre approche. Tout d'abord nous avons réalisé une étude de sensibilité sur les profils de vitesse des

[3] A_{pred} Amplification maximale et f_{pred} fréquence d'amplification maximale

sites sélectionnés afin de déterminer l'influence de chacun des paramètres d'entrée sur la courbe de réponse et de déterminer quels paramètres pouvaient être contraints par l'inversion. L'inversion des BFSR$_{lin}$ a mis en avant la difficulté d'interpréter les résultats même si des informations a priori fiable sur le profil de vitesse étaient disponibles. Cette difficulté peut s'expliquer d'une part, par le fait que les simulations numériques ne reproduisent pas correctement l'amplitude des pics de pseudo-résonance (malgré l'implémentation de l'atténuation dépendant de la fréquence) et d'autre part, par le fait que les sites sélectionnés n'ont peut-être pas tous une configuration 1-D, bien que le mode de résonance fondamentale soit en accord avec les résultats de la simulation numérique 1-D.

Néanmoins, au site FKSH14, l'inversion de BFSR$_{lin}$ a donné des résultats très cohérents avec le profil de vitesse et le log géologique. Nous avons réalisé des inversions des courbes de réponse non-linéaires issues de 3 séismes enregistrés par cette station. Nous avons constaté que les couches superficielles étaient les plus affectées par les effets non-linéaires (couches situées au-dessus du plus fort contraste d'impédance du profil de vitesse). Néanmoins, durant le plus fort séisme, l'inversion a indiqué des modifications dans une couche de sol plus profonde. Cette observation indique que la nonlinéarité prend place principalement dans les couches de sol superficielles, conformément aux conclusions de la partie précédente, mais que, pour des mouvements sismiques plus forts, elle pourrait se produire également dans des couches plus profondes et avoir un impact sur la courbe de réponse. La profondeur à laquelle la nonlinéarité prend place et influence la courbe de réponse dépendrait non seulement du site mais aussi de l'amplitude de la sollicitation. Cette observation, limitée à l'étude d'un site, ne peut être généralisée. Elle est toutefois en accord avec les conclusions de la partie précédente, issues de l'analyse statistique sur 54 sites.

ANNEXES

En annexe, nous détaillons les effets des ondes descendantes lors du calcul de la réponse du site en considérant la station de référence en fond de puits. Compte tenu de l'effet des ondes descendantes, la courbe de réponse fond de puits n'apparait pas comme la meilleure donnée pour identifier la fréquence de résonance fondamentale du site. Alors, nous avons utilisé une méthode alternative ne nécessitant pas de station de référence appelée H/V séisme qui est le rapport spectral des composantes horizontales sur verticale des enregistrements en surface.

PUBLICATIONS

Le travail présenté dans cette thèse a fait l'objet de trois publications (deux acceptées et une soumise) dans des revues scientifiques avec comité de lecture concernant la seconde partie de ce travail.

- L. F. Bonilla, K., Tsuda, N., Pulido, J., Régnier and A. Laurendeau. Nonlinear site response

evidence of K-NET and KiK-net records from the Mw 9 Tohoku earthquake. Earth Planets Space, 58, 1-7, 2011.

- J. Régnier, H. Cadet, L.-F. Bonilla, E. Bertrand, J.-F. Semblat. Assessing nonlinear behavior of soils in seismic site response:
  Statistical analysis on KiK-net strong motion data. Bulletin of Seismological Society of America, accepted in february 2013.

- J. Régnier, L.-F. Bonilla, E. Bertrand, J.-F. Semblat. Influence of the velocity profile at depth on the linear characterization of site effects: Tests on the KiK-net database, Bulletin of Seismological Society of America, submitted in january 2013.

Ce travail a été également présenté dans trois conférences internationales et trois nationales dont 4 avec actes de conférence:

- J. Régnier, L.-F. Bonilla, E. Bertrand, J.-F. Semblat. Vs30 : un paramètre pertinent pour décrire les effets de site ? Tests sur la base de données Kiknet. AFPS, Paris, 2011.

- J. Régnier, L.-F. Bonilla, E. Bertrand, J.-F. Semblat. Empirical evidence of nonlinear site response at several KiK-net stations. ESG4, Santa-Barbara, US, 2011.

- J. Régnier, L.-F. Bonilla A.M. Duval , .F. Semblat et E. Bertrand. Revisiting Vs30 as a proxy parameter for sit effects: a case of study using KiK-net data. 5ICGE (Chili), 2011.

- J. Régnier, L.-F. Bonilla, E. Bertrand, J.-F. Semblat. Empirical evidence of soil non-linear behavior effects on seismic site response, WCEE15, Lisbone, 2012.

Notation

GENERAL NOTATION

BFSR: Borehole Fourier Spectral Ratio (can be computed in this study numerically or empirically).

BFSR$_{lin}$**:** Mean linear empirical BFSR.

BFSR$^{95}_{lin}$**:** 95% confidence limit of the linear empirical BFSR.

BFSR$_{num}$**:** Numerical linear BFSR.

BFSR$_{ref}$**:** referent BFSR (either BFSR$_{lin}$ or BFSR$_{num}$)

b$_{konno}$**:** Smoothing parameter.

Fc: Cost Function.

OFSR$_{num}$**:** Numerical linear Outcrop Fourier Spectral Ratio (calculated numerically only in this study).

HV$_{lin}$**:** Mean linear Horizontal to Vertical spectral ratio at the surface sensor.

σ**:** Stress matrix

ϵ**:** Strain matrix

u: Displacement vector

u$_r$**:** Displacement wave-filed in the substratum

u$_s$**:** Displacement wave-filed in the sediment

k: Wave number.

SOIL PARAMETERS

B_{30}: Gradient of the Vs profile down to 30 m depth

B_{100}: Gradient of the Vs profile down to 100 m depth

G, G_0: Shear modulus

G_1: Viscosity

Q: Quality factor

Th: Thickness of the soil layers

Vs: Shear wave velocity

Vs_{30}: Harmonic mean shear wave velocity

α**:** Parameters for the attenuation frequency dependent

ρ**:** Density or volume mass

λ**:** First lame constants

LINEAR SITE RESPONSE PARAMETERS

f: Frequency vector.

f_0**:** Fundamental resonance frequency

f_{pred}**:** Predominant resonance frequency

A_{pred}**:** Amplitude of BFSR at the predominant resonance frequency

INCIDENT MOTION PARAMETERS

arms: Acceleration root mean square

CAV: Cumulative Absolute Velocity

Dtri: Trifunac Duration

Ia: Arias intensity

Fc: Central Frequency

PGA: Peak Ground Acceleration

PGV: Peak Ground Velocity

PGD: Peak Ground Displacement

NONLINEAR PARAMETERS PER EVENT

PNL$_{ev}$: Percentage of NonLinearity per event

Sh$_{ev}$: Shift of the predominant frequency per event

NONLINEAR PARAMETERS PER SITE

PGA$_{th}$: Threshold PGA for nonlinearity

PNL$_{site}$: Percentage of NonLinearity site-specific

Sh$_{site}$: Shift of the predominant frequency site-specific

fNL: Frequency of NonLinearity: Frequency from which the ratio between the linear and non-linear site response goes below 1.

Contents

General introduction

The Earth is in perpetual movement, which manifested itself through natural disasters including earthquakes, volcanic eruptions, landslides, droughts, floods, cyclones and fires representing the appearances of this phenomena that threaten the human population. The seismic risk is the combination of the seismic hazard and the vulnerability of the local infrastructure. Therefore, even for equivalent initial hazard levels, the impact of these natural disasters is heterogeneous, and will be greater in developing countries where the building stock is more vulnerable. For now, there is no consensus in the seismological community about reliable seismological precursors even for short-term earthquake occurrences. The only tool available for seismic risk mitigation is therefore prevention; i.e., the definition of a reliable seismic hazard map and the construction of structures that are adapted to this hazard.

The seismic motion recorded at the surface come from the seismic waves that are generated by a seismic source. These are modified according to the travel path from the source to the site and to the local site configuration, the so-called "site effects". These site effects are the consequence of seismic waves being trapped in soft subsurface material that overlies stiffer rock (Bard & Bouchon, 1985). It is widely recognised that site effects can dramatically increase the seismic motion at the soil surface. Consequently, a large part of the local variability of ground motion is associated to site effects. In ground-motion prediction equations (GMPEs), the variability is described by the standard deviation (σ_{GMPE}), and this can have a great influence on the evaluation of the probabilistic seismic hazard. The variability is divided into two main terms: the between-events variability; and the between-sites variability. The first of these describes the variability of the seismic ground motion at the surface between seismic events. The second describes the inter-site variability of the ground motion for a given couple-magnitude distance. This variability is intimately related to the site effects (Al Atik et al., 2010; Rodriguez-Marek et al., 2011). Consequently, robust evaluation of site effects is required for risk mitigation and urban planning, or when computing GMPEs at sites with different soil conditions (e.g., Cotton et al., 2006; Douglas, 2006; Douglas et al., 2009).

Site effects can be evaluated in different ways: empirically, using the analysis of earthquake recordings; or numerically, by simulating the wave propagation. The main advantage of the empir-

ical method is that it does not require knowledge of the soil properties and the site configuration, although it is limited to local studies, and in areas of low seismicity, to the analysis of weak motion. Observations of site responses show that, in one site, even for similar magnitude and epicentral distance couple, although having a similar aspect, the site response curves may be variable from one event to another (in the frequency domain). Thus, empirical site response must be performed using a significant number of recordings. Besides, the site response may be very different when considering strong seismic motions that induce large deformations in the soil layers.

The simulations can integrate simple or more sophisticated soil behaviours and geometry, one-dimensional [1-D], 2-D or 3-D site configurations, (e.g. Semblat & Pecker, 2009). Depending on their sophistication, numerical methods require detailed geotechnical and geometrical data. The detailed geotechnical description of a site is often limited and has to be evaluated with uncertainties (linked to measurement errors or spatial variability).

To compensate for these limitations, it has been proposed to use intermediate solutions based on correlations between simplified soil and earthquake parameters with site effects or surface ground motion. Nowadays, the Vs_{30} (harmonic mean shear-wave velocity in the first 30 m depth of the soil) is widely used to classify the soils for regulation codes, such as with Eurocode 8 (EC8) and the National Earthquake Hazard Reduction Programme (NEHRP). In addition, ground motion simulations in California, USA, that take site effects into account involve amplification functions that are based on Vs_{30} and rock peak ground motion (Graves et al., 2008). In Europe, Cadet, Bard, Duval, et al. (2012) developed amplification functions that are based on Vs_{30} and the soil fundamental frequency (f_0). Such intermediate approaches can be applicable at large scales but are not site-specific, and in some sites they may not provide accurate evaluations of site effects.

The present study is designed to provide improvements to site-effect evaluations. The basic idea of this study is to quantify and understand the sources of the seismic site response variabilities, both between sites and between events. Although small to medium strength earthquakes induce weak deformation in materials, they are the most common shaking that is felt by the infrastructure, and therefore they need to be taken into account for seismic hazard assessment. On the other hand, large earthquakes induce strong deformation, and although these are rare, they can have such a destructive potential that they absolutely need the be considered. This is why in this study we have considered both the linear and non-linear soil behaviour for site-effect assessment. The variability of the between-sites seismic site responses will be treated in the linear range (for weak motion only), and the between-events variability will concern the site response variability that is caused by the non-linear soil behaviour.

This thesis is organised as follows. In the first part, the fundamental concepts required for this work are defined and the database used is presented. The second part is mainly based on analyses of earthquake recordings, while the third part deals with inversions of site-response curves to find

the equivalent linear soil column that best-fits the empirical non-linear site response.

The first part of this thesis is divided into two chapters that essentially introduce the second and third parts. In the first chapter, we define the fundamental concepts of seismic wave propagation in elastic and visco-elastic media. To understand the numerical method used to compute the 1-D linear seismic site responses in the present study, we show the equations of propagation of the shear waves with vertical incidence in isotropic, homogeneous, layered, and visco-elastic media in the linear range (weak deformation). Then, we present the sensitivity analyses of the numerical method that were performed. The basic concepts presented in this first chapter represent the technical support for chapter one of the second part, which is concerned with observations of the between-sites seismic-site response variability in the linear range, and the third part, which deals with inversion techniques. The second chapter of the first part concerns the presentation of the KiK-net Japanese database (accelerometric and geotechnical) that was used, as well as the data selection and the treatments carried out to compute the empirical evaluation of the site responses in the second part.

The second part of this thesis concerns the analysis of the empirical data. It is divided into two chapters. The first deals with the analysis of the between-sites variability of the seismic-site response in the linear range. The relevancy of proxy parameters such as Vs_{30} is tested, and we have defined additional parameters that can improve site classification. The hypothesis of this first analysis is that the non-linear soil behaviour and complex site configuration (2-D, 3-D) do not affect the seismic-site response variability. Hence, we have selected sites from the KiK-net database for which the 1-D site configuration assumption is valid, and we have used recordings of weak ground motion. The second chapter focusses on the analysis of between-events site-response variability caused by non-linear soil behaviour. The main scope of this chapter is to find the incident motion, soil, and linear site-response parameters that are the most correlated to the non-linear soil behaviour and to the effects on the site response.

In the third part, inversions of the site-response curves are performed, to find the equivalent linear soil column that best fits the non-linear site response. The scope of this third part is to find the depths where the non-linear soil behaviour mostly takes place. The first chapter of this third part describes the search techniques that are used, and the second chapter shows the application of the inversion method to the empirical data at four selected KiK-net sites.

Part I

Basic concepts and data processing

Chapter 1

Linear wave propagation

Sommaire

INTRODUCTION

Earthquakes start from a sudden and brutal rupture in the earth crust that generate seismic waves. These body waves propagate from the seismic source and reach the soil surface or the foundations of the man-made infrastructure, within which they can be further propagated. While propagating in a medium, the seismic waves are attenuated with the distance. The first source of attenuation is linked to the material in which the waves propagate (intrinsic and scattering attenuation) while the second is associated to geometrical effects (geometrical spreading).

In the general introduction, we underlined that the seismic ground motion at the surface is the convolution of the seismic waves generated during an earthquake by the path traveled between the seismic source to the site and by the shallowest soil layers configuration and geotechnical properties (lithological and topographic site effects). The present study deals with the evaluation of lithological site effects. The scope of this first chapter is to introduce the theoretical calculation of these site effects. In this chapter, the wave propagation formulation adopted assumes the following 4 main hypotheses: (1) the propagation of plane waves is in one-dimension (during the propagation, the wavefronts are infinite parallel planes) with (2) vertical incidence. (3) The medium is composed by homogeneous and isotropic soil layers. (4) The soil behavior is, linear elastic or linear visco-elastic. The soil non-linear behavior will be presented in the next chapter.

First, we introduce the equation for the propagation of shear waves with vertical incidence in isotropic, homogeneous, elastic and visco-elastic media. We discuss the damping implementation and the analytical formulation of the borehole site response for a mono-layer soil column with respect to inversion purposes. Then, we discuss the relevancy and limits of the hypothesis assumed in this study. Finally, we describe the computational approach used to compute the 1-D site response with sensitivity analysis on some synthetic cases.

1.1 WAVE PROPAGATION IN LINEAR ELASTIC AND VISCO-ELASTIC MEDIA

1.1.1 The elastic-wave propagation equation

To fully describe the wave propagation into a medium, we have to express the stress and strain tensors and displacement vector, which in three dimensions means that there are 15 unknowns to be found. Using the equation of motion (Equation 1.1), the stress-strain (Equation 1.2) and the strain-displacement (Equation 1.4) relationships means that there are 15 equations to solve in three dimensions. The equation of motion is :

$$div(\boldsymbol{\sigma}) + \rho\mathbf{b} = \rho\frac{\partial^2\mathbf{u}}{\partial t^2}, \tag{1.1}$$

where, $\boldsymbol{\sigma}$ is the stress tensor, ρ is the mass density, $\mathbf{b}$ is the body-force vector, and $\mathbf{u}$ is the displacement vector. For weak deformation, soils have linear and elastic behaviours, which means that in the stress ($\boldsymbol{\sigma}$) - strain ($\boldsymbol{\epsilon}$) plane, the loading, unloading and re-loading paths are similar (Figure 1.8 (a)). Hooke's law, which linearly links ($\boldsymbol{\sigma}$) and ($\boldsymbol{\epsilon}$), illustrates this behaviour in isotropic medium:

$$\boldsymbol{\sigma} = \lambda(tr\boldsymbol{\epsilon})\mathbb{1} + 2G\boldsymbol{\epsilon}, \tag{1.2}$$

where, ϵ is the strain tensor, λ and G are the Lamé's constants (G being the shear modulus).

The seismic body waves are composed of the compressive waves (P-waves) and the shear waves (S-waves) with P-wave propagation velocity greater than that of the S-waves. The P-waves, which stands for 'primary' waves are also called the compressive waves, as suggested by their name, they move by compressing/dilating the material in the direction of propagation. The S-waves, or shear waves, move by shearing the material in a transverse plane to the direction of propagation (wave plane). Considering that more than 90% of the energy radiated from an earthquake is carried by the S-waves (Boatwright & Fletcher, 1984), we will consider only the propagation of S-waves. Also, for site-effect assessments (for the shallows soil layers), the vertical incidence of the incident waves is a reasonable hypothesis, as mentioned in the Introduction. Thus, in the following, we only consider SH-wave propagation into a 1-D medium along the vertical direction. As illustrated in Figure 1.1, the direction of motion of the SH-waves is along the X-axis in the wave plane, which is transverse to the direction axis (Z-axis).

The displacement vector $\mathbf{u}$ is therefore only a function of Z in the X-direction. This can be written as:

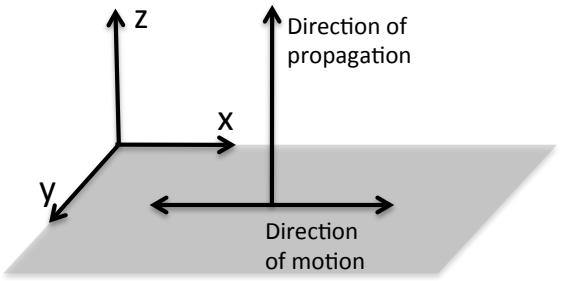


Figure 1.1: *Propagation of SH-waves in the Z-axis direction.*

$$\mathbf{u} = \begin{bmatrix} u_x(z) \\ 0 \\ 0 \end{bmatrix}. \tag{1.3}$$

The previous equations (1.1, 1.2, 1.4) can be simplified considering the previous hypothesis of 1-D SH-wave propagation, which for the strain-displacement relationship, gives:

$$\boldsymbol{\epsilon} = \frac{1}{2}[grad\ \mathbf{u} + (grad\ \mathbf{u})^t] = \begin{bmatrix} 0 & 0 & \frac{1}{2}\frac{\partial u_x}{\partial z} \\ 0 & 0 & 0 \\ \frac{1}{2}\frac{\partial u_x}{\partial z} & 0 & 0 \end{bmatrix}. \tag{1.4}$$

The trace of the strain tensor is zero, and Hooke's law becomes:

$$\sigma_{xz} = 2G\epsilon_{xz}. \tag{1.5}$$

The equation of motion that links the stress to the displacement, considering that the body-forces are zero, becomes:

$$\frac{\partial \sigma_{xz}}{\partial z} = \rho \frac{\partial^2 u_x}{\partial t^2}. \tag{1.6}$$

Replacing σ_{xz} from Equation 1.5 into Equation 1.6, this gives the wave equation for elastic linear soil behaviour given in Equation 1.7.

$$V_s^2 \frac{\partial^2 u_x}{\partial z^2} = \frac{\partial^2 u_x}{\partial t^2}, \tag{1.7}$$

where, $V_s = \sqrt{\frac{G}{\rho}}$ (G is the shear modulus, and ρ is the mass density) is the shear-wave

velocity.

1.1.2 Solution of the elastic wave equation

The solution of Equation 1.7, is a combination of the waves propagating at a velocity Vs, in opposite directions. It can be expressed as:

$$f(z,\omega) = e^{i\omega t}(Ae^{ikz} + Be^{-ikz}). \tag{1.8}$$

For vertical incidence and considering the positive direction of z from the surface to the depth, the amplitude A is associated to the upward propagating waves and the amplitude B to the downward propagating waves. The coefficient k is inversely proportional to a length and is called the wave number. In the elastic case, k is the ratio between the angular frequency and it's velocity:

$$k = \frac{\omega}{Vs}. \tag{1.9}$$

1.1.3 The visco-elastic wave equation

1.1.3.1 Background on seismic waves attenuation

Even for weak deformation, soft materials show hysteresis behaviour in the stress-strain plane (see Figure 1.8 (b)) and dissipation of energy. While pure elastic materials restore 100% of the energy due to deformation, viscoelastic or nonlinear materials do not restore 100% of the energy under deformation, but actually lose or dissipate some of this energy. The attenuation is usually divided in two terms i.e. the effective attenuation that is dependent on the material the wave propagate in and the geometrical attenuation that is associated to propitiation itself: when the body waves get further from the seismic point source, the wave plane become larger, the total energy being conserve in the whole wave plan, locally the amplitude of the waves is decreasing. The effective attenuation can be decomposed into:

Intrinsic attenuation : At a microscopic scale different process of energy dissipations have been proposed (e.g. Biot, 1956; Johnston et al., 1979; Leurer, 1997). The intrinsic attenuation can be associated to several individual mechanisms: In coarse grained soils, it is mostly caused by frictional losses between soil particles and fluid flow losses due to relative movements between the solid and the fluid phases. In fined grained soils, more complex phenomena take place mainly associated to electromagnetic interaction between water particles and macroscopic solid particles.

Scattering attenuation : in heterogeneous materials, the scattering of the waves will also be a source of energy dissipation.

The contribution of both scattering and intrinsic attenuation in the attenuation of short period S-Waves (0.08 - 0.1 s) has been widely discussed in the past (see e.g. Tselentis (1998) for an overview).

The attenuation can be represented by the quality factor (Q). Seismologists community used more commonly the quality factor, whereas geotechnical community usually use the damping ratio (ξ). Both factors are linked by the formula $Q = 1/(2\xi)$ It is widely recognized in the seismology community that seismic waves attenuation (here S-waves) in the earth's crust can be represented by the following equation:

$$A(r, f) = A_0 e^{-\pi f r/(Q_{ef} Vs)}, \tag{1.10}$$

where, A is the amplitude of the seismic waves, A_0 being the geometrical spreading factor, r is the distance, f is the frequency, Vs is the shear wave velocity and Q_{ef} is the effective quality factor of S-waves . shortciteauthorDainty-1981 (1981) gives an expression of the effective quality factor as a function of the intrinsic (Q_{in}^{-1}, inverse of the quality factor) that is frequency-independent and scattering attenuation Q_{sc}^{-1} that is frequency-dependent:

$$Q_{ef}^{-1} = Q_{in}^{-1} + Q_{sc}^{-1}. \tag{1.11}$$

For site-effect assessment, considering that the wave propagation is mostly plane waves, the wave attenuation is mainly control by the intrinsic and scattering attenuation rather than by the geometrical attenuation. In the following sections, the damping characteristic that will be used is the quality factor, denoted as Q, but refers to the effective quality factor aforementioned.

1.1.3.2 Modelisation of waves attenuation

One way of modeling the dissipation of energy is to introduce viscous behaviour, which means that the dissipation depends on the velocity (or strain rate). It is important to recall that the attenuation is caused by different internal mechanisms, and introducing viscosity in the soil behavior models explains only part of these phenomena. Different spring-Dashpot systems can be used to illustrate the viscous behaviour. The most common used are the Kelvin-Voigt and the Maxwell viscoelastic models, that are spring and dashpot in parallel or in series respectively (see (Ishihara, 1996)). The properties of a visco-elastic medium are as follow: (1) Hysteresis is observed in the stress-strain plane, (2) stress relaxation occurs: at constant strain, the stress decreases with time, (3) creep occurs: at constant stress, strain increases with time. The Kelvin-Voigt model is illustrated in Figure 1.2. The aforementioned Hooke's law for 1-D propagation of SH-waves can then be re-written to add the viscous behaviour. The stress is the sum of an elastic part that is proportional to

the strain and a viscous part that is proportional to the strain rate:

$$\sigma_{xz} = \sigma^1_{xz} + \sigma^2_{xz} = G_0\epsilon_{xz} + G_1\frac{\partial \epsilon_{xz}}{\partial t} = G_0\frac{\partial u_x}{\partial z} + G_1\frac{\partial^2 u_x}{\partial z \partial t}, \tag{1.12}$$

where, G_0 is the shear modulus and G_1 is the viscosity of the material. For visco-elastic media (Kelvin-Voigt model), the wave equation is described in Equation 1.13:

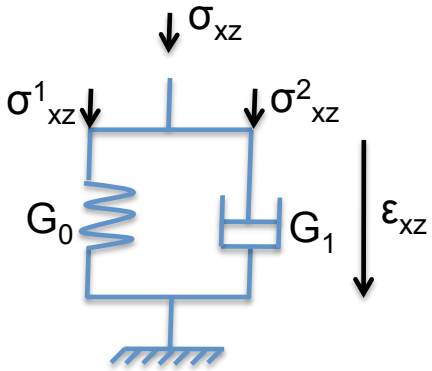


Figure 1.2: *Kelvin-Voigt rheological model.*

$$G_0\frac{\partial^2 u_x}{\partial z^2} + G_1\frac{\partial^3 u_x}{\partial z^2 \partial t} = \rho\frac{\partial^2 u_x}{\partial t^2}. \tag{1.13}$$

While this expression is slightly more complicated than the elastic wave propagation, substituting $u_x(z,\omega) = Ae^{i\omega.t}e^{-ikz}$, we find a similar expression to the elastic wave propagation equation such that:

$$G^*\frac{\partial^2 u_x}{\partial z^2} = \rho\frac{\partial^2 u_x}{\partial t^2}, \tag{1.14}$$

where, the complex shear modulus being $G^* = G_0 + i\omega G_1$, and the wave number is therefore complex and reads:

$$k^* = \frac{\omega}{\sqrt{(G_0 + i\omega G_1)/\rho}} \tag{1.15}$$

The solution of the visco-elastic wave equation is similar to the elastic case. The difference between elastic and visco-elastic linear wave propagation is that in the first case, the wave number is real, while in the second case it is complex. Considering $k^* = k_R - ik_I$, the solution of the visco-elastic wave equation is still a wave propagating with a velocity given by ω/k_R and with an

amplitude that is decreasing with the distance exponentially at a rate that depends on k_I.

1.1.4 Damping characteristics in the visco-elastic model (kelvin-Voigt type)

For an harmonic shear strain in the form of

$$\epsilon_{xz} = \frac{\partial u_x}{\partial z} = \epsilon_0 \sin(\omega.t), \tag{1.16}$$

where, ω is the angular frequency $2\pi f$ and ϵ_0 the maximal shear strain. Compared to the elastic case, the stress and strain are no longer in phase. The shear stress will be:

$$\sigma_{xz} = G_0\epsilon_0 \sin(\omega t) + \omega G_1 \epsilon_0 \cos(\omega t). \tag{1.17}$$

With this model, the stress-strain curve shown in Figure 1.8 (b) can be expected. We observe that the material shows hysteresis behaviour under cyclic loading, although the relation between σ and ϵ is linear. The damping characteristics can be represented by the quality factor (Q). Several definitions of Q, that lead to similar results, were proposed:

- In cyclic loading, Q is defined in equation 1.18 as initially proposed by Futterman in (1962).

$$Q = 2\pi \frac{W}{\Delta W}. \tag{1.18}$$

  It is defined as the ratio of the maximal energy that can be stored in a unit volume of a viscoelastic body (Equation 1.19), that could be represented in Figure 1.8 (b) by the grey triangle area:

$$W = \frac{1}{2} G_0 \epsilon_0^2, \tag{1.19}$$

  by the dissipated energy that is loss in one cycle (Equation 1.20), that is the area of the hysteresis loop in Figure 1.8 (b):

$$\Delta W = \pi G_1 \omega \epsilon_0^2. \tag{1.20}$$

- O'Connell and Budiansky (1978) have proposed a standard definition of the quality factor of the S-waves described in terms of the complex shear moduli as:

$$Q = \frac{\Re(G^*)}{\Im(G^*)}, \tag{1.21}$$

where, G^* is the complex shear modulus, $G^* = G_0 + i\omega G_1$. Hence, it is clear that both way of defining the quality factor lead to a similar result given in the following Equation (1.22):

$$Q = \frac{G_0}{\omega G_1}. \tag{1.22}$$

As mentioned in Ishihara (1996) pages 22 to 28, such models introduce a quality factor that is dependent of the frequency (inversely proportional).

1.1.5 Non-viscous model type Kelvin-Voigt

The frequency-dependence of the damping parameter has been widely discussed. A large number of laboratory experiments have shown that the damping properties are practically independent on the frequency at least in the frequency range of seismic loading (0.1- 30 Hz) (e.g. Hardin & Drnevich, 1972; Shibuya et al., 1995; Lo-Presti et al., 1997). Campbell (2009) and Morozov (2008, 2009) discussed the frequency dependence of the scattering attenuation which can be explained because part of the geometrical attenuation is accounted in the scattering attenuation.

Therefore, alternative models such as non-viscous type Kelvin-Voigt have been introduced. In this model, a rate independent dashpot is introduced. Such models do not have a physical basis, indeed viscous behavior should be rate dependent, but satisfy the laboratory observations. Although, this model overcomes the main shortcoming of the Kelvin-Voigt model it is important to recall that, with only one dashpot (one single relaxation time), this model is unable to simulate complex soil behavior phenomena. With analogy to the Kelvin-Voigt model, the Hooke's law for a non-viscous model that satisfy the frequency independent damping characteristics is written as:

$$\sigma_{xz} = (G_0 + iG_1)\frac{\partial u_x}{\partial z}. \tag{1.23}$$

This model is the one implemented in the method used in the present study. The damping characteristic, represented by the quality factor Q is given by:

$$Q = \frac{G_0}{G_1}. \tag{1.24}$$

The wave number, given in Equation 1.25 is still complex, but the denominator is independent of the frequency:

$$k^* = \frac{\omega}{\sqrt{(G_0 + iG_1)/\rho}}. \tag{1.25}$$

1.2 WAVE PROPAGATION IN A LINEAR VISCO-ELASTIC LAYERED MODEL

The propagation of waves in layered media involves the wave propagation into the layers and the propagation from one material into another. In the present study, we analyze the SH-wave propagation with vertical incidence. First we present the 1-D seismic response of a single-layer soil column, then we extend these results to multi-layer soil column.

1.2.1 Single-layer soil column response

1.2.1.1 Wave propagation

Let us take the example of a single-layer of sediment lying over a semi-infinite bedrock, such as depicted in Figure 1.3.

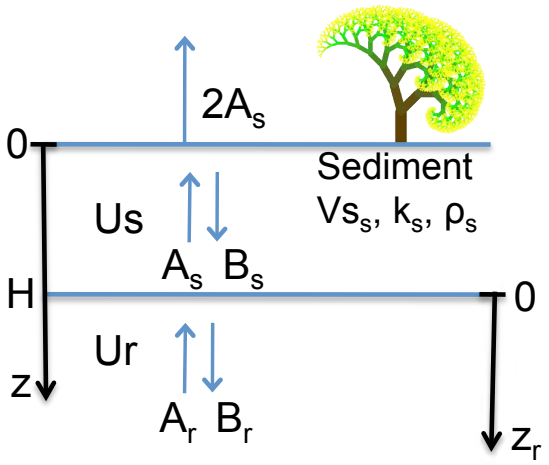


Figure 1.3: *Single-layer soil model. U_r is the substratum wave-field with A_r the amplitude of the up-going wave filed in the substratum (incident wave field) and B_r is the amplitude of the down-going wave filed in the substratum. U_s is the sediment wave-field with A_s, the amplitude of the up-going wave filed in the sediment and B_s the amplitude of the down-going wave filed in the sediment. H is the thickness of the sediment layer.*

In the bedrock, the displacement field is given by a combination of up-going and down-going harmonic and monochromatic waves:

$$u_r(z,\omega) = e^{i\omega t}(A_r e^{ik_r^* z_r} + B_r e^{-ik_r^* z_r}),$$

where, z_r is a specific reference frame to the bedrock (in Figure 1.3 z_r = 0 when $z = H$) and k^* being the complex wave number. We consider, $u_r^i = A_r e^{ik_r^*(Vs_r + z_r)}$ as the incident wave field.

Similarly, in the sediment layer, the displacement field is given by:

$$u_s(z,\omega) = e^{i\omega t}(A_s e^{ik_s^* z} + B_s e^{-ik_s^* z}).$$

To find the coefficients A_s, B_s, A_r and B_r, let us write the boundary conditions: (1) at the interface, there is continuity of both the displacement and the traction (i.e. the normal component of the stress tensor σ_{xz} as shown in Equation 1.5). Therefore, the displacement in the sediment and the bedrock must be equal at the interface: $\forall t\ u_r(z_r = 0, t) = u_s(z = H, t)$, this gives:

$$A_r + B_r = A_s e^{ik_s^* H} + B_s e^{-ik_s^* H}, \tag{1.26}$$

Similarly, the traction in the sediment and the bedrock must be equal at the interface: $\sigma_r(z_r = 0, t) = \sigma_s(z = H, t)$, this gives:

$$A_r - B_r = \frac{k_s^* G_s}{k_r^* G_r}(A_s e^{ik_s^* H} + B_s e^{-ik_s^* H}), \tag{1.27}$$

(2) The traction at the free surface is null: $\forall t\ \sigma_s(z = 0, t) = 0$. Using the wave-propagation equation in a visco-elastic medium (1.13), this gives:

$$A_s = B_s. \tag{1.28}$$

We define $C^* = k_s^* G_s / k_r^* G_r$ as the complex impedance contrast. From the following equations, we can find A_r and B_r as a function of A_s:

$$A_r = \frac{1}{2} A_s[(1 + C^*)e^{ik_s^* H} + (1 - C^*)e^{-ik_s^* H}], \tag{1.29}$$

$$B_r = \frac{1}{2} A_s[(1 - C^*)e^{ik_s^* H} + (1 + C^*)e^{-ik_s^* H}]. \tag{1.30}$$

1.2.1.2 Transfer function

We can define different transfer functions depending on the location of the reference site:

The outcrop site response is the ratio of the displacement wave field at the surface by twice the incident wave field at $z = H$ (or $z_r = 0$):

$$TF_{out} = \frac{u_s(z = 0, t)}{u_r^i(z_r = 0, t)} = \frac{2A_s}{A_r} = \frac{1}{cos(k_s^* H) + iC^* sin(k_s^* H)}. \tag{1.31}$$

The borehole-site response is the ratio of the wave field at the surface to the total wave field at the bottom of the soil column.

$$TF_{bor} = \frac{u_s(z=0,t)}{u_r(z_r=0,t)} = \frac{2A_s}{A_r + B_r} = \frac{1}{cos(k_s^* H)} \tag{1.32}$$

where, k^* is defined as in Equation 1.15 for the visco-elastic Kelvin-Voigt model of by Equation 1.25 for the non-viscous type Kelvin-Voigt model.

1.2.2 Multi-layer soil column response

The previous calculations can be generalised to a multi-layer soil column using the Haskell-Thomson method involving the propagation matrix (Thomson, 1950; Haskell, 1953). Let us consider one of the layer of this soil column with complex shear modulus (G) a shear wave velocity (V) and thickness (H). The lower boundary is denoted as 2 and the upper boundary as 1 as illustrated in Figure 1.4.

The governing equation is the wave propagation equation for the visco-elastic model given by Equation 1.13, with solution in the form of two waves propagating in opposite directions:

$$u_x(z,\omega) = e^{i\omega t}(Ae^{ikz} + Be^{-ikz}),$$

where k, the wave number is complex so is the shear modulus (for convenience, the complex terms (k and G) are not marked with a star as in the previous Section):

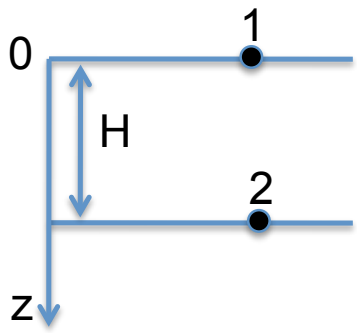


Figure 1.4: *One layer embedded in a multi-layer soil model.*

$$k = \frac{\omega}{\sqrt{(G_0 + i\omega G_1)/\rho}}, \tag{1.33}$$

$$G = (G_0 + i\omega G_1). \tag{1.34}$$

Considering the relationship between the stress and the displacement (Equation 1.12), the shear stress is:

$$\sigma_{xz} = ikG[e^{i\omega t}(Ae^{ikz} - Be^{-ikz})]. \tag{1.35}$$

At the upper boundary (1), for z = 0, we can write:

$$u_{x1} = (A + B)e^{i\omega t} \qquad \text{and} \qquad \sigma_{xz1} = ikG(A - B)e^{i\omega t}, \tag{1.36}$$

or in matrix form and ignoring the $e^{i\omega t}$ common term :

$$\begin{bmatrix} u_{x1} \\ \sigma_{xz1} \end{bmatrix} = \begin{bmatrix} 1 & ikG \\ 1 & -ikG \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix}. \tag{1.37}$$

Similarly at $z = H$ we can write:

$$\begin{bmatrix} u_{x2} \\ \sigma_{xz2} \end{bmatrix} = \begin{bmatrix} e^{ikH} & ikGe^{ikH} \\ e^{-ikH} & -ikGe^{-ikTh} \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix}.$$

Inverting the second system of equations and replacing the A and B matrix coefficients, we end up with:

$$\begin{bmatrix} u_{x2} \\ \sigma_{xz2} \end{bmatrix} = \begin{bmatrix} cos(kTh) & -kGsin(hTh) \\ \frac{1}{kG}sin(kTh) & cos(kTh) \end{bmatrix} \begin{bmatrix} u_{x1} \\ \sigma_{xz1} \end{bmatrix}. \tag{1.38}$$

The matrix above is called the Haskell-Thomson transfer matrix and relates the displacement and stress fields of the upper layer to the ones of the lower layer.

If we consider that the soil layer model is composed of N-1 layers of sediments overlying on a semi-infinite substratum such as illustrated in Figure 1.5, we can relate the displacement and stress fields at the top of each layer N_1 (with $n_1 \in$[1,N-1]) to the displacement and stress fields at the bottom of each layer N_2 above (with $N_2 < N_1$) by simply multiplying the Haskell Thomson transfer matrix of the layers in between N_1 and N_2.

Let us consider the bedrock elastic rather than visco-elastic, which is a reasonable assumption considering that the bedrock material is stiff and the seismic strain is low. The motion in the basement is characterised by a shear modulus G_B, a volume mass ρ_B and a wave number k_B that are real with $k_B = (\omega)/(\sqrt{G_B/\rho_B})$. The displacement in the basement is composed of an upward-travelling incident wave and a downward reflected wave such as illustrated in Figure 1.6:

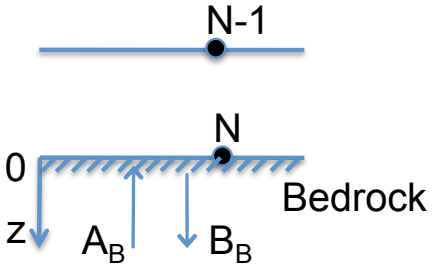


Figure 1.6: *Multi-layer soil model.*

$$u_{xB} = (A_B e^{ik_B(V_B t+z)} + B_B e^{ik_B(V_B t-z)}). \tag{1.39}$$

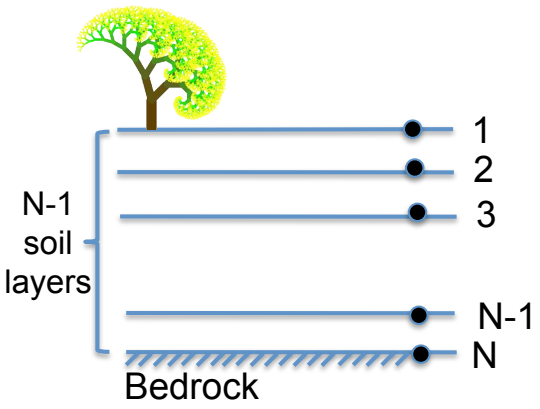


Figure 1.5: *Multi-layer soil model.*

The stress in the bedrock will be:

$$\sigma_{xB} = ik_B G_B (A_B e^{ik_B(V_B t+z)} - B_B e^{ik_B(V_B t-z)}). \tag{1.40}$$

At the interface between the bedrock and the deepest soil layer, assuming that z is equal to 0, the continuity conditions of the displacement and traction fields give:

$$u_{xN} = (A_B + B_B)e^{ik_B V_B t} \qquad \text{and} \qquad \sigma_{xzN} = ik_B G_B (A_B - B_B)e^{ik_B V s_B t}. \tag{1.41}$$

Let us consider that the incident wave field at the interface is given by the upwards wave amplitude, $u_I = A_B e^{ik_B V s_B t}$. Using the previous equation and eliminating the B_B term, we find:

$$ik_B G_B u_{xN} + \sigma_{xzN} = 2ik_B G_B u_I. \tag{1.42}$$

At the surface the traction vanishes, such that $\sigma_{x1} = 0$. Therefore, the displacement and traction at the surface are linked to the displacement and traction at the bottom of the soil column according to the following system of equations:

$$\begin{bmatrix} u_{xN} \\ \sigma_{xzN} \end{bmatrix} = \begin{bmatrix} u_{xB} \\ \sigma_{xzB} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{21} \\ B_{12} & B_{22} \end{bmatrix} \begin{bmatrix} u_{x1} \\ 0 \end{bmatrix}, \tag{1.43}$$

where B_{11}, B_{21}, B_{12} and B_{22} are the coefficients of the matrix resulting from the multiplication of the individual transfer matrices of the soil layers, u_{xB} and σ_{xzB} are the displacement and

traction in the bedrock and u_{x1} is the displacement field at the surface.

Using the system of equations (Equations 1.43) and replacing, u_{xN} by $B_{11}u_{x1}$ and σ_{xN} by $B_{21}u_{x1}$ we find:

$$ik_B G_B B_{11} u_{x1} + B_{21} u_{x1} = 2ik_B G_B u_I. \tag{1.44}$$

The outcrop transfer function is given by:

$$\frac{u_{x1}}{2u_I} = \frac{ik_B G_B}{ik_B G_B B_{11} + B_{21}}. \tag{1.45}$$

The borehole transfer function is given by the ratio between u_{x1} by u_{xB}, such that:

$$\frac{u_{x1}}{u_{xB}} = \frac{1}{B_{11}}. \tag{1.46}$$

For a monolayer case, $B_{11} = cos(k_s H_s)$ and $B_{21} = -k_s G_s sin(k_s H_s)$, with, k_s, G_s and H_s, the complex wave number, the complex shear modulus and the thickness of the sediment layer. If we replace these expressions in the borehole and outcrop transfer function defined in Equations (1.45, 1.46), we find the ones previously presented for the mono-layer case except that in this section the bedrock is considered as elastic rather than visco-elastic.

1.2.3 Validity and limits of the present hypotheses

In this section we printed the propagation of SH seismic waves in 1-D layered soil model with vertical incidence and with homogeneous, isotropic, elastic and visco-elastic material within the soil layers.

1-D assumption Site effects are concerned with the subsurface soil layers (from 10 m to 1000 m in depth). At these depths, the spatial distribution of the materials can be very complex. It depends on the geological history of the site, and naturally depends on the scale we are interested in. However, the one-dimension assumption, although important, appears reasonable with regards to the quality of the information that can usually be obtained at a site. In the present study, the empirical site-response calculations come from vertical arrays from the KiK-net Japanese database, for which only 1-D site information is available.

Homogeneous and isotropic materials The hypothesis of materials that are homogeneous and isotropic is strong but also reasonable considering the level of details of the information (sampling of the velocity profile $\approx$ 1 m) that is available for this study regarding the frequency band of interest (0.1-30 Hz, minimum wavelengths of $\approx$ 3 m) .

Vertical incident of seismic waves When the seismic waves reach an interface between two materials, part of the wave is refracted and part is reflected, and wave conversions occurs. The direction of the refracted ray (direction normal to the wavefront) changes compared to the incident ray. For a ray propagating from material (1) to material (2), if material (2) is softer (lower seismic wave velocities and density) compared to material (1), then the ray is bent upwards (see Figure 1.7). Usually the soil layers are softer near the surface. Thus, when the waves propagate from a deep seismic source, the rays are likely to have a vertical incidence when they reach the surface. Hence, the vertical incidence of the seismic waves and the plane wave are reasonable assumptions for site effect assessment.

Propagation in linear, elastic and visco-elastic materials The rheological model of soil behavior, as illustrated in Figure 1.8 in the stress-strain plane, should be in compliance with the order of magnitude of the strain as suggested by Ishihara (1996). Linear elastic model are valid for shear strain below 10^{-5} Figure 1.8 (a). Visco-elastic models should be used for strain between 10^{-5} and 10^{-3} Figure 1.8 (b). Above 10^{-3}, non-linear model should be used Figure 1.8 (c). In this chapter we are mainly interested in low shear strain for which linear elastic and visco-elastic model is valid.

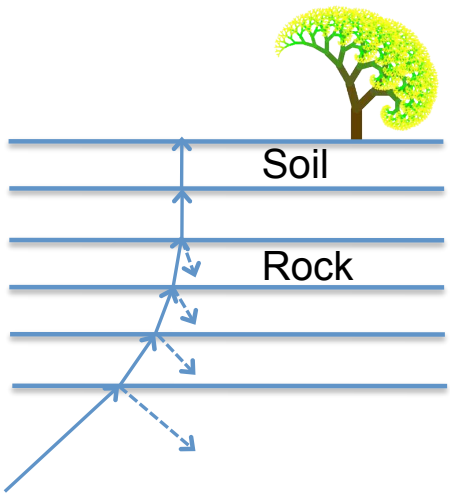


Figure 1.7: *Direction of the propagation of body waves in a layered medium with increasing rigidity with depth.*

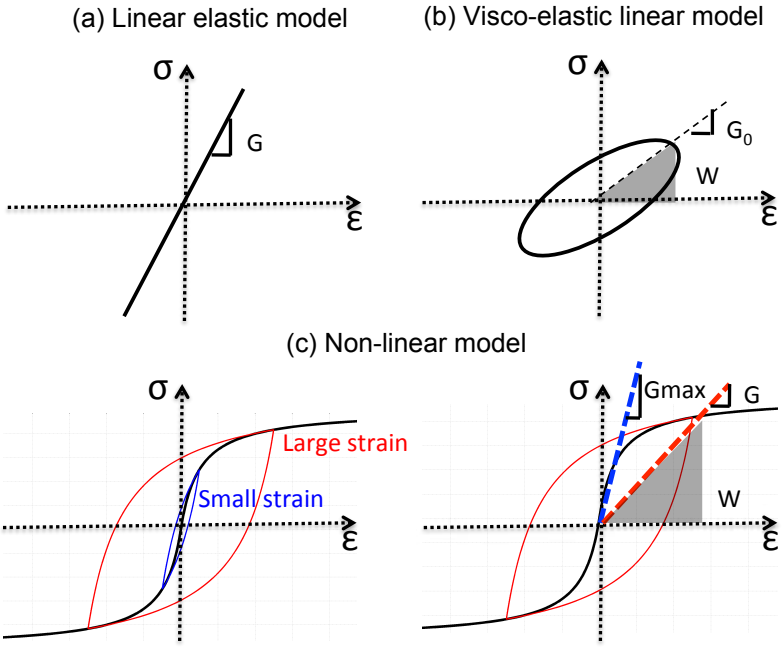


Figure 1.8: *Cyclic soil behaviour in the stress-strain plane, for the linear elastic model (a), the linear visco-elastic model (b), and the non-linear hyperbolic model (c).*

1.3 BOREHOLE-SITE RESPONSE AND IMPLICATION FOR INVERSION PURPOSES

In the present study, we are mainly interested in the borehole-site response, considering that vertical-array earthquake recordings are used to compute the site responses (under borehole conditions). In this very simple example, the linear theory indicates that the borehole-site response is independent of the bedrock parameters. The amplitude of the borehole transfer function does not depend on the impedance contrast; instead it only depends on the damping, the shear-wave velocity, and the thickness of the sediment layer (in the elastic case, the amplitude will be infinite). This is very important when dealing with inversions of transfer functions or time histories using borehole data. Indeed, if the borehole station is located just at the sediment/ bedrock interface, then using linear visco-elastic (or non viscous type Kelvin-Voigt) methods we cannot constrain the impedance contrast between the sediment and the bedrock; only information on the sedimentary layer can be inverted.

At the same time, we can show that for each couple (Vs, thickness) of the sedimentary layer linked to a constant coefficient (β) such that $H = Vs/\beta$, with similar resonance frequencies, the amplitude of the numerical transfer function is similar. For the non-viscous model type Kelvin-Voigt:

$$\begin{aligned} k^* &= \frac{\omega}{\sqrt{(G_0 + iG_1)/\rho}} \\ &= \frac{\omega}{\sqrt{(G_0(1 + i/Q)/\rho}}, \quad \text{with} \quad Q = \frac{G_0}{G_1} \\ &= \frac{\omega}{\sqrt{G_0/\rho}\sqrt{(1 + i/Q)}}, \end{aligned} \tag{1.47}$$

$\sqrt{G_0/\rho}$ is the real part of Vs (i.e the real velocity of the soil layer). Therefore, $\sqrt{G_0/\rho} = Vs = H\beta$ (for this specific case). Then, the complex wave number will be:

$$k^* = \frac{\omega}{H/\beta\sqrt{(1 + i/Q)}}. \tag{1.48}$$

Replacing this expression of k^* in Equation 1.32 gives:

$$\begin{aligned} TF_{bor,\beta} &= \frac{1}{cos(k^*H)} \\ &= \frac{1}{cos(\frac{\omega}{H/\beta\sqrt{(1+i/Q)}}H)} \\ &= \frac{1}{cos(\frac{\omega}{1/\beta\sqrt{(1+i/Q)}})}. \end{aligned} \tag{1.49}$$

In this relation, we have a trade-off between the Vs and the thickness of the sediment layer, since they are both included in the β coefficient. This trade-off is a main drawback for inversion purposes by enhancing the non-unikness of the solution. To illustrate this, let us defined the cost function, which is the difference between a transfer function calculated for a referent model (with N layers) and a transfer function calculated with a similar site configuration (with N layers) but with different Vs and thicknesses of the sediment layers:

$$Fc = \sum_{f=1}^{F} (BFSR_{ref}(f) - BFSR_{num}(f))^2, \tag{1.50}$$

where, BFSR$_{ref}$ is the measured (or reference) Borehole Fourier Spectral Ratio, BFSR$_{num}$ is the calculated BFSR, f is the frequency and F is the total number of frequency samples. Fc represents the mismatch between the reference and the calculated transfer functions. Considering the remarks in the previous paragraph, the cost function does not have a global minimum, but instead an infinity of local minima with the same amplitude.

The cost function for a monolayer soil profile is calculated for different values of Vs and thickness of the sediment layers, and while keeping the attenuation term constant. We use the (Vs_i, H_k) values, with $Vs_i \in [100, 105, ...500]$ m/s and $H_k \in [1, 2, ...50]$ m, and we compared the cost function to a reference model with sediment layers with Vs of 300 m/s and a thickness of 20 m.

Figure 1.9 illustrates the cost function in the Vs-thickness plane of the sediment layer for a monolayer soil model. There is a 'valley' in the cost function with an amplitude of zero (approximately zero, considering round-off errors) that corresponds to a Vs/thickness = 15. It can also be noted that the cost function is not symmetric on each side of this valley. On the left side of the valley (thickness < 35 m), the cost function increases and then decreases until it reaches a local minimum. In contrast, on the right side of this valley (thickness > 35 m), the cost function increases with a lower maximal amplitude, and then decreases almost monotonically. The third part of this study deals with inversions and tests on synthetic cases (see section 5.4 page 153). Several runs are carried out to analyze the random-walk in the model space followed by the inversion algorithm.

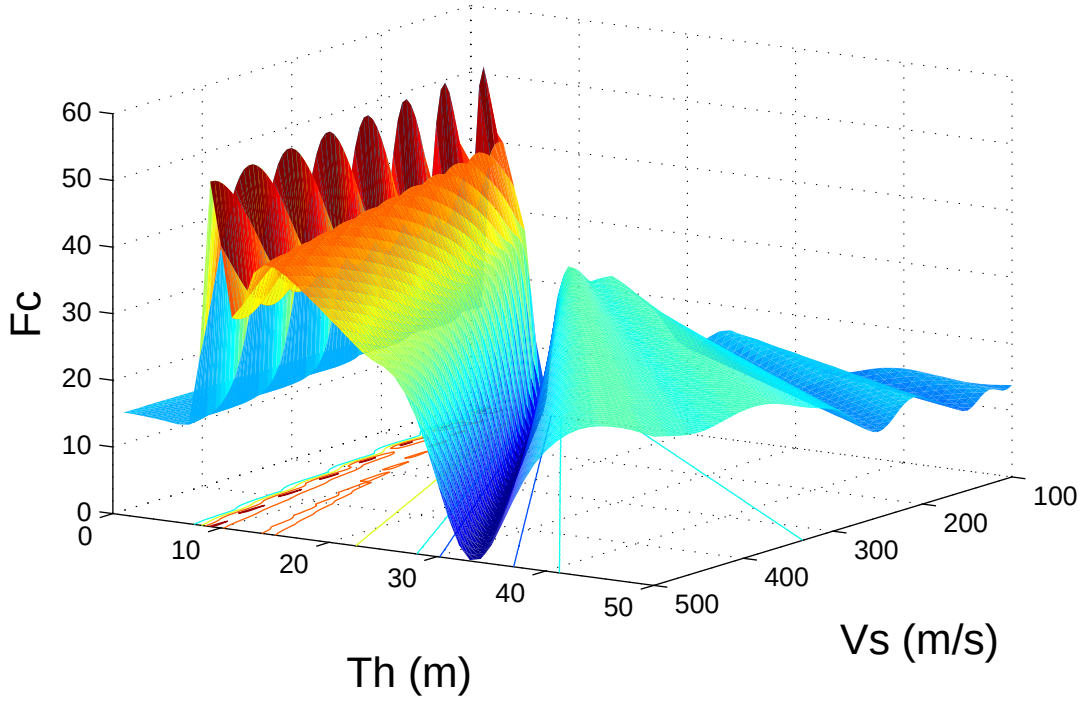


Figure 1.9: *Cost function of the borehole-site response calculated for a monolayer soil model in the Vs, thickness plane, considering that the reference model is a sediment layer with a Vs of 300 m/s and a thickness of 20 m.*

1.4 COMPUTATIONAL APPROACH

We used the well-known Haskell-Thomson method to perform the numerical simulation, which is the method presented above using a non-viscous type Kevin-Voigt model. This method is implemented in the NRATTLE code (C. Mueller and R. Herrmann), which has been included in the TREMOR program with the borehole condition case developed by F. Bonilla (2005).

1.4.1 Signal processing

To reproduce as close as possible the empirical data, we have to applied the same signal processing for both numerical and empirical data. The amplitude and the sharpness of the frequency peaks in the site-response curves depend mainly on the frequency sampling step used and on the smoothing applied to the data. We used the same frequency sampling step for all of the recordings (0.0244 Hz), and for the empirical data we applied a Konno-Ohmachi smoothing (with the parameter b = 40) to the Fourier transforms of the recordings at both surface and down-hole locations, before performing the ratio.

The first modification to the code was to impose the same sampling frequency step as for the accelerometric data. The second modification was to apply Konno-Ohmachi smoothing directly to the transfer function.

In Figure 1.10 (d), we compare the BFSR$_{num}$ (Borehole Fourier Spectral Ratio) calculated using different signal processing on the soil profile illustrated in Figure 1.10 (a,b and c). The dotted black line in Figure 1.10 (d) represents the BFSR$_{num}$ calculated without any signal processing, and it is the transfer function that comes directly from the Haskell-Thomson method. The black solid line curve in Figure 1.10 (d) is the transfer function with the Konno-Ohmachi smoothing applied directly to the transfer function, and the dotted light gray curve in Figure 1.10 (d) is the transfer function of the surface-to-borehole smoothed spectral ratio. It can be seen that there are slight differences between both of these ways of smoothing the numerical transfer function, and the first one has the advantage of applying only one smoothing, which significantly reduces the computation time.

1.4.2 Frequency-dependent attenuation

To compare the empirical and theoretical transfer functions, several methods can be used. For example, Thompson et al. (2012) used the Pearson coefficient to evaluate the "goodness-of-fit" between the two curves. For correct evaluation of this coefficient, a prior step is to find the values of the Q profile such that the amplitude of the empirical and numerical transfer function are equivalent. This step is not straight-forward. Indeed, when comparing the empirical and numerical borehole-transfer functions, we observed that the pseudo-resonance peak amplitudes (caused by destructive

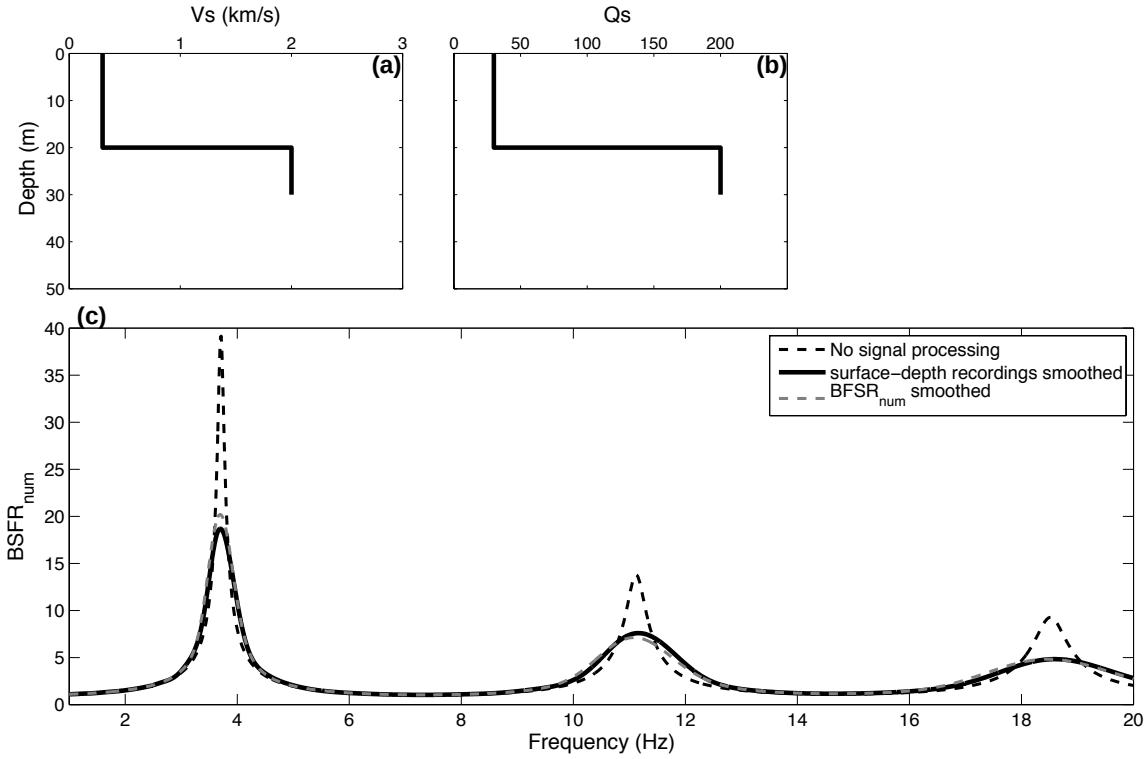


Figure 1.10: *Effects of signal processing on the evaluation of the BFSR$_{num}$ (linear condition). (a and b) The synthetic profile used (Vs, Q). (c) The BFSR$_{num}$ that comes from the different tests.*

interferences between the up-going and the down-going waves at the down-hole station, for more details see appendix B, also called the down-going waves effect) (e.g., Steidl et al., 1996; Bonilla et al., 2002; Régnier, Cadet, et al., 2013) are much higher in the numerical simulations, as compared to the observations. For given velocity and density profiles, the amplitudes of the peaks of the BFSR$_{num}$ are entirely controlled by the quality factor.

