HAL
open science

Apport des enregistrements de séismes et de bruit de fond pour l'évaluation site-spécifique de l'aléa sismique en zone de sismicité faible à modérée

Vincent Perron

To cite this version:

Vincent Perron. Apport des enregistrements de séismes et de bruit de fond pour l'évaluation site-spécifique de l'aléa sismique en zone de sismicité faible à modérée. Sciences agricoles. Université Grenoble Alpes, 2017. Français. NNT : 2017GREAU020 . tel-01687070

HAL Id: tel-01687070

https://theses.hal.science/tel-01687070

Submitted on 18 Jan 2018

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

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

Communauté UNIVERSITÉ Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES

Spécialité : **Sciences de la Terre, de l'Univers et de l'Environnement**

Arrêté ministériel : 25 mai 2016

Présentée par

Vincent PERRON

Thèse dirigée par **Pierre-Yves BARD** et
co-encadrée par **Fabrice HOLLENDER** et **Céline GELIS**

préparée au sein de l'**Institut des Sciences de la Terre (ISTerre)**
dans l'**École Doctorale Terre, Univers, Environnement**

Apport des enregistrements de séismes et de bruit de fond pour l'évaluation site-spécifique de l'aléa sismique en zone de sismicité faible à modérée

Thèse soutenue publiquement le **15 septembre 2017**,
devant le jury composé de :

Michel CAMPILLO
Professeur à ISTerre, Grenoble, Président
Étienne BERTRAND
Ingénieur-Chercheur au Cerema, Nice, Rapporteur
Fabrice COTTON
Professeur à GFZ, Potsdam, Rapporteur
Olivier SÈBE
Ingénieur-Chercheur au CEA-LDG, Paris, Examinateur
Jérôme VERGNE
Physicien à l'EOST, Strasbourg, Examinateur
Pierre-Yves BARD
Ingénieur général des ponts des eaux et des forêts à ISTerre, Grenoble, Directeur de Thèse
Céline GÉLIS
Ingénieur-Chercheur, IRSN, Fontenay-aux-Roses, Co-encadrante de Thèse
Fabrice HOLLENDER
Ingénieur-Chercheur, CEA, Cadarache, Co-encadrant de Thèse

Remerciements

Les projets de recherche ont généralement la particularité d'engendrer chez les personnes qui les mènent une alternance – presque cyclique – de sentiments de stress et de satisfaction, dont le tracé pourrait faire croire à s'y méprendre à un sismogramme. Dans cette analogie à la sismologie, le thésard est l'instrument par lequel sont retranscrites les interactions avec tous les acteurs de la thèse, qu'elles soient perceptibles ou non. J'adresserai donc mes remerciements depuis mes sources d'inspirations situées en champ proche jusqu'à celles situées en champ lointain. Ceux-ci seront à n'y pas manquer incomplets, peut-être même maladroits, mais toujours sincères. Je prends ici le risque de remerciements nominatifs, et demande à quiconque aurait été oublié de bien vouloir m'en excuser.

En champ proche, viennent tout d'abord mes encadrants de thèse. À commencer par Fabrice Hollender qui m'a encadré le plus clair de la durée de cette thèse et même au-delà. Je ne commenterai pas son flegme ou sa patience, mais je saluerai tout particulièrement sa générosité, son enthousiasme, sa passion communicante pour la recherche et son leadership. Céline Gélis qui m'a fait le plaisir de m'encadrer pendant six mois dans les locaux du BERSSIN et le reste du temps à distance, pour sa gentillesse et sa patience. Et enfin, j'ai eu le privilège d'être l'un des derniers étudiants soumis à la direction de Pierre-Yves Bard, que je remercie pour sa bienveillance et pour avoir été aussi inspirant.

À distance régionale, je tenterai de citer tous les scientifiques avec qui j'ai eu l'occasion de collaborer, que ce soit au travers de projets communs ou simplement lors de discussions informelles. Je remercie en premier lieu les membres de mon jury, Michelle Campillo pour sa présidence, les rapporteurs Fabrice Cotton et Étienne Bertrand pour leur efficacité, ainsi que les examinateurs Olivier Sèbe et Jérôme Vergne pour leurs questions très pertinentes. Un grand merci à toutes les personnes du CEA et plus particulièrement à mes collègues du SA2S, à Catherine Berge-Thierry, Clément Desbordes, Alain Dujardin, Leatitia Foundotos, Cédric Guyonnet-Benaize, Bruno Hernandez, Jean Letort, Marc Nicolas et Sarah Sbaa. Merci à tous mes collègues du BERSSIN avec entre autres Christophe Clément, Marc Cushing, Élise Delavaud, Bérénice Froment, Sébastien Hok, Maria Lancieri et Oona Scotti. Merci aux chercheurs d'ISTerre et plus spécifiquement à Emmanuel Chaljub, Cécile Cornou, Philippe Guéguen, Afifa Imtiaz, Armand Mariscal, Emeline Maufroy et Philippe Roux. Et

finalement merci à tous les chercheurs des instituts auxquelles je n'étais pas rattaché mais avec lesquels j'ai échangé, dont Gabriele Ameri, David Baumont, Régis Cottereau, Giovanna Cultrera, Stéphane Drouet, Flora Garofalo, Ianis Gaudot, Chris Van Houtte, Hiroshi Kawase, Olga-Joan Ktenidou, Aurore Laurendeau, Fernando Lopez-Caballero, Jessie Mayor, Deborah Sicilia, Angkeara Svay, Nikos Theodoulidis et Paola Traversa.

À distance télésismique, j'aimerais mentionner tous les chercheurs, qu'ils soient illustres ou non, qui ont apporté leur pierre à l'édifice de la connaissance. Ceux qui par leur travail désintéressé ont participé au progrès de la société à des périodes, toujours d'actualité, de défiance vis-à-vis de la science et plus généralement de la notion de vérité.

Je finirai par mes proches, qui n'ont pas interagi directement sur le contenu scientifique du travail de thèse, mais qui participent grandement à la « maintenance » du thésard lui-même. D'abord mes parents, mon frère, et maintenant mon neveu pour leurs affections. Sans oublier celle qui partage ma vie et donc également les inéluctables turbulences de la thèse. Merci pour leur soutien et leur relecture attentive du manuscrit. Enfin, merci à tous mes amis qui étaient là, ceux d'avant la thèse et ceux que j'ai rencontrés pendant ma thèse à Cadarache, à Paris et à Grenoble. Merci à mes encadrants et à tous les collègues qui m'ont offert leur amitié, pour tous ces bons moments que l'on a partagés.

Abstract

Site effects can greatly increase both the duration and the amplitude of the seismic solicitation imposed on structures. 2D-3D site effects induce broadband amplifications that cannot be simulated up to high frequency (>2-4 Hz) due to the limited resolution of the geological, geophysical and/or geotechnical information. Empirical site effect assessment is therefore essential for reliable observations of this complex phenomenon up to high frequency. However, such assessments often require good quality records from many earthquakes that cannot be rapidly obtained in low seismicity areas.

This work presents a comparative analysis of these empirical evaluations on two very different sites, the first in a moderate seismicity context (Provence, France) and the second in a very active context (Kefalonia, Greece). For the Provençal site, nearly 500 earthquakes were recorded in only 2½ years thanks to the use of velocimeters. The site attenuation parameter κ_0 could thus be measured both on the acceleration (κ_{0_AS}) and displacement (κ_{0_DS}) spectra. Our results show that the measurement of κ_0 is relatively reliable on rock sites only, mainly due to the too great disturbance by the amplification for sedimentary sites, even for those that are quite stiff. The standard spectral ratio (SSR) approach provides the relative site effects from the numerous weak motion recordings available at sites located in Provence and Kefalonia. The results show a strong epistemic variability due to the lighting induced by the position of the seismic source with respect to the basin. Thus, a reliable site response estimation requires many earthquakes evenly distributed around the site when the geometry is 2D-3D. The mean SSR results are then compared with those obtained from methods using the ambient noise. As expected, the H/V spectral ratio approach (HVSR) provides only the fundamental frequency for some sites. Conversely, SSR applied to ambient noise (SSRn) shows very similar results to the SSR method up to high frequency, provided that the reference site is taken in the sedimentary basin. Approaches using the noise correlation (coherence and ANIRF) reveal that the transfer function relative to a rock reference site can be estimated at least up to low frequency (<4 Hz). Methods using ambient noise are promising for a rapid evaluation of the site response and its spatial variability (microzonation), even when seismicity is low. Empirical site effects methods are therefore applicable everywhere. They provide an essential complement to numerical approaches, which remain inevitable when the coverage of available earthquakes is not homogeneous or when soils are likely to present non-linear behaviors.

These results led us to propose a methodology for the evaluation of the site-specific seismic hazard, which consists of three main stages: (i) adjustment to the rock reference site of the ground motion prediction equations (GMPEs) using, among others, κ_{0_DS}; (ii) Reliable evaluation of the SSR transfer function between at least one site in the basin and the reference site (for which the GMPEs were adjusted); (iii) carrying out this assessment from the ANIRF in the less seismically active regions and/or extension to the entire study area with the SSRn. This methodology allows a much better consideration of the site effects (especially 2D-3D) and a significant reduction of the uncertainties in the evaluations of the site-specific seismic hazard. It only required, simultaneous recording of earthquakes for at least two sites equipped with sensible velocimeters, and to carrying out temporary campaign of ambient noise measurements, if necessary.

Keywords: Seismic hazard, site-specific, site effects, low seismicity, empirical approach, kappa, ambient noise

Résumé

Les effets de site peuvent augmenter fortement la durée et l'amplitude des sollicitations sismiques imposées aux structures. Les effets de site 2D-3D induisent des amplifications sur une large-bande de fréquence qui ne peuvent pas être simulées numériquement jusqu'à haute fréquence (>2-4 Hz) du fait de la limite de résolution des connaissances géologiques, géophysiques et/ou géotechniques du sous-sol. Les évaluations empiriques des effets de site sont donc indispensables pour pouvoir observer ce phénomène complexe de façon fiable jusqu'à haute fréquence. Cependant, de telles évaluations nécessitent souvent un enregistrement de bonne qualité de nombreux séismes ce qui rend leur obtention rapide difficile dans les régions faiblement actives.

Ce travail présente une analyse comparative de ces évaluations empiriques sur deux sites très différents, l'un en contexte sismique modéré (Provence, France) et l'autre en contexte très actif (Céphalonie, Grèce). Sur le site provençal, ce sont tout de même près de 500 séismes qui ont pu être enregistrés en l'espace de 2½ ans grâce à l'utilisation de vélocimètres. Une mesure du paramètre d'atténuation de site κ_0 a ainsi pu être réalisée via l'analyse des spectres en accélération (κ_{0_AS}) et en déplacement (κ_{0_DS}). Nos résultats montrent que la mesure de κ_0 n'est relativement fiable que sur les sites au rocher, principalement du fait de la perturbation trop importante de l'amplification sur les sites sédimentaires même assez raides. La méthode des rapports spectraux sismiques (SSR) permet l'évaluation des effets de site relatif à partir des nombreux enregistrements de mouvement faible disponibles sur les sites de Provence et de Céphalonie. Les résultats montrent une forte variabilité épistémique attribuable à l'éclairage induit par la position de la source sismique vis-à-vis du bassin. Ainsi, une estimation fiable de la réponse des sites à géométries 2D-3D nécessite l'enregistrement d'un grand nombre de séismes présentant une répartition spatiale homogène. Les résultats SSR moyens sont ensuite comparés à ceux issus des méthodes utilisant le bruit ambiant. Conformément aux attentes, la méthode des rapports spectraux H/V (HVSR) ne permet que de caractériser la fréquence fondamentale pour certains sites. À l'inverse, la méthode SSR appliquée au bruit ambiant (SSRn) montre des résultats très comparables à ceux de la méthode SSR jusqu'à haute fréquence à la condition que le site de référence soit pris dans le bassin sédimentaire. Les approches par corrélation de bruit (cohérence et ANIRF) révèlent qu'une estimation de la fonction de transfert relative à un site de référence au rocher est possible au moins jusqu'à basse fréquence (<4 Hz). L'utilisation du bruit ambiant permet ainsi d'envisager l'évaluation rapide

de la réponse des sites et de sa variabilité spatiale (microzonage), même lorsque la sismicité est faible. Les méthodes empiriques d'évaluation des effets de site sont donc applicables partout. Elles apportent un complément essentiel aux approches numériques, qui restent indispensables lorsque la couverture des séismes disponibles n'est pas homogène ou que les sols sont susceptibles de présenter des comportements non-linéaires.

Ces résultats permettent de proposer une méthodologie d'évaluation de l'aléa sismique site-spécifique qui se décompose en trois étapes principales : (i) ajustement des équations de prédictions de mouvement du sol (GMPEs) sur le site de référence au rocher au moyen, entre autres, de κ_{0_DS} ; (ii) évaluation empirique fiable de la réponse d'au moins un site dans le bassin relativement au site de référence (pour lequel les GMPEs ont été ajustées) par la méthode SSR ; (iii) réalisation de cette évaluation à partir de l'ANIRF dans les régions les moins actives sismiquement et/ou extension à toute la zone d'étude à partir du SSRn. Cette méthodologie permet d'envisager une bien meilleure prise en compte des effets de site (en particulier 2D-3D) et une réduction importante des incertitudes dans les évaluations de l'aléa sismique spécifique à un site. Elle nécessite simplement l'enregistrement simultané de séismes sur au moins deux sites équipés de vélocimètres sensibles, et la réalisation de campagnes de mesure temporaire du bruit ambiant, si nécessaire.

Mots-clés : Aléa sismique, site-spécifique, effets de site, faible sismicité, méthode empirique, kappa, bruit ambiant

Table des matières

REMERCIEMENTS **3**

ABSTRACT **5**

RESUME **7**

TABLE DES MATIERES **9**

INTRODUCTION GENERALE **13**

CHAPITRE 1 CONCEPT, TRAITEMENTS ET DONNEES **19**

1.1. LES EFFETS DE SITE 19
1.1.1. Définition 19
1.1.1.1. Phénoménologie 19
1.1.1.2. Quantifications et prédictions 23
1.1.2. L'évaluation empirique des effets de site 25
1.1.2.1. Contexte historique 25
1.1.2.2. SSR 26
1.1.2.3. SSRn 27
1.1.2.4. HVSR 29
1.1.2.5. HVSRs 30
1.2. TRAITEMENT DES DONNEES 31
1.2.1. Traitements spécifiques aux séismes 31
1.2.1.1. Extraction des séismes depuis les enregistrements continus 31
1.2.1.2. Pointé des premières arrivées 33
1.2.1.3. Sélection des fenêtres de bruit et des phases du signal 35
1.2.2. Traitements applicables au champ sismique ambiant 37
1.2.3. Évaluation des spectres 37
1.3. SITES D'ETUDE 40
1.3.1. Provence 40
1.3.2. Argostoli 45
1.4. L'ENREGISTREMENT SISMIQUE EN ZONE DE SISMICITE FAIBLE A MODEREE 49

CHAPITRE 2 ÉVALUATION DE KAPPA EN ZONE DE SISMICITE FAIBLE A MODEREE **53**

2.1. CONTEXTE 53
2.2. ABSTRACT 54

2.3. INTRODUCTION 54
2.4. STUDY AREA AND DATASETS 58
2.4.1. Study area 58
2.4.2. Spectrum computation 61
2.4.3. Velocimeter versus accelerometer 62
2.4.4. Site amplification 63
2.5. KAPPA 65
2.5.1. Data processing 65
2.5.2. Results and comparison between κAS and κDS 67
2.6. ANALYSIS OF THE SENSITIVITY AND ROBUSTNESS OF K TO VARIOUS PARAMETERS 71
2.6.1. Measurement uncertainty Δκr 71
2.6.2. Frequency-dependence of the attenuation 72
2.6.3. Site amplification dependence of κ 73
2.6.4. Source dependence of κ 78
2.7. DISCUSSION 80
2.7.1. Reliability of the κ0 measurements 81
2.7.2. Variability of κr measurements 82
2.7.3. The κDS approach 84
2.8. CONCLUSIONS 85
2.9. DATA AND RESOURCES 87
2.10. ACKNOWLEDGMENTS 87

CHAPITRE 3 ÉVALUATION DE LA REPONSE DES SITES A PARTIR DES ENREGISTREMENTS DE MOUVEMENTS FAIBLES ET INCERTITUDES ASSOCIEES 89

3.1. INTRODUCTION 89
3.2. BASE DE DONNEES DE MOUVEMENT FAIBLE 90
3.1. RESULTATS SSR 90
3.2. ANALYSES STATISTIQUES 96
3.3. INCERTITUDE EPISTEMIQUE 101
3.3.1. Influence du back-azimut 102
3.3.2. Influence de la distance épicentrale 105
3.3.3. Influence de la profondeur 108
3.3.4. Influence de la magnitude 109
3.4. DISCUSSION SUR LES INCERTITUDES 111
3.5. CONCLUSIONS 115

CHAPITRE 4 ÉVALUATION DE LA REPONSE DES SITES A PARTIR DU BRUIT AMBIANT 119

4.1. CONTEXTE 119
4.2. ABSTRACT 119
4.3. INTRODUCTION 120
4.4. METHODS DEFINITION AND BACKGROUND 123
4.4.1. Spectrum analysis 123
4.4.2. HVSR 124
4.4.3. SSRn 124
4.4.4. SSRh 126

4.4.5. Cross-correlation or coherency *128*
4.4.6. ANIRF *130*
4.5. DATA AND PROCESSING 132
4.5.1. Provence site array *132*
4.5.2. Argostoli array *133*
4.5.3. Processing *135*
4.6. MICROTREMORS PROCESSING WITH AMPLITUDE-BASED METHODS (HVSR, SSRN AND SSRH) 136
4.6.1. HVSR *137*
4.6.2. SSRn *137*
4.6.3. SSRh *137*
4.6.4. A0 ..... *140*
4.6.5. Discussion *141*
4.7. MICROTREMORS PROCESSING WITH COHERENCY AND ANIRF METHODS 144
4.7.1. Phase of the reconstructed wavefields *144*
4.7.2. Amplitude of the reconstructed wavefields *146*
4.7.3. Site transfer function assessment from the ANIRF *146*
4.8. INVESTIGATING THE PARAMETERS THAT CONTROL THE SITE RESPONSE ASSESSMENT FROM MICROTREMORS 149
4.9. GENERAL CONCLUSIONS 153
4.10. ACKNOWLEDGMENTS 154

CONCLUSIONS ET PERSPECTIVES 157

BIBLIOGRAPHIE 163

ANNEXE 1 SELECTING TIME WINDOWS OF SEISMIC PHASES AND NOISE FOR ENGINEERING SEISMOLOGY APPLICATIONS: A VERSATILE METHODOLOGY AND ALGORITHM 181

ANNEXE 2 ACCELEROMETER, VELOCIMETER DENSE-ARRAY AND ROTATION SENSOR DATASETS FROM THE SINAPS@ POST-SEISMIC SURVEY (2014-2015 AFTERSHOCK SEQUENCE IN CEPHALONIA, GREECE) 197

1.1. ABSTRACT 198
1.2. INTRODUCTION 198
1.3. INSTRUMENT DEPLOYMENT AND DETAILS 201
1.4. OVERALL DATA PREPARATION, QUALITY AND AVAILABILITY 203
1.5. INITIAL OBSERVATIONS AND RESULTS 207
1.6. SUMMARY 208
1.7. DATA AND RESOURCES 209
1.8. ACKNOWLEDGEMENTS 209
1.9. REFERENCES 209

Introduction Générale

Les tremblements de terre sont la première cause de mortalité par catastrophe naturelle de ce début de siècle (https://www.unisdr.org/archive/12470). Le séisme de 2010 en Haïti compte à lui seul plus de 300 000 décès. Il n'existe à ce jour aucun indicateur permettant une prédiction fiable de la survenue de tels évènements. Ainsi, la seule protection efficace face aux tremblements de terre est la réalisation de bâtiments et d'infrastructures résistantes aux séismes (constructions parasismiques). Le dimensionnement parasismique n'est possible que si l'amplitude des mouvements du sol à laquelle sont susceptibles d'être soumises les structures est connue. Lors d'un séisme, le mouvement du sol enregistré en un point d'intérêt A dépend directement de l'énergie sismique libérée par la source en direction de A, de la propagation (atténuation, dispersion) de l'onde le long de son trajet en profondeur depuis le foyer sismique et des effets locaux imputables aux structures géologiques avoisinants le point A, communément réunis sous le terme d'effets de site (Figure 1). Ces derniers peuvent être particulièrement importants pour les remplissages sédimentaires de surface (vallée ou bassin sédimentaire) qui par leur géométrie et leur contraste d'impédance augmentent significativement l'amplitude et la durée de la sollicitation sismique à certaines fréquences. De nombreuses mégapoles à travers le monde ont été construites sur des bassins sédimentaires présentant des effets de site important (Mexico City, Mexico ; Los Angeles, USA ; Tokyo, Japon, etc.). Lors d'évènements majeurs, ceux-ci peuvent avoir des conséquences désastreuses comme lors du séisme de Michoacán (Mexique) de 1985 qui est responsable d'environ 11 000 décès principalement dans la capitale alors même que celle-ci se situait à plus de 400 km de l'épicentre. Des effets de site très importants y ont été identifiés et sont reconnus comme étant la cause de la destruction de nombreux bâtiments présentant souvent une fréquence de résonance proche de celle du bassin. La prédiction fiable des effets de site est donc essentielle, mais en pratique reste difficile à réaliser du fait de la complexité et de la variabilité du phénomène.

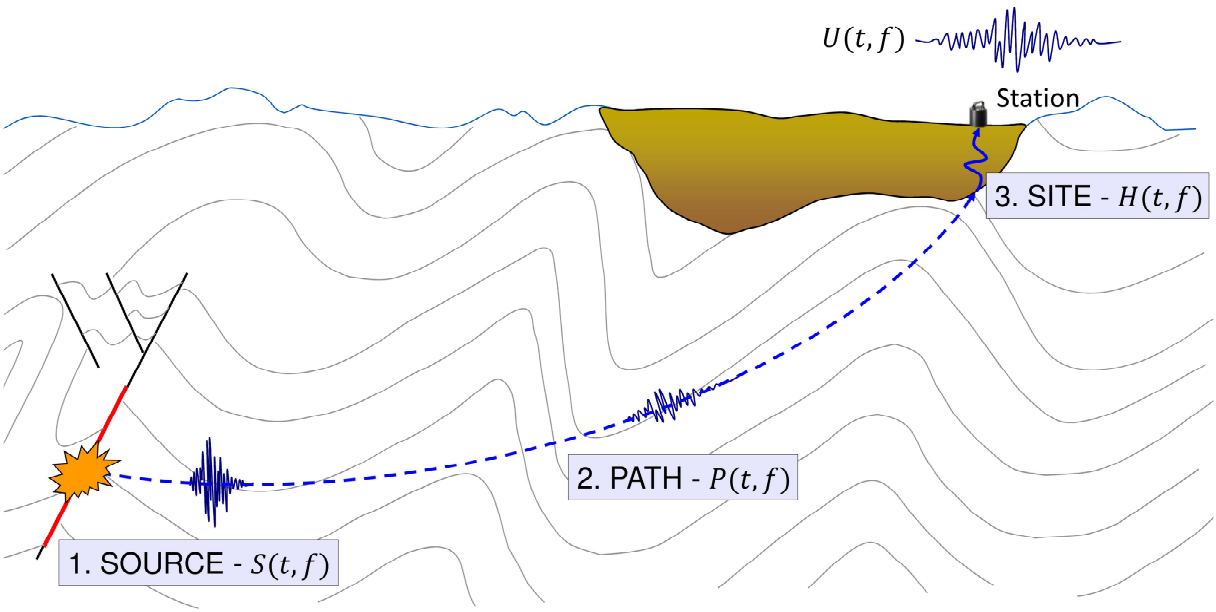


Figure 1. Illustration des différents termes du mouvement sismique : source, trajet et site (adaptée depuis Berge-Thierry et al. 2016).

Outre l'impact de la réduction d'impédance du milieu qui accroit directement l'amplitude de l'onde, les variations de la géométrie des structures géologiques ainsi que de la topographie de surface libre suivant 1, 2 ou 3 dimensions peuvent induire des phénomènes de focalisation/défocalisation de l'énergie sismique, de résonance, de génération d'ondes de surface sur les bords de bassin et de diffraction, réfraction et changement de mode des ondes de volume au niveau des interfaces qui complexifient grandement le champ d'ondes sismiques. Il en résulte une variabilité spatiale et fréquentielle de la réponse sismique des sites (variabilité site-à-site) qui est généralement très importante et directement liée aux propriétés des structures géologiques avoisinant le site. Pour un même site, les variations dans l'éclairage des structures géologiques par l'onde incidente suivant la position de la source sismique induisent une variabilité inter-évènements qui peut être très importante également. Cette dernière est accrue pour certains sites dont les sols (généralement très mous) subissent des déformations plastiques lorsqu'ils sont soumis à une forte sollicitation. On parle alors de réponse non-linéaire des sites par opposition aux comportements linéaires des sols qui sont soumis à des mouvements faibles.

Le risque sismique résulte de la combinaison d'un enjeu (villes, hôpitaux, barrages, centrales nucléaires, etc.) et d'un aléa. L'évaluation de l'aléa sismique est fondée sur l'hypothèse que les futurs séismes auront les mêmes propriétés de répartition spatio-temporelles que la sismicité passée. Une fois que les séismes passés à portée de connaissance ont été identifiés (sismicité historique et instrumentale), deux types d'approche sont classiquement utilisés pour évaluer l'aléa sismique : l'approche déterministe et l'approche probabiliste. La première consiste à réaliser des scénarios pour des séismes de référence dont on cherche à prédire quels seront les mouvements du sol qu'ils induiront sur le site d'étude. La seconde, a

priori mieux adaptée aux études de risque, s'intéresse à la probabilité qu'un niveau de mouvement de sol soit dépassé pour un site donné et sur une période de temps fixée, en se basant sur l'ensemble des séismes connus. Quelle que soit l'approche privilégiée, pour chaque combinaison de site et de séisme, la prédiction des mouvements du sol nécessite la caractérisation de l'énergie émise à la source ainsi que l'estimation de l'atténuation de l'onde le long de son trajet et des effets de site. Ceci est le plus souvent effectué via les équations de prédiction de mouvements du sol (GMPEs). Les GMPEs sont des modèles empiriques déduits de très nombreux enregistrements réalisés pour des configurations de sources, de trajets et de sites variées qui permettent la prédiction d'indicateurs associés aux mouvements du sol à partir de quelques paramètres de source, d'atténuation, et parfois de site. Lorsque les conditions de site sont prises en compte par les GMPEs, elles le sont au travers de proxys généralement très simples comme la vitesse moyenne des ondes S dans les 30 premiers mètres de sol (V_{S30}), parfois l'épaisseur des sédiments, et encore plus rarement la fréquence fondamentale du site (f_0), ou l'atténuation haute fréquence des ondes S au niveau du site (κ_0), etc.. Parfois même, en l'absence de toute mesure *in situ*, seules la géologie et la topographie sont prises en compte ; la géologie permettant de se référer aux classes de sol établies empiriquement et la pente des terrains ayant montré une certaine corrélation avec V_{S30}. De plus, ces évaluations statistiques sont difficilement réalisables jusqu'aux fortes magnitudes (M) dans les zones faiblement sismiques où les séismes majeurs (M>6) présentent des périodes de récurrence supérieure au siècle voire au millénaire. Ainsi, ces équations sont généralement établies dans des régions fortement sismiques (États-Unis, Japon, Grèce, etc.) puis ajustées dans les autres régions (ajustement host-to-target) à l'aide des proxys tenant compte des conditions de site (V_{S30}, f_0, κ_0, etc.).

Par exemple, en France métropolitaine depuis le 1er mai 2011 et l'entrée en vigueur de la nouvelle règlementation parasismique européenne (Eurocode 8), un nouveau zonage sismique est pris en compte pour le bâti conventionnel. Il repose sur une carte d'aléa sismique probabiliste « au rocher standard » (V_{S30} = 800 m/s) dont l'objectif est de décrire le mouvement sismique ayant 10 % de probabilité d'être atteint ou dépassé sur une période de 50 ans. Les effets de site sont ensuite implémentés au travers d'un coefficient défini en fonction de la classe du sol sur lequel le bâtiment à dimensionner est fondé (Legifrance 2010). Pour les installations nucléaires de base, l'évaluation de l'aléa sismique est réalisée suivant la règle fondamentale de sureté (ASN 2001) qui préconise une approche déterministe. La V_{S30} du sol sous l'installation permet alors de définir la GMPE à prendre en compte suivant que celle-ci est supérieure ou inférieure à 800 m/s. Cette règlementation impose également la réalisation d'études spécifiques lorsque les effets de site sont 2D-3D ou que V_{S30} est inférieure à 300 m/s.

Lorsque la géométrie des structures géologiques est 2D-3D, les proxys de site ne peuvent pas retranscrire toute la complexité des effets de site. Des évaluations spécifiques pour chaque site d'intérêt particulier (ville, hôpitaux, barrages, centrales nucléaires, etc.) sont alors préférables et parfois obligatoires. Celles-ci peuvent être réalisées à partir d'observations *in situ* ou par modélisation. Une évaluation empirique fiable des effets de site tenant compte des variabilités site-à-site et inter-évènements nécessite l'enregistrement

depuis un grand nombre de stations d'une quantité statistiquement représentative de séismes présentant une distribution spatiale homogène dans une large gamme de magnitudes. En pratique, la mise en place de stations selon une grille dense d'une maille de l'ordre de quelques dizaines de mètres durant un temps suffisamment long pour permettre l'enregistrement de nombreux séismes est très rarement réalisée. De plus, la distribution des séismes n'est jamais homogène, mais s'organise le long des failles et la survenue de séismes de forte magnitude est beaucoup moins fréquente que celle de séismes de magnitude plus réduite selon une décroissance décrite par la loi de Gutenberg-Richter. La prédiction empirique fiable des effets de site associés à des évènements de grande magnitude est donc très rarement réalisable en un temps suffisamment court pour permettre le dimensionnement des structures, en particulier dans les régions faiblement sismiques.

Pour cette raison, dans ces régions sismiquement peu actives, l'évaluation des effets de site incluant d'éventuels effets non-linéaires est réalisée pratiquement exclusivement à partir de simulations numériques. Ces dernières bénéficient des avancées permanentes des performances de calcul pour implémenter des modèles de plus en plus complexes allant jusqu'à tenir compte de géométries 3D. Tandis que la résolution des mailles des modèles numériques s'affine de plus en plus, elle se heurte aujourd'hui à la limite de résolution des connaissances géologiques et géotechniques des milieux qui se restreignent pour les meilleurs cas à des modèles 3D déduits de la géophysique, des affleurements géologiques de surface et de quelques forages. Ainsi la fréquence maximale permise par ces simulations 3D est réduite le plus souvent à quelques Hertz seulement (<1-4 Hz). Celles-ci ne permettent donc pas de couvrir la bande de fréquence fondamentale de l'ensemble des structures et équipements à dimensionner. De plus, les différents benchmarks réalisés en aveugle par le passé sur des cas linéaires ont montré les erreurs importantes pouvant être introduites par de telles approches lorsqu'elles ne sont pas confrontées aux résultats empiriques.

L'évaluation empirique des effets de site est donc nécessaire pour réaliser une estimation large-bande fiable de la réponse linéaire spécifique à un site et est un préambule incontournable aux simulations non-linéaires. Dans les zones de sismicité faible à modérée, une telle approche empirique est cependant difficile à réaliser en un temps suffisamment court du fait de l'occurrence moins fréquente d'évènements permettant la réalisation d'enregistrements de bonne qualité présentant un rapport signal-sur-bruit (SNR) suffisant. Des solutions doivent être donc explorées pour pallier à cette difficulté. L'objectif de ce travail de thèse est de déterminer la faisabilité et la robustesse des évaluations empiriques des effets de site dans un contexte de sismicité faible à modérée. Ce travail s'applique d'abord à évaluer la quantité et la qualité des enregistrements de séismes qui sont réalisables sur un site industriel situé dans une petite vallée sédimentaire en Provence (sud-est de la France). Les données acquises dans le bassin sédimentaire de Koutavos près de la ville d'Argostoli dans une région sismiquement active (ouest de la Grèce), sont également utilisées pour confronter nos résultats. Plusieurs méthodes empiriques sont testées et comparées afin de vérifier leur fiabilité dans l'estimation de la réponse des sites à géométrie 3D et leur potentielle dans le contexte de sismicité faible à modérée. Ainsi, nous proposons une méthodologie tenant compte au mieux des effets de site dans les évaluations d'aléa

sismique site-spécifique qui est particulièrement adaptée aux régions faiblement sismiques. Celle-ci s'articule en trois étapes principales : (i) réalisation des ajustements host-to-target permise par la mesure du paramètre κ_0 en Provence ; (ii) utilisation des mouvements faibles (issus de séismes proches généralement de petites magnitudes ou plus distants et de magnitudes plus importante) pour l'évaluation de la réponse linéaire des sites ; (iii) exploitation du champ sismique ambiant (bruit ambiant) pour pouvoir envisager une évaluation rapide de la fonction de transfert des sites.

Ces trois points sont repris dans les chapitres 2 à 4 de ce manuscrit qui s'articule comme suit :

Le chapitre 1 rappelle d'abord la définition et la phénoménologie des effets de site avant de donner un aperçu de leurs quantifications et de leurs prédictions. Après un bref rappel du contexte historique qui a permis leurs émergences, les méthodes empiriques sont présentées une par une. Les traitements réalisés pour extraire les données sismiques depuis les enregistrements continus, pour le pointé des premières arrivées d'onde P et S, pour le fenêtrage des différentes phases du signal (P, S, Coda et tout le signal), pour sélectionner les fenêtres de bruit dans les méthodes de bruit ambiant et pour l'évaluation des spectres sont détaillés également. Enfin les sites d'études et les bases de données qui ont été constituées et utilisées durant ce travail sont décrits dans une dernière partie. Cette dernière partie entamera une première discussion sur la quantité et la qualité des enregistrements qui sont réalisables en zone de sismicité faible à modérée à partir de l'exemple d'un site d'étude situé en Provence dans le sud-est de la France.

Le chapitre 2 se concentre sur la mesure empirique du paramètre d'atténuation κ_0 sur le site de Provence. Bien que celui-ci est essentiel à l'ajustement des GMPEs, il est généralement déduit de corrélations très incertaines avec V_{S30}, faute de mieux dans les régions faiblement sismiques. Nous tentons ici une mesure de ce paramètre depuis l'approche classique sur les spectres en accélération (κ_{AS}) de séismes de magnitude élevée (M>3), et par la méthode utilisant les spectres en déplacement (κ_{DS}) de séismes de petite magnitude (M<1,5). La comparaison entre les valeurs moyennes et les incertitudes obtenues depuis ces deux approches nous permet de discuter la dépendance de la mesure de κ_0 à l'amplification de site, à la part dépendante en fréquence de l'atténuation et à une éventuelle composante de source. Ce chapitre fournit ainsi des éléments de discussion concernant la faisabilité et la fiabilité d'une mesure de κ_0 en zone de sismicité faible à modérée.

Le chapitre 3 réalise l'évaluation de la réponse linéaire des sites depuis les enregistrements de mouvements faibles via la méthode standard spectral ratio (SSR) ou site-sur-référence. Cette méthode permet une évaluation relative de la fonction de transfert entre le site d'étude et un site de référence supposé moins sujet aux effets de site (généralement situé sur un rocher dur). C'est une méthode de référence pour l'étude des effets de site qui est cependant difficile à exécuter rapidement dans les régions faiblement actives. Dans un premier temps, la loi de distribution statistique des résultats SSR est évaluée. Ensuite, la variabilité inter-évènements des résultats SSR est étudiée afin d'identifier les paramètres qui influencent la moyenne et l'écart-type des fonctions de transfert SSR. Ainsi, la part

d'incertitude épistémique et aléatoire associée à la variabilité inter-évènements peut être discutée dans la perspective de comprendre la qualité et la représentativité des évaluations SSR réalisées depuis des bases de données généralement pauvres issues de régions à sismicité faible ou modérée.

Le chapitre 4 teste des approches empiriques employant le bruit ambiant pour l'évaluation de la fonction de transfert moyenne des sites, telles que la méthode des rapports spectraux H/V (HVSR), la méthode SSR appliquée au bruit ambiant (SSRn) et les méthodes utilisant les corrélations de bruit (cohérence et ANIRF). Après une revue des travaux ayant été réalisés autour des différentes approches, nous comparons ces dernières sur nos sites d'étude et confrontons les résultats avec ceux obtenus depuis les SSR. Nous proposons et testons également une méthode hybride (SSRh) combinant les méthodes SSR et SSRn et se basant sur l'hypothèse d'une influence possible des conditions géologiques du/des site(s) sur le champ sismique ambiant. Enfin, nous tenterons de substituer l'utilisation des séismes dans la méthode SSR par celle de sources virtuelles obtenues avec l'ANIRF. Ceci est d'un intérêt fondamental pour les zones où la sismicité est trop faible pour autoriser l'enregistrement en un temps suffisamment court, d'un nombre satisfaisant de séismes pour l'application de la méthode SSR. De plus, les méthodes utilisant le bruit ambiant pourraient fournir un outil précieux pour l'évaluation rapide et économique des variations spatiales de la réponse des sites (microzonage) quelle que soit l'activité sismique de la zone que l'on cherche à cartographier.

Chapitre 1

Concept, Traitements et Données

1.1. Les effets de site

1.1.1. Définition

Les effets de site définissent les modifications du champ d'ondes sismiques incident induitent par les structures géologiques superficielles. La profondeur à laquelle fait référence le mot « superficiel » dépend du site considéré. Cette profondeur se rapporte couramment à celle de l'interface entre le substratum rocheux et les dépôts sédimentaires qui le surplombent ; interface marquant bien souvent le plus fort contraste d'impédance au niveau du site. Le substratum rocheux peut se trouver à des profondeurs de plusieurs kilomètres pour les bassins sédimentaires les plus épais, ou bien en surface lorsque les formations rocheuses affleurent à la surface libre. Dans ce dernier cas, la géométrie de la surface libre (effets topographiques) contrôle souvent la réponse des sites, à l'opposé des bassins sédimentaires où les effets liés aux propriétés rhéologiques des sols (effets lithologiques) prédominent. Ces deux types d'effets sont donc généralement traités séparément. Dans cette étude, qui s'intéresse principalement aux effets de site dans les bassins sédimentaires, nous n'aborderons pas les effets topographiques spécifiquement. La section qui suit fait une présentation succincte des effets de site lithologiques et est inspirée notamment des travaux d'Anderson (2007) et de Kramer (1996) auxquels le lecteur pourra se référer pour plus de détails.

1.1.1.1. Phénoménologie

Du fait de l'augmentation de la pression isostatique, la masse volumique (ρ) et le module élastique (μ) des sols augmentent avec la profondeur. Comme l'augmentation du module élastique est plus rapide que celle de la densité, la vitesse des ondes de cisaillement, $\beta =$

$\sqrt{\mu/\rho}$, augmente avec la profondeur. En moyenne, cette vitesse diminue donc dans un remplissage sédimentaire depuis le substratum rocheux jusqu'à la surface. Cette diminution de la vitesse vers la surface augmente, par conservation de l'énergie, l'amplitude des ondes qui se propagent. Par exemple, pour le cas d'une diminution graduelle de la vitesse et de la densité depuis une profondeur Z (défini par ρ_Z et β_Z) vers la surface (défini par ρ_0 et β_0) et en négligeant l'amplification de surface libre, le gain d'amplitude de l'onde de cisaillement entre la profondeur Z et la surface est donné par (p. ex. Aki and Richards 1980)

$$G_A = \sqrt{\frac{\rho_Z \beta_Z}{\rho_0 \beta_0}} \qquad 1$$

où le produit $\rho\beta$ représente l'impédance du milieu.

La diminution de la vitesse vers la surface permet également de prédire, au travers de la loi de Snell ($\sin j_1 / \beta_1 = \sin j_2 / \beta_2$), une incidence des ondes de volume à la surface de la Terre qui est proche de la verticale lorsque les sources sont lointaines. Ceci explique pourquoi il est courant de considérer une incidence verticale des ondes dans les évaluations d'effets de site. De même, le gradient de vitesse vers la surface libre, explique pourquoi une partie de l'énergie sismique peut être piégée à la surface de la Terre sous forme d'ondes de surface : les ondes de Love de polarité SH et les ondes de Rayleigh qui sont le couplage de modes P et SV. La génération et la propagation des ondes de surface sont développées en détail dans les ouvrages de sismologie théorique (p. ex. Aki and Richards 1980). Outre l'angle d'incidence de l'onde, la loi de Snell prédit également les coefficients de transmission (A_T) et de réflexion (A_R) d'une onde plane SH passant au travers d'une interface depuis un milieu 2 vers un milieu 1 d'impédance différente. L'équation

$$A_T = \frac{2\rho_2\beta_2}{(\rho_2\beta_2 + \rho_1\beta_1)} \qquad 2$$

prévoit ainsi dans le milieu 1 un doublement de l'amplitude de l'onde incidente lorsque $\beta_2 \gg \beta_1$ et une transmission nulle lorsque $\beta_2 \ll \beta_1$. Tandis que le premier cas explique une partie de l'amplification de l'onde SH lors de son entrée dans la couche sédimentaire (Figure 2 (a)), le deuxième cas implique que l'onde faisant le chemin inverse est majoritairement réfléchie ; c'est ce phénomène qui est à l'origine du piégeage des ondes sismiques dans les couches de faible impédance (Figure 2 (b)). Ce phénomène est cependant plus complexe puisque lorsque l'incidence de l'onde est oblique à l'interface, des conversions des ondes P vers les ondes SV et inversement apparaissent, mais nous ne les aborderons pas plus en détail ici.

Le piégeage de l'onde décrit précédemment induit que l'onde peut effectuer plusieurs allers-retours entre la surface libre et les interfaces présentes en profondeur. Ceci engendre le phénomène de résonnance qui est fonction de la géométrie des couches géologiques et des contrastes d'impédances aux différentes interfaces. Pour le cas 1D simple d'une onde S à

incidence verticale dans une couche horizontale d'épaisseur b_0 et de vitesse β, le phénomène de résonnance survient aux fréquences

$$f = \frac{(2k-1)\beta}{4b_0}, \quad avec\ k = 1, 2, \ldots \qquad 3$$

Cette approche est couramment utilisée pour prédire la valeur de la fréquence fondamentale du site (f_0) ainsi que celles de quelques harmoniques supérieures pour les géométries les plus simples à une dimension. Pour les géométries plus complexes (2D ou 3D), les résonances peuvent être liées à des contrastes d'impédances correspondant à des profondeurs variées. C'est pourquoi, pour ces géométries, l'amplification concerne une large-bande de fréquence avec des niveaux mesurés qui sont supérieurs à ceux prédits par les modèles 1D.

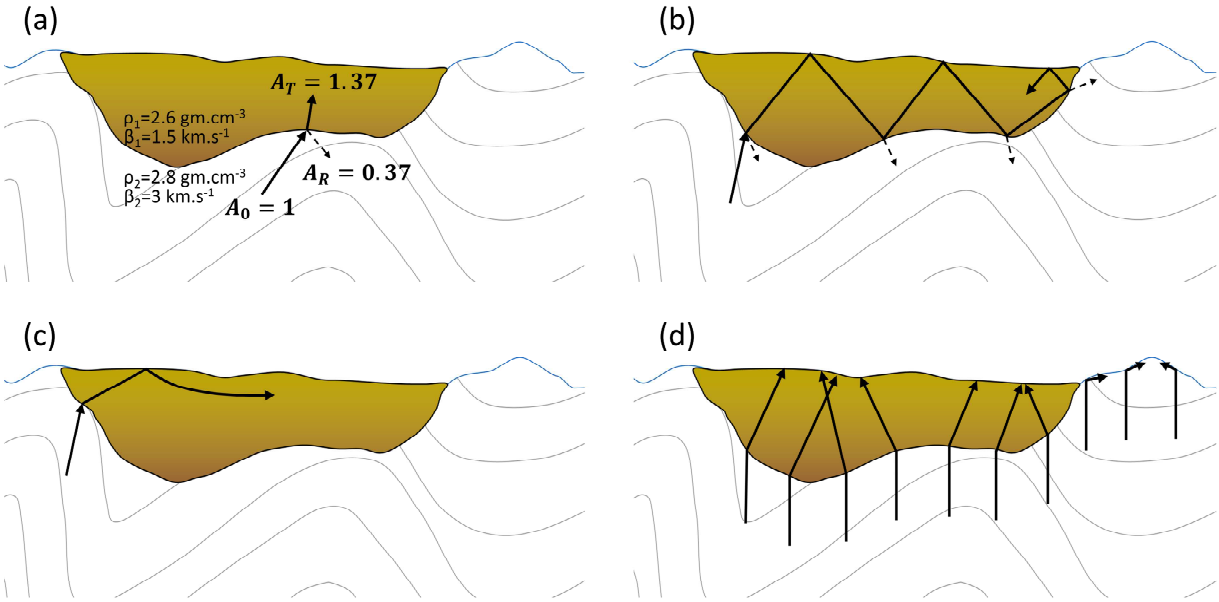


Figure 2. Illustration des processus d'effet de site dans un bassin dont le remplissage sédimentaire présente une impédance plus faible que le rocher qui l'entoure. (a) Onde SH incidente sur l'interface entre le rocher et le sédiment, avec une partie transmise et une partie réfléchie. (b) Piégeage d'une onde SH incidente dans un bassin sédimentaire, à l'origine des phénomènes de résonance à certaines fréquences. (c) Génération d'ondes de surface depuis les ondes de volume à incidence subverticale se propageant sur les bords inclinés du bassin. (d) Phénomène de focalisation/défocalisation des ondes lié à la forme de bassin en profondeur, ou à celle de la surface libre (effet topographique).

Dans les bassins sédimentaires, les géométries — généralement 2D ou 3D — sont à l'origine de propagations d'ondes complexes. Les ondes de volume à incidence verticale se propageant sur les bords inclinés du bassin, voient une interface oblique. Ceci a pour conséquence de diriger le rai sismique horizontalement à l'intérieur du bassin. Ainsi, une partie des ondes de volume arrivant dans le bassin sont converties en ondes de surface par les bords de celui-ci (Figure 2 (c)) ; la durée et l'amplitude des sollicitations dans le bassin en sont sensiblement augmentées. De plus, des trajets multiples peuvent amener les différentes ondes à arriver simultanément au même endroit ou à en déserter d'autres. Ces effets de concentration/déconcentration augmentent la variabilité spatiale du mouvement sismique, ce qui par conséquent peut augmenter significativement l'amplitude des sollicitations localement (Figure 2 (d)).

Dans les bassins sédimentaires comme ailleurs, les ondes sismiques s'atténuent lors de leur propagation. Outre l'atténuation géométrique qui, par conservation d'énergie, induit une décroissance en $1/r$ pour les ondes de volume et en $1/\sqrt{r}$ pour les ondes de surface, plusieurs mécanismes dépendant du milieu de propagation entrent en jeu et sont habituellement décrits au travers du facteur de qualité Q. Premièrement, l'atténuation anélastique (absorption), décrite par le paramètre Q_i, fait référence à l'ensemble des pertes d'énergie survenant par friction sur les fractures, amortissement visqueux par les fluides présents dans les pores, dégagement thermique lors de la compression/décompression des matériaux, etc. Deuxièmement, le phénomène de diffusion (scattering), quantifié au travers du paramètre Q_{sc}, traduit l'effet des hétérogénéités du milieu. Celles-ci engendrent de petites irrégularités du module élastique et de la densité qui agissent tels des points diffractant en diffusant les ondes qui les traversent. La diffusion a pour effet de complexifier le trajet de l'onde, redistribuant ainsi son énergie dans le temps et dans l'espace. Ce phénomène est à l'origine des arrivées tardives observables sur les sismogrammes et référencées sous le terme de « coda sismique ».