Figure 1.11 (b) illustrates the comparison of the empirical and numerical linear BFSRs (Borehole Fourier Spectral Ratios), as well as the linear earthquake H/V spectral ratio (that is the earthquake Horizontal to Vertical spectral ratio) at the surface and the numerical Outcrop Fourier Spectral Ratio (OFSR) at station IWTH23 (the velocity profile at this site is illustrated in Figure 1.11 (a)). We calculated the BFSR$_{num}$ for various Q (the value is constant along the depth). The OFSR indicates a similar peak frequency around 12.4 Hz, compared to the H/V at the surface. Around this frequency, the peak on the empirical BFSR is around 14 Hz and wider. The BFSR$_{num}$ indicates a series of four peaks below and above this frequency, showing the down-going waves effect (i.e. pseudo-resonance peaks). Whatever the value of Q used, the sharp peak at 5.4 Hz in the BFSR$_{num}$

has a large amplitude.

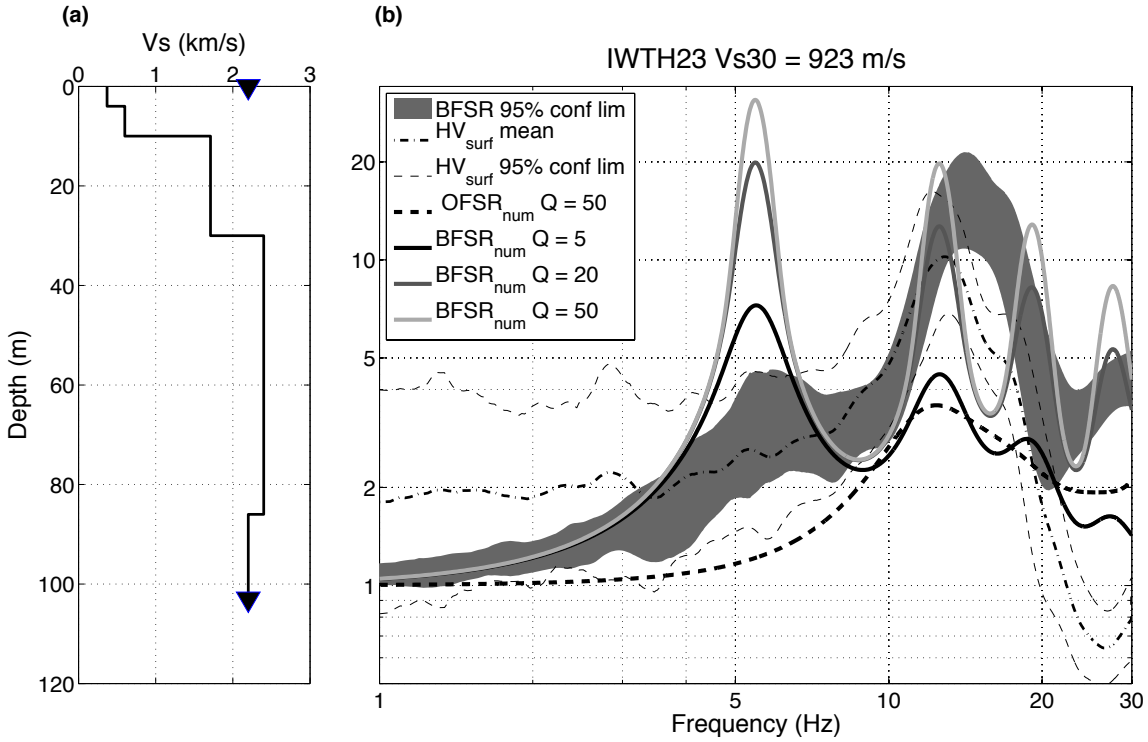


Figure 1.11: *(a) Shear-wave velocity profile of station IWTH23 down to 100 m. (b) Comparison of: (1) the empirical linear $BFSR_{emp}$ (95% confidence limits illustrated by the dark area); (2) the smoothed $BFSR_{num}$, computed for different values of Q (constant along the soil profile); (3) the smoothed OFSR, computed for a constant Q along the profile; and (4) the H/V at the surface (mean and 95% confidence limits represented by the black solid and dotted lines, respectively).*

Some studies have suggested that taking into account a frequency-dependent quality factor is a way to better model the wave scattering, and consequently the borehole observations (e.g., Satoh et al., 2001; Thompson et al., 2009; Assimaki & Steidl, 2007). However, as mentioned by Morozov (2008), the physics behind this assumption is poorly understood and not in accordance with the model chosen. Indeed, the frequency-dependence of the quality factor is positive (Q increases with frequency) whereas, we expect the short wavelengths to be more affected by small-scale heterogeneities (more attenuated) compared to larger wavelengths. Others explanations of this discrepancy can be put forward, simply considering the validity of the assumptions in the numerical implementation:

1. The incidence of the waves is vertical,

2. the soil layers are homogeneous and isotropic,

3. the site configuration is 1-D and we consider plane wave propagation,

4. only SH-wave propagation is considered.

Let us begin with the easy one. The vertical incidence cannot be a valuable explanation to the discrepancy since a large amount of earthquake recordings are used to calculate the empirical site response, and that the 95% confidence limit of the empirical site response at the first peak is still largely above the numerically calculated amplitude.

Heterogeneities in the soil layers would affect more the high frequencies which will not be in accordance with the observations. The empirical site response is computed by composing the two horizontal components of the recordings. Therefore, the anisotropy of the materials should not be a main cause of the discrepancy between observation and numerical simulation.

Now, the one-dimensional site configuration assumption might be one of the explanation of the discrepancy between numerical and empirical borehole site response. Indeed, the lowest frequencies have wavelengths that are comparable to the scale of a sedimentary basin. Therefore, low frequencies may be affected by possible 2-D or 3-D effects, leading to correlations (or anti-correlations) between up going and down going waves less pronounced. At the same time, another assumption of the numerical simulation is the propagation of SH-waves only, preventing from taking into account wave conversions at the interfaces as potential energy losses.

Despite, the physical explanation of the discrepancy between observations and numerical simulations is not straightforward, we choose in this study to adopt a frequency-dependent attenuation to reduce the differences. Hence, the third modification (after applying similar frequency sampling and smoothing at the numerical and empirical Fourier spectra) that we apply to the initial method is the implementation of frequency-dependent attenuation of the form detailed in Equation 1.51:

$$Q = Q_0 . f^{\alpha}, \tag{1.51}$$

where, f is the frequency, Q_0 is the frequency-independent quality factor at 1 Hz, and α is a coefficient modeling the attenuation decrease with frequency. Standard values of α are between 0.6 and 0.8 (e.g. Satoh et al., 2001; Yoshimoto et al., 1992).

1.5 SENSITIVITY ANALYSIS

To understand the wave propagation simulation and the effects of the aforementioned assumptions and modifications of the numerical method, sensitivity analysis is a widely use tool. It can be defined as the impact of the individual input parameters on the solution and can be addressed either

locally or globally. Local sensitivity analysis asseses the influence of each parameter on the mean of the results. Global sensitivity analysis estimates the sensitivity of the variance of the solution to the variance of the input parameters. This aims to quantify the relative importance of each input parameter uncertainty in the uncertainty of the model output, to answer the following question: "If we can eliminate the uncertainty in one of the input factors, which factor should be chosen to reduce the variance of the output?" (Saltelli et al., 2004). In the present thesis, as we are interested not only in understanding the influence of each parameter on the solution, but also in the propagation of the uncertainties from the input to the output, we will focus on global sensitivity analysis.

1.5.1 Description of the method

If we assume a specific model, correlation methods can be used. Without a hypothesis for the model, variance-based methods can be used. These methods are based on the exploration of the input-parameter space using Monte-Carlo techniques. Among these methods, three classical approaches can be found:

Sobol (Sobol, 1993): Let us assume that the mathematical model, which in our case is the vertical propagation of SH-waves in a layered model, is described by the function $Y(X = X_1, ..., X_n)$, for instance Y can be the BFSR(f) (the spectral amplification at specific frequencies) or the resonance frequencies f_0, f_1, where X is defined in a unit n-dimensional cube $K^n = \{X \mid 0 \leq X_i \leq 1, i = 1, .., n\}$, with X_i can be the normalised function of the Vs, the thickness, the density, the attenuation of each soil layer, or even the Konno-Ohmachi smoothing parameter. Using the numerical method previously defined, we can calculate the values of Y for each X. Sobol (1993) has shown that there is a unique decomposition of Y into summands, as described in Equation 1.52 for each function of Y that is integrable in K^n.

$$Y(X_1, ...X_n) = Y_0 + \sum_{i=1}^{n} Y_i(X_i) + \sum_{i=1}^{n}\sum_{j>i}^{n} Y_{ij}(X_i, X_j) + ... + Y_{1,2,...,n}(X_1, X_2, ..X_n), \tag{1.52}$$

where, Y_0 is a constant, and the integrals of the terms $Y_{i_s,..,i_p}$, $1 \leq (i_s, ..., i_p) \leq$ n with respect to each of their own variables $(X_{i_s}, ..., X_{i_p})$ are zero ($\int_{K^n} Y_{i_s,...,i_p} dX_j = 0 \quad j \in \{i_s, .., i_p\}$). Y_0 is defined as in Equation 1.53, and all of the terms $Y_{i_s,..,ip}$ are orthogonal to $Y_{j_s,..,j_p}$ if the $(i_s, ..i_p)$ and are different from $(j_s, ...j_p)$:
$\int_{K^n} Y_{i_s,...,i_p} Y_{j_s,...,j_p} dx = 0 \quad (i_s, ..., i_p) \neq (j_s, ..., j_p)$.

$$Y_0 = \int_{K^n} Y(X) dX, \tag{1.53}$$

If we consider that the X parameters are random variables, the function called D defined as the difference between the integral of the square of the function Y and the square of the function Y_0, can be considered as the variance of the function (see Equation 1.54):

$$D = \int_{K^n} Y^2(X)dx - Y_0^2. \tag{1.54}$$

By squaring Equation 1.52 and integrating over the parameter space, we obtain the result describe in Equation refeq-1-21 (considering that the square of the function Y and the square of the terms of the summand are also integratable):

$$D = \sum_{i=1}^{n} D_i + \sum_{i=1}^{n}\sum_{j>i}^{n} D_{i,j} + ... + D_{1,2,...,n} \tag{1.55}$$

where, $D_i = \int_{K^n} Y_i^2(X_i)dX_i$,
$D_{i,j} = \int_{K^n} Y_{ij}^2(X_i, X_j)dX_i dX_j$,
$D_{1,2,...,n} = \int_{K^n} Y_{1,2,...,n}^2(X_1, X_2, ..X_n)dX_1...dX_n$

It has been shown that the Sobol first-order sensitivity measures can be expressed as the ratio of D_i to the total variance of the result (D). D_i represents the so-called variance conditional expectation, which can also be written as $Var[E(Y|X_i)]$, and it is the complementary term to $E[Var(Y|X_i)]$, which is called the residual and represents the average of the variance of the function Y when X_i is fixed. The variance conditional expectation is a suitable measure of the importance of X_i; so, if D_i resembles D, X_i, it is very important.

$$S_i = \frac{Var[E(Y/X_i)]}{D}. \tag{1.56}$$

The main idea of the Sobol (Sobol, 1993) approach is to decompose Y into summands of increasing dimensionality, as specified in Equation 1.52. The summands are functions that involve each parameter or a combination of parameters. The variance of the function Y is given by the integration of the function over all of the variables space, as specified in Equation 1.54. While this method gives a very accurate result for each index independently, it can be ineffective when a large number of variables need to be tested, considering the large number of simulations required (for each parameter to be tested separately). This approach is interesting when a first selection of the main parameters has been realised (it quantitatively gives the sensitivity of the results to different parameters and to specific combinations of parameters, if wanted). Hence faster methods have been developed.

The Fourier amplitude sensitivity test . The Fourier amplitude sensitivity test (FAST) approach

is one of the most robust global sensitivity analysis techniques (Cukier et al., 1973; Saltelli et al., 1999; Xu & Gertner, 2008). In this approach, the parameters to be tested are variables at the same time when performing the simulation. The variability of each input parameter (X_i) follows a harmonic drawn with a specific frequency (w_i), as specified in Equation1.57. The output is therefore considered as a periodic function too, with different periodicities for the specific frequencies. If an input parameter X_j has a large influence on the solution, then the amplitude of the Fourier spectrum of the solution at w_j will be large. The Fourier spectrum of the model output is calculated, and the first-order sensitivity index of each input parameter (X_i) is given by the amplitude of the Fourier spectrum at the specific frequencies (w_i) and the associated harmonics relative to the total sum of the Fourier transform. To perform this calculation, the frequency and associated harmonics of an input parameter have to be different from the other input parameters, which therefore limits the number of input variables to be tested, and implies a larger number of simulations.

$$X_i(s) = \frac{1}{2} + \frac{1}{\pi}\arcsin(\sin(w_i s)), \tag{1.57}$$

$\forall i \in [1, n]$ if $s \in [-\pi, \pi]$ then $X_i \in [0, 1]$. This search-curve was proposed in (Saltelli et al., 1999), because it gives a satisfactory uniform distribution of X_i between 0 and 1. The solution Y is expanded into a Fourier series:

$$Y(X) = \sum_{j=-\infty}^{+\infty} A_j \cos(js) + B_j \sin(js), \tag{1.58}$$

where, the Fourier coefficients are defined as:

$$A_i = \frac{1}{2\pi}\int_{-\pi}^{\pi} Y(X(s))\cos(js)ds, \tag{1.59}$$

$$B_i = \frac{1}{2\pi}\int_{-\pi}^{\pi} Y(X(s))\sin(js)ds. \tag{1.60}$$

Over the domain of the integer frequencies $j \in Z =]-\infty, +\infty[$. The spectrum of the Fourier transform of the solution is given by $TF_j = A_j^2 + B_j^2$. The portion of the solution variance associated to the i^th parameter ($X_i, i \in [1, n]$) is the amplitude of the Fourier spectrum at the frequency w_i and the associated harmonics ($w_{i,p}$).

$$D_i = 2\sum_{p=1}^{\infty} TF_{f_i,p}. \tag{1.61}$$

The total variance is given by:

$$D = 2\sum_{j=1}^{\infty} TF_j. \tag{1.62}$$

Saltelli et al. (1999) showed that the sensitivity factor given in Equation 1.63 is equivalent to the first-order sensitivity indices defined by Sobol:

$$S_i^{FAST} = \frac{D_i}{D} \tag{1.63}$$

To find the S_i^{FAST} $\forall i \in [1, p]$ (p is the number of input factors), the FAST method requires only N simulations (which depend on the number of input factors tested; see Table 1.1), whereas with Sobol, a different sample was required for each input factor; i.e., $N * p$ Table 1.1 indicates the frequencies that can be used and the number of simulations that are required depending on the number of input factors.

Table 1.1: *Number of simulations and frequencies (for the input factors drawn) to be used in sensitivity analysis FAST depending on the number of input tested*

Number of	Sample size	Frequency
input factors	N	ω_i
5	625	[11,21,27,35,39]
6	393	[1,21,31,37,45,49]
7	697	[17,39,59,69,75,83,87]
8	1001	[25,55,77,97,107,113,121,125]

Random balance design FAST (Tarantola et al., 2006; Mara, 2009). Recent developments in the implementation of FAST by use of the random balance design technique (RBD-FAST) have allowed significant reductions in the computational cost. The drawback of this improvement is that only individual first-order sensitivity indices can be computed.

In this approach, all of the parameters are drawn according to the same frequency which reduced significantly the number of simulations to be performed. Each sample of the input factor X_i, which represents N realizations of X_i, is randomly mixed (with an aleatoric order for the input factor X_i called O_i). For each input factor (X_i), the result is sorted according to O_i, and the Fourier transform of the result is calculated. As all of the other variables except X_i are random, their effects in the Fourier spectrum can be considered as noise. The peak

in the Fourier transform at the common frequency can be associated to the importance of variable X_i. This approach is very interesting, as it significantly minimizes the number of simulations to be performed, and it does not depend on the number of input factors tested.

1.5.2 Application to synthetic cases

In the present study, we adopt the RDB-FAST global sensitivity analysis implemented in Matlab code that was originally written by Lopez-Caballero et al. (2011).

Illustration of the RDB-FAST global sensitivity analysis: To illustrate the application of the method, we analyzed the sensitivity of the calculated fundamental resonance frequency peak (f_0) to the soil parameters (Vs, thickness of the layers, attenuation) and specifically to the Vs of the sediment layer (Vs_1). We used a monolayer synthetic case and calculated the numerical BFSRs using the TREMOR code. In Figure 1.12 (a), the values of the normalised input factors, randomly mixed along the 250 simulations, are represented (Vs_1, Q_1, Th_1, ρ_1, the velocity, quality factor, thickness and density of the sediment layer). In Figure 1.12 (c) the solution, here f_0, obtained over the 250 simulations is shown. Figure 1.12 (b), the input factors are sorted according to the aleatoric order associated to the Vs_1. We see that Vs_1 has a harmonic behaviour along the 250 simulations performed, whereas the other parameters are completely random. Figure 1.12 (d) illustrates the values of f_0 sorted according to aleatoric order associated to the Vs_1. We can see in Figure 1.12 (d) that the solution has a clear harmonic behaviour and when we compute the Fourier transform of this solution (Figure 1.12, (e)), we can see that at the common frequency (here 2) the amplitude of the Fourier spectrum is high, indicating that the variance of Vs_1 has a large influence on the variance of the fundamental resonance frequency.

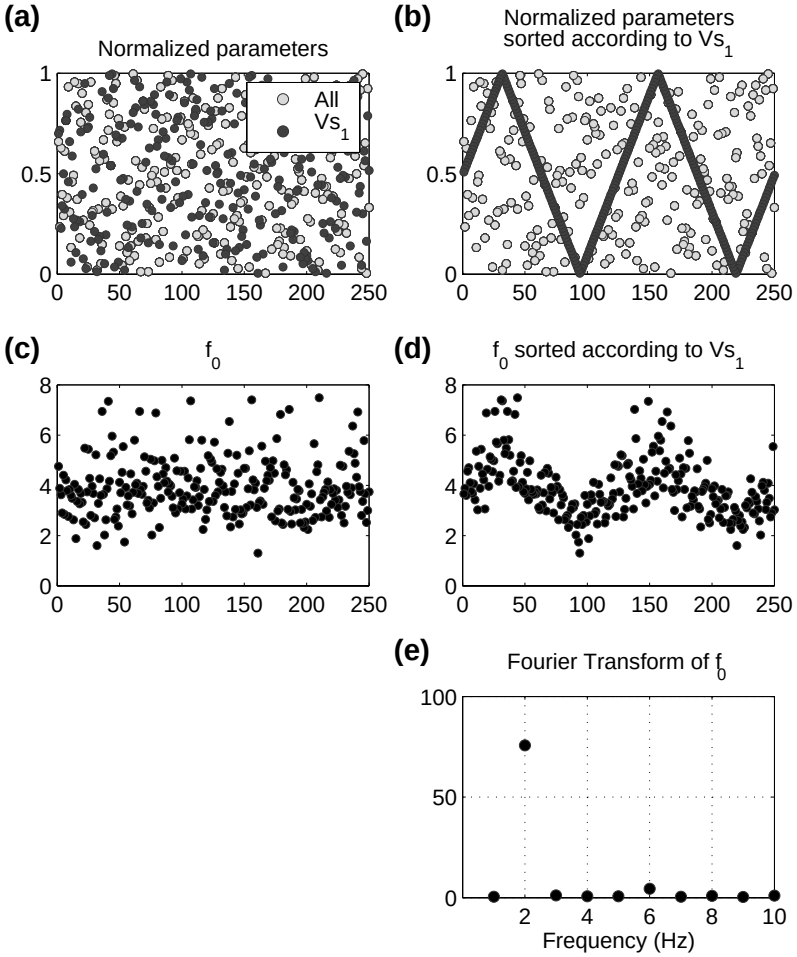


Figure 1.12: *Method of the RGB-FAST approach.(a) Input factors randomly mixed and normalized by their maximum value, along the 250 simulations performed (Vs_1, Q_1, Th_1, ρ_1, the velocity, quality factor, thickness and density of the sediment layer). (b) Input factors sorted according the aleatoric order of Vs_1. (c) The fundamental resonance frequency (f_0) resulting from the 250 simulations. (d) f_0 sorted according the aleatoric order of Vs_1. (e) Fourier transform of f_0 values*

The monolayer case : The mean monolayer profile is composed of a sediment layer of 20 m in depth and a Vs of 300 m/s, which is underlain by a bedrock with a Vs of 2000 m/s, and also considering that the reference station is a down-hole station at 10 m below the sediment-to-bedrock interface. The quality factors are a tenth of the Vs, and the α parameters in both layers are -0.5. We use a Konno-Ohmachi smoothing with a coefficient of 40.

We tested the global sensitivity of the frequency peak and the amplitude associated with all the parameters involved in the method; i.e., the Vs, thickness, ρ (the mass density), Q, and α of each layer (sediment and bedrock), as well as b_{konno}. We used a coefficient of variation of 30% for all of the parameters, except for the mass density, for which we used 10% (this parameter is less variable for actual soils) assuming a lognormal distribution. Figure 1.13 illustrates the results of this sensitivity analysis. In 1.13 (a), the $BFSR_{num}$ for all the random simulations are plotted in light gray curves with the $BFSR_{num}$ of the mean soil profile in black line. In Figure 1.13 (b), the sensitivity of the resonance frequencies is shown, with the associated amplitudes in Figure 1.13 (c). In Figure 1.13 (d), values of all the parameters during the 250 simulations performed are indicated. The variability of the resonance peak frequencies is mainly controlled by the variability of the Vs and thickness pair of parameters of the sediment layer. We can see that, as expected, the bedrock properties have low influence on the borehole site response, as discussed in the previous section. The amplitudes of the peaks are mostly controlled by the smoothing coefficient, the Q, and the α of the sediment layer. The relative influences of Q and α depend on the values of the resonance frequency peaks. At the fundamental resonance frequency, Q has more influence than α, and this influence is inverted at the harmonic resonance frequency peaks. This observation is in agreement with the definition of the attenuation term: at high frequencies the power parameter has a larger influence on the result than the product term.

The two-layer case : We applied this approach to a synthetic case represented by a soil column composed of two layers of sediments and a semi-infinite space bedrock. The mean characteristics of the soil column are given in Table 1.2. We used a coefficient of variation of 30% for all of the parameters, except for the volume mass, for which we used 10%. We determined the global sensitivity of the first third of the frequency peaks and the associated amplitudes for all of the parameters involved in the method; i.e., the Vs_i, H_i, Q_i and ρ_i of each sediment layer. Knowing that when the downhole sensor is located at the sediment-to-bedrock interface, the borehole transfer function is independent of the bedrock parameters, we did not test the bedrock properties.

Figure 1.14 (a) illustrates the transfer of the mean synthetic case and the 250 simulations that were performed. Figure 1.14 (b) and (c) show the sensitivities of the resonance frequencies and the associated amplitudes. Figure 1.14 (d) illustrates the values of all of the parameters

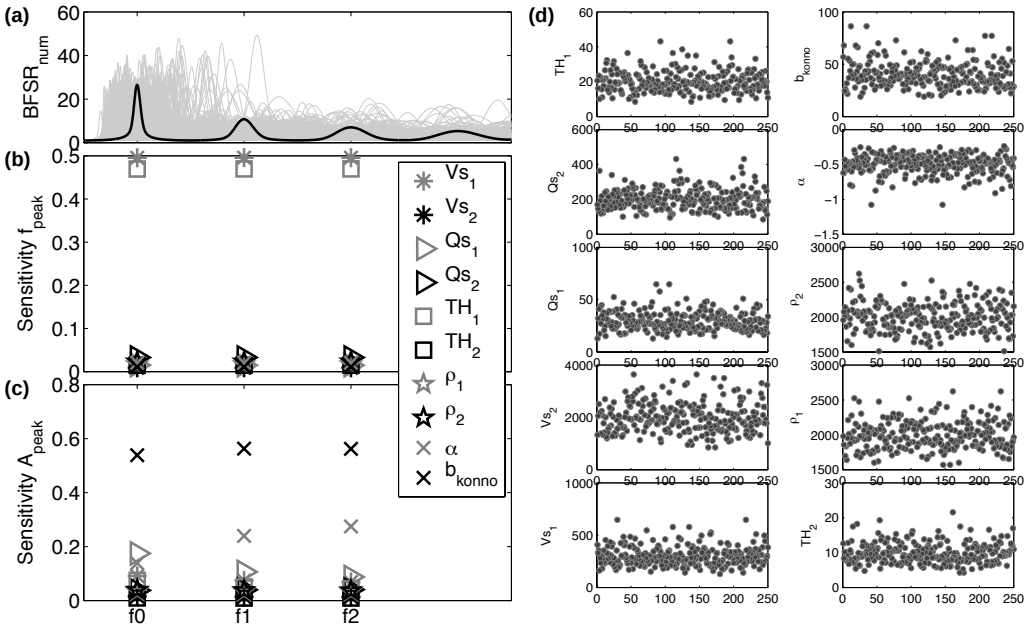


Figure 1.13: *(a) Borehole transfer function of the mean synthetic case, monolayer case, (black curve) with the 250 simulations (grey curves). (b) First-order sensitivity coefficients of the different input parameters on the first of their frequency peaks. (c) First-order sensitivity coefficients of the different input parameters on the associated amplitudes. (d) Values of the input parameters during the 250 simulations.*

Table 1.2: *Characteristics of the synthetic cases used to apply the sensitivity analyzes*

	Vs	H	ρ	Q
	m/s	m	t/m^3	
Layer$_1$	200	20	1.8	30
Layer$_2$	600	60	1.8	30
bedrock	2000	–	2	200

during the 250 simulations that were performed. We observe that the variability of the resonance frequency peaks is mainly controlled by the variability of the Vs and the two thickness parameters of the sediment layers. The first peak is mainly controlled by the parameters of the second layer, whereas the second and third peaks are more dependent on the parameters of the first layer. The variability of the peak amplitudes are mostly controlled by the Q and the

Vs of the sediment layers. The amplitude of the first peak depend mainly on the Q, whereas the amplitudes of the second and third peaks were controlled by the Vs of the layers.

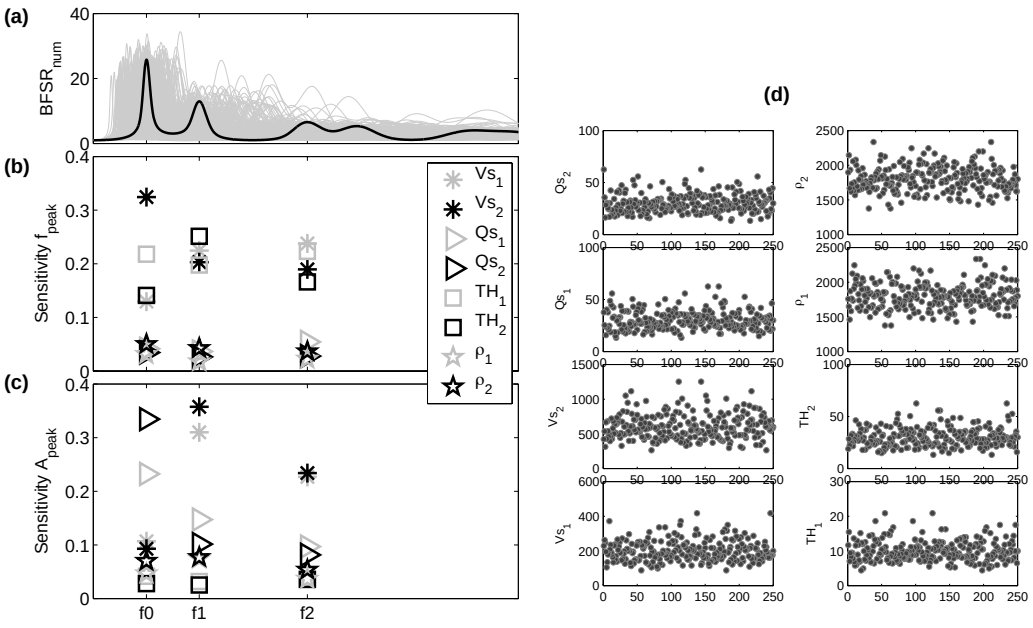


Figure 1.14: *Similar to Figure 1.13 but for a two layers of sediment synthetic case.*

1.5.3 Conclusions

The application of the sensitivity analyses on two synthetic cases showed that the modifications made on the computational approach may have strong influence on the results. Indeed, the amplitudes of the frequency peaks were mainly controlled by the Konno-Ohmachi smoothing parameter. In the monolayer case, excepted for the smoothing parameter, the attenuation parameters (Q and α) had a strong influence. Their relative influences depended on the frequency. When looking at the two-layer cases, the velocity of the sediment layers have influence on the peak amplitudes. It showed that the contrast between the two sediments layers have influence on peak amplitudes whereas the contrast between sediment and bedrock do not. This analyses also showed that the mass density had a very restrictive influence on the results.

Chapter 2

KiK-net database: data-selection, signal processing and site-response calculations

Sommaire

INTRODUCTION

A large part of this study is devoted to observations of earthquake recordings. As mentioned in the main Introduction, the aim of this study was to apply statistical analysis to the sources of the between-site and between-event site-response variability. To achieve this, we needed to find a large database, in terms of both the intensity of the incident motion recorded, and the geotechnical variability of the site. The Kiban-Kyoshin Network (KiK-net) in Japan fits these requirements precisely for this study. KiK-net was initially implemented following the very destructive Kobe earthquake. At present, the KiK-net is composed of 688 stations, with high-quality surface and downhole digital 3-component accelerometers. Among the KiK-net sites, 668 shear and compressive wave-velocity profiles have been collected (see Data and Resources). These velocity profiles were obtained from downhole PS logging measurements. Most of the borehole stations are located between 100 m and 200 m in depth. Figure 2.1 shows the Vs_{30} values of the KiK-net sites. Although most of the KiK-net stations are located on rock or on thin sedimentary sites (Fujiwara, 2004), Figure 2.1 indicates that two thirds of the sites have a Vs_{30}<550 m/s. In the Eurocode 8 regulations, sites with $Vs_{30} <$ 800 m/s are classified as sites prone to site effects, which confirms that the KiK-net database is very interesting for the analysis of site effects.

In this section, we present the KiK-net database and the data selection that was made. Then, we present the signal processing that was performed before the empirical site response was computed. We next introduce the methods that can be used to compute the empirical site response, or the equivalent curves, and detail the calculations performed in the present study. Finally, the last sections are devoted to presentation of the soil and site-response proxy parameters, and to the incident motion parameters that are used in the following chapters.

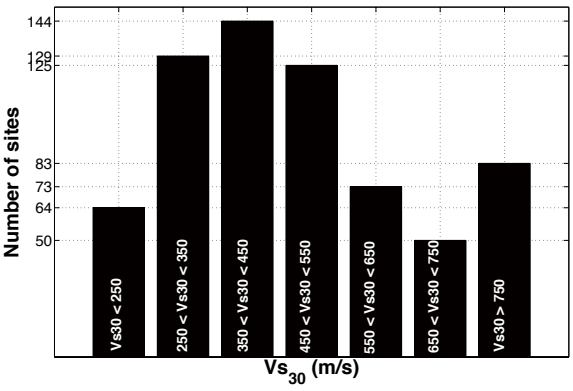


Figure 2.1: *Distribution of the KiK-net sites, according to the Vs_{30} (m/s) of the sites.*

2.1 DATA SELECTION AND SIGNAL PROCESSING

2.1.1 Data selection

For all of the KiK-net sites, we collected the accelerometric data that were recorded between 1996 and 2009 and showed a magnitude (M_{JMA})>3, and a hypocentral depth and epicentral distance<150 km. In addition, we collected the seismic events that were associated with the Tohoku earthquake, which occurred on 11 March, 2011, without any epicentral distance criteria. More than 46,000 (six component) recordings were analysed. The distributions of the selected recordings according to magnitude, distance, and peak ground acceleration (PGA) at depth are shown in Figure 2.2. The PGA of the recordings is specified by the colour scale on the right of Figure 2.2. Most of the recordings at the downhole sensor had a PGA<20 cm/s^2 (46,494 recordings), and 370 recordings had a PGA>50 cm/s^2.

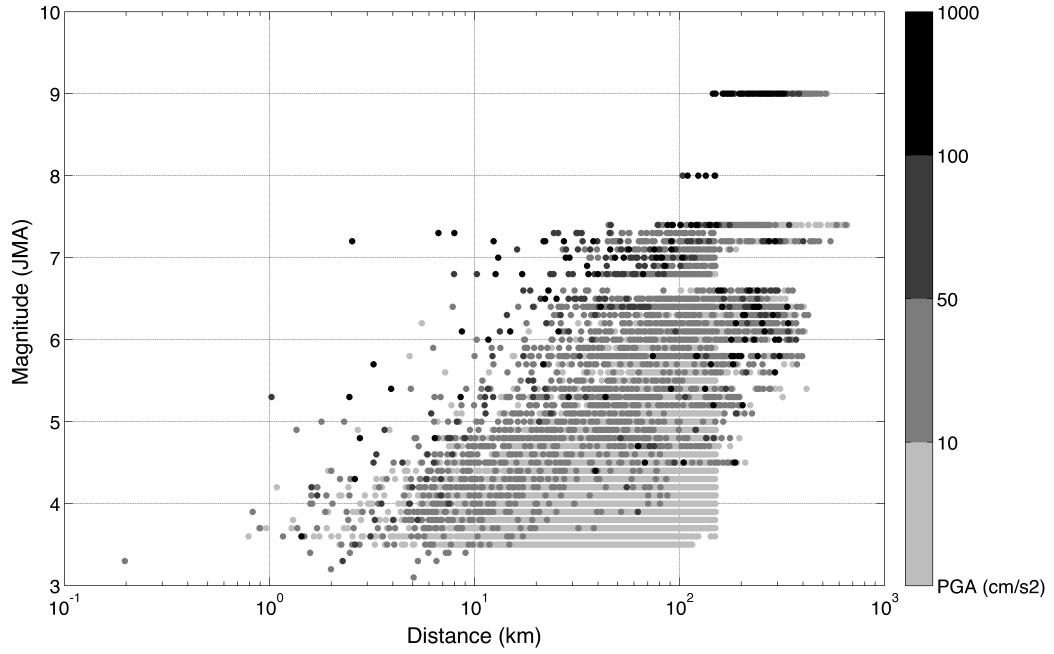


Figure 2.2: *Magnitude (M_{JMA}) and epicentral distance coupled along with the surface PGA of the selected recordings (> 46,000) from all of the sites of the KiK-net database.*

2.1.2 Signal processing

To avoid any bias in the signal processing, the only processing that was applied was a baseline correction of the time histories. The P-wave arrivals and the signal end (end of the coda waves) were automatically picked, as was the pre-event noise. The algorithm used for this automatic selection was based on the calculation of the ratio of the long-term average (LTA) to the short-term average (STA), which is commonly used for earthquake location (e.g., Withers et al., 1998). We chose a LTA of 5 s, a STA of 1 s, and a threshold of 0.5. The selection procedure was carried out according the following steps:

1. **Find the number of events present in the recording** . We found that in some of the KiK-net recordings, several events were present in one recording; in this case, we selected the most energetic event. We calculate the envelope of the signal, which is the module of the analytical function of the signal. The analytical function corresponds to the sum s(t)+i.H[s](t), where, s(t) is the signal and H[s](t) is the Hilbert transform of the signal. After filtering (2^{th}-order lowpass Butterworth filter with cutoff frequency of 0.25 Hz), we picked the local minima of the envelope. Minima before and after the time at which the maximal amplitude occurred (t_{max}) were identified.

   - Case 1: If no local minimum was detected, then we can conclude that only one event was recorded on the file.
   - Case 2: If a local minima was find before t_{max}, this means that either a first event was present before the most energetic event, or that a minimum local was found in the earthquake itself (e.g., between phases with strong amplitudes; i.e., P-wave, S-wave or surface wave arrivals).
   - Case 3: If a local minima was picked after the maximal amplitude, this means that an event is present in the recording after the most energetic event.

2. **Find the beginning (T_0) and the end (T_{end}) of the signal** . This step depends on the aforementioned cases:

   - Case 1: The LTA and STA were calculated all along the recording. A loop was run on the recording time ($t_0, t_1, ...t_n, ...t_{end}$). This loop stopped when a time for which the LTA/STA ratio was$<$0.5 was found and the two criteria below were respected, or when the end of the recording was reached. The two criteria below ensured that the picking was reliable. Thus with the assumption that $LTA/STA(t_n) < 0.5$:

     Criteria 1. The trigger must not be due to a small variation in the pre-event noise. Amplitudes in the noise can be very weak, and consequently the LTA/STA ratio can be very variable. We compared the signal amplitude at the trigger time with

the maximal amplitude of the signal, whereby the amplitude of the signal had to be at least>1/100th the maximal amplitude: $s(t_n)$>$s(t_{max})/100$

Criteria 2. The recordings must have a pre-event noise time window larger than the LTA: t_n>LTA

If T_0 was found, then the end of the signal was picked when $LTA(t > T_0)$ fell below twice the LTA at the beginning of the recording. If not, then $T_{end} = t_{end}$.

- Case 2: The beginning and end of the signal were detected similarly to case 1, but from the minimum local instead of the beginning of the recording. If no beginning was found, this meant that the local minimum was picked during the earthquake. If so, the first case procedure was applied.
- Case 3: The beginning of the signal was picked according to the case 1 procedure. The end of the signal corresponded to the first local minimum after the maximal amplitude.

Figure 2.3 illustrates the example of a recording for which several events were detected. The green curve corresponds to the envelop of the signal, the purple crosses to the local minima, and the dotted red window to the selected signal window. All of the recordings with a surface PGA>50 cm/s^2 were selected manually.

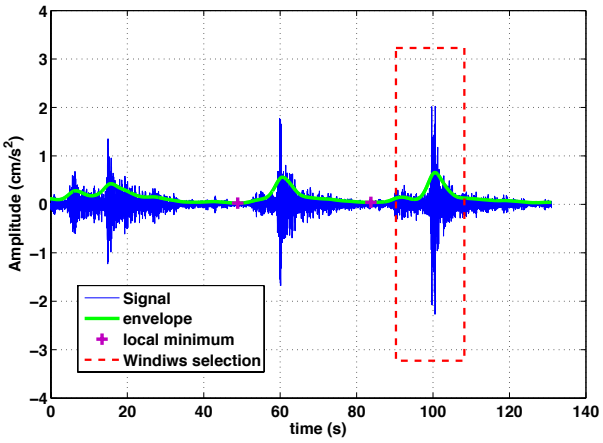


Figure 2.3: *Example of automatic detection of the beginning and end of an earthquake.*

All of the signals picked had the same time sampling, although the length of the selected window was variable from one earthquake to another. Hence, when performing the Fourier transform of the selected signals, the frequency sampling should be different. In the present study, we chose to

use the same frequency sampling for all of the signals, as 0.0244 Hz, by interpolation or decimation of the Fourier spectra. The interpolation was performed in the time domain by adding zeroes to the selected signals before performing the Fourier transform (this case occurred most of the time for the pre-event noise signal). Decimation was performed on the Fourier spectra.

We calculated the Fourier spectra of the three components of the whole selected signal and the pre-event noise (NS Fourier spectrum, for the North-South component; EW Fourier spectrum, for the East-West component; and V Fourier spectrum, for the vertical component). To each Fourier spectrum, we applied Konno-Ohmachi smoothing, using a coefficient b of 40 (Konno & Ohmachi, 1998). We combined the two horizontal components of the motion by computing the quadratic mean of the Fourier spectra ($\sqrt{(NS^2 + EW^2)/2}$).

2.2 EMPIRICAL SITE-RESPONSE CALCULATIONS

2.2.1 Presentation of the empirical method

As mentioned in the first chapter, there have been a number of different definitions of site responses, which have depended on the location of the reference site:

The outcrop transfer function. This is half the up-going wave transfer function (defined as the ratio of the outcrop displacement to the incident displacement at the downhole location), and it is the ratio of the outcrop displacement at the site to the outcrop displacement at a rock equivalent to the substratum rock.

The borehole transfer function. This is the ratio of the outcrop displacement to the incident and down-going wave displacement at the downhole location.

To evaluate the site responses empirically (the outcrop transfer functions), the standard method is to compute the spectral ratios between the signals simultaneously recorded on the sediments and at a reference site located nearby, which is usually a rock site (e.g., R. D. Borcherdt, 1970). The main issue to overcome is the selection of the reference site. Alternative methods were developed to overcome this last issue: vertical arrays of accelerometers for which the downhole sensor represented the reference station; horizontal-to-vertical spectral ratios (HVSRs) of the earthquake recordings at the surface; the so-called earthquake H/V spectral ratio (e.g. Langston, 1979; Lermo & Chavez-Garcia, 1993); or even global linearised inversions that have been used on large datasets of earthquake recordings at several sites to simultaneously determine the source, path and site-effect terms (e.g. Drouet et al., 2008) without reference sites. In the following, we present in more detail, (1) the standard spectral ratio, with emphasis on the specific case of vertical arrays and (2) the earthquake H/V spectral ratio methods that were used in the present study.

2.2.1.1 Standard spectral ratio

The surface ground motion is the convolution of the source path and the site effects. In the frequency domain, this gives :

$$S_{ij}(f) = O_j(f) \,.\, P_{ij}(f) \,.\, H_{ij}(f) \,.\, I_i(f) \tag{2.1}$$

where, S_{ij} is the Fourier transform of the seismic signal of the earthquake j recorded at site i, O_j is the source effect, P_{ij} is the path from the source to the site, H_{ij} is the local site response, and I_i the response of the recording material. With the hypothesis that the distance between the studied site and the reference site is small compared to the site to seismic-source distance: $P_{ij} = P_{ij,ref}$ et $O_j = O_{j,ref}$. After correction of the material response, we can considered that $I_i = I_{i,ref}$. The spectral ratio of the site to the reference Fourier transform gives the site-effects term

$$\frac{S_{ij}(f)}{S_{ij,ref}(f)} = H_{ij}(f) \tag{2.2}$$

The empirical site response is usually evaluated by using a spectral ratio between simultaneous recordings on sediments and on a nearby rock site (the so-called reference site). When this technique is applied, the main issue to overcome is the selection of a reliable reference site. The reference site must not amplify seismic waves, and should be close enough to the studied site so that the path travelled from the seismic source remains equivalent for both sites. The selection of a nearby reference site to calculate the empirical site response is a complicated task, as has been shown in several studies (Steidl et al., 1996; Duval et al., 1996; Drouet, 2006; Cadet, Bard, Duval, et al., 2012).

Vertical arrays of accelerometers overcome the reference-site issue. Indeed, the downhole station located on the bedrock at depth represents the reference station. Such a configuration has some disadvantages, mainly due to the down-going wavefield (Bonilla et al., 2002). Indeed, the borehole site response can be different from the outcrop site response, even if the downhole station is located at an equivalent rock site. At any depth, the particle motion contains the incident wavefield and the reflections from the free surface and from the different layer interfaces in the soil column. In the frequency domain, the destructive interference between the incident wavefield and the down-going waves can produce holes in the ground-motion spectrum (Steidl et al., 1996). Consequently, a direct spectral ratio between the surface and the total motion at depth generally produces pseudo resonances where these holes are present. This phenomenon is known as the down-going wave effect. Figure 2.4 illustrates these down-going wave effects. When the reference station is located just at the interface between the sediments and the rock, then the frequency peaks of the borehole transfer function (Figure 2.4, A/C) are in agreement with those of the outcrop (Figure 2.4, A/B). The amplitude of the peaks are nevertheless different, as already explained in the chapter dealing with

the one-dimensional site-response numerical calculation (chapter 1). When the reference station is located below (or above) the interface, then the frequency peaks of the borehole transfer function (Figure 2.4, A/C_{bis}) are different from those of the outcrop (Figure 2.4, A/B). In addition, when performing standard spectral ratios of both of the outcrop recordings, the free surface effect is similar to both the site and the reference station; however, in the case of a downhole reference station, the free surface effect is frequency dependent.

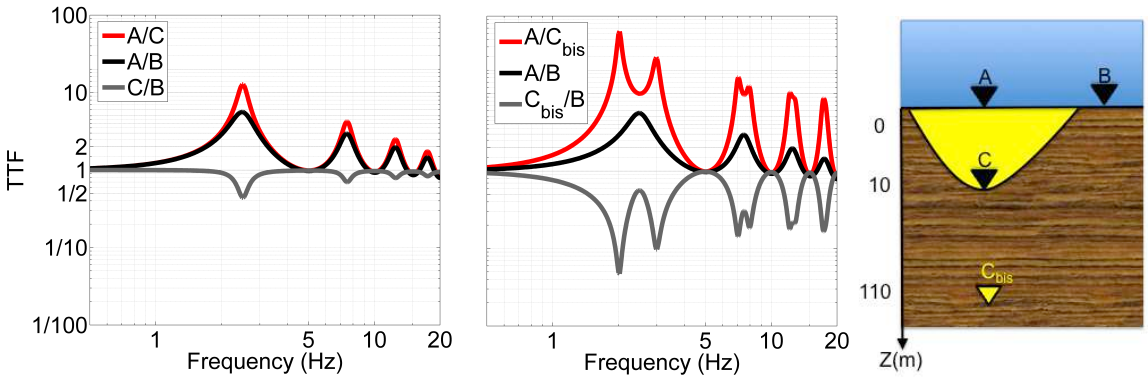


Figure 2.4: *Illustration of the down-going wave effect on the site response, depending on the location of the downhole station*

2.2.1.2 Earthquake H/V spectral ratio

Earthquake H/V spectral ratios or HVSRs of the earthquake recordings at the surface are an alternative method to calculate at least the fundamental resonance frequency of the site (f_0) without the need for a reference station. This method was introduced by Lermo and Chavez-Garcia (1993) following a study by Langston (1979). Since then, several studies have shown that (under the one-dimensional structure hypothesis) the earthquake H/V spectral ratio indicates the fundamental frequency of the site, and sometimes the higher modes of soil resonance (e.g., Theodulidis & Bard, 1995). Nevertheless, the peak amplitude does not indicate the amplification level (Bonilla et al., 1997).

2.3 COMPUTATION OF THE EMPIRICAL SITE RESPONSE

2.3.1 Borehole site response

In the present study, we used vertical arrays of accelerometers to calculate the empirical borehole site response. Some techniques were developed to correct the spectral ratio for the so-called

down-going wave effects (e.g. Kokusho, 2004; Cadet, Bard, & Rodriguez-Marek, 2012). As already shown in the first chapter when dealing with the improvement of the numerical method, the frequency peaks of pseudo-resonance on the empirical borehole site response are much more attenuated compared to the borehole site response coming from linear simulations. This shows that the down-going waves have less effect on the downhole recording then expected. Considering the strong assumption intrinsic to the down-going wave-correction techniques (the one-dimensional configuration), we chose not to apply any correction to the empirical borehole transfer function.

Hence, in this study, the borehole Fourier spectral ratio (BFSR) represents the empirical site response calculated per event. We calculated the BFSR at frequencies for which the signal-to-noise ratio was higher than a threshold value, called SN_{th} (both at the surface and at depth). SN_{th} is proportional to the ratio of the length of the signal window to the length of the pre-event noise window, and it is 3 when both windows have the same length).

$$SN_{th} = 3.\frac{Lw_{signal}}{Lw_{noise}}, \tag{2.3}$$

where, Lw_{signal} and Lw_{noise} are the lengths of the signal and noise windows, respectively.

For all of the sites with a Vs profile (668), we first characterised the empirical linear site response of the station, which was represented by the mean and standard deviation (in log units) of the BFSR computed with recordings for which the PGA at the downhole depth was$<$10 cm/s^2. In the present study, the mean empirical linear site response will be denoted as $BFSR_{lin}$, and the 95% confidence limit of the BFSR as $BFSR_{lin}^{95}$.

2.3.2 earthquake H/V spectral ratio

Similarly, we calculated the earthquake H/V spectral ratios of the surface sensors by dividing the combined horizontal Fourier spectra by the vertical Fourier spectrum, designated as the H/V, according to Equation 2.4.

$$H/V = \frac{\sqrt{NS^2 + EW^2}}{\sqrt{2}.V}, \tag{2.4}$$

where, EW is the Fourier transform of the East-West component, NS is the Fourier transform of the North-East component, and V is the Fourier transform of the vertical component of the recording.

For all of the sites, we computed the linear earthquake H/V spectral ratio of the surface station, which was represented by the mean and standard deviation (in log units) of the H/Vs computed with recordings for which the PGA at the downhole depth was$<$10 cm/s^2. In the present study, the mean surface earthquake H/V spectral ratio was denoted as HV_{lin}, and the 95% confidence limit of H/V

as HV_{lin}^{95}.

2.4 DEFINITION OF THE SOIL AND LINEAR SITE-RESPONSE PROXY PARAMETERS

The purpose of this section is to define the soil and site-response parameters that are available for the KiK-net sites and that were used to explain the between-site and the between-event (in this study, the non-linear soil behaviour effects on site response) site-response variability. These parameters are summarized in Table 4.3.

Shear-wave velocity profiles are related to the stiffness of the material, and therefore this parameter has an important role, as it controls the ground-motion amplification, signal duration, and spatial variability (Shearer & Orcutt, 1987; Cranswick et al., 1990; Semblat et al., 2005). However, precise knowledge of this geotechnical parameter is difficult to obtain, especially down to the depths of the seismic bedrock. Thus, proxy parameters, such as the mean shear-wave velocity in the first 30 m (i.e., the Vs_{30}), that might be easier to obtain are attractive options. The Vs_{30} was first introduced by R. Borcherdt (1992; 1994). Nowadays, the Vs_{30} is widely used to classify soils in regulation codes, such as with Eurocode 8 (EC8) and the National Earthquake Hazard Reduction Programme (NEHRP). In addition, ground motion simulations in California, USA, that account for site effects involve amplification functions that are based on Vs_{30} and rock peak ground motion (Graves et al., 2008). In Europe, Cadet, Bard, Duval, et al. (2012) developed amplification functions that were based on Vs_{30} and the soil fundamental frequency (f_0). Considering the wide use of Vs_{30}, several geophysical/ geotechnical techniques have been developed for its estimation (e.g., Castellaro & Mulargia, 2009). The uncertainties in Vs_{30} evaluation have also been characterised for various techniques (Moss, 2008).

The relevancy of Vs_{30} is, however, still under discussion in the seismology community, because it does not reflect the complexity of the shear-wave velocity profile (e.g., Castellaro et al., 2008; Cadet, 2007). In the past, several studies tried to correlate the site amplification at high frequency with Vs_{30} (e.g., R. Borcherdt et al., 2004); however, recent studies have demonstrated that this correlation is poorly constrained (Castellaro & Mulargia, 2009; Idriss, 2009). Frankel et al. (2002) and Park and Hashash (2004) underlined the influence on site amplification of the shear-wave velocity distribution at greater depths for deep sedimentary sites. Nevertheless, in practice, the earthquake engineering community appears to agree that Vs_{30} can perform well when classifying sites (Boore, 2004; Idriss, 2009).

Alternative soil classifications to the Vs_{30} have been proposed. Abrahamson and Silva (2008) proposed Vs_{30} and Z_{1000} (the depth at which Vs reaches at least 1000 m/s) as the parameters for characterising site effects in their GMPE model. Indeed, they used Vs_{30} to be consistent with current regulations, and Z_{1000} to distinguish between shallow, medium and deep soil sites. Luzi et

al. tested the relevance of seven different parameters (other than Vs_{30}): the average shear-wave velocity to the bedrock (Vs_{bed}), the average shear-wave velocity for different depths (Vs_H), the depth-to-bedrock and the resonant frequency obtained from the surface HVSR (using earthquake recordings and ambient-vibration measurements), f_0 obtained from numerical simulations, and f_0 obtained from HVSRs performed with response spectra instead of with Fourier spectra. They assessed the relevance of different sets of parameters for site-effect assessment in GMPEs, and they reported that the Vs_{30} and the f_0 are the most relevant pair of parameters, as also indicated by Cadet, Bard, Duval, et al. (2012). These studies underlined the importance of taking information complementary to the Vs_{30} in terms of depth of investigation, and also in terms of information type. For the purpose of the present study, we defined a parameter that characterised the behaviour of the Vs profile with depth. The KiK-net data provided the descriptions of the shear-wave velocity profiles at each instrumented site. Making use of this information, we proposed to use the gradient of the shear-wave velocity profile.

2.4.1 Vs profile proxy parameters

We defined proxy parameters for the shear-wave velocity profile. We used the well-known Vs_{30} and the gradient of the shear-wave velocity profile. This gradient was defined as the slope of the linear regression between the logarithm in base 10 of shear-wave propagation velocity and the logarithm in base 10 of the depth (Equation 2.5).

$$\log_{10} Vs(z) = B_{Zmax}.\log_{10}(z) + A_{Zmax} \pm \sigma_{Zmax}, \tag{2.5}$$

where, B_{Zmax} is the gradient of the Vs profile calculated from the regression between Vs and Z, up to Zmax, A_{Zmax} is the origin ordinate of the regression, and σ_{Zmax} is the standard deviation associated to the linear regression.

The gradient was calculated from the surface to a maximal depth. First, we chose 30 m (B_{30}), to be consistent with the depth of investigation needed to obtain Vs_{30}. While Vs_{30} gives an idea of the stiffness of the soil, B_{30} indicates the slope of the Vs profile. This parameter lies between 0 and 1. If B_{30} is 0, the velocity is constant with depth, and for a B_{30} >0, the velocity increases globally with depth. The larger the B_{30} is, the more rapidly the velocity increases with depth. We found that the B_{30} was enough to separate (for a limited Vs_{30} range of the sites) sites that had strong impedance contrast close to the surface (in the first 30 m of depth). We also calculated the gradient for two other depths: 50 m and 100 m in depth. Here, 100 m in depth is the maximum common depth to all of the KiK-net sites, and 50 m represents a middle way between 30 m and 100 m.

2.4.2 Linear site-response proxy parameters

In addition, we used parameters from the empirical linear site response, such as the fundamental resonance frequency (f_0), the predominant resonance frequency (f_{pred}), and the associated amplitude (A_{pred}) of the site. The fundamental resonance frequency is the frequency of the first significant peak of the outcrop site response curve. The predominant frequency is the frequency for which the site response curve has its maximum value.

In the database section, we emphasize the down-going wave effect on the borehole site response, to show that the evaluation of the fundamental resonance frequency can be difficult when using borehole site responses. As underlined in the section for the presentation of the empirical site-response method, the earthquake H/V spectral ratio is an alternative method that can provide at least the fundamental resonance frequency of the site. We picked the fundamental resonance frequency of the site on the earthquake H/V spectral ratio (mean and 95% confidence limit), checking that the peak amplitude was significantly>2 (t-test). The overall methodology developed to pick the fundamental resonance frequency with emphasis on the down-going wave effect is presented in Appendices B and C.

For the overall KiK-net sites that have a Vs profile (668), we calculated the fundamental resonance frequency at sites that had recorded enough earthquakes to have relevant evaluation (at least 3 recordings), which represented 649 sites. In addition, no amplification was found on 61 sites. Finally, we evaluated the fundamental resonance frequency for 588 sites.

Table 2.1: *Names of the soil and site-response parameters*

Vs_{30}	m/s	∈[100-2100]	Harmonic mean shear wave velocity
f_0	Hz	∈[0.1-30]	Fundamental resonance frequency
f_{pred}	Hz	∈[0.1-30]	Predominant resonance frequency
A_{pred}	-	∈[2-120]	Amplitude of $BFSR_{lin}$ at f_{pred}
B_{30}	-	∈[0-1]	Gradient of the shear-wave velocity profile to 30 m in depth
B_{50}	-	∈[0-1]	Gradient of the shear-wave velocity profile to 50 m in depth
B_{100}	-	∈[0-1]	Gradient of the shear-wave velocity profile to 100 m in depth

2.4.3 Analyses of the soil and site-response parameters

2.4.3.1 Distributions of the parameters

The distributions of the soil and site-response parameters are illustrated in Figure 2.5. The maximum amplitude of the BFSR_{lin} and the Vs_{30} have distributions close to log-normal, whereas the gradients have distributions close to a normal one (we used the statistical test called the Jarque-Bera test, which has the null hypothesis that the parameter comes from a normal distribution, with a 1% significant level). Half of the sites had maximal amplification (under borehole conditions)>10, indicating sites with potentially strong site effects. The distribution of the predominant resonance frequency was more spread between 0.1 Hz and 30 Hz compare to that for the f_0, for which half of the sites had a f_0 below 3 Hz. The distributions of the gradient calculated with a greater depth are less flat and narrow around 0.5.

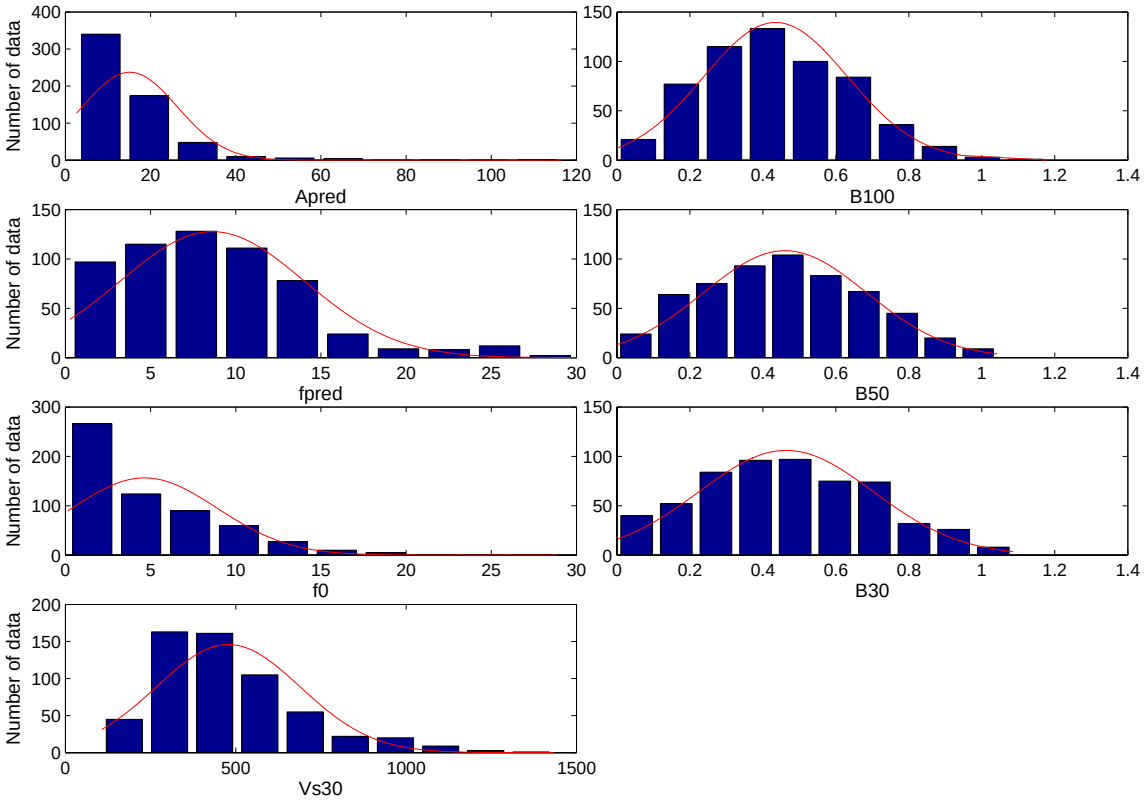


Figure 2.5: *Distribution of the soil and site-response parameters calculated for the KiK-net sites*

2.4.3.2 Correlations between parameters

We calculated the coefficients of correlation between the soil and site-response parameters defined above and given in Table 2.2. We can note that the correlation between Vs_{30} and all of the other soil and site-response parameters was very limited. The greatest correlations were with f_0 and f_{pred}, although these parameters were computed using different sources of information, here, the site-response curves (or earthquake H/V spectral ratios). Except for the gradient parameters calculated at different depths, the correlation coefficients were very low for all of the parameters. The results indicate that the soil and site-response parameters proposed in the present study are not correlated to Vs_{30}, and nor one to the other, which indicates that they might represent additional information for site characterisation or even for site-effect assessment.