Les dépôts sédimentaires mous sont plus atténuants que les formations rocheuses rigides. Dans les bassins, bien qu'amplification et atténuation s'opposent, le premier phénomène prend le pas sur le second entrainant des facteurs d'amplifications pouvant atteindre 20 par rapport à un rocher sain à certaines fréquences. Néanmoins, l'atténuation du site peut avoir un effet observable dans la décroissance du spectre à haute fréquence qui est décrite par le paramètre κ_0. κ_0 étant la composante de site déduite des mesures individuelles pour chaque séisme de la pente de décroissance linéaire du spectre de Fourier des ondes S obtenue dans un espace log-linéaire (κ_r). κ_r est alors interprété comme étant la part indépendante en fréquence de l'atténuation le long du rai sismique. Cette interprétation est cependant encore contestée aujourd'hui. Il n'en reste pas moins que κ_0 est très largement utilisé dans les GMPEs pour contraindre la forme des spectres à haute fréquence. Une revue plus complète de ce paramètre est fournie dans le chapitre 2 de ce manuscrit.

Enfin, les sols non consolidés peuvent présenter des réponses non-linéaires lorsqu'ils sont soumis à de fortes sollicitations. L'élasticité linéaire formulée par la loi de Hooke n'est alors plus respectée et le sol enregistre des déformations permanentes (ou prolongées). Il en

résulte en général une diminution à haute fréquence et une augmentation à basse fréquence des facteurs d'amplification, liées aux décalages des pics de résonnance vers les basses fréquences. Cet aspect ne sera cependant pas abordé durant ce travail qui est dédié aux effets de site linéaires associés aux mouvements faibles.

1.1.1.2. Quantifications et prédictions

Un grand nombre d'indicateurs peuvent être utilisés pour décrire les effets de site. Des observations macrosismiques telles que des témoignages et le relevé des dégâts générés par un tremblement de terre peuvent être employés, entre autres, pour caractériser d'éventuels effets de site. De même, les observations instrumentales telles que la mesure du pic de déplacement (PGD), de vitesse (PGV) et d'accélération (PGA), de la forme d'onde, du contenu fréquentiel et de la durée du mouvement du sol peuvent fournir une information qualitative sur la distribution spatiale du phénomène. Cependant, pour pouvoir réaliser une observation quantifiée des effets de site, il est indispensable de s'affranchir de l'influence de la source sismique et de la propagation entre la source et le site (Figure 1). En ingénierie, les paramètres clés pour la tenue des structures sont généralement l'amplitude et la fréquence (f) des sollicitations qui sont décrites à travers le spectre de Fourier ou parfois la réponse d'oscillateurs à un degré de liberté (spectre de réponse). L'effet de la fatigue des matériaux induite par la durée de la sollicitation est assez souvent négligé. Ainsi, pour un séisme donné et une fois les corrections de la réponse instrumentale réalisées, le mouvement sismique (U) enregistré à une station de mesure peut être simplement représenté par le produit, dans le domaine de Fourier, d'un terme de source (S), d'un terme de propagation (P) et d'un terme de site (H) :

$$U(w) = S(w) \times P(w) \times H(w), \quad avec\ w = 2\pi f \qquad 4$$

La fonction de transfert du site, $H(w)$, fournit les facteurs d'amplification du site à chacune des fréquences. $H(w)$ est sans unité et correspond donc à une évaluation relative de l'effet de site vis-à-vis d'une position de référence. Cette dernière est théoriquement l'interface entre le substratum rocheux et le sédiment comme discuté au début de ce chapitre. En pratique, la référence peut correspondre à la profondeur du profil de vitesse utilisé pour calculer $H(w)$ ou bien, lors des évaluations empiriques, il peut s'agir de la position d'un site de référence localisé en profondeur dans un forage ou à la surface sur une structure géologique moins sujette aux effets de site.

Pour les sites présentant une géométrie très simple (typiquement un milieu tabulaire à une dimension) et un comportement linéaire, $H(w)$ peut être obtenu assez facilement par simulation numérique. Dans ce cas, un profil vertical 1D de la vitesse des ondes P et S, de la densité et des facteurs de qualité jusqu'au substratum rocheux est suffisant pour établir une bonne prédiction de la réponse des sols. Généralement, seul le profil de vitesse des ondes S est connu et les autres paramètres sont déduits des valeurs de ce dernier. Cependant, bien souvent le profil de vitesse des ondes S n'est disponible jusqu'en proche surface, la réponse

de site est alors approximée à partir de proxys très simples comme la vitesse moyenne des ondes S dans les 30 premiers mètres de sol (V_{S30}) et f_0. Parfois, lorsque même ces proxys ne sont pas connus, seules la géologie et la topographie sont prises en compte ; la géologie permettant de se référer aux classes de sol établies empiriquement et la pente des terrains ayant montré quelques corrélations intéressantes avec V_{S30}.

Lorsque les effets de site sont 2D-3D comme dans les bassins sédimentaires, un profil de vitesse unique et les proxys de site ne permettent pas de reproduire toute la complexité du phénomène (résonnances multiples, génération d'ondes de surface, etc.). Il convient alors de réaliser localement des études de site spécifiques. Les effets de sites étant maintenant assez bien compris, des modélisations numériques 2D-3D fiables sont réalisables. La qualité des résultats qui en sont déduits est alors directement dépendante de la fiabilité du modèle géologique et physique 2D-3D. La construction de ces modèles 2D-3D nécessite une connaissance détaillée de la géométrie des structures géologiques et de la distribution spatiale de leurs paramètres physiques. En pratique, cette connaissance est généralement limitée par le cout de réalisation de forages ou de campagnes de caractérisation géotechnique et géophysique des sites. Ainsi, la limite de résolution des modélisations numériques ne se trouve plus dans les capacités de calcul des ordinateurs, mais dans la résolution des connaissances géologiques, géotechniques et géophysiques des milieux. Aujourd'hui, les modélisations 3D les plus performantes permettent rarement de réaliser des prédictions de la réponse linéaire des sites fiables au-delà de 1-4 Hz (p. ex. Maufroy et al. 2016). De plus, les modélisations 3D (même linéaires) peuvent être notablement incorrectes lorsqu'elles sont réalisées en aveugle (sans ajustement avec les observations expérimentales) comme l'ont montré différents benchmarks réalisés sur le sujet (p. ex. Boore 2004 ; Chaljub et al. 2010, 2015 ; Maufroy et al. 2015a ; Régnier et al. 2016).

Les méthodes empiriques sont essentielles pour l'estimation fiable de la réponse linéaire des sites, en particulier ceux à géométrie 2D-3D. Plusieurs méthodes empiriques existent et sont détaillées dans la section 1.1.2. Celles-ci permettent d'obtenir des informations variées sur la réponse des sites, certaines ne donnant qu'une estimation fiable de f_0 tandis que d'autres fournissent une bonne évaluation de $H(w)$. Dans certaines conditions, les approches empiriques fournissent une information fiable aussi bien sur la réponse moyenne des sites que sur la variabilité associée. Ceci est vrai jusqu'à très haute fréquence (>30 Hz), permettant ainsi de couvrir toute la gamme de fréquences d'intérêt pour le génie civil (1-10 Hz) et pour les éventuels équipements de sureté (dans les centrales nucléaires, les barrages, etc.) ayant des fréquences propres élevées (>5-10 Hz). Ces approches empiriques apportent donc un complément essentiel aux modélisations numériques. Elles requièrent cependant la mise en place d'une instrumentation sur la zone d'étude. La durée d'acquisition de cette instrumentation peut aller de plusieurs semaines à plusieurs années (suivant l'activité sismique de la région, le niveau de bruit local et la sensibilité des instruments) pour les méthodes nécessitant l'enregistrement de bonne qualité (SNR>3) d'un grand nombre de séismes. Ce temps d'acquisition peut toutefois être considérablement réduit (de quelques heures à quelques semaines) avec les méthodes empiriques utilisant l'information portée par le champ sismique ambiant. Le champ sismique ambiant est composé de toutes les

vibrations du sol, improprement appelé bruit parfois, qui sont enregistrables partout et tout le temps avec une instrumentation suffisamment sensible pour les détecter.

1.1.2. L'évaluation empirique des effets de site

1.1.2.1. Contexte historique

Les premières constatations des effets de site furent réalisées lors d'évaluations de l'intensité macrosismique (témoignage, descriptions des dégâts, etc.). Dès le début du XIXe siècle, Macmurdo (1824) note que « buildings situated on rock were not by any means so much affected… as those whose foundations did not reach to the bottom of the soil » lors du séisme de 1819 à Cutch en Inde. Quelques années plus tard, Mallet (1862) relève des effets de la géologie locale sur les niveaux de dommage dans son rapport sur le séisme de Naples de 1857. Par la suite, plusieurs auteurs observeront des variations de l'intensité du mouvement sismique en fonction des conditions de sols (p. ex., Milne 1908; Wood 1908; Reid 1910).

La sismologie instrumentale est une science relativement jeune puisque les premiers sismomètres « modernes » furent développés à partir des années 1880 au Japon par l'équipe de John Milne, James Alfred Ewing et Thomas Gray (Reitherman 2012). Les premiers réseaux capables d'enregistrer des mouvements forts susceptibles de provoquer des dégâts importants furent quant à eux déployés en 1932 aux USA par l'« US Coast and Geodetic Survey » (Trifunac and Todorovska 2001) et à partir de 1951 au Japon (Takemura et al. 1995). Ces derniers permirent le développement de la sismologie de l'ingénieur qui s'intéresse aux effets des séismes en surface par opposition à la sismologie classique qui vise à étudier l'intérieur de la Terre. Ainsi quelques mois seulement après l'installation du réseau accéléromètrique aux États-Unis, le séisme destructeur de Long Beach de 1933 put être enregistré et pour la première fois les ingénieurs purent observer la nature des mouvements forts : l'amplitude, la durée et les fréquences caractéristiques des vibrations. Parallèlement à l'étude des séismes eux-mêmes, la première moitié du XXe siècle voit la réalisation des premiers travaux s'intéressant au champ sismique ambiant (Ishimoto 1937). Gutenberg (1927) propose des facteurs d'amplification site-dépendant évalués depuis des enregistrements du champ sismique ambiant sur des sites aux géologies variées. Cependant, c'est seulement à partir des années 50 que l'étude du champ sismique ambiant jusqu'aux hautes fréquences émerge réellement (p. ex., Kanai et al. 1954 ; Kanai and Suzuki 1955 ; Aki 1957 ; Kanai and Tanaka 1961).

Avant les années 70, les études empiriques relatives aux effets de site sont essentiellement qualitatives et ne proposent pas d'approches permettant de séparer complètement la composante de site des autres composantes du mouvement sismique, interdisant ainsi toute mesure quantifiée du phénomène. Depuis lors, la sismologie de l'ingénieur a connu un essor très important motivé par l'augmentation des enjeux liés à la réalisation de projets de grande envergure (centrales nucléaires, gratte-ciels, grands barrages, plateformes pétrolières offshore). Dans le même temps, les réseaux sismologiques se densifient et effectuent le

passage de l'analogique vers le numérique, fournissant ainsi toujours plus de données pouvant être gérées et traitées grâce aux progrès simultanés des ordinateurs. Plus spécifiquement, les effets de site ont connu un regain d'intérêt très net suite au séisme de Michoacán de 1985 dont les conséquences sur la ville de Mexico leur sont très largement imputées. Cet essor se traduit par un nombre croissant d'articles publiés chaque année sur le sujet et par l'introduction de plusieurs méthodes d'évaluation empirique des effets de site. Nous ne tenterons pas ici de présenter toutes les méthodes empiriques, mais seulement celles qui ont été utiles à la réalisation de ce travail. De la même façon, une liste exhaustive des travaux existant pour chacune des méthodes n'est pas envisageable en raison du nombre très important d'études ayant porté sur le sujet. Toutefois, pour ce qui concerne les méthodes utilisant le champ sismique ambiant, le chapitre 3 fournit une revue plus complète de la littérature s'y rapportant. Ainsi, les méthodes utilisant la corrélation de bruit comme la cohérence et l'« ambient noise impulse response function » (ANIRF) ne sont pas présentées ici, mais au chapitre 3, ces dernières n'étant pas spécifiquement dédiées à l'étude des effets de site.

1.1.2.2. SSR

C'est Borcherdt (1970) qui le premier propose une approche empirique capable d'isoler la composante de site du mouvement sismique. Il introduit la méthode « standard spectral ratio » (SSR) ou « site-sur-référence » en français qui consiste à effectuer le rapport entre les spectres évalués depuis le signal sismique enregistré à une station située sur le site que l'on cherche à étudier « s » et celui obtenu depuis une seconde station qui sert de référence « r »

$$SSR_{s/r}(w) = \langle \frac{\{|\overline{U}_{is}(w)|\}}{\{|\overline{U}_{ir}(w)|\}} \rangle \qquad 5$$

où $|\overline{U}_{is}(w)|$ et $|\overline{U}_{ir}(w)|$ sont l'amplitude du spectre de Fourier (FAS) d'une composante du signal sismique i enregistrée à la station située sur le site d'étude et à la référence respectivement. $\{\cdot\}$ dénote le lissage Konno and Ohmachi (1998) et $\langle\cdot\rangle$ indique l'application de la moyenne géométrique (logarithmique) sur la distribution des n évènements, i allant de 1 à n.

Le site de référence est classiquement situé sur une zone peu sensible aux effets de site (rocher dur) et est localisé à proximité du site que l'on cherche à étudier. Lorsque la distance épicentrale (Re) est bien supérieure à la distance entre les deux sites de mesure ($Re > 10\Delta d$) alors on fait l'hypothèse que le signal incident en profondeur (terme de source et de trajet) est le même pour les deux sites. De cette façon, le rapport des spectres tirés des enregistrements sismiques réalisés simultanément sur les deux sites fournit la fonction de transfert relative entre le site et la référence comme illustrée sur la Figure 3.

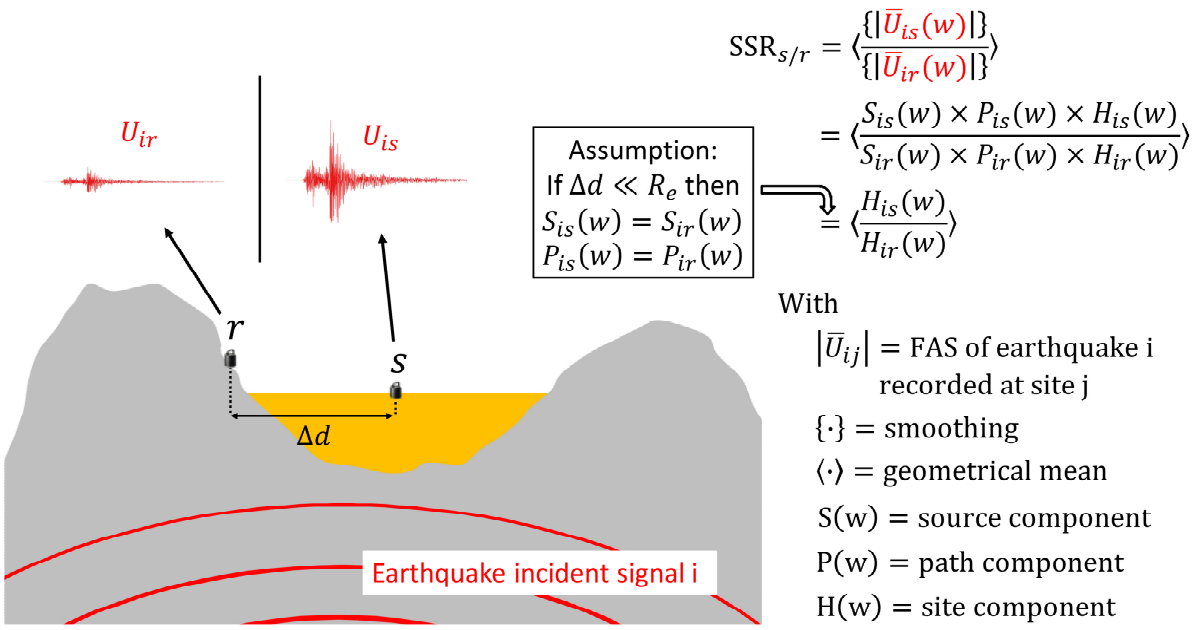


Figure 3. Illustration de la méthode SSR pour le cas d'un bassin sédimentaire.

Les principales limites de cette approche sont liées à la difficulté de sélection d'un site de référence fiable (peu sujet aux effets de site) et suffisamment proche du site d'étude (Steidl et al. 1996; Yu and Haines 2003 ; Drouet 2006 ; Cadet et al. 2012) et à la nécessité de réaliser un nombre statistiquement représentatif d'enregistrements de séismes présentant un bon rapport signal-sur-bruit (SNR>3) sur les deux sites de mesures. Il résulte du premier point que cette approche est particulièrement adaptée à l'étude des effets de site dans les bassins sédimentaires tandis que le dernier point exclut une utilisation de la méthode SSR pour la réalisation rapide de microzonages en particulier dans les zones faiblement sismiques où un nombre très important de capteurs devrait être déployé durant un temps généralement long (parfois plusieurs années). Quoi qu'il en soit de nombreuses études ont montré la fiabilité de cette approche qui est très largement utilisée pour évaluer la réponse des sites (p. ex. King and Tucker 1984 ; Aki 1993 ; Field 1996 ; Bonilla et al. 1997 ; Faccioli and Vanini 2003; Bindi et al. 2009 ; Thompson et al. 2012). De plus, la méthode SSR est régulièrement considérée comme méthode de référence pour valider des résultats empiriques issus d'autres approches (Rovelli et al. 1991; Malagnini et al. 1993 ; Field and Jacob 1995; Satoh et al. 1998 ; Horike et al. 2001 ; Cultrera et al. 2014) ou pour confronter les résultats issus des modélisations numériques (Jongmans and Campillo 1993; Cornou and Bard 2003; Triantafyllidis et al. 2006; Maufroy et al. 2015 ; Maufroy et al. 2017).

1.1.2.3. SSRn

Une dizaine d'années après l'introduction de la méthode SSR par Borcherdt (1970), son équivalent pour le champ sismique ambiant, le SSRn, fut proposé pour la première fois par Irikura and Kawanaka (1980) pour détecter la présence de failles en profondeur. Rapidement, le SSRn fut repris dans plusieurs études pour tenter d'évaluer la réponse des

sites (Kagami et al. 1982, 1986). Comme pour la méthode SSR présentée dans la section précédente, le SSRn effectue le rapport des spectres composante par composante entre une station située sur le site que l'on cherche à étudier et une seconde station localisée sur un site de référence moins sujet aux effets de site :

$$SSRn_{s/r}(w) = \langle \frac{\{|\bar{u}_{is}(w)|\}}{\{|\bar{u}_{ir}(w)|\}} \rangle \qquad 6$$

où $|\bar{u}_{is}(w)|$ et $|\bar{u}_{ir}(w)|$ sont la FAS d'une composante d'une fenêtre temporelle de champ sismique ambiant i enregistré simultanément à la station située sur le site d'étude et à la référence respectivement. $\langle\cdot\rangle$ correspond à l'application de la moyenne géométrique sur la distribution des n fenêtres, i allant de 1 à n.

D'une façon similaire à la condition de distance épicentrale minimum requise pour la méthode SSR, beaucoup d'études employant la méthode SSRn depuis le champ sismique ambiant ont été réalisées à basse fréquence sous le Hertz (microséismes) s'assurant ainsi de l'utilisation de sources de bruit océanique distantes du site d'étude (p. ex. Okada et al. 1991 ; Yamanaka et al. 1993 ; Ferritto 1995 ; Field 1996). De cette façon, l'hypothèse selon laquelle la composante de source et de trajet du bruit est la même pour le site et la référence est vérifiée. Au-delà du Hertz, le champ sismique ambiant est dominé par l'activité humaine (microtremors). Ainsi, lorsque la fréquence augmente, les sources de bruit à l'origine des microtremors sont plus proches du site d'enregistrement et leurs positions sont bien souvent inconnues rendant l'hypothèse précédente invérifiable. Cependant, des solutions ont été développées comme la sélection minutieuse de nombreuses fenêtres de bruit qui lorsqu'on les moyenne tend à annihiler la signature de chacune des combinaisons de source et de trajet. Ainsi, si l'énergie moyenne reçue (hors effets de site) au site et à la référence est similaire, alors les composantes de source et de trajet des microtremors peuvent être négligées. Cette vision déterministe du champ sismique ambiant, qui est celle autour de laquelle s'est construite la méthode SSRn, considère chaque source indépendamment. La méthode SSRn est illustrée sur la Figure 4 pour le cas simpliste d'une seule source de bruit déterministe. Pour plus de réalisme, de nombreuses sources de bruit devraient être représentées puisque le succès de la méthode jusqu'aux hautes fréquences dépend de la multiplicité des sources et de l'homogénéité de leurs distributions spatiales.

Les études ayant évalué le SSRn à partir des microtremors donnent des résultats mitigés qui dépendent des conditions expérimentales (géologie du site, niveau et proximité des sources de bruit, distance entre le site et la référence) et des traitements effectués (choix des instruments, critère de sélection des fenêtres, bande de fréquence évaluée, etc.). En l'absence d'un consensus quant au succès du SSRn dans l'évaluation de la fonction de transfert du site et du fait du succès grandissant de la méthode HVSR dans les années 90, la méthode SSRn a été graduellement abandonnée.

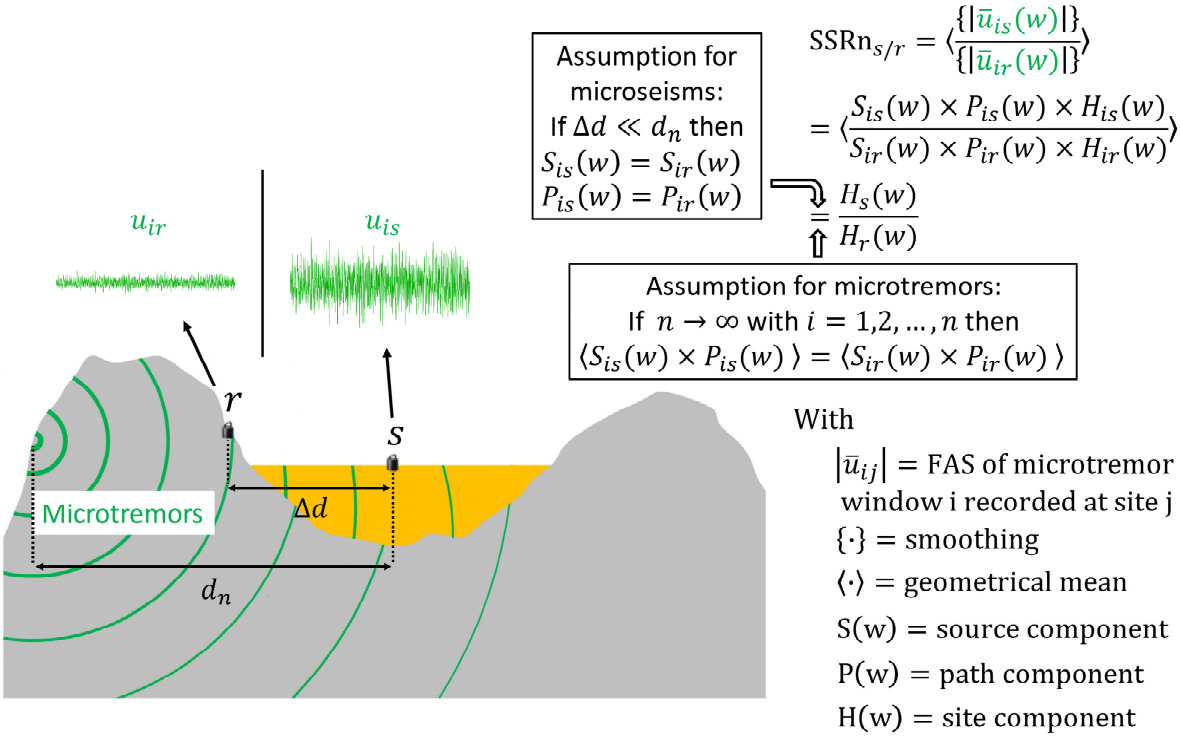


Figure 4. Illustration de la méthode SSRn pour le cas simplifié d'une seule source de bruit se propageant dans un bassin sédimentaire.

1.1.2.4. HVSR

Quelques années après l'introduction du SSRn, la méthode HVSR, parfois dénommée « Nakamura's technics » ou méthode H/V, est proposée pour la première fois en anglais à la communauté scientifique par Nakamura (1989). Cependant, cette dernière fut introduite d'abord au Japon par Nogoshi and Igarashi (1971) à partir d'une étude initiale de Kanai and Tanaka (1961). Ce sont finalement les travaux de Lachet and Bard (1994) puis de Kudo (1995) et plusieurs autres auteurs qui permirent de mieux comprendre et de généraliser l'utilisation de cette approche. Dès lors, un nombre très important de travaux ont été réalisés sur le sujet (Bonnefoy-Claudet et al. 2006). Cette popularité s'explique principalement par la facilité, la rapidité et le faible cout d'utilisation de la méthode permis par l'emploi de microtremors et par l'absence de nécessité d'un site de référence. En effet, la méthode HVSR consiste à faire simplement le rapport des spectres des composantes horizontales sur la composante verticale de microtremors enregistrés sur chacune des stations de mesures indépendamment.

$$HVSR(w) = \langle \frac{\{|\bar{u}_{iH}(w)|\}}{\{|\bar{u}_{iV}(w)|\}} \rangle \qquad 7$$

où $|\bar{u}_{iH}(w)|$ peut être la FAS d'une direction horizontale particulière ou de la moyenne des composantes horizontales et $|\bar{u}_{iV}(w)|$ est la FAS de la composante verticale de la fenêtre de microtremor i.

Il est couramment admis aujourd'hui que cette approche permet de prédire la fréquence fondamentale du site, f_0, pour les ondes S se propageant dans un site sédimentaire 1D. Ceci s'explique à la fois par l'ellipticité des ondes de Rayleigh qui engendre une chute de la composante verticale à la fréquence f_0 et par la phase Airy de l'onde de Love qui correspond à un saut d'énergie sur les deux composantes horizontales à f_0 qui conduisent tous deux à un pic lors du rapport H/V (p. ex., Lachet and Bard 1994; Lermo and Chávez-García 1994; Satoh et al. 2001; Bonnefoy-Claudet et al. 2006). Sa fiabilité dans l'estimation de f_0 a été démontrée à la fois théoriquement et dans la plupart des évaluations empiriques qui en ont été faites pour certaines conditions de site (contraste d'impédance important et géométrie assimilable à du 1D). Un long débat eu lieu concernant la capacité de cette méthode à estimer les facteurs d'amplification ou les fréquences de résonnance des harmoniques supérieures à f_0. Cela entraina, entre autres, la réalisation d'un workshop sur le sujet qui conclut à une incapacité de la méthode à estimer les facteurs d'amplifications et les modes de résonance supérieurs (SESAME team 2004 – http://sesame.geopsy.org/Delivrables/Del-D23-HV_User_Guidelines.pdf). Ce workshop conduit également à de nombreuses recommandations quant à la réalisation de la mesure HVSR ainsi qu'à la mise à disposition du logiciel de calcul HVSR Geopsy (http://www.geopsy.org/).

1.1.2.5. HVSRs

La méthode des rapports spectraux H/V sur séismes fut introduite par Lermo and Chávez-García (1993) suivant une étude de Langston (1979) visant à évaluer la structure superficielle de la croûte terrestre à partir du HVSRs des ondes P de signaux télésismiques. Elle est également largement inspirée de la méthode HVSR appliquée aux microtremors qui fut introduite antérieurement et qui est présentée dans la section précédente. En comparaison avec la méthode SSR, la méthode HVSRs a l'avantage de ne pas nécessiter de station de référence et peut donc être appliquée partout avec une seule station de mesure. Cette approche consiste à réaliser le rapport spectral des composantes horizontales sur la composante verticale d'une fenêtre d'ondes S faisant l'hypothèse que la composante verticale est moins sensible aux effets de site.

$$HVSRs(w) = \langle \frac{\{|\bar{U}_{iH}(w)|\}}{\{|\bar{U}_{iV}(w)|\}} \rangle \qquad 8$$

où $|\bar{U}_{iH}(w)|$ peut être la FAS d'une direction horizontale particulière ou de la moyenne des composantes horizontales et $|\bar{U}_{iV}(w)|$ est la FAS de la composante verticale du signal i.

Cependant, l'hypothèse sous-jacente selon laquelle la composante verticale ne présente pas d'amplifications liées au site est très discutable (Oth et al. 2009). Ainsi, la méthode HVSRs ne permet pas de reproduire complètement les résultats qui sont obtenus avec la méthode SSR (Bonilla et al. 1997). Les fonctions d'amplification tirées du HVSRs sous-estiment généralement grandement les niveaux d'amplification vis-à-vis de ceux obtenus avec la méthode SSR (Riepl et al. 1998 ; Raptakis et al. 1998 ; Parolai et al. 2004) et même de ceux issus de la méthode HVSR réalisée sur les microtremors (Mucciarelli and Gallipoli 2004). De plus, les harmoniques supérieures à la fréquence fondamentale ne sont pas toujours observables. Finalement, tout comme la méthode HVSR, la méthode HVSRs ne produit de façon fiable que la fréquence fondamentale du site f_0 pour certaines configurations de site, généralement 1D (Haghshenas et al. 2008; Cultrera et al. 2014).

1.2. Traitement des données

Les méthodes empiriques (détaillées dans la section précédente) sont toutes évaluées à partir des spectres de Fourier tirés des fenêtres de bruit ambiant ou de signaux sismiques. La sélection et la vérification de ces fenêtres sont des étapes essentielles qui font l'objet de cette partie. La sélection des fenêtres pour les méthodes requérant des signaux sismiques est sensiblement différente de celle des méthodes ne nécessitant que du bruit ambiant. Les premières demandent une identification et une extraction des enregistrements de séismes depuis les données acquises en continu. S'en suit la nécessité de pointer les temps des premières arrivées d'onde P et d'onde S pour pouvoir réaliser le fenêtrage des différentes phases du signal (ondes P, ondes S, coda, tout le signal, etc.). Le fenêtrage des enregistrements de bruit est plus simple à réaliser puisque les fenêtres ciblées, à condition de respecter certains critères, peuvent être tirées aléatoirement depuis les enregistrements continus et aucun pointé n'est nécessaire. Finalement, quelle que soit la fenêtre d'intérêt plusieurs vérifications sont requises afin d'assurer la qualité des évaluations des spectres de Fourier qui en découlent. En effet, les enregistrements des mouvements du sol par les sismomètres peuvent être corrompus par toutes sortes de défauts (p. ex. problème d'horloge, mauvais échantillonnage, mauvaise orientation des composantes, signaux erratiques, etc.) qu'il faut pouvoir repérer et corriger afin d'éviter d'introduire des biais dans les résultats des évaluations empiriques.

1.2.1. Traitements spécifiques aux séismes

1.2.1.1. Extraction des séismes depuis les enregistrements continus

L'identification des signaux sismiques depuis les enregistrements continus est réalisée à partir des informations tirées des catalogues sismiques. Pour la France métropolitaine, les informations présentes dans les bulletins du réseau national de surveillance sismique (ReNaSS - http://renass.unistra.fr/recherche), du centre sismologique euro-méditerranéen (CSEM - http://www.emsc-csem.org/Earthquake/?filter=yes), de Sismoazur du laboratoire Géoazur (http://sismoazur.oca.eu/catalog/) et de l'« italian seismological instrumental and

parametric database » (ISIDe - http://iside.rm.ingv.it/iside/standard/index.jsp) sont utilisées. Les évènements communs à plusieurs catalogues sont identifiés par comparaison automatique sur des intervalles de recherche des valeurs de magnitude, position (latitude, longitude) et de date de l'évènement. Pour chaque évènement, les temps d'arrivée théorique des ondes P et S à la station de mesure ($T_{Pthéo}$ et $T_{Sthéo}$) sont calculés à partir des temps origines des séismes fournis par les bulletins, de la position relative des stations de mesures vis-à-vis de l'hypocentre des séismes et du modèle de vitesse globale IASP91 (https://ds.iris.edu/ds/products/emc-iasp91/). Lorsque plusieurs évènements aboutissent à des temps d'arrivée théoriques à la station qui sont similaires et qui concordent avec une arrivée réelle dans les données recueillies, alors l'attribution à l'évènement le plus plausible est décidée au cas par cas. Pour le site grec, du fait d'un très grand nombre de séismes ayant eu lieu durant la séquence sismique de 2014 sur l'ile de Céphalonie et pour éviter tout recoupement hasardeux, seul le bulletin NOA de l'« institute of geodynamics » (G.I. - http://www.gein.noa.gr/en/) rattaché à l'observatoire national d'Athènes est utilisé.

Plusieurs dizaines de milliers de séismes survenant partout dans le monde sont répertoriés chaque année. Une fois un premier filtre magnitude/distance épicentrale effectué, seuls quelques milliers voire quelques centaines d'évènements présents dans les catalogues sont réellement susceptibles d'avoir été enregistrés chaque année sur un site d'étude donné, les autres étant trop distants et de trop petites magnitudes.

Une vérification de type STA/LTA est réalisée sur le module des trois composantes autour du temps théorique d'arrivée de chaque séisme potentiellement enregistrable à la station cible. Ainsi, seuls sont conservés les enregistrements présentant un pic dans un intervalle de temps court autour de $T_{Pthéo}$ dépassant de plus de trois fois le module moyen calculé sur un intervalle long d'une heure. Cette sélection est illustrée sur la Figure 5. L'intervalle de temps de la fenêtre courte autour de $T_{Pthéo}$ ainsi que l'intervalle de fréquence considéré sont définis en fonction de la distance épicentrale comme détaillé dans le Tableau 1.

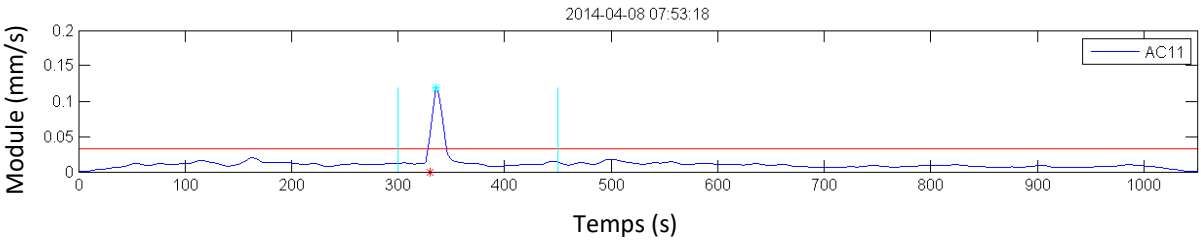


Figure 5. Exemple de détection d'un évènement par la méthode STA/LTA réalisée sur le module des trois composantes (courbe bleue) d'un enregistrement. L'intervalle de temps court est indiqué par les traits verticaux bleu clair, l'étoile bleu clair montre le pic tandis que celle en rouge donne le temps d'arrivée théorique à la station ($T_{Pthéo}$). Le trait horizontal rouge indique la limite de trois fois le niveau de bruit moyen évalué sur une fenêtre longue d'une heure.

L'inconvénient de cette approche est qu'elle interdit l'utilisation des séismes de très petite magnitude ($M_l < 1{,}3$) et très locaux, pouvant être présents dans les enregistrements, mais qui ne sont pas référencés dans les catalogues. Ces petits évènements locaux ne bénéficient cependant pas, de fait, des métadonnées renseignées par les catalogues,

rendant ainsi leur utilisation plus limitée. De plus, la réalisation d'une recherche d'évènements de type STA/LTA sur l'ensemble des enregistrements continus est chronophage et présente le risque d'engendrer un grand nombre de faux déclenchements qui devraient être vérifiés visuellement un par un par la suite. Cette étape de vérification est réalisée lors du pointé et est détaillée dans la section suivante.

Tableau 1. Intervalles de temps (relatif à $T_{Pthéo}$) et de fréquence en fonction de la distance épicentrale utilisée pour l'identification des évènements sismiques depuis les enregistrements continus

Re (km)	≤500	]500 1000]	]1000 2000]	]2000 4000]	]4000 8000]	≥8000
$\Delta t/T_{pthéo}$ (min)	[-0,5 2]	[-0,5 3]	[-0,5 5]	[-0,5 12]	[-0,5 20]	[-0,5 30]
Δf (Hz)	[2 10]	[1 7]	[0,5 6]	[0,1 5]	[0,05 3]	[0,03 3]

1.2.1.2. Pointé des premières arrivées

Le pointé manuel des premières arrivées d'onde P (T_P) et S (T_S) est un travail fastidieux. Depuis longtemps, plusieurs méthodes automatiques ont été proposées (p. ex. Baer and Kradolfer 1987 ; Zhang et al. 2003 ; Küperkoch et al. 2010) mais ces dernières égalent difficilement la finesse d'analyse qui est permise par un opérateur. De plus, le pointé est bien souvent la seule étape durant laquelle les signaux sont vérifiés, un par un, visuellement.

Pour réaliser ce double travail de pointé et de vérification des signaux, une routine Matlab a été développée. Elle permet l'affichage sur une fenêtre graphique des trois composantes du signal, du spectrogramme moyen aux trois composantes et elle propose de nombreux boutons de contrôle ou de prise de décisions. Un exemple de son utilisation est donné par la Figure 6.

Cet outil permet d'identifier rapidement à l'aide du spectrogramme d'éventuels bruits transitoires ou toutes autres perturbations pouvant survenir à n'importe quelle fréquence. Le spectrogramme est également utile pour distinguer rapidement les fréquences caractéristiques de l'évènement que l'on cherche à pointer. Une fois les bornes du filtre ajustées pour maximiser le rapport signal-sur-bruit, le pointé des ondes P et S peut être réalisé directement sur l'interface. Le temps marquant la fin du signal (T_{end}) peut également être pointé en se basant sur les informations fournies par le spectrogramme. Ceci a l'avantage de tenir compte des durées du signal à toutes les fréquences, de la présence éventuelle de bruit ou d'une réplique après l'évènement ciblé et de ne pas être dépendant de la longueur de l'enregistrement considéré comme peuvent l'être les approches automatiques basées sur l'énergie cumulée (p. ex. l'intensité d'Arias). L'opérateur a la possibilité de conserver les éventuels pointés réalisés dans les métadonnées du fichier évènementiel ou de supprimer l'enregistrement si celui-ci n'est pas satisfaisant.

Une fois les pointés effectués pour toute la base de données, une vérification visuelle rapide de la présence d'éventuelles erreurs doit être réalisée. La Figure 7 montre un exemple de cette vérification pour la base de données d'Argostoli.

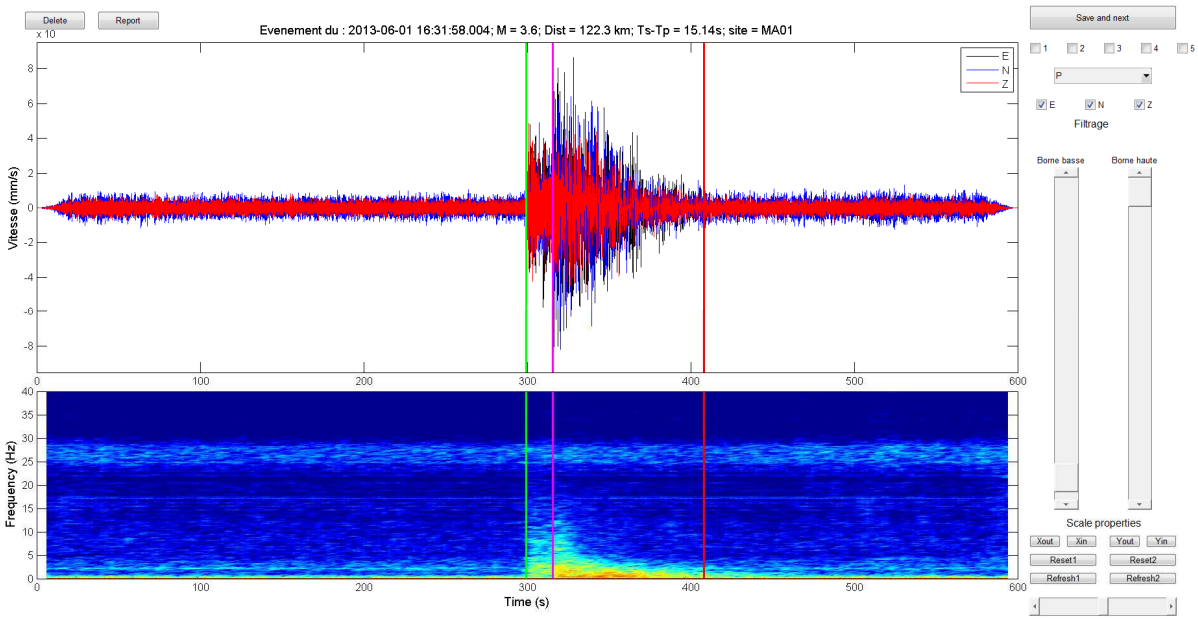


Figure 6. Exemple d'utilisation du programme de pointé et de vérification des signaux sismiques.

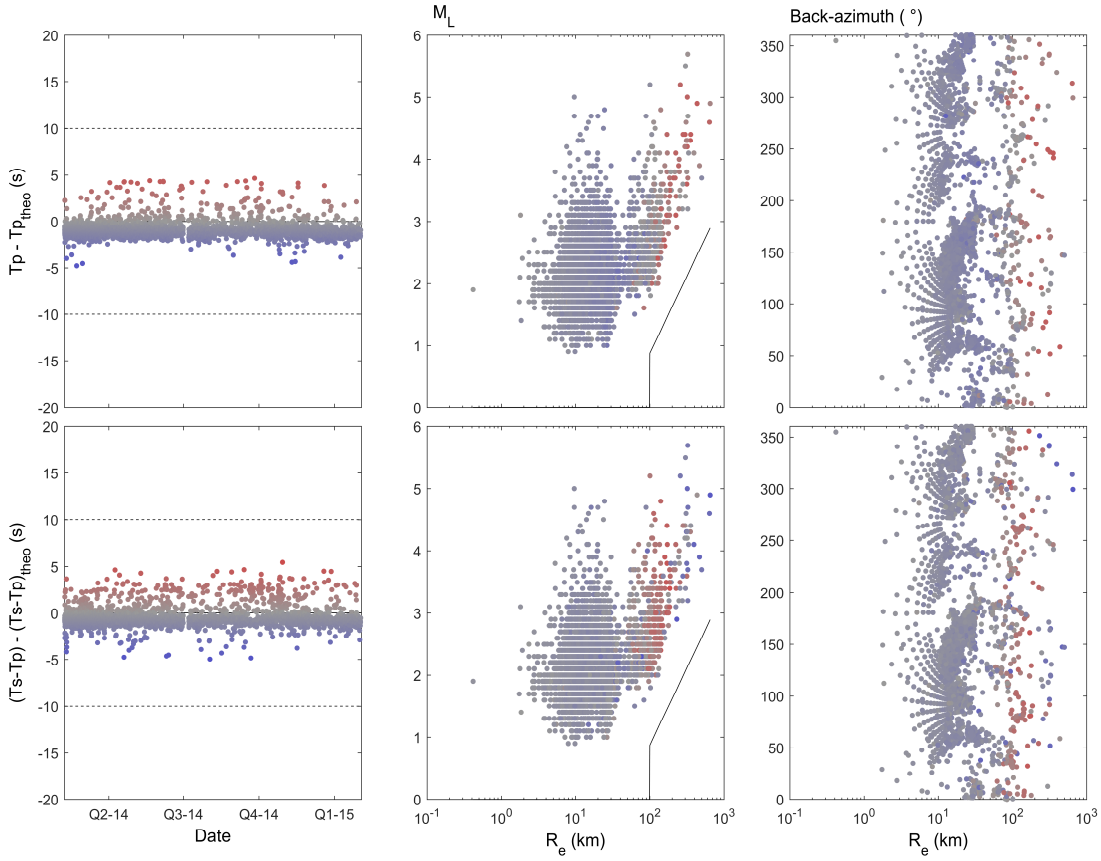


Figure 7. Vérifications des valeurs de pointé manuel T_P et $T_S - T_P$ vis-à-vis de ceux estimés à partir du modèle de vitesse IASP91. Les écarts (échelle de couleur) sont reportés sur une

représentation magnitude/distance et back-azimut/distance afin de faire ressortir les grandes tendances du modèle de vitesse qui pourraient expliquer les écarts observés.

Les pointés manuels de T_P y sont comparés à ceux prédits par le modèle de vitesse globale IASP91 ($T_{Pthéo}$). Les écarts autour de 0 sont marqués par un code couleur du bleu (écart négatif) vers le rouge (écart positif). Les séismes correspondants sont ensuite reportés avec leurs couleurs sur une représentation magnitude/distance et back-azimut/distance. Ainsi, des tendances peuvent être relevées, comme sur la Figure 7 (panneaux centraux), où il apparait nettement que le modèle de vitesse sous-estime les temps d'arrivée des évènements les plus proches (en bleu) et surestime ceux des évènements localisés aux alentours des 100 km de distance épicentrale (en rouge). Un travail similaire est effectué pour les valeurs de $T_S - T_P$. De cette façon, les écarts de pointé attribuable au modèle de vitesse ou à une erreur lors du pointé manuel peuvent être identifiées. Les erreurs de pointé manuel les plus importantes sont ainsi évitées. Ceci est particulièrement important lorsque, comme ici, plusieurs opérateurs différents ont réalisé le pointé.

1.2.1.3. Sélection des fenêtres de bruit et des phases du signal

Cette partie est un résumé de l'article Perron et al. (2017) auquel le lecteur pourra se référer en annexe 1 de ce manuscrit pour plus de détails. Un algorithme semi-automatique est proposé pour évaluer les fenêtres des différentes phases du signal (ondes P, ondes S, coda et signal complet) et la fenêtre de bruit. L'évaluation de ces fenêtres est un prérequis essentiel au calcul des spectres de Fourier à partir des différentes phases du signal. Le spectre de bruit est quant à lui nécessaire à la détermination du rapport signal-sur-bruit à chacune des fréquences. La plupart des études empiriques utilisent la phase d'onde S et considèrent simplement une fenêtre de durée constante, identique pour tous les évènements. La fenêtre de bruit est le plus souvent prise juste avant l'arrivée des ondes P. Ici, nous proposons une approche plus complexe qui tient compte de façon plus réaliste de la durée réelle des signaux et qui teste plusieurs fenêtres de bruit.

Avant le passage en Fourier, les bords des fenêtres temporelles de dimension finie doivent être apodisées afin de minimiser les effets de bords liés au fenêtrage temporel. Dans notre approche, les fenêtres sont dimensionnées pour que l'apodisation se fasse à l'extérieur des phases de signal afin d'éviter de raboter des parties de signaux pouvant être utiles. Ainsi, les durées des phases P (D_P), S (D_S), de coda (D_C) et de tout le signal (D_{All}) sont définies par :

$$D_P = \frac{T_S - T_P}{1 - t_X} \qquad 9$$

$$D_S = \frac{\frac{1}{f_c} + (T_S - T_P)}{1 - 2t_X} \qquad 10$$

$$D_C = T_{end} - (3.3T_S - 2.3T_P) \quad 11$$

$$D_{All} = \frac{T_{end} - T_P}{1 - tx} \quad 12$$

où t_X (fixé à 5 % ici) représente le taux d'apodisation et f_c la fréquence coin de la source. La durée de la phase d'ondes P est donc bornée par l'arrivée des ondes S, tandis que celle de la phase d'onde S est définie par un terme f_c^{-1} représentatif du temps de la rupture à la source et d'un terme en $(T_S - T_P)$ tenant compte de l'allongement du signal durant la propagation. Le terme de source est négligeable tant que la magnitude reste modérée ($M_l < 5$). Au-delà, f_c peut être approximée par la magnitude en suivant quelques relations (p. ex. Brune 1970 ; Hanks and Kanamori 1979) et en faisant quelques hypothèses localement sur les valeurs de chute de contrainte ($\Delta\sigma = 10\ bars$) et de vitesse moyenne des ondes S ($\beta = 3500\ m/s$).

Finalement, les intervalles des différentes phases sont définis par :

$$I_P = [T_P - D_P \times t_X\ ;\ T_S] \quad 13$$

$$I_S = \left[T_S - D_S \times t_X\ ; \min\begin{Bmatrix} T_S + D_S(1 - t_X) \\ T_{end} \end{Bmatrix}\right] \quad 14$$

$$I_C = [3.3T_S - 2.3T_P\ ;\ T_{end}] \quad 15$$

$$I_{All} = [T_P - D_{All} \times t_X\ ;\ T_{end}] \quad 16$$

où I_P, I_S, I_C et I_{All} sont les intervalles en temps des phases d'ondes P, d'ondes S, de coda, et de tout le signal respectivement. Pour ce qui concerne l'intervalle de bruit, plusieurs fenêtres sont testées avant et après l'enregistrement du séisme. La méthode de sélection de la fenêtre de bruit ne sera pas montrée plus en détail ici, mais l'idée principale est de proposer un code pouvant fonctionner même lorsque des durées courtes de bruit sont disponibles avant l'arrivée des ondes P, permettant ainsi de s'adapter à des bases de données de qualité très variable.

1.2.2. Traitements applicables au champ sismique ambiant

Le champ sismique ambiant a pour avantage d'être présent partout et en tout temps. Cependant, pour être pleinement utilisables les fenêtres de bruit doivent être sélectionnées avec attention afin d'éviter d'introduire des bruits transitoires ou tout autres perturbateurs susceptibles de biaiser les résultats des évaluations qui en découlent. En effet, le champ sismique ambiant peut être perturbé à basse fréquence par des séismes de forte magnitude très lointains (téléséismes), par des orages, etc., et aux fréquences plus élevées par l'activité humaine principalement (p. ex. trafic routier, bruit de pas). Plusieurs études ont montré l'importance des traitements appliqués aux données de bruit, en particulier pour l'émergence de la fonction de Green dans les approches de corrélation de bruit (Bensen et al. 2007; Baig et al. 2009 ; Seats et al. 2012 ; Melo et al. 2013 ; Bowden et al. 2015 ; Yoritomo and Weaver 2016).

Dans un soucis d'homogénéité, nous avons choisi de suivre la même procédure pour toutes les approches empiriques relatives au champ sismique ambiant. Dans notre travail nous disposons d'enregistrements continus de plusieurs mois qui nous autorisent à effectuer une sélection minutieuse des fenêtres de bruit. La procédure consiste à tirer aléatoirement des fichiers de données continues (généralement d'une heure) puis à les tester selon plusieurs critères pour décider de les utiliser ou non. Les données de chaque fichier sont d'abord segmentées en fenêtre de bruit d'une minute. Cette durée est suffisante pour assurer une bonne résolution spectrale jusqu'à la plus basse fréquence d'intérêt ici (~0,25 Hz). Seules les fenêtres d'une minute ne présentant pas de pic dépassant le seuil de 10 fois l'écart-type évalué sur le fichier complet sont conservées (Denolle et al. 2013). Si plus de 70 % des fenêtres ont été supprimées, alors le ficher complet est rejeté. Dans le cas contraire, le spectre moyen et le coefficient de variation associé sont calculés depuis les fenêtres d'une minute sélectionnées. Si ce coefficient de variation moyen à toutes les fréquences est supérieur à 100 % ou si le maximum du coefficient de variation entre 0,2 et 15 Hz est supérieur à 150 % alors le fichier est rejeté. À la fin du processus, ce sont plusieurs milliers de fenêtres d'une minute de bruit ambiant qui sont sélectionnées.

1.2.3. Évaluation des spectres

Une fois les fenêtres de bruit ou de signal sismique sélectionnées, les données temporelles peuvent être passées dans le domaine de Fourier. Avant cela, quelques traitements préliminaires sont toutefois nécessaires. Tout d'abord, la tendance linéaire des signaux doit être retirée afin de s'assurer que la moyenne est nulle sur la fenêtre concernée. Comme nous l'avions évoqué dans les sections précédentes, l'apodisation est une étape incontournable pour effectuer la transformée de Fourier d'un signal de dimension finie. Celle-ci est réalisée au travers d'une décroissance cosinus des bords du signal (fenêtre de Tukey) dont l'expression analytique est

$$w(x) = \begin{cases} \frac{1}{2}\left(1 + \cos\left(\frac{2\pi}{t_X}\left(x - \frac{t_X}{2}\right)\right)\right), & 0 \leq x < \frac{t_X}{2} \\ 1, & \frac{t_X}{2} \leq x < 1 - \frac{t_X}{2} \\ \frac{1}{2}\left(1 + \cos\left(\frac{2\pi}{t_X}\left(x - 1 + \frac{t_X}{2}\right)\right)\right), & 1 - \frac{tx}{2} \leq x < 1 \end{cases} \qquad 17$$

Outre de possibles problèmes d'échantillonnages qu'il faut pouvoir détecter et corriger par interpolation, il n'est pas rare que des trous (absence de données) soient présents au sein de la fenêtre dont on souhaite calculer le spectre. Lorsque l'on travaille avec le bruit, la grande quantité de fenêtres disponibles autorise à rejeter toute fenêtre présentant des trous. Il en va différemment lorsqu'il s'agit de signaux sismiques, où des solutions doivent être imaginées pour préserver le maximum d'enregistrements. Les stratégies utilisées divergent en fonction de la taille des trous : les trous inférieurs à 10 % de la longueur totale de la fenêtre dans la limite de 1000 échantillons sont comblés par interpolation linéaire ; sinon, le trou est comblé par des 0 et le signal avoisinant le trou est apodisé par une fenêtre de Tukey. En réalité, le schéma utilisé est légèrement plus complexe puisqu'il tient compte de la position des trous entre eux pour choisir de les traiter individuellement ou bien ensemble en supprimant les portions très courtes de signal qui les séparent.

Pour permettre l'utilisation conjointe des spectres issus des différentes phases du signal et pour autoriser l'évaluation du spectre de bruit sur une fenêtre suffisamment grande pour être statistiquement représentative, il faut que les spectres évalués sur des fenêtres de longueur variée puissent être comparables. Pour ce faire, le module de la transformée de Fourier discrète est normalisé par la racine du nombre total d'échantillons (densité spectrale d'amplitude - DSA)

$$|\overline{U}(k)| = \frac{\sqrt{dt}}{\sqrt{n_f}} \left| \sum_{n=0}^{n_f - 1} U(n) \cdot e^{-i\omega k \frac{n}{n_f}} \right| \qquad \text{pour } 0 \leq k < n_f \qquad 18$$

où U est le signal temporel, $|\overline{U}|$ est la DSA, n est l'indice d'échantillonnage en temps, n_f est le nombre total d'échantillons, dt est le pas de temps entre deux échantillons et k est l'indice d'échantillonnage en fréquence. Parfois, une étape de dérivation/intégration du signal peut être requise et est réalisable dans le domaine des fréquences. Ainsi, pour obtenir un spectre en accélération depuis un signal en vitesse il suffit, dans l'équation 18, de diviser la transformée de Fourier discrète par $i\omega$ et inversement de multiplier par $i\omega$ pour réaliser l'intégration des accélérations vers les vitesses.

Pour certaines applications, il peut être plus utile de travailler sur la moyenne des composantes (généralement horizontales) plutôt que sur chacune des composantes individuellement. Plusieurs formulations existent, des plus simples comme la moyenne

arithmétique ou géométrique, aux plus complexes comme celles introduites par Boore (2006). Pour ce travail, c'est la moyenne quadratique qui a été privilégiée, car cette simple formulation $(^1/_2\,(\overline{U}_{EW}^2 + \overline{U}_{NS}^2))^{1/2}$ assure l'indépendance de la moyenne à l'orientation des composantes. Cependant, nous avons choisi une formulation légèrement différente qui permet de conserver également l'information de la phase des deux composantes (Steidl et al. 1996). Ainsi, la somme des enregistrements temporels des deux composantes horizontales est effectuée avant le calcul de la DSA

$$\overline{U_H} = \frac{\overline{U_{NS} + iU_{EW}}}{\sqrt{2}} \qquad 19$$

où U_H, U_{NS} et U_{EW} sont les signaux temporels de la moyenne des composantes horizontales, de la composante NS et de la composante EW respectivement.

Bien souvent, les spectres sont lissés pour en retirer les petites variations et éviter ainsi les divisions proches de 0 dans les méthodes empiriques utilisant les rapports spectraux (p. ex. Eq (5) à (8)). Ce lissage est réalisé ici suivant la méthode proposée par Konno and Ohmachi (1998), qui a l'avantage d'être défini dans un espace logarithmique et dont l'expression analytique est

$$W_B(f, f_m) = \left[\frac{\sin\left(log_{10}\left(\frac{f}{f_m}\right)^b\right)}{log_{10}\left(\frac{f}{f_m}\right)^b}\right]^4 \qquad 20$$

où b et f_m sont le coefficient de largeur de bande et la fréquence centrale respectivement. Plus la valeur de b est petite et plus le lissage est prononcé. Dans l'ensemble de ce travail, nous avons fixé la valeur $b = 30$. Parallèlement au lissage, les DSA qui sont initialement échantillonnées dans l'espace linéaire sur quelques milliers de points, sont rééchantillonnées selon une échelle logarithmique de 200 échantillons de façon à les alléger et pour satisfaire à la représentation logarithmique communément utilisée avec les fréquences.

Finalement, il est essentiel de prendre en compte la compétition existant entre la résolution des spectres à basse fréquence et la longueur de la fenêtre utilisée pour les calculer. Ainsi un seuil N, correspondant au nombre minimum de longueurs d'onde comprises dans la fenêtre temporelle de durée D, peut être utilisé pour déterminer la fréquence minimum

$$f_{min} = \frac{N}{D} \qquad 21$$

en deçà de laquelle la résolution spectrale n'est pas suffisante. Bien que nos analyses semblent montrer que $N = 3$ est un critère suffisant pour assurer une bonne résolution fréquentielle pour certaines bases de données (annexe 1), nous avons choisi de définir un minimum de 10 longueurs d'onde par fenêtre temporelle ($N = 10$) pour toutes nos analyses par sécurité.

1.3. Sites d'étude

Deux sites sont étudiés pour ce travail de thèse : le premier se trouve en Provence dans le sud-est de la France et le second près de la ville d'Argostoli sur l'ile de Céphalonie en Grèce. Celui en Provence est un bon exemple de cas réel de site industriel situé dans un contexte de sismicité faible à modérée, tandis que celui d'Argostoli est un bon cas d'étude de site soumis à une sismicité active. Ce dernier permet de tester les limites de validité des approches empiriques, en particulier lorsqu'elles sont transposées dans des zones plus faiblement sismiques. De plus, la confrontation des résultats sur deux sites d'études renforce la confiance dans les conclusions qui en sont tirées. Sur le site d'Argostoli deux campagnes de mesures ont eu lieu à deux périodes de temps distinctes : une première campagne a été réalisée entre 2011 et 2012 pour le projet NERA puis une seconde de début 2014 à 2015 associée au projet Sinaps@ et motivée par la forte activité sismique de l'ile suite à la crise sismique initiée fin janvier 2014. Le réseau vertical permanent Argonet, qui suit une géométrie déduite de la campagne Sinaps@, est opérationnel sur le site grec depuis juillet 2015. Deux bases de données, celle de Provence et celle de la campagne Sinaps@ à Argostoli, ont été constituées durant ce travail de thèse. Les différentes bases de données sont utilisées de façon disparate tout au long de ce manuscrit.

1.3.1. Provence

La Provence est une région du sud-est de la France de sismicité faible à modérée, située aux abords de la chaine des Alpes. Le site d'étude est une zone industrielle située à l'aplomb d'une petite vallée creusée au Miocène et comblée de sédiments. Sept sites de mesure (P1 à P7) y sont présents avec des distances inter-stations variant de quelques dizaines de mètres à quelques kilomètres. La Figure 8 présente la carte géologique de la zone d'étude ainsi que les positions des sites de mesure et les profils géologiques et de vitesse des ondes S relevés au niveau des sites P1 et P4 par mesures en forages (cross-hole, down-hole et P-S suspension logging) et par méthode de réseau de mesure du bruit ambiant à la surface. Les sites P1 et P2 sont localisés sur le substratum rocheux affleurant qui est constitué de calcaire crétacé très rigide. Le profil de vitesse en P1 montre une inversion de vitesse entre 20 et 30 mètres de profondeur due à la présence d'une zone fracturée. La paléovallée est un remplissage sédimentaire de dépôt miocène et pliocène rigide encaissé dans le substratum crétacé et surmonté de dépôts quaternaires plus récents et plus mous d'épaisseurs variables suivant la position dans la vallée. Cette dernière est orientée N135°E et s'évase vers le nord-ouest. Au niveau du site P4, la vallée est profonde d'une centaine de mètres et large d'environ 500 mètres. Le profil de vitesse montre cependant un fort contraste de vitesse dès

50 mètres de profondeur à l'interface entre les sables et les brèches (Figure 8). La V_{S30} a été estimée à partir de mesures géophysiques à : 2100, 1800, 440 et 720 m/s pour les sites P1, P2, P3 et P4 respectivement. Le site P1 est localisé dans une cave sismique à environ 3 m de profondeur contrairement aux autres sites qui sont tous situés dans des abris à la surface libre.

Chaque site de mesure est équipé d'un vélocimètre large-bande CMG6-TD Güralp dont la réponse instrumentale est plate entre 0,03 et 107 Hz. L'expérience a cependant montré que ces instruments n'ont pas une bonne sensibilité en dessous de 0,1 Hz. Ils incluent un numériseur intégré dont la fréquence d'échantillonnage est établie à 100 Hz. Sur les sites P2 et P3, un accéléromètre CMG5-TDE de Güralp est colocalisé avec le vélocimètre. Ce réseau opère depuis février 2012 pour les premiers instruments avec des durées d'enregistrement très variables suivant le site de mesure. Les sites P1, P2 et P3 sont ceux qui ont connu la plus longue période d'enregistrement s'étalant sur presque deux ans et demi jusqu'à l'établissement de la base de données en juin 2014. La Figure 9 montre les différentes périodes d'enregistrements (en vert) pour chacun des sites de mesure. Certaines périodes sans enregistrement (en rouge) correspondent à des problèmes instrumentaux pendant lesquels les capteurs n'étaient pas fonctionnels, mais la plupart de ces périodes sont synchrones pour tout le réseau vélocimétrique et se rapportent aux déploiements des instruments sur d'autres sites d'étude comme à Argostoli lors de la mission post-sismique qui débuta en février 2014.

Pas moins de 500 séismes ont pu être extraits sur cette période d'enregistrements continus à partir de la méthodologie décrite dans la partie 1.2.1.1.. Une instrumentation accéléromètrique préexistait de 2000 à fin 2011 sur les sites P2 et P3 et était opérée par le laboratoire de détection et de géophysique (LDG) du CEA. Ces instruments, qui fonctionnaient en déclenchement, ont permis l'enregistrement de 237 évènements supplémentaires. Deux vélocimètres 5s Lennartz ont également fonctionné de façon intermittente de la fin 2010 au début 2012 ajoutant 66 évènements à la base de données principalement sur le site P1.

L'ensemble des épicentres des 802 séismes enregistrés dans la base de données ainsi que la position du site d'étude sont représentés sur la Figure 10 (a). Cette base de données se compose entre autres des crises sismiques de Jausiers survenant après les séismes du 26 février 2012 ($M_l = 4{,}5$) et du 7 avril 2014 ($M_l = 5{,}2$). Ces deux crises sont approximativement colocalisées avec des épicentres à environ 120 km dans la direction N50°E par rapport au site d'étude et sont visibles sur la Figure 10 (a) et (b). Outre les nombreux évènements associés à ces deux crises, de nombreux séismes régionaux – exclusivement crustaux (<30 km de profondeur) – ont été enregistrés et sont originaires pour la plupart de la région Alpine située à l'ouest du site d'étude. Plusieurs téléséismes ont également été relevés dans les enregistrements.

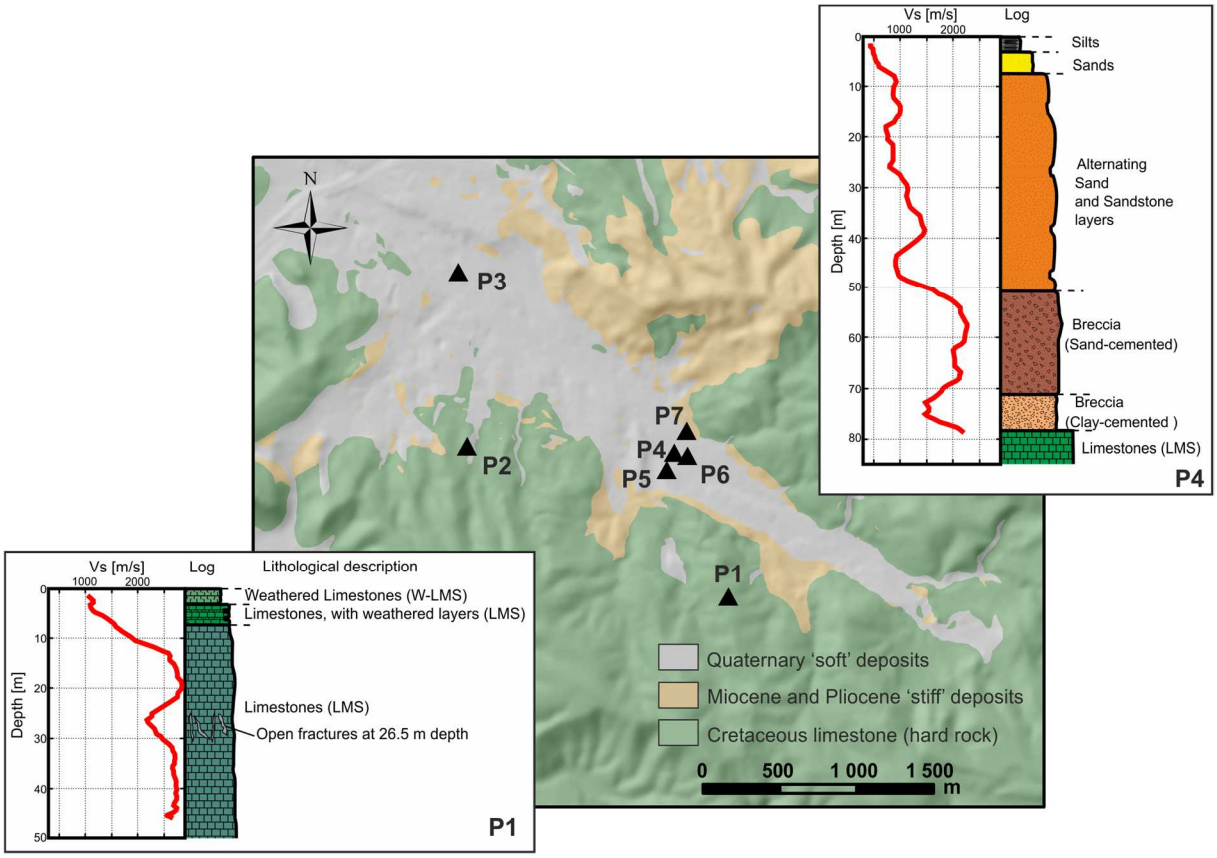


Figure 8. Carte géologique du site d'étude en Provence comprenant la position des sites de mesures P1 à P7 ainsi que les profils géologiques et de vitesse des ondes S relevés en forage au niveau des sites P1 et P4.

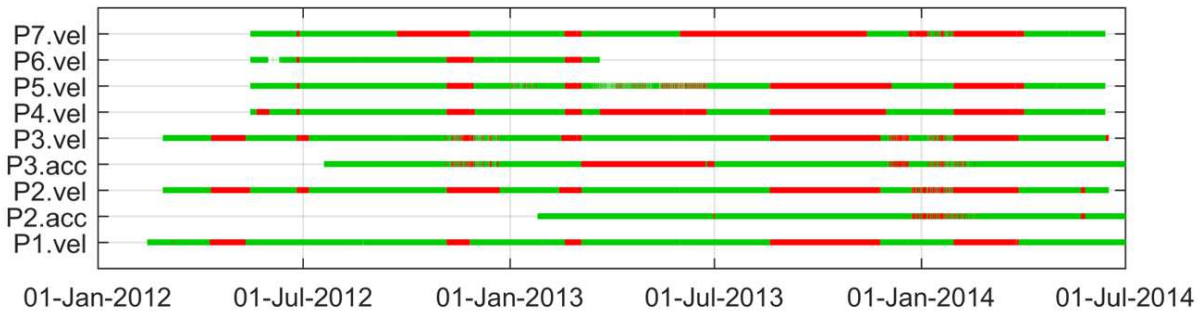


Figure 9. Périodes d'enregistrements (en vert) et sans enregistrement (en rouge) de l'instrumentation accélérométrique (acc) ou vélocimétrique (vel) continue récente présente sur les sites P1 à P7.

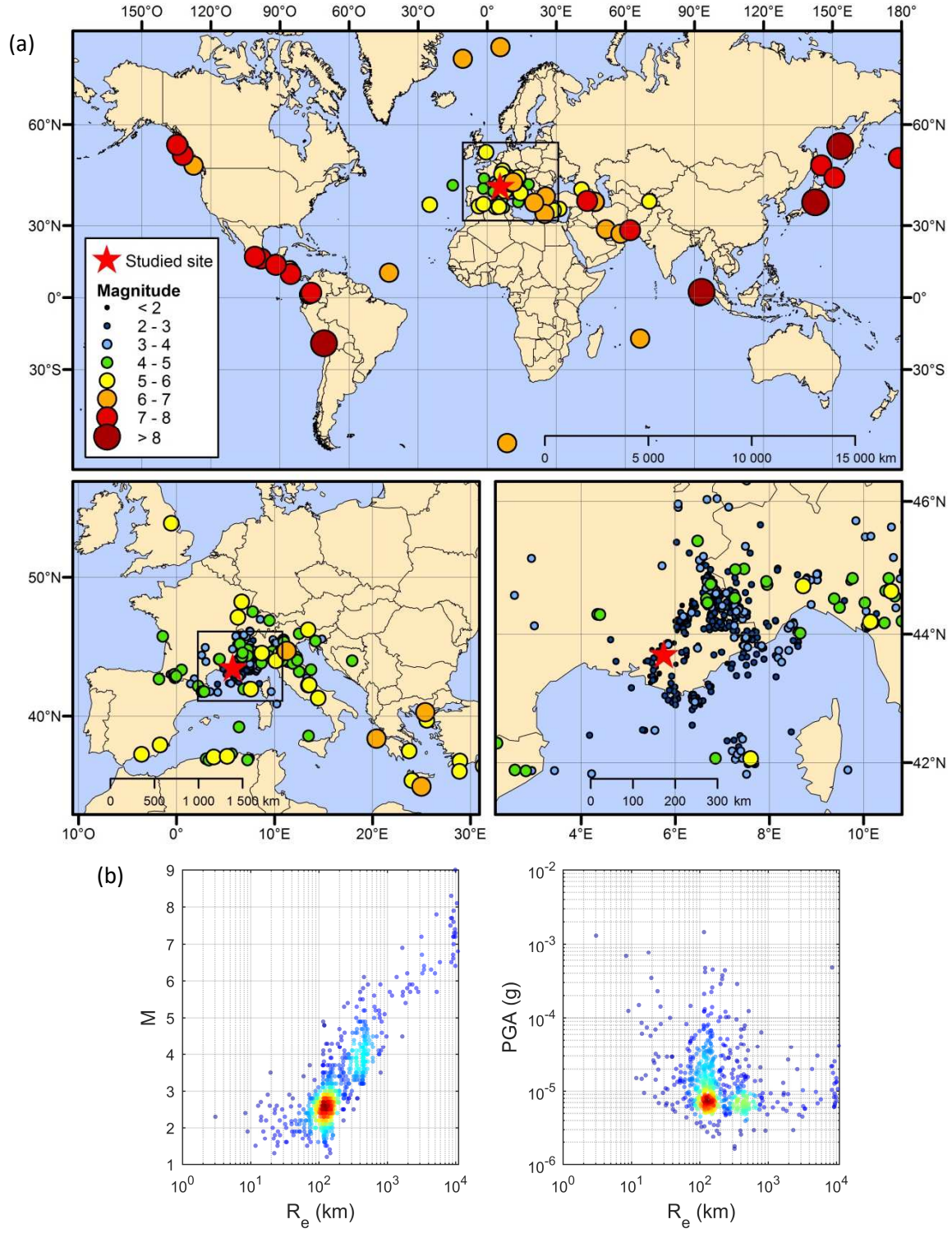


Figure 10. (a) Cartes à différentes échelles des épicentres des séismes enregistrés au site d'étude en Provence (étoile rouge). (b) Magnitude et PGA en fonction de la distance épicentrale pour les 779 évènements enregistrés par les instruments installés sur les sites rocheux P1 et P2 entre 2000 et 2014. L'échelle de couleur du bleu vers le rouge indique une densité croissante d'évènements.

La quantité de séismes qui a été enregistrée sur chacun des sites est variable puisque les périodes d'enregistrement varient d'un site à l'autre (Figure 9). Un exemple de la base de données constituée à partir du vélocimètre localisé sur le site au rocher P2 est présenté sur la Figure 11. Le rapport signal-sur-bruit relevé à 6 différentes fréquences y est présenté en fonction de la magnitude locale et de la distance à l'épicentre des séismes. Il apparait assez clairement que la majorité des évènements enregistrés sont régionaux (R_e<1000 km). Ceux-ci présentent les meilleurs SNR essentiellement dans la bande de fréquence intermédiaire (1–8 Hz). Plusieurs téléséismes sont tout de même visibles et apportent suffisamment d'énergie pour sortir du bruit seulement aux plus basses fréquences (<2-4 Hz). À haute fréquence (16 Hz), seuls les évènements les plus locaux (R_e<120 km) permettent la réalisation d'enregistrements présentant un SNR élevé. Dans l'ensemble, on notera les bons rapports SNR observés sur ce réseau bien qu'aucun mouvement fort n'y a jamais été relevé Figure 10 (b).

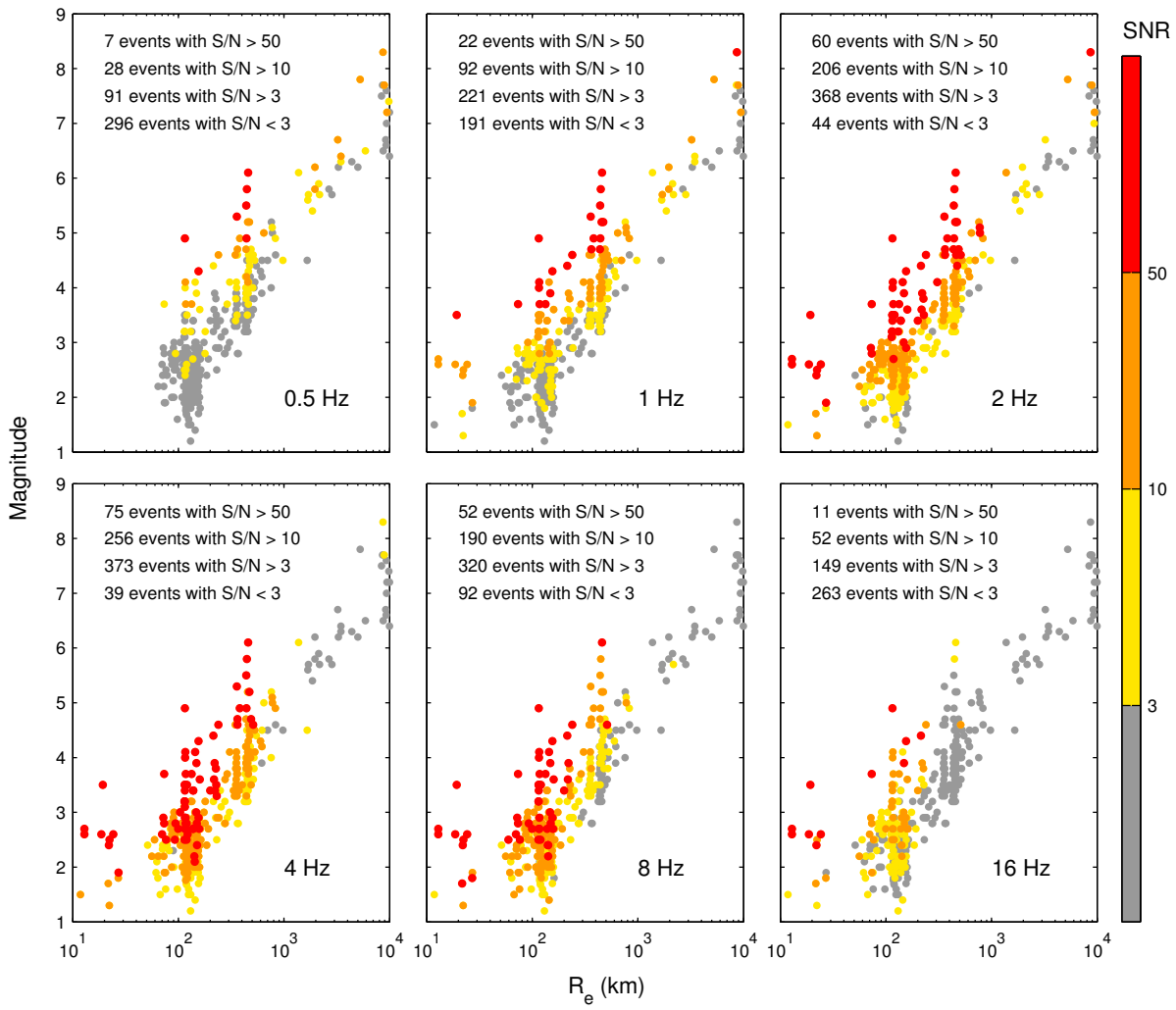


Figure 11. Rapport signal-sur-bruit (SNR – échelle de couleurs) à six fréquences (0,5, 1, 2, 4, 8 et 16 Hz) représenté en fonction de la magnitude locale et de la distance épicentrale pour 412

évènements enregistrés par le vélocimètre du site P2 entre février 2012 et juin 2014. De 0 à 1000 km le SNR est estimé depuis les ondes S tandis qu'au-delà de 1000 km celui-ci est évalué à partir des ondes P afin de maximiser l'énergie reçue depuis les évènements télésismiques.

1.3.2. Argostoli

Le nord-ouest de la Grèce est l'une des zones les plus actives de la région méditerranéenne et est celle qui enregistre la plus forte activitée en Europe. Elle est à la jonction des plaques eurasienne et africaine dont le taux de convergence est de quelques centimètres par an. L'ile de Céphalonie se situe à la terminaison nord-ouest de la subduction égéenne, à l'est de la faille de Céphalonie. Cette faille décrochante dextre de grande longueur réalise la transition entre la plaque africaine en subduction et la plaque continentale apulienne (Figure 12). L'ile de Céphalonie est soumise à de nombreux séismes, parfois de forte magnitude comme lors du séisme destructeur M_w 7,2 de 1953 ou plus récemment celui de Leucade M_w 6,2 de 2003, le doublet de Lixouri M_w 6,0 et 6,1 de 2014 et celui de Leucade M_w 6,5 de 2015. Cette forte activité sismique de l'ile combinée à un contexte géologique et géotechnique qui est proche de ce que l'on peut trouver ailleurs en Europe et qui est localement favorable à l'émergence de non-linéarité, explique l'intérêt pour la communauté scientifique d'y réaliser des mesures.

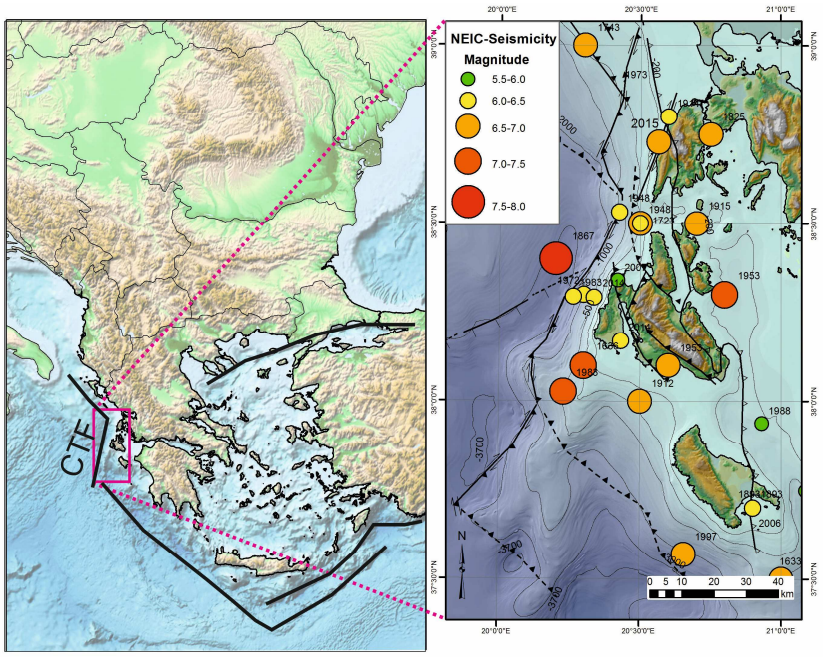


Figure 12. Caractéristiques géodynamiques de l'ile de Céphalonie (d'après Cushing et al. 2016)

Le site d'étude principal se situe dans le parc de Koutavos au sud-est de la ville d'Argostoli. Il s'agit d'un bassin sédimentaire de plus d'un kilomètre de large et épais d'une centaine de mètres sur sa partie la plus profonde. Ce bassin se caractérise par son asymétrie avec une bordure peu pentue sur son flanc sud-ouest et une bordure raide du côté nord-est où se concentrent les dépôts quaternaires les plus mous (Figure 13 (a)). Le substratum rocheux

est un calcaire crétacé qui, du fait de sa fracturation, présente des vitesses de propagation relativement faible pour ce type de formation (V_{S30}~830 m/s).

Plusieurs campagnes de mesure se sont succédées sur ce site expérimental. Les premières mesures ont été réalisées dans les années 90 par l'université d'Athènes (Protopapa et al. 1998). Ces données ne seront cependant pas utilisées ici. Dans le cadre du projet européen NERA (network of European research infrastructures for earthquake risk assessment and mitigation, projet FP7 n° 262330, www.nera-eu.org), 62 sismomètres furent déployés entre septembre 2011 et avril 2012 par un consortium de quatre laboratoires européens (Institut de la Terre, Grenoble, France, ISTERRE ; Instituto Nationale di Geofisica e Vulcanologia, Rome, Italie, INGV ; GeoForschungsZentrum, Potsdam, Allemagne, GFZ; Institute of Engineering Seismology and Earthquake Engineering, ITSAK). Pour notre travail (chapitre 4) seuls onze vélocimètres 5s Lennartz installés en travers du bassin par l'INGV ont été utilisés et sont indiqués sur la Figure 13 (a) de N1 à N11, N1 étant la station de référence placée sur le rocher affleurant. Ces données ont été utilisées principalement pour l'analyse du champ sismique ambiant puisque les analyses SSR avaient déjà été réalisées par Cultrera et al. (2014) dans un précédent travail.

En 2013, une campagne de caractérisation géologique et géophysique a été réalisée dans le cadre du projet PIA-SINAPS@ (www.institut-seism.fr/projets/sinaps/). Celle-ci permit, entre autres, d'établir la carte géologique présentée en Figure 13 (a) de ce document (Cushing et al. 2016). Au début de l'année 2014, toujours dans le cadre du projet SINAPS@, une campagne de mesure post-sismique fut décidée à la suite du doublet de séismes de Lixouri (Hollender et al. 2015). Cette campagne, à laquelle j'ai participé, a permis l'installation au rocher d'un réseau dense de 21 vélocimètres dédié principalement à l'étude de la variabilité spatiale du mouvement sismique à l'échelle du bâtiment (Imtiaz et al. 2017 ; Svay et al. 2017). En parallèle, cinq accéléromètres furent déployés, principalement dans le bassin de Koutavos, dans le but d'enregistrer d'éventuels mouvements forts et non-linéarités associées. Finalement, un rotatiomètre a également été testé aussi bien au centre du réseau dense que dans le bassin de Koutavos (Sbaa et al. 2016). L'ensemble des bases de données associées à ces trois instrumentations ont été constituées durant ce travail de thèse, cependant seules les données issues de deux sites du réseau accéléromètrique ont été utilisées pour ce travail : le site A2 qui est le site sédimentaire qui a eu la plus longue période de fonctionnement et le site de référence A1 qui est localisé au même endroit que N1, utilisé durant la campagne NERA (Figure 13 (a)). Les capteurs sont deux CMG5 de Güralp associé à une numérisation Taurus paramétrée avec une fréquence d'échantillonnage de 200 Hz. Finalement, ce sont plus de 6000 évènements qui ont été enregistrés entre février 2014 et juillet 2015, date à laquelle le réseau permanent Argonet – également initié par le projet SINAPS@ – a été mis en fonctionnement en lieu et place des deux sites A2 et A1.

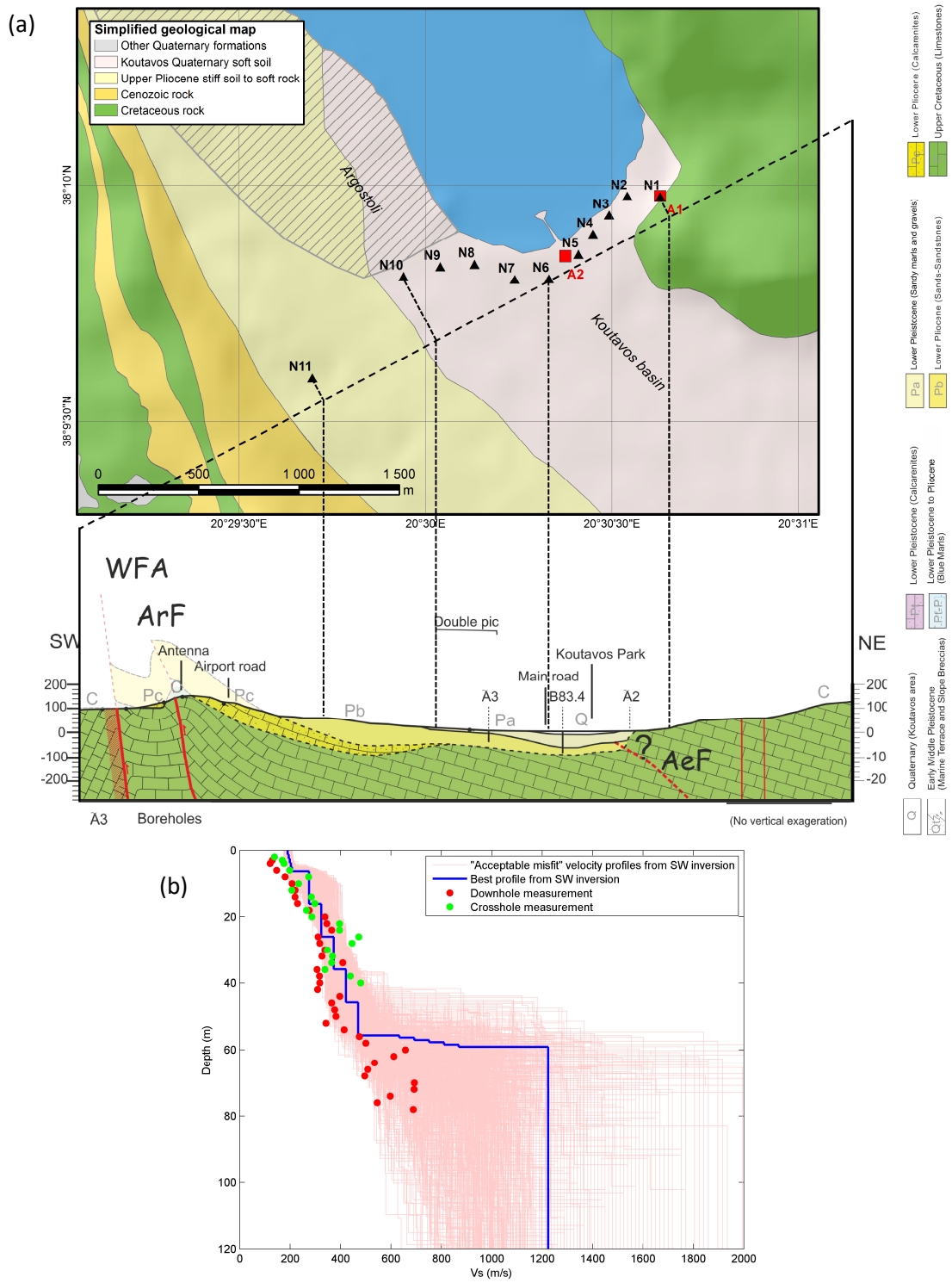


Figure 13. (a) carte et coupe géologique du bassin sédimentaire de Koutavos. Les sites de mesures de la campagne NERA sont numérotés de N1 à N11, tandis que ceux de la campagne post-sismique et du réseau Argonet sont référencés par A1 et A2. (b) profil de vitesse des ondes S avec la profondeur évalué en forage (cross-hole et down-hole) et par inversion des ondes de surface au niveau du site A2.

En plus des instruments EpiSensor Kinemetrics présents à la surface avec une numérisation Taurus, ce réseau bénéficie de trois forages au niveau du site A2 qui accueille trois instruments de puits à différentes profondeurs, bien que les données de puits n'ont pas été utilisées ici. Ces forages, dont le plus profond descend à 83,5 m de profondeur, ont également permis la réalisation des mesures (cross-hole et down-hole) du profil de vitesse d'ondes S avec la profondeur présenté en Figure 13 (b). Le réseau Argonet a permis l'enregistrement, entre autres, du séisme de Leucade M_w 6,5 de 2015. Une sélection de 277 séismes enregistrés par l'accéléromètre de surface en A2 et à la station de référence, toujours en A1, ont été utilisés pour compléter la base de données déjà très riche tirée de la période post-sismique. Au final, ce sont 6442 évènements qui ont pu être exploités et qui se concentrent majoritairement à une vingtaine de kilomètres à l'est du bassin de Koutavos (Figure 14). Le PGA maximal enregistré pour cette base de données est de 0,1 g ce qui n'a pas permis d'observer de non-linéarité. L'annexe 2 est un article soumis à la revue *Seismological Research Letters* qui détaille le travail et les bases de données qui ont été acquises durant cette campagne post-sismique.

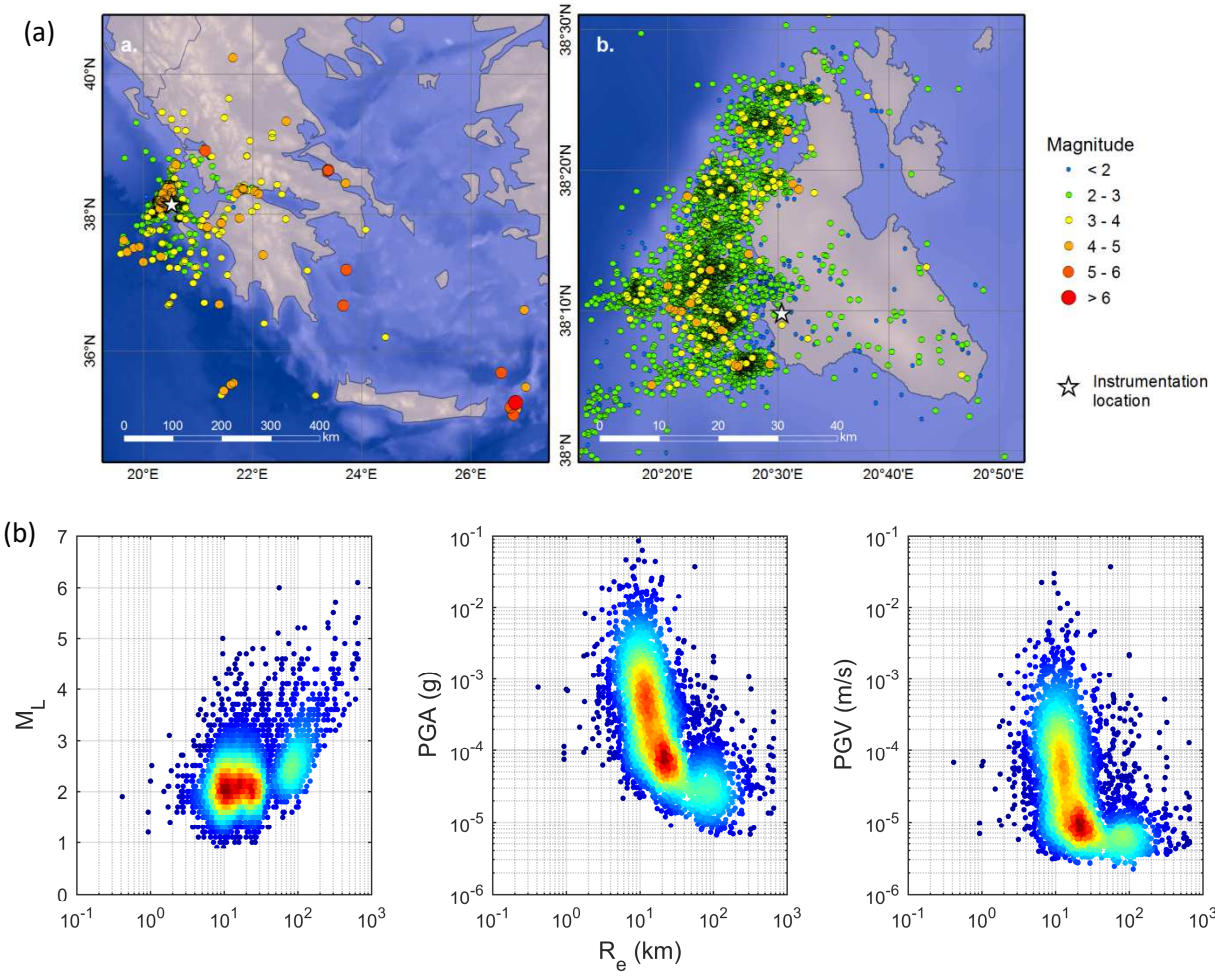


Figure 14. (a) Carte des épicentres des séismes enregistrés au site d'étude A2 par l'accéléromètre (étoile). (b) Magnitude locale, PGA et PGV en fonction de la distance épicentrale pour 6165

évènements de la base de données post-sismique et 277 évènements de celle de l'Argonet enregistré à la station au rocher A1. L'échelle de couleur du bleu vers le rouge indique une densité croissante d'évènements.

1.4. L'enregistrement sismique en zone de sismicité faible à modérée

La section précédente illustre bien le contraste existant, en termes de quantité et de qualité des enregistrements qui sont réalisables, entre les zones de sismicité faible à modérée comme dans le sud-est de la France et les zones fortement actives comme l'ile de Céphalonie. Ce contraste est d'autant plus marqué qu'en période de crise sismique comme celle de 2014, l'activité de l'ile est exceptionnellement élevée. Il n'en reste pas moins que le contexte de sismicité faible à modérée de la France est un argument qui est parfois utilisé pour ne pas instrumenter certains sites d'intérêt. Bien qu'il soit indéniable que la faible activité sismique est une difficulté pour l'application de méthodes nécessitant parfois une grande quantité d'évènements (SSR, Kappa, inversion généralisée, etc.), une part importante de cette réticence peut également s'expliquer par certains choix opérationnels qui sont classiquement réalisés.