Table 2.2: *Correlation coefficients between the soil and site-response parameters*

	Vs_{30}	f_0	f_{pred}	A_{pred}	B_{30}	B_{50}	B_{100}
Vs_{30}	1	0.59	0.46	-0.11	0.13	-0.03	-0.23
f_0	-	1	0.56	0.01	0.38	0.32	0.12
f_{pred}	-	-	1	-0.19	0.17	0.11	-0.05
A_{pred}	-	-	-	1	0.23	0.31	0.36
B_{30}	-	-	-	-	1	0.84	0.48
B_{50}	-	-	-	-	-	1	0.78

2.5 DEFINITION OF THE SEISMIC MOTION INTENSITY PARAMETERS

To find the pertinent intensity parameter for the study of the between-event site-response variability, and more specifically, for the non-linear soil behaviour and effects on site response, we selected six different intensity parameters that are related to different characteristics of seismic motion: maximum values, energy, duration and frequency content. In this section, we present the intensity parameters that were used, and we analyse these variabilities among the whole selected ground motion.

In the literature, different parameters have been defined to characterise the intensity of seismic motion. Kramer (1996) gave a definition of these parameters. Several studies have related the intensity parameters to the degree of damage and found that the cumulative absolute velocity (CAV), PGV and Arias intensity are relevant parameters for damage assessment (Cabañas et al., 1997); (Hernández, 2011). For non-linear soil-behaviour prediction, Assimaki et al. (2008) worked on the ratio between the incident motion central frequency and the fundamental resonance frequency of the site. The following intensity parameters were calculated over the whole record, both at the surface and at the downhole sensor location. We calculated the intensity parameters for the three components of motion, the values that we used were the quadratic means between the horizontal components:

- PGA: Peak ground acceleration (cm/s^2)
- PGV: Peak ground velocity (cm/s). To calculate the time history of the velocity, we integrated the acceleration after filtering between [0.1-25] Hz with a band-pass Butterworth filter in both the forward and reverse directions (order 3).
- PGD: Peak ground displacement (cm). To calculate the time history of the displacement, we integrate the acceleration twice after filtering between [0.1-25] Hz with a band-pass Butterworth filter, in both the forward and reverse directions (order 3).
- Arias intensity: I_a measures the energy of the signal and does not depend on the duration of the signal (Arias, 1970). I_a is a velocity (cm/s).

$$I_a = \frac{\pi}{2g} \int_0^\infty a(t)^2 dt \tag{2.6}$$

  where, a(t) is the acceleration of the seismic signal.

- The cumulative absolute velocity (CAV; cm/s) is the integration of the absolute value of the acceleration of the whole seismic signal.

$$\text{CAV} = \int_0^\infty |a(t)| dt, \tag{2.7}$$

- Trifunac duration, D_{tri} (s),

$$D_{tri} = D(\epsilon_1, \epsilon_2) = T(\epsilon_2) - T(\epsilon_1), \tag{2.8}$$

where, T(ϵ_1) and T(ϵ_2) are the times for which the cumulated energy of the signal reaches ϵ_1 = 5 and ϵ_2 = 95 % of the total cumulated energy of the signal.

- Acceleration root mean square a_{rms} (cm/s^2),

$$a_{rms} = \sqrt{\frac{\int_0^{D(\epsilon_1,\epsilon_2)} a(t)^2 dt}{D(\epsilon_1, \epsilon_2)}}, \tag{2.9}$$

- Central frequency (Fc). Fc is the frequency for which the power spectral density of the signal is the most concentrated. It is calculated using the ratio of the spectral moments defined in Equation 2.11.

$$\lambda_n = \int_0^{\infty} f^n.(TF(a(t)^2))df \tag{2.10}$$

$$Fc = \sqrt{\frac{\lambda_2}{\lambda_0}}, \tag{2.11}$$

where, a(t) is the acceleration of the signal, TF(a(t)^2) is the Fourier transform of the square of a(t), and f is the frequency.

Table 2.3: *Names of the soil and site-response parameters*

Abbreviation	Unity	Range at surface	Range at depth	Name
PGA	cm/s^2	∈[0.04-1966]	∈[0.01-906]	Peak ground acceleration
PGV	cm/s	∈[0.0013-67]	∈[0.008-53]	Peak ground velocity
PGD	cm	∈[0.1-40]	∈[0.0013-50]	Peak ground displacement
IA	cm/s	∈[0.1-31]	∈[0-6.8]	Arias intensity
a_{rms}	cm/s^2	∈[0-3.6]	∈[0-2.3]	Acceleration root mean square
Dtri	s	∈[0.06-219]	∈[0-208]	Trifunac duration
CAV	cm/s	∈[0-124]	∈[0-30]	Cumulative absolute velocity
Fc	Hz	∈[0.6-24]	∈[0.4-24]	Central frequency

2.5.1 Analyses of incident seismic motion parameters

2.5.1.1 Distributions of all of the parameters

The distributions of the intensity parameters at the surface and at the downhole sensor locations are illustrated in Figures 2.7 (a) and (b). The distributions for all of the parameters are almost log-normal (the x-axis scale is logarithmic in these Figures). The log-normal distribution is in agreement with the natural rate of occurrence of the earthquakes, depending on their magnitude and considering that the very small earthquakes are not recorded by the KiK-net accelerometric stations. According to the USGS, observations from 1900 indicate that the occurrence of earthquakes according to their magnitude follow a Gutenberg-Richter distribution, as shown in Figure 2.6.

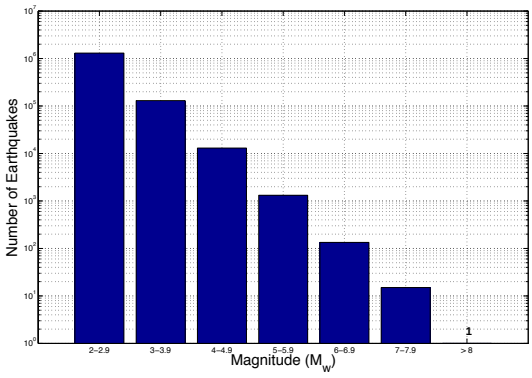
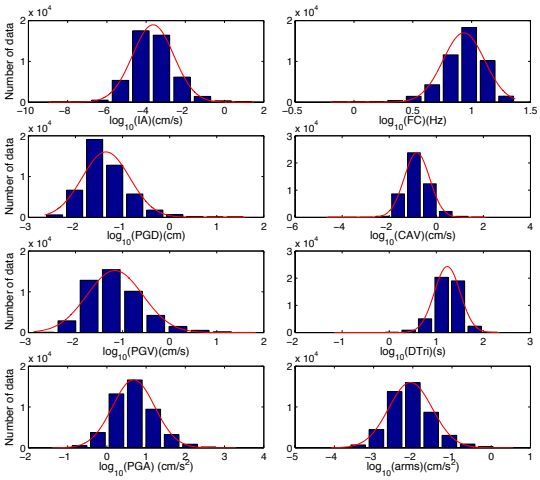
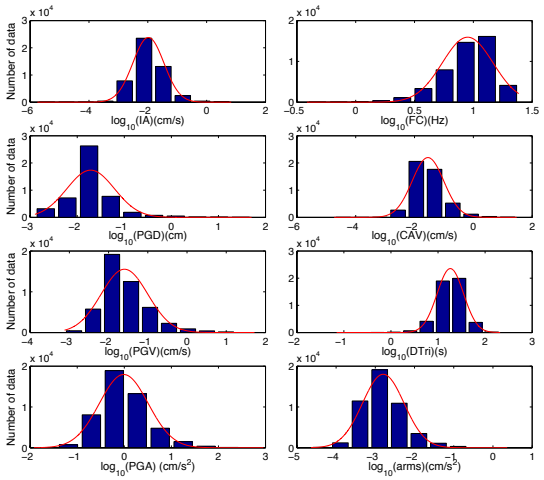


Figure 2.6: *Distribution of the number of earthquakes that occurred in one year, according to their magnitudes.*



(a) at the surface sensor location.



(b) at the downhole sensor location.

Figure 2.7: *Distribution of the incident-motion parameters calculated.*

The distribution of the number of sites according to the number of earthquakes recorded with a PGA at the downhole sensor depth $\geq$ 50 cm/s^2 is illustrated in Figure 2.8. For the whole of KiK-net, 54 sites have recorded at least two earthquakes with PGA at the downhole sensor depth $\geq$ 50 cm/s^2. The names of the 54 KiK-net are given in the Figure 2.8.

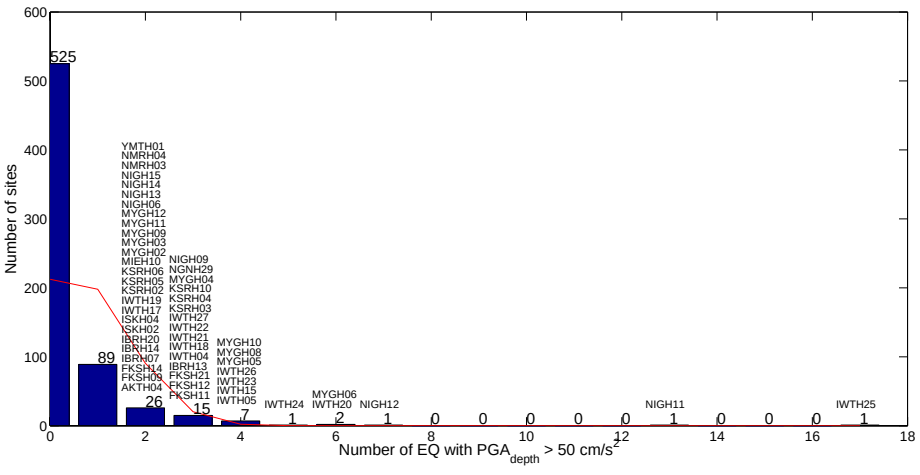


Figure 2.8: *Distribution of the KiK-net sites according to the number of recordings that had a PGA at the downhole sensor>50 cm/s^2*

2.5.1.2 Correlations between parameters

The coefficients of correlation between each intensity parameter were calculated, both at the surface and at the downhole sensor location. These are given in Tables 2.4 and 2.5. The parameters: PGA, PGV, PGD, CAV, a$_{rms}$ and I$_a$, were well correlated one with the other. However, the Trifunac duration and the central frequency did not appear to be correlated to the other intensity parameters, and not with each other. At depth and at the surface, the correlations were almost the same.

2.6 CONCLUSIONS

We presented the dataset used in our work. We also shown the empirical site-response calculations and parameters (soil, site-response and incident-motion) that we used to analyse the between-site and between-event site-response variabilities. We have shown that in the KiK-net database, the soil and site responses have normal or log-normal distributions that guarantee the reliability of the sample of KiK-net sites to represent large cases of site configuration. Similarly, the intensity parameters of the ground motion are log-normal distributed, which represents the observations of the earthquake rate of occurrence depending on the magnitude

(http://earthquake.usgs.gov/earthquakes/world/). We can note here that the correlations between PGA, PGV and PGD with the arias intensity are higher at the down-hole sensor than at the surface. It can be explain by the fact that the arias intensity is an integration of the whole signal compared to the PGA, PGV and PGD that are maximal values only, consequently this parameter could be more sensitive to site effects. However, we would have expected similar trends with the CAV, but it is not the case. Although, this observation is not well understand, it should be considered since arias intensity begin to be a more widely intensity parameter used in GMPEs.

Table 2.4: *Correlation coefficients between intensity parameters of the recordings at the surface*

	PGA	PGV	PGD	Ia	a_{rms}	CAV	Dtri	Fc
PGA	1	0.77	0.52	0.70	0.96	-0.02	0.74	-0.07
PGV	-	1	0.82	0.69	0.74	0.17	0.80	0.16
PGD	-	-	1	0.57	0.48	0.25	0.74	-0.10
IA	-	-	-	1	0.64	0.10	0.87	-0.04
a_{rms}	-	-	-	-	1	-0.04	0.67	-0.07
Dtri	-	-	-	-	-	1	0.20	-0.35
CAV	-	-	-	-	-	-	1	-0.10

Table 2.5: *Correlation coefficients between the intensity parameters of the recordings at depth*

	PGA	PGV	PGD	Ia	a_{rms}	CAV	Dtri	Fc
PGA	1	0.77	0.52	0.82	0.97	0.01	0.69	-0.03
PGV	-	1	0.87	0.81	0.72	0.22	0.87	0.14
PGD	-	-	1	0.67	0.47	0.27	0.81	-0.10
IA	-	-	-	1	0.77	0.12	0.83	-0.07
a_{rms}	-	-	-	-	1	0.00	0.61	-0.04
Dtri	-	-	-	-	-	1	0.25	-0.11
CAV	-	-	-	-	-	-	1	-0.49

Part II

Analysis of empirical seismic site response variability

Chapter 3

Inter-site site-response variability

Sommaire

ABSTRACT

Local geology can strongly affect seismic ground motion. One standard parameter used in soil classification for site effects is the average shear-wave velocity in the first 30 metres; namely the Vs_{30}. The higher the Vs_{30}, the stiffer the material is and the less amplification is expected. However, this index does not account for the complexity of the velocity profile, and especially its variability at depth. Indeed, some sites may have similar Vs_{30} values, but very different velocity profiles. In the present study, we propose to assess the influence of the Vs profile on site amplification. In addition to Vs_{30}, we propose to consider the gradient of the Vs profile (calculated for a depth ranging from 30 m to 100 m). The gradient characterises the velocity evolution with depth: a lower gradient value means low velocity increases with depth; a higher gradient value indicates strong impedance contrast in the shallow layers. In addition, we consider the fundamental resonance frequency of the soil (f_0), which has been shown to be a relevant parameter for site-effect assessment, and which is independent of the Vs profile. Using the KiK-net database, we selected 351 sites with one-dimensional linear numerical modelling close to the empirical site response (weak motion, peak ground acceleration $\leq$ 10cm/s^2). We analyse here the variability of the Vs profiles and the empirical borehole site responses of these selected sites through the Vs_{30}, the gradient, and the fundamental resonance frequency. We find that although the gradient is calculated using shallow information, as just the Vs_{30}, it can separate at least the sites characterised by strong shallow velocity contrasts from the sites that have a similar Vs_{30}, and can provide a robust evaluation of the site response of such sites. For a given Vs_{30}, the gradient calculated from the deeper Vs information (100 m in depth) and the f_0 can be used to distinguish between three types of sites: deep sedimentary sites, medium sites, and sites with high Vs contrast at shallow depth. Considering that the corresponding site response curves are very different for these three kinds of sites, the Vs profile at depth appears to influence the site response. The combined use of Vs_{30}, the velocity gradient, and the fundamental resonance frequency allows robust comparisons of sites to evaluate the influence of the Vs profile at depth, and therefore the relevance of the Vs_{30} for site-effect assessment and site classification purposes.

This chapter is mainly based on the work that was submitted to the BSSA in January 2013, under the title: "Influence of the velocity profile at depth on the linear characterisation of site effects: Tests on the KiK-net database" (Régnier, Bonilla, et al., 2013).

3.1 INTRODUCTION

In ground-motion prediction equations (GMPEs), the variability is described by the standard deviation (σ), and it has a great influence on the evaluation of the probabilistic seismic hazard. The variability is divided into two main terms: the between-event variability, and the within-event variability. The latter describes the inter-site variability of the ground motion for a given couple-magnitude

distance. This variability is intimately related to site effects (Al Atik et al., 2010; Rodriguez-Marek et al., 2011). Consequently, a robust evaluation of site effects is required for risk mitigation and urban planning, or when computing GMPEs at sites with different soil conditions (e.g., Cotton et al., 2006; Douglas, 2006; Douglas et al., 2009).

As mentioned in the previous chapter, Vs_{30} is a widely used parameter for site classification. The relevancy of this parameter is, however, still under discussion in the seismology community, because it does not reflect the complexity of the shear-wave velocity profile (e.g., Castellaro et al., 2008; Cadet, 2007). Alternative soil classification to the Vs_{30} using parameters that constrain the depth of the substratum of the fundamental resonance frequency of the sites have been proposed (e.g. Abrahamson & Silva, 2008; Cadet, Bard, Duval, et al., 2012; Luzi et al., 2011) .

In the present study, we propose to determine the influence of the Vs profile at depths greater than 30 m, and consequently to determine the relevance of Vs_{30} as a proxy for site-response analysis. We chose the well-characterised KiK-net database of boreholes in Japan, where the surface-to-borehole spectral ratios can be computed, to obtain the empirical site responses (borehole condition). The KiK-net database includes 668 sites for which the P-wave and S-wave velocity profiles are available down to the depth of the borehole station.

The main hypothesis of the present study is that the variability of the lateral soil properties (which involve two-dimensional [2-D] or 3-D site-configuration effects) does not significantly affect the seismic site response. Hence, we selected sites for which the linear empirical borehole site response evaluations ($BFSR_{lin}^{95}$ and RF_{lin}^{95}) are close to the 1-D linear numerical simulations ($BFSR_{num}$ and $OFSR_{num}$). We determine the variability of these shear-wave velocity profiles and the site responses among the site classes, as defined by the Vs_{30}. We also define the gradients of the shear-wave velocity down to 30 m and 100 m (B_{30}, B_{100}, respectively). These parameters provide information about the velocity evolution with depth. For a given Vs_{30}, a greater gradient value will mean a greater velocity contrast in the first 30 (or 100) m. Lower gradient values indicate low increases in the velocity with depth. Also, using the site response curves, we determine the fundamental resonance frequencies of the sites (f_0). The combined use of Vs_{30}, f_0 and the shear-wave velocity gradient allows robust comparisons of sites, to evaluate the relevance of Vs_{30} for site classification purposes.

3.2 SELECTION OF THE KIK-NET SITES WITH 1-D SITE CONFIGURATION

The basic idea of this section is to select the sites for which the empirical site response was as close as possible to the computed 1-D linear site response. We compared the empirical and numerical borehole site responses ($BFSR_{lin}$, $BFSR_{num}$, respectively), and we compared the numerical outcrop site responses (OFSRs) with the earthquake H/V spectral ratios at the surface (RF_{lin}).

To compute the $BFSR_{num}$ and $OFSR_{num}$, we used the Haskell-Thomson method, as explained

in the first chapter. The numerical simulation required the profiles of the shear-wave velocity, the density, and the quality factor (Q) with depth. For the 668 sites, the shear-wave and compressive-wave velocity profiles were available through the KiK-net website. However, neither the density nor the attenuation factor profiles were available. In the present study, considering that the volume mass had a very low influence (see sensitivity analyses, in the last section of the chapter 1), we chose a constant value of the volume mass along the depth of 2t/m^3. The Q factor had a strong influence on the amplitude of the transfer function. The choices of the values of Q are discussed in following paragraph.

Thompson et al. (2012) used the calculation of the Pearson coefficient to evaluate the 'goodness-of-fit' between the two curves. For a correct evaluation of the Pearson coefficient, a prior step is to find the values of the Q profile such that the amplitude of the empirical and numerical transfer function are equivalent. This step is not straightforward. Indeed, when comparing the empirical and numerical borehole transfer functions, we observed that the pseudo-resonance peak amplitudes (e.g., Steidl et al., 1996; Bonilla et al., 2002; Régnier, Cadet, et al., 2013) were much greater in the numerical simulations than in the observations. For a given velocity and density profile, the amplitude of this peak for the BFSR$_{num}$ was entirely controlled by the Q factor (see chapter 1, in the subsection on the frequency-dependent attenuation).

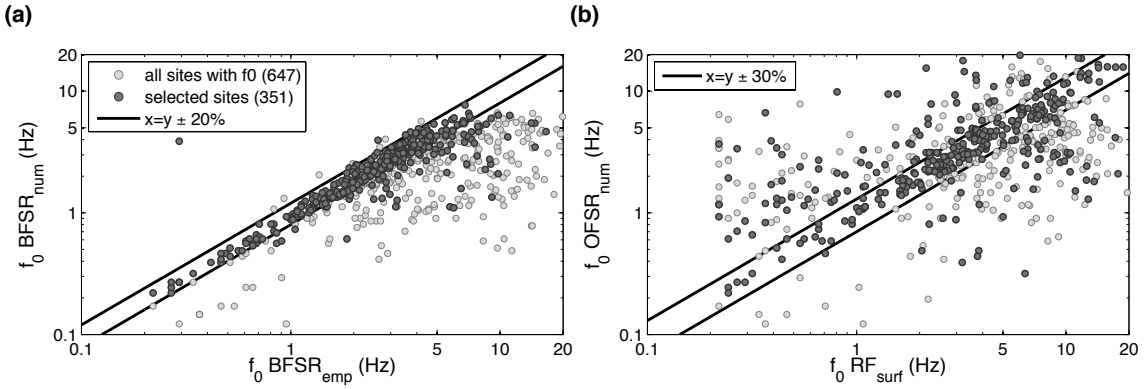


Figure 3.1: *(a) Comparison of the fundamental frequency picked for BFSR$_{num}$ and picked for BFSR$_{emp}$, calculated using weak motion (peak ground acceleration at the downhole station $\leq$ 10 cm/s^2). (b) Comparison of the first significant peaks picked on the OFSR and on the mean of the earthquake H/V spectral ratio at the surface.*

The difficulty of comparing the amplitude of the observations with the numerical simulation led us to consider only the peak frequency for the comparison that depended only on the known parameters (Vs and the thickness of the sediment layers), according to the sensitivity analysis. Figure 3.1(a) illustrates the comparison of the frequency of the first significant peak of the BFSR$_{num}$

with the $BFSR_{emp}$ peak (the frequency was picked for 647 sites among the 688 total). We found that for 270 sites, these two frequencies were close, with 20% variability. In Figure 3.1(a), it can be seen that most of the time the fundamental frequency picked on the numerical curve was lower than the fundamental frequency picked on the empirical transfer function. This observation shows that the down-going wave effect was generally less pronounced on the empirical data, as discussed in the previous paragraph. Figure 3.1(b) illustrates the comparison of the frequency of the first significant peak of the earthquake H/V spectral ratio at the surface with the OFSR peak. In the sites rejected by the previous selection, we compared the first significant peak of RF_{surf} with the OFSR peak. We added to the selection the sites for which these frequencies were close (within 30% variability). Considering these two criteria of selection on the first significant peak frequency, we ended up with 351 sites (we chose a 20% difference between the numerical and empirical data as a compromise between the reduction in the difference between the numerical and empirical site responses and the number of sites selected; at a 10% difference, the number of sites was reduced to 189).

As mentioned in the signal processing chapter, we do not apply any correction to the signals. Especially, we do not correct the borehole transfert function such that the selected sites have a common reference (i.e. a similar Vs at the down-hole station). This could be achieved by convolving the empirical borehole transfer function with the numerical transfer function of a generic rock profile (Boore & Joyner, 1997).

3.3 VARIABILITY OF THE VS PROFILES AND SITE RESPONSES ACCORDING TO VS_{30} ONLY

We explored the variability of the shear-wave velocity profiles between the sites with similar Vs_{30} ranges ($\leq$250, 250-350, 350-450, 450-550, 550-650, 650-750, 750-850, $>$850 m/s). We observed that below 20 m in depth, the shear-wave velocity profiles were highly variable among the Vs_{30} ranges (Figure 3.2). From this simple observation, we noted that the Vs_{30} parameter did not explain the complexity of the shear-wave velocity, especially at depth.

We also explored the variability of the mean $BFSR_{emp}$ among the same Vs_{30} classes illustrated in Figure 3.3. The variability of the site responses was large, whatever the Vs_{30} range. This variability appeared to be even greater for lower Vs_{30}. Nevertheless, it was clear that the confidence limits were different depending on the Vs_{30} range, and the frequency peak, for which the amplification was maximal, increased with increasing Vs_{30}.

3.4 ON THE USE OF ADDITIONAL PARAMETERS TO V_{S30}

According to the previous observations we used additional parameters to Vs_{30} for site-effect assessment and site classification purposes. We used the soil and site response parameters that were

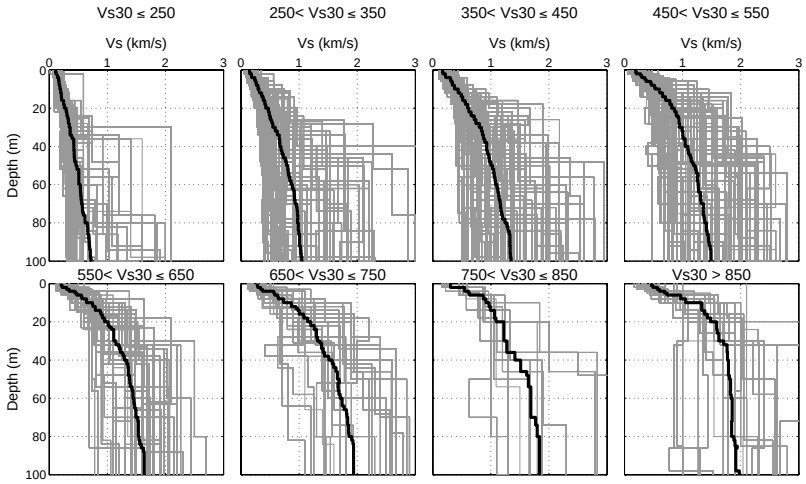


Figure 3.2: *Variability of the shear-wave velocity profiles down to 100 m in depth for different ranges of Vs_{30}. The thick lines represent the means of the Vs profiles for each Vs30 bin; the thin lines represent the individual Vs profiles.*

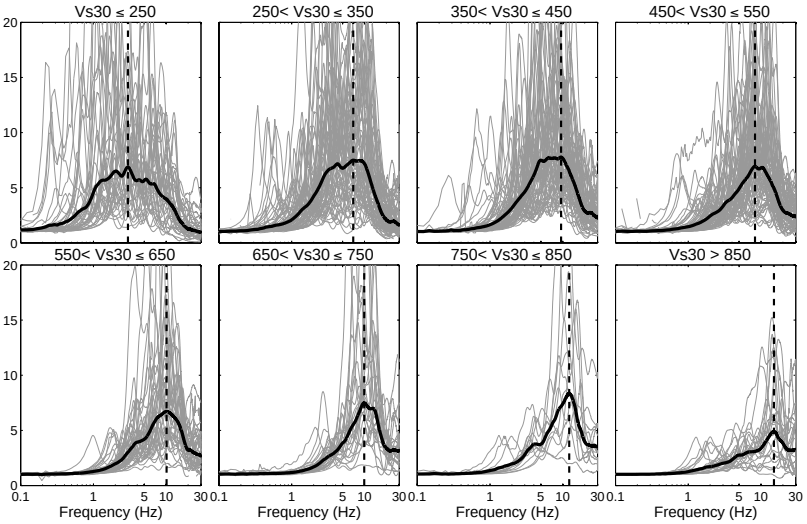


Figure 3.3: *Variability of the mean $BFSR_{emp}$ for different ranges of Vs_{30}. The thick line represents the means of the $BFSR_{emp}$ for each Vs30 bin; the thin lines represent the individual $BFSR_{emp}$.*

defined in chapter 2; namely (f_0, f_{pred}, A_{pred}, B_{30} and B_{100}). Considering that B_{50} was well correlated to both B_{30} and B_{100}, we choose not to use it.

For a specific Vs_{30} range, we investigated the variability of the shear-wave velocity profiles and the $BFSR_{emp}$ for different values of B_{30}, B_{100} and f_0 . We calculated the linear correlations between $log_{10}(Vs_{30})$ with B_{30}, B_{100} and $log_{10}(f_0)$. The results of the linear regressions are summarised in Table 3.1. We used the mean linear correlation and associated standard deviation curves to separate the sites according to the additional parameters for each Vs_{30} range. We separated the sites for each Vs_{30} range into three groups according to the values of B_{30}, (B_{100}, and f_0) (see Figure 3.4): (1) B_{30}/ B_{100}/ f_0 greater than the mean plus half the standard deviation of the linear regression between B_{30}/ B_{100}/ f_0 and $log_{10}(Vs_{30})$; (2) B_{30}/ B_{100}/ f_0 between the mean $\pm$ half the standard deviation; and (3) B_{30}/ B_{100}/ f_0 lower than the mean minus half the standard deviation.

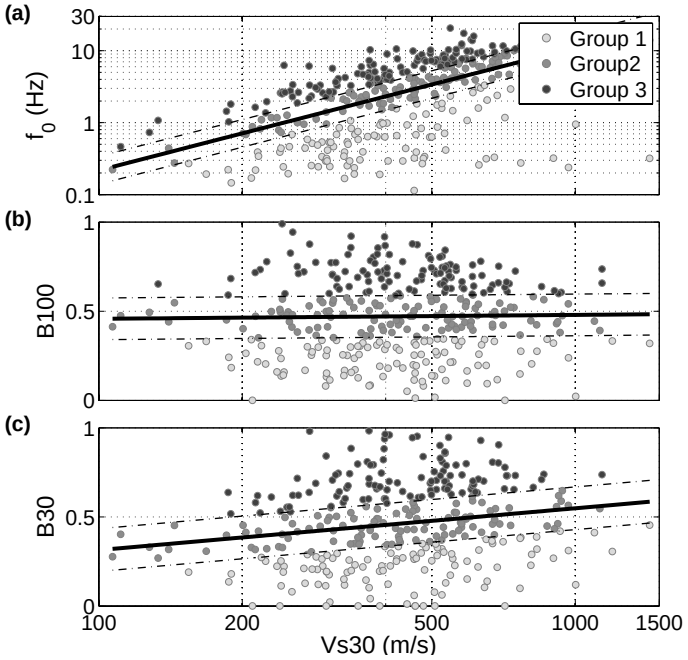


Figure 3.4: *Linear regression between the additional parameters: (a) $log_{10}(f_0)$, (b) B_{30}, and (c) B_{100} with $log_{10}(Vs_{30})$. The dotted lines are the means $\pm$ 0.5 standard deviation. The darker points are the values of the additional parameters that are above the mean $\pm$ 0.5 standard deviation, the middle grey points are the values of the mean $\pm$ 0.5 standard deviation, and the lighter grey points are the values of the additional parameters that are below the mean $\pm$ 0.5 standard deviation.*

Table 3.1: *Coefficients of the linear regression between $\log_{10}(Vs_{30})$, B_{30}, B_{100} and f_0*

$X = a + b\log_{10}(Vs_{30}) \pm \sigma$			
X	a	b	σ
B_{30}	-0.25	0.27	0.22
B_{100}	0.87	-0.17	0.2
$\log_{10}(f_0)$	-3.86	1.62	0.44

3.5 ILLUSTRATION OF THREE KIK-NET SITES

3.5.1 Description of the sites

The relevance of Vs_{30} for the characterisation of site responses was tested on a small set of stations for which the Vs_{30} values were relatively close. The characteristics of the three selected sites are summarised in the Table 3.2. Figure 3.5 (left) illustrates the shear-wave velocity profiles of the KiK-net sites with Vs_{30} close to 400 $\pm$ 20 m/s. Once more, despite their close Vs_{30}, the velocity profiles are very variable below 20 m in depth. The gradient (here down to 30 m; B_{30}) distinguishes well between the profiles with high and low gradients. Among these sites, we selected three that were very different at depth according to B_{30}.

Table 3.2: *Characteristics of the three sites selected to illustrate the variabilities of the seismic site responses for sites with very close Vs_{30}*

Site Name	Vs_{30}	f_0	B_{30}	B_{100}	N_{ev}	M_{JMA}	D_{epi}	Depth
	m/s	Hz					km	km
IKRH01	405	0.3	0.12	0.19	15	3.9-6.1	61-138	1-134
TKSH05	380	2.8	0.39	0.6	26	3.5-4.7	14-100	5-45
NGSH01	398	4.1	0.94	0.69	11	3.7-5.4	50-97	5-15

3.5.2 Comparison of the empirical transfer function

We compared the mean and 68% confidence limits of the $BFSR_{emp}$ of these three selected sites, as shown in Figure 3.5 (upper right). Figure 3.5 (lower right) illustrates the frequencies for which the mean of one site response was statistically different from the other two. We used a statistical test (ttest2 in Matlab software) on the logarithm of $BFSR_{emp}$. This test rejected the null hypothesis that the variables are independent, normally distributed, and with similar means, against the alternative that the means are different at the 5% significance level. This test indicated the frequencies for which the curves are different. These data indicate clearly that the site responses are completely different for the three stations, and consequently the information on the Vs_{30} is far from enough to fully characterise the sites. Through this simple example, we have illustrated that Vs_{30} cannot be used to characterise site responses of specific sites, but is the Vs_{30} enough to classify the sites? The next section is devoted to the analysis of all of the selected KiK-net sites.

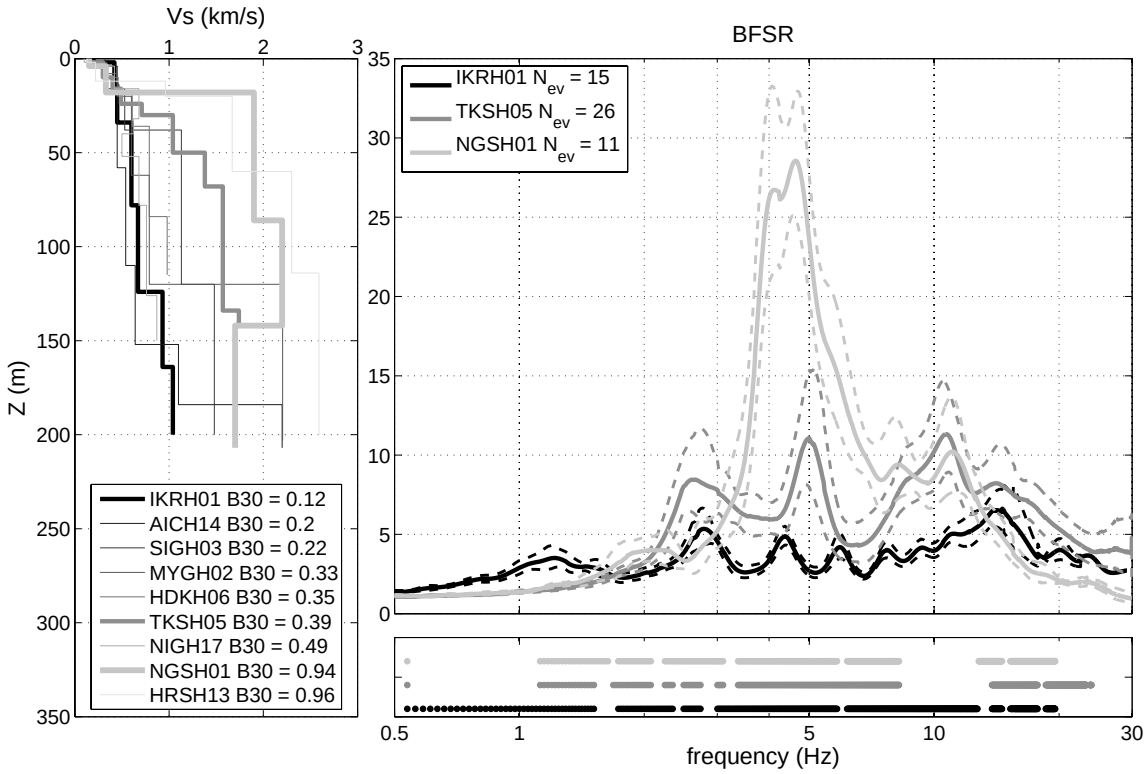


Figure 3.5: *Left: Comparisons of the shear-wave velocity profiles of the KiK-net sites with Vs_{30} of 400 $\pm$ 20 m/s, depending on the value of B_{30} of the Vs profile (grey scale). Upper right: Comparison of the borehole site responses of three sites with different B_{30} means and 68% confidence limits. Lower right: Frequencies for which the means of the site responses of one site are statistically different from the other two.*

3.6 ANALYSES OF THE VS PROFILE AND BFSR$_{EMP}$ VARIABILITIES OF THE 1-D SELECTED SITES

3.6.1 Variabilities of the Vs profiles

Figure 3.6 illustrates the variabilities of the Vs profiles for different Vs$_{30}$ ranges and according to the B$_{30}$. In each Vs$_{30}$ range, the sites were sorted according to the B$_{30}$, as explained in section *On the use of the additional parameters*, on the use of the additional parameters. For each group, we calculated the means and standard deviations of the Vs profiles. The shaded areas in Figure 3.6 show the 68% confidence limits of the Vs profiles in group 2 (section *On the use of the additional parameters*, B$_{30}$ between the mean linear regression of B$_{30}$ with log_{10}(Vs$_{30}$) $\pm$ half the standard deviation). Group 2 is composed of the mean values of B$_{30}$, and should represent the mean behaviour of the Vs profiles at each Vs$_{30}$ range (the so-called 'reference' group). The other two groups (section 4.2., 1 and 3) for the high and low values of B$_{30}$ (Figure 3.6, light and dark grey curves, respectively) represent the two extreme Vs profile behaviours. Group 3 is composed of sites with a low gradient value that are typical of deep sedimentary sites. Group 1, on the other hand, is composed of sites with a high gradient value, which indicates a large impedance contrast at shallow depth. These observations are homotetic over the Vs$_{30}$ ranges. Except for the first Vs$_{30}$ range, B$_{30}$ distinguishes well between group 1 and group 3 of the Vs profiles along the depth, down to 30 m.

3.6.2 Variability of linear site responses

Figure 3.7 illustrates the BFSR$_{emp}$ variabilities for different Vs$_{30}$ ranges and according to the B$_{30}$. Similar to Figure 3.6, for each Vs$_{30}$ range, the BFSR$_{emp}$ was separated into three main groups, with the means and standard deviations calculated. The shaded area represents the 68% confidence limits of the BFSR$_{emp}$ of group 2 (see section 4.2.). The dotted black lines represent the maximum of the means of group 2 (reference group). The dark grey curves (group 1) are the means and 68% confidence limits of the BFSR$_{emp}$ of the sites that have a high gradient value (i.e., a large impedance contrast in the first 30 m of depth). The light grey curves (group 3) are the means and 68% confidence limits of the BFSR$_{emp}$ of the sites that have a low gradient value (i.e., deep sedimentary sites). We see that the mean amplitude of amplification of group 3 is below the mean of the reference group; i.e., below the black dotted line, and starts at frequencies below 1 Hz. Group 3 is characterised by a low amplification spread in a large frequency band, and starts at low frequencies, whereas group 1 is characterised by large amplification at higher frequencies (the mean amplitude of the amplification is above the black dotted line). For each Vs$_{30}$ range, B$_{30}$ separates well the BFSR$_{emp}$ of groups 1 and 3. The BFSR$_{emp}$ of group 2 is very similar to that of group 1.

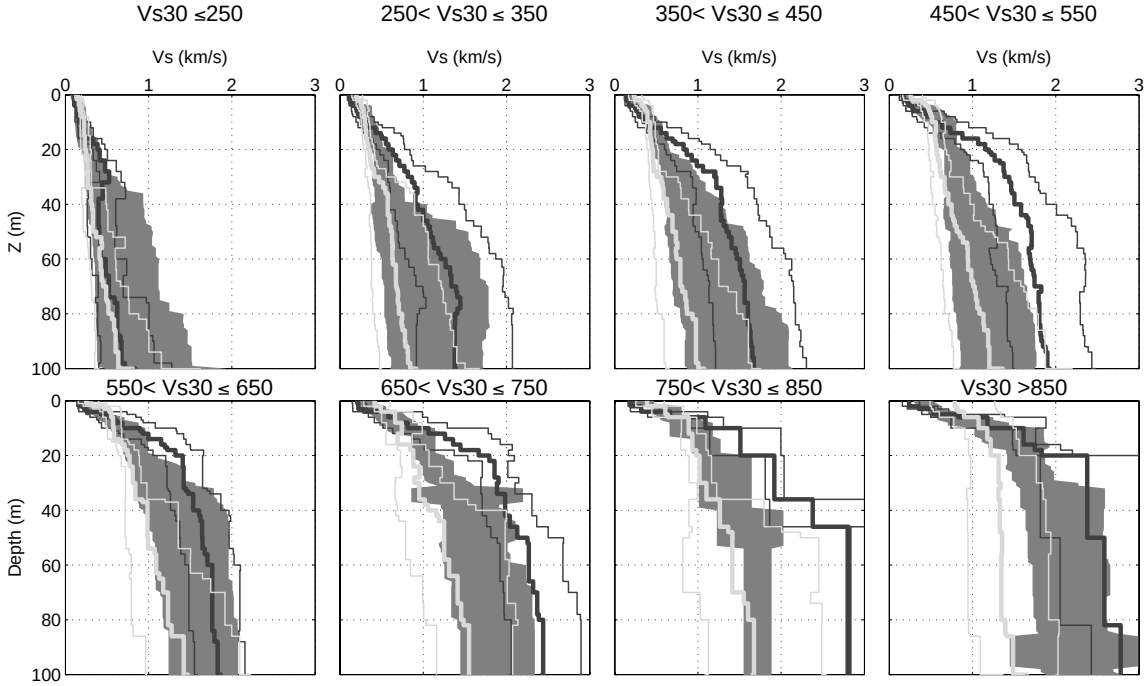


Figure 3.6: *Variability of the shear-wave velocity profiles down to 100 m in depth for different ranges of Vs_{30} and considering three groups of sites that depend on the B_{30}. The light grey curves represent the means and 68% confidence limits of the Vs profiles of the sites in group 1 (B_{30} below the mean minus half the standard deviation of the linear regression between B_{30} and $log_{10}(Vs_{30})$). The grey shaded areas represent the 68% confidence limits of the Vs profiles of sites in group 2 (B_{30} between the mean minus and plus half the standard deviation of the linear regression between B_{30} and $log_{10}(Vs_{30})$). The dark lines represent the means and 68% confidence limits of the Vs profiles of sites in group 3 (B_{30} above the mean plus half the standard deviation of the linear regression between B_{30} and $log_{10}(Vs_{30})$).*

3.6.3 Comparisons between additional parameters

The additional parameters B_{30}, B_{100} and f_0 are not correlated either to Vs_{30} nor one to the other, and they can thus give complementary information. Figure 3.8 (left to right) shows the comparisons of the variabilities of the Vs profiles and $BFSR_{emp}$ according to B_{30}, B_{100} and f_0 for a given Vs_{30} range, here as 250 m/s to 350 m/s. For f_0, instead of separating sites depending on the values of f_0 compared to the linear regression $\pm$ half the standard deviation, we chose a constant value, which is the mean of the linear regression $\pm$ half the standard deviation over the Vs_{30} range. Indeed, as f_0 is highly variable even in a reduced range of Vs_{30} (the linear regression between $log_{10}(f_0)$ and $log_{10}(Vs_{30})$ has a steep slope), taking the linear regression as the limit implies mixing sites with very heterogeneous f_0 .

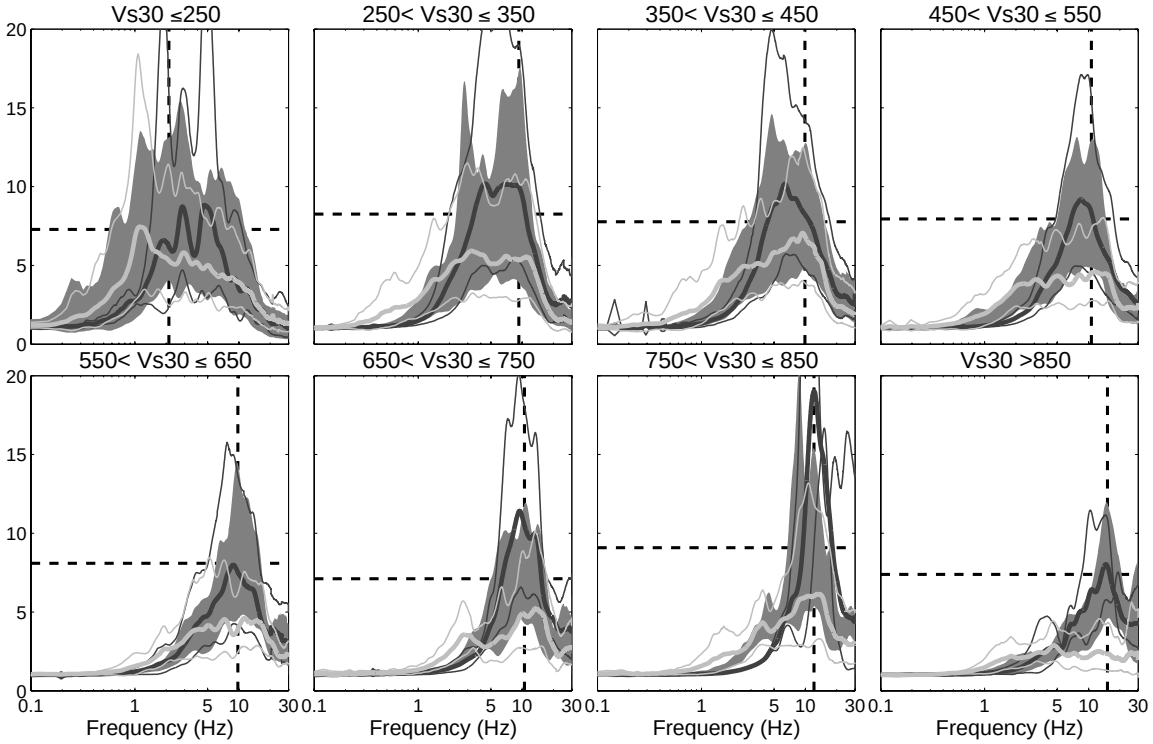


Figure 3.7: *Variability of the BFSR$_{emp}$ mean for different ranges of Vs$_{30}$ and considering three groups of sites that depend on the values of B$_{30}$. The light grey curves represent the means and 68% confidence limits of the BFSR$_{emp}$ for sites in group 1. The grey shaded areas represent the 68% confidence limits of BFSR$_{emp}$ for sites in group 2. The dark lines represent the means and 68% confidence limits of the BFSR$_{emp}$ of sites in group 3. The dotted line represents the frequencies for which the mean of the BFSR$_{emp}$ of group 2 is maximal*

For each additional parameter, Figure 3.8 (a) illustrates the distributions of their values with Vs$_{30}$. Figure 3.8 (b) is composed of two representations: on the left there are the Vs profile variabilities that depend on the values of the additional parameters; and on the right there are the depths for which the mean of each Vs profile group is statistically different from the other two. Figure 3.8 (c) is also composed of two representations: the upper one shows the variability of the BFSR$_{emp}$ according to the values of the additional parameters; and the lower one shows the frequencies at which the means of the BFSR$_{emp}$ of each group are statistically different from the other two.

We quantified these observations by calculating the coefficients of variation (CVs), which are the ratios of the means to the standard deviations. Figure 3.9 has a similar configuration to Figure 3.8, but illustrates the comparisons of the CVs of the Vs and BFSR$_{emp}$ of each group. In Figure 3.9 (b) and (c), the CVs of each group (depending on the values of the additional parameter) are

compared with the global CV; i.e., the CV of the whole set of sites with Vs_{30} between 250 m/s to 350 m/s.

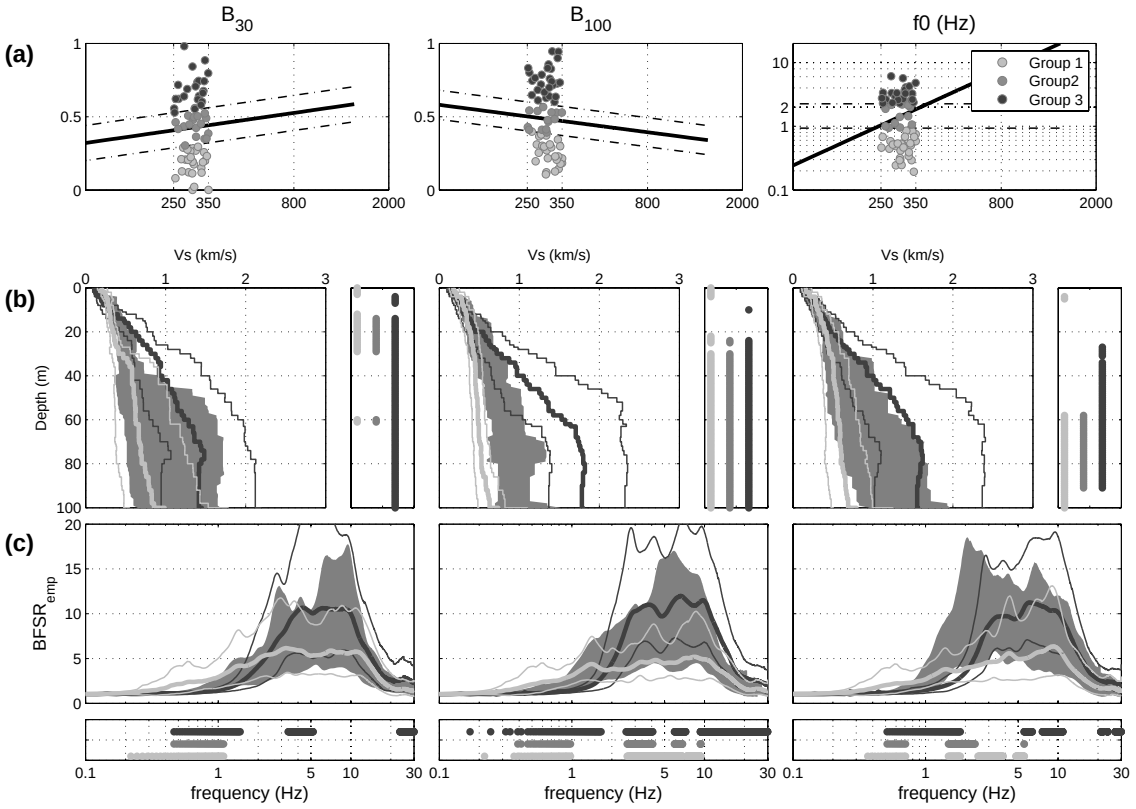


Figure 3.8: *Comparisons of the variabilities of the Vs profiles and the $BFSR_{emp}$, considering B_{30}, B_{100} and f_0, for a given Vs_{30} range, here of 250 m/s to 350 m/s. (a) Each column represents the values of each additional parameters versus Vs_{30} of the sites used. (b) Left: The Vs profile variability depends on the value of the additional parameters. Right: The depth for which the mean of each Vs profile group is statistically different from the other two. (c) Top: The variability of the $BFSR_{emp}$ according to the values of the additional parameters. Bottom: The frequencies at which the means of the $BFSR_{emp}$ of each group are statistically different from the other two.*

3.6.3.1 Variabilities of the Vs profiles

B_{30}: B_{30} separates well the profiles down to 30 m in depth. It reduces the CVs above 40 m in depth globally, but below this depth, the CVs are higher for groups 2 and 3, compared to the global CVs.

B_{100}: Compared to B_{30}, B_{100} separates the Vs profiles at depth even better. From 20 m to 100 m in depth, the means of the three Vs profile groups are statistically different. The reductions in

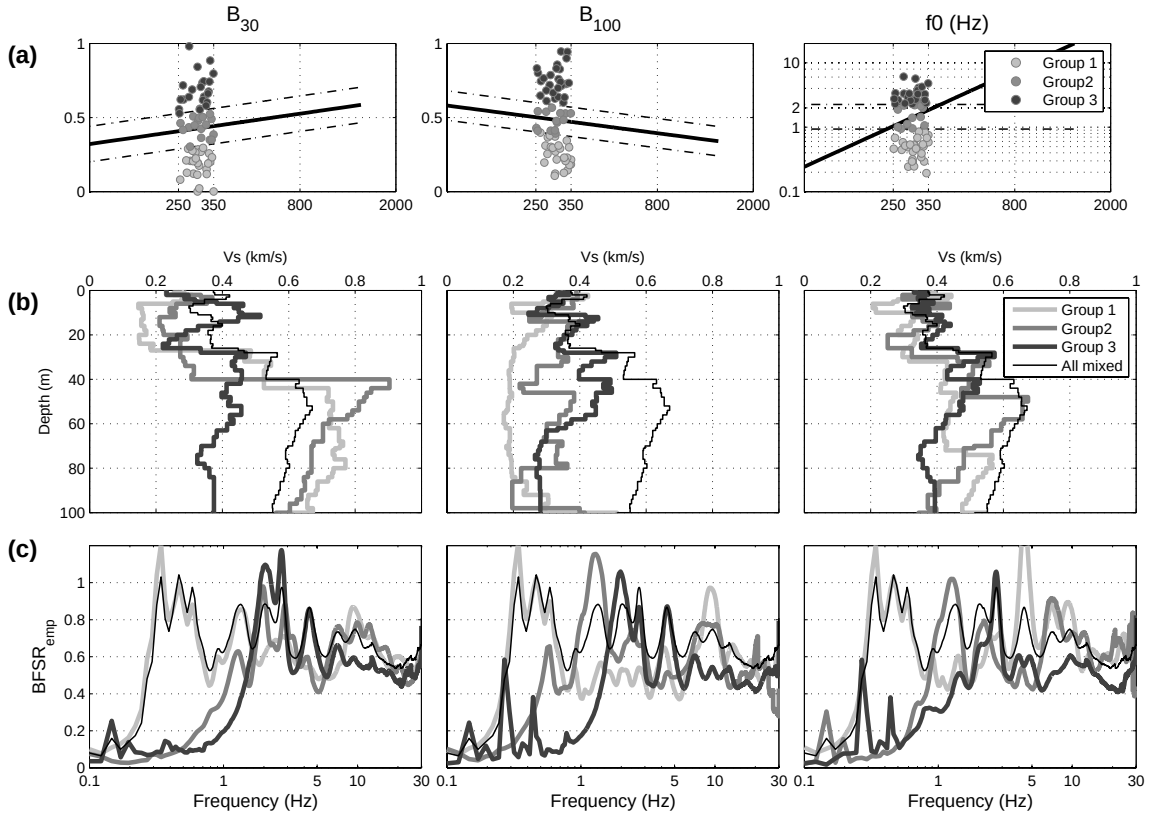


Figure 3.9: *Comparisons of the coefficients of variation of the Vs profiles and the BFSR$_{emp}$, considering Vs$_{30}$ alone (black thin curves) and three groups of sites that depend on B_{30}, B_{100} and f_0, for a given Vs$_{30}$ range, here as 250 m/s to 350 m/s.*

the CVs on the Vs profiles are clear from 30 m in depth for all three of the groups, compared to the global CVs.

f$_0$: f$_0$ separates the Vs profiles of group 1 from 20 m in depth, and from 60 m in depth for the other two groups. The CVs are reduced for depths ranging from 40 m to 100 m.

Whatever the additional parameter used, the Vs profiles are separated into three groups that have specific behaviours with depth: Group 1 shows a large impedance contrast in the shallow layers (before 60 m in depth), group 3 shows a deep sedimentary soil column, and group 2 is a reference group that is a mix between the Vs profiles of groups 1 and 3.

3.6.3.2 Variabilities of the BFSR$_{emp}$

B$_{30}$: We expected that as a parameter that was calculated with shallow Vs information, B$_{30}$ would separate the BFSR$_{emp}$ at high frequencies; however, the separation is statistically valid for the three groups of sites only for frequencies between 0.7 Hz to 1.1 Hz. Group 1 is statistically different from the other two also at medium and high frequencies, although in a limited frequency bandwidth (3 Hz to 5 Hz, and 22 Hz to 30 Hz). Group 3 is statistically different at frequencies below 0.7 Hz. The reductions in the CVs (compared to the global CV) are important below 2 Hz for groups 1 and 3, and significant for group 1 above 5 Hz, but not clear for group 3.

B$_{100}$: B$_{100}$ appears to be the most powerful parameter for separating the site response curves into the largest frequency bandwidths for all three groups (0.4 Hz to 1 Hz, and 2.2 Hz to 10 Hz). Group 1 is statistically different from the other two at almost all frequencies. The CVs of group 3 are reduced from 0.8 Hz, except for a limited frequency bandwidth close to 10 Hz. For group 1, the CVs are also reduced at all frequencies, except between 1.5 Hz to 3 Hz. For group 2, the CVs are reduced, except at two frequency bandwidths between 1 Hz to 3 Hz and around 10 Hz.

f$_0$: As far as f_0 is concerned, the separation of the BFSR$_{emp}$ is valid for the three groups in a limited frequency bandwidth (0.4 Hz to 0.6 Hz, 1.5 Hz to 2 Hz, and 5 Hz to 6 Hz), although group 1 is different from the other two at high frequencies, and group 3 is different at medium frequencies (1.5 Hz to 7 Hz). The CVs are reduced for group 3 at all frequencies. For group 1, the reductions are not significant, except between 1 Hz to 3 Hz, and for group 2, the reductions are large below 1 Hz.

Similar to the Vs profiles, the three additional parameters separate the BFSR$_{emp}$ into three groups that have specific behaviours: group 1, large amplification from 3 Hz; group 2, medium to large amplification from 1 Hz; and group 3, low amplification from 0.5 Hz. These behaviours are more or less pronounced depending on the additional parameters used to classify the sites. The parameter f_0 is the one that shows these three BFSR$_{emp}$ behaviours most clearly.

3.7 CONCLUSION AND DISCUSSION

The scope of this chapter was to determine the influence of the 1-D Vs profiles on site responses. We selected 351 sites from the KiK-net database, such that the empirical linear site responses were close to the numerical 1-D linear evaluation. This selection was based on comparisons of the fundamental resonance frequencies. We analysed the variabilities of the Vs profiles and the BFSR$_{emp}$ (the empirical borehole Fourier spectral ratio) with Vs$_{30}$, and used the additional parameters that

reflect the Vs evolution with depth. We used the gradient of the Vs profile calculated using only the first 30 m (B_{30}) of the Vs profiles, to be consistent with the depth of the investigation required to obtain Vs_{30}. We also used the maximal depth common to all of the selected Kik-net site profiles; i.e., 100 m (B_{100}). In addition, we used a third parameter that was independent of the Vs profile, the fundamental resonance frequency (f_0), which is calculated using earthquake recordings.

We found that B_{30} can separate the Vs profiles at shallow depths into three groups; the reductions in the CVs for all of the Vs profiles was nevertheless significant only for group 1 (sites with shallow high velocity contrast). The reductions in the $BFSR_{emp}$ CVs were important for group 3, and were limited to a frequency range (below 1 Hz) for group 2. B_{30} is a parameter that did not require more information than Vs_{30} to be obtained (same depth of investigation), and it is used as a complementary parameter to Vs_{30}. B_{30} can improve evaluations of site responses, especially for sites with high velocity contrast in the superficial layers.

In the next section, we showed that this parameter is also a relevant parameter for non-linear site response assessment. We found that B_{30} is well correlated to a site-specific peak ground acceleration threshold that triggered non-linear behaviour, which suggests that sites with high shear-wave velocity contrast close to the surface can trigger non-linear behaviour at low input motion peak ground acceleration values.

When using B_{100}, the coefficient of variation for the Vs profiles and for the $BFSR_{emp}$ were even more reduced. The Vs profiles were well separated into three groups of sites with specific behaviours of the Vs profile with depth. Each of these groups was characterised by a specific borehole site-response curve, which suggests that the Vs profiles at depth have a significant influence on the site responses.

We also found that f_0 can distinguish between the site-response curve behaviours for a given Vs_{30}. The relevance of this parameter in association with Vs_{30} for site-effect assessment has already been underlined in several studies, as mentioned in section 2 (Introduction).

We mentioned earlier that we do not correct the empirical transfer function to a common reference. Nevertheless, we tested a more drastic selection of KiK-net sites having a Vs at the down-hole station greater than 1500 m/s and performed the same comparison of inter-site Vs and BFSR variabilities. However, we observed that this selection did not reduce the variability in each sites group which confirm our first choice.

To conclude, with B_{100}, we have shown that the distributions of the Vs profiles at depth have strong influences on the linear site responses. The Vs profile is, however, difficult to obtain precisely down to this depth, and it would be expensive to do so. Thus, other site characterisations are of interest. In the present study, we proposed B_{30}, which does not require more effort than Vs_{30} to be obtained. B_{30} turns out to be a relevant complementary parameter to Vs_{30} for site characterisation and site-effect assessment. We also underlined the similar potential for f_0, which can be obtained from low-cost, non-invasive methods and which is independent of Vs_{30}.

Chapter 4

Inter-event site-response variability

Sommaire

ABSTRACT

This chapter investigates the effects of the non-linear behaviour of soils on site responses, through various earthquake recordings from the KiK-net database in Japan. This network is composed of more than 688 surface/ borehole instruments, where a characterisation of the shear and compressive wave velocity profiles down to the borehole depth are available. We selected events with a peak ground acceleration (PGA) at the downhole station of < 10 cm/s^2 to characterise the linear soil behaviour by computing the surface to downhole spectral ratios at each site. All events with a greater PGA are believed to trigger non-linear soil behaviour.

To describe the effects of non-linear soil behaviour on site responses for each event, we used the percentage of modification of the site-response curve (either amplification or de-amplification) compared to the linear evaluation (PNL_{ev}, percentage non-linearity) and the associated shift frequency (Sh_{ev}). These parameters are used to estimate the probability that non-linear site responses are significantly different that their linear counterparts. We find that regardless of the site, this probability is important even for low input motion PGA (>30 cm/s^2 at the downhole sensor). This indicates that non-linear soil behaviour must be taken into account in site-response evaluation for moderate to strong motion.

In addition, for 54 sites of the KiK-net database that had recorded at least two strong events (PGA at the downhole station $>$ 50 cm/s^2), we define four additional parameters that characterise the effects of non-linear soil behaviour on site responses for each site: (1) A PGA threshold (PGA_{th}), defined as the PGA value for which the PNL_{ev} >10%, (2) a site-specific PNL for a PGA of 50 cm/s^2 (PNL_{site}), (3) a site-specific shift of the predominant frequency for a PGA of 50 cm/s^2 (Sh_{site}) and (4) a frequency from which we observe de-amplification between nonlinear and linear site responses (fNL). We observe that non-linear soil behaviour can increase the amplification at frequencies below fNL. We find that fNL lies inbetween the fundamental and the predominant resonance frequencies of the site response, and that sites with a Vs contrast close to the surface can trigger non-linear behaviour at a lower input motion PGA threshold. These results suggest that non-linear soil behaviour occurs mostly in the superficial soil layers. Furthermore, by investigating the non-linear soil behavior on earthquake H/V spectral ratios at the surface, we find that they can give satisfactory results (equivalent to the analysis of borehole site responses) for the evaluation of the fNL frequency and shift frequency (Sh_{site}), which indicates that part of the results obtained in this study can be extended to other databases without downhole sensors.

This chapter is mainly based on a study that was submitted to the BSSA in December 2012, under the title: "Assessing non-linear behavior of soils in seismic site response: Statistical analysis of KiK-net strong motion data" (Régnier, Cadet, et al., 2013).

4.1 INTRODUCTION

It is widely recognised that seismic waves can be locally amplified due to soil layering and basin geometry (e.g., Bard & Bouchon, 1985). These so-called site effects can dramatically increase the seismic motion at the surface and the subsequent damage to man-made structures. The precise evaluation of site effects is therefore of great importance for the earthquake engineering community.

Increasing numbers of ground-motion observations from low to high amplitudes contribute to the improving of the knowledge of the physical phenomena of wave propagation, and for the modelling of sediment responses (e.g., Field et al., 1997). The linear theory of wave propagation is valid in a large number of seismological problems that involve small deformations in stiff geomaterials. However, the prediction of site effects during strong motion can involve soft materials and large deformation. Under such conditions, laboratory tests (e.g., Ishibashi & Zhang, 1993) and vertical-array recordings have demonstrated that the soil behaviour in the stress-strain space is non-linear and shows hysteresis (e.g., Zeghal & Elgamal, 1994; Assimaki et al., 2008).

Empirical site responses are usually evaluated using the spectral ratio between the simultaneous recordings on sediments and on a nearby rock site (the so-called reference site). When applying this technique, the main issue to overcome is the selection of a reliable reference site. The reference site must not amplify seismic waves, and it should be close enough to the studied site so that the travel path from the seismic source remains equivalent for both sites. Vertical arrays of accelerometers with a downhole reference site overcome this distance difficulty. Nevertheless, it is imperative to note that downhole data (recorded at the bottom of a borehole) includes some spurious interference that is mainly due to the down-going wavefield (e.g., Bonilla et al., 2002; Cadet, Bard, & Rodriguez-Marek, 2012). In areas of high seismic activity, vertical arrays of accelerometers have provided direct evidence of non-linear soil behaviour. In their retrospective of studies on strong ground motion, I. Beresnev and Wen (1996) showed that the effects of nonlinear soil behaviour on seismological observations are significant. More recently, evidence that there is non-linear soil behaviour was shown by comparative site response curves computed from weak and strong motions (e.g., Wen, 1994; Iai et al., 1995; Satoh et al., 1995; Sato et al., 1996; Aguirre & Irikura, 1997; Field et al., 1997; Noguchi & Sasatani, 2008; Wen et al., 2011). Accelerometric data have also been used to determine the shear modulus degradation and damping curves, both in Taiwan (e.g., Glaser & Baise, 2000; Zeghal et al., 1995) and Japan (e.g., Pavlenko & Irikura, 2003; Kokusho, 2004; Pavlenko & Irikura, 2006). Finally, non-linear soil behaviour has also been directly observed in acceleration time histories, where the alternance of dilatancy and the generation of pore water pressure can induce spikes that are directly visible in acceleration time histories (e.g., Bonilla et al., 2005, 2011)

In low seismicity areas, surface strong ground-motion recordings are limited in number, or are even non-existent. However, non-linear soil behaviour should also be taken into account to produce accurate ground-motion predictions. In the present study, our approach was to observe, in a statistical way, the relationships of soil, site responses, and incident motion parameters, with the effects of non-linear soil behaviour on seismic site responses. The aim of this section is four-fold: (1) to define the parameters that characterise the effects of non-linear soil behaviour on site responses (for each event and as site specific, considering all of the recordings at a site); (2) to define the parameters that characterise the site and the incident motion (relevant for non-linear site-response assessment); (3) to define the correlations between these two sets of parameters; and (4) to define the level of intensity of the incident motion where these phenomena have significant effects on site responses.

We chose the well-characterised KiK-net boreholes in Japan to empirically evaluate non-linear site responses. After describing the database, the method to compute the site responses is presented. As an example, to understand the effects of non-linear soil behaviour on site responses, we present these effects during the 11 March, 2011, Tohoku great earthquake. Then we extend these observations to all of the available seismic data. In the first section, we identify the parameters that describe the effects of non-linear soil behaviour on site responses for each event and for each site (the non-linear parameters). Secondly, we identify the parameters that can explain these effects: i.e., the soil, the linear site response, and the incident ground-motion parameters.

Finally, we perform two statistical analyses. In the first, we use all of the combined KiK-net sites, and we correlate the non-linear parameters for each event with the incident ground-motion parameters (i.e., PGA). The second statistical analysis is carried out with the 54 KiK-net sites that have recorded strong events. Here, we use multivariate statistical analysis to qualitatively evaluate the links between the non-linear parameters for each site and the soil and linear site-response parameters, followed by univariate regression to quantitatively evaluate the relationships initially found.

Inter-event linear site-response variability is due to source and path effects that are associated with complex site responses. This complexity of the site response can be due to several parameters: the site configuration is not 1-D or the materials are not isotropic. The main hypothesis of this study is that the linear characterisation, which is performed with a large number of recordings, contains all of the variability defined above. In other words, the difference between site responses computed from strong motion and linear site responses is mainly due to non-linear soil behaviour.

4.2 ELEMENTS OF SOIL NON-LINEAR BEHAVIOUR

The linear theory discussed in the previous chapter is adequate for a large number of problems in seismology, specifically when the materials in which the waves propagate are rigid and the deformation is weak. However, site-effect assessment implies seismic waves that propagate in soft soil, and for strong incident ground motion there will be high deformation in the material. Non-linear soil behaviour modifies key parameters of the seismic-site response (i.e., shear modulus, damping ratio), and therefore it might have a strong influence on the evaluation of site effects. The scope of this chapter is to present elements of the non-linear soil behaviour, the way it is accounted for in numerical simulations, and the expected effects it can have on site responses.

4.2.1 Non-linear soil behaviour characterization

In geotechnical engineering, non-linear soil behaviour is well established. The traveling of the seismic waves in the soil induce an additional shear stress to the initial equilibrium of the soil (for a 1-D layered soil column, the soil particles are only subjected to vertical effective stress). The shear stress induce a shear strain which is accompanied by a volumetric strain. For an elastic material, the volumetric strain is zero. However, when the volumetric strain is no longer zero then the material is no longer elastic. The soil response under cyclic loading (which represents the seismic loading) depend on the properties of the cyclic loading (History of the cylic loading, maximal amplitude) and on the soil composition. Geotechnical engineering usually divided into two kind of soil depending of the cohesion: cohesionless soil are granular soil (sand or gravel) which strength mainly depend on friction and cohesive soil, such as clay, whose particles adhere to each other by means of adhesive and cohesive forces. The dynamic soil behaviour depend on the type of soil:

- cohesionless soil: In loose coheisonless saturated soil subjected to cyclic loading in undrained condition (earthquake loading and low permeability of the soil) liquefaction may occur. The process of liquefaction is characterise by a rise in pore water pressure against cycles in a cyclic undrained test which eventually create a decrease of the effective stress. The cyclic degradation of the dense sand is not as severe as that for the loose sand. This process depends on the initial effective stress field and on the geologic age of the formations (for more details see (Kramer, 1996), pages 348). When the soil is not saturated, there is no pore pressure but the soil may suffer from volume changes. Cyclic shearing densifies on medium dense to dense sand. Other aspects of cohesionless soil under cycle loading can be encounter such as particle crushing (changes the grain-size distribution of a sample see (Marketos & Bolton, 2006)), low compressibility soil and cavitation.

- cohesive soil: The major difference between cohesionless and cohesive soils is with permeability, for cohesive soils this is several orders of magnitude less than that for a cohesionless

soil. For a saturated cohesive soil two limiting cases are considered in geotechnical design, the so-called undrained and drained conditions (short term and long term). During earthquake loading, it is short term (i.e. undrained) conditions that are usually used. The degradation of the shear modulus occurr from a thereshold strain of 5.10^{-5}, whereas it starts immediately for cohesionless soil (Anderson & Richart Jr, 1976). Vucetic and Dobry (1991; 1992) showed that the degradation mainly depend on the plasticity index (PI) and the threshold strain from which non-linear behavior begin increase with PI.

Many laboratory tests has been developed to determine the stress-strain relationships of materials under cyclic loading, among which the following are the most used worldwide (more details on each of the previously described trials can be found in the book by Kramer (Kramer, 1996), pages 216 to 225) :

Cyclic triaxial test : Under drained or undrained conditions, the soil dynamic parameters (e.g., shear modulus, damping ratio) can be extracted from the hysteresis curves. Liquefaction or cyclic mobility can also be studied in the case of undrained trials.

Shearing test : This test determines the shear modulus of the soil and the hysteresis loops in the case of cyclic loading (quasi-static loadings). Under undrained conditions, this test determines the potential of a soil to liquefy.

Resonant column : This test measures the shear modulus and damping through the calculation of the resonance frequency of the soil sample. This test measures the decrease in the movement amplitude in free oscillations once the sample is loaded at the resonance frequency.

(Ishihara, 1996) (1996) synthesizes in a table the in situ and laboratory tests that can be performed depending on the strain level. Part of this table has been reproduced in Table 4.1

The laboratory tests show that, for strong cyclic loading, the relationship between (σ) and (ϵ) is no longer linear. The shear modulus becomes dependent on the deformation. At the same time, for strong deformation, the material can have plastic behaviour, meaning that permanent deformation can occur. This plastic behaviour induces a hysteresis behaviour, with the path of unloading and re-loading different from the loading path.

Several studies have used laboratory tests to investigate the evolution of hysteresis curves. They have defined constitutive models that can reproduce the observed hysteresis curves, with analytical formulation of the backbone curve and the laws that describe the loading and un-loading path. Figure 1.8 (c) in page 45 illustrates the comparison between the linear (blue) and non-linear (red) behaviours. In Figure 1.8, the non-linear soil behaviour (i.e. the analytical formulation of the backbone curve) is represented by the hyperbolic model (Ishihara, 1996; Duncan & Chang, 1970), with the application of the extended Masing rules (details of these models are given in the following

Table 4.1: *Material properties measurements depending on the strain.*

Magnitude of strain		10^{-6} 10^{-5}	10^{-4} 10^{-3}	10^{-2} 10^{-1}
Phenomena		Wave propagation, vibrations	Craks, differential settelment	Slide, compaction, liquefaction
Mechanical model		Elastic	Elasto-plastic	Failure
Soil parameters involved		Shear modulus, damping poison ratio		internal friction angle, cohesion
In-situ measurements	Seismic wave method	<——>		
	In-situ vibration test	<————>		
	Repeated loading test	<————————>		
Laboratory measurements	Wave propagation, precise test	<——>		
	Resonant column precise test	<————>		
	Repeated loading test	<————————>		

section). During cyclic loading, there is a reduction in the shear modulus (which is given by the tangent of the backbone curve) and an increase in the hysteresis damping, given according to the following equation:

$$\xi = \frac{W_d}{4\pi W}$$

where, W_d is the area of the hysteresis loop. The damping ratio (equal to 1/(2Q)) is the ratio of the elastic energy created in one cycle cycle (represented by the triangle area in Figure 1.8(c)) with the dissipated energy per cycle symbolised by the hysteresis loop area (Figure 1.8 (c)).

Simplified curves mimicking the main effects of non-linear soil behaviour have also been defined. These simplified curves represent the degradation of the shear modulus and the increase in the attenuation evolution with deformation. Ishibashi and Zhang (1993) used laboratory measurements to provide an analytical formulation for the aforementioned curves for sandy and clayey soil that depended on plasticity index and the effective stress in the soil.

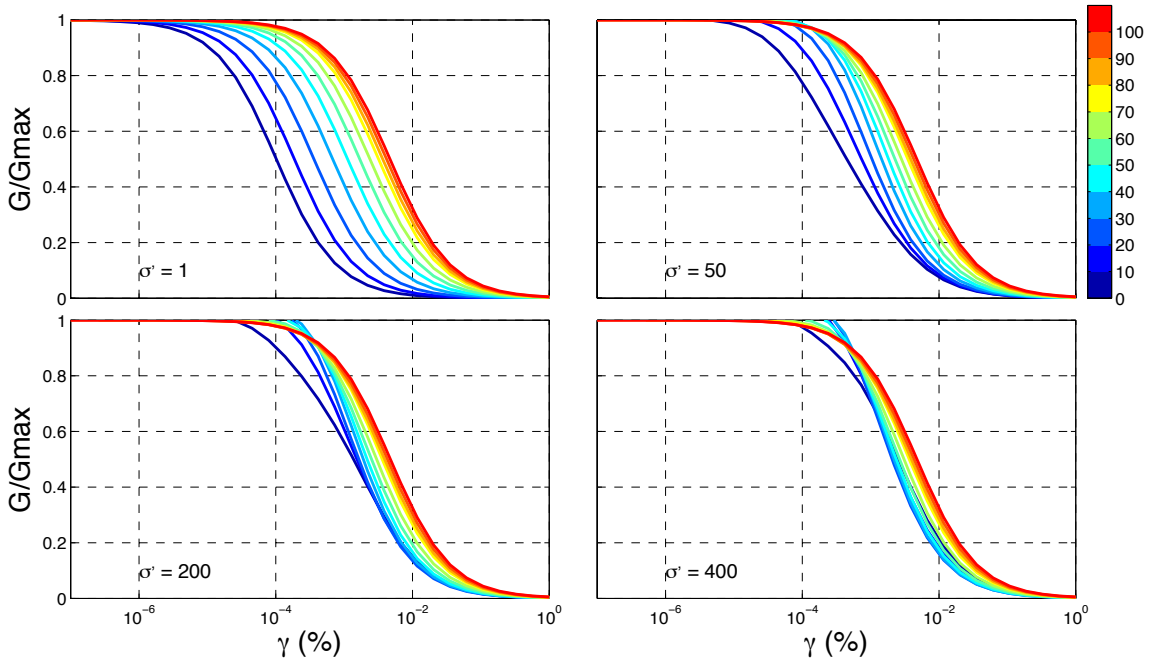


Figure 4.1: *Dependance of the shear modulus degradation curves with plasticity index and confining pressure in the Ishibashi and Zhang (1993) curves.*

4.2.2 Non-linear soil behaviour modelling

4.2.2.1 Visco-elastic equivalent linear models

The visco-elastic equivalent linear model is a visco-elastic linear model (similar to that which was presented in the previous chapter) that is integrated into an iterative loop. This technique estimates the response of a soil column by computing the shear moduli and damping ratio compatible with the deformation at the middle of the layers (Schnabel et al., 1972). The iterations stop when

the shear modulus converges with a pre-defined tolerance value (Kramer, 1996; Ishihara, 1996).

This model is widely use by the engineering community as it is easy to implement and it has a restricted number of input parameters. Nevertheless, this model is not based on constitutive model of soil with a physical model and a large number of studies have shown that this method is limited to medium deformation and that it can underestimate the non-linear soil behaviour effects that occur at large strain (e.g. Kramer, 1996). Also, this model cannot simulate more complex phenomena, such as liquefaction or cyclic mobility. These are ones of the reasons why non-linear models were developed, which can instantaneously adjust the material properties to the level of strain and the loading history, as defined by constitutive models that link (σ) to (ϵ) .

4.2.2.2 Non-linear models

For non-linear models, the true hysteresis soil behaviour is simulated by the use of constitutive models that mimic the experimental hysteresis curves or the shear modulus degradation and damping curves. The constitutive models, that rely on mechanicals models such as generalized Maxwell model, are defined by an analytical formulation of the backbone curve shown in Figure 1.8c, and by rules that describe the loading, re-loading and un-loading paths. In practice, the constitutive models are implemented in non-linear numerical simulations using mechanical models.

4.2.2.2.1 Constitutive models

Backbone curves

backbone curve of Figure 1.8c represents the path walk in the stress-strain during the initial loading. A large number of analytical functions have been developed to simulate the laboratory test results, from simple models to more sophisticated ones. In this section, five backbone curves are presented with gradually increasing complexities.

Multilinear : In the multilinear model, the relationship between (σ) and (ϵ) comes directly from the shear modulus degradation curve. This is a linear function that is in part defined by the following equation: $\sigma_i = G_i.\epsilon_i$, where G_i is the ith value of the shear modulus linked to the deformation ϵ_i. The damping is given by the following equation (4.1):

$$\xi = \frac{2}{\pi}.\left(\left[\sum_{j=2}^{i}(G_j.\epsilon_j + G_{j-1}.\epsilon_{j-1})(\epsilon_j - \epsilon_{j-1})\right]/G_i.\epsilon_i^2 - 1\right) \tag{4.1}$$

This model can be calibrated to perfectly fit the experimental shear modulus degradation curve. However, the damping is usually overestimated compared to laboratory measurements.

Hyperbolic (Ishihara, 1996), (Duncan & Chang, 1970): In the hyperbolic model, the relationship between (σ) and (ϵ) follows a hyperbolic law that is limited at high strain (in both directions).

For weak deformations, the tangent to the backbone curve is G_{max}, the maximum shear modulus. At high strain, the backbone curve is a constant that is equal to σ_{max}, which indicates the maximum material resistance. Equation 4.2 shows the analytical formulation of this model.

$$\sigma = G_{max}.\epsilon.(1 + \frac{\epsilon}{\epsilon_r}) \tag{4.2}$$

where, $\epsilon_r = \sigma_{max}/G_{max}$ is the reference deformation, which would be reached at the maximum material resistance if the material had a purely linear behaviour. The damping is given in Equation 4.3. The limited number of parameters of this model (G_{max} et σ_{max}) do not allow it to perfectly follow the experimental hysteresis curves, and the damping is usually overestimated at high strain.

$$\xi = \frac{4}{\pi}.\frac{1}{1 - G/G_{max}}.[1 + \frac{G/G_{max}}{1 - G/G_{max}}.ln(\frac{G}{G_{max}})] - \frac{2}{\pi} \tag{4.3}$$

Modified hyperbolic (Matasovic & vucetic, 1993), (Matasovic & vucetic, 1995): This model, which is also known as the MKZ model, is a hyperbolic model with one additional parameter, (β, s et σ_{max}). The backbone curve is defined by Equation refeq-4-6:

$$\sigma = G_{max}.\epsilon.(1 + \beta.\frac{\epsilon}{\epsilon_r})^s \tag{4.4}$$

Ramberg-Osgood (Faccioli & Vitiello, 1973), (Streeter et al., 1974): This model is very close to the MKZ model, and it is defined by three parameters, (ϵ_r, α, r). The backbone curve is given in Equation 4.5, and the damping by Equation 4.6.

$$\sigma = \sigma_{max}.\frac{\epsilon/\epsilon_r}{1 + \alpha.|\frac{\sigma}{sigma_{max}}|^{r-1}} \tag{4.5}$$

$$\xi = \frac{2}{\pi}\frac{r-1}{r+1}(1 - \frac{G}{G_{max}}) \tag{4.6}$$

Generalised hyperbolic (Hayashi et al., 1992): Finally, the generalised hyperbolic model is defined by four parameters. Thus, the hysteresis curves from this model will be closer to the experimental data. Nevertheless, there is still the need to be very careful when defining the values of the parameters, as the parameters have a physical meaning. The backbone curve equation can be found in Assimaki et al. (2008).

Definition of the loading/ un-loading paths

Constitutive models that are currently used in non-linear simulations of seismic-wave propagation (e.g. Hardin & Drnevich, 1972; Pyke, 1980) are based on the extended Masing rules (Masing, 1926). Initially, the Masing rules comprised two criteria (1 and 2), although these are usually extended with additional criteria (3 and 4):

1. For each part of the reverse loading, the shear modulus is equal to the maximal tangent of the initial loading (the backbone curve);

2. The un-loading and re-loading paths are duplicates of the backbone curve transformed and scaled by a constant factor (K_h) of two during the complete loading;

3. If the loading or un-loading curves go past the maximal deformation that the material experiment the curves intercept the backbone curve, then the path follows the backbone curve;

4. If the loading or un-loading curves cross a previous cycle, then the path follows this previous cycle.

These criteria are illustrated in Figure 4.3, which is taken from the Kramer book (Kramer, 1996). From A to B in Figure 4.3, for the initial loading, the stress/ strain curve follows the backbone curve. Then, from B to C the stress is reversed. The stress/ strain curve follows the un-loading curve until point D. Then, at point D, the stress is again reversed, and the stress/ strain curve follows the re-loading path until it intercepts the backbone curve at point E, and follows the path from E to F. In the following, the stress/ strain curve follows the loading and un-loading path as defined by the first two rules of the Masing criteria.

Other loading, re-loading and un-loading paths have been proposed, among which we can cite the following: (Pyke, 1979), (Li & Moshell, 1993) and (Bonilla et al., 1998).

In Assimaki et al. (2008), the five models previously presented were implemented and tested for a finite-difference method. They found that if several constitutive models are used to model the non-linear surface ground motion, then the deviation between the observations and the simulations was around 15% to 20% whatever the level of the incident motion. They also showed that true non-linear simulations give closer results to observations than the equivalent linear method.

4.2.2.2.2 Mechanical models

The mechanical models used to represent non-linear soil behaviour can be defined by a set of perfectly plastic elastic strings (Iwan, 1967), the number and properties of which are adjusted to follow the constitutive model. This physical model can also be defined with visco-elasticity using a set of Maxwell models in parallel (Gélis & Bonilla, 2012).

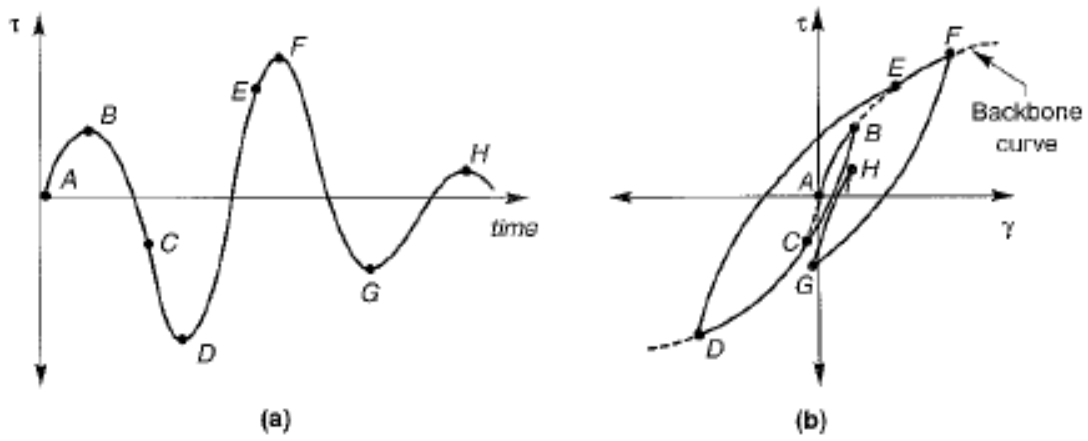


Figure 4.2: *Extended Masing model. (a) Loading history, (b) hysteresis curves. From Kramer (1996).*

4.3 QUANTIFICATION OF THE NON-LINEAR EFFECTS

In this section, the relevant parameters to quantify the effects of non-linear soil behaviour on site responses are presented (the non-linear parameters). We first present the effects of this phenomena that are expected according to non-linear wave propagation theory. Then, we illustrate these effects through a comparison of the site responses computed using the main event and those using the aftershocks, of the Tohoku earthquake.

4.3.1 Expected effects of non-linear behaviour on site responses

For strong ground motion, below the deformation level that triggers soil failure (shear stress above the shear strength), the soil is expected to behave non-linearly. In this section, we shall consider first the effect of soil non-linear behavior on ground motion at the surface in time domain then, in site response in the frequency domain.

4.3.1.1 Time domain surface ground motion

In cohesion less soil in undrained condition, cyclic mobility (in dense soil) or even liquefaction (in loose soil) may occur. Such soil failure can created large permanent deformation. Bonilla et al. (2005) showed that soil nonlinearity can be directly observed in acceleration time history. They showed that for cohesionless soil, spikes can be observed in the time domain that created large accelerations at high frequency. The occurrence of these spikes is explain by a partial recovery of

the shear strength in the soil during cyclic loading which is related to dilatancy and pore pressure generation.

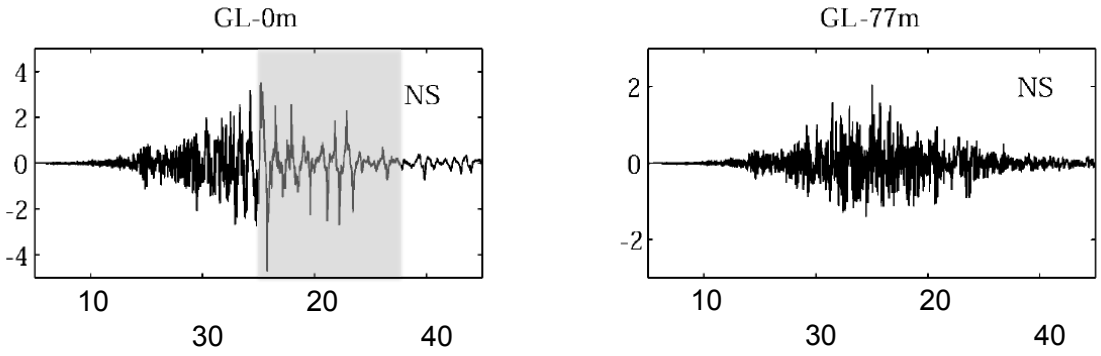


Figure 4.3: *Surface (GL-0 m) and borehole (GL-77 m) acceleration time histories for a dense sand deposit during the 1993 Kushiro-oki earthquake. The gray area indicate "the the spiky repetitive waveform that dominates the surface" from Bonilla (2005).*

At the seams time, some studies have shown that because the soil have a limited shear resistance, the input motion that can be transmitted to the ground surface is also limited (e.g. PECKER, 2005). Indeed, when failure occurs at any depth in the soil, the incident seismic motion is filtered and is no longer transmitted too the upper soil layers.

4.3.1.2 Frequency domain is site responses

In the constitutive models defined above the main effects of nonlinear soil behaviour are the decrease of the shear modulus and the increase of the damping ratio with increasing deformation.

The apparent shear wave velocity is linked to the shear modulus and is decreasing as well with increasing deformation. Consequently, the peak frequencies of the site-response curves are shifted to lower frequencies. Indeed, considering a one-dimensional site configuration composed of one layer of sediment lying over a semi-infinite space of rigid bedrock, the resonance frequencies are linked to the shear-wave velocity by the well-known formula f_n = (2n+1).Vs/4H, where f_{n-1} is the $(n)^{th}$ resonance frequency of the layer, H is the depth of the layer, and Vs is the shear-wave velocity.

The amplitude of the site-response curve varies according to two opposing phenomena. The first of these, which generally tends to dominate, is the increase in the damping ratio with deformation, which induces a decrease in the amplitude; the second is an increase in the impedance contrast, which is linked to the decrease in the shear-wave velocity in the sediment layer, which will induce an increase in the site-response amplitude. While the second effect of soil non-linear behaviour on site response amplitude will influence the outcrop site response, as shown in the first chapter, the borehole site response (when the down-hole sensor is located at the sediment/substratum interface

depth) is independent from the impedance contrast between sediment and substratum. Therefore, for BFSR, the non-linear effect expected on amplitude is a decrease.

These expected effects are based on a simplified model of non-linear soil behaviour. These do not map more complex phenomena, such as soil dilatancy, pore-water pressure increase, or even strain hardening during earthquakes (Gélis & Bonilla, 2012).

4.3.1.2.1 Observations of non-linear behaviour during the Tohoku earthquake

We used the data from a limited number of KiK-net sites to investigate the non-linear soil behaviour during the Tohoku earthquake (Bonilla et al., 2011). The Tohoku earthquake occurred on the afternoon of 11 March, 2011, off the Pacific coast of Tohoku, Japan, with Mw 9. This was one of the largest earthquakes in the world that has been widely recorded in the near vicinity of the source (NIED, 2011). Due to the KiK-net network, this event provides an important amount of strong ground-motion data. We selected four KiK-net sites that recorded the main event, IBRH11, IWTH21, IBRH16 and MYGH04, which had Vs_{30} of 242 m/s, 521 m/s, 626 m/s and 850 m/s respectively.

Figure 4.4 shows the locations of the sites and the epicentres of the earthquake series used. Figure 4.5 shows the comparison of the borehole site-response evaluations (BFSRs) computed using the recordings of the main event (Figure 4.5, thick line), with BFSR_{lin} and $\text{BFSR}^{\sigma}_{lin}$ calculated using weak motions from aftershocks (the 68% and 95% confidence limits are illustrated in dark and light grey, respectively). The main event produced PGAs of 821 cm/s^2, 375 cm/s^2, 546 cm/s^2 and 504 cm/s^2 for the IBRH11, IWTH21, IBRH16 and MYGH04 stations, respectively. The PGA of the chosen aftershocks was limited to 20 cm/s^2. Thus, in Figure 4.5, we can see that the site response is significantly different when it is derived from the main event as compared to the aftershocks. For these four stations, there was a systematic decrease in the peak frequencies that was associated with a decrease in their amplitude, except for station IBRH16, for which the ampli-

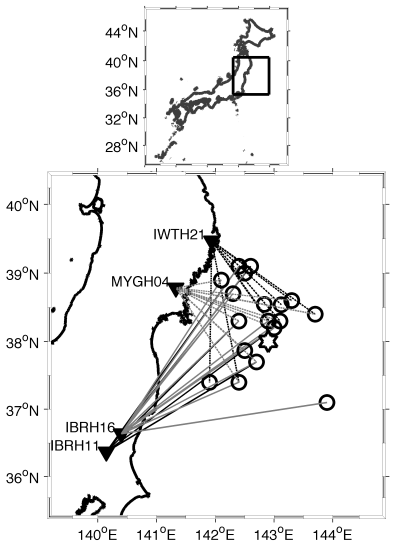


Figure 4.4: *Location of the four KiK-net sites and locations of the epicentres of the earthquakes used to compute the linear borehole site responses (black circles), and the main event of the Tohoku earthquake (white star).*

tude of the first peak was slightly increased. The maximum frequency shift occurred at the MYGH04 station, although the PGA of the main event was not the strongest (Figure 4.5, bottom right). It can be noted that BFSR computed using the recordings of the main event were amplified as compared to $BFSR_{lin}$ at frequencies below the predominant frequency, and were de-amplified above these frequencies, as is very clear at station MYGH04. The frequency at which de-amplification occurs increases as the station Vs_{30} increases. For example, station IBRH16 (Vs_{30} = 626 m/s) de-amplifies at around 7 Hz, whereas station MYGH04 (Vs_{30} = 850 m/s) de-amplifies at around 12 Hz. The observations during the Tohoku earthquake at the four selected KiK-net sites support the theory discussed in the second chapter.

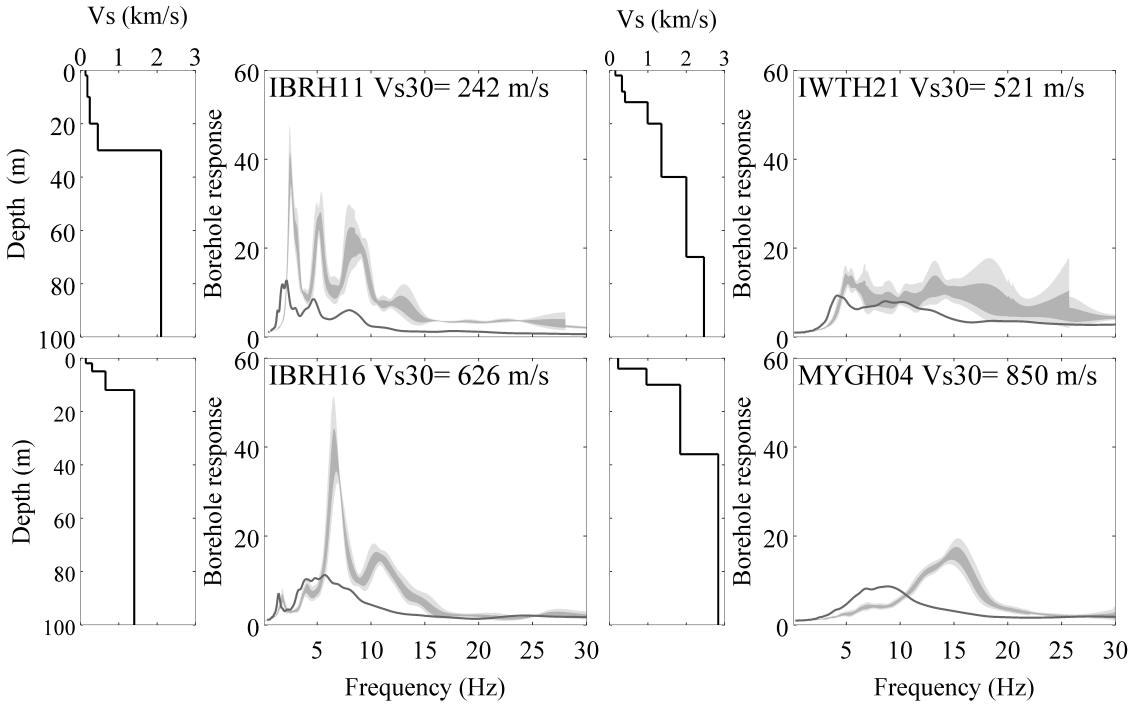


Figure 4.5: *Comparison of the borehole site responses computed from the aftershocks of the Tohoku earthquake (PGA $<$ 20 cm/s^2), the dark grey area represents 68% of the observations around the mean, and the lighter area 95%, with the borehole site response computed with the Tohoku main event represented by the thick dark curve (modified from Bonilla et al. 2011*

4.3.1.2.2 Example of the evolution of the BFSRs at IWTH23 KiK-net site

Figure 4.6 shows the evolution of the BFSRs according to the PGA (at the downhole sensor) at site IWTH23. It is clear here that the site-response curve is strongly modified according to

the PGA of the incoming motion. The following section is devoted to the quantification of these effects.

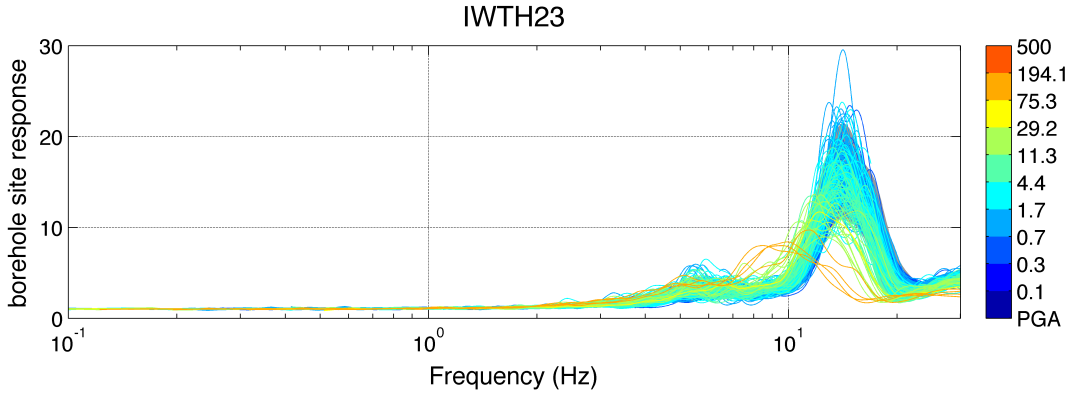


Figure 4.6: *Breakdown of the borehole site responses at IWTH23, according to the input motion PGA (cm/s^2)*

4.3.2 Characterisation of non-linear site responses

To quantify the modifications between linear and non-linear site responses, different evaluations can be performed. In a previous study, Field et al. (1997) computed the ratio between the linear and non-linear amplification functions. Then after the work of Noguchi et al. (Noguchi & Sasatani, 2008), Wen et al. (Wen et al., 2011) proposed the DNL parameter (as the degree of non-linearity of a site response) given in Equation 4.7. This parameter is calculated in the frequency range [0.5-20] Hz.

$$DNL = \sum_{i=N_1}^{N_2} \left| log\left(\frac{BFSR_s(i)}{BFSR_{lin}(i)}\right) \right| .(f_{i+1} - f_i) \tag{4.7}$$

where, BFSR_s is the empirical site response for strong motion, and BFSR_{lin} is the mean empirical site response for weak motion, f is the frequency, N_1 is the first index of the frequency that is above 0.5 Hz, and N_2 is the last index of the frequency that is below 20 Hz.

In line with of former studies, we define: (1) two parameters calculated for each event for each site; and (2) four parameters that characterise the non-linear behaviour of a given site, considering all of the events recorded. These parameters are summarised in Table 4.3.

4.3.3 Non-linear event parameters

The non-linear event parameters quantify the effects of the non-linear soil behaviour on site responses calculated for each event, as compared to the linear site characterisation. The two non-linear parameters are defined as:

Sh_{ev}, the frequency shift between $BFSR_{lin}$ and BFSR . This is calculated by performing a cross-correlation of the $BFSR_{lin}$ and the BFSR computed from the recording of a given event and taking the lag for which the cross-correlation reaches its maximum. The value of this parameter is illustrated in Figure 4.7. In the case of a clear peak, the Sh_{ev} represents the shift of the predominant peak amplitude frequency.

PNL_{ev}, the percentage of non-linearity . This parameter represents the percentage of modification of the site-response curve between the BFSR of each event and the linear site characterisation. This is represented by the area between the BFSR computed from the recording of a given event and the $BFSR_{lin}^{95}$ site response divided by the area below the $BFSR_{lin}$, considering a frequency log scale between 0.3 Hz and 30 Hz. This parameter is illustrated in Figure 4.7 and given in Equations 4.8 and 4.9. In this calculation, we compute the modifications considering both the amplification and de-amplification values between BFSR and $BFSR_{lin}^{95}$. Compared to the DNL proposed by Wen et al. (2011), the PNL_{ev} takes into account the variability of the linear site-response curve and is normalised by the mean linear site-response curve, to give an absolute estimation of the non-linear behaviour of the soil independently of the linear site-response amplitude.

$$A = \sum_{i=N_1}^{N_2} \begin{cases} (BFSR(i) - BFSR_{lin}^{+}(i)).\log_{10}(\frac{f_{i+1}}{f_i}) & \text{if} \quad BFSR(i) \geq BFSR_{lin}^{+}(i) \\ (BFSR_{lin}^{-}(i) - BFSR(i)).\log_{10}(\frac{f_{i+1}}{f_i}) & \text{if} \quad BFSR(i) \leq BFSR_{lin}^{-}(i) \\ 0 & \text{otherwise} \end{cases} \tag{4.8}$$

$$PNL_{ev} = 100.\frac{A}{\sum_{i=N_1}^{N_2} |BFSR_{lin}|.\log_{10}(\frac{f_{i+1}}{f_i})} \tag{4.9}$$

where, $BFSR_{lin}^{-}$ is the lower boundary of $BFSR_{lin}^{95}$, $BFSR_{lin}^{+}$ is the upper boundary of $BFSR_{lin}^{95}$, f is the frequency, N_1 is the first index of the frequency that is above 0.3 Hz, and N_2 is the last index of the frequency that is below 30 Hz.

Figure 4.7 shows the comparisons of the $BFSR_{lin}$ and the $BFSR_{lin}^{95}$ at IWTH23 with the BFSR, as calculated from the recordings of the main Tohoku event (PGA at depth of 137 cm/s^2, grey line;

the mean is shown by the black line, and the 95% confidence limits by the grey shaded area). The PNL_{ev} and Sh_{ev} for this event are also illustrated in Figure 4.7.

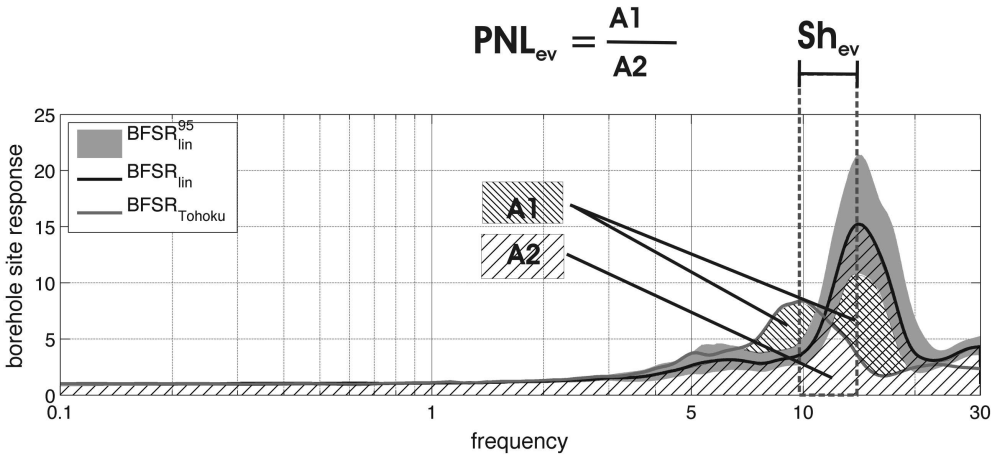


Figure 4.7: *Illustration of the calculation of the PNL_{ev} (percentage of non-linearity) and Sh_{ev} values, with recordings at station IWTH23.*

4.3.4 Definition of the relevant seismic motion intensity parameters for non-linear site-response assessment

For each KiK-net site, we calculated the linear correlation coefficient between each of the intensity parameters (defined in the 2.3) and the PNL_{ev} parameters. We selected KiK-net sites with at least 10 events with a maximum PGA recorded at depths greater than 30 cm/s^2 (226 sites). For each site, we calculated the coefficient of correlation between the PNL_{ev} and each intensity parameter. Finally, we calculated the mean correlation coefficient over the sites for each intensity parameter, and is given in Table 6.1. In the present study, we consider that the best intensity parameter for the observation of non-linear soil behaviour effects on the empirical site response is the one that has the highest mean correlation coefficient for PNL_{ev}. From this calculation, we find that the best intensity parameters are the acceleration root mean square (a_{rms}) and the PGA. Nevertheless, as the PGA is a widely used parameter, especially in seismic hazard assessment, it appeared more convenient to use it in the continuing aspects of this study. We also calculated the mean correlation coefficients between the intensity parameters and the DNL. We find that the best correlation is also with the PGA, but with a lower value (0.32).

Table 4.2: *Mean over the whole KiK-net sites, of the correlation coefficient between the intensity parameters of the recordings at depth and PNL$_{event}$*

Intensity parameters	PGA	PGV	Ia	a_{rms}	CAV	Dtri	Fc
Mean correlation coefficients with PNL$_{ev}$	0.52	0.42	0.48	0.52	0.40	0.04	0.03

4.3.5 Non-linear site parameters

Using the correlations between the non-linear event parameters with respect to PGA, we defined for each site three site parameters that characterise the effects of the non-linear soil behaviour on site responses (PNL$_{site}$, PGA$_{th}$ and Sh$_{site}$, the meanings of these parameters are given above). Furthermore, using numerical simulations, Yu et al. (1993) studied the differences between linear and non-linear site responses. They separated the site responses into three frequency bands, for which they observed different effects on the site responses. In the low frequency band, no effects were seen; in the medium frequency band, de-amplification compared to the linear evaluation was observed; and in the high frequency band, amplification was seen. Delépine et al. (2009) also underlined the frequency dependence of the effects of non-linear behaviour. From our observations of the linear to non-linear site response ratio, we note two main patterns: an amplification of the non-linear site response compared to the linear case, below a given frequency; and a de-amplification above this given frequency. Thus, the fourth non-linear site parameter is the frequency that separates these two behaviours in the site-response curve (fNL). These non-linear parameters are summarised in Table 4.3.

PNL$_{site}$ and PGA$_{th}$ Figure 4.8(a) shows all of the BFSRs per event (grey lines), as calculated at site IWTH23. The corresponding PNL$_{ev}$ values are plotted in Figure 4.8(b), along with the non-linear regression between the PNL$_{ev}$ and the PGA at this station, according to Equation 4.10. From the recordings at IWTH23, we selected three events with low, medium and high PGAs at depth (0.3, 29 and 105 cm/s^2, respectively). The corresponding site responses and the associated PNL$_{ev}$ are plotted with the thicker curves in Figure 4.8(b). We can note that, as expected, the PNL$_{ev}$ increases with the PGA.

$$PNL_{ev} = a \tanh(\log(PGA) + b) + 1) \tag{4.10}$$

Considering the lack of strong-motion data, it can be seen that the curve at large PGAs is

not very well constrained. Nevertheless, this expression agrees with the models that are used to describe the degradation of the shear modulus and the increase in the damping with the deformation, and it gives lower residuals compared to the linear models. The first non-linear site parameter is called PNL_{site}.This corresponds to the value of PNL_{ev} at a threshold value of PGA (here, 50 cm/s^2), according to Equation 4.10. I. A. Beresnev et al. (1995) proposed a review on the observation of non-linear soil behaviour. Based on the studies of Darragh and Shakal (1991) and Chin and Aki (1991) among others, Beresnev et al. (1995) indicated that: "In summary, geotechnical testing of soils and the limited seismological data obtained so far suggest that non-linear soil behaviour may become significant when surface accelerations exceed 0.1 to 0.2 g.". The recording at the downhole sensor contains the incident and down-going wavefields. To select the PGA threshold, we simply divided the PGA at the surface by a factor of two. We are aware that under this simplification, two assumptions are made: the free surface effect is similar at all frequencies, and the down-going wavefield does not affect the PGA.

The second non-linear site parameter is the threshold value of PGA at a PNL_{ev} of 10% (PGA_{th}). PGA_{th} is the acceleration for which we expect significant effects of the non-linear soil behaviour on the site response. In the present study, we estimate that 10% of modification inbetween the linear and non-linear site responses (outside the 95% confidence limits of the linear evaluation) is representative of a significant modification. This choice is a compromise between a sufficiently large value that represents a significant modification, and a low value, to be able to estimate the PGA_{th} at a large number of sites (At some sites, the non-linear regression between the PNL_{ev} and the PGA may not exceed 10%).

The last two other non-linear site-specific parameters are related to modifications of the site-response peak frequencies.

Sh_{site} is the value of Sh_{ev} when the PGA is 50 cm/s^2. This parameter is found according to linear regression between Sh_{ev} and PGA.

fNL We calculated for each strong motion (a PGA recorded at the downhole sensor >50 cm/s^2), the ratio between $BFSR_{lin}$ and BFSR for the strong events. We calculated a mean and standard deviation in log units of this ratio. We take the frequency from which the lower boundary (95% confidence limit) of the ratio becomes higher than unity; this represents the frequency from which we observe de-amplification from linear to non-linear site-response evaluations (Figure 4.9). It is essential to remember that the shift of the predominant peak frequency in the site response during strong motion implies that below this frequency (fNL), the site response computed with strong events is likely to be amplified compared to linear evaluation.

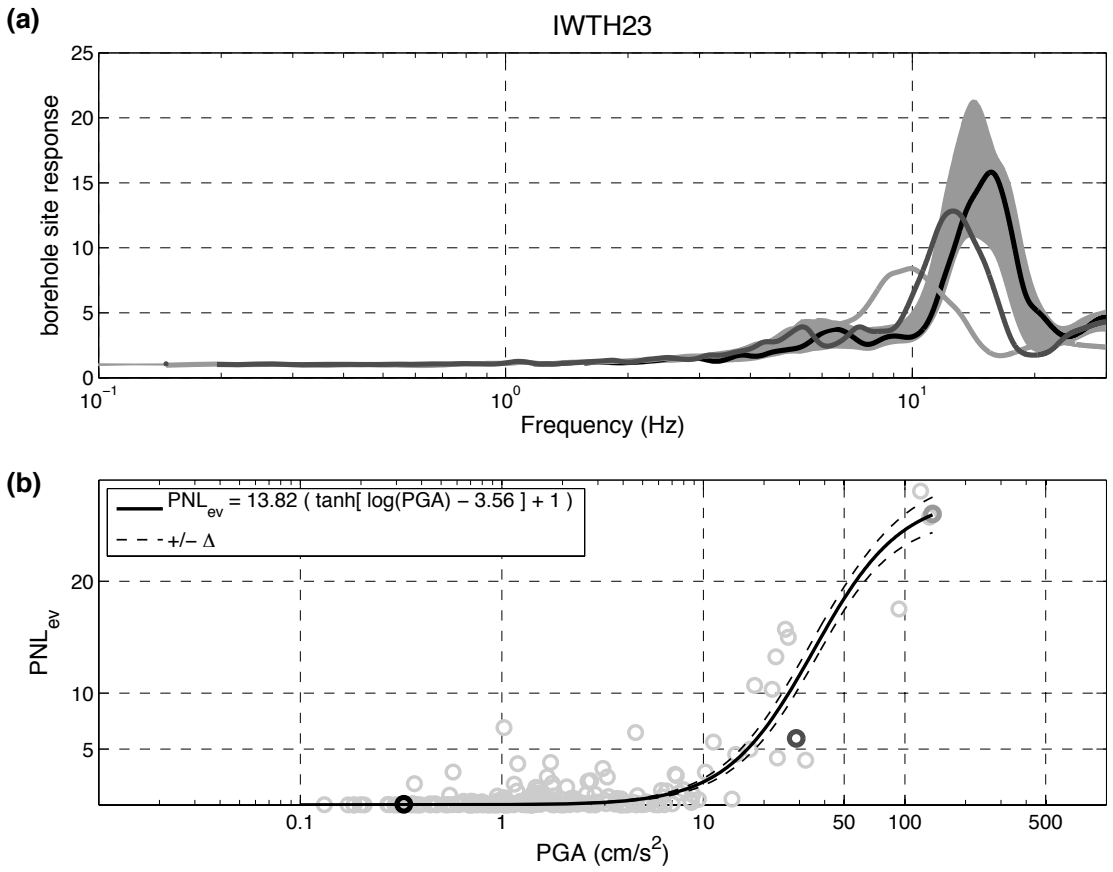


Figure 4.8: *(a) The borehole site response at site IWTH23, with three curves highlighted: one that comes from the recordings with low PGA at depth (black curve, PGA $<$ 10 cm/s^2); one with medium PGA (dark grey curve, PGA at depth = 40 cm/s^2); and one with high PGA (light grey, PGA $>$ 100 cm/s^2). (b) The corresponding PNL$_{ev}$ that were calculated with the linear regression between PGA and PNL$_{ev}$. PNL$_{site}$ I the value of the linear regression for PGA of 50 cm/s^2. PGA$_{th}$ is the PGA that is associated with a PNL of 10%, according to the linear regression.*

4.4 STATISTICAL ANALYSES OF THE EVENT/ SITE PARAMETERS

In the previous section, we saw that PGA is one of the best intensity parameter candidates to give the possibility that the corresponding recording triggers the non-linear soil behaviour. Indeed, the correlation between PGA and PNL$_{ev}$ performed at each site is one of the highest. In this section, we investigate this relationship in greater detail, mixing all of the sites. Secondly, we explore the relationships between the non-linear (PNL$_{site}$, PGA$_{th}$, Sh$_{site}$ and fNL) and site (Vs$_{30}$, B_{30}, f_0, f_{pred} and A$_{pred}$) parameters, to determine the interactions among them.

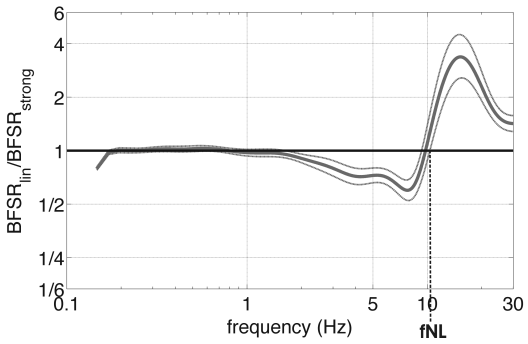


Figure 4.9: *Illustration of the calculation of the fNL, with the ratio of the linear to non-linear site responses at station IWTH23.*

Table 4.3: *List of the parameters used in this chapter*

Input motion, Soil and site Response Parameters	per event	PGA	cm/s^2	peak ground acceleration
	Per site	Vs_{30}	m/s	Mean shear-wave velocity
		B_{30}		Gradient of shear-wave velocity profile
		f_0	Hz	Fundamental resonance frequency
		f_{pred}	Hz	Predominant resonance frequency
		A_{pred}		Amplitude of $BFSR_{lin}$ at f_{pred}
Non-linear, parameters	for each event	PNL_{ev}	%	percentage of soil non-linearity for each event
		Sh_{ev}	Hz	Shift for each event
	Per site	PNL_{site}	%	Percentage of soil non-linearity for each site
		Sh_{site}	Hz	Shift
		PGA_{th}	Hz	PGA threshold of non-linearity
		fNL	Hz	Frequency of non-linearity

4.4.1 Effects of PGA on non-linear parameters for each event

We use the entire set of KiK-net data to answer the following question: Can the incident motion parameter (namely, the PGA) be used to indicate the possibility that the corresponding recording triggers the non-linear soil behaviour? To do so, we investigated the relationship between PNL_{ev} and PGA.

Figure 4.10(a) shows the PNL_{ev} and the PGA recorded at the downhole sensor for each recording at all of the KiK-net sites (grey circles). For the PGA bin specified in Table 4.4, we calculated the mean (Figure 4.10(a), filled squares) and standard deviation of the PNL_{ev}, in log units. In the bottom panel of Figure 4.10(b), the probability density function of the PNL for each PGA bin has been plotted, assuming a log-normal distribution. We checked that the PNL_{ev} values were log normally distributed using a Lilliefors test (Lilliefors, 1967). The PNL_{ev} increases with increasing PGA. As the PGA increases, the probability density function peakedness (kurtosis parameter) also increases, which indicates a decrease in the variance of the PNL_{ev}. For each PGA bin, we also calculated the probability that the PNL_{ev} reaches 10% (P($PNL_{ev} > 10\%$ / PGA)), as given in Table 4.4. What is remarkable here is that independent of the site, the probability of a significant modification (up to 10% outside the linear site-response variability) between the linear and the non-linear site responses is >20% for PGA recorded at downhole sensor values >30 cm/s^2.

Table 4.4: *Values of the mean and standard deviations of the PNL_{ev} per PGA bin at the downhole sensor. The last line indicates the probability that the PNL_{ev} is >10% per PGA bin.*

PGA bin (cm/s^2)	0.1-0.3	0.3-0.7	0.7-1.7	1.7-4.4	4.4-11	11-29	29-75	75-194	194-500
PNL_{ev} mean	0.009	0.013	0.025	0.050	0.0136	0.0970	3.616	8.204	12.914
PNL_{ev} σ	2.6	2.6	2.7	2.7	2.6	1.9	1.3	1.1	0.9
P(PNL>10% /PGA) (%)	0.3	0.6	1.3	2.3	4.6	11.3	21.6	41.9	60

The frequency shift (Sh_{ev}) between $BFSR_{lin}$ and BFSR according to the PGA recorded at the downhole sensor of the corresponding event is shown in Figure 4.11(a). Only values of Sh_{ev} where the PNL was >zero are shown (where the differences between the mean linear site response and the event site response are greater than the variability of the linear site response). The mean delay increases with increasing PGA. The delay is centered around zero, with a small variance for PGA below 10 cm/s^2 in depth. When the PGA increases, the mean delay increases, along with a decrease in the kurtosis parameter of the probability density function (Figure 4.11(b)).