Historiquement, l'essor de la sismologie de l'ingénieur a été permis par le développement des réseaux accéléromètriques partout dans le monde (cf. section 1.1.2.1), les vélocimètres large-bande étant généralement réservés à la sismologie classique du fait de leur grande sensibilité et parce qu'ils saturent lorsque l'amplitude du mouvement devient trop importante. À l'inverse, les accéléromètres sont généralement paramétrés pour enregistrer des mouvements forts (de plusieurs g), mais ils ne disposent alors pas d'une sensibilité élevée. Sur les sites provençaux P2 et P3, la colocalisation d'un vélocimètre et d'un accéléromètre (Figure 9) donne l'opportunité de comparer la qualité des enregistrements réalisés à partir de chaque instrument. La Figure 15 montre la répartition des évènements selon quatre intervalles de rapport signal-sur-bruit (SNR<3, 3≤SNR<10, 10≤SNR<50, SNR≥50) pour le vélocimètre CMG6-TD de Güralp et l'accéléromètre CMG5-TDE du même constructeur paramétré pour enregistrer des mouvements allant jusqu'à 2 g sur le site P2 et pour une même période d'enregistrement.

La Figure 15 montre très clairement le très grand écart de sensibilité existant entre le vélocimètre large-bande et l'accéléromètre de dynamique élevée. Il en résulte que pour toutes choses étant égales par ailleurs, le vélocimètre enregistre 10 à 20 fois plus de signaux de bonne qualité (SNR>3) que l'accéléromètre en moyenne à basse fréquence (<5 Hz). Pour ce qui concerne les signaux de très bonne qualité (SNR>10), cet écart dépasse même un facteur de 50. À très haute fréquence (>17 Hz), les deux instruments donnent des résultats assez équivalents. La très nette différence observée entre accélérométrie et vélocimétrie s'explique évidemment par la paramétrisation (2 g) des accéléromètres utilisés (qui reste cependant dans la gamme de grandeur classiquement utilisée) et également par le

relativement faible niveau de bruit ambiant du site provençal qui implique que le SNR est principalement contrôlé par la sensibilité des instruments. Il n'en reste pas moins que les vélocimètres permettent l'enregistrement de bonne qualité d'un bien plus grand nombre d'évènements que les accéléromètres.

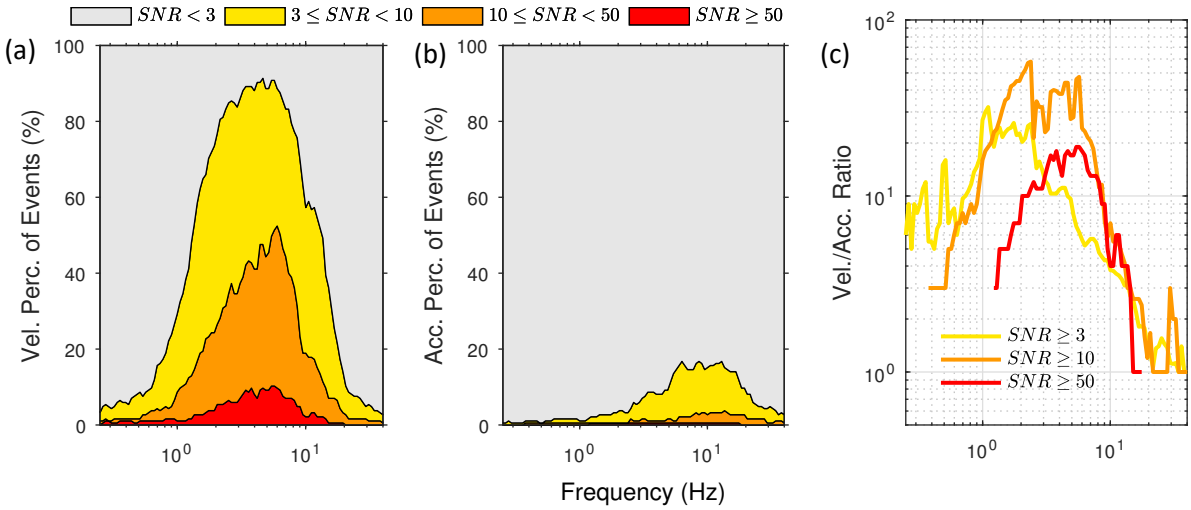


Figure 15. Répartition de 190 évènements (%) selon quatre intervalles de rapport signal-sur-bruit (SNR) pour un vélocimètre (a) et un accéléromètre (b) ayant enregistré durant une même période de temps sur le site P2. (c) rapport pour chaque intervalle de SNR du nombre d'évènements enregistrés par le vélocimètre et l'accéléromètre.

Outre le choix de l'instrument lui-même, le mode d'enregistrement – « continu » ou « sur déclenchement » – est également essentiel. L'enregistrement sur déclenchement a été longtemps privilégié, car il permet de limiter grandement l'espace nécessaire au stockage des données sismiques. Pourtant ce mode d'enregistrement a le défaut de n'enregistrer que les évènements les plus importants, dépassant le seuil de déclenchement. Ainsi, tout enregistrement de séisme ne dépassant pas le seuil de déclenchement fixé est perdu. À titre d'illustration, de 2000 à 2011 l'instrumentation accéléromètrique sur déclenchement opérée par le LDG a permis l'enregistrement de 235 séismes contre 298 évènements extrait durant les deux années et demie d'enregistrements continus du réseau accéléromètrique actuel. Évidemment, l'ensemble des évènements extraits depuis l'instrumentation récente ne sont pas d'aussi bonne qualité que ceux obtenus durant les onze années d'enregistrement sur déclenchement. Néanmoins, cela signifie que de nombreux évènements qui auraient pu être enregistrés si le mode de fonctionnement avait été en continu ne l'ont pas été. De plus, l'argument parfois défendu de l'espace requis par l'enregistrement continu n'a plus de justification réelle aujourd'hui où les avancées de l'informatique permettent le stockage et la gestion de grandes bases de données. À titre d'exemple, une année d'enregistrement continu pour les trois composantes d'un instrument paramétré avec une fréquence d'échantillonnage de 100 Hz requiert moins de 15 gigaoctets d'espace lorsque les fichiers sont compressés dans le format constructeur de Güralp.

Le choix souvent effectué de l'enregistrement accéléromètrique sur déclenchement explique en grande partie les réticences vis-à-vis de l'instrumentation en France. Par exemple, les règles de sureté pour l'instrumentation règlementaire des installations nucléaires imposent aux exploitants l'enregistrement certain de tout évènement entrainant des accélérations supérieures à 0,01 g au niveau du site (ASN 1984 - https://www.asn.fr/Reglementer/Regles-fondamentales-de-surete/RFS-relatives-aux-REP/RFS-I-3.b.-du-08-06-1984). Avec ce seuil de déclenchement, aucun séisme n'aurait été enregistré pendant les 14 années d'instrumentation du site provençal (Figure 10 (b)). Bien que ce site d'étude ne soit pas représentatif de l'ensemble du territoire métropolitain qui est en moyenne sismiquement moins actif que la Provence, les données qui y ont été acquises – en particulier depuis 2012 – montrent la grande quantité de séismes qui peuvent être enregistrés lorsque l'instrumentation est adaptée. De plus, pour les sites les moins actifs sismiquement, la Figure 11 montre que plusieurs téléséismes sont enregistrables partout avec de bons SNR jusqu'à des fréquences de 4 Hz environ, permettant ainsi de couvrir les fréquences fondamentales de la plupart des sites sédimentaires métropolitains les plus profonds (>~100 m). Une instrumentation adaptée signifie d'abord et avant tout que les instruments employés possèdent une réponse instrumentale plate dans la gamme de fréquences du site qui est ciblée et qu'ils enregistrent en continu. De plus, nous préconisons très fortement l'utilisation de vélocimètre large-bande afin d'assurer l'application des méthodes empiriques d'évaluation des effets de site en un temps suffisamment court. Ainsi, bien que les accéléromètres soient d'un intérêt premier pour l'enregistrement de séismes potentiellement destructeurs, ils ne doivent pas se substituer à l'instrumentation vélocimétrique large-bande pour l'étude des mouvements faibles dans les régions faiblement sismiques.

Chapitre 2

Évaluation de kappa en zone de sismicité faible à modérée

2.1. Contexte

Dans les régions faiblement sismiques, les évènements de forte magnitude (M>6) ont des périodes de retour de plusieurs décennies voir plusieurs siècles. Il n'est alors pas possible de constituer des GMPEs représentatives pour les séismes majeurs qui sont pourtant d'un intérêt premier pour la tenue des installations. Les GMPEs établis dans les régions fortement actives (États-Unis, Japon, Grèce, etc.) sont ainsi utilisées en remplacement. Ces équations empiriques sont pourtant évaluées pour les conditions de site des sites où elles ont été établies (host). Il convient donc de les ajuster aux conditions de site des régions pour lesquelles on souhaite les utiliser (target). Ceci est réalisé en pratique à l'aide des ajustements « Host-to-target ». Ces ajustements sont calculés à partir des proxys de sites (p. ex. V_{S30} et κ_0) des sites d'enregistrements host et target. Si V_{S30} peut être mesurée facilement sur n'importe quel site à partir de sondage géotechnique ou géophysique, il n'en va pas de même pour κ_0 qui nécessite généralement l'enregistrement de bonne qualité de nombreux séismes locaux à régionaux (R_e<200 km) de magnitude modérée à forte. Dans les régions fortement actives sismiquement, κ_0 peut être obtenu après quelques mois d'enregistrement seulement. À l'inverse, dans les régions plus faiblement actives, κ_0 est difficilement mesurable sur site et est généralement déduit de corrélations très incertaines que pourrait avoir ce paramètre avec V_{S30}.

Ce chapitre reprend l'article « Robustness of kappa ($\boldsymbol{\kappa}$) measurement in low-to-moderate seismicity areas: insight from a site-specific study in Provence, France » (Perron et al. 2017a) publié dans la revue « Bulletin of the Seismological Society of America ». L'objectif

principal de ce travail est de vérifier la mesurabilité de κ_0 dans un contexte de sismicité faible à modérée. La signification physique et la fiabilité des mesures de κ_0 sont discutées dans le contexte d'enregistrements sismiques de qualité et de quantité limitées.

2.2. Abstract

Determination of the site component of κ (κ_0) is important in the implementation of host-to-target adjustments for estimation of seismic hazard at hard-rock sites. Its evaluation through the classical approach of Anderson and Hough (1984) (κ_{0_AS}) faces specific difficulties in low-to-moderate seismicity areas, as the quantity and bandwidth of the usable data are generally limited. In such a context, κ measurements might have higher sensitivity to site amplification, frequency-dependent attenuation, the earthquake source and the instrumental equipment. Here, the κ_{DS} (displacement spectrum) approach of Biasi and Smith (2001) is compared with the κ_{AS} (acceleration spectrum) approach for three sites in an industrial area in Provence (southeastern France). A semi-automatic procedure is developed to measure individual values of κ_r that reduces inter-operator variability and provides the associated uncertainty. We show that this uncertainty is mainly dependent on the bandwidth used to determine κ_r. A good agreement is found between κ_{0_AS} and κ_{0_DS} for the two hard-rock sites, which yielded of ~30 ms. This highlights the κ_{DS} approach that is well adapted to low-magnitude events recorded at rock sites, and the use of velocimeters in low-to-moderate seismicity areas. The comparisons between these approaches are also used to infer the reliability of κ measurements by addressing their sensitivity to site amplification, frequency-dependent attenuation, and the earthquake source. First, the impact of site amplification on κ_0 estimates is shown to be very important and strongly frequency dependent for stiff-soil sites, and nonnegligible for hard-rock sites. Secondly, frequency-dependent attenuation cannot be ruled out for κ, as indicated by comparison with the literature quality factor (Q) for the Alps. Finally, a source component for κ_{AS} is questionable from the comparison of κ_{r_AS} evaluated for a cluster of events that shared the same path and site components.

Key words: kappa, site specific, low-to-moderate seismicity, reliability

2.3. Introduction

The kappa (κ) parameter describes the high-frequency spectral shape of ground motion. This parameter was introduced by Anderson and Hough (1984) as the linear decay in a log-linear space of the acceleration high frequency Fourier amplitude spectrum (FAS) of the horizontal component of the shear waves. For a given record at epicentral distance (R_e), κ (denoted κ_r) can be defined as:

$$A(f) = A_0 \exp(-\pi \kappa_r f), \quad f_1 < f < f_2 \qquad 22$$

where f_1 and f_2 are the frequency bounds between which the decay of the spectrum amplitude ($A(f)$) is approximately linear in a log-linear space. κ_r can be decomposed in terms of site (κ_0), source (κ_S), and path ($\tilde{\kappa}$) components:

$$\kappa_r = \kappa_0 + \kappa_S + \tilde{\kappa}(R_e) \qquad 23$$

Anderson and Hough (1984) assumed that κ_r can only be explained by the attenuation of the path and the site when it is measured above the corner frequency (f_c); i.e., where the acceleration spectrum of the source is assumed to be flat in the Brune (1970) model. This ω^{-2} source model was initiated by Aki (1967) and remains a reference model to date. According to the original model that neglected the source component ($\kappa_S = 0$), the distance-independent part of κ_r was attributed to κ_0; i.e., to the S-wave attenuation due to the geological structure beneath the recording site (Hanks 1982; Anderson and Hough 1984; Hough and Anderson 1988). The distance-dependence term that represents the attenuation of the S-wave along the propagation in the crust from the source to the site can be described by many different models. Generally, the linear assumption $\tilde{\kappa}(R_e) = m_\kappa \times R_e$ proposed by Anderson and Hough (1984) is a reasonable approximation (Douglas et al. 2010; Ktenidou et al. 2013). Then, when the source term is neglected, Equation (23) can be written as:

$$\kappa_r(R_e) = \kappa_0 + m_\kappa \times R_e \qquad 24$$

In this model, the site term κ_0 and the path term m_κ (following the notation of Douglas et al. 2010) can be simply separated by linear regression, where the first term is the intercept at zero epicentral distance ($\kappa_r(0)$) and the second term is the slope of the $\kappa_r(R_e)$ linear trend with epicentral distance. In Equation (24), κ_r and κ_0 are expressed in seconds (s) whereas m_κ is expressed in s/m, with the epicentral distance R_e expressed in meters.

The site component, κ_0, has many applications in hazard seismology, as it helps to constrain the high-frequency spectral shape of the predicted seismic signals at a specific site. This is particularly important for low-attenuating hard-rock sites, where the ground motion can be underestimated at high frequencies. κ_0 has thus been used as an input parameter in stochastic simulations (Boore 1986, 2003; Beresnev and Atkinson 1997; Graves and Pitarka 2010) and in the functional forms of some ground-motion prediction equations (GMPEs; e.g., Anderson et al. 1996; Laurendeau et al. 2013). However, the vast majority of GMPEs are developed using data from accelerometric networks in seismically active regions. Thus, the representativeness of the GMPEs for hard-rock sites is not ensured, as surface accelerometric stations are rarely installed on hard-rock sites. The host-to-target adjustments take into account differences in site properties (i.e., for the time-averaged V_S within the first 30 m [V_{S30}], and κ_0) to adapt the GMPEs from the host soft rock or rock where they are developed, to the target hard-rock sites where they are needed (Campbell 2003, 2004; Cotton et al. 2006; Van Houtte et al. 2011; Delavaud et al. 2012; Boore and Campbell 2017;

Ameri et al. 2017). Although an estimation of κ_0 is very often available in active host areas, its determination is more difficult in target areas of low-to-moderate seismicity. When no seismological recordings are available, κ_0 is generally deduced from the κ_0/V_{S30} correlation, even if the scatter of this correlation is very large. However, this introduces large uncertainties in seismic hazard assessments. It is thus of great interest to determine reliable site-specific values of κ_0 from seismic recordings, which can be relatively challenging in low-to-moderate seismicity areas.

Since the first definition of κ by Anderson and Hough (1984), many studies have proposed different techniques to determine κ_r or κ_0. Ktenidou et al. (2014) provided a comprehensive review of the methods at the time, and provided the notations that are followed here. The evaluation of κ_0 through the original definition based on the acceleration spectrum (κ_{0_AS}) is difficult in low-seismicity areas, because of the lack of local earthquakes with magnitudes >3. Indeed, the lower the magnitude, the higher the f_c and the lower the highest frequency with good signal-to-noise ratio (SNR >3), which leads to a smaller usable width of the frequency window ($\Delta f = f_2 - f_1$) for the κ_{r_AS} measurement. Due to this difficulty, only one study has explicitly reported an estimation of κ_{0_AS} for mainland France (Douglas et al. 2010), and to do so, they joined the individual κ_{r_AS} measured from many sites of the same type and in the same region.

The approach proposed by Biasi and Smith (2001) represents an alternative for low-seismicity areas, which estimates κ_{r_DS}, and then κ_{0_DS}, on the horizontal components of the FAS computed from the direct shear-wave part of the displacement seismogram. The displacement spectrum of the source flattens up to f_c, which allows the measurement of κ_{r_DS} for low-magnitude events and at lower frequencies (i.e., below f_c). This relies on the assumption that the stress drops for the smallest earthquakes are similar to those of the large earthquakes, which implies high f_c values (Kilb et al. 2012). In contrast to κ_{r_AS}, the lower the magnitude of the earthquake, the larger the Δf for the measure of κ_{r_DS} on the record. This is why this approach was initially proposed for very small magnitude events (M <1). Finally, local-to-regional earthquakes (R_e <200 km) listed in the catalog for low-to-moderate seismicity areas have magnitudes mainly between 1 and 3, which is not in the ideal magnitude range for either the κ_{AS} and κ_{DS} approaches. This difficulty can lead to higher sensitivity of the results to the site amplification, the frequency-dependent attenuation, and the earthquake source, due to the small Δf that are used for κ_{r_AS} and κ_{r_DS}.

Indeed, as the physics of κ are not clear, and wide uncertainty is generally associated with its measurement, this also results in a multiplicity of possible interpretations. The regain of interest in this parameter over the last decade has recently led to numerous studies of its dependence on various parameters and to a reduction in the associated uncertainties (Campbell 2009; Van Houtte et al. 2011; Kilb et al. 2012; Ktenidou et al. 2013, 2015; Edwards et al. 2015; Parolai et al. 2015). First, since the origins of κ, some studies have attributed part of the decay to source effects (Papageorgiou and Aki 1983; Aki 1987; Papageorgiou 1988, 2003; Gariel and Campillo 1989), while a few studies have suggested that there might be both source and site components for κ_r (Tsai and Chen 2000; Purvance

and Anderson 2003). Although the site-effects interpretation is commonly accepted at present and the majority of recent studies of κ neglect the source term, an influence of the source on κ_r is likely if there is any divergence from the ω^{-2} source model. Moreover, the f_c criterion that allows for the neglecting of the source influence is difficult to respect, as its estimation is very uncertain, especially when the value of the stress drop for the target region is not known. Secondly, one of the most dubious assumptions concerning κ is its frequency independence. This assumption is an implicit part of the choice of a linear model to measure κ_r between f_1 and f_2 of the acceleration spectrum (Eq. (22)). κ_r is tied to another attenuation parameter: the effective quality factor of the S-wave, Q_{ef} (Futterman 1962; Knopoff 1964). Campbell (2009) provided a good overview of the relationship between Q and κ. Since Q_{ef} was introduced, it has been widely accepted as frequency dependent at least in part. The model proposed by Aki (1980) and Dainty (1981) divided Q_{ef} into two parts: a frequency-independent intrinsic attenuation part (Q_i), and a frequency-dependent scattering part (Q_{sc}), given by Equation (25):

$$\frac{1}{Q_{ef}} = \frac{1}{Q_i} + \frac{1}{Q_{sc}} \qquad 25$$

Based on the frequency-dependent t* model of (Cormier 1982), Hough et al. (1988) and Hough and Anderson (1988) linked Q and κ with a general frequency-independent model. This model described the attenuation along the ray path as:

$$\kappa_r(r) = \int_{path} \frac{1}{Q_i(z)\, V_S(z)} dr, \qquad 26$$

where $Q_i(z)$ is the frequency-independent component of Q_{ef}, and V_S is the shear-wave velocity at depth z within the profile. This model assumes that Q_i and V_S are laterally homogeneous, and that Q_{sc} does not affect the evaluation of κ when it is inversely proportional to the frequency (Warren 1972; Rovelli 1982; Anderson 1986). However, the frequency independence assumption for κ is dubious, as depending on the size of the heterogeneities, a frequency-dependent scattering contribution cannot be excluded (Edwards et al. 2015; Parolai et al. 2015; Ktenidou et al. 2015). This might impact upon κ, depending on the frequency band in which it is defined, which will lead to different results when using different approaches (e.g., high frequency κ_{AS}, low frequency κ_{DS}, κ_{BB} broadband inversion). Finally, another frequency-dependent phenomenon can modify the spectrum and therefore the evaluation of κ: the site amplification. Indeed, the spectral modulations induced by site effects can change the slope of the decay and thus modify the κ_r estimates, depending on the selected frequency windows. Moreover, the modification of the spectrum shape can hide the frequency interval where the decay should be linear in the absence of site amplification, and thus alter the identification of the “true” frequency band where κ_r should be measured

(Hough et al. 1999; Parolai and Bindi 2004; Van Houtte et al. 2014; Edwards et al. 2015). The smaller the Δf, the greater the influence of the site amplification on κ should be.

The objective of the present study is to evaluate the applicability of reliable determination of site-specific κ in the low-to-moderate seismic context of mainland France. After a short description of the study area in terms of its geology and the datasets, some recommendations are provided for implementation of the instrumentation, and the site effects are evaluated. First, the semi-automatic procedure used to measure κ_r is introduced, and a detailed comparison is given between the κ_{AS} and κ_{DS} approaches on hard-rock sites. Secondly, the sensitivity of κ to frequency-dependent attenuation, site amplification, and the earthquake source are investigated. Finally, the reliability and variability of the κ measurements are discussed in the context of low-to-moderate seismicity areas.

2.4. Study area and datasets

2.4.1. Study area

The industrial area under study is in Provence, close to the Alps (southeastern France). The Alps are one of the most active seismic regions in mainland France, although the associated seismic activity is low to moderate (Guéguen et al. 2007a; Sanchez et al. 2010). Figure 16 shows the location of the study site and the event epicenters from the database used. The main database is composed of seismic data that were recorded between February 2012 and June 2014, with the recording of nearly 500 earthquakes by several velocimeters (Güralp CMG6-TD). During this 2-year period, two seismic sequences occurred, after the February 26, 2012 (M_L = 4.5) and April 7, 2014 (M_L = 5.2) earthquakes of Jausiers. These two sequences are approximately co-located at R_e = 120 km and at an approximately N50°E azimuth from the recording area (Figure 16). In the framework of this κ study, only the three sites where the seismic records are the most abundant are considered (Figure 17, P1, P2, P3). Sites P2 and P3 have two accelerometers (Güralp CMG5-TDE) as well as the velocimeters. All of the sensors record continuously with a 100 Hz sampling frequency and a flat response beyond the Nyquist frequency (50 Hz). Seismic events were extracted from the continuous data using the earthquake bulletin information provided mainly by the Euro-Med Seismological Centre. When information was missing for an earthquake, the information used was from the *Réseau National de Surveillance Sismique* (French National Seismic Surveillance Network), Géoazur, or the Italian Seismological Instrumental and Parametric Database. These catalogs were also used to determine the earthquake parameters (e.g., magnitude, location, among others), where the magnitudes were mainly local (M_L). Two accelerometers in triggering mode completed the database, with 300 additional events recorded from 2000 to 2011 at sites P2 and P3. This initial instrumentation was managed by the Laboratory for Detection and Geophysics (CEA, France), which also provided the associated earthquake parameters. Differences between the catalogs are assumed to be negligible compared to the uncertainty associated to κ. Finally, more than 800 events were

recorded, with epicentral distances from 3 km to >10000 km. Some teleseisms were also recorded, although the vast majority of events were within epicentral distances of 500 km. All of the recorded regional earthquakes were crustal events (depth, <30 km) and corresponded to weak motions. Almost all of these had local magnitudes <4, and were north-east of the recording site. However, the number of events recorded by each site varied due to differences in the recording durations, and this was very dependent on the application (Tableau 2). For κ estimation, only the best records from the closest events ($R_e < 180$) are used, to provide good SNRs over a broad enough frequency band, and to ensure that the propagation is only in the crust.

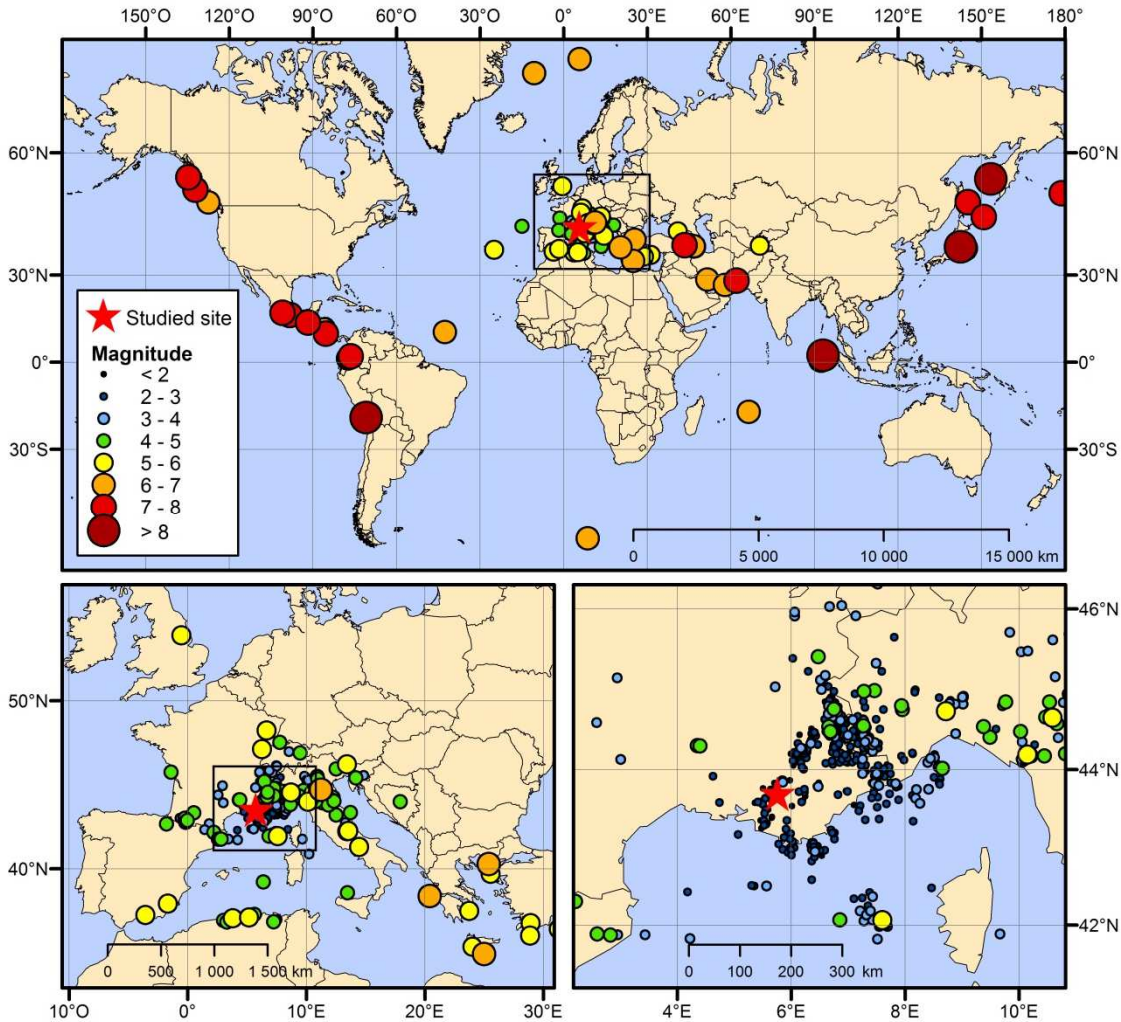


Figure 16. Maps of the earthquake epicenters recorded at the studied site (red star), at three different scales.

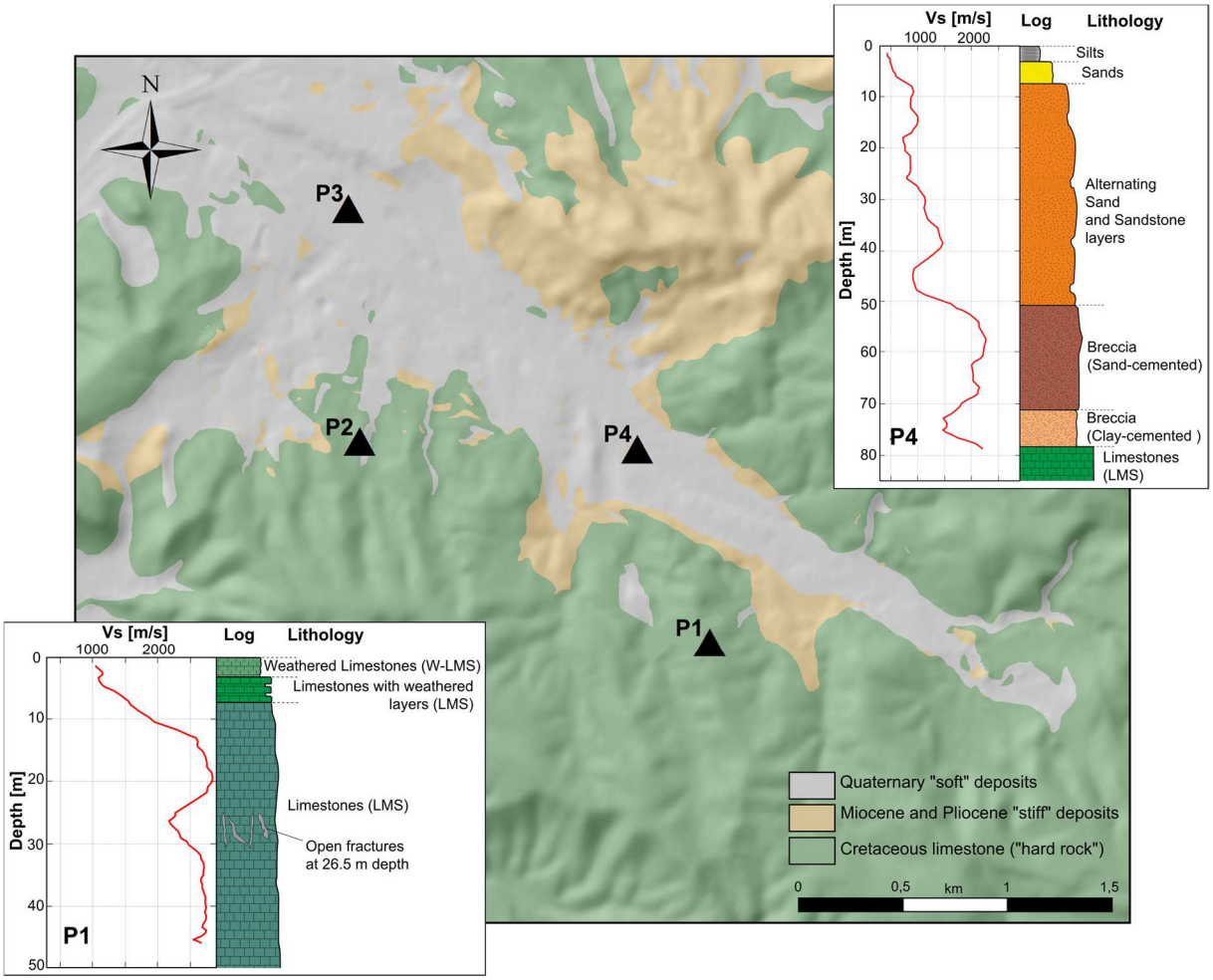


Figure 17. Geological map of the recording area. Sites P1 and P2 are located on hard rock, while sites P3 and P4 are located on stiff soils. At sites P1 and P4, boreholes allowed the recording of the velocity profiles with depth (V_S, V_P) for the different techniques: the cross-hole, down-hole, and P-S suspension logging methods.

Tableau 2. Number of events available according to the site and to the application.

Site	Number of events available			
	Total	For the $SSR_{/P1}$	For κ_{AS}	For κ_{DS}
P1	453	-	33	37
P2	678	371	35	39
P3	686	350	48	18
P4	246	205	-	-

Sites P1 and P2 are located on outcropping massive Cretaceous limestone. Site P3 and a further site, site P4, are located within a relatively small paleovalley (a few hundred meters wide, 50-150 m deep) that is filled with stiff Miocene sand and sandstone, and softer quaternary deposits. Based on the geophysical measurements for sites P1, P2, P3, and P4, V_{S30} is evaluated at 2100 m/s, 1800 m/s, 440 m/s, and 720 m/s, respectively. Sites P1 and P2 are thus classified as the 'hard-rock' class, whereas sites P3 and P4 are in the 'very dense soil' class, according to the *National Earthquake Hazards Reduction Program* classification. The sensor at site P1 is set up in a seismic vault buried at 3 m in depth, while the sensors of sites P2, P3, and P4 are at the surface. Figure 17 shows the locations of these four sites on a simplified geological map. For sites P1 and P4, three cored boreholes had been drilled, which provided a lithological description of the substratum, as well as *in-situ* shear-wave velocity measurements performed using cross-hole, down-hole, and P-S suspension logging methods. Site P1 was one of the sites used by the InterPACIFIC Project to perform a comparative benchmark of invasive and non-invasive methods for site characterization (Garofalo et al. 2016). No κ evaluation was carried out for site P4, due to too low a number of well-recorded seismic events. However, site P4 is included here due to the availability of borehole and *in-situ* V_S measurements, which are representative of the local basin features.

2.4.2. Spectrum computation

The FAS are mandatory to compute κ_r and for site effects assessment through horizontal-to-vertical spectral ratio and standard spectral ratio (SSR) approaches. A common procedure was followed to determine the FAS from the earthquake recordings (Perron et al. 2017b). A visual check and manual picking of the P-wave and S-wave first arrivals (T_P, T_S) were performed for one site of the network (generally P1) for each earthquake. It is assumed that the differences in the time arrivals between these sites are negligible due to the very short inter-station distances compared to the epicentral distances. In addition to T_P and T_S, the end of the signal (T_{end}) was also visually picked, based on a time-frequency analysis (spectrograms), to take into consideration the SNR criteria and to detect potential post-event perturbations at every frequency (e.g., aftershocks, transient noise, among others). Only the direct S-wave window was considered for κ estimation, while the entire signal was used to assess the SSR. A 5% cosine taper was applied at the edge of each time window, and the windows were extended to apply the cosine taper out of the target window. The S-wave duration was defined by a specific, and relatively simple, scheme (Perron et al. 2017b) that took into account the expansion due to the propagation (approximated by $T_S - T_P$) and the source (through $1/f_C$). In the low-to-moderate seismicity context of Provence, the S-wave window is mainly controlled by the propagation term, as the source term is negligible for magnitudes <5. A minimum nominal duration of 5 s was used to constrain the spectral resolution at low frequencies. The influence of window length on the spectrum was tested, which led to only small changes, in agreement with previous observations (Anderson and Hough 1984; Tsai and Chen 2000; Douglas et al. 2010). To obtain length-independent FAS, the Fourier transforms were normalized by the square root of the number of samples, which led to the computation of the FAS density (FASD). The FASD is important only for the SNR

computation when the noise and the signal windows are not of the same duration. The north–south and east–west components were combined, to obtain a single orientation-independent component, as follows:

$$S(H) = \frac{S(N + iE)}{\sqrt{2}} \qquad 27$$

This evaluation of the horizontal mean component is equivalent to the quadratic mean in the frequency domain [$S(H) = \sqrt{(S(E)^2 + S(N)^2)/2}$]. Nevertheless, this complex representation of horizontal motion allowed it to be applied to the time domain, and maintained the phase between the components (Steidl et al. 1996). A criterion of a minimum of 10 wavelengths contained in the signal was applied to define the minimum frequency (f_{min}), which is determined according to the duration of the time window (Δt): $f_{min} = 10/\Delta t$. For κ_{AS} or κ_{DS}, the spectra (in acceleration or displacement) were obtained from the velocity spectra by multiplication or division by $2\pi i f$ in the Fourier domain.

2.4.3. Velocimeter *versus* accelerometer

As there is co-localization of accelerometers with velocimeters at sites P2 and P3, these data were used for comparisons between these two types of sensors, in terms of the quantity and quality of the records. The quality of a dataset impacts directly on the achievability of the κ measurement. Figure 18 shows the comparison between the number of accelerometer and velocimeter recordings that satisfied similar quality criteria at the same rock site (site P2) over the same period of time. The quality criteria are based on the SNR at each frequency and for each recording, with different threshold values considered for the SNR (i.e., 3, 10, 50). Figure 18 thus shows the percentages of the recorded events for which the (frequency-dependent) SNR fall within the corresponding ranges (i.e., SNR <3; 3 ≤ SNR < 10; 10 ≤ SNR < 50; SNR ≥50). This shows that the velocimeter recordings provide more to many more usable events, especially below 20 Hz (sometimes >50-fold for 10 ≤ SNR < 50), and the available frequency ranges are mainly from 0.25 Hz to 15 Hz, which is below the high frequency range generally required for κ_{AS}. Above 15 Hz, the two types of sensors give similar results, even if the number of recorded events with SNR >3 is relatively low due to the lack of local earthquakes. It should be noted, however, that these data were obtained for a given accelerometer model *versus* a given velocimeter model, as the purpose here is not to achieve complete instrumental comparisons. However, although the use of accelerometers is justified for strong ground-motion recording (as they do not saturate), these experimental results demonstrate the interest in using velocimeters for site-specific studies in low-to-moderate seismicity areas, to record enough earthquakes within a reasonable time period. If only accelerometers were available, few κ_r only and no κ_0 evaluations would have been possible for this study, especially for κ_{DS}, which was evaluated mainly at low frequencies (i.e., below 15 Hz).

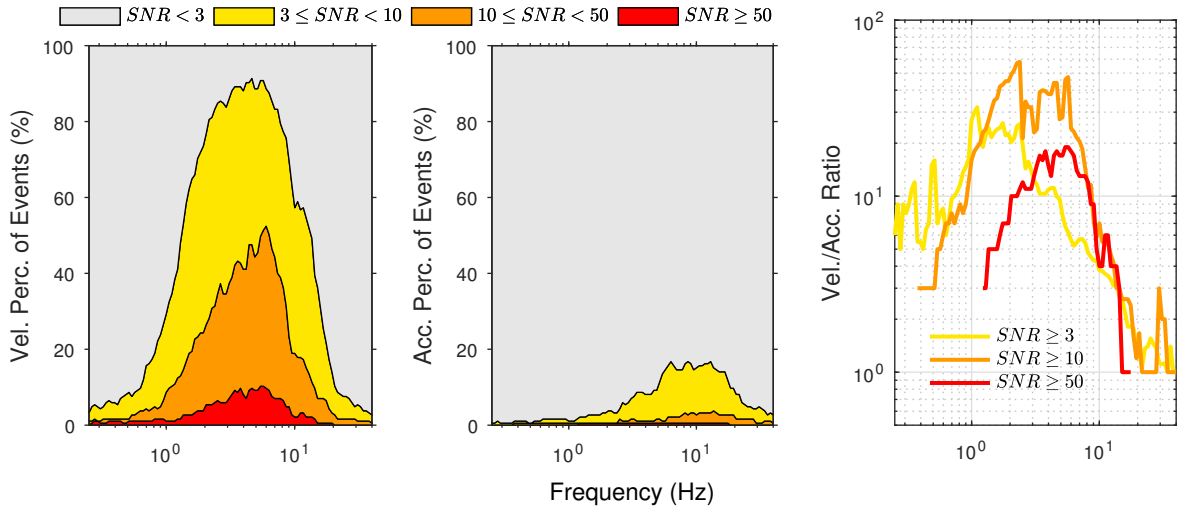


Figure 18. Left and central panels: Comparisons of the percentage of velocimeter (left) and accelerometric (middle) recordings that satisfy four ranges of signal-to-noise ratios (SNRs; as indicated), as a function of the frequency. Both instruments recorded in continuous mode, at the same site (P2), and over the same period of time. The S-wave windows from a total of 185 earthquakes were considered. The right panel shows the ratio between these velocimeter and accelerometer recordings that satisfy the same SNR criteria.

2.4.4. Site amplification

Seismic ground motion can be modified (most often amplified) by the near-surface geological structure anywhere at the Earth surface. This phenomenon is referred as the site effects, which have been widely observed for alluvial deposits, with amplitudes and frequency bands that vary greatly from site to site as a function of their geometry and mechanical properties. However, it is often neglected for hard-rock sites, because their amplitude is expected to be much lower, and to be shifted to high frequencies only (i.e., beyond 5-10 Hz). These frequency-dependent phenomena have to be evaluated on a site-specific basis, as they might significantly contaminate the measurement of κ_r based on the apparent spectral slope.

The records of numerous earthquakes for each site allows the inference of the relative transfer function using the SSR approach (Borcherdt 1970). One important precondition for using the SSR technique is the availability of a nearby reference (i.e., rock) site with negligible site effects. This approach consists of computing the ratio between the FAS from the earthquake recorded at both the site and the reference site. The FAS were processed following the procedure described in the 'spectrum computation' section, and were smoothed following the Konno and Ohmachi (1998) procedure, with a *b*-value of 30. For each frequency, the median was estimated from all of the earthquakes with SNR >3. Figure 19 shows the SSR data for the mean horizontal components and the vertical components for sites P2, P3, and P4, using P1 as the reference site. The theoretical one-dimensional (1D) site transfer function at P1 is shown in Figure 19 (black dotted curves). This transfer function was computed with a velocity profile that used the measured velocity profile down to 46 m, where Vs reached approximately 2800 m/s (Figure 17). This was then completed down to

8 km in depth (Vs = 3600 m/s), with a generic velocity profile to account for crustal amplification. The 1D reflectivity model (Kennett 1974) was used to reproduce the response of horizontally stratified layers excited by a vertically incident SH plane wave. An infinite Q was used for the computation to consider only the site amplification.

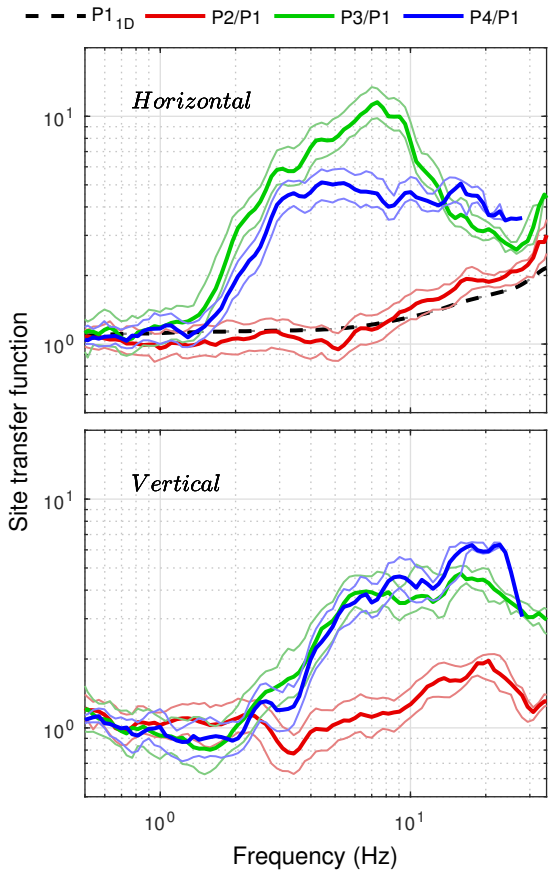


Figure 19. Median and percentile 16% and 84% of the standard spectral ratios estimated from the earthquakes recorded at sites P2 (red), P3 (green), and P4 (blue), according to the reference site P1, for the horizontal mean (top) and the vertical (bottom) components. The theoretical transfer function estimated from the velocity profile at site P1 is given by the black dotted curve for the horizontal mean component.

Site P3 shows significant site amplification above 2 Hz (e.g., up to 12-fold at 7 Hz on the horizontal component), while site P4 shows more moderate amplification (up to 5-fold at 4 Hz). These amplifications appear to be mostly related to the first 55 m of soil, according to the V_S profile shown in Figure 17. According to the P2/P1 SSR, as well as the theoretical 1D transfer function computed at site P1, the amplification at the rock sites is much lower. This is due to the weathered zone that affects the limestone within the first few meters beneath the surface.

In addition to these lithographic effects, the topography of the free surface near the site can also modify the spectral shape, and thus the evaluation of κ_r, especially on rock sites where the lithographic effects are limited and the topography is important. However, negligible

influence of the topography was noted for each site through the frequency-scaled curvature approach proposed by Maufroy et al. (2015c).

2.5. Kappa

2.5.1. Data processing

Once the horizontal mean FAS or FASD have been processed (Eq. (27)), κ_r can be determined following the methodology proposed by Ktenidou et al. (2013). The slope of the spectral decay is measured by the linear regression from the acceleration FAS for κ_{r_AS} and from the displacement FAS for κ_{r_DS}. An example of the κ_{r_AS} measurements is given in Figure 20. Care was taken to be sure that the frequency window within the slope that was measured had SNR >3. In the same way, special attention was paid to the frequency window, to be sure that it was above f_c for κ_{r_AS} and below the f_c for κ_{r_DS}. This f_c checking is essential for the assumption that the result is independent of the shape of the source spectrum in the Brune (1970) model. Direct visual evaluation of f_c was carried out on the displacement spectra of a few earthquakes, and comparisons were made with the value proposed by Drouet et al. (2010) for the Alps. Then, the initial bounds of the frequency window (f_{1ini}, f_{2ini}) were manually picked, respecting the SNR and f_c criterion, and for the most linear decay.

A semi-automatic procedure was developed for more precise and repeatable selection of the lower and upper bounds (f_1, f_2) of this frequency window. The aim is to reduce the variability between operators and to determine the uncertainty associated to each measure of κ_r. This procedure is illustrated in Figure 20: an uncertainty range (δf = ±2 Hz) is defined around each bound of this manually selected frequency window (f_{1ini}, f_{2ini}), and κ_r is estimated from the linear regression slopes over all of the frequency interval combinations ($f_{1ini} \pm \varepsilon_1 \delta f$, $f_{2ini} \pm \varepsilon_2 \delta f$), with ε_1 and ε_2 as random numbers between 0 and 1 (Figure 20, yellow lines). In this way, the precision of the κ_r estimate can be quantified with various statistical parameters (e.g., minimum, maximum, mean, standard deviation). The best κ_r estimation is defined as that which minimizes the root mean square (*RMS*):

$$RMS = \sqrt{\frac{\sum_i (FAS(f_i) - FAS_{fit}(f_i))^2}{N}}, \qquad f_1 \leq f_i \leq f_2 \qquad 28$$

where $FAS(f_i)$ is the S-wave FAS, $FAS_{fit}(f_i)$ is the regression prediction at the ith frequency, and N is the number of samples between the f_1 and f_2 bounds of the tested slope. Only windows wider than 10 Hz were taken into account, to ensure the minimum reliability for the κ_r estimation. Finally, the best fit κ_r estimate (Figure 20, red line) that minimizes RMS is taken with its associated uncertainty, which corresponds to the difference between the

maximum and minimum values of κ_r (Figure 20, blue lines) obtained in the tested slope set ($\Delta\kappa_r = \kappa_{r_{max}} - \kappa_{r_{min}}$).