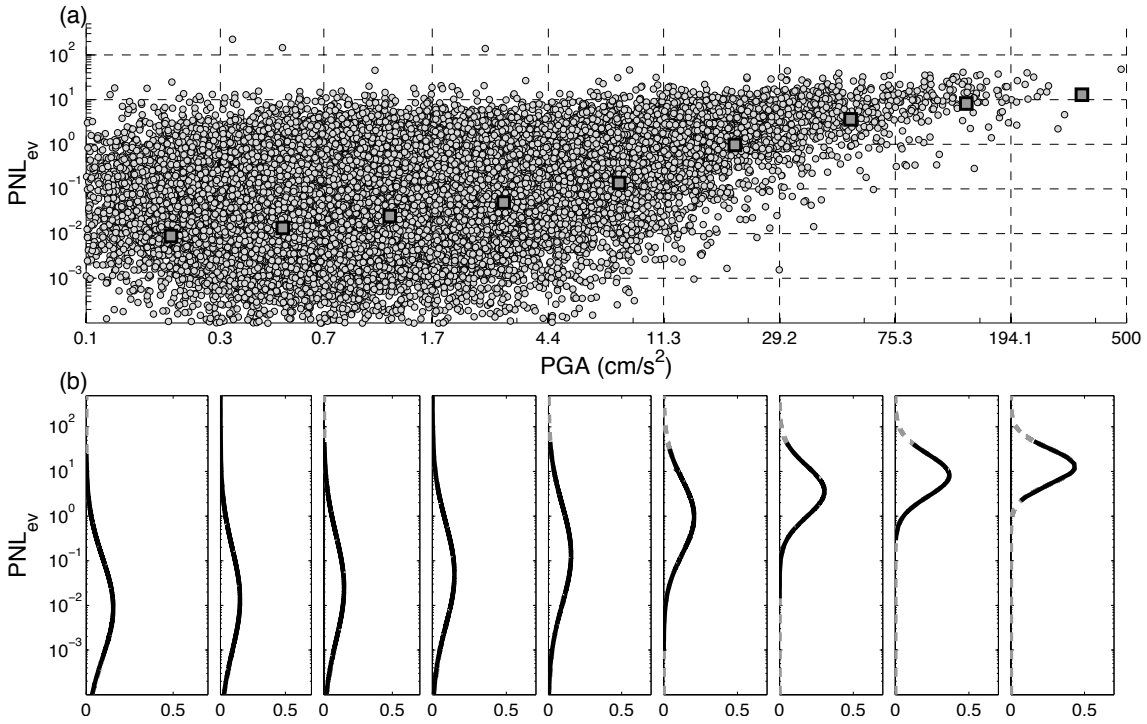


Figure 4.10: *(a) The grey circles represent the PNL_{ev} values as a function of the PGA at depth, the squares represent the mean PNL_{ev} (in log units) for the PGA bin that are displayed on the abscissa (b) Probability density functions assuming a log-normal distribution of the PNL_{ev} for different PGA at depth bins.*

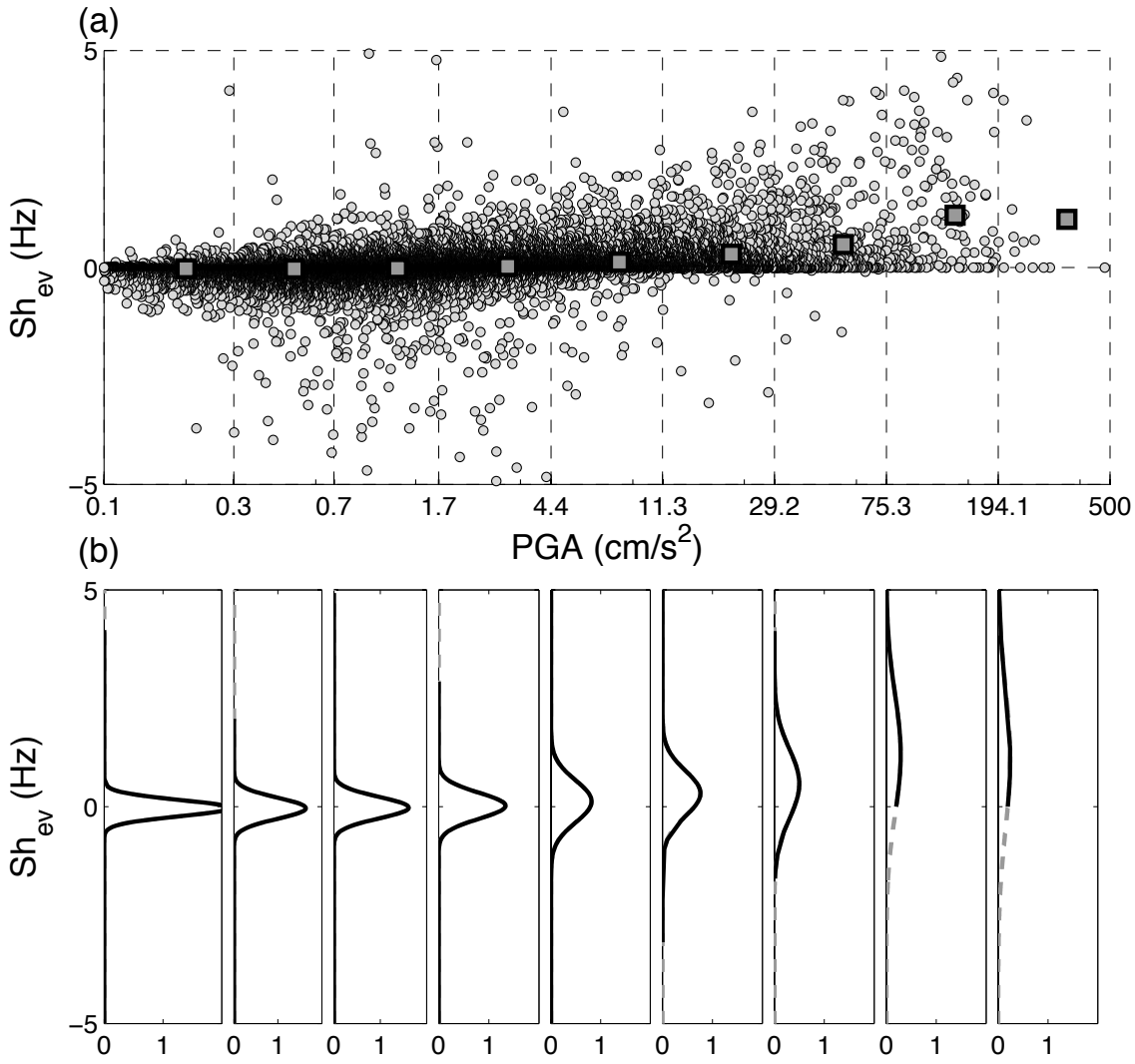


Figure 4.11: *(a) The grey circles represent the Sh_{ev} values as a function of the PGA at depth, the squares represent the mean Sh_{ev} for the PGA bin display on the abscissa. (b) Probability density function assuming a normal distribution of the Sh_{ev} for different PGA at depth bins.*

4.4.2 Site parameters

The second part of these statistical analyses investigates the relationships between the soil and the site-response parameters (Vs_{30}, f_0, f_{pred}, A_{pred} and B_{30}) with the non-linear site parameters (PNL_{site}, Sh_{site}, PGA_{th}, fNL). To take into account the relatively large number of variables, we

used a multivariate statistical analysis to observe the main trends. This analysis was followed by quantitative evaluation of the correlated parameters using regression analysis.

4.4.2.1 Canonical correlation

Canonical correlation analysis (CCA) (e.g., Raykov & Marcoulides, 2008) is a general framework of multivariate statistical analysis methods. Regression analysis, multivariate analysis of variance, and discriminant analysis are particular cases of this method. Consider two sets of variables called A and B. A consists of p members (p>1) and B consists of q members (q >1). The variables in A (or B) cannot be considered as dependent variables. The aim of the CCA is to analyse the interrelationships between these two sets of variables, to 1) examine the independence of the two sets of data; and 2) to reduce the number of data. The CCA procedure is composed of the following steps: The first step consists of finding the linear combination of the variables of set A (called Xcan1) and the linear combination of the variables of set B (called Ycan1), such that their correlation (p11) is the highest possible across all of the combinations. Xcan1 and Ycan1 are called the first pair of canonical variates, and their correlation is the first canonical correlation. The next step is to find the second pair of canonical variates that are not correlated with the first pair and that have the highest canonical correlations. This process is repeated until we find the number of canonical variates equal to the minimum of the variable numbers in set A or B (d).

The next step is to determine the number of canonical variates that are relevant. To do this, we examine the significance of the canonical correlations using a statistical test with a null hypothesis, H_0. For the canonical variate n, n $\in$ [1,d], H_0 is such that all of the canonical correlations that correspond to the canonical variates from n to d are zero. The statistical test used here is based on the calculation of the Wilks' lambda statistics. This coefficient is then transformed to a statistic that has approximately an F-distribution, for which the p values are more easily extracted. In the present study, if the p value is lower than 5%, then the null hypothesis is rejected and we can conclude that at least the canonical variate n is significant.

The final step is to interpret the meaning of the canonical variates. In the present study, to interpret the canonical variates we can use the correlation between each variable of set A (or B) with the canonical variates Xcan1 (or Ycan1). The higher the correlation is with Xcan1, the better the concerned variables in set A explain the variables in set B that have the greatest correlation with Ycan1.

Consider set A that consists of p = 5 soil and site-response parameters (Vs_{30}, f_0, f_{pred}, A_{pred} and B_{30}), and set B that represents the effect of non-linear soil behaviour at one site (PNL_{site}, PGA_{th}, Sh_{site}, fNL). The substantive question in this analysis is to understand which soil parameters are associated with which non-linear soil behaviour effects on site responses.

The formal question that can be asked with CCA is as follows: Are there high correlations between linear combinations of soil parameters with linear combinations of non-linear behaviour

effects on site responses? If so, what is the minimum number of such pairs that can be found to nearly completely represent the cross correlations of the soil parameters with non-linear effects with the site responses? The interpretation of canonical variates gives an idea how the soil parameters are related to the non-linear effects on site responses.

Among the KiK-net sites, we selected the 54 sites that had recorded at least two events with PGA at the downhole sensor >50 cm/s^2. The site locations and associated Vs_{30} are shown in Figure 4.12. The selected sites are located in 11 different districts in Japan, mainly on the east Pacific coast. Among these, we were able to calculate all of the non-linear parameters for 34 sites. For the remaining 20 sites, the low non-linear behaviour allowed us to calculate the PGA_{th} (the PNL_{ev} at these sites did not exceed 10%; none of the strong events recorded induced a significant modification of the site response). The non-linear parameters for these KiK-net sites are given in Tables D.1 and D.2 in the appendix D.

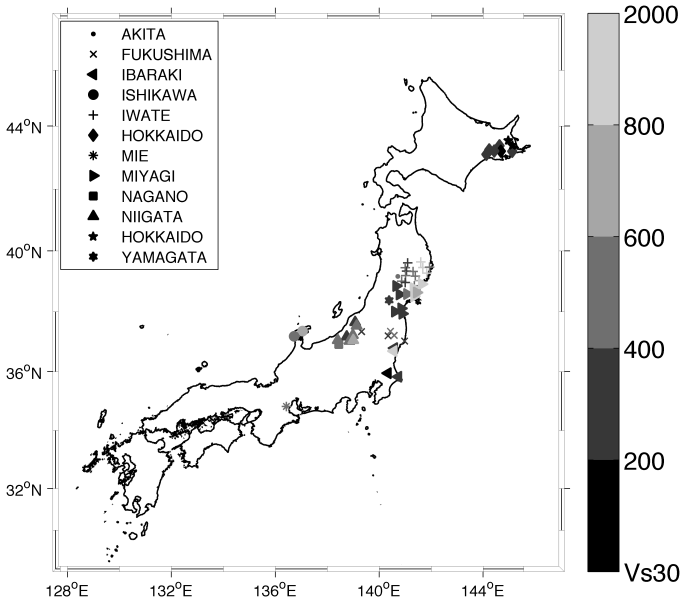


Figure 4.12: *Locations of the KiK-net sites that have recorded at least two earthquakes with PGA at depths >50 cm/s^2.*

The CCA requires the variables to be log-normal distributed to proceed. The soil and site response parameters, along with the non-linear parameters of the selected sites, are not approximately normal (See figure 2.5 page 75 for the site response parameters. The distributions of the non-linear site parameters are not shown). To improve the normality assumption, we applied a Box-Cox transformation to all of the variables (Box & Cox, 1964).

We performed the canonical correlations on sets A and B that were previously defined, to find four canonical variates. Among these four, only the first two canonical variates were significant according to their p values. Figure 4.13 shows the results of the first canonical correlation. The correlation between the canonical variates Ycan1 and Xcan1 with the non-linear parameters and the

soil and site response parameters indicates that the fNL and the Sh_{site} are well explained mostly by the Vs_{30}, the predominant frequency, and the fundamental resonance frequency of the site. Figure 4.14 shows the second canonical variates Ycan2 and Xcan2, with a coefficient of correlation of 76%. This represents a second possibility to correlate the variables in set A and set B. In Figure 4.14, we can see that A_{pred} and B_{30} are correlated with PNL_{site} and SH_{site}, whereas they are negatively correlated with PGA_{th}.

According to the canonical correlations performed above, we see the following trends: 1) the fNL and Sh_{site} are well explained by Vs_{30}, f_0 and f_{pred}; and 2) PNL_{site} and PGA_{th} are well explained by A_{pred} and B_{30}.

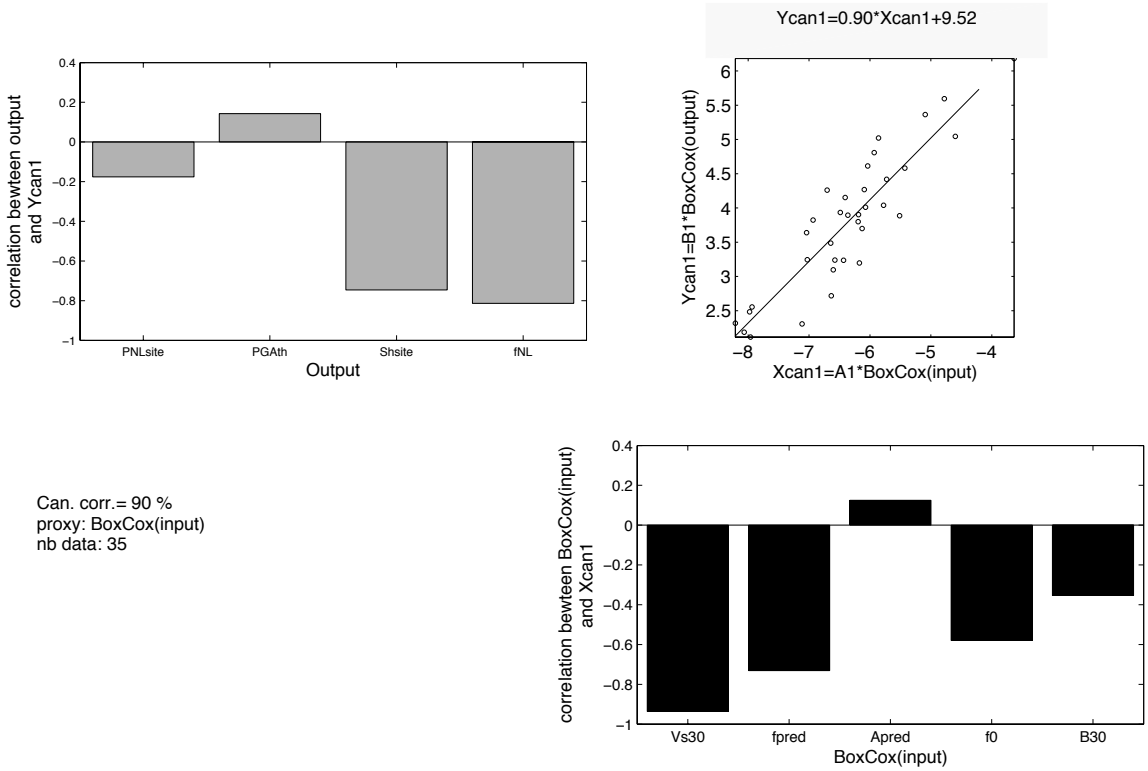


Figure 4.13: *First canonical variates of the canonical correlation of the site non-linear parameters with the soil and site-response parameters.*

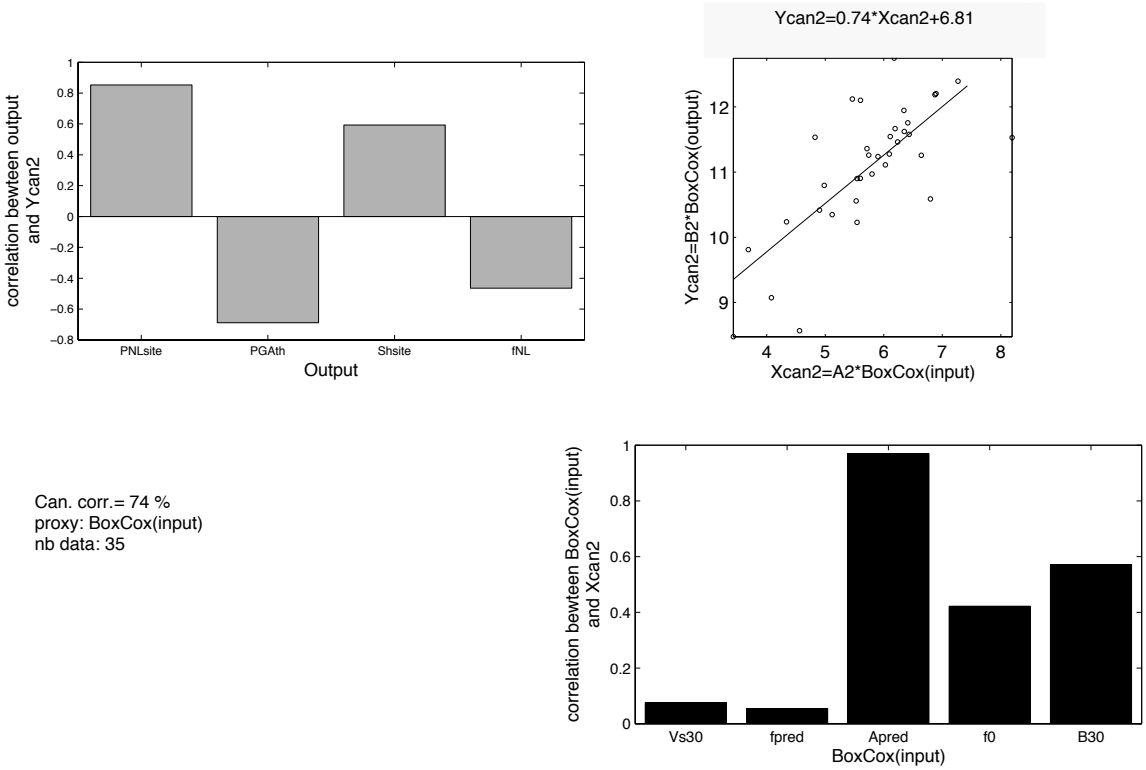


Figure 4.14: *Second canonical variates of the canonical correlation of the site non-linear parameters with the soil and site-response parameters*

4.4.2.2 Regression analyses

The following section is devoted to the quantification of the trends highlighted by the canonical correlations, using simple linear regression analysis.

Although fNL clearly increases with Vs_{30} (Figure 4.15(a)), the linear correlation between fNL and Vs_{30} is not strong, as the coefficient of correlation is 0.35. In Figure 4.15(b), we observe that fNL is greater than or equal to the fundamental resonance frequency of the site (Figure 4.15(b), light grey circles). FNL is calculated using the ratio of the BFSR, whereas the fundamental resonance frequency is deduced from the earthquake H/V spectral ratio at the surface. Thus, to check that the previous observation was not because the BFSR miss the fundamental frequency, we also calculated the fNL on the earthquake H/V spectral ratio at the surface (Figure 4.15(b), light grey triangle). Similarly, the fNL on the earthquake H/V spectral ratio curve is âL'ěf_0. For a large number of sites, the frequencies de-amplified by the non-linear effects are above the fundamental resonance frequency. The predominant frequency (Figure 4.15(b), dark grey circles) of the BFSR is better correlated to fNL, and is usually higher (with a coefficient of correlation for the robust fit of 0.82).

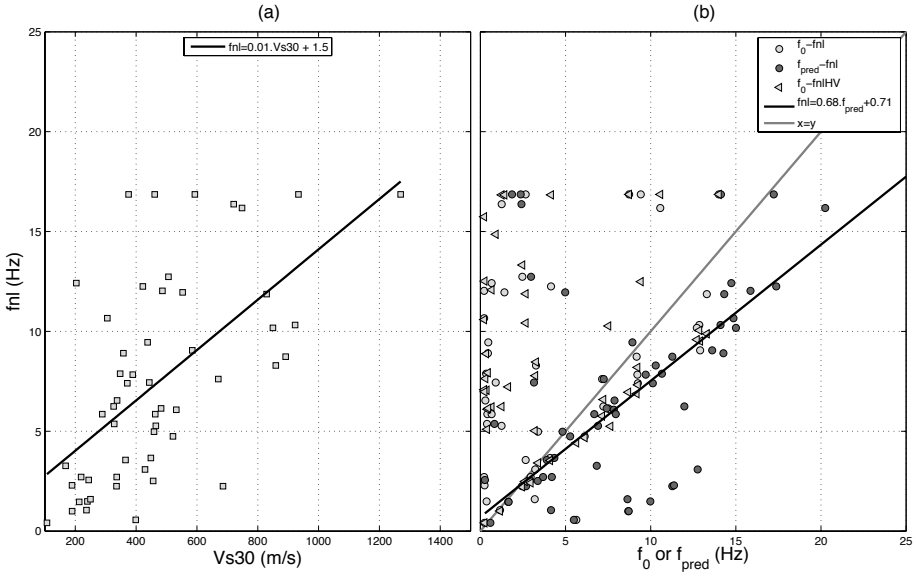


Figure 4.15: *(a) Linear regression of fNL with Vs_{30}. (b) Link between fNL (computed from BFSR of the earthquake H/V spectral ratio) with f_0 and f_{pred}.*

Figure 4.16 shows the correlation (robust fit) of PNL_{site} and PGA_{th} with the maximum amplitude of the BFSR_{lin} (A_{pred}) and the gradient of the Vs profile (B_{30}). In Figure 4.16 (top), the grey circles indicate the sites for which there is no PGA_{th}. PGA_{th} is correlated to A_{pred}, although the correlation coefficient is a low 0.51 (excluding the site IBRH07, which represents an outlier, with $\text{A}_{pred} > 100$). The correlation with the B_{30} is similar (0.37). The sites for which there is no PGA_{th} are associated to sites for which A_{pred} is very low. The PNL_{site} is well correlated to A_{pred}, with a correlation coefficient of 0.74. The correlation between B_{30} and PNL_{site} is relatively low (0.41).

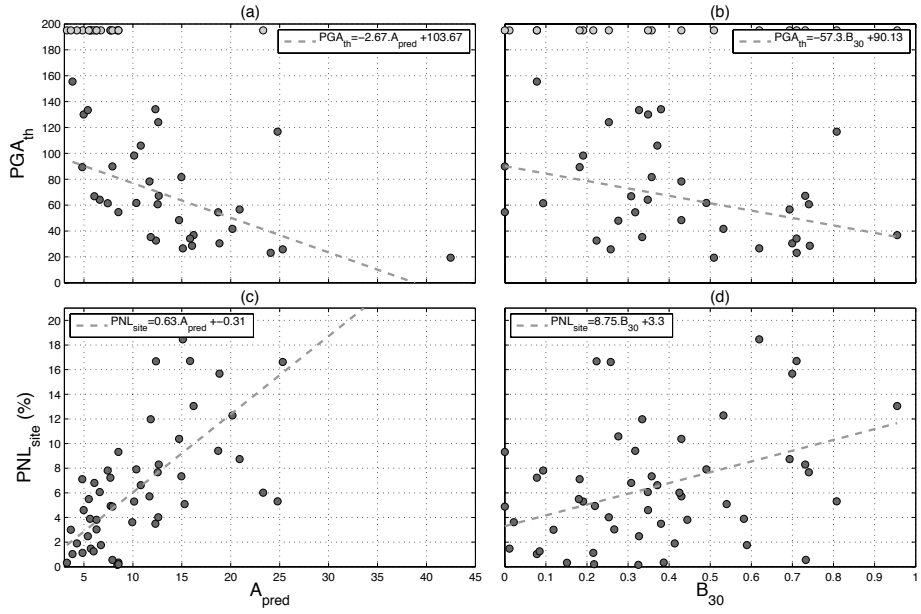


Figure 4.16: *(a) Linear regression of PGA_{th} with A_{pred}. (b) Linear regression of PGA_{th} with B_{30}. (c) Linear regression of PNL_{site} with A_{pred}. (d) Linear regression of PNL_{site} with B_{30}. Top, (a) and (b): the light grey circles represent the sites for which the PGA_{th} could not be computed (the non-linear regression between PNL_{ev} and PGA did not reach 10%)*

In addition to the calculations of the probability that PNL_{ev} is >10% knowing the PGA at depth regardless of the site, we also investigated the effects of knowing one parameter of the site, namely the maximum amplification of the linear borehole site response (A_{pred}). We calculated the PNL_{site} for six other values of PGA (20, 30, 40, 60, 70 and 80 cm/s^2).

Figure 4.17 illustrates the linear regression between PNL_{site} and A_{pred} according to the PGA threshold used. The dark grey curves in Figure 4.17 correspond to the linear regression between the PNL site and the A_{pred} for low PGA thresholds, whereas the light grey curves in Figure 4.17 correspond to high values of the PGA threshold. We find that whichever PGA threshold is used, the linear regression has the same trend. As expected, for a given site with a specific A_{pred}, the PNL_{site}

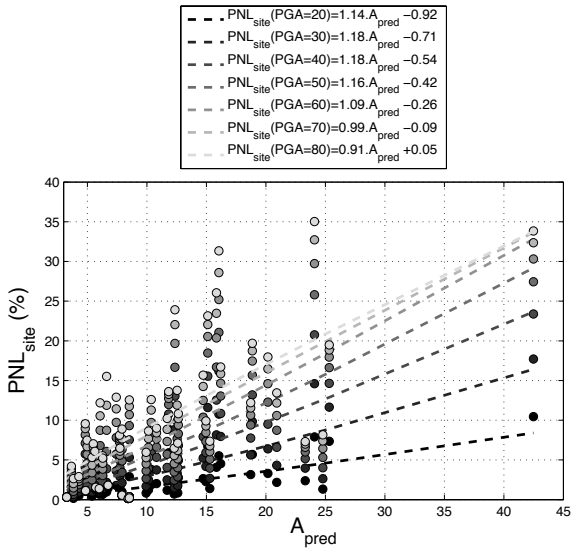


Figure 4.17: *Linear regression of PNL_{site} with A_{pred}, for PNL_{site} calculated using seven different PGA thresholds, from 20 cm/s^2 to 80 cm/s^2.*

is lower when using a small PGA threshold value, and it is higher otherwise. For instance, for a site with A_{pred} close to 15, we would expect 3% modification between the linear and non-linear site responses at a PGA of 20 cm/s^2, and for a PGA of 80 cm/s^2, we would expect 12% modification. For each PGA threshold and for two different ranges of A_{pred} (above or below 15), we calculated the aforementioned probabilities, as given in Table 4.5. For a PGA threshold between 60 cm/s^2 and 80 cm/s^2, the probability that the PNL_{site} is >10% is around 20% for a site with $A_{pred} <$ 15, and around 70% otherwise. For sites with large site effects, the probability that the non-linear soil behaviour will affect the site response is more than three times greater than for sites with low amplification.

Figure 4.18 shows the correlation between Sh_{site} and f_{pred}. We separated the data for which the associated amplitude A_{pred} was above or below 10. We find a high correlation (0.66) between the two parameters when A_{pred} is >10.

Table 4.5: *Probability that the PNL_{site} is higher than 10% depending on the PGA at the downhole station and the A_{pred} of the borehole linear site response.*

PGA cm/s^2	A_{pred}-bin	P($PNL_{site} > 10\%$) %
≥ 20	< 15	0
	≥ 15	6.5
≥ 30	< 15	0
	≥ 15	25.4
≥ 40	< 15	8.9
	≥ 15	45.6
≥ 50	< 15	15
	≥ 15	51.9
≥ 60	<15	18.9
	≥ 15	71.6
≥ 70	< 15	23.5
	≥ 15	72.5
≥ 80	< 15	26.3
	≥ 15	73.3

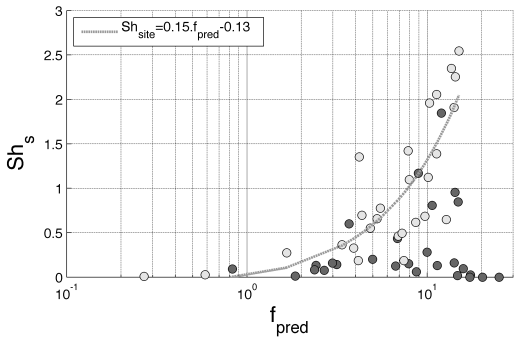


Figure 4.18: *Linear regression between Sh_{site} and f_{pred}. The light grey circles represent the sites for which A_{pred} is >10, the dark grey circles represent the sites for which A_{pred} is âL'd'10. The grey dotted line represents the linear regression of Sh_{site} with f_{pred} for sites with A_{pred} >10*

4.5 OPTIMAL PARAMETERS TO ASSESS SOIL NON-LINEARITY

Recently, Assimaki and Li (2012) investigated the relationships between soil and site response parameters by comparing non-linear, equivalent-linear, and linear numerical evaluations of site responses with linear empirical evaluations. They found that the intensity of non-linear effects at a given site during a specific ground motion is a function of Vs_{30} and the amplitude at the fundamental resonance (site parameters) and the characteristics of the incident motion parameters. In the present study, using empirical data, we first analysed the effects of incident motion intensity on non-linear site responses regardless of the site. Secondly, we examine the site-specific non-linear behaviour considering the soil characteristics and the linear site-response parameters. The scope of the second part of this study is similar to the study performed by Assimaki and Li (2012). However, given the wealth of data in Japan, we estimated the effects of non-linear soil behaviour empirically, instead of doing numerical simulations.

The statistical analysis between PGA and PNL_{ev} shows that regardless of the site, the probability that there is a significant departure from a linear site response is >20% for PGA values recorded at the downhole station greater than between 30 cm/s^2 to 75 cm/s^2 (see Table 4.4). For example, the French national seismic zonation indicates that for southern France, the surface PGA that needs to be taken into account for risk mitigation is 160 cm/s^2 at rock sites, for a return period of 475 years (articles R563-1 and R563-8 of the French Environmental Code). Taking this value to depth by simply dividing by a factor of 2 (effects of the free surface), and neglecting the down-going wavefield, we obtain an incident PGA of 80 cm/s^2. For such an input, the probability that the PNL_{ev} is >10% increases to 40% (see Table 4.4). This means that even for low seismicity countries, this probability is quite high, and it represents a quantitative indicator that non-linear soil behaviour should be taken into account when site responses are evaluated for medium to strong motion.

We used multivariate statistical analyses to define the relationships between the non-linear parameters for each site (PNL_{site}, Sh_{site}, PGA_{th}, fNL) and the soil and site response parameters available for all of the KiK-net sites (Vs_{30}, f_0, f_{pred}, A_{pred}, B_{30}). We find that the fNL and the SH_{site} are well explained by Vs_{30}, f_0 and f_{pred}, with a positive correlation. At the same time, this frequency lies inbetween f_0 and f_{pred}. These observations suggest that the soil non-linearity mainly shifts the peak resonance frequencies above the fundamental resonance frequency. Thus, the deepest velocity contrasts are less affected by the non-linear behaviour than the shallow ones, which indicates that the non-linear soil behaviour occurs mostly in the superficial layers. Consequently, characterisation of the non-linear behaviour of a soil column can be achieved through investigations of non-linear soil parameters in the layers close to the surface only.

We also find that the PGA threshold from which we observe significant non-linear soil behaviour and the PNL_{site} are well correlated to the amplitude of the predominant BFSR peak and

to the shear-wave velocity gradient (B_{30}) (negatively and positively, respectively). The higher the amplitude of the predominant peak, the greater the expected effects of non-linear soil behaviour on site responses. This can be explained because the deformation in the layer associated with the predominant peak is linked not only to the incident motion intensity, but also to the amplification due to the impedance contrast. The regressions that were found in the present study are based on the calculations of borehole site responses, and hence the correlations defined here with A_{pred} cannot be used directly with linear outcrop site responses to find a PGA threshold, although same trends should be observed. Similarly, PNL_{site} increases with increasing B_{30}. A high value of B_{30} indicates a high velocity contrast in the first 30 m depth of soil. This observation is in agreement with what has already been emphasised about the fNL lying inbetween f_0 and f_{pred}, which suggests again that the non-linear soil behaviour occurs mostly in the subsurface layers.

In addition, instead of using borehole site responses, we also calculated the non-linear site parameters from earthquake H/V spectral ratio surface records.

- We calculated for each earthquake the corresponding earthquake H/V spectral ratio, using the surface recordings;
- We calculated the mean and standard deviation for recordings that have a PGA at depth <10 cm/s^2;
- We compared this linear characterisation with all of the earthquake H/V spectral ratios, to get the PNL_{ev}^{rf} and Sh_{ev}^{rf}. Using the correlation with PGA, we found PNL_{site}^{rf} , PGA_{th}^{rf} and Sh_{site}^{rf};
- We also compared the linear characterisation with the earthquake H/V spectral ratios computed using recordings with PGA >50 cm/s^2. Then, we calculated the ratios of the linear to non-linear earthquake H/V spectral ratios to find fNL_{rf}.

Figure 4.19 shows the comparisons between the non-linear site parameters calculated using BFSR and using earthquake H/V spectral ratios at the surface. Figure 4.19(a) shows that the PNL_{site} calculated from the receiver-function curves is overestimated at low values and underestimated at high values of PNL_{site}, as compared to the PNL_{site} calculated from BFSR. The underestimation can be explained because at low frequencies, even for strong motion the variability of the surface earthquake H/V spectral ratio is very high and the non-linear soil behaviour effects on site responses are indistinguishable from the receiver-function intrinsic variability. Furthermore, it has been shown that the receiver-function curves cannot give correct estimations of the amplitude, because the vertical component can have its own response curve, due to significant S-to-P converted waves (Bonilla et al., 2002). The vertical component response can be non-linear as well, which leads to different modifications in the earthquake H/V spectral ratio amplitudes, with increasing input motion PGA compared to the BFSR ones. As shown in Figure 4.19(b), similar trends are seen with the PGA_{th}. In Figure 4.19(b), the PGA threshold is calculated for PNL_{ev} as 5%, instead of 10%, because the

PGA_{th} at 10% calculated from the earthquake H/V spectral ratios does not exist for a large number of sites. The frequencies (fNL) and shift of frequency (Sh_{site}) are very similar when calculated from either the earthquake H/V spectral ratio curves or the BFSR, as shown in Figures 4.19(c) and 4.19(d). These close estimates indicate that the results from BFSR are not associated with pseudo-resonances. This also indicates that fNL and Sh_{site} can be deduced from surface recordings only.

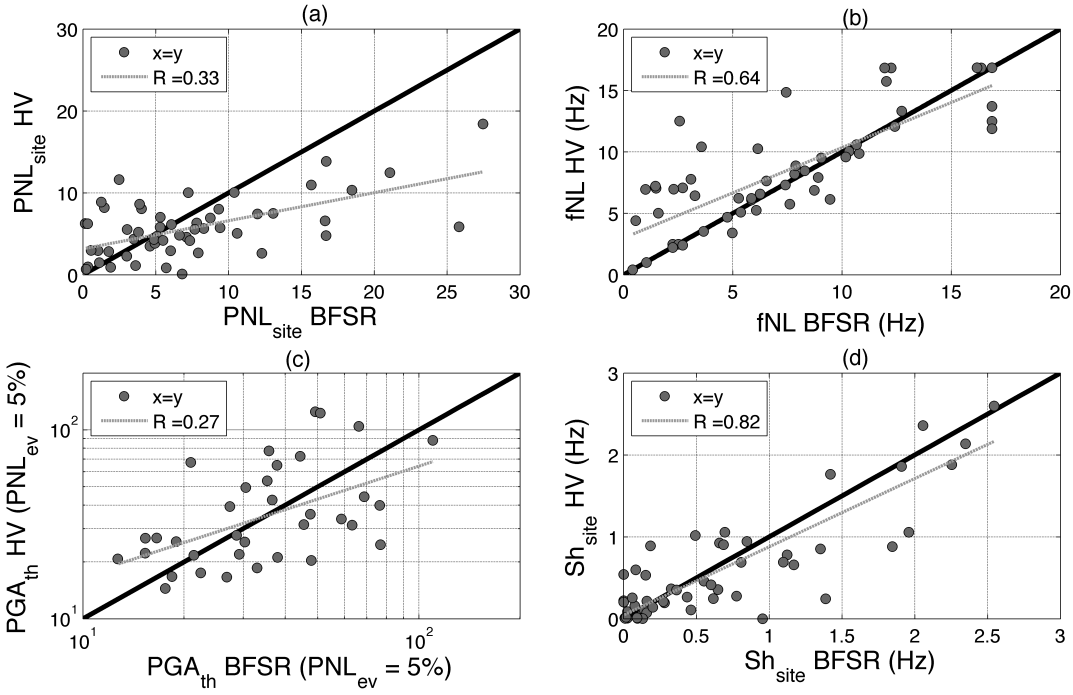


Figure 4.19: *Comparison of the non-linear parameters calculated from BFSR and the earthquake H/V spectral ratio at the surface. (a) Comparison of the PNL_{site}. (b) Comparison of the fNL. (c) Comparison of the PGA_{th}. (d) Comparison of the Sh_{site}. In each panel, the dark line represents the x = y line, and the grey dotted line is the linear regression between the parameters.*

4.6 CONCLUSIONS

The aim of this chapter is to understand the relationships between the parameters that characterise the effects of non-linear soil behaviour on site responses, and the parameters that characterise either the site or the incident ground motion. We first defined the parameters that describe the effects of the non-linear soil behaviour on the site response for each event: the percentage of modification of the site response curve compared to the linear evaluation (PNL_{ev}), and the shift of the curve (Sh_{ev}). We find that the PGA (at the downhole sensor) is the intensity parameter that is the most correlated to the non-linear parameters for each event.

In addition, for the 54 sites of the KiK-net database that recorded at least two strong events with a PGA at depth >50 cm/s^2, we defined the parameters that characterise the effects of non-linear soil behaviour on the site response for each site. A PGA threshold (PGA_{th}), a percentage of modification between the linear and the non-linear site responses (PNL_{site}) for a PGA of 50 cm/s^2, a shift of the peak frequency for a PGA of 50 cm/s^2 (Sh_{site}), and a frequency for which we observe de-amplification between the non-linear site response and the linear site response (fNL). From observations of the linear to non-linear site response ratio, we note that the shift of the predominant peak frequency in the site response during strong motion implies that below fNL, the site response computed with a strong event is likely to be amplified compared to linear evaluations.

For GMPE or strong ground-motion simulation studies, this study indicates the relevant soil and site response parameters for integration of non-linear soil behaviour. Also, the results show that regardless of the site, depending on a PGA threshold (even for moderate input motion: PGA 75 cm/s^2 at the downhole sensor) there is a large probability (>40%) that the site response behaves non-linearly, as indicated in Table 4.4. These results suggest that when working on ground-motion estimation, non-linear soil behaviour should be taken into account when the incident wavefield has a PGA from 75 cm/s^2.

We used multivariate statistical analyses to define the relationships between the non-linear parameters and the soil and site-response parameters available for all of the KiK-net sites. We find that non-linear behaviour occurs mostly in the superficial layers. This conclusion is supported by fNL lying inbetween f_0 and f_{pred}. This indicates that the non-linear soil behaviour induces a shift of the peak frequencies mostly above f_0, showing that the shallowest velocity contrasts are more affected by the non-linear soil behaviour. Also, the shear-wave velocity gradient (B_{30}) correlates well with the PGA threshold. This suggests that sites with high shear-wave velocity contrast close to the surface can trigger non-linear behaviour at low input motion PGA values. We also find that sites with high amplification (A_{pred}) are likely to be more affected by non-linear soil behaviour (considering that PNL_{site} is positively correlated to A_{pred}, and PGA_{th} is negatively) The results obtained here are not limited to the Tohoku earthquake, which has a very particular signature, as the whole dataset from 1996 was used. At the same time, the 54 sites selected for the multivariate

analysis are spread across different districts, and therefore the results are not specific to a given area. As we find that the earthquake H/V spectral ratio can give satisfactory results for the evaluation of the frequency fNL and the shift frequency (Sh_{site}), this makes these results extendable to other databases without downhole sensors.

Part III

Inversion of the borehole transfer function

ABSTRACT

In the previous chapter, we indicated that the non-linear soil behaviour occurs mostly in the subsurface soil layers. This conclusion can have a large influence when dealing with the characterisation of non-linear properties of a soil. However, the main challenge is to find the depths where most of the soil non-linear behaviour takes place. In this chapter, we invert the empirical borehole transfer function. The main objectives of this inversion are to obtain the elastic soil properties (the Vs and Q profiles) from the weak-motion data and to compare these with those obtained from the inversion of strong-motion data. We selected four KiK-net sites with different Vs profiles to apply this method. First, we analysed the global sensitivity of the frequency resonance peak and associated amplitudes to each soil parameter to determine their influence on the computed borehole site response, and consequently to find the soil parameters that can be well constrained during inversion. Then, we performed the inversion on the BFSR_{lin}. However, the results of the inversions are very difficult to interpret. This difficulty arises mostly because the numerical simulations cannot reproduce the low pseudo-resonance peak amplitudes observed, and also with the selection of sites that may not have a simple site configuration (one dimensional). For the FKSH14 site, the numerical simulation reproduced the BFSR_{lin} very well. Thus, we performed inversion of the non-linear BFSR at this site, and we found that for the three earthquakes that induced large modifications between the linear and non-linear site responses, the soil properties of the first three layers of the soil were affected the most. At the same time, during the strongest earthquake, the inversion results indicate that the fourth soil layer is also affected. This observation confirms that at this site at least, the non-linear soil behaviour mostly takes place in the subsurface soil layers, as suggested by the results of the previous chapter, although for very strong motions that induces large deformations, there is non-linear behaviour in deeper soil layers. These results are, for now, limited to one site, and they cannot be generalised, although they are in a very good agreement with the conclusions of the previous chapter, which provided a statistical analysis of the non-linear behaviour for 54 sites.

INTRODUCTION

As shown in the previous chapter, the study of the effects of non-linear soil behaviour on site responses indicates that the frequencies affected by these non-linear effects lies between the fundamental resonance frequency and the predominant resonance frequency, and that soil columns with high velocity contrasts close to the surface are prone to strong non-linear behaviour. These results appear to indicate that the soil non-linearity takes place mainly in the superficial soil layers. These conclusions can have a large influence when dealing with site characterisation. Indeed, measurements of non-linear soil behaviour are mainly performed using laboratory tests that require undisturbed soil samples. Consequently, such measurements are very expensive, especially at depth. Defining a depth below which non-linear soil behaviour does not have a large influence

on site response is thus a high priority. To go further in that direction, numerical simulations of wave propagation in a soil column that reproduce the non-linear site response can be used to check further the previous conclusions, and maybe to find the depths where the maximal soil non-linear behaviour occurs.

The non-linear soil column can be represented either as (i) a 'true' non-linear soil column, with the description of the non-linear behaviour of the material on each soil layer; or as (ii) an equivalent linear soil column (which suggests that only modifications of the shear modulus and damping of the layers are applied to an initial elastic soil column). In the present study, the non-linear soil column represents the second of these cases. An inversion approach associated to a linear one-dimensional (1-D) numerical simulation can be used to determine a family of solutions (equivalent linear soil columns) by reducing the deviation of the numerical simulation results to the observed non-linear site response (the so-called cost function of the problem). The aim of the inversion is to find the global minima of the cost function.

Several studies have already investigated the inversion of earthquake recordings, in both the time and frequency domains, to characterise the non-linear soil column. Satoh et al. (2001) estimated the strain-dependent non-linear characteristics by inversion of the Vs and Q profiles for different earthquake recordings. They used 1-D numerical simulations that were based on the well-known Haskell-Thomson method with frequency dependent attenuation. De Martin et al. (2010) analysed the non-linear soil response of a borehole station in Japan. To find the equivalent linear soil column, they used a genetic algorithm with a specific objective function based on the shift of the frequency peaks of the borehole Fourier spectral ratio (BFSR). They also used 1-D numerical simulations that were based on the well-known Haskell-Thomson method. They found a maximum reduction in the Vs at 35 m in depth. To find the velocity and attenuation structure, Assimaki and Steidl (2007) used a hybrid scheme composed of a genetic algorithm associated with a wavelet-domain cost function and local hill-climbing with frequency-domain inversion, to obtain the equivalent linear dynamic soil properties. Pavlenko and Irikura (2002) (2006) studied the non-linear soil behaviour during the 1995 Kobe and 2000 Tottori earthquakes recorded by KiK-net. They found a non-linear soil column described by stress-strain relationships.

In this chapter, we have used a 1-D numerical simulation based on the Haskell-Thompson theory with frequency-dependent attenuation, as presented in the first section. We used a simulated annealing-downhill simplex hybrid global inverse algorithm (Liu et al., 1995) to: (1) adjust the Vs logging that comes from the PS logging measurements to the linear site response; and (2) invert the equivalent linear Vs profile. The inverse algorithm is presented in the first part of this chapter, with tests on synthetic cases. Then, the application of the inverse algorithm is shown for empirical data at four selected KiK-net sites.

Chapter 5

The simulated annealing downhill simplex hybrid global inverse algorithm

Sommaire

INTRODUCTION

The aim of an inversion algorithm is to find the global minimum of a cost function. A cost function represents the differences between observations or measurements and the results of a set of numerical model. Finding the minima of the cost function provides the model that best simulates the observations or measurements. In our case, we are looking for the soil model; i.e. the shear-wave velocity and attenuation profiles with depth (1-D), by inverting the BFSR. Direct search methods using inversion algorithms that follow a path such that the cost function is decreased at each iteration are efficient for the finding of a local minimum, but these can fail when there are several local minima. Stochastic search methods, such as genetic algorithms or simulated annealing, on the other hand, give promising results for finding of global minima of the cost function, although these can be time consuming and they are dependent on the initial model. In the present study, we used a hybrid method that combines simulated annealing and a downhill simplex optimisation algorithm. The combination of both of these methods increases the efficiency of the search method and minimises the dependence of the results on the initial model (as a simplex of initial models is used). First, we present the two methods separately, and then we present the combined process used in the present study. The code used was developed by Pengcheng Liu (Liu et al., 1995).

5.1 SIMULATED ANNEALING

The purpose of simulated annealing is to find the global minimum of a cost function that might have several local minima. This probabilistic search method was initially proposed by Kirkpatrick et al. (1983) and Cerny (1985), and it models the physical process of heating a material and slowly cooling it, such that eventually its structure is frozen at a minimum-energy configuration. While the kinetic energy of the molecules of a gas in thermal equilibrium has a definite value, the energy of the individual molecules varies with time, changing from collision to collision. When dealing with a large number of molecules, as for a gas at equilibrium, the probability distribution of the energy is well known and is defined by the Boltzmann distribution. The Boltzmann distribution depends mainly on the temperature of the system. Instead of being a deterministic value, the energy of the system is defined as a random field with a probability that depends on the temperature. This statistical definition allows the system to have high energy although the temperature is low, with a small probability. During the process, if the initial heating temperature is not high enough, or if the material is cooled too quickly, then its structure can stabilize to a semi-crystalline state that does not correspond to a minimum energy configuration.
In practical terms, the method performs a random walk (from the initial model to the best model) in the parameter space that respects the Metropolis criterium (see next paragraph), to find the global minimum of a complex cost function. In analogy with the physical process, the energy of the system refers to the cost function, with the parameter governing the process that is defined as the

temperature of the system. Let us define the basic elements of simulated annealing:

- Fc, the cost function, is defined on S (the finite set). $S^* \in S$ are the set of global minima of the function Fc.
- $\forall i \in S, S(i) = S - \{i\}$ represents the neighbours of i.
- $\forall i \in$S and $\forall j \in S(i)$, we define the positive coefficients q_{ij} (that represent the probability of j being in $S(i)$), such that $\sum_{\forall j \in S(i)} q_{ij} = 1$, and also if $j \in S(i)$, then $i \in S(j)$.
- We define a function T:$\mathbb{N} \to \mathbb{R}^+$, which is called the cooling schedule. $T(t)$ is the temperature at a time t.
- We define an initial state $x(0) \in S$

The simulated annealing algorithm consists of a discrete-time inhomogeneous Markov chain. A Markov chain is a random process that states that the conditional probability distribution for the system at the next step only depends on the present state (and not on previous states). It is inhomogeneous, which means that the probability:
$P(X(n+1) = j|X(n) = i) \neq P(X(n) = j|X(n-1) = i)$. The initial model is defined randomly in the input parameter space (M^k, where k is the number of input parameters, and M is the bound space of each parameter), called $x(0)$. For the current state $x(t) = i$, we randomly chose $j \in S(i)$ with the probability q_{ij}. The Metropolis criterium that defines the transition between the steps $x(t)$ and $x(t+1)$ is defined as follows:

$$\begin{array}{ll} \text{If} \quad Fc(j) \leq Fc(i), \text{ then} & x(t+1) = j \\ \text{If} \quad Fc(j) > Fc(i) \text{ then} & \begin{cases} x(t+1) = j \text{ if } r < e^{-[Fc(j)-Fc(i)]/T(t)} \\ x(t+1) = i \text{ otherwise} \end{cases} \end{array} . \tag{5.1}$$

Formally, the probability that the j model is accepted is defined by the following equation:

$$P(x(t+1) = j \mid x(t) = i) = q_{ij} e^{-\frac{1}{T(t)} \max\{0, Fc(j)-Fc(i)\}}. \tag{5.2}$$

The Metropolis criterium indicates that if the cost function of the new model j, $Fc(j)$, is lower than the current model, $Fc(i)$, then the model j is accepted. If the mismatch is greater than the current model, the new model can be accepted if r (a random value between 0 and 1) is lower than the probability $e^{-[Fc(j)-Fc(i)]/T(t)}$, which depends on the deviation between the cost functions of i and j, and on the temperature of the system at time t.
The search technique involves three parameters that need to be defined: $T(t = 0)$, the first temperature; T_f, the final temperature; and a cooling parameter, a. The cooling schedule used in the

combined algorithm is defined by the following equation:

$$T(t) = \frac{T_0 - T_f}{2} \cos\left(\pi a^{0.25}\right) + \frac{T_0 + T_f}{2} \quad \text{with} \quad a = \frac{N_{trial}}{N},$$

where N_{trial} is the number of external loop iteration at time t and Ns is the maximal number of external loop iteration allowed. More details about this method can be find in Berstimas and Tsitsikis (1993). Applications of this method to characterise seismic sources (apparent source time function) were shown in Courboulex (1995), with inversions of the transfer function to find the velocity profiles of the crust below southern alps, Italy and Algeria areas in Bertrand (2000), and inversions of the surface waves using fast simulated annealing for inversion of surface waves using phase-velocity spectra Ryden and Park (2006).

5.2 DOWNHILL SIMPLEX METHOD

The downhill simplex method, or the Nelder-Mead (simplex) method, was initially introduced by (Nelder & Mead, 1965). The optimization algorithm does not require any assumption on the cost function, and especially the differentiability. This method relies on splices, which are polytopes of $k + 1$ dimensions, where k is the number of input parameters to be inverted. In one dimension, the simplex is a segment; in two dimensions, it is a triangle. This method is a heuristic technique. It initiates with a simplex of $k + 1$ dimension (X_1,..., X_k, X_{k+1}) that are $k + 1$ models that are randomly chosen in the boundary spaces of the input parameters. The first step consists in the calculation of the cost function of each model, and their sorting. For instance, let us assume that $Fc(X_1)$ <... $Fc(X_k)$ < $Fc(X_{k+1})$. The second step of this method involves the reflection, extension or contraction of one of the simplex elements that requires specific parameters to find the new models (α_{DS}, γ_{DS} and ρ_{DS}, for the reflection, extension and contraction, respectively). Standard values of these parameters are α_{DS} =1, γ_{DS} =2 and ρ_{DS} =-1/2. After the first step, the method reflects the worst model (here, X_{k+1}). Three results are possible:

- If the cost function of the reflected model is between the best model and second worst model, then, the worst model is replaced by the reflected one, and the step one is repeated until convergence is achieved.
- If the reflected solution is better than the initial best solution, then the method extends the reflected point. If the extended point is better than the best initial point, then the worst initial model is replaced by the extended one in the simplex, and the method repeats step one until convergence is reached. If the extended point is worse than the best initial one, the method replaces the worst initial model with the reflected one.
- If the reflected point is worse than the second worst model, then the method calculates a

contracted model. If the contracted model is better than the initial worst model, the method replaces it in the simplex, and step one of the method is repeated. If not, the method operates a reduction of the simplex by replacing all of the models except for the best one by the extension of the best model, and repeats step one.

5.3 THE HYBRID METHOD

The hybrid method used here is a combination of the two methods described above. This algorithm is composed of two nested loops. The internal loop (with an iteration number designated as N_{iter}) follows the steps of the downhill simplex method, but with stochastic implementations for finding the new models, and it is repeated until it reaches convergence (external loop, with an iteration number designated as N_{trial}). At each step of the internal loop, the reflected, extended or contracted models are evaluated according to the temperature of the system. The condition for the proposed model to replace the worst model of the simplex is based on the Metropolis criterium, although it is independent of the temperature ($P_{i,j} = P(X(t+1) = j \mid X(t) = i)$). The algorithm is illustrated in Figure 5.1. In the different steps of the process, the algorithm finds new models that depend on the worst model, the barycentre model of the simplex, and the temperature of the system, according to the following equation:

$$X_{new}(p) = X_0(p) + C(X_0(p) - X_{k+1}(p)) + \left(\frac{M_{max}(p) - M_{min}(p)}{R_{ed}} T(N_{trial}) \right), \qquad (5.3)$$

where, p is the index of the parameter to be inverted. X_{new} can be X_r, X_e, X_c, the reflected, extended or contracted new models, respectively, C can be the α_{DS}, γ_{DS}, and ρ_{DS} associated constant parameters. X_{k+1} is the worst model of the simplex. M_{max}(p) and $M_{min}(p)$ are the limits of the input space parameter p. R_{ed} is a random coefficient between 0 and 1. $T(N_{trial})$ is the temperature of the system at the iteration N_{trial}. The probability $P_{i,j}$ of acceptance of a new model in the simplex is given by the following equations:

$$P_{i,j} = V_{conv}^{b_{conv}}, \qquad (5.4)$$

where, V_{conv} is a pre-defined parameter, called accepting rate. Smaller accepting rate values (0.4-0.6) speed the convergence but must be used for simple problem, whereas for more complicated inversion, accept ion rate between 0.6 and 0.8 should be used. Here, we chose the middle value proposed for complicated problem (0.7), considering that the cost function may be complicated. b_{conv} represents the normalised difference between the cost function of the new model compared to

the current model:

$$b_{conv} = \frac{Fc(j) - Fc(i)}{\overline{Fc}(1, .., .k + 1)}. \tag{5.5}$$

The external loop repeats the internal one until it converges. The convergence of the algorithm is based on the following stopping criterion (see Equation 5.6), that was introduced by (Woods, 1985). Preliminary testing has indicated that it is a useful stopping criterion of the simplex downhill method.

$$C = \frac{Fc(X_{n+1}) - Fc(X_0)}{[Fc(X_0) + Fc(X_{n+1})]/2}. \tag{5.6}$$

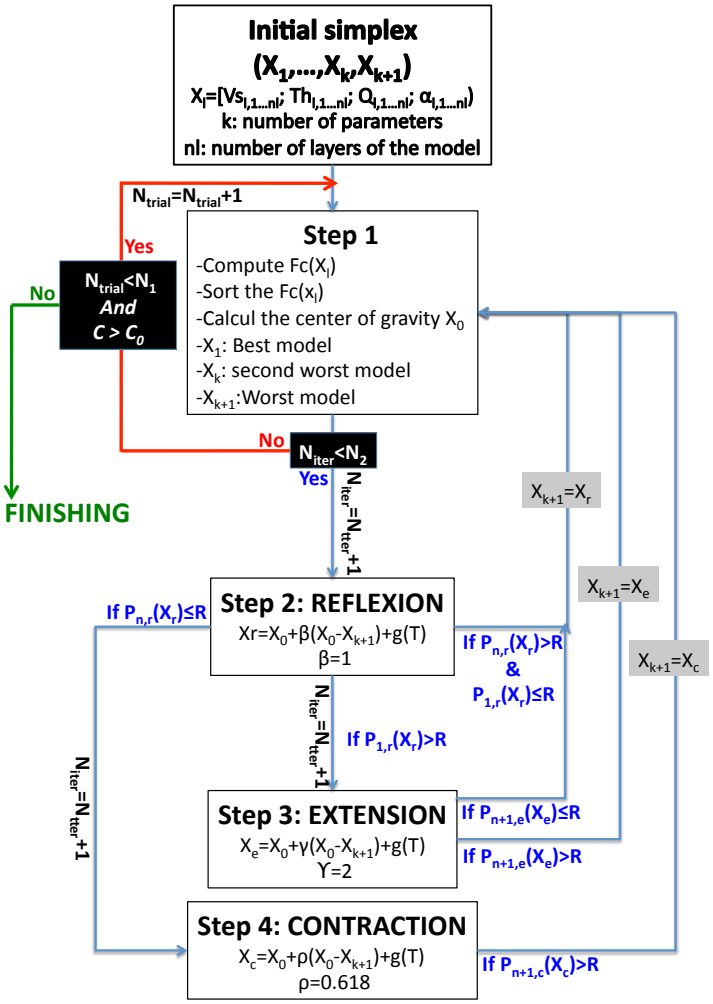


Figure 5.1: *Schematic of the inversion algorithm*

5.4 INVERSIONS ON SYNTHETIC CASES

The forward problem is the calculation of the borehole site response as presented in the chapter one. We used the well-known Haskell-Thomson method (Thomson, 1950; Haskell, 1953) using a non-viscous type Kevin-Voigt model. In the first chapter, we discuss the analytical formulation of the borehole transfer function for inversion purposes (section 1.3 page 46), we underline the following shortcomings:

- No information on the soil properties below the downhole sensor can be inverted.
- For a mono-layer case, Vs profiles having similar ratio of the Vs and thickness of the sediment layer give exactly similar transfer function.

In the same section, the cost function has been already defined, but for convenience will be recall here. The cost function is the difference between the transfer function calculated for a referent model or an empirical transfer function and a transfer function calculated with a the site parameters proposed by the inversion algorithm:

$$Fc = \sum_{j=N_1}^{N_2} \left(BFSR_{ref}(j) - BFSR_{num}(j)\right)^2 \tag{5.7}$$

Where, BFSR$_{ref}$ is the borehole Fourier spectral ratio (BFSR), as measured or referent, and BFSR$_{num}$ is the calculated BFSR. Fc represents the mismatch between the referent and the calculated transfer functions. In the present study, we choose to calculate the overall mismatch between the BFSRs at 250 discrete frequencies regularly sampled on the frequency interval [f(N_1)-f(N_2)]. The frequency interval is adapted to each site depending on the frequency band amplified in the site response.

Considering the second shortcoming of the borehole transfer function inversion and the definition of the cost function, the cost function in the Vs-thickness plane is characterised by a valley, which direction follows the Vs to Thickness ratio of the referent model and with similar amplitude at the very bottom of the valley. Therefore, the inversion will not be able to discriminate between soil profiles having similar Vs to thickness ratio, because of the trade-off between these two parameters. Only a combination of these two parameters will be solved by the inversion.

5.4.1 Test of the initial model effect

To test the inversion method, a first step is to perform the analysis on synthetic cases. We ran an auto-test by computing a numerical transfer function from a synthetic case (forward calculation) before inverting it. Let us recall that the first step of the inversion technique is to define a simplex of randomly chosen models. In this synthetic test (monolayer case with: Vs = 300 m/s, Thickness of

the sedimentary layer equal to 20 m, quality factor = 80 and volume mass equal to 2000 kg/m^3), we imposed one of the initial models of the simplex, and the others are randomly chosen, as proposed for the previously described algorithm. We tested an initial model that corresponded exactly to the solution, and one initial model that was located outside the valley of the cost function in the Vs-Thickness plane (see Figure 1.9 on page 48). We also performed tests on the convergence of the inversion algorithm depending on different combinations of the maximum number of iterations to be performed for the internal and external loops.

5.4.1.1 One of the initial models is the true solution

When the initial model is the true solution, the best solution of the simplex is always this initial model. In Figure 5.2, we show the results of the inversion method. The upper graphs show the input parameters with depth (Vs, Q_0 and α, the attenuation parameter term: $Q = Q_0(f^\alpha$, note that Q_0 stand for the quality factor at 1 Hz). The middle graph shows the associated transfer function, and the bottom graph shows the residuals between the true solution and the calculated transfer function with the frequency. In each graph, we plotted the true model, all of the model of the initial simplex, the best model of the simplex every five iterations of the external loop, and finally the best-fit solution. In Figure 5.2, all of the curves are superimposed, except for those that come from the initial simplex, which means that the best model of the simplex is the initially imposed model at each iteration. The inversion retrieves the true model perfectly.