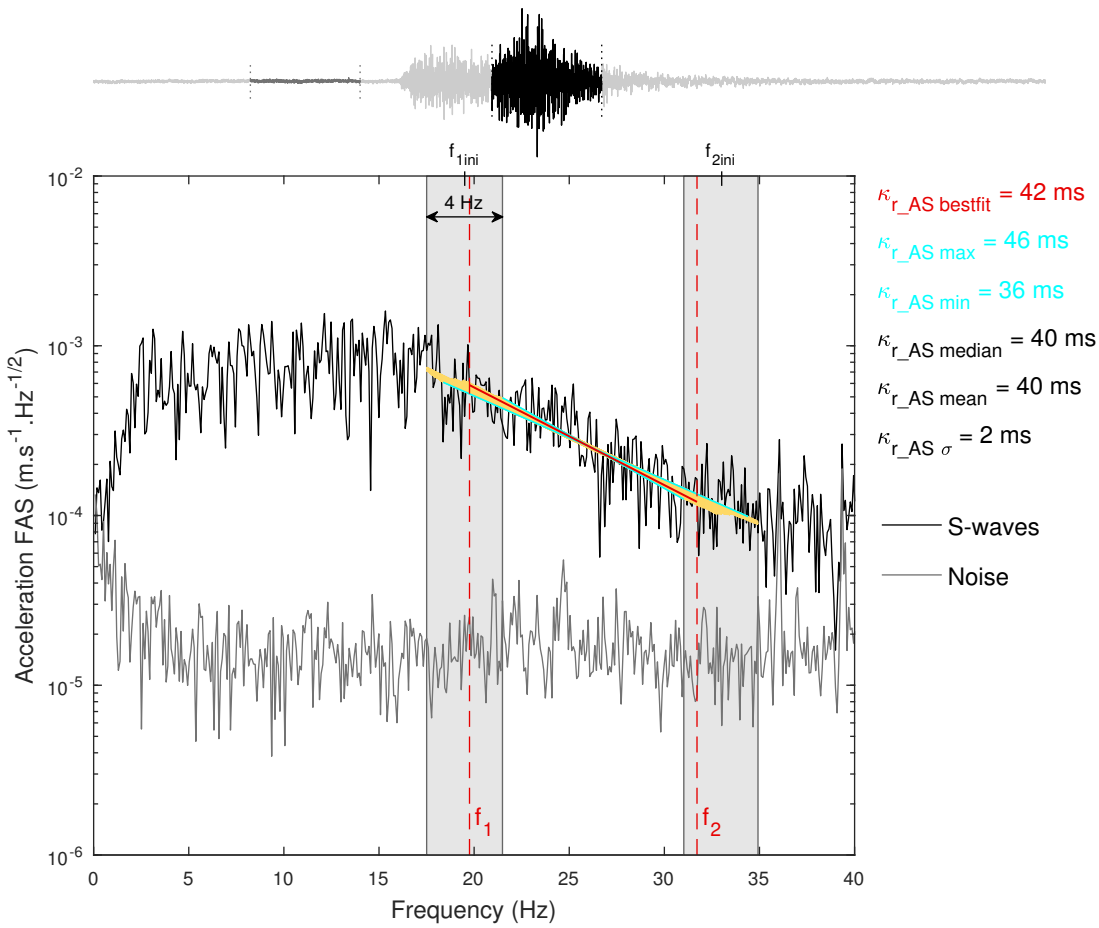


Figure 20. Example of κ_{r_AS} evaluation for site P2. Top: E-W accelerogram with the corresponding windows taken for the noise and S waves. Bottom: Horizontal mean component of the Fourier amplitude spectrum density (FASD) for the noise and for the S waves given with the procedure of evaluation of κ. The two initial frequencies picked by the operator (f_{1ini}, f_{2ini}) are used to define the two frequency windows (vertical gray bands) where the semi-automatic procedure is implemented. Between these two bounds, all of the combinations of the slope of linear regression are tested (yellow lines) to find the best in terms of the residuals of the regression ($\kappa_{r_AS\ best\ fit}$; red line) and the minimum and maximum slopes ($\kappa_{r_AS\ min}$, $\kappa_{r_AS\ max}$, blue lines).

Once every κ_r had been estimated with their associated uncertainties, κ_0 was computed following the chosen distance-dependence model ($\kappa_r(R_e)$). Here, the simple linear regression (Eq. (24)) was considered, with each κ_r value weighted by the inverse of its associated uncertainty. κ_0 can also be approximated by the individual κ_r measurements that correspond to short R_e distances, on the assumption that the path component is negligible when the earthquake occurs within a few tens of kilometers around the site (Ktenidou et al. 2013). Thus, in addition to the classical κ_0 "intercept value" evaluation, another estimate

$\kappa_{0<30\,km}$ was also computed as the mean of the κ_r values from events with R_e <30 km. This approach avoids a too large sensitivity to the slope of the distance dependence model, but it can lead to slightly higher κ_0 estimates.

2.5.2. Results and comparison between κ_{AS} and κ_{DS}

The different κ definitions imply differences in the range of the magnitudes and frequencies that are considered for its computation. Figure 21 shows the distribution of events used to determine κ_r from the acceleration (Figure 21, black filled circles) and displacement (Figure 21, gray circles) FAS according to the magnitude, depth, and back-azimuth. All of the events were crustal (depth, <20 km) and the back-azimuthal repartition shows preferential orientation close to N50°E for both of these approaches.

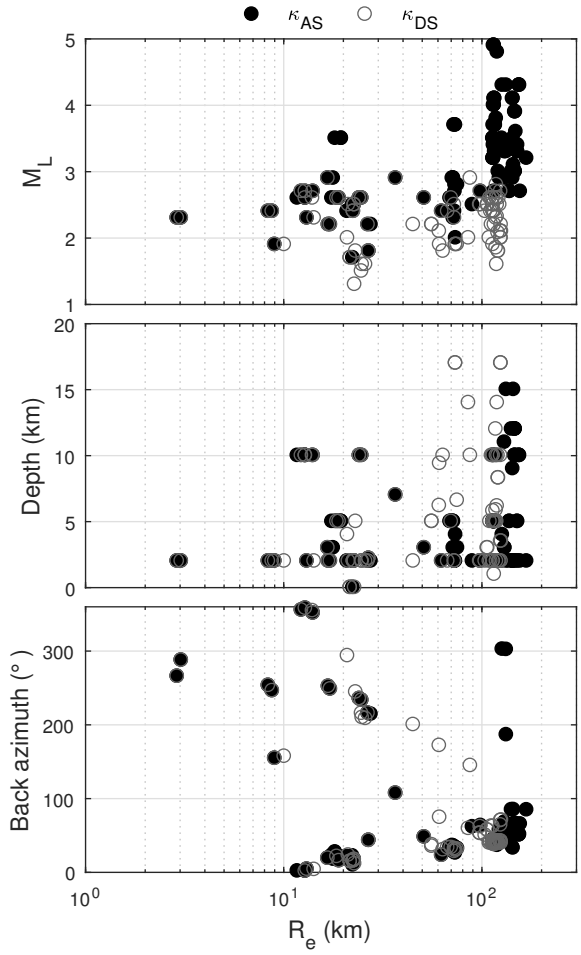


Figure 21. Comparisons of the dataset used for the κ_{AS} method (black filled circles) and the κ_{DS} method (gray circles), in terms of the magnitude (M_L), depth, and back-azimuth, according to the epicentral distance (R_e).

In comparison with previous studies (Kilb et al. 2012), events down to relatively low magnitudes for κ_{r_AS} (M <3) were used here, as well as relatively high magnitudes for κ_{r_DS}

(M >1.5). However, for κ_{r_AS}, most of the events have magnitudes >3, and those that are <3 are close enough to provide energy of 30 Hz or more, which allows for very high frequency κ estimations.

For magnitudes between 2 and 3, these acceleration and displacement approaches have both been realized. A comparison of the data obtained with these approaches for two earthquakes is given in Figure 22. These provide relatively similar value considering the strong uncertainty associated to each κ_r measurement. The frequency ranges for these approaches are not the same, as κ_{r_AS} involves higher frequencies than κ_{r_DS}, and the frequency window widths (Δf) are also slightly higher, in general, for κ_{r_AS}.

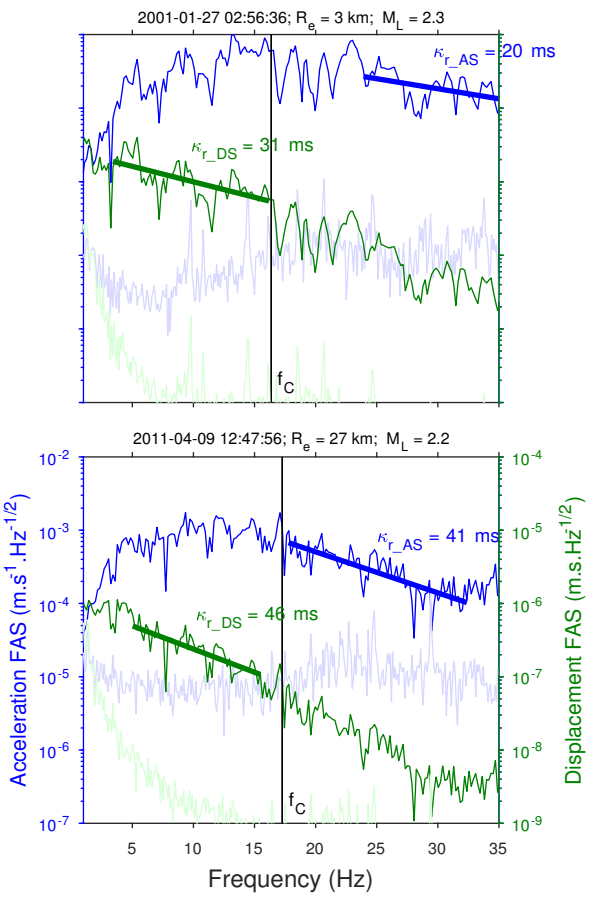


Figure 22. Comparisons of the individual κ_{r_AS} (blue) and κ_{r_DS} (green) estimations for two earthquakes at site P2. The vertical black line on each plot represents the picked corner frequency of the source (f_c).

Figure 23 shows the comparison between the κ_0 evaluations for these acceleration and displacement approaches, and provides a summary of the main features of the results. Here, the recordings at rock sites P1 and P2 are processed together to provide the maximum events to estimate the statistics for each approach, and to derive a regional m_κ from both sites (Anderson and Humphrey 1991; Ktenidou et al. 2013). The first expected result is that the number of usable events is somewhat lower with the acceleration method than with the

displacement method. Moreover, the event extraction methodology from the national catalogs that is followed here, imposes a lower limit on the exploitable magnitude range, which penalizes the displacement method. Indeed, local events of very small magnitude (M <1.5) that are not listed in the catalogs are not processed, even though they are particularly suitable for this approach. In addition, the acceleration method benefits from the 10 years of pre-existing triggered instrumentation for site P2, as only the higher magnitude events were recorded. As shown in Figure 22 and Figure 23, the measurement frequency range (characterize by the distribution of the central frequency $f_{mean} = {}^{1}/_{2}\,(f1 + f2)$), is definitely higher for κ_{r_AS} than for κ_{r_DS}, while the measurement bandwidth Δf are also a little higher for κ_{r_AS}. The event-to-event variability in the individual κ_r estimates, together with the associated uncertainties, are higher for the displacement approach, especially at large epicentral distances. This also leads to larger uncertainties in the estimation of κ_0 and m_κ for the displacement approach.

The discrepancy between κ_{r_AS} and κ_{r_DS} increase with increasing R_e (Figure 23) due to the much lower m_κ slope for the acceleration approach (m_{κ_DS} > 3 m_{κ_AS}). This large difference might explain why compared to κ_{0_AS}, κ_{0_DS} is lower, whereas $\kappa_{0_{DS<30\,km}}$ estimated from the nearest events is greater than $\kappa_{0_AS<30\,km}$. However, the dependence of κ_r on R_e is discussed later in terms of the Q-values obtained in previous studies in this region. Nevertheless, the κ_0 are very similar for both approaches, as close to 30 ms on average for the hard rock of the studied site. This is relatively high for sites with V_{S30} ≈ 2000 m/s, in comparison with those commonly proposed in the literature based on V_{S30}/κ_0 correlations (Ktenidou et al. 2014, 2015), although it still remains within the (large) uncertainties associated to such correlations. This is consistent with the κ_{0_AS} = 26 ms obtained by Douglas et al. (2010) for the Alps. However, the study of Douglas et al. (2010) is not fully comparable with the observations of the present study, as they used mean κ_{0_AS} from many rock sites under different site conditions, which are not likely to have been all as hard as the present site.

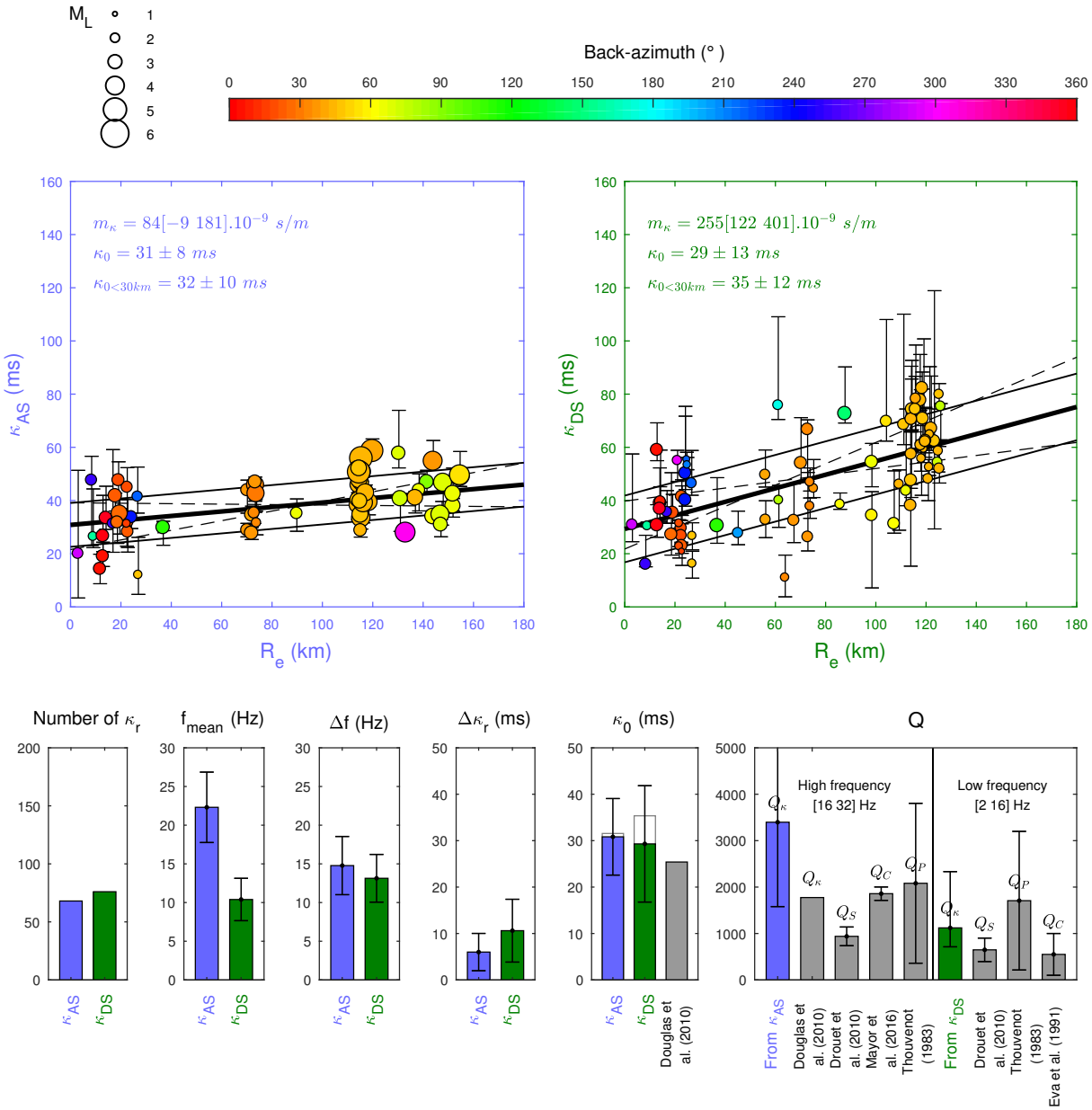


Figure 23. Comparisons between κ_{AS} (top left, purple) and κ_{DS} (top right, green) evaluations at the rock sites P1 and P2, taken together. Top: κ_r represented with the uncertainties ($\Delta\kappa_r$) as a function of the epicentral distance (R_e). The linear regression (thick solid line) gives the slope (m_κ) for the distance-dependence model, with the associated uncertainties (dashed lines) and the zero intercept (κ_0) with its uncertainties (thin solid lines). The $\kappa_{0<30\ km}$ approximation as the mean of the κ_r for R_e <30 km is given as well. Bottom: Statistics (generally as means $\pm$ one standard deviation) associated to each method: left to right, the number of individual estimations of κ_r, central values (f_{mean}) and widths (Δf) of the frequency ranges used to determine κ_r, $\Delta\kappa_r$, κ_0, and $\kappa_{0<30\ km}$ (gray) and finally the Q-values deduced from κ_r and compared with those available in the literature for the Alps region.

2.6. Analysis of the sensitivity and robustness of κ to various parameters

2.6.1. Measurement uncertainty $\Delta\kappa_r$

The data obtained from the acceleration and displacement approaches provide the opportunity to determine the sensitivity of the individual κ_r estimation uncertainties ($\Delta\kappa_r$) to various parameters where individual κ_r values are computed, such as the local magnitude (M_L), the epicentral distance (R_e), and the frequency window mean (f_{mean}) and width (Δf). As explained above (Figure 20), this uncertainty corresponds to the variability of the spectral regression slope over all the considered frequency intervals ($\Delta\kappa_r = \kappa_{r_{max}} - \kappa_{r_{min}}$).

Figure 24 shows $\Delta\kappa_r$ as a function of M_L, R_e, f_{mean}, and Δf for the acceleration and displacement approaches. While some trends can be seen between $\Delta\kappa_r$ and mainly M_L, f_{mean}, and Δf, the general trend differs greatly when considering $\Delta\kappa_{r_AS}$ and $\Delta\kappa_{r_DS}$. Moreover, a strong trade-off is suspected between M_L or f_{mean} and Δf. Indeed, when the magnitude is high, then f_c is low and the SNR is often good up to high frequencies, which provides a wide frequency range to measure κ_{r_AS} (high Δf). In contrast, low f_c values constrain the evaluation of κ_{r_DS} to the low frequency range, which restricts Δf. In the same way, an increase in f_{mean} for the acceleration approach indicates generally decreased Δf, while this is the opposite with κ_{DS}. These differences between the two approaches for the Δf trade-off with M_L and f_{mean} appear to explain the differences in the behaviors of these parameters with $\Delta\kappa_r$, whereby it is finally Δf that primarily controls of the uncertainty on κ_r. However, the apparent dependence of $\Delta\kappa_r$ on Δf is probably increased by the choice of a constant width (±2 Hz) for the investigated frequency band, which impacts more on a short window than a long one. Nevertheless, the minimum width of 10 Hz reduces this bias. After removing the parts due to the trade-off between Δf and the M_L and f_{mean} trends, the data (not shown) are convincing in terms that the dependence of the κ_r uncertainty on M_L, R_e, and f_{mean} is negligible, as this can be almost totally explained by Δf. $\Delta\kappa_r$ actually exhibits exponential decay with increasing Δf. This sensitivity to Δf is likely to be associated to several physical factors (site amplification, frequency-dependent attenuation, and source effects on κ_r). These are discussed in the next sections.

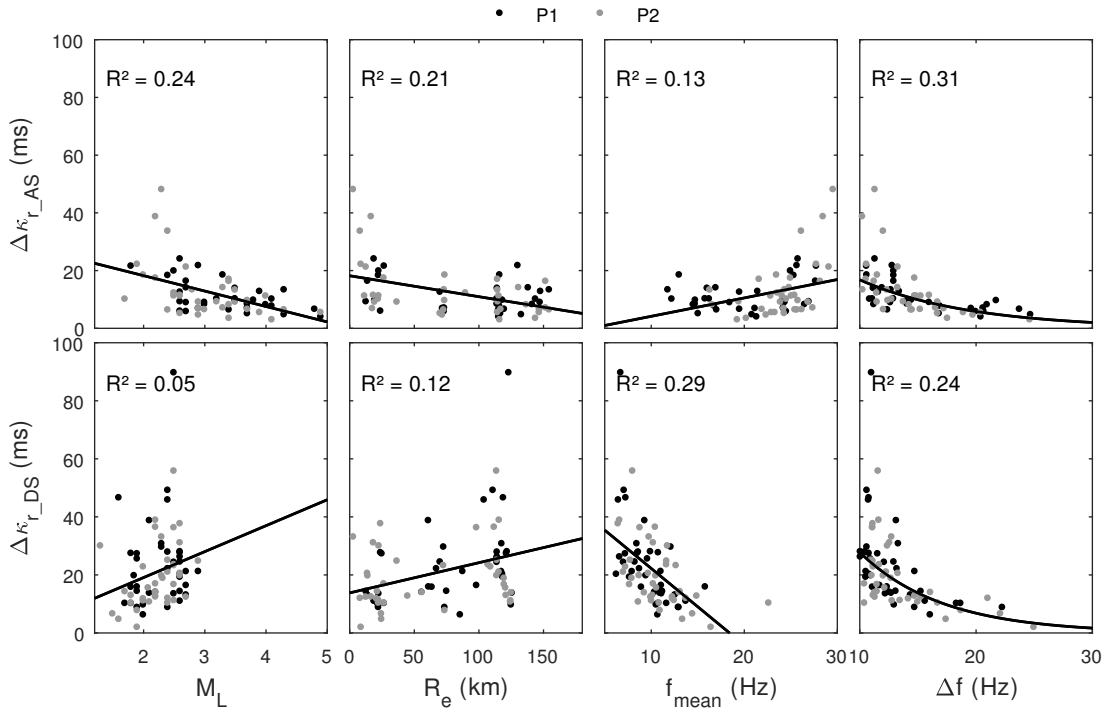


Figure 24. Evaluation of the dependence of the κ_r uncertainty ($\Delta\kappa_r$) on local magnitude (M_L), epicentral distance (R_e), frequency window mean (f_{mean}), and width (Δf) used to assess κ_r, shown for the κ_{AS} (top) and κ_{DS} (bottom) approaches. For each plot, the linear trend is represented with its corresponding determination coefficient (R^2). For Δf, an exponential model is preferred to the linear trend. Black and gray circles represent the results for sites P1 and P2, respectively.

2.6.2. Frequency-dependence of the attenuation

As indicated in the Introduction, κ is assumed to be related to the frequency-independent component of Q, thus ignoring the scattering component of the attenuation. When considering propagation in the crust only, Equation (26) can be simplified into (Hough et al. 1988; Ktenidou et al. 2015):

$$Q_i = \frac{1}{V_S\, m_\kappa} \qquad 29$$

where Q_i describes the intrinsic crustal attenuation only, and V_S is the mean shear-wave velocity in the crust. To avoid making any assumption in Equation (29), this Q estimate from the m_κ values is referred to as Q_κ. Figure 23 includes comparisons between Q_κ and Q from previous studies in the Alps. For this, a shear-wave velocity of V_S = 3500 m/s was assumed, which is a standard value for the crust. The m_{κ_AS} value of Douglas et al. (2010) is also translated into the Q_κ value, and compared with the other values of Q at high frequencies. Mayor et al. (2016) estimated a value of Q_c from the coda of between 16 Hz and 32 Hz, while

Eva et al. (1991) proposed a Q_c value between 2 Hz and 16 Hz. Thouvenot (1983) proposed a $Q_P(f)$ model from P waves recorded during a few active, deep-sounding experiments, and in a different way, Drouet et al. (2010) also established a $Q_S(f)$ model from a generalized inversion technique scheme on the S wave phase of earthquake recordings. Here, values at high (16-32 Hz) and low (2-16 Hz) frequencies are calculated from these two $Q(f)$ models. The high-frequency Q deduced from these models are compared to the m_{κ_AS} evaluations, while the low-frequency Q are compared to those deduced from the present m_{κ_DS} value.

The values from previous studies show large scatter, which is not surprising, as they were evaluated from different techniques in different phases of the signal (i.e., P waves, S waves, coda waves) and for different locations in the Alps. Moreover, Q_c is primarily controlled by the absorption (Q_i) (Aki and Chouet 1975), while Q_P and Q_S provide access to the full attenuation Q_{ef} that also includes the scattering (Campbell 2009). In Figure 23, the Q_P estimates present higher values, especially at low frequencies, while Q_S and Q_c estimates are comparable. The evaluation of Q_κ at lower frequencies from m_{κ_DS} is in good agreement with previous studies, as it is within most of the variability ranges. At high frequencies, the very small m_{κ_AS} values lead to Q-values much larger than those reported in the literature. This inconsistency with previous studies can be explained again by the large differences between each approach. Even considering Douglas et al. (2010), who followed the same κ_{AS} procedure, this is not fully comparable, as they used many stations from many locations in the Alps to determine m_{κ_AS}.

Limited impact on the uncertainty is expected for the source and path components, as all of the events were crustal low-to-moderate magnitude earthquakes that are mainly from the same narrow azimuthal range (Figure 21). Nevertheless, a strong three-fold discrepancy appears between m_{κ_AS} and m_{κ_DS} (Figure 23). A possible explanation for this is that m_{κ_AS} refers to a higher frequency range than m_{κ_DS}. This difference in the frequency range in which m_κ is measured might explain the discrepancy between the slopes obtained from each of these methods, as Q is widely accepted to increase as the frequency increases. However, the difference between the previous estimates of high and low frequency Q is significantly lower than the difference inferred here from the m_κ data. Thus, the frequency dependence of Q might partly explain the discrepancy observed between m_{κ_AS} and m_{κ_DS}, but probably not all.

2.6.3. Site amplification dependence of κ

According to most studies, κ_0 is linked to the S-wave attenuation due to the geological structures beneath the site. In GMPEs or host-to-target adjustment techniques, κ_0 reflects only the attenuation, while the amplification is generally taken into account mainly through V_{S30}. However, attenuation and amplification are impacting the same frequency range and are difficult to separate in practice. κ_0 measurements without due consideration to site amplification may thus be significantly biased in an unpredictable way: Sedimentary basins generally exhibit large amplifications that are strongly frequency dependent over a wide

frequency range (Figure 19), while the presence of a weathered zone on rock sites can also produce high frequency amplification. These site effects modify the FAS shapes and can thus bias the κ_r evaluation. The site amplification is expected to modify κ_0 mainly and m_κ only slightly, as every κ_r are biased approximately in the same manner, as long as they are evaluated for a similar frequency range. Various studies have made the assumption that reliable evaluation of κ_r is possible, as long as the analysis frequency windows are chosen out of the fundamental resonance frequency range of the site (f_0) and in a sufficiently wide frequency range (Hough et al. 1999; Parolai and Bindi 2004; Ktenidou et al. 2013). However, this assumption is doubtful when the site amplification is complex (2D, 3D) and/or broadband, and it is difficult to respect this in a low-to-moderate seismicity context, in which the spectral windows used to evaluate κ_{r_AS} are generally narrower compared to those available in higher seismicity areas. For instance, sites P3 and P4 show a broadband amplifications (Figure 19) that might have different impacts on κ_r due to the difference in the spectral shape, even if these sites are located near to each other in the same valley. Another approach consists of evaluation of κ_r from recordings that have been initially deconvolved (corrected) from the site transfer function. Recently, taking into account only the amplification, deconvolution by the theoretical 1D transfer function was tested, but did not provide convincing results (Van Houtte et al. 2011; Ktenidou et al. 2013). The main difficulties of such an approach are the availability of a well-known velocity profile for theoretical computation, the validity of the 1D approximation, and the potential introduction of some uncertainty associates to the transfer function on κ. In contrast, when using empirical approaches (e.g., SSR, generalized inversion technique), the difficulty is to separate the amplification and the attenuation.

Figure 25 is designed to show the correction function that is needed to correct the FAS before computing κ_r values, to account for site effects. For sites P2 and P3, these correction functions are given using the inverse of the relative site transfer functions estimated from the SSR approach at the sites, with site P1 taken as reference (Figure 25, black curves). For site P1, the correction function is the inverse of the theoretical transfer function that is computed from the 1D reflectivity model (Kennett 1974) (Figure 25, gray curve), based on the *in-situ* velocity profile available at this site. To understand how the site response influences the κ_r evaluation, the linear trends of the correction function are shown in Figure 25. The slopes of these trends quantify the 'corrections' that will modify the κ_r evaluations, which are denoted as κ_{corr_AS} and κ_{corr_DS}. The slopes are computed on the site correction functions for the frequency windows defined by the mean of Δf and f_{mean} used for the κ_{r_AS} and κ_{r_DS} determinations (Figure 25, blue, green solid lines, respectively). The mean of $f_{mean} \pm$ standard deviation (σ) are also given, to infer the frequency dependence induced on κ_r by the site effects (Figure 25, blue, green dashed lines).

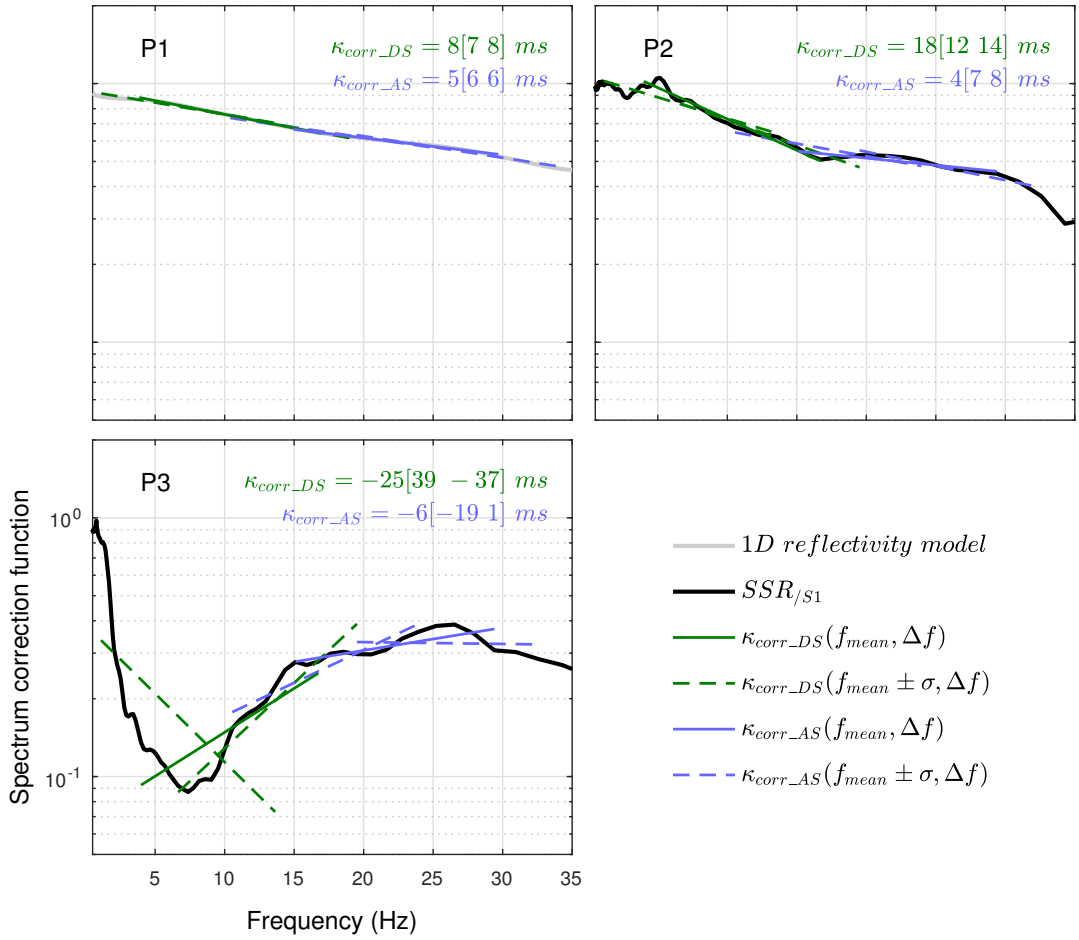


Figure 25. Perturbation of κ induced by site amplification for sites P1 to P3. The spectral correction functions are estimated as the inverse of the site transfer functions. At sites P2 and P3, the empirical site transfer functions were computed from the standard spectral ratio (SSR; black) according to the reference site P1. At site P1, the theoretical site transfer function was computed through the 1D reflectivity model approach, based on the *in-situ* velocity profile. The linear trend was computed from the transfer function in the frequency range defined by $(\mathbf{f_{mean}}, \Delta\mathbf{f})$ and $(\mathbf{f_{mean}} \pm \boldsymbol{\sigma}, \Delta\mathbf{f})$ for $\boldsymbol{\kappa_{AS}}$ (purple) and $\boldsymbol{\kappa_{DS}}$ (green), where $\Delta\mathbf{f}$ is the mean width, $\mathbf{f_{mean}}$ is the mean, and $\boldsymbol{\sigma}$ is the standard deviation of the central frequency of the frequency windows used to determine $\boldsymbol{\kappa_r}$. The $\boldsymbol{\kappa}$ values deduced from these linear trends are also indicated.

From the correction inferred from the theoretical site amplification for site P1, it can be seen that due to the shallow weathered zone, the velocity gradient within the first meters in depth also induces amplification that can bias the measures of κ (+5 to +8 ms), even if this is a hard-rock site. Site P2 is also a hard-rock site, and it is very similar to site P1, and thus small κ differences are expected between these two sites. The SSR transfer function between sites P2 and P1 shows two main linear trends with different bias for κ_{r_AS} (+4 to +8 ms) and κ_{r_DS} (+12 to +18 ms), but do not increase the frequency variability by much. For the soil site P3, the strong relative amplification below 15 Hz induces very variable and important modifications to κ_{r_DS} that depend on the frequency window (+39 to -37 ms). At higher

frequencies, the transfer function is flatter, which leads to κ_{r_AS} evaluations that are less dependent on the site responses.

The real influence of the site amplification on κ is shown in Figure 26, through a comparison between sites P1, P2, and P3. To understand how the site effects interact with κ, κ_r evaluations that are made from the FAS deconvolved by the theoretical (for site P1) and empirical (for sites P2, P3) site transfer functions are also shown in Figure 26. It should be noted here that the whole processing procedure, which included the frequency bound (f_{1ini}, f_{2ini}) picking, was performed after the deconvolution. This is important, because the site amplification can also change the apparent linearity of the FAS and lead operators to select mistaken frequency bounds for κ_r evaluation. In Figure 26, the κ_{AS} results are given on the left, with the κ_{DS} results on the right. For each panel in Figure 26, the individual κ_r measurements are represented according to the epicentral distance, where the diameters of the symbols are proportional to the magnitudes, and their color indicates the back-azimuth of the corresponding event.

First, for the κ_{AS} method, the results at the two rock sites (sites P1, P2) show similarities in terms of both κ_{0_AS} and m_{κ_AS}, as the discrepancy is within the variability of the measurements. This is consistent with the spatial and geological proximity of these two sites. However, we observed some significant differences for a few of the individual κ_{r_AS} evaluations between these two sites. Indeed, modification of the FAS by the site effect can lead to higher frequency evaluations for site P2 than site P1, due to the slope change in the SSR at around 17 Hz (Figure 26). The displacement method shows greater discrepancy between these two sites. The slope m_{κ_DS} for site P1 is almost twice that of site P2, and the site component κ_{0_DS} is a little lower for site P1 than site P2, although this difference can be easily explained by the differences in the slope. Indeed, the slope-independent $\kappa_{0_{DS}<30\,km}$ is similar for sites P1 and P2. Moreover, as expected, $\kappa_{0_DS<30\,km}$ is higher at site P1 than site P2, contrary to κ_{0_DS}. Differences in the slopes between sites P1 and P2 cannot be explained physically, as the regional attenuation must be the same for all of the sites and is expected to be proportional to the regional Q_i (Ktenidou et al. 2015). This can be attributed to the large scatter on the individual κ_{r_DS}, to the lack of measured points at short epicentral distances for site P1 to constrain the slope, and to the differences in the input dataset. At stiff-soil site P3, κ_{0_AS} is very similar to the values obtained for the rock sites. At first glance, a higher value of κ_{0_AS} might be expected for this stiff-soil of site P3, as such sites are classically more attenuating than rock sites. However, the influence on κ_{0_AS} of the shallow stiff-soil basin might be limited, as this parameter is assumed to infer attenuation down to deep geological structures (Ktenidou et al. 2015).

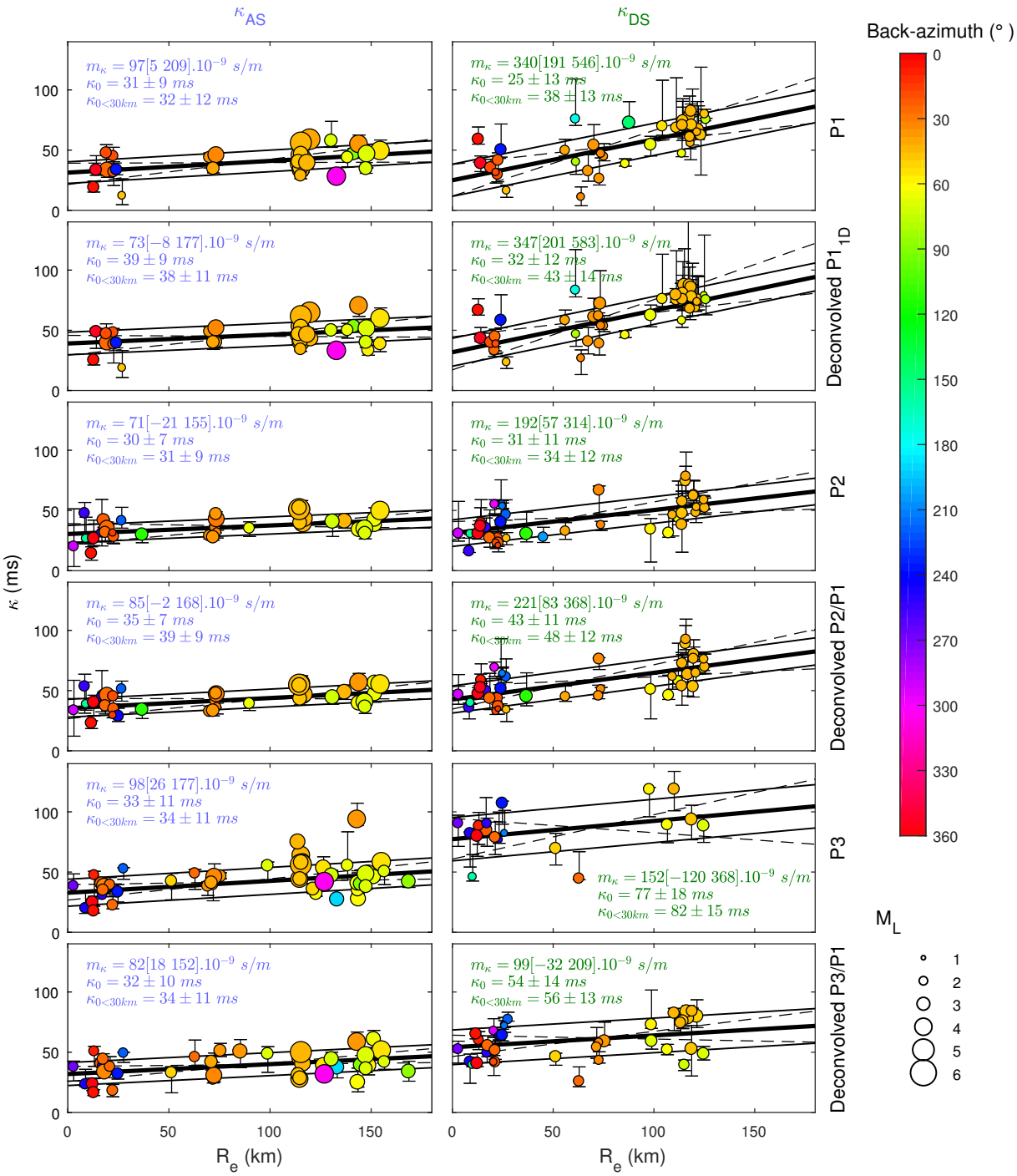


Figure 26. Estimation of κ_r, κ_0, $\kappa_{0<30km}$, and m_κ (slope) from the κ_{AS} (left) and κ_{DS} (right) approaches for the two hard-rock sites (P1, P2) and the stiff-soil site (P3). At sites P2 and P3, the results are obtained after deconvolution of the recordings by the relative transfer functions estimated by SSR, with site P1 as reference. For site P1, the deconvolution is realized from the theoretical 1D reflectivity model site amplification function. κ_0 and m_κ are estimated from linear regression, where each κ_r is weighted by the inverse of its variability ($\Delta\kappa_r$).

The influence of site transfer function deconvolution on κ_0 might be roughly predicted by the κ_{corr_AS} and κ_{corr_DS} obtained from the slopes of the site correction function (Figure 25). Deconvolution of the hard-rock sites P1 and P2 provides κ_{0_AS} and κ_{0_DS} results that are close to those predicted by the site correction function given in Figure 25. As predicted, the deconvolved κ_{0_AS} are only slightly changed, as they are within the variability band of the raw estimation. For site P3, the κ_{0_AS} modification (-1 ms) does not agree well with the prediction (-6 ms). Moreover, the scatter in κ_{r_AS} and κ_{r_DS} is slightly reduced, and the number of individual κ_{r_DS} available increases after the deconvolution. These observations suggest that the site effects disturb the linearity of the spectrum decay for site P3, which led the operator to remove some events or to improperly select the initial frequency bound for other events. At sites P2 and P3, the deconvolution was computed from the SSR transfer function relative to site P1. This correction should provide κ results at the site that are very close to those obtained at the reference site, as the site is thus placed according to both the amplification and attenuation conditions of the reference. Then, strong similarities are expected between the deconvolved sites P2 and P3, and site P1. This convergence toward site P1 is not realized in Figure 26, especially for κ_{DS}. These observations probably do not agree because of differences in the datasets used for the different sites, and of the introduction of uncertainties by the SSR deconvolution. We do not expect such differences to be due to differences in the deep structure, as suggested by Ktenidou et al. (2015), as the deep structure (i.e., beyond, at most, 100 m in depth) should be the same for all sites. The impact of site and of crustal amplification on κ_0 is estimated through deconvolution of the 1D theoretical transfer function at site P1, similar to what was done by Van Houtte et al. (2011) and Ktenidou et al. (2013). Figure 26 shows that κ_{0_AS} and κ_{0_DS} increased notably after the deconvolution, mainly due to the site amplification rather than to the crustal amplification, as the transfer function is widely control by the former. Thus, the site amplification cannot be neglected, even for hard-rock sites.

To conclude this section, it can be seen that the site-effect influence can be high and variable, depending on the frequency band. Moreover, in agreement with recent studies (Van Houtte et al. 2014; Edwards et al. 2015; Laurendeau et al. 2017), the site amplification can explain a part of the observed κ variability, as it is frequency dependent and each individual κ_r is measured for different frequency windows. Thus, κ should be considered carefully, as a site amplification component cannot be excluded even for rock sites, especially for low-seismicity context when κ_r are estimated from limited spectral windows.

2.6.4. Source dependence of κ

The assumption of a negligible source contribution for κ relies on the validity of the ω^{-2} source model (Brune 1970). Any variation from this model or any bad consideration of the f_C criteria can impact upon the measurement of κ. To evaluate the validity of this assumption, the two recorded seismic sequences of Jausiers are considered. The Jausiers cluster of points is shown in Figure 21 and Figure 23 close to the 120-km epicentral distance and at

approximately the N50°E azimuth. All of the events are co-located, so the records share at least the same site and path components.

Figure 27 shows the linear trend for the clusters between the κ_{r_AS} individual values estimated at sites P1 (black filled circles) and P2 (gray filled circles), and the local magnitudes. The associated coefficient of determination (R^2) is also shown. Although there are not enough data points to form any conclusions here, in Figure 23 and Figure 27, higher magnitudes appear to correspond to higher κ_{r_AS}. For κ_{r_DS}, this trend cannot be seen in Figure 23 and is not represented in Figure 27, due to the too narrower range of the magnitudes that is available with the displacement method. An initial possible explanation is that κ is dependent on the magnitude, in agreement with some previous studies that have argued for its source dependence (Papageorgiou and Aki 1983; Aki 1987; Papageorgiou 1988, 2003; Gariel and Campillo 1989; Wen and Chen 2012). A second explanation is that the decrease in κ_{r_AS} is due to the shortening of Δf for decreasing magnitudes (Figure 27), which makes its measurement less robust, as observed through the increase in $\Delta\kappa_{r_AS}$ for decreasing Δf (Figure 24). Indeed, the κ_{r_AS} estimate at low magnitudes can be more sensitive to the bad consideration of the f_c criterion. Indeed, because of the source spectrum shape in acceleration that increases up to f_C and is then flat, if f_1 is taken below f_C, this would result in an underestimation of κ_{r_AS} (Boore and Campbell 2017; Ktenidou et al. 2017). Moreover, it should be noted that the source model is not bilinear, but has a smooth transition around f_C that is described by the gamma parameter. This means that f_1 should be taken a few Hertz above f_C to avoid any influence of the sloped part of the source spectrum. However, in practice, the limited bandwidth that is available to measure κ_{r_AS} provides such a precaution, especially for low magnitude events. In addition, f_c is difficult to determine due to site effects that modify the spectrum, and then potentially hide the correct value. For the displacement approach, if f_2 exceeds f_C, this should result in an increase in κ_{r_DS}. The stress drop of small magnitude events is very uncertain, which makes this latter approach very sensitive to f_C. This phenomenological difference between the two approaches might explain, at least in part, why κ_{r_DS} generally exceeds κ_{r_AS} (Ktenidou et al. 2017).

However, the possible influence of the frequency windows for κ_{r_AS} is not clear in Figure 27. It appears that lower f_{mean} and narrower Δf correspond to lower κ_{r_AS}, although the correlations are not very good, especially for Δf. The linear regression shows low correlation between f_{mean} and κ_{r_AS} (R^2 = 0.27), while no correlation is seen for Δf (R^2 = 0.04). The investigation of the trade-off between f_{mean} and Δf with the magnitude is given on the right in Figure 27. The correlation between the magnitude and f_{mean} is not clear (R^2 = 0.18), while that with Δf is evident (R^2 = 0.61). This latter parameter is not correlated with κ_{r_AS}, so the possible bias of the lower magnitudes due to the associated frequency windows appears not to explain very well the apparent magnitude dependence of κ_{r_AS}. Only a limited part of the influence of the trade-off between the magnitude and the frequency range where κ_{r_AS} is evaluated can be explained, and this is highly uncertain.

No influence of the depth was found, as this parameter was only slightly variable between the events, and because this information was extracted from the national bulletin, and was thus affected by relatively large uncertainty. Moreover, this result is in agreement with Edwards et al. (2011), who showed that the linear trend between attenuation and distance indicates limited depth dependence for κ_r.

To conclude this section, among the different explored source parameters, the best correlation with κ_r appears to be for the magnitude, and this appears to be explained by the source dependence, rather than by bias on the lower magnitudes due to the overlap of the frequency window with f_C. However, the correlation remains rough, and the uncertainties on κ_{r_AS} and the local magnitude estimations are too large to be conclusive on this point.

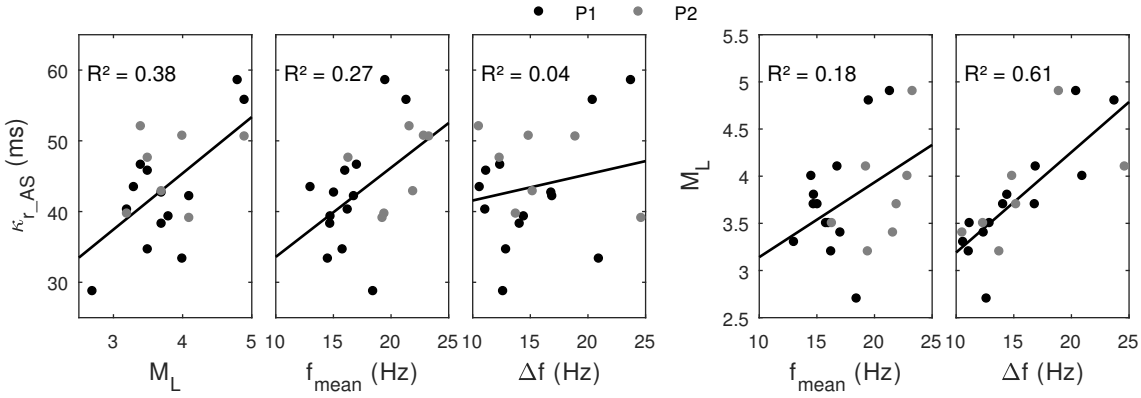


Figure 27. Evaluation of the source dependence of κ from the correlation between κ_{r_AS} and the local magnitude (M_L) for earthquakes approximately in the same position (a cluster of events located at approximately 120 km epicentral distance, and azimuthal direction N50°E). The potential trade-off between M_L and the frequency windows chosen to measure κ_{r_AS} is illustrated by the correlation between M_L and the central frequency (f_{mean}) and width (Δf) of the frequency window (left). The influence of the frequency window on κ is illustrated by the correlation of κ_{r_AS} with f_{mean} and Δf (right). Black and gray circles define κ_{r_AS} from sites P1 and P2, respectively; the linear trend is represented by the corresponding determination coefficient (R^2).