5.4.1.2 One of the initial models shows a thinner sediment layer

Then, we tested an initial model that shows a higher velocity in the sediment (400 m/s) and a lower thickness (15 m) than the actual thickness. The cost function in the Vs-thickness plane is shown in Figure 5.3(a), along with the path that the inversion technique used to find its final solution. The red points in Figure 5.3(a) represent the initial model and the seven other randomly chosen models that composed the initial simplex (two layers with four parameters, Vs, Th, Q_0 and α, which makes eight parameters to be inverted). Considering that the attenuation is a compromise between Q_0 and α, here we deliberately limited the boundary of the α parameter around 0.7. The white star in Figure 5.3(a) represents the true solution, and the black and white points are the best models of the simplex at the different iterations of the inversion external loop (from black to white: increasing iterations, N_{trial}). Figure 5.3(b) shows the path walked by the algorithm in three dimensions, and the white star represents the true solution. The cost function value of the final solution (10^{-6}) is greater than the cost function that was found using the true model as solution (10^{-12}), although the bottom of the cost function valley (i.e. the Vs and thickness couple for which the ratio is similar to the true model, here 15) have similar amplitude. In Figure 5.3(b), it can be noted that the algorithm lies in the valley of the cost function (but not at the very bottom), and tries different locations in the valley that are close in terms of the cost function but not equivalent; it

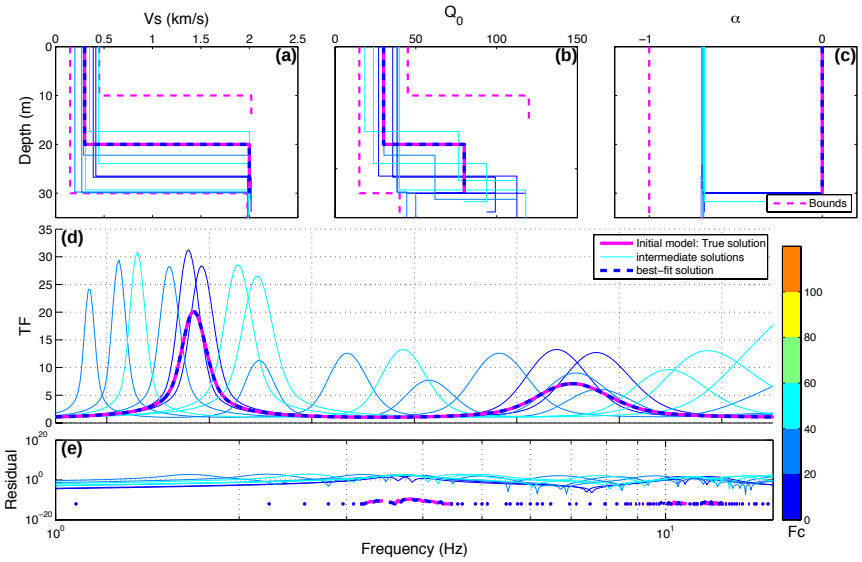


Figure 5.2: *Inversion method on synthetic cases : results for the monolayer soil model, with the true model as one of the initial models (a) Inverted Vs profiles, showing the synthetic (true solution) Vs profile (purple line), the boundary of the inversion (purple dotted lines; here we chose a CV of 60%), and the examples of the Vs profiles tested by the inversion (thin lines) with colours according to the values of the associated cost function. The best-fit family of the Vs profiles are also shown (blue line; the colour depends on the Fc). (b) As for (a), for the quality factor. (c) As for (a), for the attenuation coefficient. (d) The BFSR from the soil column described in (a), (b) and (c), compared to the BFSR from the true soil model (purple curve). (d) The residuals of the BFSR compared to BFSR from the true soil model .*

finally stops at a local minimum, and does not reach the true solution.

The soil profile family of the best-fit models is shown in Figure 5.4. The best-fit solution is very close to the true solution, with similar frequency peaks, but it has slightly different Vs profiles (the Vs of the sedimentary layers is 272 m/s, instead of 300 m/s, and the thickness is 18 m, instead of 20 m).

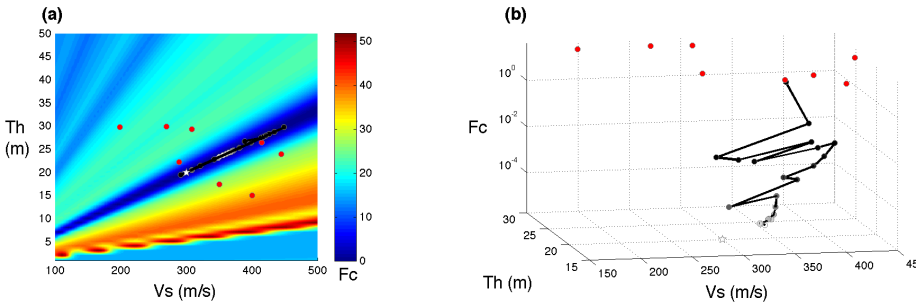


Figure 5.3: *(a) Cost function in the Vs-Th plane with the location of the initial models tested by the inversion (red points) and the models tested during the simplex downhill hybrid algorithm. The white star indicates the true solution, the gray scale of the points from black to white corresponds to the iteration number of the inversion algorithm. (b) Path followed by the inversion algorithm from the initial simplex (red points) to the final best-fit solution (white point), passing through the intermediate best solutions of the simplex (gray points).*

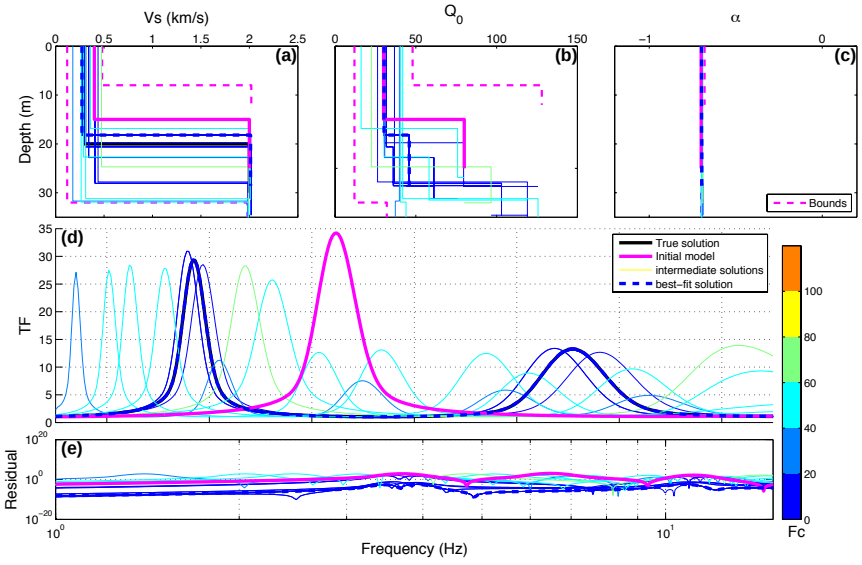


Figure 5.4: *Similar to Figure 5.2 but for a monolayer soil model without the true solution as one of the initial models.*

5.4.1.3 conclusion

In the synthetic cases presented above, the inversion method succeeds in retrieving approximately the true solution. For both of the synthetic cases, whatever the number of iterations used (100 or 1000), the inversion method did not converge. The convergence criterion for the algorithm is based on the ratio between the maximum deviation of the cost function in the simplex ($\Delta(Fc(X_1); Fc(X_n))$, X_1 being the best model of the simplex (lowest cost function) and X_n the worst model) and the mean of the cost function of the best and worst models of the simplex ($(Fc(X_1) - Fc(X_n))/2$). In synthetic cases, the valley of the cost function is flat at the bottom and when looking at a logarithmic scale (see Figure 1.9 page 48), the valley is very narrow, which makes the numerator of the convergence criterion equation (that can be approximate to the slope of the valley) much higher than the denominator (that is very low close to the bottom of the valley). The narrowness (associated to the high slope of the edges) of the cost function, prevents the convergence criterion equation from decreasing down to the pre-defined tolerance, here chosen as 1%. The following section is devoted to the analysis of the number of internal iterations on the convergence and on the cost function of the 'best-fit' solution.

5.4.2 Test of convergence

We tested the effects of changing the iteration number of the internal loop of the inversion method. For a similar number of external loop (1000), we tested 100, 500 and 1000 as the maximum numbers of the internal loop. Figure 5.5(a) shows the evolution of $Fc(X_0)$; i.e., the cost function of the best solution of the simplex at each external iteration. Using larger numbers of internal iterations naturally appears to be more time consuming, although this gives the best-fit solution with a lower cost function. In Figure 5.5(b), we show the evolution of the stopping criterion (C; see Equation 5.6). We observe that the stopping criterion did not drop below 0.5. We chose a relative tolerance of 0.01 (1%) in these computations. This means that convergence is not likely to be reached however many iterations are performed (internal and external). This can be explained by the narrowness of the minima in the synthetic cases, as explained in the previous paragraph. For the inversion of the empirical data, we will see in the following section that the convergence is achieved.

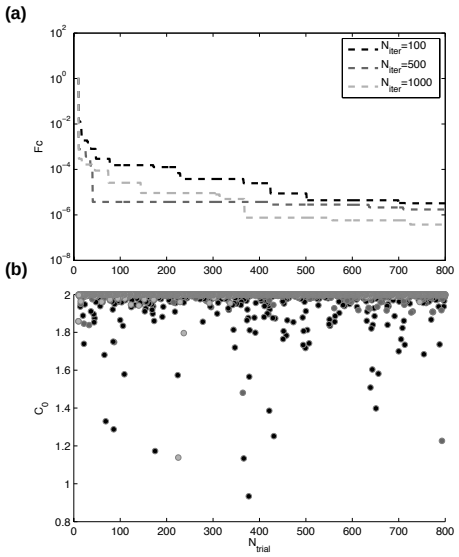


Figure 5.5: *(a) Value of the cost function according to the iteration of the external loop, for three different maximal values of the internal loop. (b) Evolution of the convergence criterion according to the external loop iteration,colour coded as for (a).*

In addition, we also compared the cost function and the stopping criterion of the inversion method with three different maximum numbers of internal loops, according to the time of computation. For a comparative time of computation, the method implemented with 100 maximal internal loops is applied, with a maximal number of external loops of 5000. The data are illustrated in Figure 5.6. We observed that the more efficient scheme in terms of the cost function for a given time, is the one implemented with the largest maximal number of internal loops. The method that was implemented with 100 or 500 maximal internal loops is fastest, but it appears to saturate at a threshold cost function value.

5.4.3 Conclusions on the synthetic cases

To conclude, while the narrowness of the cost function valley (in the Vs-thickness plan) prevented the inversion algorithm to converge, it facilitated the obtention of a solution that could be close to the true one. We showed that the main short coming of the inversion of borehole transfer function is the flatness of the bottom valley that prevent the inversion from discriminate between soil profiles having similar Vs to thickness (of the sediments) ratio. For the inversion on empirical data, we found that using a large number of internal loop, although time consuming, help decreasing the cost function. Implementing an initial model pre-defined in the initial simplex could be interesting

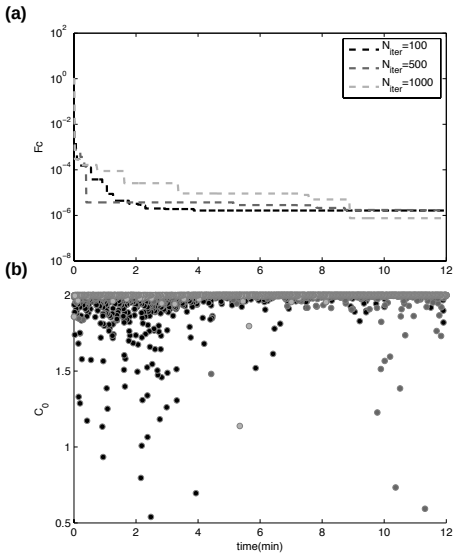


Figure 5.6: *(a) Value of the cost function according to the computation time for three different maximal values of the internal loop. (b) Evolution of the convergence criterion according to the computation time.*

and the choice of the initial model should not have strong influence on the result except when the initial model is the true one.

In this section, we choose to analyse synthetic only mono-layer soil column cases, considering that the cost function in two dimensions (Vs and thickness) can be represented. Therefore, the overall working of the inversion algorithm can be understood. More complicated synthetic cases involving multilayer soil column will represent more than four parameters to be inverted and the cost function will not be easily shown.

This inversion algorithm gives the best-fit solution rather than a family of soil profiles. However, considering the trade-off between the Vs and the thickness of the sediment layers it could be interesting to consider family of soil profiles solutions. When inverting the empirical data we will show the best fit solution (which is the result of the algorithm used in the present study) but we will show as well the intermediate solutions to have an overview of the family of profiles that have low cost function amplitude.

Chapter 6

Inversion of empirical data

Sommaire

6.1 INTRODUCTION

We divided the inversion of the empirical data into two steps. In the first step, we inverted the linear site response, to improve the Vs and to define the quality factor profile. In the second step, we performed the inversion of the non-linear site response by fixing the thickness of the soil layers and adjusting only the Vs and Q profiles. We selected KiK-net sites such that the linear site response was as close as possible to a 1-D configuration. Hence, we used the KiK-net Vs profile as one of the initial models of the simplex, and we gave some freedom to the Vs and the thickness for the inversion.

6.2 DESCRIPTION OF THE SELECTED KIK-NET SITES

6.2.1 Selection of the KiK-net sites

In the KiK-net database, the first task was the identification of sites, that were characterised by strong non-linear behaviour. The selected sites had to fulfill the following criterion:

1. The sites recorded at least two strong events with the peak ground acceleration (PGA) at the downhole station higher than 50 cm/s^2.

2. The sites are characterised by strong non-linear behaviour.

3. The sites show a relative simplicity in the geometry of the geological structure.

To fulfill the first two criteria, we selected the sites from among the 54 sites that were used in chapter 5 (non-linear site responses) with large PNL$_{site}$ (>10%) (percentage of non-linearity). To achieve the third criterion here, we crossed the previous selection with the selection of the 1-D site configurations that were performed in chapter 4 (linear site-response variability). We ended up with 11 sites, the characteristics of which are detailed in Table 6.1. From among these 11 sites, we selected four sites that have different soil columns (different Vs$_{30}$ and B$_{30}$, gradient of the Vs profile), namely FKSH12, FKSH14, KSRH10 and MYGH04.

Table 6.1: *Characteristics of the preliminary selection of KiK-net sites*

Station	Vs_{30}	f_{pred}	A_{pred}	f_0	B_{30}	PNL_{site}	$N_{eq}(PGA > 50cm/s^2)$
AKTH04	459	4.83	14.7	3.4	0.43	10.4	2
FKSH12	449	4.34	42.5	4.1	0.51	27.5	3
FKSH14	237	4.15	12.4	1.2	0.22	16.7	2
IBRH07	107	0.59	116.4	0.2	0.28	10.6	2
IBRH13	335	11.27	25.3	2.6	0.26	16.6	3
IWTH04	456	3.37	20.2	2.9	0.53	12.3	3
IWTH22	532	7.83	18.9	7.6	0.7	15.7	3
IWTH23	923	14.1	15.1	12.9	0.62	18.5	4
MYGH02	399	5.49	11.8	5.6	0.33	12	2
MYGH04	850	15.01	15.8	12.7	0.71	16.7	3
NIGH06	336	4.2	16	3	0.74	21.1	2

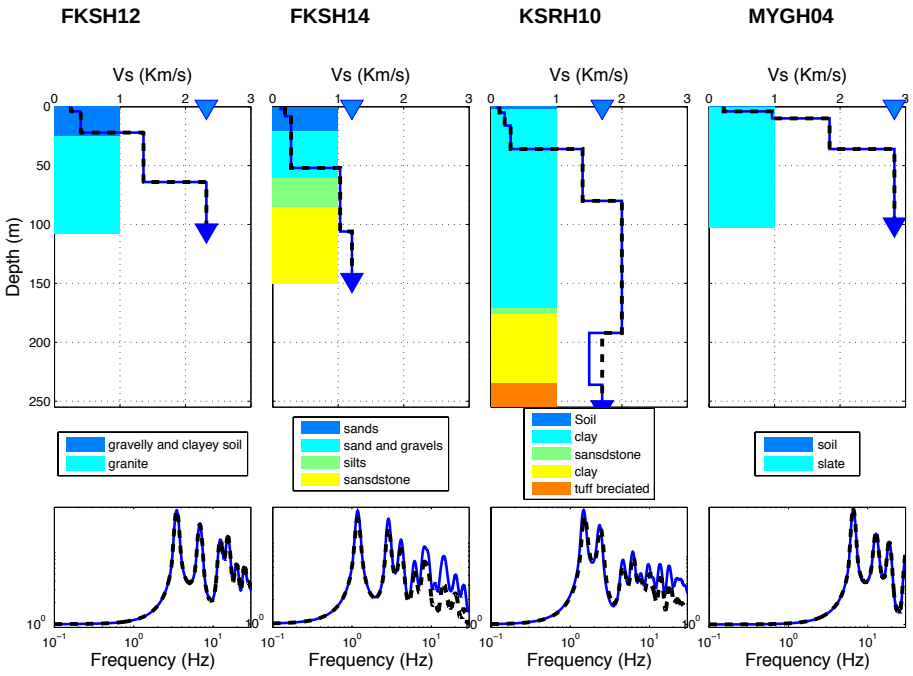


Figure 6.1: *Initial and simplified proposed Vs profiles and geological logging of the selected sites: FKSH12, FKSH14, KSRH10 and MYGH04. The bottom graphs show the comparisons of the $BFSR_{num}$, calculated using the initial and simplified Vs profiles.*

6.2.2 Description of the KiK-net sites

FKSH12: This site is in the Fukushima district. The topography of the site is relatively flat. The soil column is composed of 25 m of gravelly and clayey soil, located on 80 m of granite at least (Figure 6.1). This station has recorded more than 220 earthquakes, among which three have produced PGA at the down-hole sensor over 50 cm/s^2. Figure 6.2(a) illustrates the Vs profile of the station. Figure 6.2(b) shows how that the numerical simulations fit the observations at low strain (the first frequency peaks are close by 5.8 %). At higher strain, the deviation between the linear site response and the non-linear site response is large (see Figure 6.2(c)), and the first peak at around 5 Hz in the linear characterisation is shifted to 3 Hz to 4 Hz when the site response is computed with strong-motion recordings.

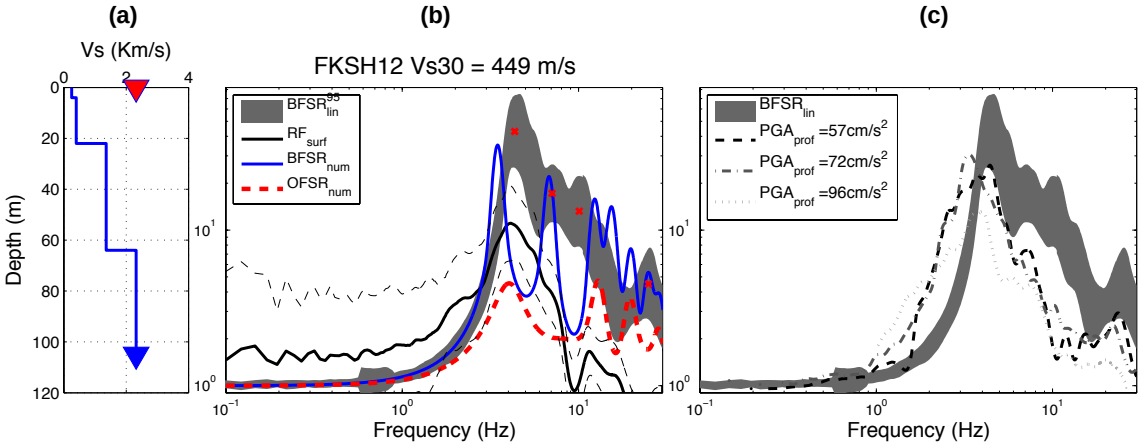


Figure 6.2: *(a) Vs profile of the FKSH12 station. (b) Comparison of the empirical and numerical site response evaluations. (c) Comparison of the BFSR$_{lin}$ and the BFSR$_{emp}$ performed using strong-motion recordings.*

FKSH14: This site is also in the Fukushima district, on the east coast of Japan. The soil is composed of 20 m of sands at the top of the soil column, followed by 40 m of sands and gravels, and then 25 m of silts, and finally 80 m of sandstone (Figure 6.1). This station has recorded more than 300 earthquakes, for which two of them generated PGA at the downhole station over 50 cm/s^2. Figure 6.3(a) illustrates the Vs profile of the station. Figure 6.3(b) shows how that the numerical simulations fit the observations at low strain (the first frequency peaks are close by 1 %). At higher strain, the observations indicate that the deviation between the linear site response and the non-linear site response is large (see Figure 6.3(c)), and the first peak at around 1 Hz in the linear characterisation is shifted to 0.9 Hz, and the high frequencies from 3 Hz are largely attenuated when the site response is computed with strong-motion recordings.

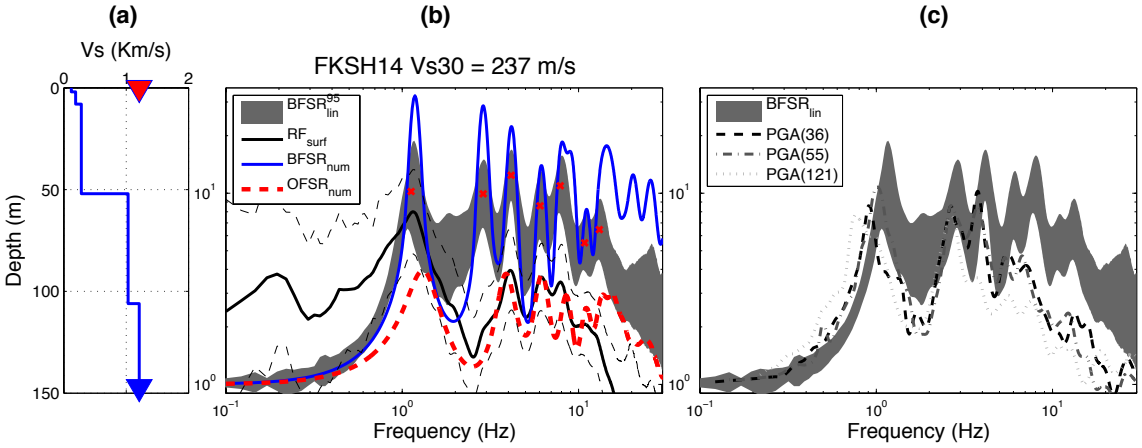


Figure 6.3: *(a) Vs profile of the FKSH14 station. (b) Comparison of the empirical and numerical site response evaluations. (c) Comparison of the BFSR$_{lin}$ and the BFSR$_{emp}$ performed using strong-motion recordings.*

KSRH10: This site is in the Hokkaido district, on the east coast of Japan. The soil is mostly composed of clay (250 m) on a base of breciated tuff (Figure 6.1). This station has recorded more than 160 earthquakes, and two of these generated PGA at the downhole station over 50 cm/s^2. Figure 6.4(a) illustrates the Vs profile of the station. Figure 6.4(b) shows how that the numerical simulations fit the observations at low strain (the first frequency peaks are close by 3 %). At higher strain, the observations indicate that the deviation between the linear site response and the non-linear site response is large (see Figure 6.4(c)), and the first peak, at around 1.6 Hz in the linear characterisation, is shifted to 1.1 Hz, with the high frequencies from 6 Hz largely attenuated when the site response is computed with strong-motion recordings.

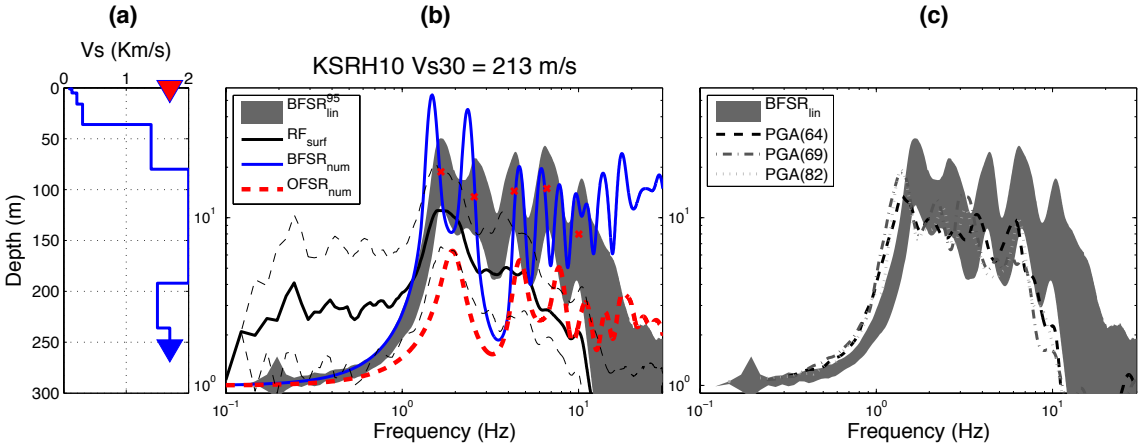


Figure 6.4: *(a) Vs profile of the KSRH10 station. (b) Comparison of the empirical and numerical site response evaluations. (c) Comparison of the BFSR$_{lin}$ and the BFSR$_{emp}$ performed using strong-motion recordings.*

MYGH04: This site is in the Miyagiken district. The station is on rock (slate) with a very thin surface layer of sedimentary deposits (4 m) (Figure 6.1). This station has recorded more than 400 earthquakes, three of which generated PGA at the downhole station over 50 cm/s^2. Figure 6.5(a) illustrates the Vs profile of the station. Figure 6.5(b) shows how that the numerical simulations fit the observations at low strain (the first frequency peaks are close by 1.5%). At higher strain, the observations indicate that the deviation between the linear site response and the non-linear site response is large (see Figure 6.5(c)), and the peak at around 15 Hz in the linear characterisation is shifted up to 7 Hz when the site response is computed with strong-motion recordings.

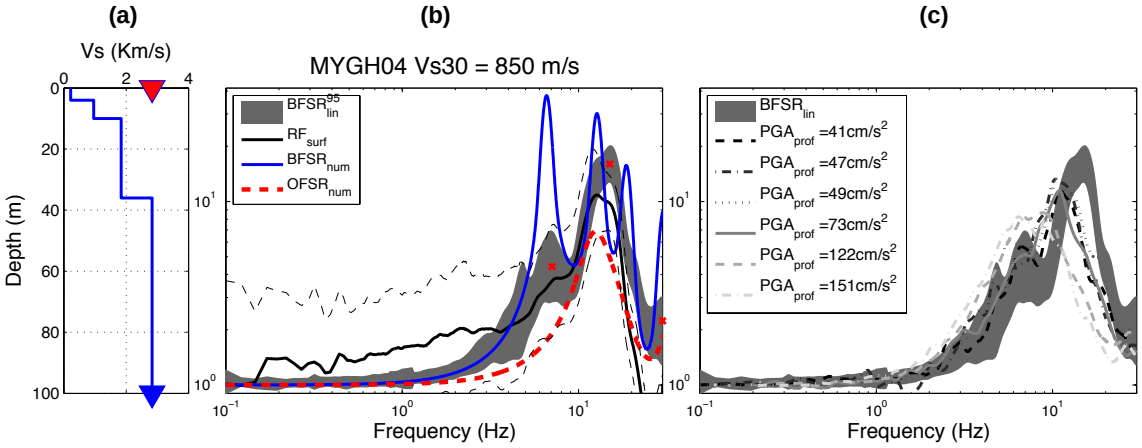


Figure 6.5: *(a) Vs profile of the MYGH04 station. (b) Comparison of the empirical and numerical site response evaluations. (c) Comparison of the BFSR$_{lin}$ and the BFSR$_{emp}$ performed using strong-motion recordings.*

6.2.3 Sensitivity analysis

We analysed the global sensitivity of the site-response peak frequencies of the selected sites. We only investigated the global sensitivity of the frequency peaks and the associated amplitudes to the Vs and Q of each soil layer. The data from the sensitivity analyses for the four sites, FKSH12, FKSH14, KSRH10 and MYGH04, are illustrated in Figures 6.6, 6.7, 6.8 and 6.9, respectively.

FKSH12: We analysed the global sensitivity of the first three peaks, f_0, f_1 and f_2, in the mean BFSR$_{lin}$ at 4.3 Hz, 7 Hz and 10 Hz, respectively (see Figure 6.2, red crosses). As shown in Figure 6.6, these peak frequencies are mainly controlled by the Vs of the second soil layer, and the first two frequency peaks also depend on the values of the Vs of the third and fourth

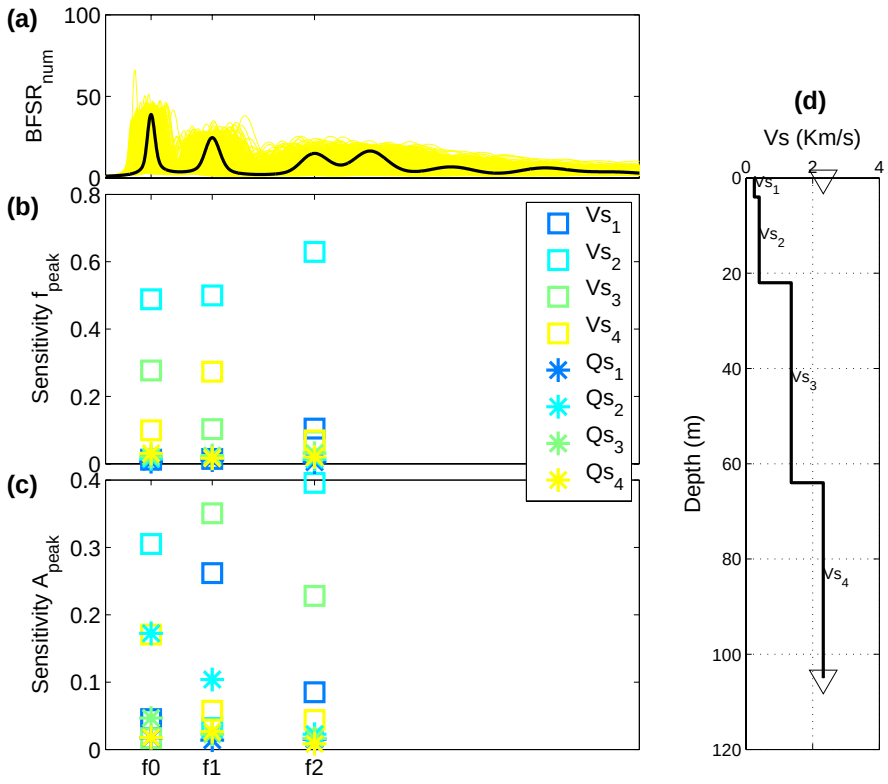


Figure 6.6: *(a) $BFSR_{num}$ for the mean velocity profile from KiK-net logging of FKSH12 (black), and the simulations tested (yellow). (b) The first sensitivity index of the peak frequencies of Vs and Q of the soil layers. (c) The first sensitivity index of the associated amplitude frequencies of Vs and Q of the soil layers. (d) The Vs profile at the FKSH12 station.*

soil layers. The dependence of the associated amplitudes on any specific parameter is less clear (the sensitivity was never greater than 0.4). The amplitude of the first peak is controlled by the parameters of the second soil layers (Vs and Q) and the Vs of the fourth soil layer. The Vs of the first soil layer only has an impact on the amplitude of the second peak, with less weight than the Vs of the third soil layer. At the peak frequencies, the Vs of the shallowest soil layers and the Q profile (except for the second soil layer) cannot be constrained.

FKSH14: We analysed the global sensitivity of the first five resonance peaks, f_0, f_1, f_2, f_3 and f_4, in the mean BFSR$_{lin}$ at 1.1 Hz, 2.9 Hz, 4.1 Hz, 6.0 Hz and 7.9 Hz, respectively. As shown in Figure 6.7, the frequencies of all of the peaks except the second one depend mainly on the Vs of the third soil layer. The second peak frequency depends on the Vs of the deepest soil layers. The third soil layer represents the thickest soil layer with low velocity above the largest impedance contrast of the soil column. This means that f_0, f_2 and f_3 are likely to be 'true' resonances (fundamental and harmonic) of the site that are associated with the velocity contrast at 50 m in depth, f_1 is a pseudo-resonance, and f_4 is a mix between harmonics of f_0 and the pseudo-resonances. Again, the dependence of the amplitude on any specific parameter is quite low. The amplitude of the first peak is mainly controlled by the parameter of the third soil layer. The second peak amplitude is related to the Vs of the second soil layer. The amplitude of the third peak is linked to the Vs of the fifth soil layer. For inversion purposes up to 10 Hz, the parameters of the third, second and the deepest soil layers should be well constrained.

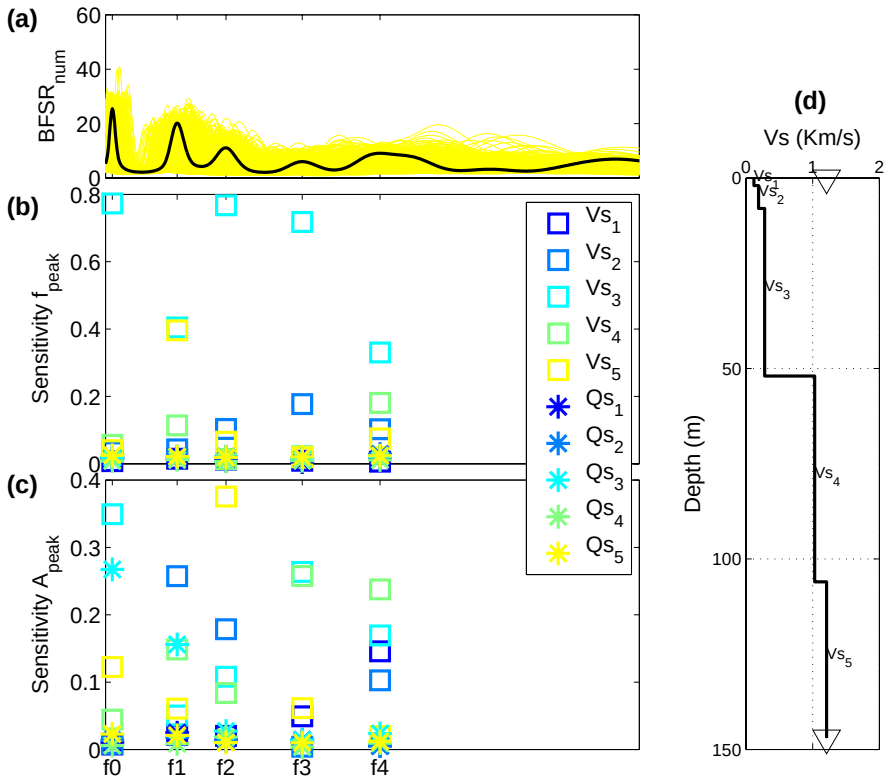


Figure 6.7: *(a) $BFSR_{num}$ for the mean velocity profile from KiK-net logging of FKSH14 (black), and the simulations tested (yellow). (b) The first sensitivity index of the peak frequencies of Vs and Q of the soil layers. (c) The first sensitivity index of the associated amplitude frequencies of Vs and Q of the soil layers. (d) The Vs profile at the FKSH14 station.*

KSRH10: We analysed the global sensitivity of the first five resonance peaks, f_0, f_1, f_2, f_3 and f_4, in the mean BFSR$_{lin}$ at 1.6 Hz, 2.6 Hz, 4.3 Hz, 6.6 Hz and 10.0 Hz. This soil column is composed of eight soil layers, which implies the analysis of 16 parameters for the sensitivity. As shown in Figure 6.8, the first two frequency peaks depend on the Vs of the fourth soil layer, which represents, as already observed for the two previous sites, the thickest soil layer before the largest impedance contrast at 40 m in depth. The first peak depend also on the Vs of the two deepest soil layers. As shown in the Figure 6.4, although the first two peaks are observed on the BFSR, only one peak is observed in the OFSR. Therefore, f_0 and f_1 appear to be a combination between the fundamental resonance frequency (as shown by the outrop site response) and a pseudo-resonance frequency peak. The third frequency peak depends mainly on the Vs of the third and fourth soil layers. This peak can be related either to a harmonic of f_0 (contrast at around 40 m in depth) or to a specific resonance of a shallower soil layer, or more likely to a combination of both of these phenomena. The last two peaks are related to the Vs of the deepest soil layers. The amplitude of the first peak is associated with the fourth soil layer parameters. The other amplitudes depend on the Vs of the deepest soil layers. For inversion purposes up to 10 Hz, the parameters of the fourth and deepest soil layers should be well constrained.

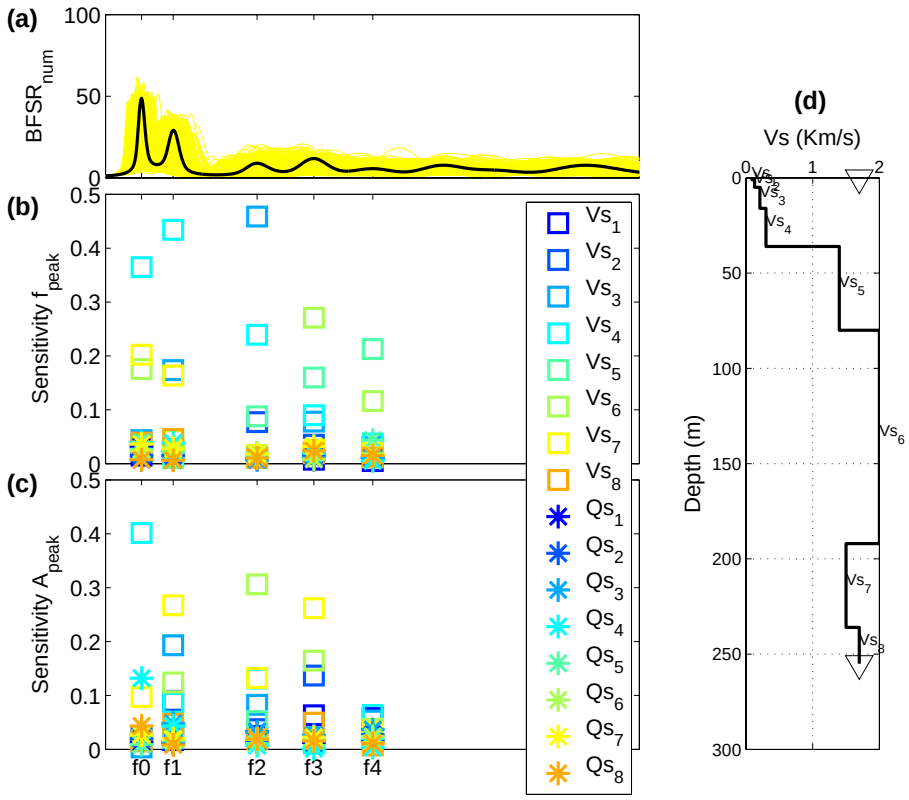


Figure 6.8: *(a) $BFSR_{num}$ for the mean velocity profile from KiK-net logging of KSRH10 (black), and the simulations tested (yellow). (b) The first sensitivity index of the peak frequencies of Vs and Q of the soil layers. (c) The first sensitivity index of the associated amplitude frequencies of Vs and Q of the soil layers. (d) The Vs profile at the KSRH10 station.*

MYGH04: We analysed the global sensitivity of the first three resonance peaks, f_0, f_1 and f_2, in the mean BFSR$_{lin}$ at 7 Hz and 15 Hz; the two peaks in the numerical simulations at around 15 Hz are not dissociated in the empirical site response. As shown in Figure 6.9, the first peak frequency depends on the Vs of the deepest soil layer, which indicates that it is a pseudo-resonance. The second peak is associated with the Vs of the first soil layer, and it is related to the impedance contrast at 4 m in depth. The third peak in the numerical BFSR (f_2) is linked to the deepest soil layers, and is likely to be related to pseudo-resonances. These observations imply that the peak at 15 Hz in the BFSR$_{emp}$ is a combination of the resonance due to the impedance contrast at 4 m in depth and to pseudo-resonances. The amplitudes of the peak frequencies depend mainly on the Vs of the deepest soil layer.

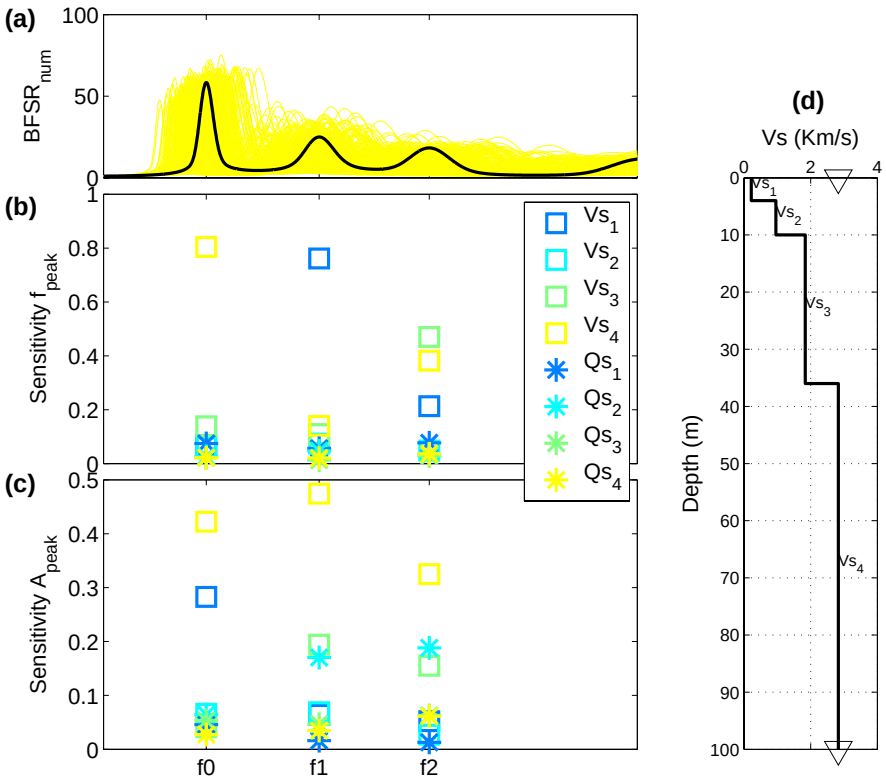


Figure 6.9: *(a) BFSR$_{num}$ for the mean velocity profile from KiK-net logging of MYGH04 (black), and the simulations tested (yellow). (b) The first sensitivity index of the peak frequencies of Vs and Q of the soil layers. (c) The first sensitivity index of the associated amplitude frequencies of Vs and Q of the soil layers. (d) The Vs profile at the MYGH04 station.*

The global sensitivity analyses performed on these four sites highlight some common aspects:

- The inversion of the transfer function below 10 Hz cannot constrain the very shallowest Vs.
- The soil layers above the largest impedance contrast can be well constrained by the inversion, for both Vs and Q.
- The pseudo-resonance peak inversion can provide information on the velocity of the deepest soil layers, close to the sensor depth.
- The Q of all of the soil layers except the one just above the greatest impedance contrast cannot be constrained by inversion.

6.2.4 Simplification of the Vs profiles

The aim of this section is to investigate the location of the non-linearity, by comparating the inversion of the linear and the non-linear transfer function. Considering the previous comments and to minimise the number of parameters to be inverted, the velocity profiles are simplified. The velocity profiles of the FKSH12 and MYGH04 sites are not modified (considering that the soil model have only four layers), the first thin soil layer at the FKSH14 site is merged with the second soil layer, and at the KSRH10 site, the first two and the last two soil layers are merged. The Vs profiles that were simplified with the associated BFSR$_{num}$ are illustrated in Figure 6.1, and these indicate that below 10 Hz to 15 Hz the simplifications do not change the site response significantly, although they minimise the number of parameters to be inverted.

6.3 INVERSION OF LINEAR DATA

We first inverted the linear transfer function of the selected sites to adjust the Vs profiles from the PS logging to the linear observations. We used the BFSR$_{lin}$ (mean of the BFSR computed with recordings with PGA $<$ 10 cm/s^2). In this work, we do not take into account the variability of the BFSR$_{lin}$. The parameter α of the frequency-dependent attenuation (see equation 1.51 on page 52) was fixed at 0.7, which represents a standard value (e.g. Satoh et al., 2001; Yoshimoto et al., 1992). The bound space of the input parameters was defined by the coefficient of variation, that give the values of the lower and upper bound with respect to the mean value that come from the Vs profile coming from the KiK-net database (of simplified profiles) such as:

$$\begin{aligned} P_{lo,ij} &= P_{m,ij} * (1 - CV_i/100) \\ P_{up,ij} &= P_{m,ij} * (1 + CV_i/100) \end{aligned} \tag{6.1}$$

Where, $P_{lo,ij}$ is the lower bound of the parameter i of the soil layer j. $P_{up,ij}$ is the upper bound of the parameter i of the soil layer j. $P_{m,ij}$ is the mean value of the parameter i of the soil layer j. CV_i is the coefficient of variation in percentage of the parameter i (that is constant for each soil layers). The properties of the soil below the down-hole station cannot be variable in the inversion algorithm used in the present study. Thus, in order to not created an impedance contrast at the downhole station that is not existent, the last soil layer properties were not variable.

For the four sites we invert the linear borehole site response on different frequency band:

- For FKSH12, we inverted between 1 to 20 Hz (the clear frequency peaks being at 4.3, 7 and 10 Hz). For
- For FKSH14, we inverted between 0.2 to 10 Hz (the clear frequency peaks being at 1.1 Hz, 2.9 Hz, 4.1 Hz, 6.0 Hz and 7.9 Hz).
- For KSRH10, we inverted between 0.5 to 8 Hz (the clear frequency peaks being at 1.6 Hz, 2.6 Hz, 4.3 Hz and 6.6 Hz).
- For MYGH04, we inverted between 1 to 20 Hz and then between 10 and 20 Hz (the clear frequency peaks being at 7 Hz and 15 Hz).

The Table 6.2 synthesizes the parameters of the inversions performed at each of the four selected sites, i.e. the frequency band inverted, the number of inversions performed, if they are illustrated in the manuscript and the soil parameters inverted with the associated coefficient of variation.

FKSH12: We tested four different inversions. For the first two inversions, only the Vs in the soil layers were inverted, with two different coefficients of variation (CVs) (the Q and the thickness were almost fixed). For the third inversion, the Q was variable as well, and in the last inversion, the three parameters were variable. We inverted the BFSR_{lin} from 1 Hz to 20 Hz. In the first inversion, we tested an inversion with a CV on Vs of 40%, with 5% for the other parameters (see Figure 6.10). The best-fit solution indicated a Vs of the second and third soil layers equal to the limit of the Vs boundary. Thus, we increased the variability and tested the inversion with a CV on Vs of 60% (see Figure 6.11). In this case, the velocity of the second soil layer was close to that of the initial soil layer (reduced, compare to the first inversion), although the velocity of the third soil layer again reached the boundary limit, reducing the impedance contrast between the third and fourth soil layers to 1.0. This last observation is in agreement with the geological log of the site (see Figure 6.1), which indicates the same material for the two deepest soil layers. The two inversions conducted by varying the same parameter but with different CVs led to different families of Vs profiles, which indicates that the parameter boundaries have a strong influence on the inversion results. Adding variability to the Q profile (Figure 6.12) did not influence the results of the inversion,

Table 6.2: *Parameters of the inversion on linear data performed at the four selected KiK-net sites FKSH12, FKSH14, KSRH10 and MYGH04*

Site	Frequency band (Hz)		Figure number	Parameters inverted with associated CV
FKSH12	[1-20]	Inversion 1	Figure 6.10	Vs (40%) Q (5%) Th (5%)
		Inversion 2	Figure 6.11	Vs (60%) Q (5%) Th (5%)
		Inversion 3	Figure 6.12	Vs (60%) Q (60%) Th (5%)
		Inversion 4	Figure 6.13	Vs (60%) Q (60%) Th (60%)
FKSH14	[0.2-10]	Inversion 1	Figure 6.14	Vs (60%) Q (5%) Th (60%)
		Inversion 2	Figure 6.15	Vs (60%) Q (60%) Th (60%)
KSRH10	[0.5-8]	Inversion 1	Figure 6.16	Vs (60%) Q (5%) Th (5%)
		Inversion 2	-	Vs (60%) Q (60%) Th (5%)
		Inversion 3	-	Vs (60%) Q (5%) Th (60%)
		Inversion 4	Figure 6.17	Vs (30%) Q (60%) Th (5%)
MYGH04	[1-20]	Inversion 1	Figure 6.18	Vs (60%) Q (60%) Th (5%)
	[10-20]	Inversion 2	Figure 6.19	Vs (60%) Q (60%) Th (5%)

because the Q profile does not have a great influence of the site response amplitude as shown by the sensitivity analysis. The Q profile resulting from this inversion is not well constrained and may be more considered as a result of numerical effects to minimise the cost function. In contrast, adding variability to the thickness led to a family of profiles with thinner second and third soil layers, with the Vs of the first two soil layers close to their initial values (see Figure 6.13).

When looking at the comparison between the numerical site responses both outcrop and borehole with the BFSR_{lin}^{95} and the RF_{lin}^{95} (Figure 6.3 (b)), we can observe that the first peak in the BFSR_{num} moves back to the low frequency band compared to the first peak of the BFSR_{lin}^{95}. The first peak of the the BFSR_{lin}^{95} is equal to the first peak of the OFSR_{num}. The first peak in the BFSR_{num} is a combination between the first resonance frequency of the site and a pseudo-resonance peak, whereas, the first peak of the BFSR_{lin}^{95} is not affected by the pseudo-resonance. Hence, the difference between the numerical and empirical site response is not attributed to an error on the Vs profile but to a pseudo-resonance that is not observed in the empirical site response. Therefore, the Vs profiles family that is solution of the inversion is not reliable. Consequently, the inversion method cannot work in the linear range, nor in the non-linear one.

We can note that choosing the mean of the BFSR_{lin} may smooth some frequency peaks and may be a limitation of the present work. However, at this site the discrepancy between the

BFSR$_{lin}$ and BFSR$_{num}$ is rather a matter of the amplitude of pseudo-resonance peaks being stronger in numerical simulations than in the observations.

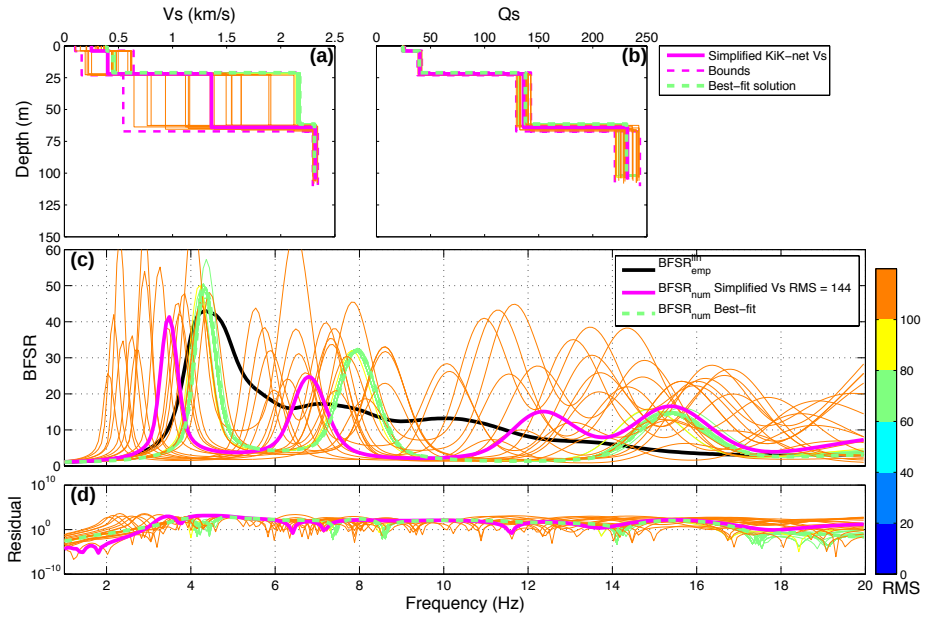


Figure 6.10: *Inversion of the BFSR$_{lin}$ at the FKSH12 site from 1 Hz to 20 Hz. (a) Inverted $V s$ profiles, showing the simplified KiK-net $V s$ profile (purple line; here it is the same as the initial KiK-net $V s$ profile), the boundary of the inversion (purple dotted lines; here we chose a CV of 60%), and the examples of the $V s$ profiles tested by the inversion (thin lines) withcolours according to the values of the associated cost function (here called RMS). The best-fit family of the $V s$ profiles are also shown (green line; thecolour depends on the RMS). (b) As for (a), for the quality factor. (c) The BFSR from the soil column described in (a) and (b), compared to the empirical BFSR$_{lin}$ (black curve). (d) The residuals of the BFSR compared to BFSR$_{lin}$.*

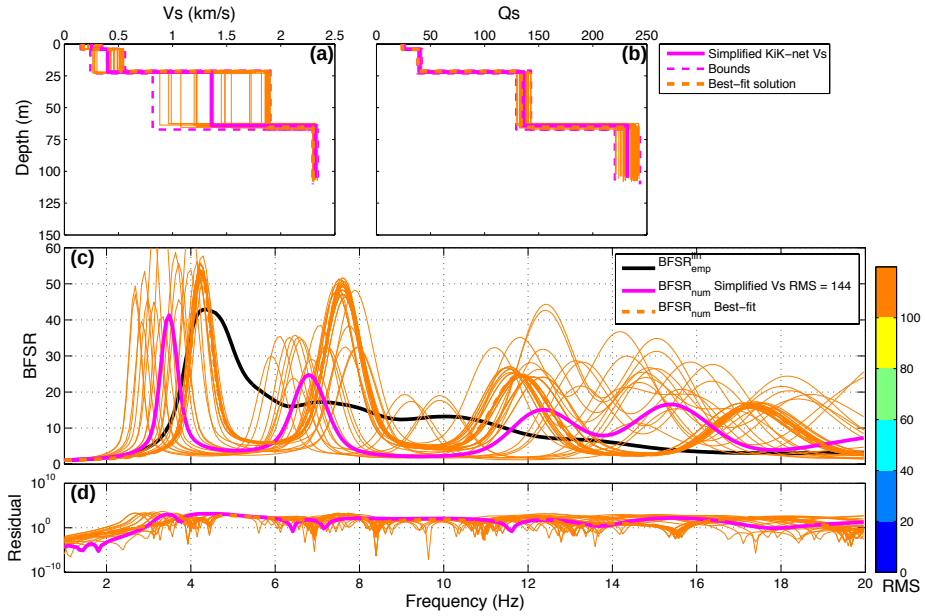


Figure 6.11: *As for Figure 6.10, with a V s CV of 40%.*

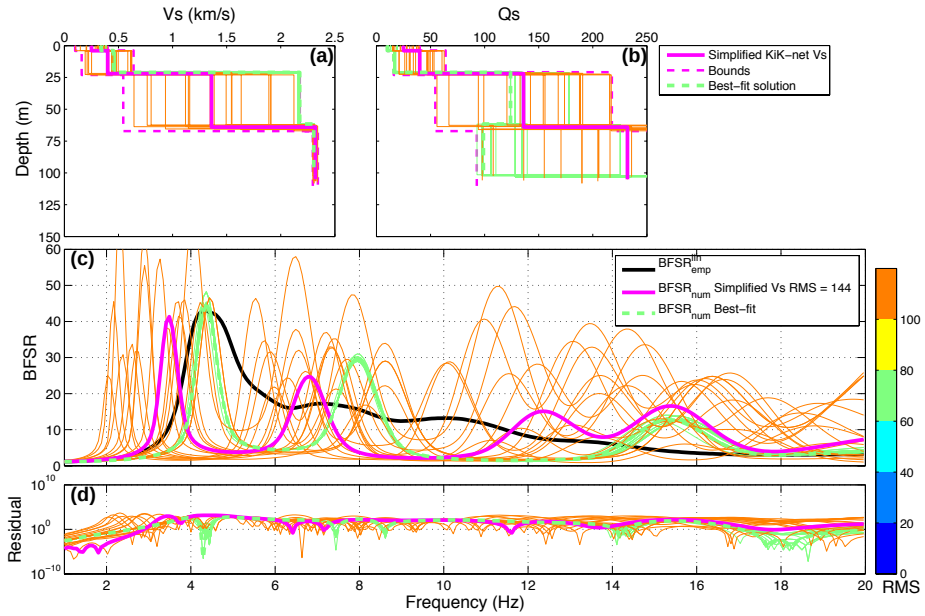


Figure 6.12: *As for Figure 6.10, with V s and Q CVs of 60%.*

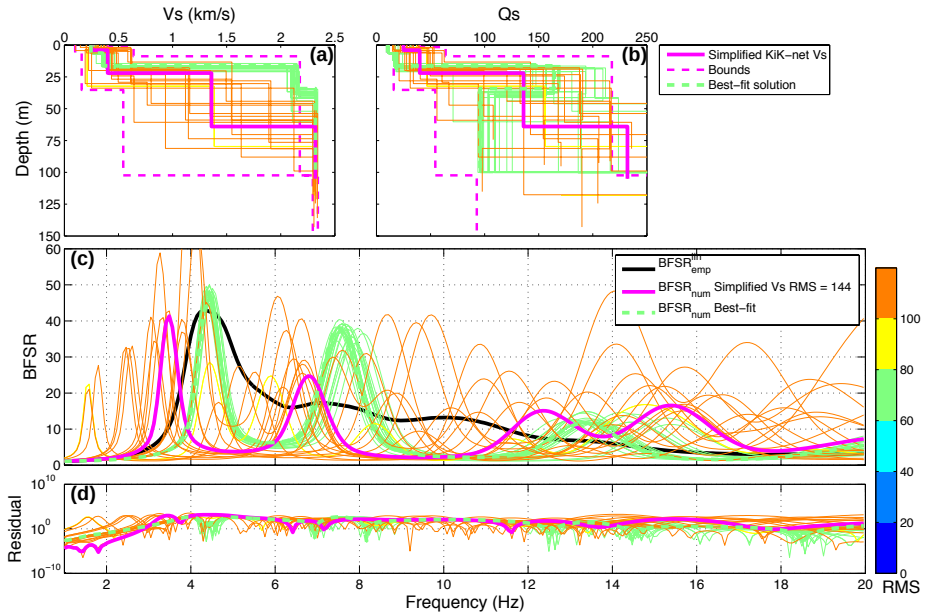


Figure 6.13: *As for Figure 6.10, with Vs, Q and thickness of the soil layers CVs of 40%.*

FKSH14: Two inversions were tested here. In the first inversion, we tested the variability of Vs and the thickness of the soil layers with a CV of 60% for both of the parameters (see Figure 6.14); the frequency peaks and forms were remarkably well reproduced. Compared to the Vs profile from the KiK-net database, the thickness of the third soil layer (the second soil layer for the simplified model) is increased, and the thickness of the fourth soil layer is decreased. Except for the fourth soil layer (the third soil layer for the simplified profile), for which the Vs is reduced, the Vs are in good agreement. We compared these data with the geological logging of this station, and we found that the depth proposed by the best-fit profile matches the depth of the geological logging (see Figure 6.1). Hence, the family of profiles proposed by the inversions appeared to be very coherent with the in-situ measurements. In the second inversion, we allowed the Q parameter to vary, and we found that reducing the Q made the amplitude match to the empirical data even better; see Figure 6.15. However, as shown in the sensitivity analysis, the Q as not a large influence on the site response amplitude and therefore may not be well constrained. Compared to the previous site, the numerical simulation and observations site response are in agreement and the frequency peaks in the BFSR$_{lin}$ are clear even at high frequencies, which give even more information. The inversion succeeded in retrieving the elastic properties of the soil and proposed an adapted soil profile that appears to be in accordance with geological information. For the inversion of the non-linear data, the

best-fit linear model found by the inversion will be used as the reference soil model (thickness, Vs and Q boundaries).

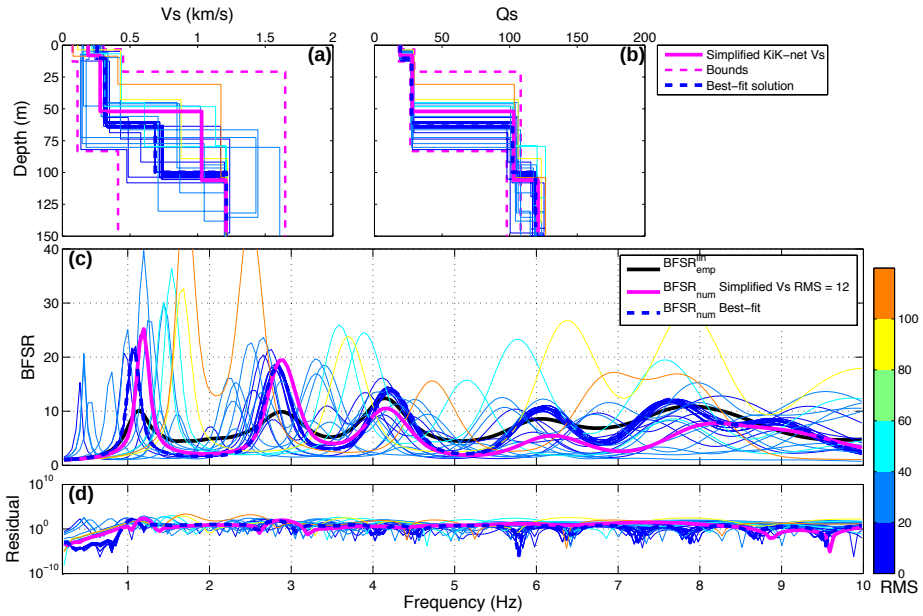


Figure 6.14: *Inversion of the $BFSR_{lin}$ at the FKSH14 site from 0.2 Hz to 10 Hz. The Figure layout is similar to that of Figure 6.10. Here, the Vs and thickness of the soil layers CVs are 60%.*

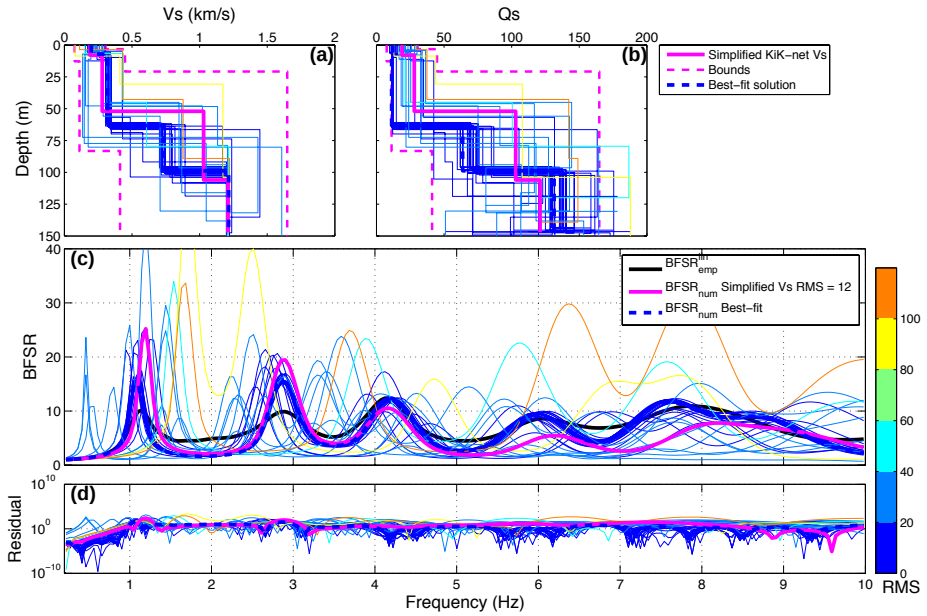


Figure 6.15: *Inversion of the $BFSR_{lin}$ at the FKSH14 site from 0.2 Hz to 10 Hz. The Figure layout is similar to that of Figure 6.10. Here, the Vs, Q and thickness of the soil layers CVs are 60%.*

KSRH10: We tested four inversions here. In the first inversion, only the Vs was variable, with a CV of 60% (see Figure 6.16). The best-fit of the inversion method consisted of a family of Vs profiles where the site response indicated a low-frequency peak that was absent from the observations. When allowing the Q to be variable as well, or the thickness, the results were similar, and these are not shown here. To force the inversion to avoid this unsatisfactory family of Vs profiles, we reduced the Vs CV from 60% to 30% and kept the Q variable as well. The velocity profiles were closer to the initial Vs profile and the frequency peaks matched the observations. The cost function of this soil profiles family are greater than the cost function of the unrealistic soil profile previously found. Indeed, the associated amplitudes of the first two peaks are much larger than the observed amplitudes (see Figure 6.17). For the inversion of the non-linear data, this site is not relevant, as careful attention is needed for the definition of the Vs boundaries, which can led to unrealistic soil models.

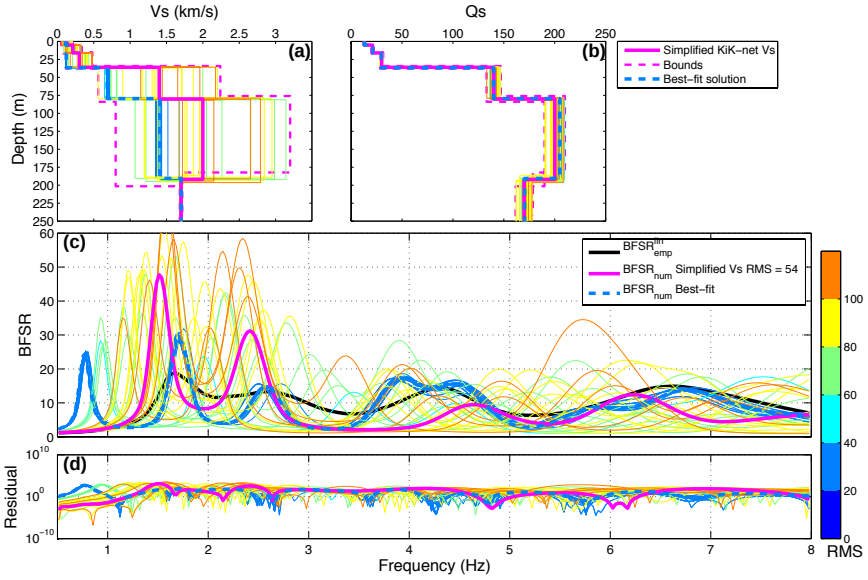


Figure 6.16: *Inversion of the $BFSR_{lin}$ at the KSRH10 site from 0.5 Hz to 8 Hz. The Figure layout is similar to that of Figure 6.10. Here, the CV of the Vs of the soil layers is 60%.*

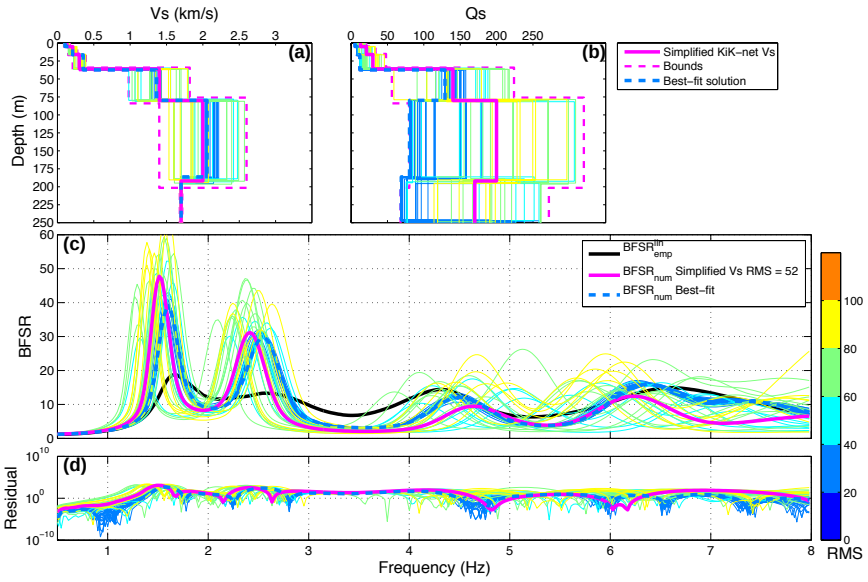


Figure 6.17: *Inversion of the $BFSR_{lin}$ at the KSRH10 site from 0.2 Hz to 10 Hz. The Figure layout is similar to that of Figure 6.10. Here, the CV of the Vs is 30%, and the CV of Q is 60%.*

MYGH04: We tested two inversions here. First, we inverted the transfer function from 1 Hz to 20 Hz. We found that the first peak amplitude in the numerical evaluation was much higher than the observed peak amplitude. Reduction of the quality factor of the last soil layer did not improve this discrepancy (see Figure 6.18). Thus, the following inversions were performed on the transfer function from 10 Hz to 20 Hz. The large peak at around 15 Hz is more spread for the observations compared to the results of the inversions. This spread peak observed in the empirical BFSR might be caused by two merged peaks, as indicated by the numerical transfer function from the initial Vs profile. Nevertheless, the inversion method cannot reproduce this behaviour (see Figure 6.19).

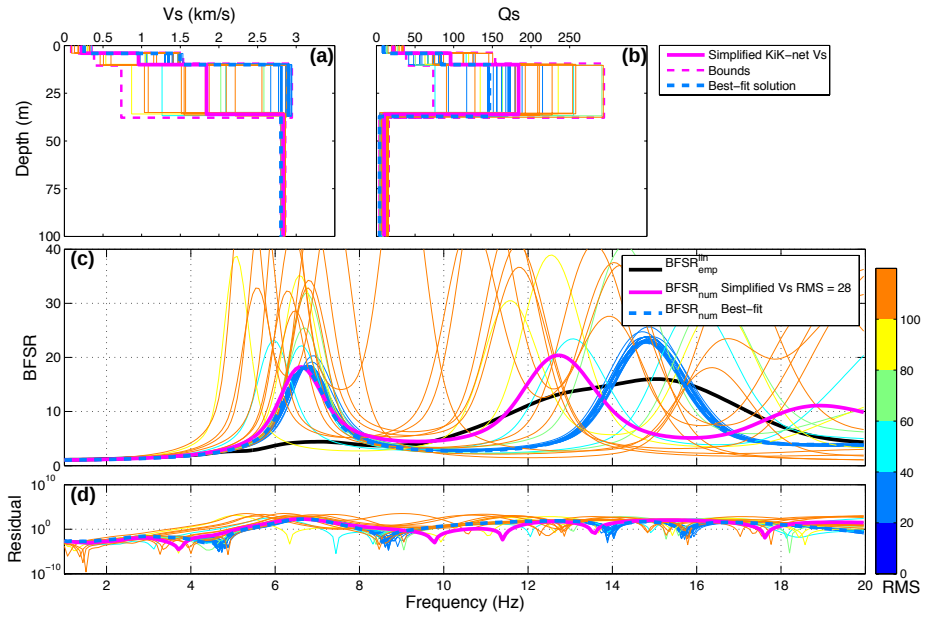


Figure 6.18: *Inversion of the $BFSR_{lin}$ at the MYGH04 site from 1 Hz to 20 Hz. The Figure layout is similar to that of Figure 6.10. Here, the CVs of the Vs and Q of the soil layers are 60%.*

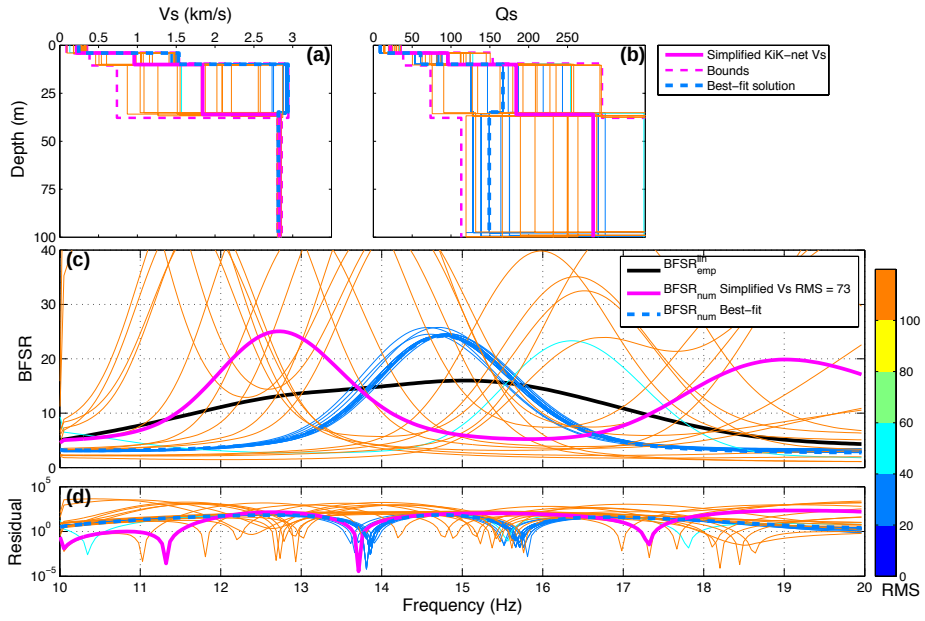


Figure 6.19: *Inversion of the $BFSR_{lin}$ at the MYGH04 site from 10 Hz to 20 Hz. The Figure layout is similar to that of Figure 6.10. Here, the $V s$ and Q $CV s$ are 60%.*

The inversions of the linear transfer functions, for which we had very good knowledge of the a-priori results, were applied to improve the velocity profiles of the observations. The results indicated that even with very good a-priori knowledge, the inversions appeared difficult to obtain a reliable solution. The principal difficulties encounter per site are the followings:

- FKSH12: The first peak in the OFSR fit the one of the BFSR_{lin}: Whereas, the first peak in the BFSR_{num} is lower (the second peak higher).
- FKSH14: The numerical simulation fits very well not only the first peak but also the higher modes.
- KSRH10: The best fit solution propose unsatisfactory velocity profiles.
- MYGH04: The first peak amplitude in the BFSR_{num} is much higher than the one observed in the BFSR_{lin}.

These difficulties might have come from the following observations:

- The numerical simulations did not reproduce well the pseudo-resonance peak amplitudes. These peak amplitudes are usually much higher in the numerical BFSR than in the empirical

BFSR. We tried to improve this discrepancy by adding a frequency-dependent attenuation. Nevertheless, the amplitude deviation is still important.

- The inversion method needs to have a-priori knowledge of the soil column. Although, as shown here, this is not a sufficient condition.

- There is a trade-off between the thickness and the Vs of the sediment layers to adjust the observed BFSR. Fäh et al. (2001; 2003) among others, suggest to apply inversion at sites where the sediment thickness is known from additional geotechnic surveys which is the case in the KiK-net sites. Consequently, we should only inverse one of this two parameters.

- The selection of the site was performed by a comparison of the fundamental resonance frequencies of the numerical and empirical BFSRs. The correspondence of the first peak appears to be insufficient to decide if the site has a 1-D configuration. Therefore, careful attention is needed to find sites such that the numerical and empirical BFSRs are in agreement not only on the first peak but on a larger frequency bandwidth.

To improve the inversion method, and especially to answer the first above-mentioned observation, one way would be to modify the cost functions, working mostly on the frequency peaks and associated amplitudes, as proposed by De Martin (2011). At the same time, It appears necessary to select sites with "true" 1-D site configuration and with observed site responses characterised by clear frequency peaks with high amplitude, such as at the FKSH14 site.

For the inversions of non-linear data, we show the results of the inversion performed for the FKSH14 site, for which the linear data inversion was the best.

6.4 INVERSION OF NON-LINEAR DATA

At the FKSH14 site, two earthquakes were recorded where the PGA at the downhole sensor was higher than 50 cm/s^2. Also, a third recording at this site induced a PNL > 10% with a PGA > 30 cm/s^2. The characteristics of these three earthquakes are summarised in Table 6.3.

Figure 6.20 illustrates the modification of the frequency peaks of the empirical BFSRs and the associated amplitudes compared to the linear evaluation in percentage. Compared to the linear site response, the first peak is shifted to the low frequency range by 10%, 20% and 40%, for earthquakes 2, 3 and 1 respectively and the fourth peak is shifted by 60% for the earthquake 1. The other peaks are shifted by 10%. The associated amplitudes, on the other hand, are greatly affected and the modification increases with frequency. The first peak amplitude is modified by 30%, 18% and 5% for earthquakes 3, 2 and 1, respectively, whereas the fourth peak (that can be identified in all of the BFSR$_{lin}$) suffers more than 120% modification for the the earthquakes 1 and 3 (not

Table 6.3: *Characteristics of the three strongest earthquakes recorded at the FKSH14 site*

EQ	Name EQ KiK-net	M_{JMA}	D_{epi}	Depth	PGA_{depth}	PGA_{surf}
			Km	Km	cm/s^2	cm/s^2
1	FKSH141103111446	9	202	24	121	372
2	FKSH141104111716	7.1	28	10	55	128
3	FKSH141104121407	6.3	24	10	36	110

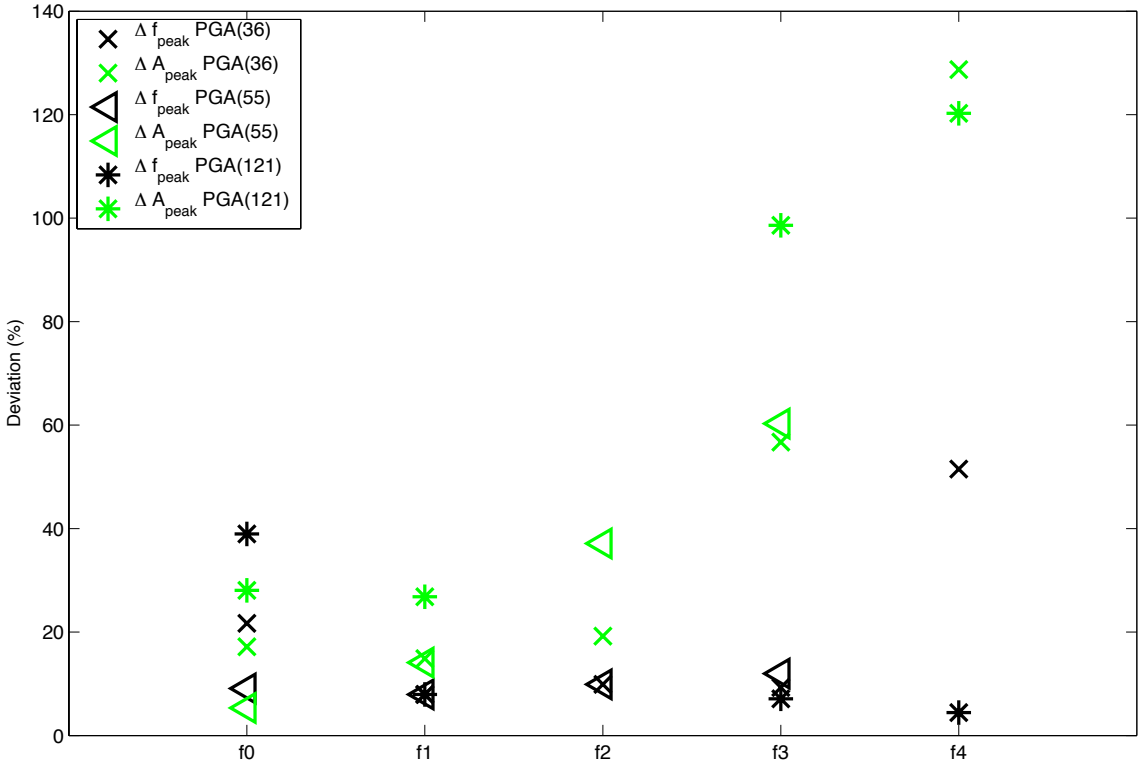


Figure 6.20: *Modification of the frequency peaks (black) and associated amplitudes (green) between the mean linear $BFSR_{emp}$ and the non-linear BFSR computed for the three largest earthquakes recorded at the FKSH14 site.*

identified in the BFSR computed using the recording of the earthquake 2).

First, we tested the inversion of the non-linear BFSR from 0.2 Hz to 10 Hz (figure not whown here), and here we found that only the first three peaks were well constrained. According to the sensitivity analyses (see Figure 6.7), the velocities of the second, third and fourth soil layers can be constrained. Considering that the first soil layer is merged with the second soil layer in the simplified model, all of the Vs soil layers can be solved. However, only the quality factor of the third soil layer (the second soil layer of the simplified model) has a large influence on the result, and consequently it can be well constrained.

Thus, the following inversions were tested on the non-linear BFSR from 0.2 Hz to 5 Hz containing the first three peaks. These are illustrated in Figures 6.21, 6.22 and 6.23 for the earthquakes 1, 2 and 3 respectively. For the three earthquakes, the inversion method gave equivalent linear soil columns, with Q decreased all along the depth, and more so in the first three soil layers (as two soil layers in the simplified Vs profile). For earthquake 1, the velocities of the equivalent linear soil column were decreased when compared to the initial Vs profile. As indicated from the data given in Table 6.4, the velocities of the first two soil layers, Vs_1 and Vs_2 (as Vs'_1 of the first soil layer in the simplified model), were decreased by 150%, the third soil layer (as the second soil layer of the simplified model) by 31%, and the fourth soil layer (as the third soil layer of the simplified model) by 23%. For earthquake 2, Vs_1 and Vs_2 (Vs_1') were decreased by 43%, Vs_3 (Vs_2') by 13%, and the last soil layer velocities were not modified. For earthquake 3, Vs_1 and Vs_2 (Vs_1') were decreased by 16%, Vs_3 (Vs_2') by 20%, and again the last soil layer velocities were not modified.

Table 6.4: *Velocity in the equivalent linear soil column (inversion of the BFSR from earthquakes 1, 2 and 3, between 0.5 Hz and 5 Hz), compared to the linear soil column from the best-fit inversion of the BFSR$_{lin}$.*

KiK-net soil column	Simplified soil column	Linear	EQ_1	EQ_2	EQ_3
		m/s	m/s	m/s	m/s
PGA_{depth} (cm/s^2)		< 10	121	55	36
Vs_1	Vs_1'	198	26	127	169
Vs_2					
Vs_3	Vs_2'	319	232	281	260
Vs_4	Vs_3'	735	580	733	733
Vs_5	Vs_4'	1210	1210	1210	1210

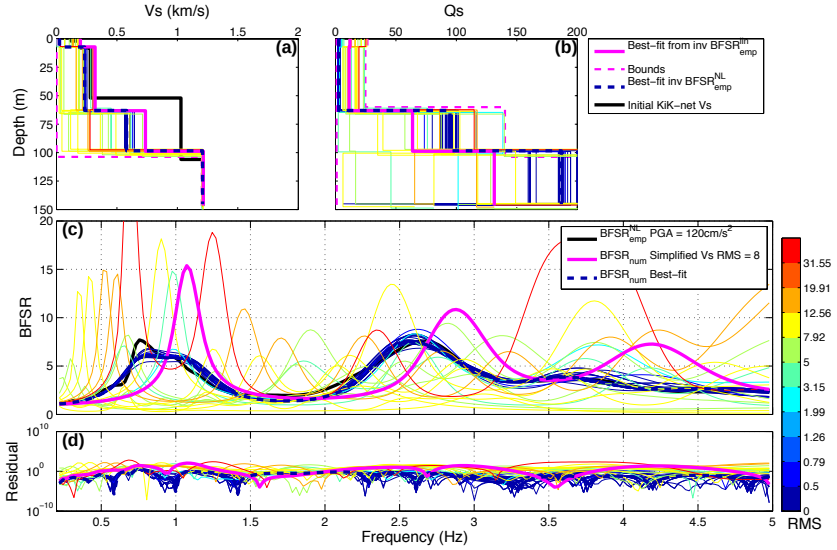


Figure 6.21: *Inversion of the non-linear BFSR at the FKSH14 site for earthquake 1 (PGA at the whole sensor of 121 cm/s^2), from 0.2 Hz to 5 Hz. (a) Inverted Vs profiles, showing the best-fit of the BFSR$_{lin}$ simplified KiK-net Vs profile (purple), the boundary of the inversion (purple dotted lines; here we chose a CV of 60%), and the examples of the Vs profiles tested by the inversion (thin lines) withcolours according to the values of the associated cost function (here called RMS). The initial Vs profile is also shown (black curve), along with the best-fit family of the Vs profiles (green line; the colour depends on the RMS). (b) As for (a), for the quality factor. (c) The BFSR from the soil column described in (a) and (b), compared to the empirical non-linear BFSR (black curve). (d) The residuals of the BFSR compared to BFSR$_{lin}$.*

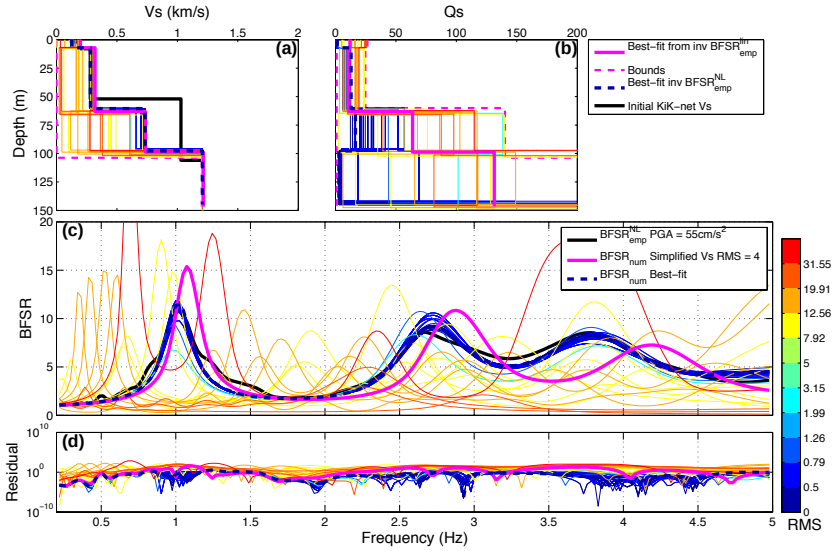


Figure 6.22: *Inversion of the non-linear BFSR at the FKSH14 site for earthquake 2 (PGA at the whole sensor of 55 cm/s^2), from 0.2 Hz to 5 Hz. The Figure layout is similar to that of Figure 6.21.*

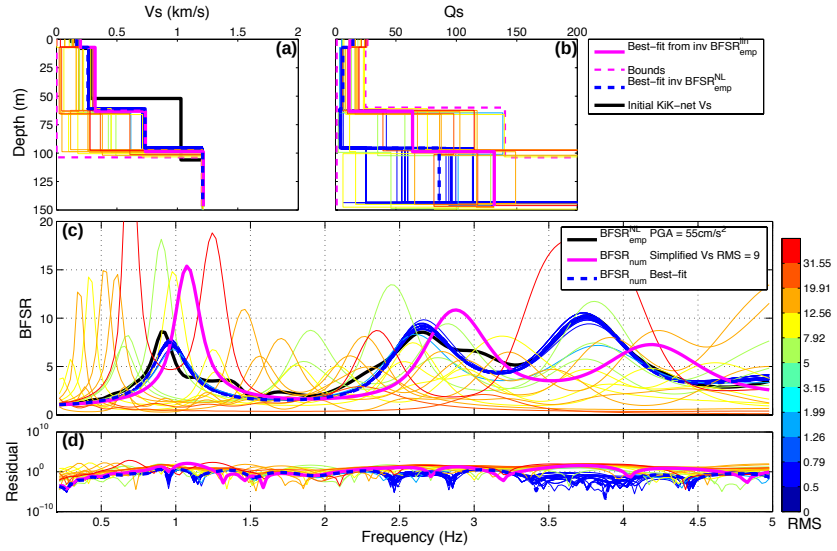


Figure 6.23: *Inversion of the non-linear BFSR at the FKSH14 site for earthquake 3 (PGA at the whole sensor of 36 cm/s^2), from 0.2 Hz to 5 Hz. The Figure layout is similar to that of Figure 6.21.*

6.5 DISCUSSION ON THE INVERSION OF NON-LINEAR DATA

6.5.1 Degradation of the shear modulus during the strongest earthquakes

We also calculated the global decrease of the shear modulus of the whole soil column depending on the level of strain between the bottom and the surface for each earthquake. We calculated the decrease of the shear modulus assuming that the shear modulus is link to the fundamental resonance frequency (f_0) of the site by the following formula:

$$G = (f_0 4H)^2 \rho, \tag{6.2}$$

where H, is the depth of the borehole sensor location, ρ a mean density of the soil column. We used the relation that linked, in the linear range and for a 1-D monolayer soil column, the fundamental resonance frequency to the shear wave velocity and thickness of the sediment layer ($f_0 = Vs/4H$) and the definition of the Vs, $Vs = \sqrt{G/\rho}$. Assuming equivalent linear soil properties for stronger incident seismic motions, the ratio of the shear modulus during a given earthquake by the maximal shear modulus (that is the shear modulus for weak motion) can be written as:

$$\frac{G}{G_{max}} = \frac{f_{0,ev}^2}{f_{0,lin}^2}, \tag{6.3}$$

where, $f_{0,ev}$ is the fundamental resonance frequency picked on a the site response computed with the recording of a given earthquake and, $f_{0,lin}$ is the fundamental resonance frequency picked on the average site response using all the recordings with PGA at the downhole station lower than 10 cm/s^2. In the appendix C, we showed the difficulty in finding the fundamental resonance frequency. It appears even more difficult to pick $f_{0,eq}$ for each earthquake. Therefore, in the present study, to find $f_{0,eq}$, we choose to use a parameter that was defined in the chapter 5 to characterize the shift of the frequency in between the linear site response characterisation (BFSR$_{lin}$) and the BFSR computed for a given earthquake, that is Sh$_{ev}$ such that:

$$f_{0,ev} \approx f_{0,lin} - Sh_{ev}. \tag{6.4}$$

The decrease of the shear modulus is then equivalent to:

$$\frac{G}{G_{max}} \approx \frac{(f_{0,lin} - Sh_{ev})^2}{f_{0,lin}^2}. \tag{6.5}$$

The shear strain of the whole column has been calculated for each event as the ratio of the maximum of the deviation between the displacement time histories at the surface with the one at

the down-hole station divided by the depth of the down-hole station:

$$\gamma(\%) = \max \frac{D_s(t) - D_b(t)}{Z_{bor}}, \tag{6.6}$$

where, γ, is the shear strain of the whole soil column, $D_s(t)$ and $D_b(t)$ are the displacement time histories at the surface and at the down-hole station. The displacement time histories were obtained by integrating two times the acceleration time history, applying to the input time history, before each integration, a bandpass Butterworth filter with cut-off frequencies [0.1-25] Hz, that was applied in both directions (forward and reverse) to avoid phase difference.

At FKSH14, the results of these calculations, illustrated in Figure 6.24, show that the decrease of the shear modulus occur at low shear strain (the ratio reach values lower than 0.6 between 0.001 and 0.01 %). For the earthquake 2, the decrease of the shear modulus is 70% whereas, it is close to 50 % for the two other earthquakes. These observations are in agreement with the decrease of the velocities found by the inversion, that are larger for the earthquakes 1 and 3 compared to the earthquake 2.

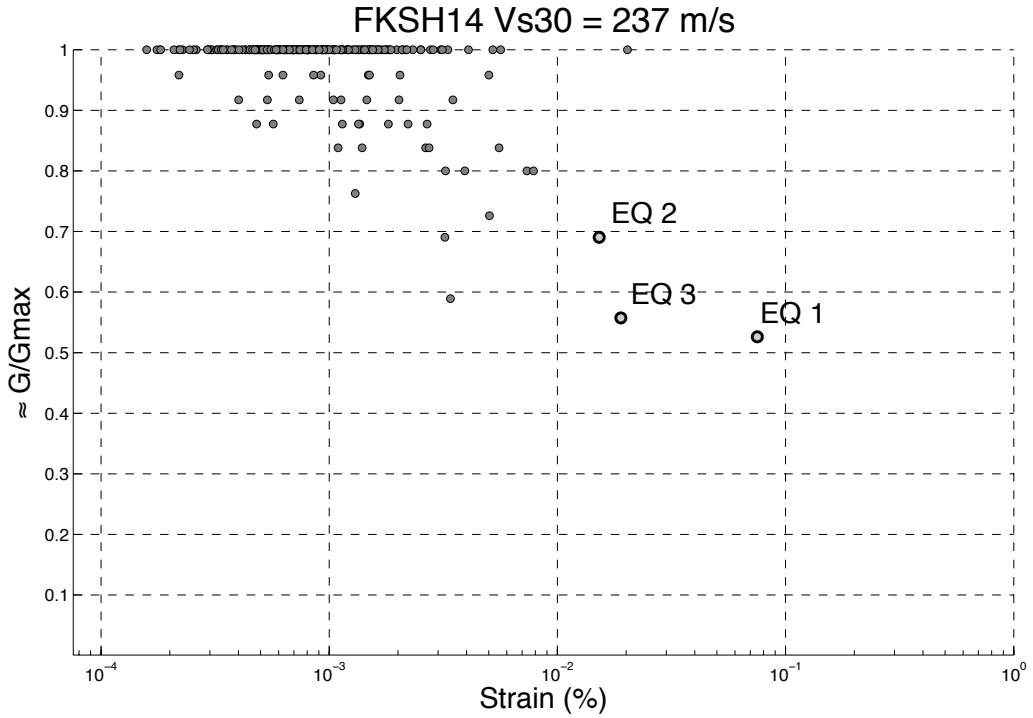


Figure 6.24: *Decrease of a equivalent shear modulus with the shear strain of the whole soil column for each earthquake recorded at the station FKSH14.*

6.5.2 Comparison of the inversion results with the previous chapter's conclusions

The inversion of the empirical BFSRs deduced from these earthquakes indicated that the velocity of the first three soil layers (as the first two soil layers of the simplified Vs profile) were systematically decreased compared to the weak motion. For the strongest earthquake, the velocity of the fourth soil layer, Vs_4, was also decreased. It is difficult here to tell if this result is an artifact of the inversion to improve the fitting, or if the apparent velocity of the fourth layer was decreased during the strongest earthquake. In the second case, it means that the depth from which the non-linear soil behavior occur and affect the site response depend not only on the soil column but also on the incident motion intensity.

At FKSH14, the first resonance frequency peak (f_0) is affected by the non-linear soil behavior. This observation is shown in Figure 6.3 and because fNL (the frequency of non linearity, from which we observed a decrease in amplitude in the non-linear site response compare to the linear counterpart, see chapter 4 page 101 for more details) is below f_0. Besides, the sensitivity analyses showed that this first peak is associated to the impedance contrast between the soil layer 3 and the soil layer 4. Therefore, it was expected that if f_0 was affected then the non-linear soil behavior would mostly took place in the first three soil layers.

The results of the previous chapter that were based only on borehole transfer functions comparison indicated that the non-linear soil behaviour mostly occur in the superficial soil layers. This conclusion is in good agreement with the results of the inversion.

We can see that clear results can be obtained only when the linear simulations are in very good agreement with the BFSR_{lin}. Even if the KiK-net database contains more then 688 sites, the number of sites that satisfy this requirement is very limited. However, the results obtain from the inversion on one site only could be extended to others KiK-net sites. Inversion may not be necessary to assess the depth from which non-linear soil behaviour do not influence significantly the site response. Using the comparison of the transfer functions (linear and non-linear) and the value of fNL compared to f_0 or f_{pred}, with a sensitivity analysis of these site response parameter to the soil column, may be enough. For instance, if f_0 is affected, then we can expect that the soil non-linear behavior occur in the layers above the impedance contrast associated to f_0. On the contrary if fNL is above f_0 and below f_{pred} then, the non-linear soil behavior may occur in soil layers above the impedance contrast associated to f_{pred}, soil layers located in between the impedance contrast linked to f_0 and f_{pred} should not have large non-linear soil behaviour.

6.5.3 Comparison of the inversion results with other studies

The results of the present studies are in agreement with what was found by Aguirre and Irikura (1997). They inverted the Vs profile using a genetic algorithm technique to determine the velocity structure before, during, and after the mainshock (Hyogo-ken Nanbu Earthquake) in port island ver-

tical array. Similarly to the present study, they found that the Vs of the superficial layers was mostly modified (the second layer between 5 m to 16 m depth) after the mainshock by 20% lower. At the same time, in area close to the fault plane, Pavlenko and Irikura (2006) observed noticeable manifestations of soil nonlinearity in the upper 15-25 m of the soil profiles. The authors suggested that the behavior of soft subsurface soils is described by other, more "nonlinear" stress-strain relations. In addition, De Martin et al. (2010) used genetic algorithm to inverse the Vs profile and compared it to the one inverted using weak motion such as in the present study. They found that the reduction of the Vs occurred mostly around 35 m depth. They also found that the increase of damping factors was not always consistent with the decrease of the Vs. Finally, Assimaki and Li (2012), showed that the Vs_{30}, which is a superficial soil parameter, was a relevant parameter to analyze non-linear site responses, but also needed to be completed by impedance contrast information and seismic sources parameters.

6.6 CONCLUSION

This chapter has deals with the inversion of the empirical borehole transfer function. The main objectives of the inversion were to adjust the elastic soil properties (the Vs and Q profiles) from the weak-motion data and to find the equivalent linear soil columns from the inversions of the strong-motion data. Using the results of the inversion, the final aim of this study was to find the depths where the non-linear soil behaviour mostly takes place. We selected four KiK-net sites with different soil columns to apply this methodology. The empirical and numerical site responses of the selected sites were close to one another, and the strong ground motion recorded at these site showed clear non-linear soil behaviour. First, we analysed the global sensitivity of each of the soil columns, to find the soil parameters that have a large influence on the computed borehole site response, and consequently that can be well constrained during the inversion. These analyses appeared to be a very important step to be realized before the inversion of non-linear data, to avoid over-interpretation of the inversion results, and consequently the obtaining of misleading results. We found that the very superficial soil layers have a small influence on the frequency peaks and their associated amplitudes. Consequently, the inversion of the properties of these layers were not well constrained. This shows that when dealing with inversion, the resolution of the data must be adapted to the information that is inverted. To analyse shallow soil layers, there is the need to use vertical arrays with intermediate sensors, to increase the resolution. Then, we performed the inversions on the $BFSR_{lin}$ and found, that even with very good a-priori knowledge of the soil column, the inversions were difficult to interpret. This difficulty can be partially explained: although the numerical simulations had frequency-dependent attenuation, they did not reproduce the amplitude of the pseudo-resonance peaks well; and although the selected sites had their fundamental resonance frequencies numerically and empirically close to each other, they might not have a simple 1-D site

configuration. For the FKSH14 site, the numerical simulations reproduced the BFSR$_{lin}$ very well. Thus, we performed the inversion of the non-linear BFSR$_{emp}$ at this site, and we found that for the three earthquakes that induced large modifications in the linear and the non-linear site responses, it was the soil properties of the first three layers of soil that were mostly affected. Nevertheless, during the strongest earthquake, the best-fit equivalent linear soil column from the inversion method indicated that the fourth soil layer was also affected. This observation confirms that the non-linear soil behaviour mostly takes place in the subsurface soil layers, although the depth where the non-linear soil behaviour occurs and influences the site response also depends on the amplitude of the incident motion. These data are limited to one site for now, and they cannot be generalized, although they are in a good agreement with previous studies and with the conclusions of the previous chapters, where the non-linear soil behaviour was analysed statistically for 54 sites.

Conclusions and perspectives

CONCLUSIONS

The scope of the work done on this thesis was to enhance the site-effect assessment by understanding the sources of the site-response variabilities and proposing solutions to reduce it. We divided the site-response variability analyses into between-site and between-event variabilities, such as in GMPEs. This study was mainly based on the analyses of a large dataset of sites and earthquake recordings. We chose the KiK-net database in Japan, considering its large variability of sites, its large number of earthquake recordings, and its sensor configurations in vertical arrays. For all of the KiK-net sites where a Vs profile was available (668), we calculated the empirical borehole site responses by computing the spectral ratio between the surface and the down-hole recordings and the H/V earthquake spectral ratios, for all of the selected earthquake recordings (more than 46,000). We characterised the linear site responses by calculating the linear borehole site responses, the so-called BFSR_{lin}^{95} (mean and 95% confidence limit), and the linear H/V earthquake spectral ratios at the surface, HV_{lin}^{95}. We used all of the recordings for which the PGA at the downhole sensor were < 10 cm/s^2. Based on this definition, the following hypotheses formed the basis of the present study:

- (1) BFSR_{lin}^{95} is computed with recordings with a PGA at the downhole sensor < 10 cm/s^2, whereby the non-linear effects do not affect the borehole site response. This evaluation can therefore be compared to numerical linear evaluation.

- (2) Considering that BFSR_{lin}^{95} is computed using a large number of earthquakes, all of the site-response variability caused by complex site configurations associated to various seismic-source locations was already contained in BFSR_{lin}^{95}. Therefore, the deviation between BFSR_{lin}^{95} and the BFSR computed with strong events would be mostly caused by non-linear soil behaviour.

We defined the soil, the linear site response, and the incident motion parameters that are relevant for the site-effect assessment and site classification, and to characterise the intensity of the incident motion. The analyses of the parameter distributions confirmed that the database was pertinent for the analysis of site effects and representative of large site configurations. Also, for the analyses of the non-linear soil behaviour, we found that more than 50 KiK-net sites have recorded at least two earthquakes with PGA at the downhole sensor depth > 50 cm/s^2.

Between-site site-response variability

Finding the classification (i.e. the pertinent site parameters) that will reduce the site-response variability in each soil class is a high stake for seismic hazard assessment. We proposed additional parameters to Vs_{30} (usually used in regulation codes or GMPEs) to reduce the between-site Vs profiles and site-response variabilities. We selected 351 sites from the KiK-net database, such that the empirical linear site responses were close to the numerical 1-D linear evaluation. This selection

was based on comparisons of the fundamental resonance frequencies. We analysed the variabilities of the Vs profiles and the BFSR_{lin} (the mean linear empirical borehole Fourier spectral ratio) with Vs_{30}, and used additional parameters that reflect the Vs evolution with depth. We used the gradient of the Vs profile calculated using only the first 30 m of the Vs profile (B_{30}), to be consistent with the depth of the investigation required to obtain Vs_{30}. We also used the maximal depth common to all of the selected KiK-net site profiles; i.e., 100 m (B_{100}). In addition, we used a third parameter that was independent of the Vs profile, the fundamental resonance frequency (f_0), which was calculated using the earthquake recordings.

We showed that the distributions of the Vs profiles at depth have strong influences on the linear site responses. Indeed, we showed that B_{100} is a relevant additional parameter to Vs_{30}, both for site classification and for site-effect assessment. The Vs profile is, however, difficult to obtain precisely down to a large depth (100 m depth), and would be expensive. Thus, other site characterisations are of interest. B_{30} requires no more effort than Vs_{30} to be obtained, and turned out to be a relevant complementary parameter to Vs_{30} for site characterisation and site-effect assessment. We also underlined the similar potential for f_0, which can be obtained from low-cost, non-invasive methods, and is independent of Vs_{30}.

Between-event site-response variability

Small earthquakes that induce weak motion are occuring frequently. Therefore, they need to be taken into account for seismic hazard assessment. Strong earthquakes, although rare, have demonstrated their destructiveness recently, as in L'Aquila in 2009, Haiti in 2010 and Tohoku in 2011, among others, which shows that such scenarios need to be considered as well. Observations of earthquake recordings during the Tohoku earthquake suggested that the soil behaved non-linearly during this strong earthquake. We analysed the between-event site-response variability considering only the effects of non-linear soil behaviour. When the soil behaves non-linearly, we expect the resonance frequency peaks to be shifted to the low frequency range, and the associated amplitude to be modified according two opposite phenomena. The non-linear soil behaviour makes the apparent Vs decrease in the soft soil layers, and thus the impedance contrast between the sediment and substratum layers should increase, as well as the amplification. On the other hand, the non-linear soil behaviour induces an increase in the attenuation, which induces a decrease in the amplification.

We try to understand the relationships between the parameters that characterise the effects of non-linear soil behaviour on site responses and the parameters that characterise either the site (Vs_{30}, B_{30}, f_0, f_{pred} and A_{pred}) or the incident ground motion (PGA). We first defined the parameters that describe the effects of the non-linear soil behaviour on the site responses for each event: a percentage of modification of the site-response curve compared to the linear evaluation (PNL_{ev}), and a shift of the curve to low frequency range (Sh_{ev}). We found that the PGA (at the downhole

sensor) is the intensity parameter that is the most correlated to the non-linear parameters for each event. Also, the results showed that regardless of the site, depending on a PGA threshold (even for moderate input motion: PGA 75 cm/s^2 at the downhole sensor), there is a large probability (>40%) that the site response behaves non-linearly.

In addition, for 54 sites of the KiK-net database that recorded at least two strong events (with a PGA value at depth >50 cm/s^2), we defined parameters that characterise the effects of non-linear soil behaviour on the site responses for each site: a PGA threshold (PGA_{th}), a percentage of modification between the linear and the non-linear site responses (PNL_{site}) for a PGA of 50 cm/s^2, a shift of the peak frequency for a PGA of 50 cm/s^2 (Sh_{site}), and a frequency from which we observed de-amplification between the non-linear site response and the linear site response (fNL). From observations of the linear to non-linear site-response ratio, we noted that the shift of the predominant peak frequency in the site response during strong motion implies that below fNL, the site response computed with strong events is likely to be amplified compared to the linear evaluations.

We used multivariate statistical analyses to define the relationships between the non-linear parameters and the soil and site response parameters available for all of the KiK-net sites. The results indicated that the non-linear soil behaviour occurs mostly in the superficial layers. The non-linear results for the soil obtained in the present study are not limited to the Tohoku earthquake, which has a very particular signature, as the whole dataset from 1996 was used. Also, the 54 sites selected to perform the multivariate analysis are spread across different districts, and therefore the results are not specific to a given area. We found that the H/V earthquake spectral ratio can give satisfactory results for the evaluation of the frequency fNL and the shift frequency (Sh$_{site}$), making these results extendable to other databases without downhole sensors.

Inversion of borehole site responses

Considering the results of the previous chapter, we used inversion techniques to confirm that the non-linear soil behaviour mostly occurs in the superficial layers, and to find the depths for which we can consider that this non-linear behaviour will not have a strong influence on the site responses. We selected four KiK-net sites with different soil columns to apply this methodology. The empirical and numerical site responses of the selected sites were close to one another, and the strong ground motion recorded at these sites showed clear non-linear soil behaviour.

First, we analysed the global sensitivity of each of the soil columns, to find the soil parameters that have a large influence on the computed borehole site response, and consequently that can be well constrained during the inversion. These analyses appeared to be a very important step

to consider before processing to the inversion of non-linear data, to avoid over-interpretation of the, and consequently to avoid the obtaining of misleading results. We proved that the very superficial and very thin soil layers have a small influence on the frequency peaks and their associated amplitudes. Consequently, the inversions of the properties of these layers were not constrained.

When performing the inversions on the $BFSR_{lin}$, we found that even with very good a-priori knowledge of the soil column, the inversion results were difficult to analyse. This difficulty can be partially explained: velocity profiles with equivalent Vs-to-thickness ratios give exactly the same borehole transfer function (shown in chapter 1). Although the numerical simulations had frequency-dependent attenuation, they did not reproduce the amplitude of the pseudo-resonance peaks well, and although the selected sites had their fundamental resonance frequencies numerically and empirically close to each other, they might not have a simple 1-D site configuration.

The inversion of the non-linear $BFSR_{emp}$ at the FKSH14 site confirmed that the non-linear soil behaviour mostly takes place in the subsurface soil layers. We also showed that the depths where the non-linear soil behaviour occurs and influences the site responses also depend on the amplitudes of the incident motion. These data are limited to one site for now, and they cannot be generalised, although they are in a good agreement with the conclusions of the previous chapters, where the non-linear soil behaviour was analysed statistically for 54 sites.

Considering the difficulty to interpret the inversion of borehole transfer function and the agreement of the conclusions using inversion or comparison of linear and non-linear BFSR, we may suggest that comparison of BFSR associated to sensitivity analysis of the soil column would be enough to evaluate the depths where soil non-linear behaviour mostly take place.

TECHNICAL PERSPECTIVES

1-D site selection

In the present work, we analysed the between-site site-response variability by focusing on sites for which the 1-D site configuration assumption was valid. We selected sites such that the fundamental resonance frequency of the empirical borehole transfer function was close to the numerical one. Thompson et al. (2012) defined a taxonomy for vertical arrays that depends on the between-event variability and the fit to 1-D simulations. Here, we did not take into consideration the between-event variability, as was proposed by Thompson et al. (2012). For over 100 sites, they found 69 with low between-site variabilities and 31 with high between-site variabilities, and when the between-event variability was large, the authors concluded that the 1-D assumptions were no longer valid.

Considering the following conclusions, it may be interesting to re-process the data for the analysis of the between-site variability using only the sites with matching fundamental resonance frequencies and with low between-event variability. For the selection of sites for the inversion, it could also be interesting to: (1) add the between-event variability as a selection parameter; and (2) fit not only the first significant peak, but also the second or third one(s), depending on the resolution required.

Improvement of ground-motion prediction

In terms of the assumptions of numerical simulation modelling, we have shown whatever the 1-D site configuration assumption is, the linear simulations do not reproduce well the pseudo-resonance peak amplitudes. In previous studies (e.g. Satoh et al., 2001; Thompson et al., 2009), the authors suggested that the discrepancy mainly comes from the scattering attenuation that is not taking into account in the Haskell-Thomson method. The authors proposed to use frequency-dependent attenuation to satisfy the deviation. Although we think that the physical basis of the explanations of the discrepancy between numerical simulation and observation with the scattering of down-going waves is not straight-foward, we implemented the frequency-dependent attenuation in the numerical code used. The discrepancy between the empirical and the numerical transfer functions remains high. This observation marks the limitations of the 1-D linear method, even when compared with the empirical site response computed with weak motion.

These methods are further limited when dealing with stronger deformation. We showed that the non-linear soil behaviour has a strong impact on the site responses, and therefore on the surface ground motion. Contrary to what is expected, we showed that the non-linear soil behaviour can lead to amplification, as compared to linear site-effect evaluations in a low frequency range. Therefore, to precisely evaluate ground motion at the surface, it is essential to take this phenomenon into consideration.

The quantification of soil non-linearity and our ability to predict non-linear soil behaviour still remains a challenging task. The benchmark conducted in Turkey Flat (Stewart et al. 2008) on 1-D non-linear wave propagation software was aimed at clarifying the implementation of non-linear analyses. The main conclusions related to the code use and the parameters, for which the recommendations highlighted drew attention to the way to specify the input motion and the damping type implemented in the code. According to the benchmark, the factors that mostly explain the bias between observations at the surface and simulations come from: (1) errors in the input parameters (mostly the shear-wave velocity profile); (2) configuration of the site that is not totally 1-D; and (3) inaccuracies in the non-linear models that cannot accurately and simultaneously reproduce the shear-modulus degradation and the damping-ratio curves.

Performing a new benchmark on non-linear numerical simulations appears to be an interesting perspective. Indeed, using the results and conclusions of the present study, we can find KiK-net sites that satisfy the first two aforementioned drawbacks, and that have recorded both weak and strong

motions. The third drawback can be removed by in-situ measurements on the selected KiK-net sites.

We have presented five constitutive models. The more sophisticated the models are, the more precisely they will reproduce the experimental data. However, the increasing number of parameters makes the models more difficult to use. The first step of the benchmark would be an overview of the constitutive models and the way that they are implemented and calibrated on the tested numerical models. This overview would be followed by another one on the parameters that should be measured, with in-situ or laboratory measurements. The second step, before the validation with earthquake data, would be a standardisation between the codes and between the operators. The originality of this benchmark would lie in the coupled use of simulation, laboratory measurements, and strong-motion records. The main results of this benchmark would be:

- finding the influence of the operator on the results when calibrating the model to the experimental data;
- finding the influence of the numerical method on the results;
- recommending the type of laboratory measurements to be made;
- proposing a comprehensive methodology for taking account of the non-linearity.

Missing data for site-effect assessment

To go further in the evaluation of numerical modelling, we emphasised in the previous paragraph the need for detailed site characterisation. In the Japanese context, where a large number of strong earthquakes have been recorded, non-linear soil properties have not been characterised, or only using earthquake data. Such measurements can be used to confirm the conclusions of this study and to test the numerical simulation tools. Considering that the seismotectonic context of Japan is different from that of European, it appears essential at the same time to increase the number of seismological observations of strong motion in Europe.

Performing new measurements is essential, although time consuming and expensive. Therefore, gathering existing data appears to be a complementary task. The results obtained here in a large sample of sites can be tested in a reduced sample of sites that have more detailed site characterisation, to confirm the results obtain here, especially on the location of the non-linear soil behaviour.

Improving the inversion of borehole data

We discussed the limitation of the numerical method to simulate the amplitude of the pseudo-resonance peaks. The cost function used in the inversion algorithm is based on the calculation

of the distance between the observed and the numerical BFSR. The numerical BFSR indicated very large pseudo-resonance peaks compared to the observations. The cost function was mainly influenced by this discrepancy, and therefore the rest of the curve was not well adjusted. One way to solve this problem would be to force the inversion method to fit all of the peaks of the curves by modifying the cost function to adjust the frequency peaks, as proposed by De Martin et al. (2010).

Finally, in the third part of this study, we highlighted the importance of adapting the interpretation of the inversion results to the input data quality and the assumptions of the forward modelling method. Sensitivity analysis of the forward modelling to synthetic cases close to the solution appear to be a powerful tool to find the resolution and the parameters that can be well constrained in the inversion. One way to increase the resolution of borehole data would be to use multi-sensor vertical arrays. A first test of sensitivity in synthetic cases should be realised, followed by inversion of empirical data.

HIGHLIGHTS

The following list highlights the main results and conclusions of the present study, along with their potential applications. The fields concerned are the site characterisation for site-effect assessment, GMPEs, and strong ground-motion simulation studies. We showed:

- which parameters must be measured to improve the site-effect assessment, in both the linear and non-linear ranges,
- which additional parameters to Vs_{30} can be used to improve the site classification by reducing the between-site variability of the site responses,
- that the PGA (Peak Ground Acceleration) is a relevant parameter for non-linear site effect assessment,
- that, whatever the site, non-linear soil behaviour affects the site response curves from moderate acceleration (75 cm/s^2 at the down-hole station),
- that only information on the non-linear soil behaviour of the superficial layers is enough to fully assess the non-linear site responses,
- that the depth from which non-linear soil behaviour has no influence on site response depends on the site and on the intensity of the seismic input motion,
- that careful attention needs to be taken when inverting data from vertical arrays. Sensitivity analyses are a powerful tool to evaluate the resolution of the inversion considering the available information and the soil parameters that are well solved during the inversion,

- that the combined used of sensitivity analysis with comparison of transfer function, without inversion, could be enough to assess the depth where the non-linear soil behaviour mostly take place.

References

REFERENCES

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

Aguirre, J., & Irikura, K. (1997). Nonlinearity, liquefaction, and velocity variation of soft soil layers in Port Island, Kobe, during the Hyogo-ken Nanbu earthquake. *Bulletin of the Seismological Society of America*, *87*(5), 1244-1258.

Al Atik, L., Abrahamson, N., Bommer, J. J., Scherbaum, F., Cotton, F., & Kuehn, N. (2010). The variability of ground-motion prediction models and its components. *Bulletin of the Seismological Society of America*, *81*(5), 794-801.

Allen, T., & Wald, D. (2009). On the use of high-resolution topographic data as a proxy for seismic site conditions (Vs30). *Bulletin of the Seismological Society of America*, *99*(2A), 935-943.

Anderson, D., & Richart Jr, F. (1976). Effects of stratining on shear modulus of cl. *Transportation research board*, *102*(GT9), 975-987.

Arias, A. (1970). A measure of earthquake intensity. *Seismic design for nuclear power plant, MIT press Cambridge*, 438-483.

Assimaki, D., & Li, W. (2012). Site and ground motion-dependent nonlinear effects in seismological model predictions. *Soil Dynamics and Earthquake Engineering*(32), 143-151.

Assimaki, D., Li, W., Steidl, J., & Schmedes, J. (2008). Quantifying nonlinearity suceptibility via site response modelling uncertainty at 3 sites un Los Angeles basin. *Bulletin of the Seismological Society of America*, *98*(5), 2364-2390.

Assimaki, D., & Steidl, J. H. (2007). Inverse analysis of weak and strong motion downhole array data from the Mw 7.0 Sanriku-Minami earthquake. *Soil Dynamics and Earthquake Engineering*, *27*, 73-92.

Bard, P.-Y., & Bouchon, M. (1985). The two-dimensional resonance of sediment-filled valleys. *Bulletin of the Seismological Society of America*, *75*, 519-541.

Beresnev, I., & Wen, K. (1996). Nonlinear soil response, a reality? *Bulletin of the Seismological Society of America*, *86*(6), 1964-1978.

Beresnev, I. A., Wen, K.-L., & Yeh, Y. T. (1995). Nonlinear soil amplification: Its corroboration in

Taiwan. *Bulletin of the Seismological Society of America*, *85*(2), 456-515.

Berstimas, D., & Tsitsikis, J. (1993). Simulated annealing. *Statistical science*, *8*(1), 10-15.

Bertrand, E. (2000). *Sismologie large-bande : des ondes de volume aux structures de la croute et du manteau superieur.* Unpublished doctoral dissertation, Sophia Antipolis.

Biot, M. A. (1956). Theory of propagation of elastic waves in a fluid-saturated porous solid. i. low-frequency range. *The Journal of the Acoustical Society of America*, *28*(2), 168-178. Available from `http://link.aip.org/link/?JAS/28/168/1`

Boatwright, J., & Fletcher, J. B. (1984). The partition of radiated energy between p and s waves. *Bulletin of the Seismological Society of America*, *74*(2), 361-376. Available from `http://www.bssaonline.org/content/74/2/361.abstract`

Bonilla, L. F., Archuleta, R. J., & Lavallée, D. (2005). Hysteretic and dilatant behavior of cohesionless soils and their effects on nonlinear site response: Field data observations and modeling. *Bulletin of the Seismological Society of America*, *95*(6), 2373-2395.

Bonilla, L. F., Lavallée, D., & Archuleta, R. J. (1998). Nonlinear site response: Laboratory modeling as a constraint for modeling accelerograms. In K. Irikura & H. Kudo (Eds.), *Proceeding of the second international symposium on the effects of surface geology on seismic motion* (Vol. 2).

Bonilla, L. F., Steidl, J. H., Gariel, J., & Archuleta, R. J. (2002). Borehole response studies at the Garner Valley downhole array southern California. *Bulletin of the Seismological Society of America*, *92*, 3165-3279.

Bonilla, L. F., Steidl, J. H., Lindley, G. T., G.Tumarkin, A., & Archuleta, R. J. (1997). Site amplification in the San Fernando valley, California: Variability of site-effect estimation using the S-wave, coda, and H/V methods. *Bulletin of the Seismological Society of America*, *87*(3), 710-730.

Bonilla, L. F., Tsuda, K., Pulido, N., Regnier, J., & Laurendeau, A. (2011). Nonlinear site response evidence of K-net and KiK-net records from the Mw 9 Tohoku earthquake. *Earth Planets Space*, *58*.

Boore, D. M. (2004). Estimating Vs30 (or NEHRP site classes) from shallow velocity models

(depths 30 m). *Bulletin of the Seismological Society of America*, *94*(2), 591-597.

Boore, D. M., & Joyner, W. B. (1997). Site amplifications for generic rock sites. *Bulletin of the Seismological Society of America*, *87*(2), 327-341.

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

Borcherdt, R., & Glassmoyer, G. (1992). On the characteristics of local geology and their influence on ground motions generated by the Loma Prieta earthquake in the San Francisco bay region, California. *Bulletin of the Seismological Society of America*, *82*(2), 603-641.

Borcherdt, R., Westerlund, R., Dietel, C., & Glassmoyer, G. (2004). The integrated surface and borehole, geotechnical strong-motion, soil response arrays in San Francisco, California. In *International workshop for site selection, installation and operation of geotechnical strong motion arrays.*

Borcherdt, R. D. (1970). Effects of local geology on ground motion near San Francisco bay. *Bulletin of the Seismological Society of America*, *60*, 29-61.