2.7. Discussion

The site component of κ, κ_0, is widely used in hazard seismology to constrain the high-frequency spectral shape in stochastic simulations, in some GMPEs, and for host-to-target adjustment. The underlying interpretation is that κ_0 represents the S-wave attenuation by the geological structure beneath the site. However, since the physics of κ are not fully captured, it is important to discuss to which extent the κ_0 estimates may be biased by the limitations of data available in low seismicity areas, and the associated origins of the large variability observed in κ_r measurements. These two issues are discussed separately, even though some physical phenomena may affect simultaneously the bias and scatter of κ_r measurements, before a final discussion on the validity of the κ_{DS} approach, as it seems well suited to low-to-moderate seismic areas.

2.7.1. Reliability of the κ_0 measurements

In GMPEs or host-to-target adjustment, κ_0 only reflects the frequency independent attenuation. However, the underlying physics are still debated. As mentioned before, the site amplification, the frequency dependence of attenuation, and the earthquake source might bias κ_0 estimates by systematically moving up or down individual κ_r measurements.

The effects of the site amplification have been reported in several instrumental and simulation studies. Ktenidou and Abrahamson (2016) observed negative apparent κ_0 on many hard-rock sites, which they attributed to biasing effects of site amplification. It has been often considered that κ_r may be reliably estimated when the frequency interval over which the spectral decay is measured is broad enough and does not include the fundamental resonance frequency of the site (f_0) (Hough et al. 1999; Parolai and Bindi 2004; Ktenidou et al. 2013). This assumption was supported by simple 1D simulations (Parolai and Bind, 2004) and is easy to implement in practice as f_0 can be easily determined through the horizontal-to-vertical spectral ratio approach computed either from microtremors (e.g., Nakamura 1989; Kudo 1995) or from earthquake recordings (e.g., Lermo and Chávez-García 1993). As mentioned earlier, there are many cases however where complex, broadband site effects hamper κ_0 measurements. We show that even for hard rock sites, high frequency amplification systematically biases κ_r in a similar manner for a given approach (Figure 25), resulting in a significant impact on the accuracy of κ_0 (by about 25-30%), with only weak changes for the regression slope with epicentral distance m_κ (Figure 26). For soil sites, the site amplification influence is large, and almost impossible to correct, when f_0 is included in the analysis frequency windows for both κ_0 and m_κ (κ_{DS} at site P3, in Figure 25 and Figure 26), and it is prejudicial otherwise (κ_{AS} at site P3, in Figure 25 and Figure 26).

In a similar way, the frequency dependence of the attenuation may impact the value of κ_0 at least through the measurement frequency interval. Boore and Campbell (2017) provided an illustrative example of the large variability of κ_0 obtained for the same site (Pinyon Flat Observatory, California) with different approaches. The frequency-independence assumption was formulated during the introduction of κ by Anderson and Hough (1984) and later by Hough and Anderson (1988), with reference to several previous studies (Warren 1972; Rovelli 1982; Anderson 1986). Anderson et al. (1996) also reported a negligible influence of Q_{sc} from numerical simulation for velocity and Q profiles with pluri-hectometer thick layers. However, the frequency dependence of Q_{sc} and even sometime of Q_i was actually shown by various studies (e.g., Aki and Chouet 1975; Calvet et al. 2013; Mayor et al. 2016). When introducing their model (Eq. (26)), Hough and Anderson (1988) already indicated that even a slight frequency dependence of Q_{ef} will yield a smaller value of κ. Edwards et al. (2015) also recently supported the frequency-dependence interpretation through a comparison of κ results from different approaches involving different frequency bands. They also observed that the high-frequency spectral decay was generally not well explained by the linear κ_r attenuation model (Eq. (22)), but rather by a curved or bi-linear model. Parolai et al. (2015) showed a nonnegligible role of scattering attenuation from numerical simulations, especially for small levels of intrinsic attenuation. They proposed a nonlinear model for the high

frequency decay due to the introduction of scattering when the FAS are determined from several-second-width time windows in the S waves. Ktenidou et al. (2015) attributed the discrepancy between borehole and surface κ_0 measurements to the scattering, which was recently confirmed by Pilz and Fäh (2017) showing that the scattering contribution to κ_0 should not be neglected. In the present study, we observe a strong three-fold discrepancy between m_{κ_AS} and m_{κ_DS} (Figure 23) that might be explained at least partly by the difference in the frequency range between each definition of κ. The impact of the frequency dependence is more difficult to assess on κ_0, but as the attenuation decreases with increasing frequency, it might reduce κ_0, as predicted by Hough and Anderson (1988). Moreover, when approximating the distance-dependence model through the linear assumption, any variation in m_κ will result in a variation in κ_0. It is thus essential to compare the κ_0 values with the average of κ_r values for the closest events (e.g., $\kappa_{0<30\ km}$).

Finally, various studies have argued for source dependence of κ since it was first defined (Papageorgiou and Aki 1983; Aki 1987; Papageorgiou 1988, 2003; Gariel and Campillo 1989; Wen and Chen 2012). While the site interpretation is at present commonly accepted for the distance independent part of κ, source-induced biases, due to deviations from the ω^{-2} model or misapplication of the f_C criteria are possible. Seismic clusters are particularly suitable to study the source dependence, as κ_r measurements of these events only differ by their source component. In the present study, κ_{r_AS} values from the Jausiers cluster show a trend for decreasing with decreasing magnitudes. No conclusive interpretation is however possible for the role of the source on κ, due to the scarcity of the data.

To sum up, the accuracy of κ_0 in low-to-moderate seismicity areas appears to be primarily controlled by the site amplification, especially for soil sites, and secondly by the approximation made with the frequency and source independence assumption. One must note however that the last two effects on κ_0 are difficult to quantify, and might be stronger than expected. Nevertheless, in this study we found a robust κ_0 estimate of 30 ms with both κ_{AS} and κ_{DS} for the hard-rock sites that is consistent with the high attenuation indicated for the Provence region (Mayor et al. 2016) and with the κ_0 obtained by Douglas et al. (2010) for rock sites in the Alps. Moreover, even if this high κ_0 is in the upper part of the very scattered κ_0/V_{S30} correlation, this is in agreement with recent studies that indicate higher κ_0 values for hard-rock sites than was initially suggested (Ktenidou et al. 2015; Ktenidou and Abrahamson 2016; Boore and Campbell 2017; Laurendeau et al. 2017).

2.7.2. Variability of κ_r measurements

Most studies reporting κ measurements indicate strong scatter when the κ_r values are represented according to the epicentral distance and sometimes an important variability in individual evaluation of κ_r themselves. Both the frequency dependence of the attenuation itself and that introduced by the site amplification can increase the frequency dependence of κ_r. This can explain a part of the variability between κ_r values, as they are evaluated over a variable frequency window. Edwards et al. (2015) showed that site amplification can have a strong influence on κ_{r_AS}, which depends on the frequency window considered, even for a

hard-rock site. Van Houtte et al. (2014) observed an important variability of κ_{r_AS} with the component orientation, and they attributed this to site effects. In the present study, we observe that when f_0 is included in the analysis frequency range, the site amplification greatly increases the frequency dependence of κ_r (κ_{DS} at site P3 Figure 25), and only slightly increases the scatter between the κ_r values (κ_{r_DS} at site P3 in Figure 26). In the same way, the implicit approximation of a lateral homogeneity for the regional Q_{ef} included in the distance-dependence model (Eq. (24)) is certainly not exact. However, no obvious influence of the back-azimuth of the source is observed on κ_r here (Figure 23 and Figure 26), although it is not easy to separate it from the distance dependence. In the present study, we also evaluate the individual uncertainty of each κ_r ($\Delta\kappa_r$) through the variability of the spectral decay slope over varying frequency intervals. $\Delta\kappa_r$ is found to be primarily controlled by the width of this frequency interval (Δf). Small-scale variations in the FAS are thus very likely to perturb κ_r measurements for short Δf.

The effects of the source can be important, as highlighted through the high κ_{r_AS} scatter that was sometimes observed between events that belonged to the same cluster (Kilb et al. 2012; Ktenidou et al. 2013). Kilb et al. (2012) attributed this to the variability of the near-source properties and the f_c values. In the present case, the former interpretation cannot be supported, as the magnitudes are small and the clusters are far enough apart to avoid near-field effects. For a subset of the Jausiers cluster events used in Figure 27, Figure 28 shows the influence of using variable (Figure 28, top) or constant (Figure 28, bottom) frequency windows over which κ_{r_AS} are evaluated. Using a constant frequency window for every event of this cluster, where the back-azimuth varies by less than 8° and the epicentral distance by less than 5%, should greatly reduce the scatter in the κ_{r_AS} values, as almost no difference is expected between the κ_{r_AS} values for the site, the path, and the frequency dependence. Surprisingly, even if no clear correlation can be found between κ_{r_AS} and the magnitude, the strong scatter on the κ_{r_AS} values (30-60 ms) observed indifferently with variable or constant frequency windows, appears to be an unambiguous link to the source for this hard-rock site. The constant frequency window that can be used for every κ_{r_AS} is relatively narrow (Δf = 10 Hz), which led to an increase in $\Delta\kappa_{r_AS}$ compared to that obtained with wider and more variable windows. However, the influence of Δf might not be preponderant here, and cannot explain the strong scatter observed between the κ_{r_AS} values. Moreover, this is more likely to be due to variable deviations from the ω^{-2} model than to be linked to incorrect consideration of the f_C criteria. This means that the dominance of the source on the variability between κ_{r_AS} values is probably not specific to this study, or to low seismicity areas.

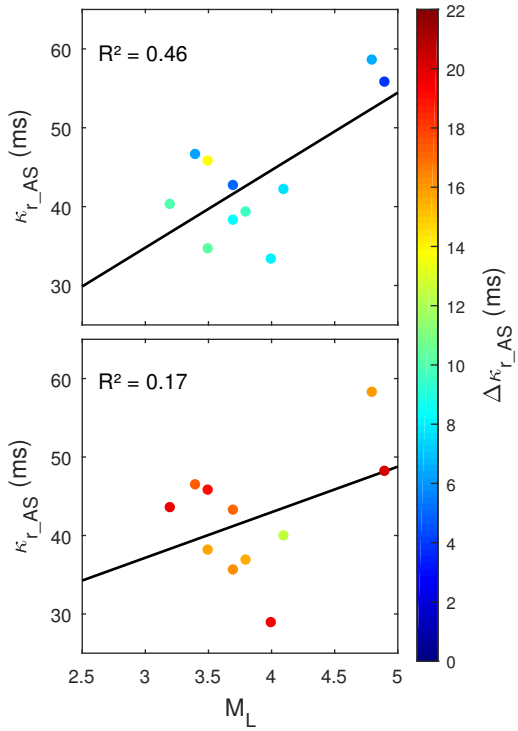


Figure 28. Evaluation at site P1 of the source dependence of κ_{r_AS} for a subset of the events from the Jausier cluster used in Figure 12. Top: Each κ_{r_AS} value is estimated on the wider frequency window available. Bottom: Same events, but with the κ_{r_AS} values calculated over the constant frequency window of 11.3 Hz to 21.3 Hz. The color scale shows the variability associate to each κ_{r_AS} evaluation ($\Delta\kappa_{r_AS}$).

2.7.3. The κ_{DS} approach

As discussed above, the evaluation of κ is variable and sometimes unreliable. This is particularly true when various approaches are used to measure κ, and in low-seismicity areas where κ_r are evaluated over narrower and more variable frequency windows. Nevertheless, in such a context, all possible approaches have to be tested to improve the current practice, which consists of the deduction of κ_0 from the very uncertain κ_0/V_{S30} correlation (Kottke 2017). The Biasi and Smith (2001) approach is very promising, as it is adapted to low magnitude events that are generally the only events that can be recorded in low-seismicity areas over a reasonably short period of time. Moreover, the flatness of the displacement source spectrum below f_C is better understood than the ω^{-2} fall-off above f_C. This will lead to a potentially stronger influence of the source for κ_{AS} than for κ_{DS}. However, this presumed stronger robustness of κ_{DS} with respect to source spectral shape has not been observed in our results : κ_{r_DS} and κ_{r_AS} exhibit a comparable scatter (Figure 23 and Figure 26). Actually, κ_{DS} is likely to be more sensitive to the site amplification than κ_{AS} for soil sites, as it is evaluated in the frequency range of 3 Hz to 15 Hz which definitely overlaps with site resonance frequencies. For rock sites, the amplification is lower and is often at higher frequencies due to the very superficial velocity contrast that is generally induced by the

weathered zone. Moreover, the crustal amplification correction realized from the generic hard-rock profile has value basically for the κ_{DS} frequency range. Instead, at higher frequencies, the small-scale information of the velocity profile is generally unknown, which prevents the correction of the site amplification for κ_{AS} (Ktenidou and Abrahamson 2016).

The use of velocimeters is strongly recommended for κ_{DS} because the accelerometers present a much lower sensitivity at low frequencies (<15 Hz; Figure 18). In the present study, we used seismicity catalogs that did not include events for magnitudes below ~1.5. These events are the most suitable for the κ_{DS} approach as they allowed very high f_c (15 to 50 Hz, depending on the stress drop). Thus the κ_{DS} approach can be improved by detection and use of very small and generally local earthquakes from continuous recordings realized at the study site (although this was not done here), especially if the local level of noise perturbation is low. Evaluation of the magnitude and epicentral distance, which are traditionally given by the catalogs, can be difficult, but it is not fully required for κ_r measurement. Indeed, these parameters can be easy avoided by considering a constant f_c equal to that for the earthquake with the lowest magnitude in the catalogs, and by neglecting the distance-dependence term for these local events, or inferring it through the approximation that R_e is proportional to the travel-time difference $T_S - T_P$ measure for each record.

The κ_{DS} approach has been rarely tested. Previous studies have shown generally higher κ_r and κ_0 with the displacement approach than for the acceleration approach, even though both methods are applied to the same records (Kilb et al. 2012; Ktenidou et al. 2017). Kilb et al. (2012) did not observe this tendency for every site, while Ktenidou et al. (2017) found a clear and strong discrepancy from very limited bandwidth data recorded in a low seismicity area. Both studies attributed this to the effects of the smooth transition zone around f_C that is strongly suspected to systematically reduce κ_{r_AS} and increase κ_{r_DS}. Ktenidou et al. (2017) showed that measuring κ_{r_DS} below $f_C/2$ (and symmetrically κ_{r_AS} above $2f_C$) greatly reduces this bias. Its influence might be, however, higher for κ_{DS}, as f_C is much more uncertain for low-magnitude events. Here, we found almost the same results for κ_{0_AS} and κ_{0_DS} for the hard-rock sites. In contrast, the results differed greatly for the soil site, although these differences are very likely to be a consequence of site amplification. In the same way, a systematic discrepancy is observed between the m_{κ_DS} and m_{κ_AS} slopes of the distance-dependence linear model, which can be attributed in part to the frequency dependence of the attenuation. However, there is good agreement between m_{κ_DS} and the regional Q-values from the literature. The κ_{DS} approach thus appeared to be very well adapted for measurement of κ_0 at rock sites. However, further investigations are required to completely understand what controls the reliability and variability of the κ_{DS} measurement.

2.8. Conclusions

The κ parameter is one of the most used and least understood parameters in hazard seismology. This is a 'clue parameter' for host-to-target adjustment, for evaluation of the

hazard for hard-rock sites. Site-specific evaluation of κ_0 is essential, although it is generally difficult at the target site in low-to-moderate seismicity areas. This is because the classical approach with acceleration (Anderson and Hough 1984) requires high magnitude events to ensure low f_C and good SNR up to high frequencies.

In the present study, the dataset used is based on continuous recordings at two hard-rock sites and one stiff-soil site in Provence, France. These were chosen to carry out site-specific κ determination using the classical Anderson and Hough (1984) approach (κ_{AS}) and the approach proposed by Biasi and Smith (2001) (κ_{DS}), which is suitable for low-magnitude events. This evaluation was possible after only a few years of monitoring due to the use of velocimeters, which allowed the recording of much higher numbers of quality events, in comparison with the use of accelerometers. This is particularly true for κ_{DS}, which is measured mainly below 15 Hz, where accelerometers are less sensitive.

Measuring reliable κ_0 values is not easy, as the physics behind κ are not clear, and the uncertainties associated to this parameter remain high. The choice and the application of the method itself can impact upon the variability of κ. For instance, an important variability is introduced in terms of the operator subjectivity in the choice of the frequency window used to determine κ_r. To reduce this inter-operator variability, a semi-automatic procedure was developed here for the frequency window selection that also has the advantage that it provides the uncertainty associated to each individual κ_r. This uncertainty is shown to be mainly dependent on the width of the frequency window. We observe a systematic shift of every κ_r, due to modifications of the spectrum shape by the site amplification, that results in bias for κ_0, even for rock sites. For some sites, this bias might be strongly frequency dependent and prevent the correct determination of κ_0. This appears to be the case for the stiff-soil site, where a strong two-fold discrepancy is observed between κ_{0_AS} and κ_{0_DS}. Moreover, this frequency-dependent phenomenon increase both the variability of each individual κ_r estimation, and the scatter between the κ_r evaluations. The assumption that the attenuation is independent of the frequency made with the definition of κ is questionable. The attenuation is widely accepted to be frequency-dependent, at least for its scattering parts. This influence of the scattering on κ cannot be ruled out and might influence both κ_0 and the slope m_κ of the linear dependence on the epicentral distance. However, only an effect on m_κ is observed, through a strong and systematic three-fold discrepancy between both of these approaches. The comparison of records from the same cluster of events allows the investigation of the relative effects of the source only. We found that the scatter between the κ_{r_AS} evaluations is clearly and strongly dominated by the source spectrum variability, while the magnitude-dependence of κ is suspected, but not clearly established.

In the low-to-moderate seismicity context, the κ uncertainty issue is strengthened due to the narrower spectral windows available. Here, there was high impact of the site amplification on κ, that lead us to discourage its evaluation for soil sites. However, for hard-rock sites that are less affected by site amplification, both of the κ_{AS} and κ_{DS} approaches produced consistent results. The site-specific values of κ_0 were around 30 ms (without site amplification correction) for the hard-rock sites in this study area. This value, which is in the upper part of

the κ_0–V_{S30} correlation, is consistent with the high attenuation indicated for the Provence region (Mayor et al. 2016) and with the κ_{0_AS} obtained by Douglas et al. (2010) for rock sites in the Alps. Moreover, it is in agreement with recent studies that have shown higher κ_0 for hard-rock sites than was initially suggested (Ktenidou et al. 2015; Ktenidou and Abrahamson; 2016, Boore and Campbell, 2017).

The κ_{DS} approach is thus a very promising alternative to the classical approach for sites in a low-to-moderate seismicity context, as this can be carried out using events with smaller magnitudes. This provides a suitable solution for rapid and easy site-specific evaluation of κ_0, with a potential better accuracy for rock sites than the classical κ_{AS} approach. In the present study, we used a seismicity catalog that might not include the smallest magnitude events. Thus, the κ_{DS} approach can be improved by detection and use of very small and local earthquakes that are not provided by seismic bulletins.

2.9. Data and Resources

The seismograms used in this study were collected using a local network that is operated by the French Alternative Energies and Atomic Energy Commission (CEA). Earthquake bulletin information was provided mainly by the Euro-Med Seismological Centre (http://www.emsc-csem.org/#2). If information was missing for an earthquake, information from the *Réseau National de Surveillance Sismique* (http://renass.unistra.fr/), Géoazur (http://sismoazur.oca.eu/), or the Italian Seismological Instrumental and Parametric Database (http://iside.rm.ingv.it/iside/standard/index.jsp) was used.

2.10. Acknowledgments

This study was conducted within the framework of the Cadarache Seismic Hazard Integrated Multidisciplinary Assessment (CASHIMA) Research Programme that is funded by the CEA, the Laue-Langevin Institute (ILL), and the International Thermonuclear Experimental Reactor (ITER) organization. We are thankful to Christopher Berrie and to the two anonymous Reviewers for their careful suggestions and corrections.

Chapitre 3

Évaluation de la Réponse des Sites à Partir des Enregistrements de Mouvements Faibles et Incertitudes Associées

3.1. Introduction

L'évaluation empirique de la fonction de transfert relative du site se fait le plus souvent à travers la méthode standard spectral ratio (SSR) aussi appelée site-sur-référence en français. La méthode SSR est généralement utilisée pour estimer les effets de site spécifiquement pour un site d'intérêt (centrale nucléaire, barrage…) et parfois pour établir localement des cartes de réponse des sols à une sollicitation sismique (microzonage). Cette méthode requiert la présence de zones d'affleurements rocheux où l'effet de site peut être négligé et situées à proximité du site d'étude, ainsi que l'enregistrement de nombreux séismes de bonne qualité (SNR >3) dans une large-bande de fréquence (cf. 1.1.2.2).

Alors que cette approche est couramment employée dans les régions fortement sismiques (Japon, Grèce…), où de nombreux enregistrements de bonne qualité peuvent être réalisés dans une période de temps relativement courte (quelques semaines à quelques mois), elle requiert un temps plus long dans les régions plus faiblement sismiques (parfois plusieurs années). L'objectif principal de ce travail est d'évaluer la quantité, la qualité et les caractéristiques des enregistrements qui sont nécessaires à une évaluation moyenne fiable de la fonction de transfert relative du site par la méthode SSR. Pour ce faire, la détermination de la loi de distribution des résultats SSR est réalisée depuis la base de données du site

d'étude en Provence (cf. 1.3.1) ainsi que depuis la base de données post-sismiques et Argonet à Argostoli (cf. 1.3.2). En effet, ce n'est qu'une fois que la loi de distribution sera connue que l'incertitude statistique associée au SSR pourra être établie de façon précise. Cette incertitude qui décrit la variabilité inter-évènements du SSR, est le plus souvent très importante (>50 %) et est généralement attribuée à la variabilité naturelle et irréductible du phénomène (incertitude aléatoire). Pourtant une portion importante de cette incertitude est fortement suspectée d'être liée à un manque de connaissance (incertitude épistémique) et pourrait être réduite en considérant des paramètres supplémentaires (PGA, magnitude et position des séismes, etc.). La réduction des incertitudes épistémiques associées aux évaluations de l'aléa sismique est l'un des enjeux majeurs de ces dernières années du fait des surcouts et/ou de l'accroissement des risques qu'ils induisent pour le dimensionnement des infrastructures. Nous profitons donc également des bases de données exceptionnellement grandes à notre disposition (en particulier à Argostoli) pour tenter de quantifier la part aléatoire et épistémique des incertitudes mesurées à partir de la méthode SSR. Ce travail apporte un complément aux études récentes déjà réalisées sur le sujet (Boore 2004; Ktenidou et al. 2016; Maufroy et al. 2017).

3.2. Base de données de mouvement faible

Les bases de données utilisées pour le calcul des SSR sont des sous-ensembles des bases de données issues du réseau de Provence et de celui d'Argostoli. En effet, seuls les séismes enregistrés simultanément sur le site d'étude et à la référence ont été utilisés. De même, une distance épicentrale (Re) minimale de 10 km a été retenue en première approche pour assurer une bonne similarité des trajets entre les sites et leurs références. Finalement, seuls les évènements présentant un SNR >3 sur une largeur de bande de plus de 10 Hz sont pris en considération.

La Figure 29 et la Figure 30 présentent les bases de données utilisées pour estimer les SSR à Argostoli et en Provence respectivement. À Argostoli ce sont plus de 6000 séismes qui sont disponibles pour cette étude. La plus grande majorité correspondent à la séquence sismique de janvier-février 2014 située à environ 20 km au nord-ouest du site d'étude. En Provence, suivant le site considéré, moins d'une centaine à plus de 200 séismes ont été utilisés. Ces évènements sont majoritairement distants (>100 km) et se concentrent au nord-est du site d'étude.

3.1. Résultats SSR

La méthode SSR fournit la fonction de transfert relative entre deux sites à toutes les fréquences. La fonction de transfert du site A2 par rapport au site de référence A1 sur le réseau d'Argostoli et celle des sites P2, P3, P4 et P6 vis-à-vis de P1 pour le réseau de Provence ont ainsi pu être établis sur les fenêtres d'ondes S. La Figure 31 présente les résultats SSR pour la composante longitudinale, transversale, horizontale moyenne et

verticale à Argostoli. À chaque fréquence, seuls les évènements pour lesquels le critère SNR>3 est respecté ont été utilisés dans le calcul de la moyenne géométrique. Chaque fréquence possède donc sa propre statistique que l'on peut décrire à l'aide du nombre total d'évènements disponibles et de l'écart-type logarithmique (σ). Ces deux paramètres sont donc également fournis sur la Figure 31 pour chacune des fréquences.

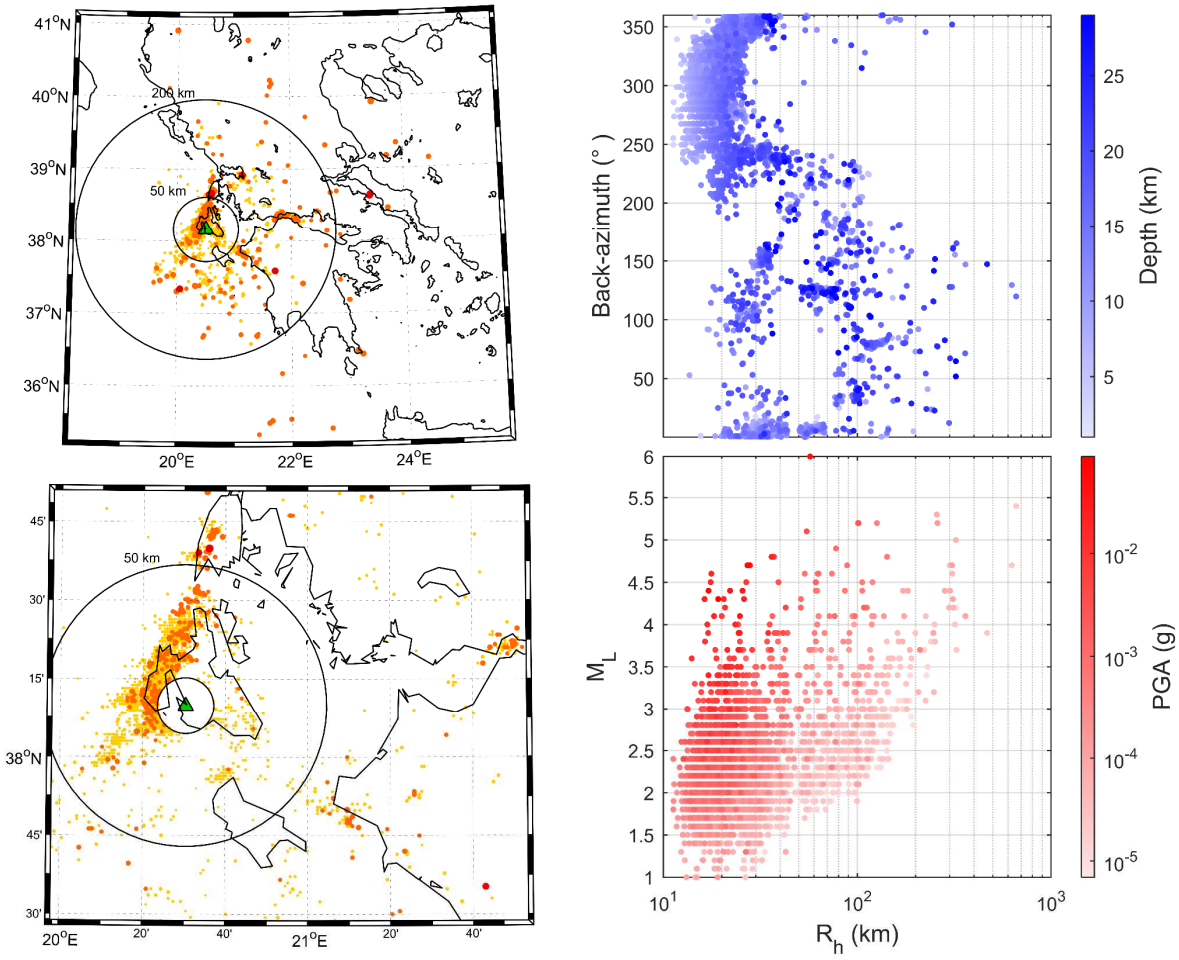


Figure 29. Base de données d'Argostoli utilisée pour la méthode SSR. À gauche : deux cartes présentent la position des épicentres des séismes par rapport au site d'enregistrement. Des cercles concentriques indiquent les distances épicentrales de 10, 50 et 200 km. À droite : le back-azimut et la profondeur des foyers (en haut) ainsi que la magnitude locale et le PGA sont représentés en fonction de la distance hypocentrale.

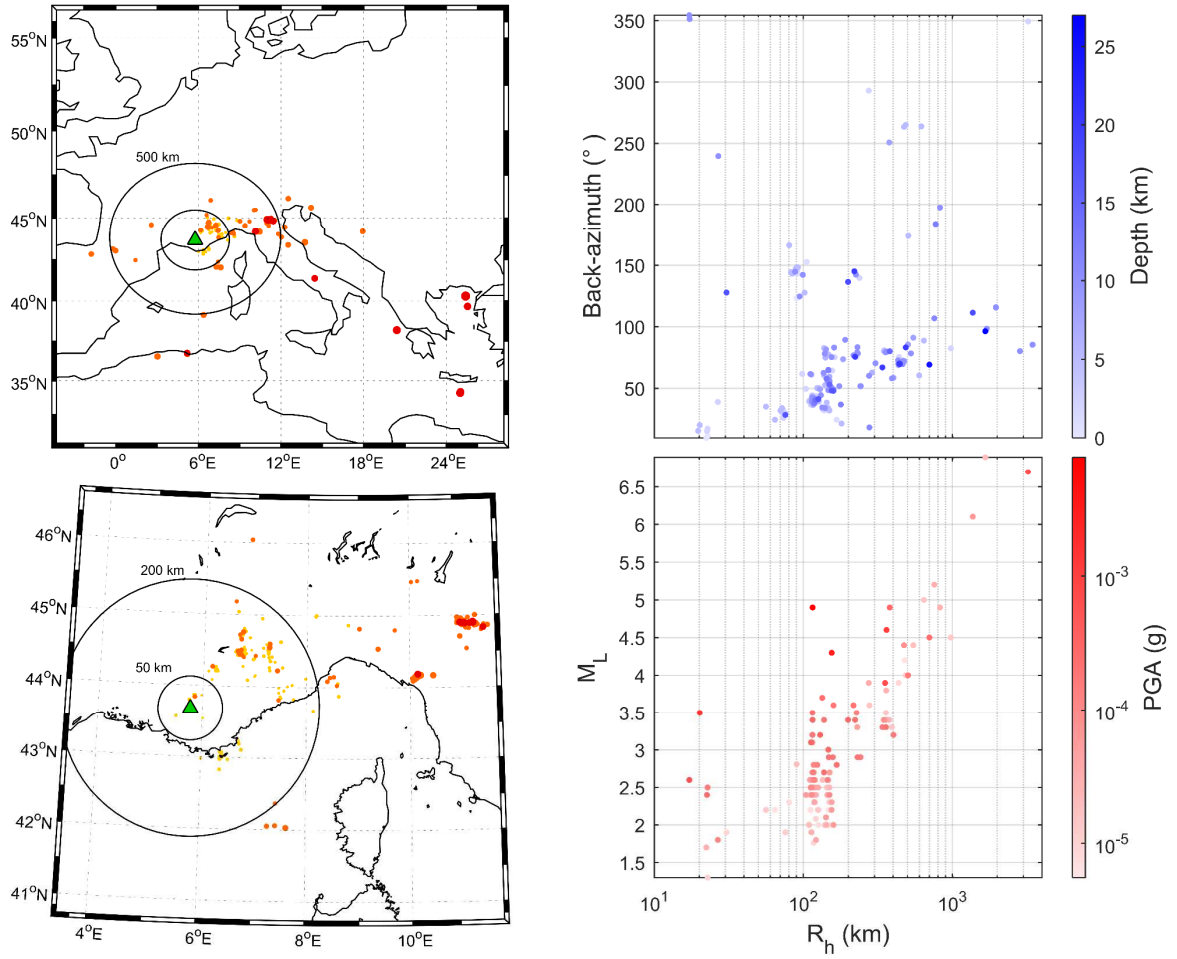


Figure 30. Identique à la Figure 29, mais pour la base de données du site P3 en Provence. Les cercles concentriques indiquent ici les distances épicentrales de 50, 200 et 500 km.

Les résultats montrent que l'amplification dans les directions horizontales est très importante, avec des facteurs dépassant même 10 à certaines fréquences. La composante verticale indique des niveaux d'amplification plus faibles que l'horizontale, mais qui ne sont absolument pas négligeables. Le grand nombre d'évènements disponibles à Argostoli permet de quantifier très précisément l'incertitude associée à la mesure SSR pour ce site. Celle-ci est assez élevée puisqu'à toutes les fréquences 10^{σ} est proche de 1,5 soit un coefficient de variation (CV) de 50 %. En effet, le coefficient de variation peut être obtenu simplement par la relation $CV = (10^{\sigma} - 1) * 100$. La variabilité des résultats sur la composante horizontale moyenne et verticale est du même ordre de grandeur, tandis que les composantes longitudinale et transversale prises individuellement présentent des variations plus fortes que la moyenne quadratique de ces dernières (composante horizontale moyenne). Sur la Figure 31, il apparait notablement que le SSR moyen sur la composante longitudinale diffère de celui de la composante transversale. En effet, les variations des résultats SSR suivant l'orientation des composantes traduisent les variations de la réponse du bassin dans les différentes directions de l'espace. Celles-ci peuvent varier fortement en

fonction des modes de résonance du bassin et de la génération d'ondes de surface sur les bords de celui qui présente souvent une orientation privilégiée. Cette variabilité azimutale de la réponse des sites est étudiée plus en détail avec la Figure 32 qui montre les variations relatives par rapport à la moyenne quadratique de la réponse horizontale du site en fonction de l'azimut pour le site Argostoli. La couleur rouge indique la direction longitudinale au bassin tandis que la couleur bleue donne la direction transversale. Tout écart autour de la valeur 1 indique que la réponse moyenne du site est plus forte (>1) ou moins forte (<1) dans une direction particulière que celle fournie par la moyenne quadratique des composantes horizontales. On peut déjà noter la pertinence de l'utilisation de la moyenne quadratique (équation 19) qui se situe bien au centre des résultats évalués suivant toutes les directions de l'espace. De même, le choix communément fait de ne représenter que la composante longitudinale et transversale se justifie aussi puisque que ces deux directions correspondent bien aux extrema observés ici. Ainsi, la comparaison des résultats SSR selon ces deux directions seulement fournit une information suffisante sur la variabilité de la réponse des sites en fonction de l'azimut. L'amplitude de cette variabilité azimutale est forte à Argostoli (±30 %), probablement dûe à l'ampleur des phénomènes 2D ou 3D à l'intérieur du bassin.

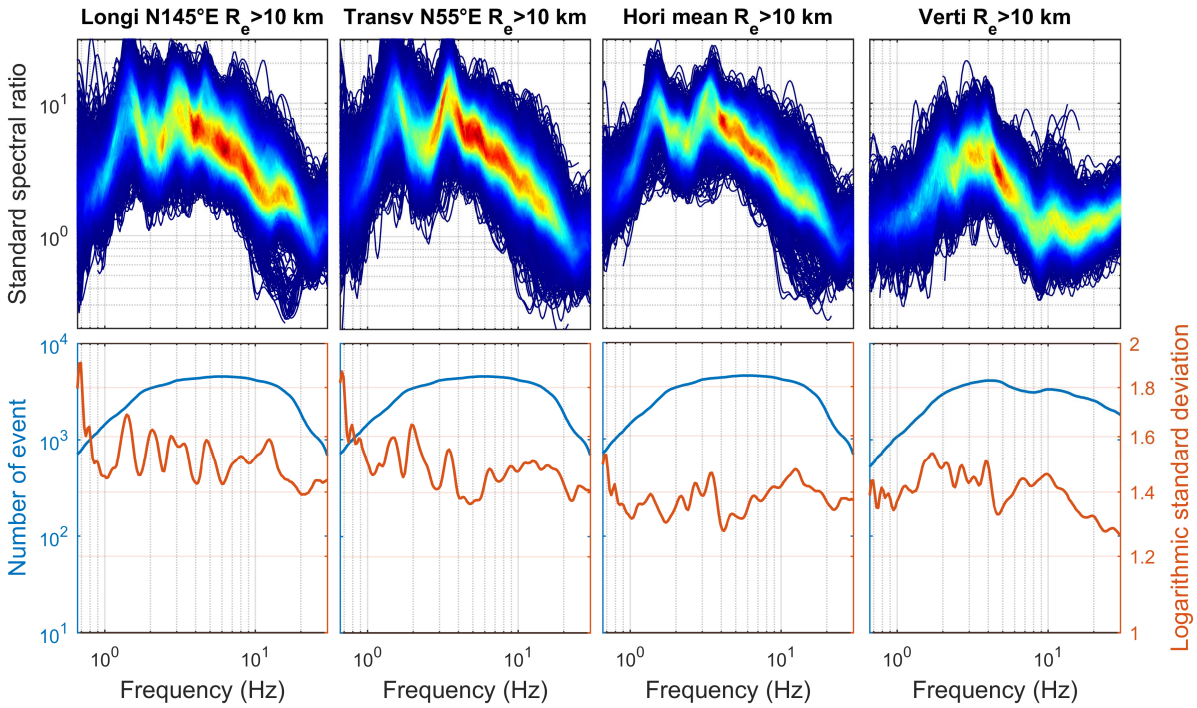


Figure 31. Fonction de transfert SSR à Argostoli (en haut) pour (de gauche à droite) la composante longitudinale, transversale, horizontale moyenne et verticale entre le site sédimentaire A2 par rapport au site de référence A1 situé au rocher. L'échelle de couleur indique la densité de point. Le nombre d'évènements sismiques utilisés ainsi que l'écart-type logarithmique en fonction de la fréquence sont également représentés (en bas).

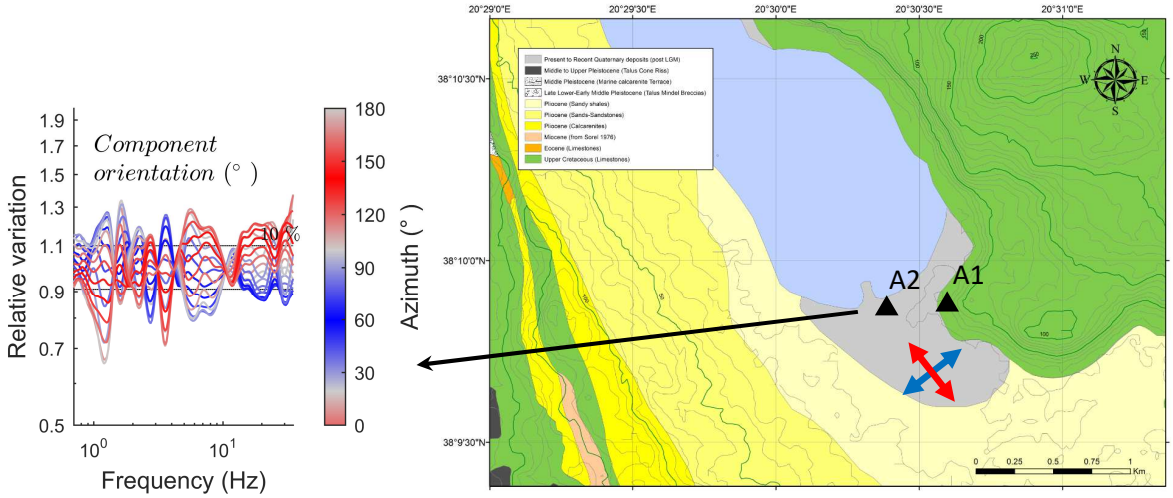


Figure 32. Variabilité azimutale de la réponse du site A2 par rapport à A1 à Argostoli. Les variations relatives de la réponse du site sont indiquées en fonction de l'azimut et de la fréquence. La direction longitudinale au bassin est en rouge tandis que celle transversale est en bleue.

La Figure 33 fournit les résultats SSR des composantes horizontales moyennes pour 4 sites en Provence. Premièrement, il est intéressant de noter la grande variabilité spatiale et fréquentielle de la réponse entre les différents sites. Cette variabilité est bien illustrée par le grand écart entre les fonctions de transfert SSR des sites P4 et P6 pourtant situés à quelques dizaines de mètres l'un de l'autre seulement. Ensuite, une amplification relative est également visible à haute fréquence sur le site P2, alors même qu'il s'agit d'un rocher dur proche de celui présent sur le site de référence P1. Cette différence pourrait s'expliquer par la présence sous le site P2 d'une couche superficielle plus altérée que pour P1. De même, le fait que la station en P1 soit située dans une cave sismique à 3 m de profondeur tandis que la station P2 est en surface peut expliquer, pour les très hautes fréquences (> 20 Hz), une partie de l'écart observé. Cela illustre la difficulté inhérente au choix des stations de référence. La faible dispersion des résultats SSR en Provence se traduit par une incertitude ($CV \approx 20\%$) qui est bien inférieure à celle observée à Argostoli. Cette incertitude est du même ordre de grandeur pour les 4 sites d'enregistrements en Provence. Cela suggère que le niveau d'incertitude est défini au premier ordre par la géométrie générale du site localement, mais dépend, probablement, également du jeu de données considéré. De plus, le niveau d'incertitude est indépendant de celui de l'amplification relevée sur les fonctions de transfert des différents sites.

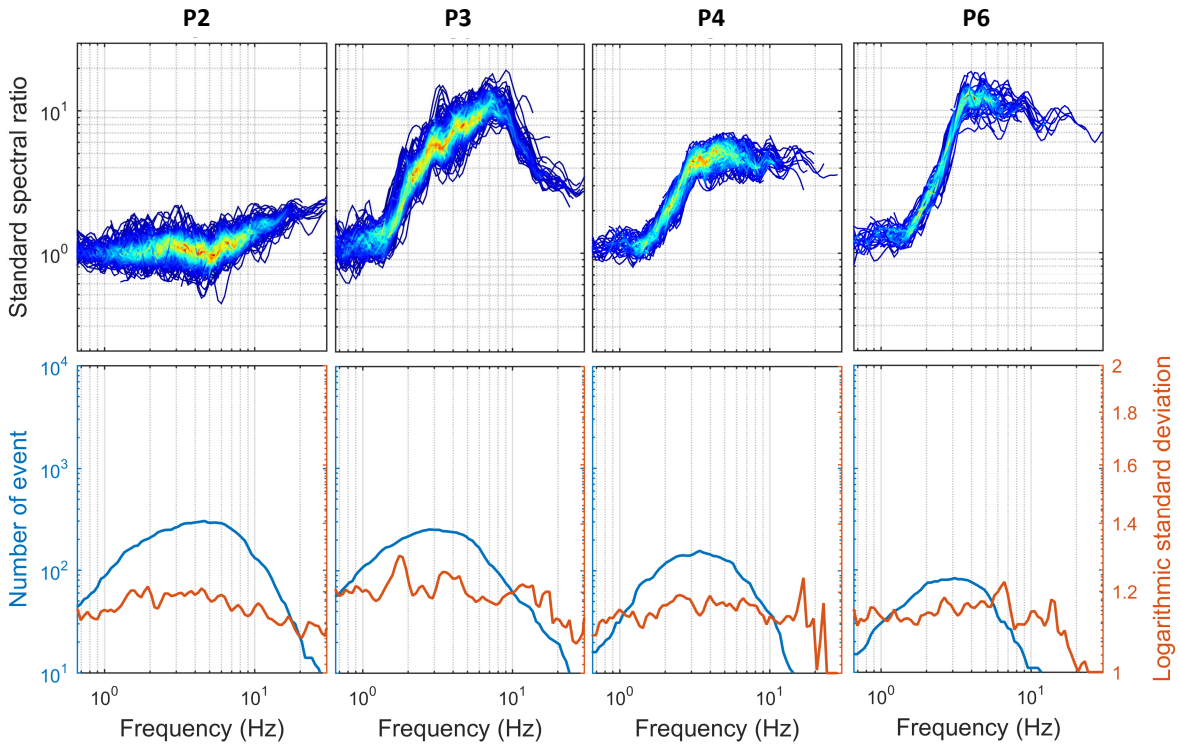


Figure 33. Identique à la Figure 31, mais pour les sites P2, P3, P4 et P6 relativement au site P1 du réseau de Provence et pour la composante horizontale moyenne uniquement.

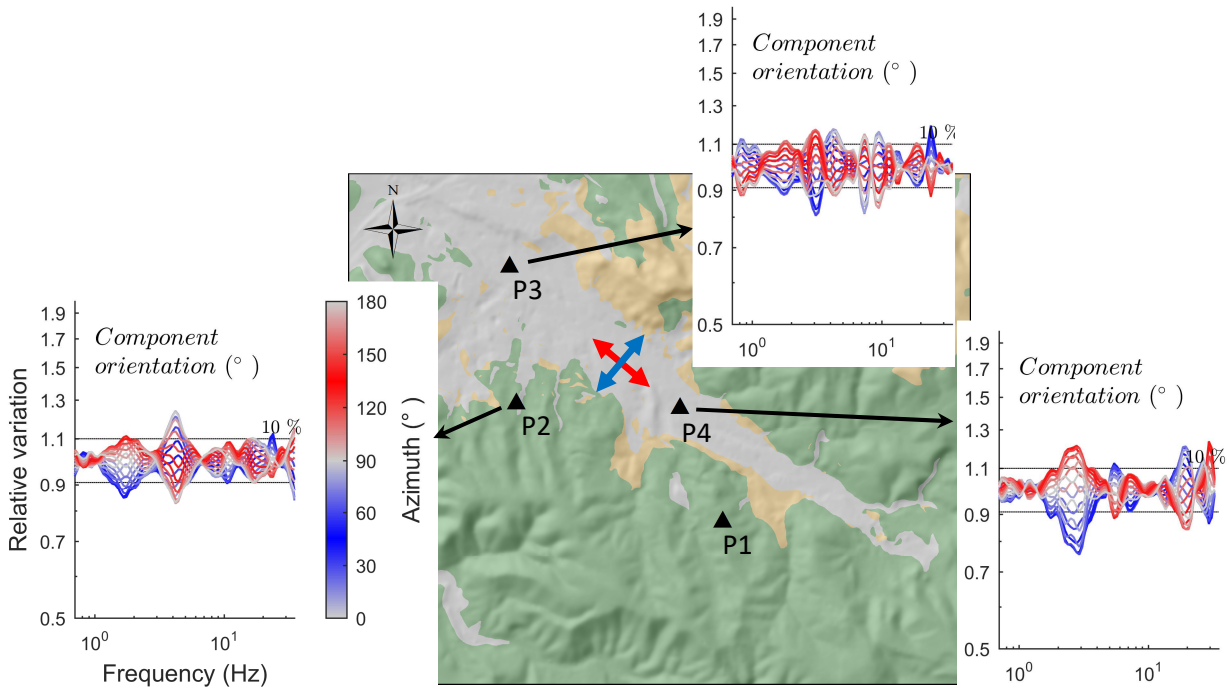


Figure 34. Identique à la Figure 32, mais pour la variabilité azimutale de la réponse des sites P2, P3, et P4 par rapport à P1 en Provence.