Box, G., & Cox, D. (1964). An analysis of transformations. *Journal of the royal statistical society*, *26*(2), 211-252.

Cabañas, L., Benito, B., & Herraiz, M. (1997). An approach to the measurement of the potential structural damage of earthquake ground motions. *Earthquake Engineering and Structural dynamics*, *26*, 79-92.

Cadet, H. (2007). *Utilisation combinée des méthodes basées sur le bruit de fond dans le cadre du microzonage sismique*. Unpublished doctoral dissertation, Université Joseph Fourier, Grenoble.

Cadet, H., Bard, P.-Y., Duval, A.-M., Bertrand, E., & Rodriguez-Marek, A. (2012). Site effect assessment using KiK-net data - Part 2 – Site amplification prediction equation SAPE based on f0 and Vsz. *Bulletin of Earthquake Engineering*, *10*(2), 451-481.

Cadet, H., Bard, P.-Y., & Rodriguez-Marek, A. (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. (2009). Estimates of shear-wave Q and Îž0 for unconsolidated and semiconsolidated sediments in Eastern North America. *Bulletin of the Seismological Society of America*, *99*(4), 2365-2392. Available from http://www.bssaonline.org/content/99/4/2365.abstract

Castellaro, S., & Mulargia, F. (2009). Vs30 estimates using constrained H/V measurements. *Bulletin of the Seismological Society of America*, *99*(2A), 761-773.

Castellaro, S., Mulargia, F., & Rossi, P. L. (2008, 540-543). Vs30: Proxy for seismic amplification? *Bulletin of the Seismological Society of America*, *74*(4).

Cerny, V. (1985). Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. *journal of optimization theory and application*, *45*(1), 41-51.

Chin, B.-H., & Aki, K. (1991). Simultaneous study of the source, path, and site effects on strong ground motion during the 1989 loma prieta earthquake: A preliminary result on pervasive nonlinear site effects. *Bulletin of the Seismological Society of America*, *81*(5), 1859-1884. Available from http://www.bssaonline.org/content/81/5/1859.abstract

Cotton, F., Scherbaum, F., Bommer, J. J., & Bungum, H. (2006). Criteria for selecting and adjusting ground-motion models for specific target regions: Application to central Europe and rock sites. *Journal of Seismology*.

Courboulex, F. (1995). *Inversion spatio-temporelle de la source sismique à l'aide des fonctions de green empiriques: Déconvolution par recuit simulé application à des séismes de faible magnitude*. Unpublished doctoral dissertation, Institut de géodynamique de Nice-Sophia Antipolis.

Cranswick, E., King, K., Carver, D., Worley, D., Williams, R., Spudich, P., et al. (1990). Site response across downtown Santa Cruz, California. *Geophysical research letter*, *17*(10), 1793-1796.

Cukier, R. I., Fortuin, C. M., Shuler, K. E., Petschek, A. G., & Schaibly, J. H. (1973). Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. i theory. *The Journal of Chemical Physics*, *59*(8), 3873-3878. Available from http://link.aip.org/link/?JCP/59/3873/1

Dainty, A. M. (1981). A scattering model to explain seismic Q observations in the lithosphere between 1 and 30 hz. *Geophysical Research Letters*, *8*(11), 1126–1128. Available from http://dx.doi.org/10.1029/GL008i011p01126

Darragh, R. B., & Shakal, A. F. (1991). The site response of two rock and soil station pairs to strong and weak ground motion. *Bulletin of the Seismological Society of America*, *81*(5), 1885-1899. Available from http://www.bssaonline.org/content/81/5/1885.abstract

Delépine, N., Lenti, L., Bonnet, G., & Semblat, J.-F. (2009). Nonlinear viscoelastic wave propagation: an extension of nearly constant attenuation (NCQ) models. *journal of engineering mechanics*, *135*(11), 1305-1314.

De Martin, F. (2011). Comparison of frequency and time-domain objective function for borehole station's inverse problems. In *5icege 5th international conference on earthquake geotechnical engineering.*

De Martin, F., Kawase, H., & Razavi, A. M.-F. (2010, February). Nonlinear soil response of a borehole station based on one-dimensional inversion during the Fukuoka prefecture western offshore earthquake. *Bulletin of the Seismological Society of America*, *100*(1), 151-171.

Douglas, J. (2006). Estimating strong ground motions at great depths. *ESG2006 grenoble*.

Douglas, J., Gehl, P., Bonilla, L. F., Scotti, O., Régnier, J., Duval, A.-M., et al. (2009). Making the most of available site information for empirical ground-motion prediction. *Bulletin of the Seismological Society of America*, *99*(3).

Drouet, S. (2006). *Analysis of the accelerometric data applied to seismic hazard assessment in France*. Unpublished doctoral dissertation, Université toulouse III-Paul Sabatier.

Drouet, S., Chevrot, S., Cotton, F., & Souriau, A. (2008, February). Simultaneous inversion of source spectra, attenuation parameters, and site responses: Application to the data of the French Accelerometric Network. *Bulletin of the Seismological Society of America*, *98*(1), 198-219.

Duncan, J., & Chang, C. (1970). Nonlinear analysis of stress and strain in soils. *ASCE J. of the Soil Mech. And Found.*, *96*, 1629-1653.

Duval, A.-M., Méneroud, J.-P., Vidal, S., & Bard, P.-Y. (1996). Une nouvelle méthode d'évaluation

des sols aux séismes par enregistrement de bruit de fond. *Bulletin des Laboratoires des Ponts et Chaussées*, *1*(203), 75-90.

Faccioli, E., & Vitiello, E. (1973). A finite element, linear programming methods for the limit analysis of thin plates. *International Journal for Numerical Methods in Engineering*, *5*(3), 311-325.

Fäh, D., Kind, F., & Giardini, D. (2001). A theoretical investigation of average h/v ratios. *Geophysical Journal International*, *145*(2), 535–549. Available from http://dx.doi.org/10.1046/j.0956-540x.2001.01406.x

Fäh, D., Kind, F., & Giardini, D. (2003). Inversion of local s-wave velocity structures from average h/v ratios, and their use for the estimation of site-effects. *Journal of Seismology*, *7*(4), 449-467.

Field, E., Johnson, P. A., Beresnev, I. A., & Zeng, Y. (1997). Nonlinear ground-motion amplification by sediments during the 1994 Northridge earthquake. *Letters to Nature*(390), 599-602.

Frankel, A. D., Carver, D. L., & Williams, R. A. (2002). Nonlinear and linear site response and basin effects in Seattle for the M 6.8 Nisqually, Washington, earthquake. *Bulletin of the Seismological Society of America*, *92*(6), 2090-2109.

Futterman, W. I. (1962). Dispersive body waves. *Journal of Geophysical Research*, *67*(13), 5279–5291. Available from http://dx.doi.org/10.1029/JZ067i013p05279

Gélis, C., & Bonilla, L. F. (2012). 2-d p–sv numerical study of soil–source interaction in a non-linear basin. *Geophysical Journal International*, *191*(3), 1374–1390. Available from http://dx.doi.org/10.1111/j.1365-246X.2012.05690.x

Glaser, S., & Baise, L. (2000). System identification estimation of soil properties at the Lotung site. *soil dynamics and earthquake engineering*, *19*, 521-531.

Graves, R. W., Aagaard, B. T., Hudnut, K. W., Star, L. M., Stewart, J. P., & Jordan, T. H. (2008). Broadband simulations for Mw 7.8 southern San Andreas earthquakes: Ground motion sensitivity to rupture speed. *Geophysical research letters*, *35*(22302).

Hardin, B. O., & Drnevich, V. P. (1972). Shear modulus and damping in soils. *Journal of the Soil Mechanics and Foundations Division,*, *98*, 667-692.

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

Hayashi, H., Honda, M., Yamada, T., & tatsukoa fumio. (1992). Modeling of nonlinear stress strain relations of sands for dynamic response analysis. In *World conference of earthquake engineering, madrid, spain* (Vol. 11, p. 6819-6825).

Hernández, B. (2011). Pertinence de divers indicateurs de nocivité à augurer des effets des séismes. In *AFPS 8ème colloque national, Paris, France.*

Iai, S., Morita, T., Kameoka, T., Matsungaya, Y., & Abiko, K. (1995). Response of a dense sand deposit during 1993 Kushiro-Oki earthquake. *soils and foundations*, *35*(1), 115-131.

Idriss, I. (2009). Use of Vs30 to represent effects of local site conditions on earthquake ground motions. In *Ssa annual meeting. session site effects: Vs30 and beyond.*

Ishibashi, I., & Zhang, X. (1993). Unified dynamic shear moduli and damping ratio of sand and clay. *Soils and Foundations*, *33*, 182-191.

Ishihara, K. (1996). *Soil behaviour in earthquake geotechnics*. Oxford: Clarenton Press.

Iwan, W. (1967). On a class of models for the yielding behavior of continuous and composite systems. *journal of applied mechanics*, *34*(3), 612-617.

Johnston, D. H., Toksoz, M. N., & Timur, A. (1979). Attenuation of seismic waves in dry and saturated rocks; ii, mechanisms. *Geophysics*, *44*(4), 691-711. Available from `http://geophysics.geoscienceworld.org/content/44/4/691.abstract`

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *science*, *220*(4598).

Kokusho, T. (2004). Nonlinear site response and strain-dependent soil properties. *Current science*, *87*(10), 1363-1369.

Konno, K., & 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*, 228-241.

Kramer, L. (1996). *Geotechnical earthquake engineering* (Vol. 1). Prentice-Hall International Series.

Langston, C. A. (1979). Structure under mount Rainier, Washington, inferred from teleseismic body waves. *journal of geophysical research*, *84*, 4749-4762.

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

Leurer, K. C. (1997). Attenuation in fine-grained marine sediments; extension of the biot-stoll model by the "effective grain model" (egm). *Geophysics*, *62*(5), 1465-1479. Available from `http://geophysics.geoscienceworld.org/content/62/5/1465.abstract`

Li, X., & Moshell, J. M. (1993). Modeling soil: Realtime dynamic models for soil slippage and manipulation. In *Proceeding siggraph '93 proceedings of the 20th annual conference on computer graphics and interactive techniques.*

Lilliefors, H. (1967). On the Komogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, *62*, 399-402.

Liu, P., Hartzell, S., & Stephenson, W. (1995). Non-linear multiparameter inversion using a hybrid global search algorithm: applications in reflection seismology. *Geophysical Journal International*, *122*(3), 991–1000. Available from `http://dx.doi.org/10.1111/j.1365-246X.1995.tb06851.x`

Lopez-Caballero, F., Julie, R., Bonilla, L. F., & Gélis, C. (2011). Site response analysis including earthquake input ground motion and soil dynamic properties variability. In *Afps 8ème colloque national.*

Lo-Presti, D., Amiolkowski, M., Pallara, O., Cavallaro, A., & Pedroni, S. (1997). Shear modulus and damping of soils. *geotechnique*, *47*(3), 603-617.

Luzi, L., Puglia, R., Pacor, F., Gallipoli, M. R., Bindi, D., & Mucciarelli, M. (2011). Proposal for a soil classification based on parameters alternative or complementary to Vs,30. *Bulletin of Earthquake Engineering*, *9*, 1977-1898.

Mara, T. A. (2009). Extension of the rbd-fast method to the computation of global sensitivity indices. *Reliability Engineering and System Safety*, *94*(8), 1274 - 1281. Available from `http://www.sciencedirect.com/science/article/pii/S095183200900026X`

Marketos, G., & Bolton, M. (2006). A statistical investigation of particle crushing in sand. In

International symposium on geomechanics and geotechnics of particulate behaviour (p. 247-252).

Masing, G. (1926). Eigenspannungen und verfestigung beim messing. In *Proceedings of the second international congress of applied mechanics* (p. 332-335).

Matasovic, N., & vucetic, M. (1993). *Analysis of seismic records obtained on november 24, 1987 at the Wildlife liquefaction array* (Tech. Rep.). University of California, Los Angeles.

Matasovic, N., & vucetic, M. (1995). Generalized cyclic-degradation-pore-pressure generation model for clays. *Journal of geotechnical engineering*, *121*(1), 33-43.

Morozov, I. B. (2008). Geometrical attenuation, frequency dependence of Q, and the absorption band problem. *Geophysical Journal International*, *175*, 239-252.

Morozov, I. B. (2009). Thirty years of confusion around "scattering Q"? *Seismological Research Letters*, *80*, 5-7.

Moss, R. (2008). Quantifying measurement uncertainty of Vs30. *Bulletin of the Seismological Society of America*, *98*(3), 1399-1411.

Nelder, J., & Mead, R. (1965). A simplex method for function minimization. *the computer journal*, *7*(4), 308-313.

Noguchi, S., & Sasatani, T. (2008). Quantification of degree of nonlinear site response. In *The 14th World Conference on Earthquake Engineering, Beijing, China.*

O'Connell, R. J., & Budiansky, B. (1978). Measures of dissipation in viscoelastic media. *Geophysical Research Letters*, *5*(1), 5–8. Available from `http://dx.doi.org/10.1029/GL005i001p00005`

Park, D., & Hashash, Y. M. (2004). Probabilistic seismic hazard analysis with nonlinear site effects in the Mississippi embayment. *13th World Conference on Earthquake Engineering*(1549).

Pavlenko, O. V., & Irikura, K. (2002). Changes in shear moduli of liquefied and nonliquefied soils during the 1995 Kobe earthquake and its aftershocks at three vertical-array sites. *Bulletin of the Seismological Society of America*, *92*(5), 1952-1969.

Pavlenko, O. V., & Irikura, K. (2003). Estimation of nonlinear time-dependent soil behavior in strong ground motion based on vertical array data. *Pure and Applied Geophysics*, *160*(12),

2365-2379.

Pavlenko, O. V., & Irikura, K. (2006). Nonlinear behavior of soils revealed from the records of the 2000 Tottori, Japan, earthquake at stations of the digital strong-motion network Kik-net. *Bulletin of the Seismological Society of America*, *96*(6), 2131-2145.

PECKER, A. (2005). Maximum ground surface motion in probabilistic seismic hazard analyses. *Journal of Earthquake Engineering*, *09*(spec01), 187-211. Available from `http://www.worldscientific.com/doi/abs/10.1142/S1363246905002225`

Platon, P. (1933). Texte établi et traduit par a. *Rivaud, Tel, Gallimard*.

Pyke, R. M. (1979). Discussion of "definition of terms related to liquefaction". *Journal of geotechnical engineering division*, *105*(10), 1260.

Pyke, R. M. (1980). Nonlinear soil models for irregular cyclic loadings. *Journal of Geotechnical and Geoenvironmental Engineering*, *106*, 1277-1282.

Raykov, T., & Marcoulides, A. G. (2008). *An introduction to applied multivariate analysis*. Routledge taylor and Francis group.

Régnier, J., Bonilla, L. F., Bertrand, E., & Semblat, J.-F. (2013). Influence of the velocity profile at depth on the linear characterization of site effects: Tests on the KiK-net database. *Bulletin of the Seismological Society of America*.

Régnier, J., Cadet, H., Bonilla, L., Bertand, E., & Semblat, J. (2013). Assessing nonlinear behavior of soil in seismic site response: Statistical analysis on KiK-net strong motion data. *Bulletin of the Seismological Society of America*, *Summited*.

Riepl, J., Bard, P.-Y., Hatzfeld, D., Papaioannou, C., & Nechtschein, S. (1998). Detailed evaluation of site-response estimation methods across and along the sedimentary valley of Volvi (EURO-SEISTEST). *Bulletin of the Seismological Society of America*, *88*(2), 488-502.

Rodriguez-Marek, A., Mantalva, G. ., Cotton, F., & Bonilla, F. (2011). Analysis of single-station standard deviation using the KiK-net data. *Bulletin of the Seismological Society of America*, *101*(3), 1242-1258.

Ryden, N., & Park, C. (2006). fast simulated annealing inversion of surface waves on pavement using phase-velocity spectra. *geophysics*, *71*(4), 49-58.

Saltelli, A., F. Campolongo, S. T. adn, & Ratto, M. (2004). *Sensitivity analysis in practice a guide to assessing scientific models*. john Wiley and Sons, Ltd.

Saltelli, A., Tarantola, S., & Chan, K.-S. (1999). A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, *41*(1), 39-56.

Sato, K., Kokusho, T., Matsumoto, M., & Yamada, E. (1996). Nonlinear seismic response and soil property during strong motion. *soils and foundations*, 41-52.

Satoh, T., Fushimi, M., & Tatsumi, Y. (2001). Inversion of strain-dependent nonlinear characteristics of soils using weak and strong motions observed by borehole sites in Japan. *Bulletin of the Seismological Society of America*, *91*(2), 365-380.

Satoh, T., Sato, T., & Kawase, H. (1995). Nonlinear behavior of soil sediments identified by using borehole records observed at the Ashigara valley, Japan. *Bulletin of the Seismological Society of America*, *85*(6), 1821-1834.

Schnabel, P., Lysmer, J., & Seed, H. (1972). *Shake : a computer program for earthquake response analysis of horizontally layered sites* (report). Berkeley, California: Earthquake Engineering Research Centre.

Semblat, J. F., Kham, M., Parara, E., Bard, P. Y., Pitilakis, K., Makra, K., et al. (2005). Seismic wave amplification : Basin geometry Vs soil layering. *Soil Dynamics and Earthquake Engineering*, *25*, 529-538.

Semblat, J.-F., & Pecker, A. (2009). *Waves and vibrations in soils: earthquakes, traffic, shocks, construction works* (I. Press, Ed.). IUSS Press. Available from `http://hal.archives-ouvertes.fr/hal-00356022`

Shearer, P., & Orcutt, J. (1987). Surface and near-surface effects on seismic waves—theory and borehole seismometer results. *Bulletin of the Seismological Society of America*, *77*(4), 1168-1996.

Shibuya, S., Mitachi, T., Fukuda, F., & Degoshi, T. (1995). Strain rate effects on shear modulus and damping of normally consolidated clay. *ASTM geotechnical testing journal*, *18*(3), 365-375.

Sobol, I. (1993). Sensitivity estimates for nonlinear mathematical models. *Mathematical modeling and computational experiment*, *1*(4), 407-414.

Steidl, J. H., Tumarkin, A. G., & Archuleta, R. J. (1996). What is a reference site? *Bulletin of the Seismological Society of America*, *86*, 1733-1748.

Streeter, V. L., Wylie, E. B., & Frank E. Richart, J. (1974). Soil motion computations by characteristics method. *Journal of the Geotechnical Engineering Division*, *100*(3), 247-263.

Tarantola, S., Gatelli, D., & Mara, T. (2006). Random balance designs for the estimation of first-order global sensitivity indice. *Reliability Engineering and System Safety*, *91*(6), 717-727.

Theodulidis, N., & Bard, P. Y. (1995). Horizontal to Vertical spectral ratio and geological conditions: an analysis of strong motion data from Greece and Taiwan (SMART-1). *Soil Dynamics and Earthquake Engineering*, *14*(3), 177-197.

Thompson, E., Baise, L., Kayen, R., & Guzina, B. (2009, october). Impediments to predicting site response: Seismic property estimation and modeling simplifications. *Bulletin of the Seismological Society of America*, *99*(5), 2927-2949.

Thompson, E., Baise, L., Tanaka, Y., & Kayen, R. (2012). A taxonomy of site response complexity. *Soil Dynamics and Earthquake Engineering*, *41*, 32-43.

Thomson, W. (1950). Transmission of elastic waves through a stratified solid. *journal of applied physics*, *21*, 89-93.

Tselentis, G.-A. (1998). Intrinsic and scattering seismic attenuation in W. Greece. *Pure and Applied Geophysics*, *153*, 703–712.

Vucetic, M., & Dobry, R. (1991). Effect of soil plasticity on cyclic response. *Journal of geotechnical engineering*, *117*(1).

Vucetic, M., & Matasovic, N. (1992). *Modeling of the stress-strain behavior of liquefiable sands* (Tech. Rep.). Japanese association for the development of Earthquake prediction.

Wald, D., & Allen, T. (2007). Topographic slope as a proxy for seismic site conditions and amplification. *Bulletin of the Seismological Society of America*, *97*(5), 1379-1395.

Wen, K.-L. (1994). Non-linear soil response in ground motions. *Earthquake Engineering and Structural dynamics*, *26*(6), 599-608.

Wen, K.-L., Huang, J.-Y., Chen, C.-T., & Cheng, Y.-W. (2011). Nonlinear site response of the 2010 Darfield, New Zealand earthquake sequence. In *4th iaspei / iaee international symposium:*

Effects of surface geology on seismic motion.

Withers, M., Aster, R., Young, C., Beiriger, J., Harris, M., Moore, S., et al. (1998). A comparison of select trigger algorithms for automated global seismic phase and event detection. *Bulletin of the Seismological Society of America*, *88*, 95-106.

Woods, D. (1985). *An interactive approach for solving multi-objective optimization problems*. Unpublished doctoral dissertation, Univesity Microfilms international.

Xu, C., & Gertner, G. (2008). A general first-order global sensitivity analysis method. *Reliability Engineering and System Safety*, *93*(7), 1060-1071.

Yoshimoto, K., Sato, H., & Ohtake, M. (1992). Frequency-dependent attenuation of P and S waves in the Kanto area, Japan, based on the coda-normalization method. *Geophysical Journal International*(114), 165-174.

Yu, G., Anderson, J. G., & Siddharthan, R. (1993). On the characteristics of nonlinear soil response. *Bulletin of the Seismological Society of America*, *83*(1), 218-244.

Zeghal, M., & Elgamal, A.-W. (1994). Analysis of site liquefaction using earthquake records. *Journal of geotechnical engineering*, *120*(6), 996-1017.

Zeghal, M., Elgamal, A.-W., Tang, H., & Srepp, J. (1995). Lotung downhole array. ii: Evaluation of soil nonlinear properties. *Journal of geotechnical engineering*, *121*(4), 363-378.

Appendix A

Earthquake locations at sites FKSH12, FKSH14, KSRH10 and MYGH04

In this appendix, the location of the earthquake epicenters recorded by the four KiK-net selected sites for the inversion is shown in the following Figures:

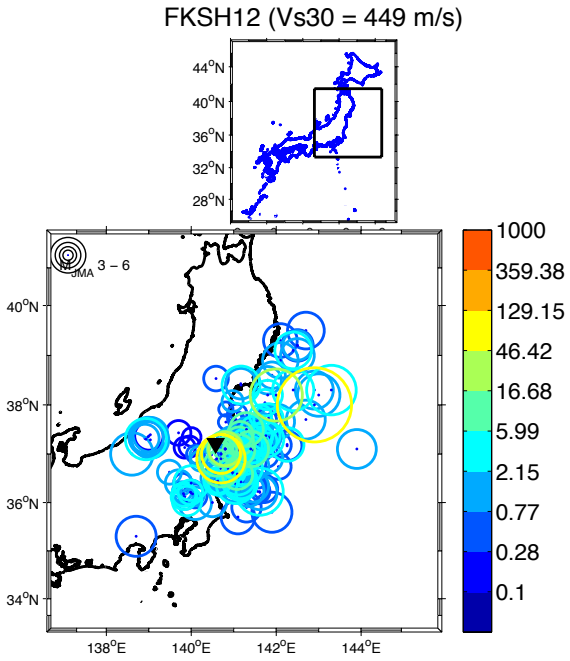


Figure A.1: *Location of the earthquakes recorded by station FKSH12, according to the PGA at the surface*

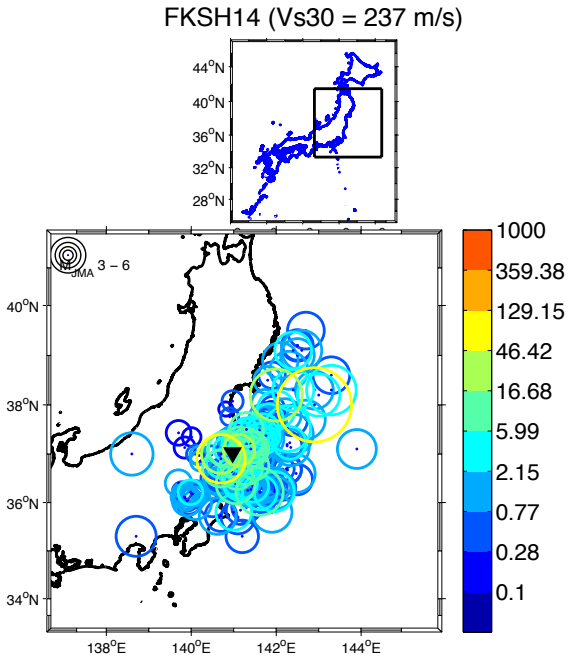


Figure A.2: *Location of the earthquakes recorded by station FKSH14, according to the PGA at the surface*

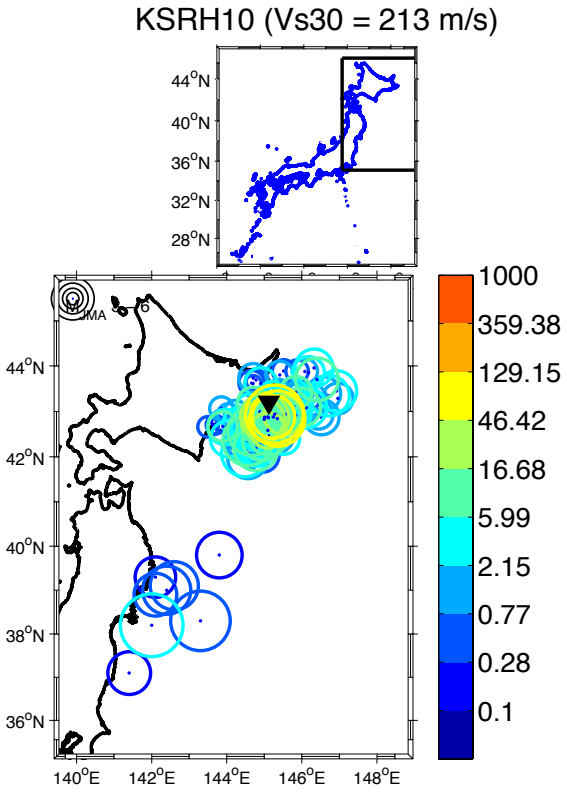


Figure A.3: *Location of the earthquakes recorded by station KSRH10, according to the PGA at the surface*

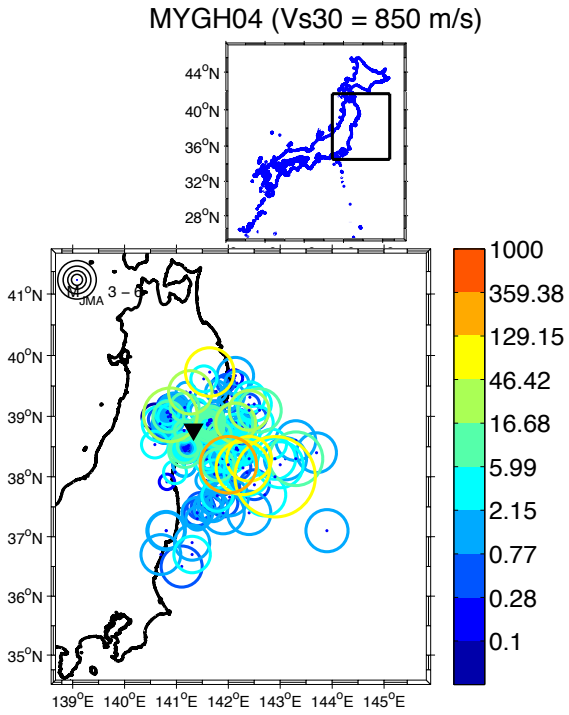


Figure A.4: *Location of the earthquakes recorded by station MYGH04, according to the PGA at the surface*

Appendix B

Illustration of the down-going wave effect on the numerical borehole site responses

We used a very simple synthetic case of a monolayer of sediment based on a bedrock. The velocity profile of this case is shown in Figure B.1. The sediment layer is 10 m in depth, with a Vs of 100 m/s, and the bedrock is a semi-infinite space with a Vs of 1000 m/s. According to the well-known formula $f_0 = Vs/4H$, the fundamental resonance of this soil is 2.5 Hz. We used the tremor code (presented in the first chapter) to calculate the site transfer functions for different locations of the reference station.

Figure B.1 illustrates the BFSR$_{num}$ for different location depths of the downhole sensor (plain lines) and the OFSR$_{num}$ (dotted line). When the downhole sensor is located just at the sediment-to-bedrock interface, both the OFSR$_{num}$ (black dotted line) and the BFSR$_{num}$ indicate the true soil resonance frequency. The most downhole sensor location is far from the sediment-to-bedrock interface, and the first resonance peak is shifted to the low frequency band. Also, below 50 m in depth, the BFSR$_{num}$ has an additional peak between the fundamental resonance (2.5 Hz) and the first harmonic (7.5 Hz). This peak comes from destructive interference between the up-going waves and the down-going waves that is generated by the refraction of the up-going waves at the sediment-to-bedrock interface. Let us consider the case when the borehole sensor is located at 60 m in depth, so 50 m below the interface (cyan inverse triangle); the second peak in the BFSR$_{num}$ (cyan dotted curve) is at 5 Hz, which is in agreement with the formula $f_0 = Vs/4H$, with a Vs of 1000 m/s and a thickness of 50 m.

There is less phase difference between the interface and the borehole sensor location when the velocity of the substratum is increase. As illustrated in Figure B.2, the influence of the location of the downhole sensor is reduced when the velocity in the bedrock is greater. The first peak resonance frequency is similar, and the higher modes are nevertheless significantly affected by the location of the downhole sensor.

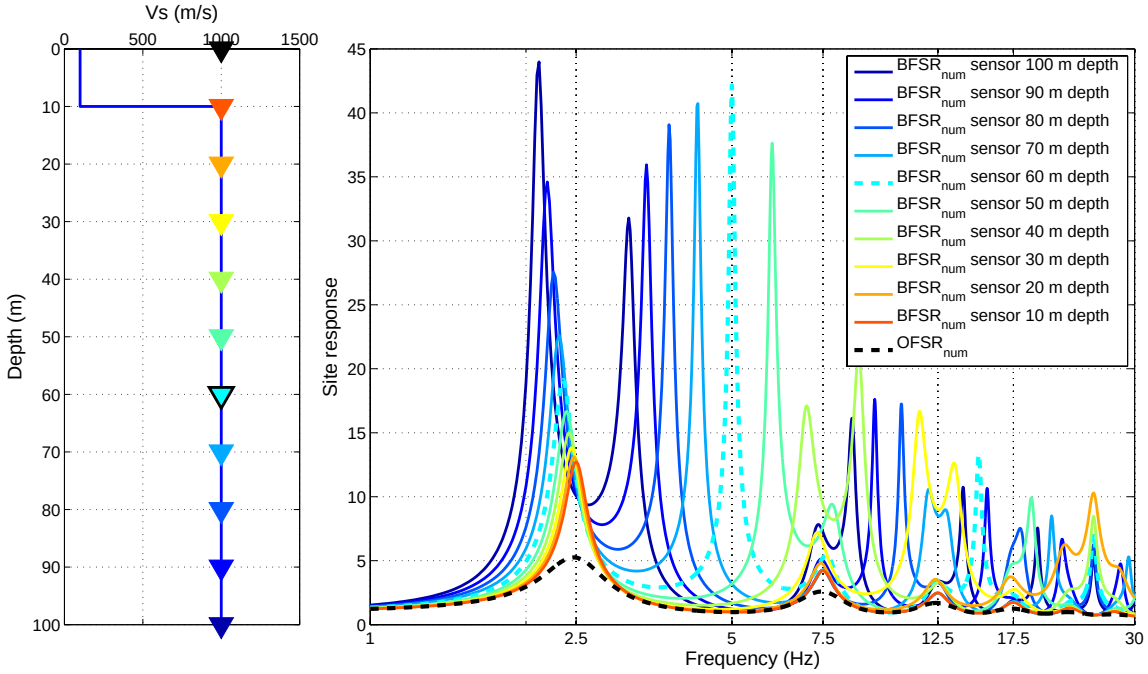


Figure B.1: *$BFSR_{num}$ according to the depth of location of the down-hole station. The black line represents the $OFSR_{num}$. The velocity in the sediment layer is equal to 100 m/s and the velocity of the bedrock is equal to 1000 m/s.*

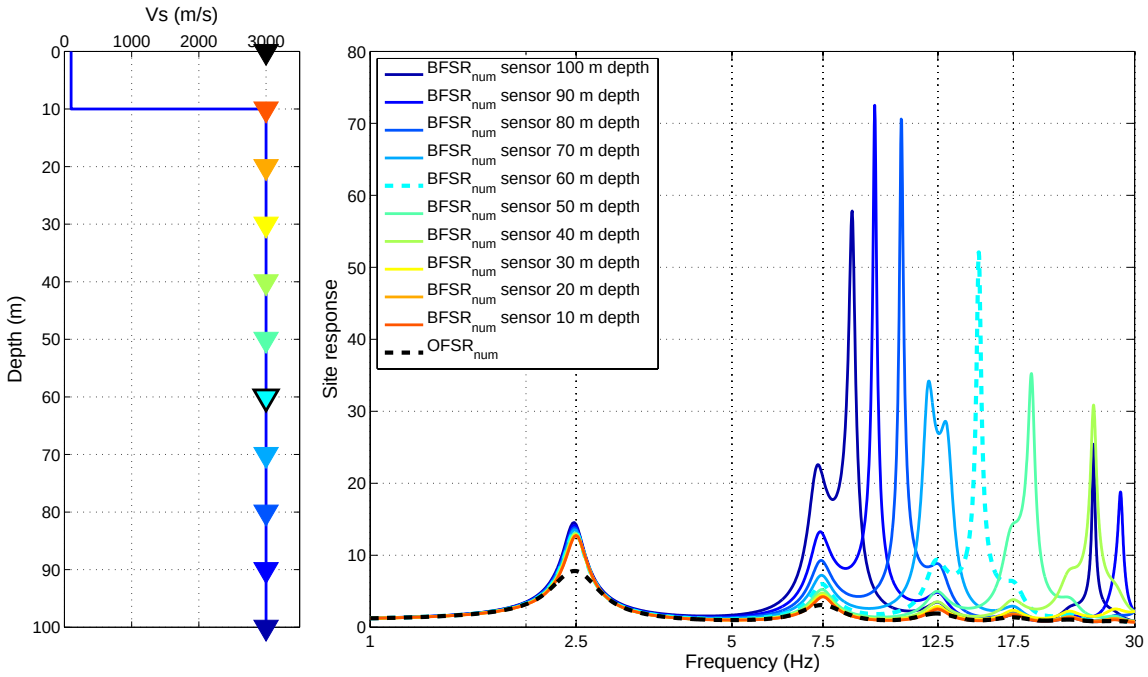


Figure B.2: *Similar figure as B.1 but with e velocity of the bedrock equal to 3000 m/s.*

Appendix C

Picking the fundamental resonance frequency

Considering the previous observations, finding the fundamental resonance frequency for the borehole site response appears difficult. The receiver function method or the horizontal to vertical spectral ratios is an alternative method to find the fundamental resonance frequency (for more detail, see chapter 3, in the section on empirical site-response calculation methods). We compare the fundamental resonance frequency picking on the BFSR and RF at the surface. The object of this section is to define a repeatable process to find the fundamental resonance frequency.

C.1 METHODOLOGY

We compare the empirical linear site-response curves (BFSR_{lin} and BFSR_{lin}^{95}) and alternative curves (RF_{lin}) with site responses computed with 1-D linear numerical simulations (BFSR_{num} and OFSR_{num}). We elaborate a code that automatically picks the first significant peak according to the following steps:

- Finding the local maximum.
- Checking if the maximum local has an amplitude significantly greater than 2. We used one sample student tests (t-tests) for the testing of the null hypothesis that the variable is normally distributed with a mean significantly greater than 2. For each local maximum, calculated on the mean of the empirical site responses (computed using a number of earthquakes N_{eq} depending on the frequencies where the signal-to-noise ratio is enough), located at the frequency f_j with an associated amplitude $A(f_j)$ and logarithmic standard deviation $\sigma_{log}(f_j)$,

we calculated:

$$t = \frac{\log_{10}(A(f_j)) - \log_{10}(2)}{\sigma_{log}(f_j)(/N_{eq}(f_j)}$$

We compared the value of t with a predefined set of values that depend on the number of earthquakes used to calculate the site response and the level of significance. Here we chose 0.5%, meaning that the test is true, the amplitude of the site responses at the maximum is higher than 2, with a 99.5 % of probability.

C.2 COMPARISON OF THE F_0

This automatic, and therefore repeatable, process was applied to all of the KiK-net sites. The automatic process was followed by a manual picking on the RF_{lin} at the surface. In the following part of this annex, the comparison between the first significant frequency peaks automatically picked for the BFSR_{lin} (f_0^{BFSR}) and RF_{lin} (f_0^{RF} auto) is shown. Then, we compare the manual picking with the automatic picking on the RF_{lin} (f_0^{RF}auto, f_0^{RF}manual).

C.2.1 Comparison between the automatically picked f_0 on the BFSR_{lin} or on the RF_{lin}

When comparing the f_0 picked automatically on the BFSR_{lin} or on the RF_{lin} (Figure C.1 (a)), we observe two groups of sites for which the frequencies f_0 are different:

- Group 1 ($f_0^{RF} < f_0^{BFSR}$) : Let us recall that receiver functions can show a low frequency peak associated to a deeper interface than the location of the downhole sensor, whereas the BFSR are obviously limited to this depth. Most of the sites in group 1 are characterised by weak Vs at the downhole sensor (the mean of the Vs at the borehole station for this group is 1300 m/s, whereas for all of the KiK-net sites it is around 1700 m/s). This means that the velocity profiles of group 1 are likely to have a strong impedance contrast below the downhole sensor depth. To illustrate group 1, we show in Figure C.2 the comparison of the empirical and numerical site-response curves.

- Group 2 ($f_0^{RF} < f_0^{BFSR}$): For this group, the first significant peak picked for the BFSR_{lin} might be associated to a pseudo-resonance that is not observed in the RF_{lin}. Also, the amplitudes of the RF_{lin} are lower than the BFSR_{lin}; thus for some sites, the first peak of the RF_{lin} might not respect the t-test on the amplitude. The comparison of the empirical and numerical curve at site AKTH05 illustrates this last (see Figure C.3).

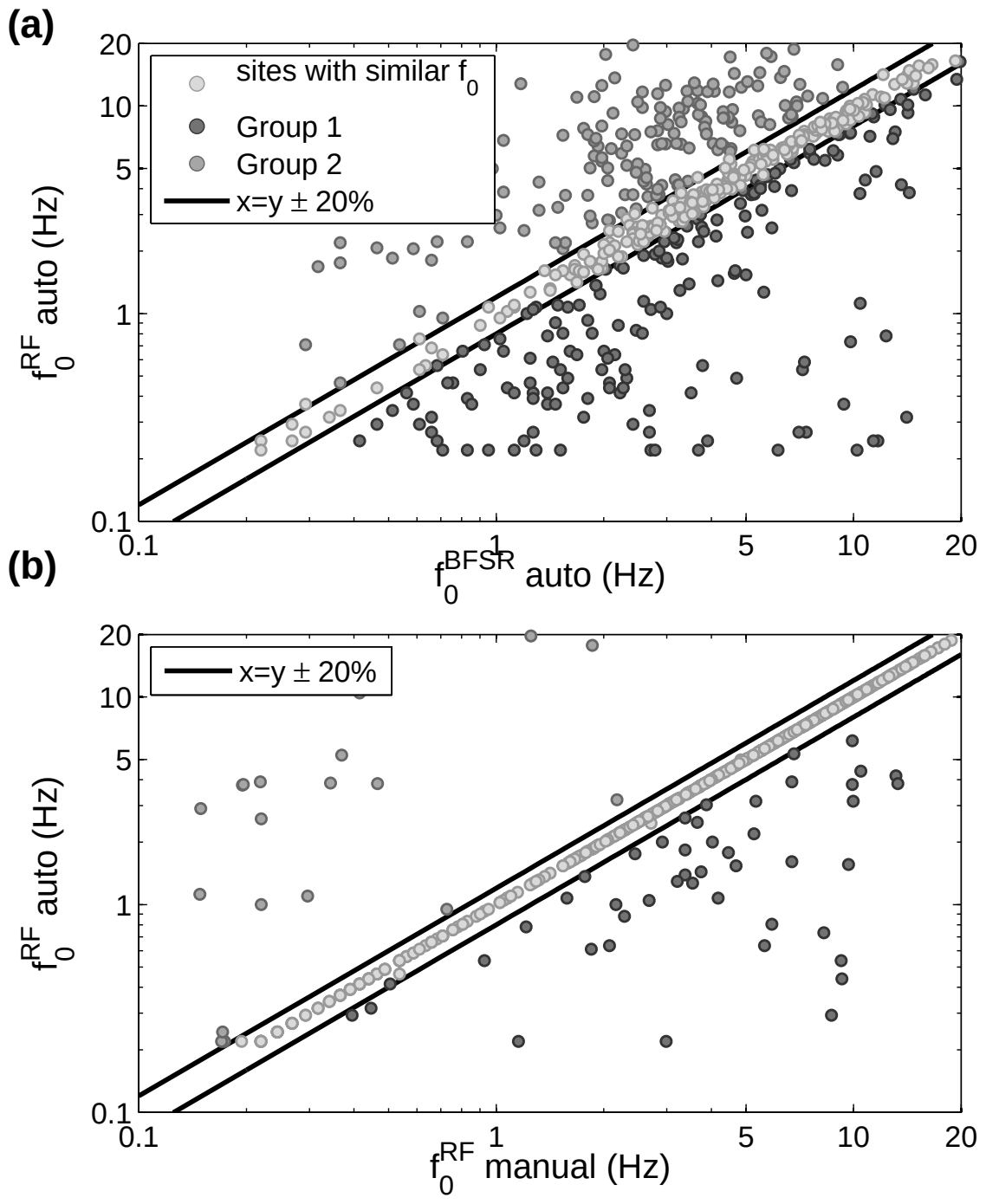


Figure C.1: *(a) Comparison between the first significant peak automatically picked on the $BFSR_{lin}$ and on the RF_{lin}. (b) Comparison of the significant peak manually and automatically picked on the RF_{lin}.*

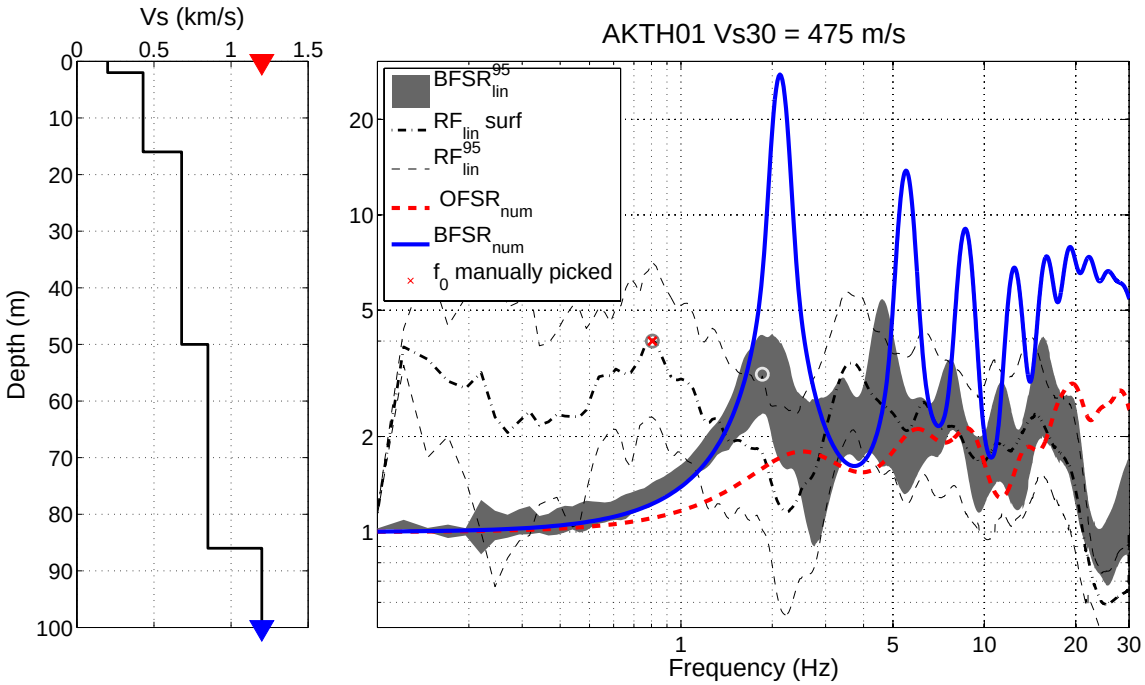


Figure C.2: *(a) Vs profile of the KiK-net site AKTH01. (b) Comparison of the empirical site-response $BFSR^{95}_{lin}$ and receiver function RF_{lin} at the surface (mean and 95% confidence limit), the numerical 1-D linear site-response $BFSR_{num}$ (borehole condition) and $OFSR_{num}$ (outcrop condition).*

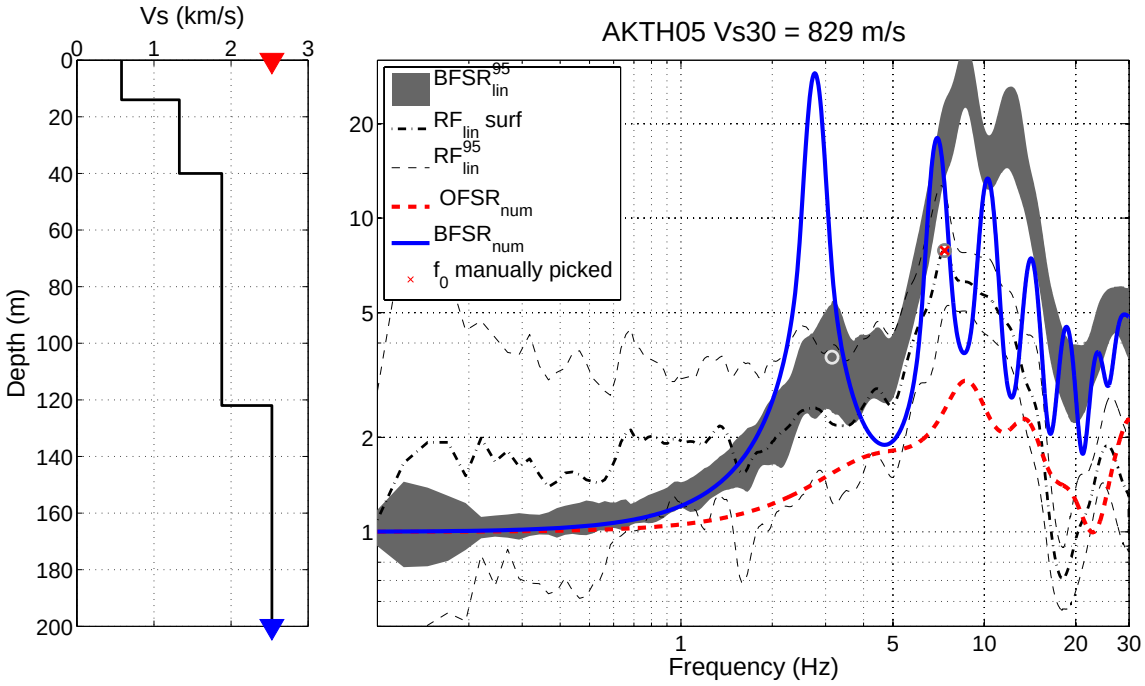


Figure C.3: *Similar to the previous Figure, but at the KiK-net site AKTH05*

C.2.2 Comparison between the automatically and manually picked f_0 on the RF$_{lin}$

The deviation between the two pickings is illustrated in Figure C.1(b).

- ($f_0^{RF} auto < f_0^{RF}$ manual) : The sites for which the fundamental resonance frequency picked automatically is below the manual picking are those that show a frequency peak with a higher amplitude at high frequency. In this case, when the peak is in agreement with the BFSR$_{lin}$ frequency peak, the operator chose the second peak. To illustrate these kinds of sites, we show the comparison of the site-response curves at site SZOH25 (see Figure C.4).

- ($f_0^{RF} auto > f_0^{RF}$ manual) : The sites for which the fundamental resonance frequency picked automatically is above the manual picking are those that indicate a clear amplification at low frequency, although the t-test rejects the null hypothesis. To illustrate these kinds of sites, we show the comparison of the site-response curves at site SZOH25 (see Figure C.5).

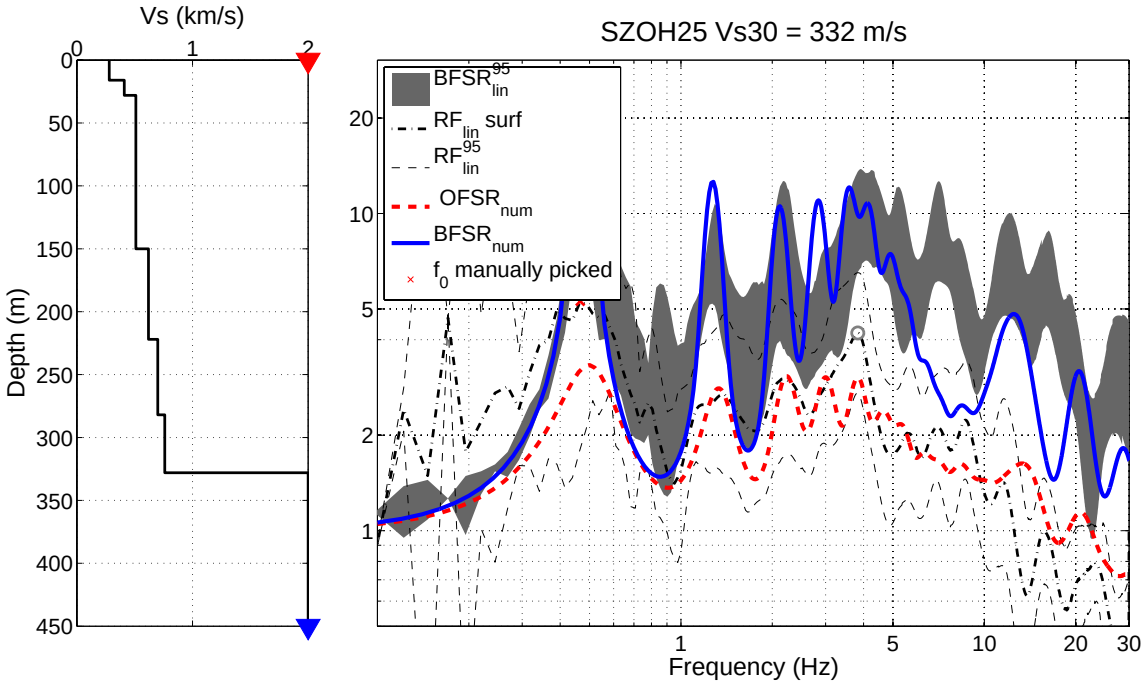


Figure C.4: *Similar to the previous Figure, but at KiK-net site SZOH25.*

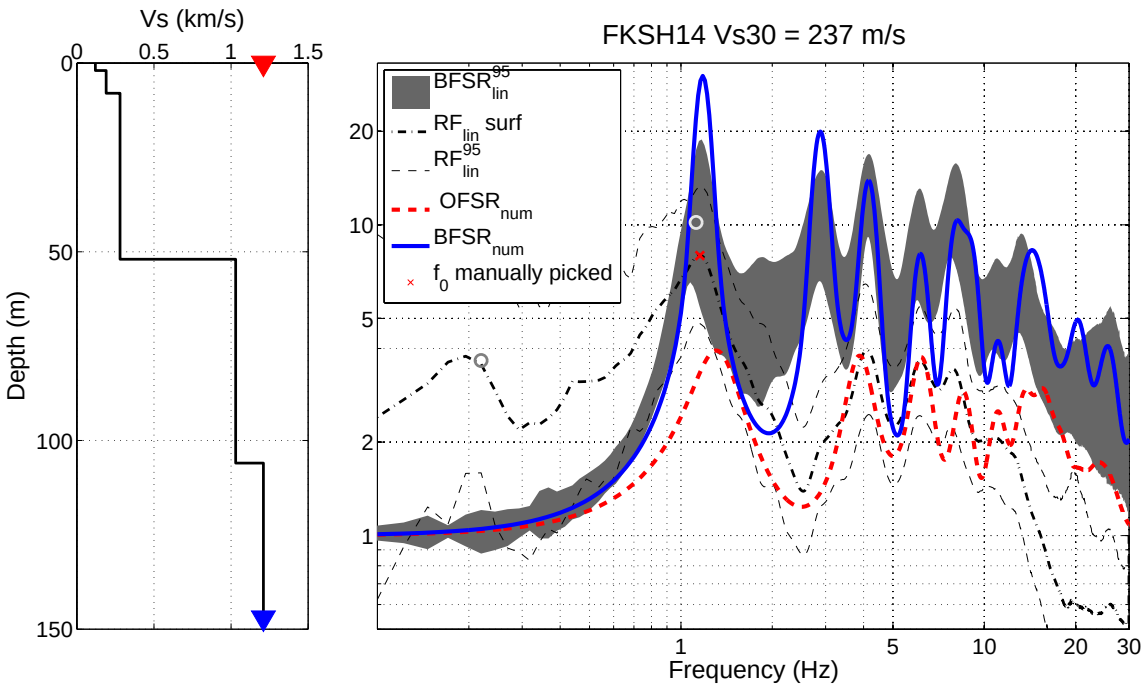


Figure C.5: *Similar to the previous Figure, but at KiK-net site FKSH14*

C.3 CONCLUSION AND DISCUSSION

In this annex, we have studied the effects of down-going waves, and especially their effects for site responses for the calculation of the fundamental resonance frequency. We show the difficulties when working with borehole data, and therefore propose an alternative to picking this frequency of the receiver function at the surface. We compared the automatic and manual picking of the fundamental resonance frequency on the receiver function at the surface. The manual picking is helped by the automatic picking, with possible corrections when the operator disagrees and finds that compared to the other site-response evaluations, another peak is more suitable for the fundamental resonance frequency of the site. For only 72 sites, the resonance frequencies differed. For the present study, we chose to use manual picking assisted by automatic picking.

Appendix D

Characteristics of KiK-net sites that have recorded at least two earthquakes with PGAs at depth >50 cm/s^2

Table D.1: *Site characteristics of KiK-net sites that have recorded at least two earthquakes with PGAs at depth >50 cm/s^2*

Station	Vs_{30}	f_{pred}	A_{pred}	f_0	B_{30}	PNL_{site}	PGA_{th}	Sh_{site}	fNL
	m/s	Hz		Hz		%	cm/s^2	Hz	Hz
AKTH04	459	4.8	14.7	3.4	0.43	10.4	48.4	0.6	5
FKSH09	585	13.6	16.2	12.9	0.96	13	36.8	2.3	9.1
FKSH11	240	10	8.5	0.4	0	9.3	54.6	0.3	1.5
FKSH12	449	4.3	42.5	4.1	0.51	27.5	19.4	0.7	3.7
FKSH14	237	4.1	12.4	1.2	0.22	16.7	32.7	0.2	1
FKSH21	365	3.9	12.6	2.7	0.25	4	124.1	0.3	3.6
IBRH07	107	0.6	116.4	0.2	0.28	10.6	47.9	0	0.4
IBRH13	335	11.3	25.3	2.6	0.26	16.6	25.8	1.4	2.3
IBRH14	829	14.3	24.1	13.3	0.71	25.8	23.2	2.3	11.9
IBRH20	244	0.3	10.1	0.2	0.19	5.3	98.3	0	2.6
ISKH02	721	2.4	4.8	1.3	0.18	7.1	89.4	0.1	16.4
ISKH04	444	3.1	5.7	0.9	0.01	1.5	NaN	0.1	7.4
IWTH04	456	3.4	20.2	2.9	0.53	12.3	41.7	0.4	2.5
IWTH05	429	12.8	15	3.2	0.36	7.3	81.7	0.6	3.1
IWTH15	338	7.9	5	0.3	0.35	4.6	130	0.2	6.5
IWTH17	1270	25.1	8.5	9.4	0.39	0.3	NaN	0	16.9
IWTH18	892	11.3	20.9	9.2	0.69	8.7	56.6	2.1	8.7
IWTH19	482	7.4	11.7	7.5	0.43	5.7	78.2	0.2	6.1
IWTH20	289	6.7	3.8	0.4	0.08	1	155.5	0.1	5.9
IWTH21	521	5.3	12.6	6.1	0.73	8.3	67.2	0.7	4.7
IWTH22	532	7.8	18.9	7.6	0.7	15.7	30.5	1.4	6.1
IWTH23	923	14.1	15.1	12.9	0.62	18.5	26.6	1.9	10.3
IWTH24	486	15.9	4.9	0.2	0.22	1.1	NaN	0.1	12
IWTH25	506	3	7.7	2.5	0.08	7.2	NaN	0.2	12.7
IWTH26	371	10.1	23.3	9.2	0.42	6	NaN	1.1	7.4
IWTH27	670	7.2	24.8	7.1	0.81	5.3	116.8	0.5	7.6
KSRH02	219	3.7	5.6	0.2	0.58	3.9	NaN	0.6	2.7

Table D.2: *Continuation*

Station	Vs_{30}	f_{pred}	A_{pred}	f_0	B_{30}	PNL_{site}	PGA_{th}	Sh_{site}	fNL
	m/s	Hz		Hz		%	cm/s^2	Hz	Hz
KSRH03	250	8.6	12.3	3.2	0.38	3.5	134.1	0.6	1.6
KSRH04	189	11.4	6.3	0.2	0.44	3.8	NaN	0.1	2.3
KSRH05	389	9.7	15.3	9.2	0.54	5.1	294.7	0.7	7.8
KSRH06	326	12	8.4	7.2	0.33	0.2	NaN	1.8	6.2
KSRH10	213	1.7	18.7	1.6	0.32	9.4	54.5	0.3	1.5
MIEH10	422	17.4	5.4	4.1	0.33	2.5	133.4	0	12.2
MYGH02	399	5.5	11.8	5.6	0.33	12	35.3	0.8	0.6
MYGH03	934	17.2	6.7	8.7	0.59	1.8	NaN	0	16.9
MYGH04	850	15	15.8	12.7	0.71	16.7	34.2	2.5	10.2
MYGH05	305	14.86	9.9	0.2	0.11	3.6	522.12	0.8	10.66
MYGH05	305	14.9	9.9	0.2	0.02	3.6	522.1	0.8	10.7
MYGH06	593	1.9	3.3	NaN	0.15	0.3	NaN	0	16.9
MYGH08	203	14.7	4.3	0.7	0.41	1.9	NaN	0	12.4
MYGH09	358	14.3	7.7	0.4	0.22	4.9	NaN	1	8.9
MYGH10	348	10.7	7.4	0.3	0.09	7.8	61.5	0.8	7.9
MYGH11	859	10.3	12.5	3.3	0.74	7.7	60.7	2	8.3
MYGH12	748	20.3	7.9	10.6	0.73	0.6	NaN	0	16.2
NGNH29	465	6.9	10.4	1.2	0.49	7.9	61.7	0.5	5.3
NIGH06	336	4.2	16	3	0.74	21.1	28.6	1.4	2.7
NIGH09	463	8	10.8	0.7	0.37	6.6	106	1.1	5.9
NIGH11	375	14.1	6	14	0.08	1.3	NaN	0.2	16.9
NIGH12	553	5	6.3	1.4	0.27	3	NaN	0.2	12
NIGH13	461	2.4	7.9	2.7	0	4.9	89.9	0.1	16.9
NIGH14	438	8.9	6.6	0.5	0.35	6.1	64.2	1.2	9.5
NIGH15	686	2.7	8.5	2.4	0.22	0.2	NaN	0.1	2.3
NMRH03	190	8.7	5.5	8.7	0.18	5.5	NaN	0.1	1
NMRH04	168	6.8	6.1	NaN	0.31	6.8	66.9	0.4	3.3
YMTH01	328	0.8	3.7	0.4	0.12	3	NaN	0.1	5.4