La Figure 34 présente la variabilité de la réponse des sites à l'orientation des composantes pour les sites P2, P3 et P4 en Provence, pareillement à la Figure 32 pour Argostoli. Les mêmes conclusions que celles qui avaient été formulées pour Argostoli s'appliquent ici également, avec cependant un niveau de variabilité plus bas. Ceci suggère que la variabilité azimutale est un bon indicateur de l'ampleur des phénomènes 2D ou 3D à l'intérieur du bassin. L'alternance bien marquée du rouge vers le bleu et inversement en fonction de la fréquence semble indiquer les changements de mode de résonance des deux bassins suivant la direction longitudinale ou transversale. Ceci est particulièrement marqué sur le site P4 en Provence. Seule la composante horizontale moyenne sera traitée dans la suite pour simplifier la lecture des résultats, et parce que les résultats tirés des composantes longitudinales et transversales au bassin ne nous ont pas semblé apporter suffisamment d'informations complémentaires.

Le critère SNR >3 est très largement utilisé dans des études très variées alors même qu'il n'a pas fait l'objet d'une véritable validation à notre connaissance. La Figure 35 montre les variations relatives du SSR évalué depuis différents intervalles de SNR par rapport au SSR évalué pour un critère SNR >3. Il apparait très clairement que les résultats issus des seuils de SNR inférieur à 3 sont très variables, tandis que ceux pour lesquels les rapports SNR >3 montrent une très grande stabilité. Cette observation conforte le choix du seuil SNR >3 pour cette étude, ainsi que pour toute étude utilisant ce critère.

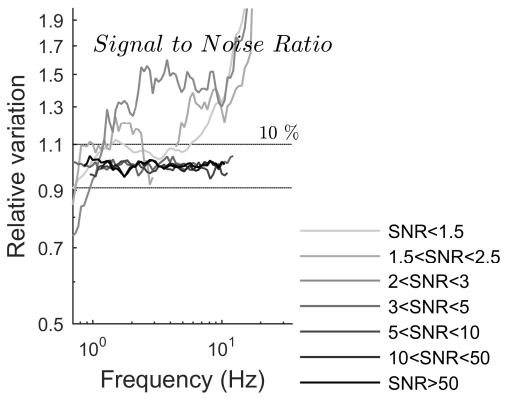


Figure 35. Influence du seuil de rapport signal-sur-bruit (SNR) sur l'évaluation du SSR sur le site P4 en Provence.

3.2. Analyses statistiques

Si un grand nombre de séismes a pu être utilisé pour réaliser les analyses SSR présentées Figure 31 et Figure 33, une question opérationnelle se pose quant au nombre minimum de séismes nécessaires pour obtenir une évaluation empirique fiable de la réponse moyenne des sites par méthode SSR. Il est souvent admis (Ktenidou et al. 2011) et nos premières observations Figure 31 et Figure 33 semblent le confirmer, que les résultats issus des méthodes par rapports spectraux suivent une distribution log-normale. Ceci a des

implications fortes puisque si la loi de distribution est connue, alors le comportement de la moyenne peut être anticipé à l'aide de l'intervalle de confiance (I_C). Par exemple, l'intervalle de confiance de 95 % ($I_{C_95\%}$) définit l'intervalle dans lequel la moyenne estimée depuis un nombre infini d'échantillons a 95 % de chance d'être contenue. Cet intervalle se formule comme suit :

$$I_C = \left[\overline{x} - t_\alpha \frac{\sigma}{\sqrt{n}} \; ; \; \overline{x} + t_\alpha \frac{\sigma}{\sqrt{n}} \right] \qquad 30$$

avec $\overline{x}$ l'espérance totale, σ l'écart-type total et n la taille de l'échantillon (nombre de séismes) ciblé. t_α correspond au coefficient de confiance qui vaut 1,96 pour un intervalle de confiance de 95 %. $I_{C_95\%}$ permet ici de prédire à partir des moyennes et écarts-types estimés depuis l'ensemble des séismes disponibles, l'intervalle dans lequel la moyenne évaluée depuis un sous-ensemble de n évènements a 95 % de chance de se trouver.

La Figure 36 (a) montre des exemples de distribution des résultats SSR à 1, 4, 8 et 16 Hz. Les données semblent s'accorder presque parfaitement selon une gaussienne lorsque l'espace est logarithmique. La distribution log-normale est évidente à Argostoli. La Figure 36 (b) teste l'influence, aux mêmes fréquences que précédemment, du nombre d'évènements n sur la robustesse des rapports SSR calculés. Pour des valeurs de n allant de 2 à 1000, 500 tirages aléatoires ont été réalisés (approche « bootstrap ») donnant accès, à chaque fois, à 500 évaluations de la moyenne depuis n échantillons (on parlera ici de « moyennes rééchantillonnées »). Ces valeurs moyennes rééchantillonnées ont ensuite été divisées par la moyenne sur l'ensemble des échantillons (on parlera de « moyenne totale ») pour obtenir leurs variations relatives. Ces variations relatives pour les 500 tirages aléatoires sont représentées par les courbes grisées en fonction du nombre d'échantillons n. Les percentiles 2,5 et 97,5 % déduits des 500 moyennes rééchantillonnées (courbes bleues) délimitent l'intervalle dans lequel 95 % des moyennes rééchantillonnées sont comprises et donc peut-être comparé directement à l'$I_{C_95\%}$ prédit par la distribution log-normale (courbes vertes). Différemment, la Figure 36 (c) fournit l'évolution, cette fois en fonction de la fréquence, de la robustesse du calcul de la moyenne rééchantillonnée à partir de 8, 15, 50 et 100 évènements tirés aléatoirement 500 fois. Le code couleur des courbes représentées est le même que celui de la Figure 36 (b). La forme que prend la distribution des résultats en fonction de la fréquence est directement dépendante de la forme de l'écart-type total (Figure 31), comme prédit par l'intervalle de confiance (équation 30).

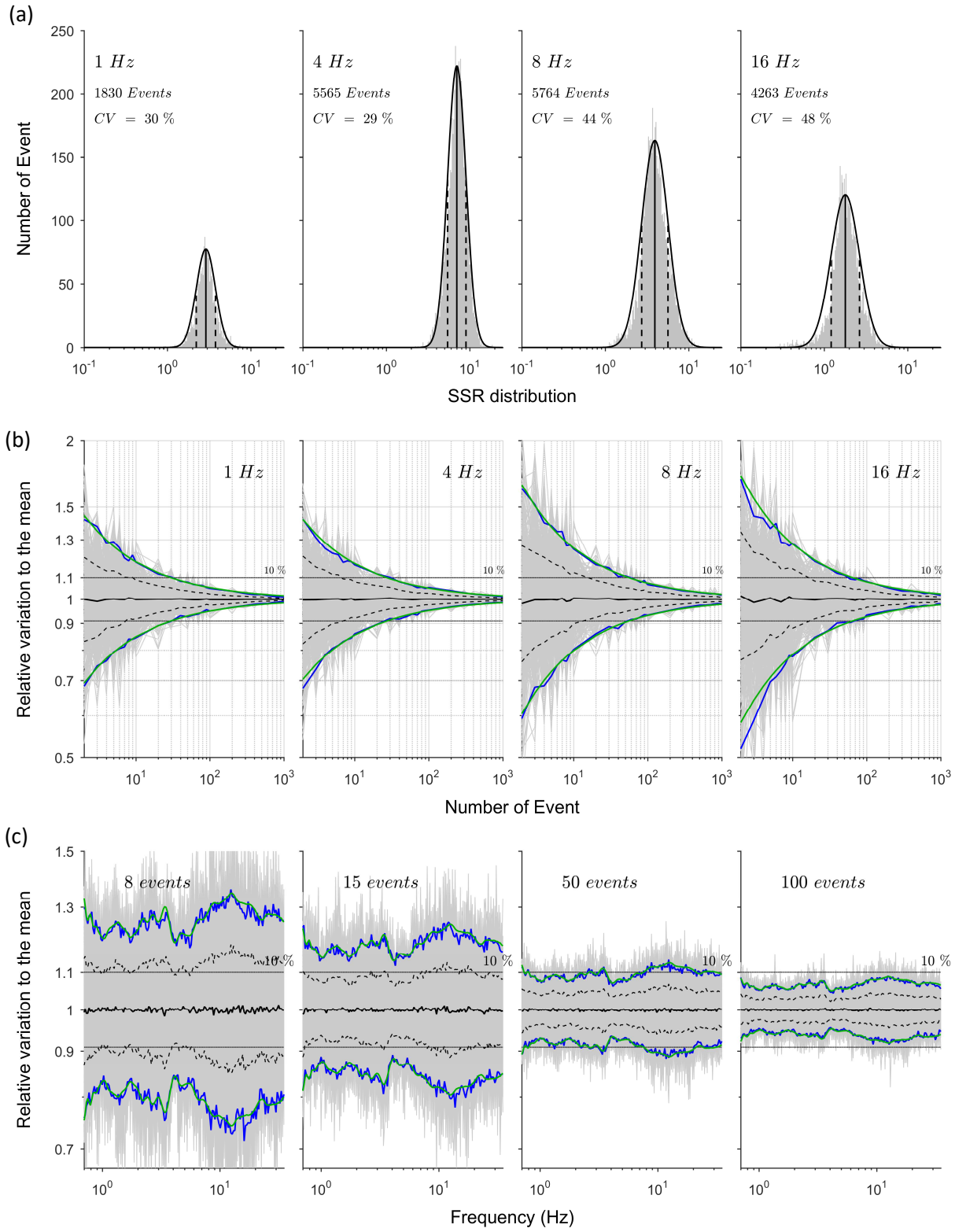


Figure 36. Évaluation de la statistique rattachée aux résultats SSR entre le site A2 et A1 à Argostoli. (a) Distribution des résultats SSR individuels à 1, 4, 8 et 16 Hz et comparaison avec la distribution log-normale correspondante. La courbe pleine et les courbes pointillées représentent

respectivement la moyenne et l'écart-type logarithmique. (b) Variation relative entre 500 tirages aléatoires de moyennes rééchantillonnées calculées depuis un nombre d'enregistrements (n) allant de 2 à 1000 et la moyenne totale déterminée depuis l'ensemble évènements à 1, 4, 8 et 16 Hz. (c) Variation relative entre 500 tirages de moyennes rééchantillonnées avec n égale à 8, 15, 50 et 100 évènements et la moyenne totale pour l'ensemble des fréquences. Sur les figures (b) et (c), les courbes noires pleines et pointillées représentent la moyenne et l'écart-type estimés depuis les 500 tirages aléatoires tandis que les courbes bleues donnent le percentile 2.5 et 97,5 %. Les courbes vertes indiquent l'intervalle de confiance 95 % estimé pour le nombre n d'évènements considéré et depuis la moyenne et l'écart-type déterminés sur l'ensemble des évènements disponibles.

La Figure 37 montre l'évolution du percentile 97,5 % ($P_{97.5\%}$) déduit des 500 évaluations de moyenne rééchantillonnée (à gauche) et de la borne supérieure de $I_{C_95\%}$ (au centre) en fonction du nombre d'évènements et de la fréquence. Cette figure fait la synthèse des résultats présentés Figure 36 puisque l'échelle de couleur y décrit la décroissance en fonction de n observée Figure 36 (b) tandis que l'évolution des résultats est décrite en fonction de la fréquence comme sur la Figure 36 (c). Le bon accord entre $P_{97.5\%}$ et la borne supérieure de $I_{C_95\%}$ est évident ici pour toutes les combinaisons de n et de fréquence. Le rapport $P_{97.5\%}/I_{C_95\%}$ (à droite) montre bien qu'à partir de 8 évènements seulement, $P_{97.5\%}$ et $I_{C_95\%}$ sont très similaires (<5 % de variation relative) et donc la représentativité statistique du SSR est assurée.

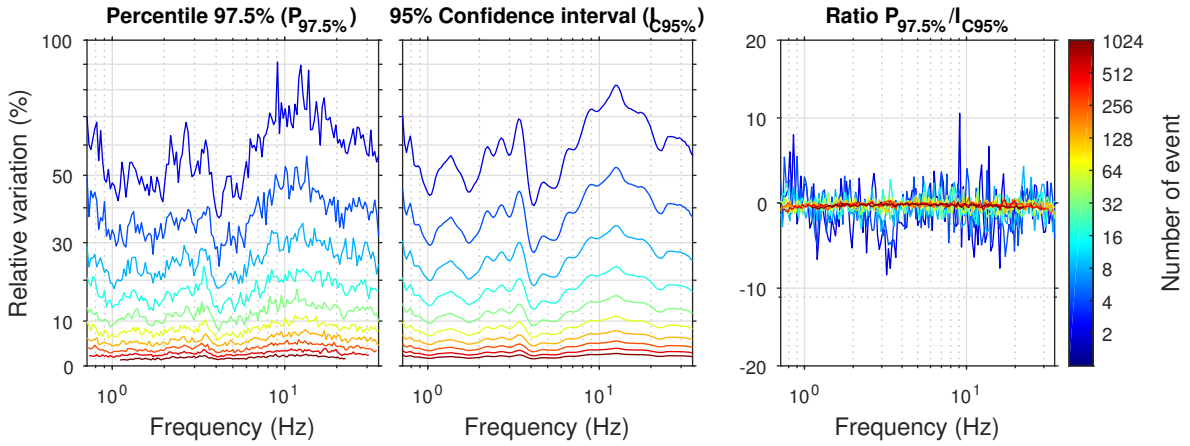


Figure 37. Comparaison du percentile 97,5 % (à gauche) déduit de l'évaluation de 500 moyennes rééchantillonnées avec la borne supérieure de l'intervalle de confiance 95 % (au centre) depuis 2, 4, 8, 16, 32, 64, 128, 256, 512 et 1024 évènements (n) tiré aléatoirement (échelle de couleur) et en fonction de la fréquence. Les courbes sont normalisées par la moyenne déterminée depuis l'ensemble des évènements disponibles. Le panneau de droite fournit l'évolution du rapport entre les deux paramètres en fonction de n et de la fréquence.

La borne supérieure de $I_{C_95\%}$ et les variations relatives des moyennes rééchantillonnées montrent une forte diminution lorsque n augmente. La limite des 10 % d'écart relatif par rapport à la moyenne totale est, par exemple, franchie à partir d'une cinquantaine d'évènements ($n = 50$) pour pratiquement toutes les fréquences (Figure 36 (c) et Figure 37).

Ce nombre de 50 évènements utilisé pour évaluer la moyenne des SSR ($n = 50$) semble donc bien adapté à Argostoli pour assurer d'avoir 95 % de chances à toutes les fréquences de se situer à moins de 10 % d'écart relatif de celui qui serait obtenu depuis une infinité de mesures.

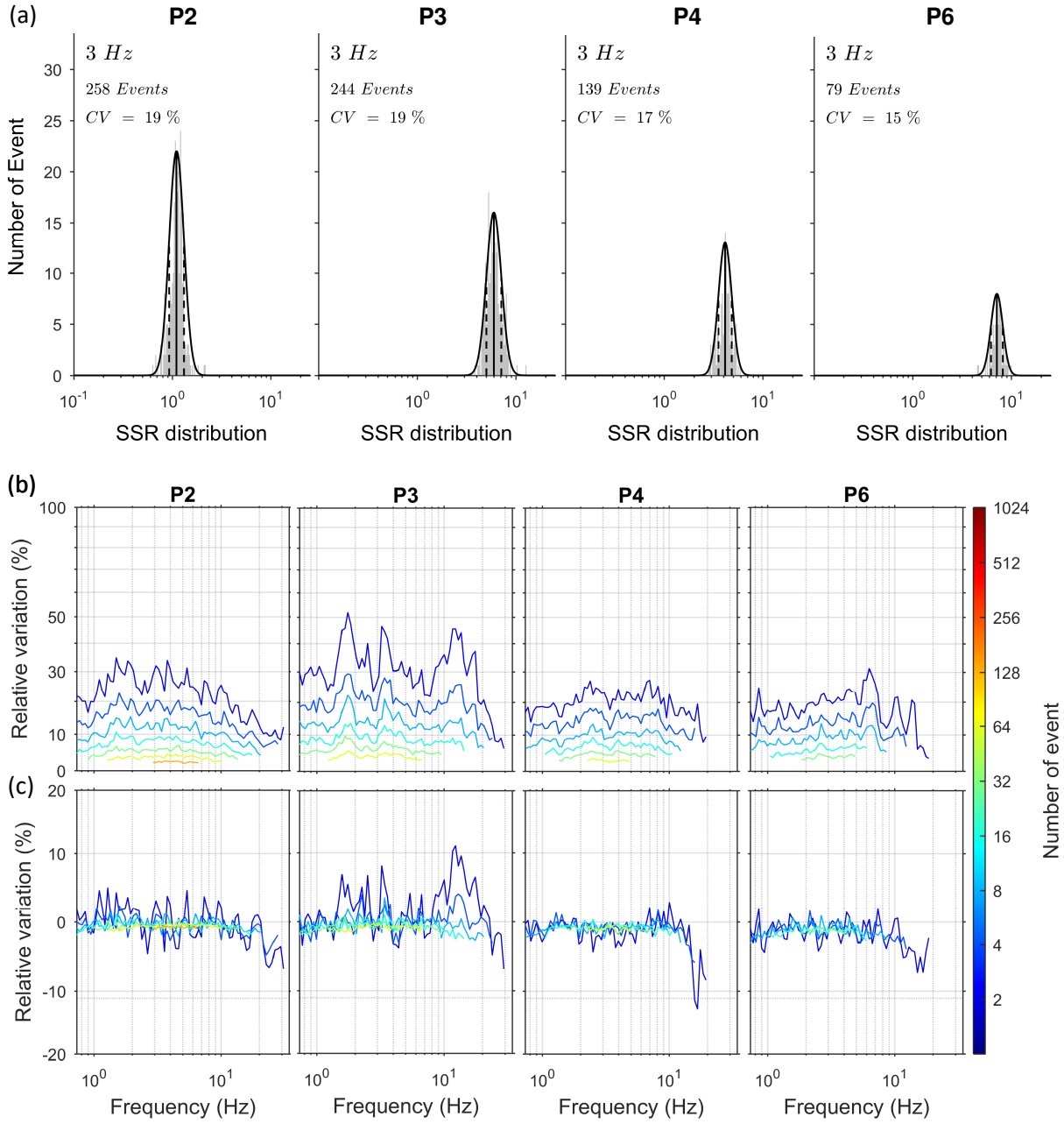


Figure 38. Évaluation de la statistique rattachée aux résultats SSR aux sites P2, P3, P4 et P6 par rapport à P1 en Provence. (a) Identique à la Figure 36 (a). (b) et (c) identique à la Figure 37 gauche et droite respectivement.

La Figure 38 (a) donne la distribution des résultats SSR à 3 Hz pour les sites P2, P3, P4 et P6 en Provence tandis que la Figure 38 (b) et (c) fait la comparaison (comme sur la Figure 37) entre le percentile 97,5 % évalué par bootstrap et les résultats prédits par la borne supérieure de $I_{C_95\%}$ en fonction de n et de la fréquence pour ces mêmes sites. La distribution des résultats SSR à 3 Hz est très proche de la distribution gaussienne. De la même façon, les valeurs de $P_{97.5\%}$ s'accordent très bien avec celles prédites par $I_{C_95\%}$. Comme pour les résultats d'Argostoli et malgré un nombre d'évènements total disponible bien plus bas, la distribution log-normale des résultats SSR est évidente pour tous les sites en Provence. $P_{97.5\%}$ montre que pour les sites de Provence le nombre d'évènements requis pour avoir une évaluation fiable de la moyenne (<10 % d'écart relatif par rapport à la moyenne totale dans 95 % des cas) n'est que de $n = 15$ (Figure 38 (b)) alors qu'il était de $n = 50$ à Argostoli (Figure 37). Cette différence est liée à la dispersion beaucoup plus réduite du SSR pour les sites de Provence par rapport à celui d'Argostoli déjà observée Figure 31 et Figure 33. Une question se pose alors sur l'origine de cet écart de variabilité observé entre les deux sites et plus généralement sur la nature des incertitudes associées aux mesures SSR.

3.3. Incertitude épistémique

On a vu dans la section précédente que la variabilité des résultats SSR peut être décrite comme un phénomène purement aléatoire qui suit une distribution log-normale. Cependant, l'écart d'incertitude observé entre le site d'Argostoli et les sites de Provence nous suggère qu'au moins une partie de cette variabilité est épistémique et dépendante du site de mesure et/ou du jeu de données utilisé. Dans cette partie, nous tentons de déterminer l'origine de cette incertitude en évaluant les SSR rééchantillonnés non pas de manière aléatoire cette fois, mais selon plusieurs paramètres liés à la source : back-azimut (bAz), distance épicentrale (R_e), profondeur (Z) et magnitude locale (M_L).

La distribution log-normale avérée des résultats SSR fait de l'intervalle de confiance un critère fiable pour détecter l'existence ou non d'une variabilité épistémique. Dans la suite de ce travail, afin de minimiser la portion de moyenne rééchantillonnée pouvant aléatoirement excéder l'intervalle de confiance, nous avons utilisé la loi de student avec un intervalle de 99 % ($I_{C_99\%}$). Cette loi est similaire à la loi normale, mais elle permet de tenir compte de la représentativité statistique du résultat en élargissant $I_{C_99\%}$ lorsque n décroit. De plus, un nombre minimum de 10 évènements ($n = 10$) est considéré pour assurer une représentativité statistique minimum. Ainsi, la probabilité pour qu'un SSR moyen évalué aléatoirement depuis n évènements excède $I_{C_99\%}$ est assurée d'être inférieure à 1 %. De cette façon, tout dépassement d'$I_{C_99\%}$ par le SSR moyen évalué depuis n évènements sélectionnés spécifiquement selon un paramètre particulier (R_e, bAz, Z ou M_L) aura très peu de chance d'être lié au hasard (<1 %) et pourra être attribué à l'influence du paramètre testé. Il est également à noter que les évènements n'ont pas une répartition spatiale et en magnitude homogène, ce qui nuit à la représentativité de la moyenne totale. Par exemple, à

Argostoli une très grande majorité des évènements est concentrée au nord-est du site et à courte distance épicentrale (< 50 km). Afin de limiter ce biais, la moyenne totale est calculée comme la moyenne logarithmique des moyennes rééchantillonnées sur le paramètre considéré. Ceci permet de donner un poids équivalent à chaque groupe de données rééchantillonnées. Seules les données du site d'Argostoli seront traitées dans cette partie du fait de la trop faible quantité d'évènements disponibles et de leurs distributions très hétérogènes pour les sites de Provence.

3.3.1. Influence du back-azimut

Les bassins sédimentaires présentant des géométries 3D fortement asymétriques comme à Argostoli peuvent induire des réponses de site très variées suivant l'éclairage qui leur est donné. Le back-azimut traduit la façon dont le bassin est éclairé par la source. La Figure 39 montre les résultats SSR lorsque les séismes utilisés sont spécifiquement sélectionnés en fonction de ce paramètre. Quatre groupes sont considérés : les séismes situés du côté sud-est dont les ondes arrivent par « l'arrière » du bassin (marron), ceux situés du côté nord-ouest propageant des ondes arrivant par son embouchure (en rouge) et les deux directions transversales (en vert et bleue). La moyenne des SSR montre une forte variabilité suivant le back-azimut qui dépasse largement la variabilité aléatoire discutée dans la section précédente et qui, même si elle est présente à toutes les fréquences, semble être plus forte au-delà de 1,5 Hz qui est la fréquence de résonance du site (Figure 39 (c3) et (c4)). Les facteurs d'amplification des pics à 1,5 et 3,5 Hz sont plus importants pour les séismes émanant de l'embouchure du bassin et par le bord sud-ouest de celui-ci. À l'inverse, les amplitudes les plus basses sont observées pour les séismes arrivant latéralement depuis le nord-est du bassin. La moyenne totale (depuis tous les évènements - en noir) montre une tendance à suivre préférentiellement les courbes rouges et vertes du fait de la surreprésentation de ces deux sous-groupes en terme de nombre d'évènements (Figure 39 (c1)). Au-delà de 4 Hz, les résultats se scindent en deux groupes, avec d'un côté les séismes émanant au nord du site (rouge et bleue) qui présentent des niveaux d'amplification plus bas et ceux situés au sud (vert et marron) qui sont plus fortement amplifiés. Les niveaux d'incertitude sont tous sensiblement du même ordre de grandeur (Figure 39 (c2)) et similaire à ceux obtenus depuis l'ensemble des évènements (en noir). Ceci indique qu'une variabilité importante persiste au sein de chaque intervalle de back-azimut. Une diminution de la largeur de ces intervalles en dessous des 60 ° utilisé ici pourrait donc permettre d'accentuer les effets observés sur la moyenne. Ceci n'a cependant pas été réalisé ici afin d'assurer un nombre minimum d'évènements disponible sur chaque intervalle, pour des questions de représentativité statistique des résultats. Cette variabilité persistante peut également être liée à l'influence d'autres paramètres tels-que M_L ou R_e.

Le bassin d'Argostoli a déjà fait l'objet de plusieurs études, en particulier dans le cadre du programme de recherche NERA (cf. section 1.3.2). Imtiaz (2015) propose, par exemple, une étude approfondie de la composition du champ d'ondes sismiques réalisée à partir de 46 évènements présentant une distribution homogène des back-azimuts et enregistrés sur un réseau dense situé au centre du bassin. La Figure 40 associe à toutes les fréquences, la

proportion d'ondes de Love et Rayleigh évaluée dans le champ d'ondes, avec le SSR réalisé au centre du réseau. Il y apparait clairement que le premier pic d'amplification à environ 1,5 Hz est majoritairement contrôlé par les ondes de Love tandis que le second pic à 3-4 Hz est attribuable au mode fondamental des ondes Rayleigh. Au-delà de 6 Hz, les deux types d'ondes sont présents approximativement dans les mêmes proportions. Ainsi, les ondes de surface dominent le champ d'ondes et contribuent significativement aux amplifications observées dans le bassin d'Argostoli.

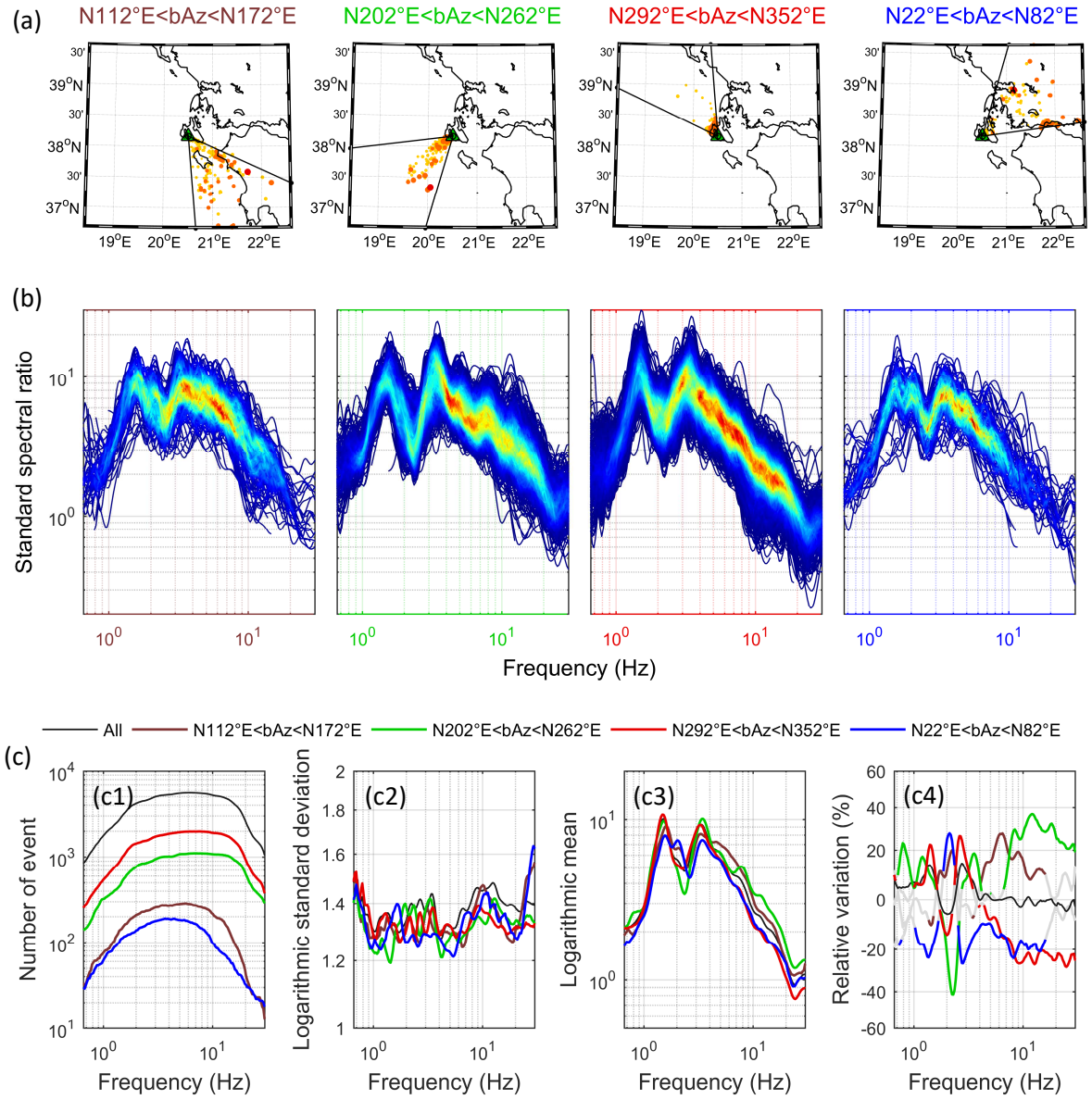


Figure 39. Influence du back-azimut (bAz) sur la variabilité des résultats SSR entre les sites A2 et A1 à Argostoli. (a) Carte des épicentres des séismes sélectionnés pour tester quatre intervalles de bAz . (b) Distribution SSR associée à chaque groupe de séismes. L'échelle de couleur indique la densité de point. (c) Paramètres statistiques tirés des quatre distributions SSR et comparés aux

résultats issus de la distribution totale depuis l'ensemble des évènements (en noir). (c1) Nombre de séismes utilisés, (c2) écart-type logarithmique (10°), (c3) moyenne logarithmique, (c4) variation relative des moyennes rééchantillonnées par rapport à leur moyenne. En (c4), pour chaque distribution, seules les variations relatives excédant $I_{C_99\%}$ sont représentées en couleur (les parties de courbes ayant plus de 99 % de chance d'être attribuable à une variation aléatoire de la moyenne rééchantillonnée sont grisées).

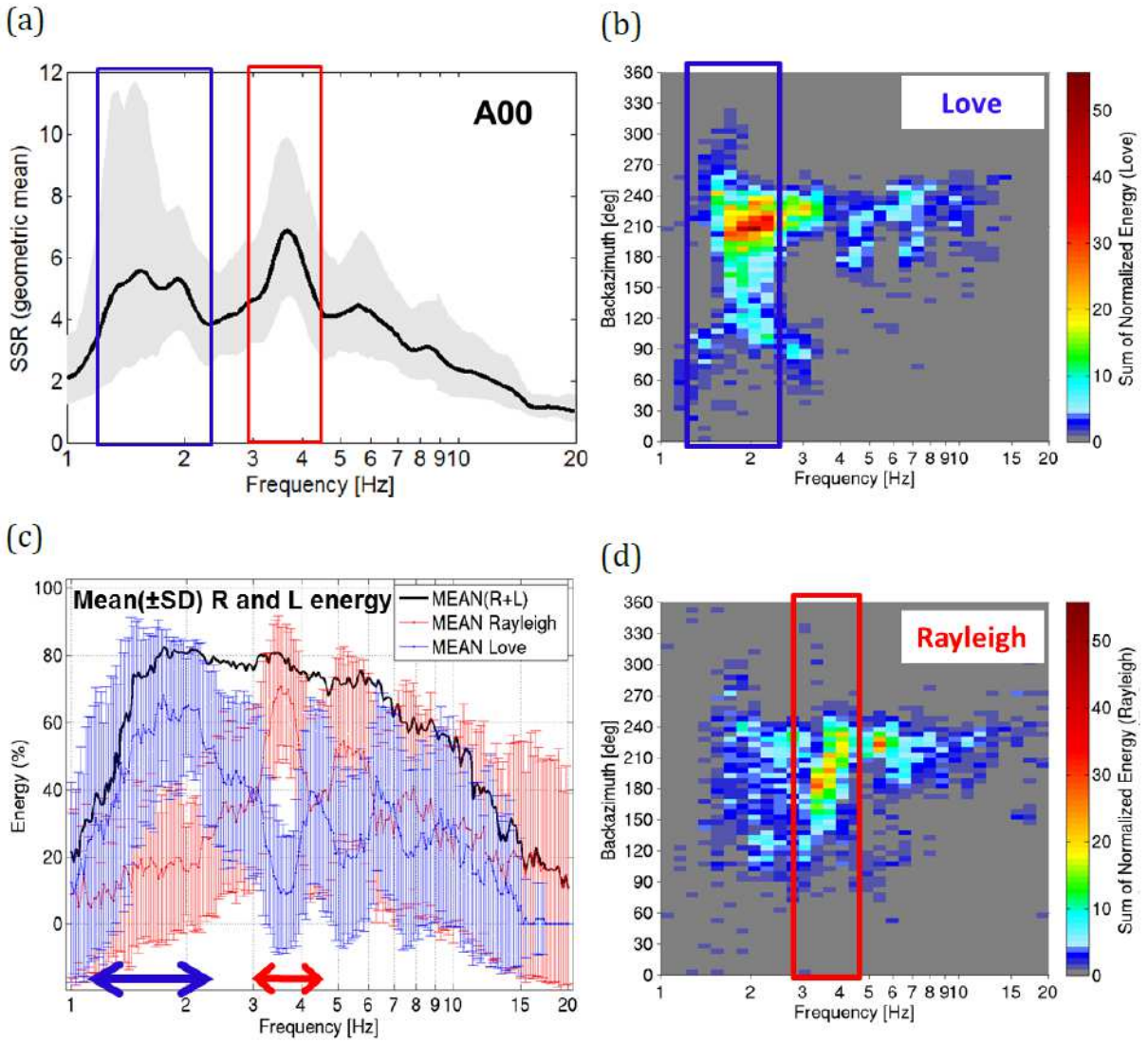


Figure 40. Association des résultats SSR avec la composition du champ d'ondes déduite des enregistrements de 46 évènements sur un réseau dense situé au centre du bassin d'Argostoli (d'après Imtiaz 2015). (a) Composantes horizontales moyennes du SSR estimé depuis 164 évènements enregistrés au centre du réseau (Cultrera et al. 2014). La moyenne logarithmique est indiquée par la ligne noire tandis que la largeur de la bande grise donne ±1σ. (c) Moyenne ±1σ de la distribution d'énergie cumulée pour chaque type d'onde. Histogramme de la distribution de back-azimut en fonction de la fréquence pour (b) les ondes de Love et (d) de Rayleigh.

Imtiaz (2015) montre dans son étude que les ondes de surface diffractées par le bord sud-ouest du bassin dominent le champ d'ondes au-delà de 1,5 Hz quel que soit le back-azimut de la source et la portion d'onde S considérée. Ceci est visible Figure 40 (b) et (d) au travers du back-azimut des ondes de Love et de Rayleigh se propageant dans le réseau. Elle observe cependant des diffractions également depuis les autres bords du bassin, mais en plus faibles proportions. D'après nos résultats, le back-azimut des séismes influence fortement la variabilité de la réponse des sites dans le bassin d'Argostoli. L'émergence d'ondes de surface par diffraction sur le bord sud-ouest du bassin pourrait être plus importante lorsque les séismes sont localisés à l'ouest (Figure 39 (c3) et (c4) : courbes rouge et verte) pour ce qui concerne des modes fondamentaux des ondes de Love et de Rayleigh (1,5 et 3-4 Hz). Pour les harmoniques supérieures (>6 Hz), la diffraction semble plus forte lorsque les séismes proviennent du sud (courbes verte et marron). La géométrie 3D asymétrique du bassin parait donc bien appropriée pour expliquer la grande dépendance des résultats SSR au back-azimut des sources.

3.3.2. Influence de la distance épicentrale

Outre le back-azimut, la distance épicentrale pourrait également jouer un rôle notable dans la façon dont le bassin est éclairé et ce particulièrement à courte distance épicentrale où l'angle d'incidence de l'onde à l'arrivée dans le bassin est plus petit. Un « trade-off » entre R_e et bAz ne peut pas être exclu dès lors que la distribution du bAz des évènements n'est pas homogène et varie selon Re. Nous étudierons donc l'influence de R_e pour des intervalles de bAz fixés.

La Figure 41 montre les variations des résultats SSR pour des séismes émanant de l'embouchure du bassin et regroupé selon quatre intervalles de R_e. Les évènements situés à très courte distance du bassin (R_e<10 km) induisent des réponses de site plus marquées (Figure 41 (c3) et (c4)) : avec une amplification plus forte à basse fréquence (<4 Hz) et plus faible à haute fréquence (>5 Hz). Ces tendances sont celles qui avaient été observées Figure 39 pour les séismes émanant de cette direction (courbe rouge). L'effet qu'a l'éclairage du back-azimut sur la réponse du bassin semble donc d'autant plus grand que les évènements sont proches. Intuitivement, il apparait envisageable que les évènements proches qui ont un angle d'incidence plus petit et un parcours plus direct entre la source et le bassin que les évènements lointains soient ceux qui marquent le plus l'effet « d'éclairage » du bassin par le back-azimut. Il est à noter également que pour les évènements dont l'épicentre se situe à seulement 5 km des sites, le critère généralement admis $Re > 10\Delta d$ (où Δd est la distance entre le site et la référence) est tout juste respecté. Cela pourrait induire des différences dans les termes de source et de propagation entre les deux sites de mesure, cependant aucun biais de ce type ne semble pouvoir expliquer les observations faites ici.

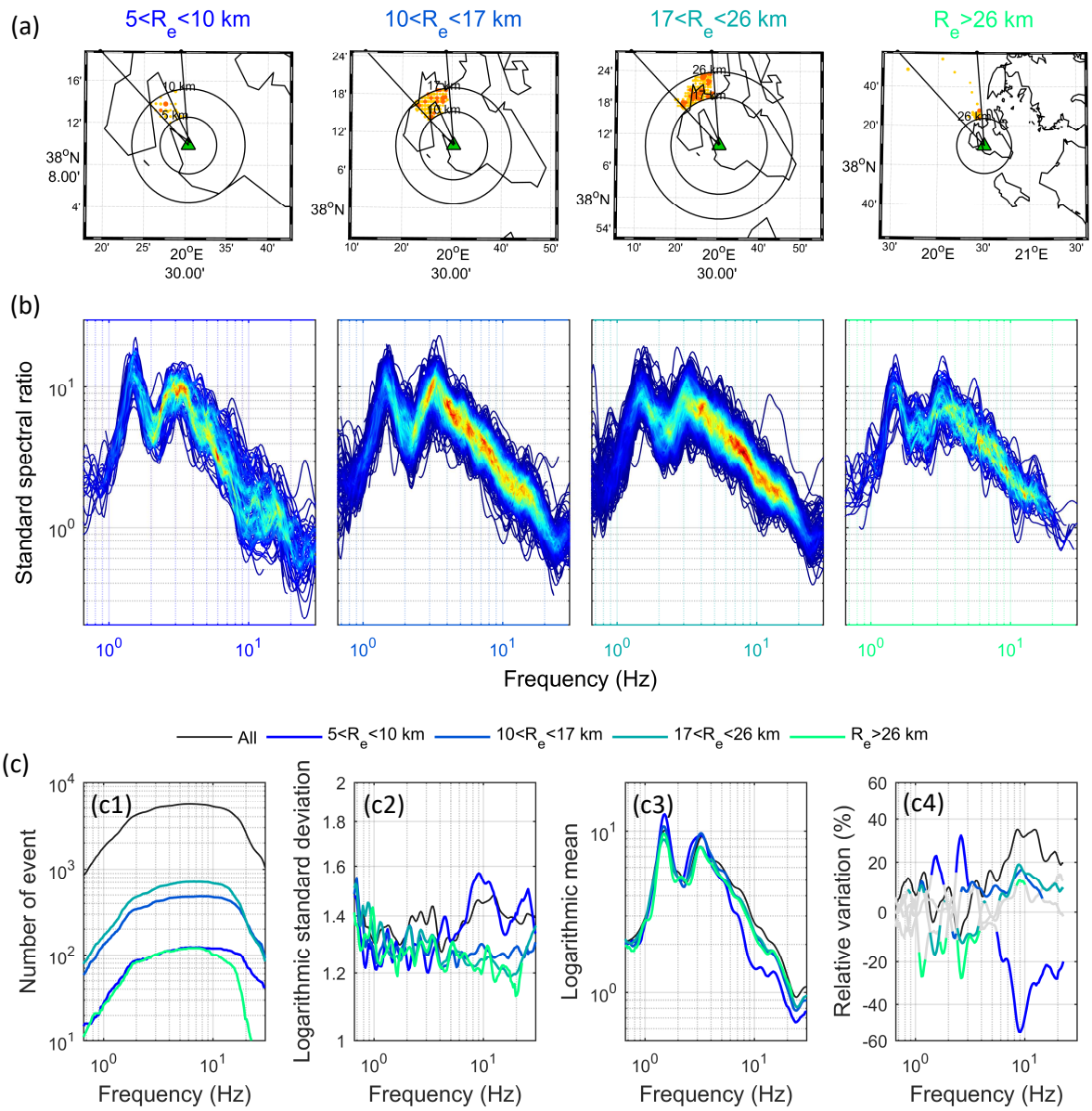


Figure 41. Influence de la distance épicentrale (R_e) sur la variabilité des résultats SSR entre les sites A2 et A1 à Argostoli. R_e est testé pour un intervalle de back-azimut fixé (N312°E-N352°E). Les différents panneaux sont identiques à ceux décrits sur la Figure 39.

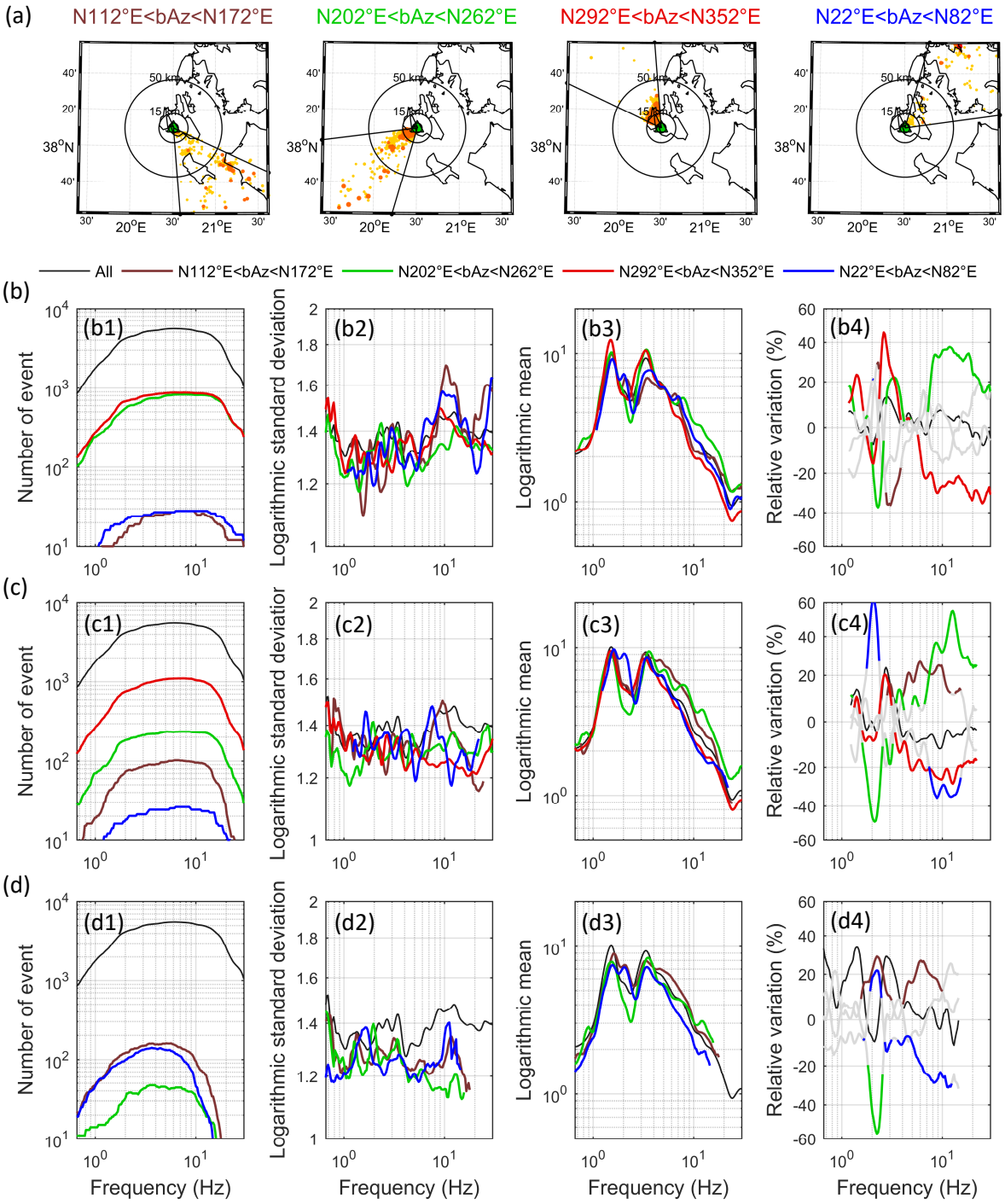


Figure 42. Influence conjuguée du back-azimut (bAz) et de la distance épicentrale (R_e) sur la variabilité des résultats SSR entre les sites A2 et A1 à Argostoli. Les quatre intervalles de bAz utilisés Figure 39 sont recoupés ici selon trois intervalle de R_e : (b) 5-15 km, (c) 15-50 km et (d) >50 km. Les panneaux sont identiques à ceux décrits sur la Figure 39.

Au-delà d'une certaine distance, on pourrait s'attendre à ce que bAz n'influence plus les résultats SSR du fait de l'incidence subverticale des ondes à l'arrivée dans le bassin, ainsi que de l'importance croissante du scattering crustal qui redistribue l'énergie incidente selon une plus grande variété d'azimuts. La Figure 41 semble aller dans ce sens comme le montrent les variations moins marquées entre les courbes correspondant aux intervalles les plus distants (R_e >10 km). La Figure 42 représente l'évolution des résultats SSR en fonction de bAz pour trois intervalles de Re. S'il est évident qu'au niveau des pics à 1,5 et 3-4 Hz l'amplitude du SSR et sa dépendance à bAz diminue fortement lorsque Re augmente, ceci n'est pas clairement visible aux autres fréquences. Par exemple, vers 2,1 Hz (entre les deux pics) l'écart entre les SSR moyens évalués depuis les évènements localisés du côté sud-ouest (vert) et nord-est (bleue) augmente même fortement pour des valeurs croissantes de Re. Les écarts-type diminuent nettement lorsque R_e augmente, probablement grâce à une plus grande similarité des trajets et de l'incidence des ondes à l'arrivée dans le bassin pour les évènements distants, et/ou à l'importance croissante du scattering crustal qui annihile la signature de l'éclairage lié à la source. De plus, les incertitudes de localisation des évènements (y compris en profondeur) par les catalogues sismiques induisent une variabilité dans le positionnement des sources qui est d'autant plus impactante que R_e est petit. Les résultats SSR à Argostoli ne nous permettent finalement pas de discriminer la distance au-delà de laquelle le bAz n'a plus d'influence.

3.3.3. Influence de la profondeur

La profondeur du foyer est susceptible d'influencer l'incidence des ondes à leur arrivée dans le bassin et ainsi peut, similairement à R_e, modifier l'éclairage de ce dernier. La Figure 44 montre les variations des résultats SSR pour des séismes proches de l'embouchure du bassin et repartis selon trois intervalles de Z.

Les observations ne mettent pas clairement en évidence un quelconque effet de la profondeur sur la moyenne (Figure 44 (c3) et (c4)). L'amplitude du second pic (à 3-4 Hz) est légèrement plus forte lorsque les séismes sont très proches de la surface sans qu'une relation de causalité ne soit clairement établie pour autant. Tous les séismes utilisés pour cette étude sont crustaux (Z <30 km) ce qui limite grandement l'intervalle de variation de ce paramètre. De plus, la grande incertitude généralement associée à la détermination de la profondeur dans les catalogues sismiques participe à l'impossibilité de conclure quant à son impact sur les résultats SSR.

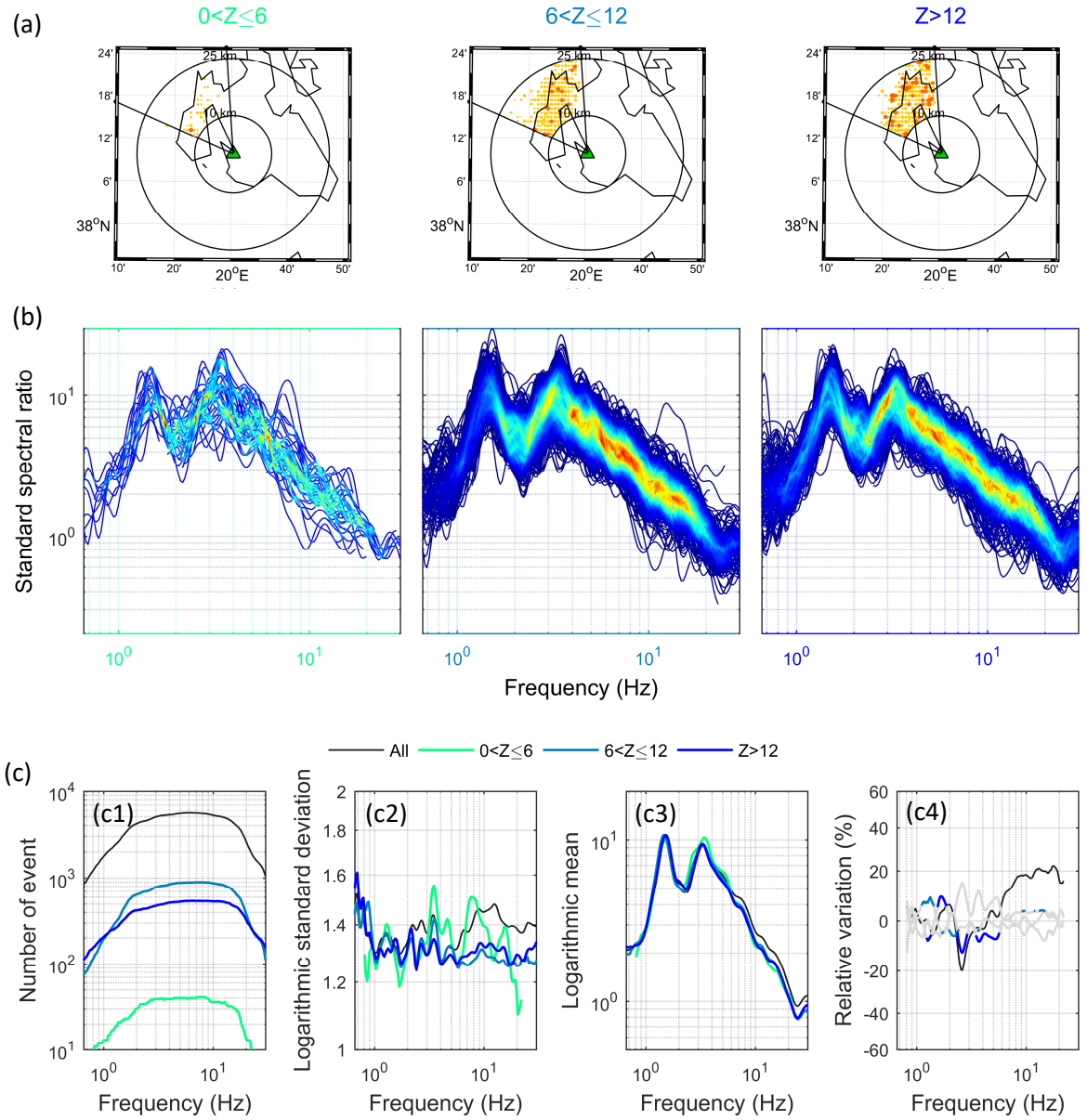


Figure 43. Influence de la profondeur (Z) sur la variabilité des résultats SSR entre les sites A2 et A1 à Argostoli. Z est testé pour un intervalle de back-azimut (N292°E-N352°E) et de distance épicentrale (10-25 km) fixée. Les différents panneaux sont identiques à ceux décrits sur la Figure 39.

3.3.4. Influence de la magnitude

L'impact de la magnitude sur la réponse du site A2 à Argostoli semble négligeable sur la Figure 44. Pour cette étude, les bases de données utilisées sont exclusivement constituées d'enregistrements de mouvements faibles (PGA < 0,1 g - Figure 29). Dans le cas de mouvements forts, il est bien connu que certains sols peuvent présenter des comportements non-linéaires dont la réponse dépend alors de la sollicitation imposée. Aucun effet non-

linéaire n'est observé ici et la réponse linéaire des sols apparait clairement indépendante de la magnitude de la sollicitation.

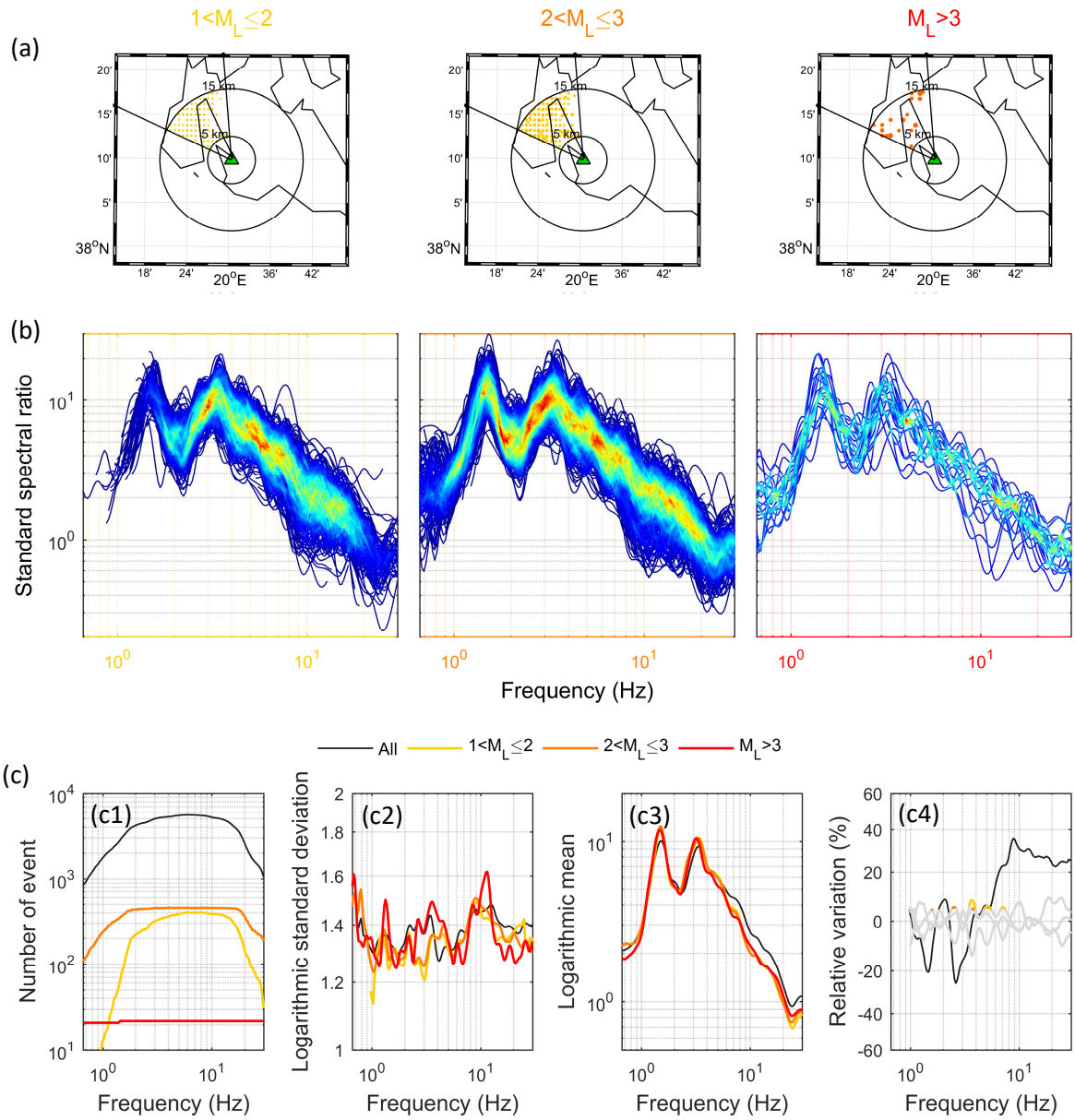


Figure 44. Influence de la magnitude locale (M_L) sur la variabilité des résultats SSR entre les sites A2 et A1 à Argostoli. M_L est testé pour un intervalle de back-azimut (N292°E-N352°E) et de distance épicentrale (5-15 km) fixée. Les différents panneaux sont identiques à ceux décrits sur la Figure 39.

3.4. Discussion sur les incertitudes

Les incertitudes associées aux estimations d'effet de site sont généralement assez grandes. C'est un fait notoire aujourd'hui que la réponse des sites est très dépendante de la complexité des structures géologiques. Pour un même évènement sismique, une très grande variabilité de réponse des sites est observée en fonction de la position géographique du site et de la fréquence (variabilité site-à-site). Nos résultats pour les sites P4 et P6 de Provence illustrent parfaitement cette variabilité spatiale et fréquentielle de l'effet de site. L'écart d'amplification observé entre ces deux sites est d'un facteur deux environ alors même qu'ils sont localisés à une cinquantaine de mètres seulement l'un de l'autre dans le bassin. Cette variabilité a été observée par de nombreux auteurs lors d'évènements de grande importance comme par exemple celui de Northridge de 1994 (p. ex., Davis et al. 2000 ; Boore et al. 2003). Au-delà des paramètres de site traditionnellement utilisés par les GMPEs (V_{S30}, κ_0), la variabilité site-à-site est également prise en compte depuis plus d'une dizaine d'années de façon statistique au travers du single station sigma (ex : Atkinson 2006). De même, cette variabilité spatiale justifie la réalisation de microzonages (ex. : Duval et al. 2013) pour les études spécifiques dans des zones d'intérêts particuliers (ville, bassin, bâtiment, etc.).

Au-delà de cette variabilité site-à-site qui est très importante, il existe également une variabilité manifeste dans la réponse d'un site à des sollicitations diverses (variabilité inter-évènements). Depuis plus de 30 ans, de nombreux travaux ont mesuré cette variabilité inter-évènements mais peu proposent de l'étudier en détail. Jarpe et al. (1988) observent une forte variabilité des SSR en Californie, qu'ils soient évalués depuis des enregistrements de mouvements forts ou de mouvements faibles. Tout comme nous, ils remarquent que les résultats SSR obtenus depuis des évènements régionaux montrent une variabilité plus faible que ceux tirés des évènements locaux, alors même que les valeurs moyennes sont très similaires. À l'inverse, Yu and Haines (2003) observent une variabilité assez réduite ($CV \approx$ 10-20 %) pour les SSR évalués à partir de 34 évènements enregistrés dans le bassin de Parkway en Nouvelle-Zélande. Thompson et al. (2012) proposent une taxonomie des réponses de sites basée sur l'étude d'une centaine de sites du réseau KiK-net au Japon. Ils obtiennent, notamment des incertitudes (moyenne quadratique des résidus) inter-évènements qui varie de 10 à 60 % suivant les sites (Figure 45). Leurs études montrent ainsi que la variabilité inter-évènements de la réponse des sites est très dépendante du site d'étude.

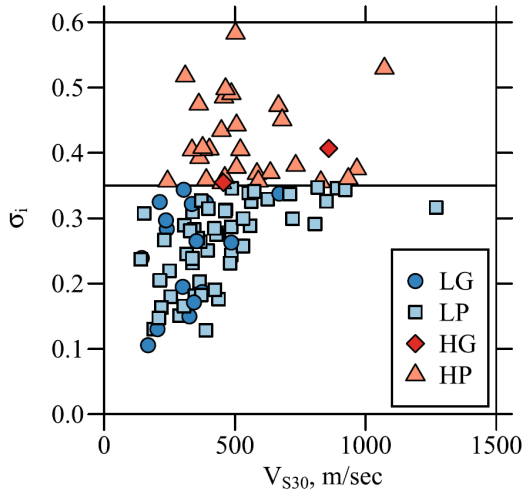


Figure 45. Incertitude inter-évènements (σ_i) en fonction de la V_{S30} du site KiK-net (depuis Thompson et al. 2012). Les symboles correspondent à la classification des sites proposée par les auteurs.

Ktenidou et al. (2016) obtiennent, dans le bassin du site expérimental de CORSSA en Grèce, des SSR très dépendants de l'orientation des composantes avec des variations de l'ordre de 100 % à la fréquence fondamentale du site et un facteur d'amplification maximal suivant la direction longitudinale du bassin. Comme nous, il apparait dans leurs résultats que les directions longitudinales et transversales retranscrivent bien la forte variabilité de la réponse des sites à l'orientation des composantes. Cette dernière semble intrinsèquement liée aux asymétries des géométries 2D/3D des structures géologiques des bassins. De nombreux travaux sur les effets de site ont été réalisés sur le site expérimental « Euroseistest ». Il s'agit d'une vallée large de 2,5 km proche de Thessaloniki en Grèce. Plusieurs stations s'alignent sur un profil à travers la vallée tandis qu'un forage est localisé en son centre et traverse les couches sédimentaires jusqu'au substratum rocheux situé à 200 m de profondeur environ. Ce dernier a permis la mesure *in situ* du profil de vitesse, ainsi que la mise en place de 6 capteurs de puits à différentes profondeurs. Un premier résultat de Raptakis et al. (1998) confirme précocement la très forte variabilité inter-évènements associée à ce site d'étude avec un coefficient de variation de l'ordre de 100 à 200 % suivant les fréquences. Récemment, Maufroy et al. (2017) y ont réalisé une étude comparant les résultats issus des SSR, des HVSRs et des simulations numériques. La simulation de 1260 évènements présentant une distribution spatiale homogène autour du site d'étude montre une forte variabilité des résultats SSR simulés qui est en accord avec les résultats expérimentaux. D'après les auteurs, cette variabilité trouve son origine dans les paramètres liés à la source (back-azimut, distance, profondeur) qui font varier l'éclairage donné au bassin et plus particulièrement qui régule la génération d'ondes de surface diffractées sur le bord nord faiblement penté de celui-ci. Ainsi, les évènements situés du côté nord sont plus fortement diffractés ce qui tend à étaler les pics, à l'inverse des évènements provenant du sud qui maximisent les facteurs d'amplification au niveau des pics. Leurs analyses montrent également que cet effet du back-azimut est plus fort lorsque l'angle d'incidence des ondes à

l'arrivée dans le bassin est le plus petit, c'est-à-dire pour les évènements les plus superficiels. La profondeur apparait dans cette étude comme un paramètre fondamental lorsque les évènements sont proches des sites. Dans notre étude, nous observons également un effet prépondérant du back-azimut sur la variabilité des résultats SSR à Argostoli. Similairement à Maufroy et al. (2017), cet effet est très probablement lié aux variations dans l'éclairage du bassin suivant la position de la source sismique qui influe sur la génération d'ondes de surface par le bord faiblement penché sud-ouest de celui-ci (Imtiaz 2015). Cependant, nous faisons également plusieurs observations contradictoires : les pics sont plus marqués pour les séismes arrivant du côté sud-ouest pourtant notablement plus diffractant d'après Imtiaz (2015) et la profondeur ne montre aucun effet tangible même à courte distance épicentrale. En revanche, nous observons un effet de la distance épicentrale. Lorsque celle-ci diminue, la variabilité des résultats SSR ainsi que les effets associés au back-azimut augmentent sensiblement. Ceci est probablement dû à des propagations d'ondes qui sont plus directes depuis la source jusqu'au site (moins de scattering) et à une diminution de l'angle d'incidence des ondes à leur arrivée dans le bassin.

Nos résultats, ainsi que les quelques travaux cités précédemment (dont la liste n'est pas exhaustive), illustrent bien la dépendance de l'incertitude inter-évènements à la complexité géologique du site au même titre que la réponse moyenne du site elle-même. Lorsque les structures géologiques présentent des géométries 2D ou 3D, des phénomènes complexes interviennent (réflexion latérale, effet de focalisation et défocalisation, génération d'ondes de surface sur les bords du bassin, etc.) qui sont sensibles à l'orientation et à l'incidence des ondes constituant l'excitation sismique. Ces phénomènes sont difficiles à interpréter comme l'illustrent les quelques contradictions relevées entre nos observations et les simulations réalisées par Maufroy et al. (2017). Il n'en reste pas moins que les effets de site sont clairement dépendants à la fois de la géométrie des structures géologiques sous-jacentes et des sources utilisées pour éclairer ces dernières. La réponse moyenne des sites et les incertitudes qui lui sont associées sont donc à la fois dépendantes du site d'étude et du jeu de données utilisé pour le caractériser. Ceci a des conséquences fortes, puisque si la distribution spatiale des séismes utilisés dans les études d'effets de site n'est pas homogène, alors la réponse moyenne des sites présentant des géométries 2D ou 3D peut être biaisée et l'incertitude associée sous-estimée. Par exemple, le site P2 en Provence montre des incertitudes inter-évènements du même niveau que ceux situés dans le bassin (Figure 33) alors même que ce site est un rocher dur ne présentant pas d'effets topographiques notables. Les incertitudes pour les sites situés dans le bassin sont donc très probablement sous-estimées du fait de la très grande concentration des évènements au nord-est du bassin qui unifie l'éclairage qui est fait de ce dernier. De même, l'écart important d'incertitude entre le site situé dans le bassin d'Argostoli ($CV \approx$ 30-50 %) et ceux localisés dans le bassin de Provence ($CV \approx$ 15-25 %) peut s'expliquer en partie par les différences de distribution spatiale des évènements disponibles pour évaluer le SSR entre ces deux sites d'étude. La Figure 46 montre la distribution des résultats SSR pour un essaim d'évènements distants (>100 km), similairement pour le site P3 en Provence et A2 à Argostoli. Une réduction très importante de l'incertitude est observée à Argostoli ($CV \approx$ 20 %), tandis que sur le site P3 de Provence cette dernière est très stable ($CV \approx$ 20 %). Les niveaux

d'incertitude obtenus depuis ces deux essaims sismiques sont très proches. De plus, il apparait difficile d'assurer une plus grande similarité de chemins et d'éclairage entre évènements que celle qui est proposée Figure 46. Il nous semble donc que cette variabilité de 20 % peut être considérée comme « aléatoire » pour les deux sites. Les écarts avec les incertitudes relevées depuis l'ensemble des évènements sont donc très probablement épistémiques.

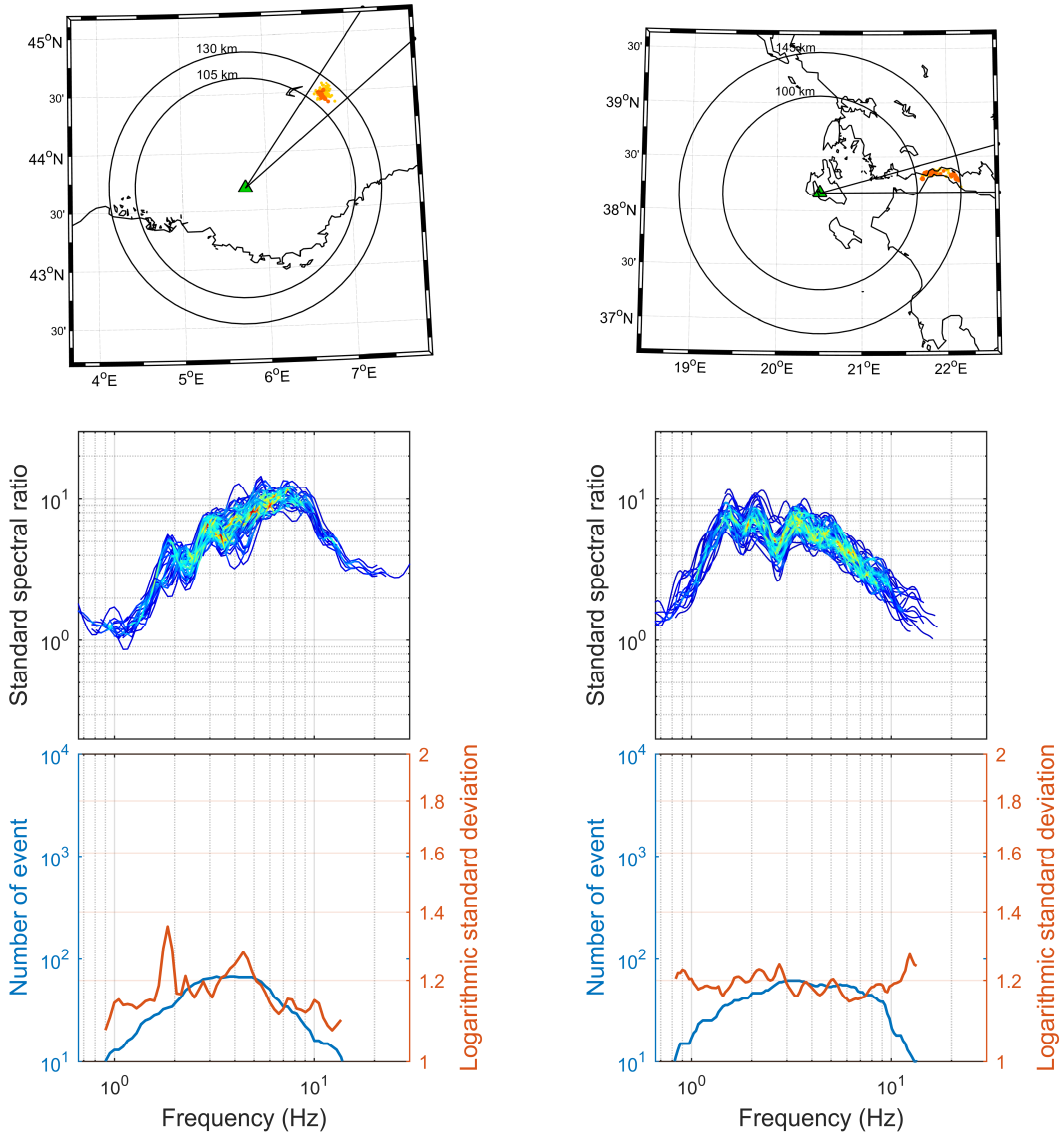


Figure 46. Évaluation des SSR au site P3 en Provence (gauche) et A2 à Argostoli (droit) depuis des évènements concentrés dans un essaim sismique distant (cartes du haut). L'échelle de couleur indique la densité de point des distributions SSR (panneaux du milieu). Le nombre d'évènements sismiques utilisé ainsi que l'écart-type logarithmique en fonction de la fréquence sont également indiqués (panneaux du bas).

Cette distinction entre incertitude aléatoire et épistémique est importante, car la première, qui est probablement du même ordre de grandeur pour tous les sites (~20 %), constitue une limite basse pour l'incertitude, tandis que la seconde explique les fortes variations d'incertitude inter-évènements observées d'un site à l'autre. Cette dernière peut être entièrement retranchée en tenant compte de la position des sources dans les évaluations d'effets de site 2D ou 3D. Ceci est d'un intérêt premier pour la prédiction des mouvements du sol (GMPEs) et pour les études de l'aléa sismique plus généralement, pour lesquelles les incertitudes associées aux sites, qui sont très grandes, pourraient être réduites. À l'inverse, on peut se demander si les études d'effets de site caractérisent bien l'incertitude totale associée au phénomène. Pour ce faire, il conviendrait de vérifier que l'incertitude épistémique est complète en s'assurant de la bonne distribution spatiale des sources. En pratique, une telle configuration est difficile à réaliser du fait des distributions spatiales généralement très hétérogènes des séismes qui s'organise le long des failles. Un séisme survenant sur une faille jusque-là peu active pourrait alors conduire à des sollicitations très différentes (de 0 à plus de 100 % suivant les configurations) de celles qui avaient été prédites. Il conviendrait donc à minima d'associer aux évaluations d'effets de site, de l'information sur la distribution spatiale des évènements utilisés, sans quoi les enseignements fournis par celles-ci seront immanquablement incomplets (p. ex. : Figure 45). Idéalement l'incertitude totale pourrait être estimée par méthode numérique similairement au travail de Maufroy et al. (2017), ou bien extrapolée à partir des dépendances au back-azimut, à la distance épicentrale et à la profondeur relevée depuis les quelques mesures SSR disponibles. Celle-ci pourrait également être déduite de la géométrie des structures géologiques, de la dépendance des SSR à l'orientation des composantes ou des mesures HVSRs (cf. section 1.1.2.5).

3.5. Conclusions

L'évaluation empirique des effets de site est une composante essentielle des études d'aléa sismique. Tandis qu'un nombre considérable d'études a été réalisé sur le sujet, un nombre plus limité d'entre elles s'est intéressé aux incertitudes qui leurs sont associées. La variabilité spatiale de la réponse des sites est généralement très importante même pour des sites proches géographiquement. Pour un site donné, la dispersion des résultats SSR autour de la moyenne est généralement assez importante ($CV \approx$ 20-100 %). Cette incertitude inter-évènements se décompose en une part aléatoire et une part épistémique. Comprendre cette dernière est essentiel pour pouvoir anticiper la quantité et les caractéristiques des évènements qui sont requis pour réaliser une évaluation fiable de la réponse des sites. En effet, l'obtention d'une base de données sismiques représentative est difficile, et ce, particulièrement dans les zones à sismicité faible ou modérée où la quantité d'enregistrements de bonne qualité qui peut être réalisée en un temps court est généralement assez limitée.

Nous bénéficions ici de bases de données de mouvements faibles très riches sur deux sites d'étude : l'un en Provence (France) avec 5 sites d'enregistrement et jusqu'à plus de 300

évènements enregistrés suivant le site et la fréquence considérée ; l'autre, près de la ville d'Argostoli (Grèce) avec deux sites d'enregistrement et plus de 6000 enregistrements disponibles. La première étape consiste à identifier la loi de distribution des résultats SSR. Nos résultats confirment l'hypothèse généralement admise d'une loi de distribution log-normale. Maintenant que celle-ci est établie, l'intervalle de confiance ($I_{C_x\%}$), qui définit l'intervalle dans lequel une moyenne évaluée depuis n évènements tirés aléatoirement a x % de chance de se trouver, peut être utilisé pour identifier des phénomènes non-aléatoires (incertitude épistémique). Pour tous nos sites de mesure, la borne supérieure « théorique » de $I_{C_95\%}$ montre un très bon accord avec le percentile 2,5 et 97,5 % des 500 SSR moyens évalués à partir de 500 tirages aléatoires de n évènements, n variant de 2 à 1000 (bootstrapping). Cela confirme à la fois la distribution log-normale de nos résultats SSR et le potentiel de l'intervalle de confiance pour identifier les paramètres qui contrôlent l'incertitude épistémique. En effet, tout dépassement de $I_{C_99\%}$ par un SSR moyen évalué depuis n évènements a moins de 1 % de chance d'être lié au hasard. Ainsi, en comparant la moyenne obtenue depuis n évènements sélectionnés spécifiquement suivant différents paramètres de source (back-azimut, distance épicentrale, profondeur ou magnitude) avec $I_{C_99\%}$, il est possible de déterminer quels paramètres contrôlent la variabilité épistémique inter-évènements de la réponse d'un site.

Nos résultats à Argostoli montrent très clairement que l'incertitude inter-évènements est principalement contrôlée par le back-azimut de la source qui modifie la façon dont le bassin est éclairé. Ceci retranscrit la géométrie 3D du bassin et plus particulièrement la sensibilité de son bord sud-ouest faiblement incliné à l'éclairage qui en est fait. Celui-ci a en effet déjà montré sa capacité à générer des ondes de surface qui dominent le champ d'ondes dans le bassin (Imtiaz 2015). L'effet du back-azimut apparait d'autant plus marqué que la distance épicentrale est courte, probablement en lien avec un trajet plus direct des ondes (moins de scattering) et un angle d'incidence à l'arrivée dans le bassin qui est plus petit. La profondeur et la magnitude de la source n'ont, en revanche, pas montré d'effet particulier sur les résultats SSR. De plus, il apparait également que la réponse des sites varie fortement suivant l'orientation des composantes. Les facteurs d'amplification sont alternativement maximum dans la direction longitudinale et transversale. Cela semble retranscrire les modes de résonance du bassin. Finalement, le critère généralement admis SNR >3 semble être justifié au vu de nos résultats.

La part d'incertitude qui est réellement aléatoire apparait finalement assez faible (de l'ordre de 20 %). À l'inverse, l'incertitude épistémique, qui est dépendante de la géométrie des sites, est généralement très importante (jusqu'à plus de 100 %) et explique les fortes variations d'incertitude inter-évènements qui sont observées d'un site d'étude à l'autre. Ainsi, l'incertitude totale est spécifique à chaque site au même titre que la réponse moyenne du site elle-même. De plus, l'incertitude épistémique est très dépendante de la couverture spatiale des séismes qui sont utilisés pour la caractériser. Ainsi, celle-ci est très probablement sous-estimée dans les études pour lesquelles la distribution spatiale des évènements est incomplète comme c'est le cas pour les bases de données de Provence. Les écarts de variabilité observés entre le site de Provence et d'Argostoli, ainsi que ceux

constatés dans la littérature, sont donc très probablement aussi bien dus à des différences de géométrie des sites qu'à des disparités dans les bases de données utilisées. Finalement, si 8 évènements suffisent à avoir une bonne représentativité statistique (en accord avec la distribution log-normale) pour les sites les plus simples à géométrie 1D, le nombre d'évènements disponibles importe moins que l'homogénéité de leurs distributions spatiales pour les sites à géométrie 2D ou 3D. Cette bonne couverture spatiale est, en pratique, difficile à obtenir et ce particulièrement pour les sites soumis à une sismicité faible à modérée. Des méthodes doivent donc être envisagées dans l'avenir pour pallier à cette difficulté. De cette façon, l'incertitude totale réelle pourra être estimée et une réduction importante de cette dernière sera possible en tenant compte de la position des sources dans les évaluations de la réponse du site. En effet, la réduction des incertitudes épistémiques est une priorité qui s'intègre dans une démarche générale d'amélioration des prédictions des mouvements de sols entamée depuis plus d'une dizaine d'années.

Chapitre 4

Évaluation de la Réponse des Sites à Partir du Bruit Ambiant

4.1. Contexte

Comme nous l'avons vu dans le chapitre précédent, la méthode SSR est la méthode de référence pour mesurer la fonction de transfert relative des sites. Cette approche nécessite cependant l'enregistrement de nombreux séismes, ce qui en pratique impose une instrumentation des sites pendant de longues périodes de temps, allant même jusqu'à plusieurs années lorsque l'activité sismique est faible à modérée. À l'inverse, le champ sismique ambiant ou bruit ambiant est présent partout et tout le temps. Celui-ci permet d'envisager l'évaluation de la fonction de transfert des sites sur des échelles de temps beaucoup plus courtes, de l'ordre de l'heure à la semaine suivant la méthode utilisée et la bande de fréquence étudiée.

Ce chapitre est une version étendue qui réunit deux articles à soumettre. L'objectif principal de ce travail est de vérifier expérimentalement le potentiel des approches utilisant le bruit ambiant. Celles-ci sont d'une importance majeure pour la mesure de la réponse des sites dans les régions faiblement sismiques et pour l'évaluation rapide de la variabilité spatiale des effets de site à l'échelle d'une infrastructure ou d'une ville (microzonage).

4.2. Abstract

Since the 1980s a great number of studies have proposed and tested various approaches to assess site effects from the ambient noise. While the earthquake-based approaches such as

the standard spectral ratio (SSR) require good quality records of numerous earthquakes, the use of ambient noise appears as a suitable possible solution for sites located in low-to-moderate seismicity areas and for microzonation applications. The noise horizontal-to-vertical approach (HVSR) exploits the relative amplitude of horizontal and vertical components of the ambient noise. Conversely, the noise standard spectral ratio (SSRn) approach estimates the amplification of seismic noise between a site and a reference station. While the HVSR is currently used in many site-specific studies, the SSRn approach has been gradually abandoned since the 1990ies. More recently, since the beginning of the 2000ies, studies have also taken advantage of the noise phase information through the cross-correlation analysis (coherency and the ambient noise impulse response function, ANIRF) to obtain information on the propagating wavefield between two sensors, possibly giving information in some cases on the ground motion local amplification due to superficial geological structures. In this study, we apply the aforementioned methods on two sites where both earthquakes and seismic noise were recorded. We aim at empirically testing these different approaches and comparing them with information obtained from the classical SSR.

After a review of these approaches, we present datasets acquired for two sedimentary basins, the first in Provence and the second in Argostoli. In agreement with the literature, we find that the HVSR does not provide more than the fundamental resonance frequency of the site (f_0) and the SSRn leads to an overestimation of the amplification factors above f_0. The discrepancy between SSR and SSRn appears to be mainly dependent on the local geological configuration. We thus introduce the hybrid standard spectral ratio approach (SSRh) that allows to improving the SSRn by adding an intermediate station inside the basin for which the SSR is known, this station being used in turn a local reference inside the basin for the SSRn computation. This approach allows to obtain similar results than the SSR ones up to high frequency (>10 Hz). We apply the ANIRF and cross-correlation techniques between pairs of stations and show that wave amplification for stations inside the basin appears clearly with the ANIRF. Applying the SSR technique on these ANIRF reconstructed waveforms approximate well the SSR curve at low frequency (<4 Hz). Based on that, we consider that the SSRn (or SSRh) should receive a renewed attention and benefit of the recent concepts developed in the framework of cross-correlations analysis, while the ANIRF might be extensively used in the future for site effect assessment purpose in complex geological configurations even if further investigations are required to increase it frequency range of validity.

Keywords: Microtremor, ambient noise processing technique, site effect, spectral ratio, cross-correlation

4.3. Introduction

Since the first understanding of modifications of the seismic signal by the geological conditions of the earth's surface (e.g. Milne 1908), it has been widely demonstrated that site effect can dramatically increase both amplitude and duration of the ground motion. It is a

source of particular concern for seismic hazard assessment since site effect increases greatly the level of hazard in many cities around the world located on sedimentary basins (e.g., Mexico City, Mexico; Los Angeles, USA; Tokyo, Japan; Grenoble, France; etc.). For example, ground motions in the basin of Mexico City have shown significantly higher amplitudes and duration than the surrounding rock sites during the Guerrero-Michoacan event of 1985 (Celebi et al. 1987). After this devastating earthquake the estimation of the site response to an earthquake solicitation has received a renewed attention. Since then an exploding number of studies have tried to assess the site effect from what is sometimes called the “ambient seismic field” and is improperly referred to here as the “ambient noise” for a question of succinctness. The advantage of using the ambient noise has been, however, long since recognized (e.g., Ishimoto 1937). Indeed, the ambient noise can be recorded everywhere at any time making its use fast, easy and inexpensive when compared to earthquake data processing. Moreover, it is of major interest to infer the spatial variability of seismic site responses (i.e., microzonation).

The site response assessment is essential for the prediction of the ground motion at a specific site (cities, facilities, dams, nuclear power plants, etc.). The ground motion (U) at a given position is the convolution of a source (S), path (P) and site effect component (H):

$$U(w) = S(w) \times P(w) \times H(w) \qquad 31$$

where w is the angular frequency. The site transfer function, $H(w)$, gives the amplification factor as a function of frequency. The most common and reliable way to assess it is the standard spectral ratio technique (SSR – Borcherdt 1970 – e.g., Field and Jacob 1995; Bonilla et al. 1997). This approach relies on the simultaneous record of the same earthquake at least at two stations close enough with respect to the earthquake distance. The station located on the best geological conditions (i.e. with the minimum site amplification level, generally a rock site) is used as the reference. The transfer functions are obtained by normalizing the Fourier amplitude spectrum (FAS) of each component recorded at the sites by the ones recorded at the reference. The SSR assumes that the incident signal at depth is the same for the site and the reference. It is commonly admitted that this assumption is true if the hypocentral distance is much larger than the site-reference distances. Thus, the normalization allows removing the source, and path components of the signal and can be summarized as follows:

$$\mathrm{SSR}_{i/r}(\mathrm{w}) = \langle \frac{\{|U_i(w)|\}}{\{|U_r(w)|\}} \rangle_{log} \qquad 32$$

where $\mathrm{SSR}_{i/r}$ is the SSR between a site i and the reference site r. For each component, $|U_i(w)|$ and $|U_r(w)|$ are the FAS evaluated from the earthquake signals at station i and r respectively. The curly brackets $\{\cdot\}$ represent spectral smoothing and $\langle\cdot\rangle_{log}$ stand for the geometric (logarithmic) mean. However this technique requires many earthquakes to be

recorded at both stations with a good signal-to-noise ratio (SNR) in a wide frequency band to be fully usable (e.g. Perron et al. 2015). Such good quality dataset is not always available depending on the recording duration, on the seismotectonic environment and on the local level of noise perturbation. For sites where such an approach is not possible, the use of ambient noise appears as a promising alternative.

The use of the ambient noise for site response assessment presents two major limitations since the noise sources locations and wave types contents are generally unknown and differ from earthquakes, so that: (1) the response of sediments to ambient noise generated and propagating mostly at the surface has to be related in some way to that for the incident seismic waves from earthquakes; (2) the source and path component of the ambient noise has to be removed in order to assess the site influence. On the first limitation, several authors (e.g. Bonnefoy-Claudet et al. 2006) showed that ambient noises are mainly dominated by surface waves since the sources are almost entirely located at the earth surface (atmospheric phenomenon, human activities, etc.). On the contrary, the earthquake signal is mainly composed by body waves in the strong-phase motion. On the second limitation, different approaches have been developed mostly in the last three decades to extract some information related to site response from ambient noise recordings. The most usual one is the horizontal-to-vertical spectral ratio (HVSR), which is a practical tool to assess the spatial variability of the fundament resonance frequency (f_0) of the site. The ambient noise standard spectral ratio (SSRn) attempt to assess the site transfer function by normalizing the noise recording at a site by the one at the rock reference station simultaneously, in a similar way than the SSR approach from earthquakes recording. More recently, methods based on the spatial correlation of the ambient noise have been used to simulate the seismic wave propagation from one receiver to another one. Both cross-correlation and ambient noise impulse response function (ANIRF – or interferometry by deconvolution) approaches have been applied. A short review of those methods is given in the next section with a special attention for the SSRn and the ANIRF approaches. For more detailed reviews, the reader is invited to refer to the following studies: Field et al. (1990) gives a good overview of the preliminary approach used to characterize the site response from the ambient noise including the first work on the SSRn; In the same way Lermo and Chávez-García (1994), Kudo (1995) and Bard (1999) provide comprehensive review of the first development and application of the SSRn and HVSR approaches; The literature on the HVSR is also given in detail by Bonnefoy-Claudet et al. (2006); More recently, Prieto et al. (2011) explored the latest application of the ambient noise for site effect evaluation through the ANIRF approach.

While some of these approaches have demonstrated their incapacity to assess the site transfer function at any frequency, others have been employed primarily in the low frequency range (<1 Hz) where the ambient noise is dominated by coherent Rayleigh wave that are generated by known and generally distant oceanic sources (microseisms). At higher frequency (>1 Hz) the ambient noise is more complex (mainly Love and Rayleigh waves) and is dominated by human activities (microtremors). Whatsoever, site effects are rarely negligible above 1 Hz even for deep basin where higher resonance modes and shallow

velocity contrast may notably amplify the ground motion. Evaluating $H(w)$ from microtremors is of major interest since it is usually the frequency range that is targeted in engineering applications (1–10 Hz). Moreover, geological and geophysical models used in numerical simulations have not had yet sufficient resolution to predict ground motion up to high frequency making generally recordings the only available way to assess the site response in this frequency range. Because the frequency range we are working on is predominately higher than 1 Hz the terms “microtremors” and “ambient noise” are used indistinctly in this study.

The aim of the present work is to experimentally estimate the ability of methods based on ambient noise (HVSR, SSRn and ANIRF) to assess the linear site transfer function for weak motion earthquake solicitations up to high frequency (~30 Hz). Moreover we develop a new approach, the hybrid standard spectral ratio (SSRh), to assess the spatial variability of the site response from a temporary network placed around a permanent or long-time recording station for which an SSR transfer function is already available. We take benefit from datasets acquired on two different sites: an industrial site located across a small sedimentary basin in the low-to-moderate seismicity context of Southeastern France (Provence) and a sedimentary basin located close to the city of Argostoli in the seismically active island of Cephalonia in Greece. We first compare results obtained from ambient noise spectral ratio approaches (HVSR, SSRn and SSRh) with the SSR results obtained from weak motion earthquakes. In a second time, we investigate information included in the phase and possibly amplitude of the microtremors through coherency and ANIRF application between different pairs of stations. Finally we discuss the dependence of our results to the local geological configuration, to the ambient noise source and to inter-stations distance. Throughout this study, we focus on real data and do not investigate theoretical aspects.

4.4. Methods definition and background

4.4.1. Spectrum analysis

Preliminary ambient noise analysis for site effects description consisted of qualitative comparisons between time histories and/or FAS recorded on different geological conditions. Such comparisons make the assumption that the site component dominates the source and path ones. Although this assumption is very dubious, various authors found that the predominant frequency of the FAS evaluated from the microtremors coincides with that from the earthquake and depends on the geology of the site (Ishimoto 1937; Kanai et al. 1954; Kanai and Tanaka 1961; Ohta et al. 1978; Kanai 1983; Field et al. 1995). In a different way, Udwadia and Trifunac (1973) found little to no correlation between ambient noise FAS and strong motion earthquake FAS that they attribute to non-stationarity of microtremors. Anyway, direct observations from the FAS at one station provides limited information regarding local site response since the source and path component are not removed. Therefore, more complex methods are required for quantitative evaluation.

4.4.2. HVSR

Another single station method applied to microtremors recording is the HVSR approach. It is based on the spectral ratio between the horizontal mean component and the vertical component of the microtremors recorded simultaneously at one single station

$$HVSR_i(w) = \langle \frac{\{|u_{iH}(w)|\}}{\{|u_{iV}(w)|\}} \rangle_{log} \qquad 33$$

where HVSR_i is the HVSR at site i and $|u_{iH}(w)|$ and $|u_{iV}(w)|$ are the FAS evaluated from the ambient noise at station i on the horizontal mean component and vertical component respectively. This approach was first introduced in Japan in the 1970s (e.g. Nogoshi and Igarashi 1971). At the end of the 1980s, Nakamura (1989) proposed its first versions in English leading some authors to call this technique the Nakamura's method. After, Kudo (1995) and others popularized its use. Nakamura (1989) interpreted the HVSR curves as the amplification of the site due to the S-wave resonance. However, it is currently admitted that the HVSR allows predicting only f_0 of S waves propagating in soil sites, mainly thanks to surface waves present in ambient seismic field. Indeed, the ellipticity of Rayleigh waves provokes the vanishing of the amplitude of the vertical component at that frequency, while the Airy phase frequency of Love waves corresponds to a bump of energy on both horizontal components. Both phenomena result in a peak in the H/V ratio close to f_0. The reliability of this technique to assess f_0 for simple 1D geological structure with strong impedance contrast has been demonstrated both theoretically and experimentally. Because the HVSR provides a fast, ease and low cost solution for f_0 microzonation, the number of studies based on this approach has dramatically increased in the last two decades. Thus, a comprehensive review is not possible here since HVSR applications now represent the vast majority of works using microtremors (Bonnefoy-Claudet et al. 2006). A long debate existed on the capability of the HVSR to assess the amplification factor of the site response. It led to the realization of a workshop on the topic leading to extensive HVSR implementations recommendations (SESAME team 2004 – http://sesame.geopsy.org/Delivrables/Del-D23-HV_User_Guidelines.pdf). The HVSR approach has been found not to be able to infer the resonance frequency at higher modes (above f_0) as well as the amplification factor at any frequency.

4.4.3. SSRn

In the 1980s, a decade after the first definition of the SSR approach given by Borcherdt (1970), the SSRn was introduced. It is the equivalent of the SSR (Eq. (32)) applied to ambient noise:

$$SSRn_{i/r}(w) = \langle \frac{\{|u_i(w)|\}}{\{|u_r(w)|\}} \rangle_{log} \qquad 34$$

where $SSRn_{i/r}$ is the SSRn between a site i and the reference site r. $|u_i(w)|$ and $|u_i(w)|$ are the FAS evaluated from the ambient noise at each component on station i and r respectively. One important preliminary condition for using the SSR and SSRn techniques to assess sedimentary site response is the availability of a nearby reference (rock) site whose site response can be considered as negligible. In a similar way of the minimal distance to the source required for the SSR, studies based on microseisms often rely on distant oceanic sources. At higher frequencies, the proximity of the source of microtremors is generally unknown; this assumption cannot be guaranteed as valid. These difficulties explain why the SSRn have been poorly studied and is mostly evaluated at low frequency for microseisms sources (<1 Hz).

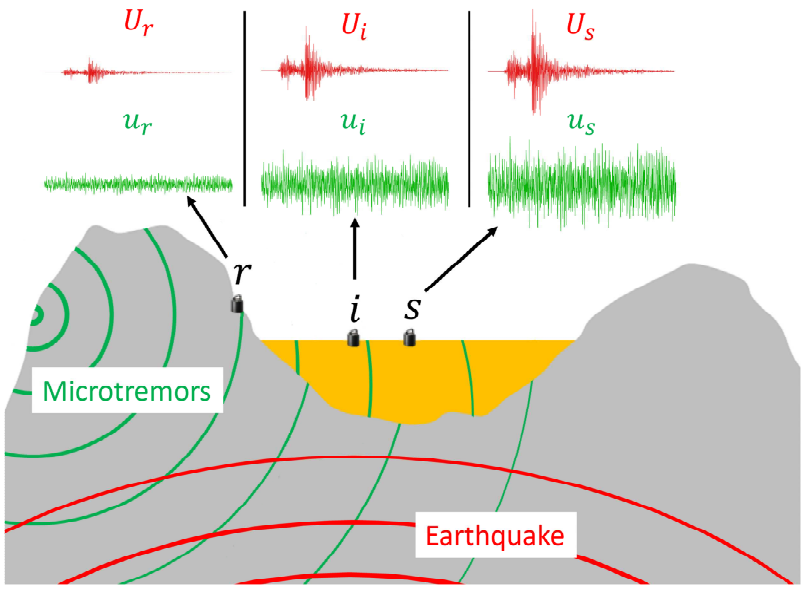
The SSRn has been tested by several authors, either to detect the presence of a fault at depth (e.g., Irikura and Kawanaka 1980) or to assess site response. The first evaluation for site effect purpose was done by Kagami et al. (1982, 1986). The authors estimated SSRn at various sedimentary locations in the period range of 1 to 10 sec. They found, in agreement with results deduced from earthquake and nuclear events, that microseisms spectral ratio amplitudes increase with increasing sediment thickness but they did not observe any systematic change in f_0. Celebi et al. (1987) applied the SSRn to the Mexico City lake-bed zone and succeeded in assessing f_0. However, the authors were not able to retrieve the amplification factors deduced from the SSR. Usually SSRn overestimated the amplification factors obtained by the SSR. They hypothesized that it was due to differences in the physics of the sources since microtremors are mostly linked to the traffic within the lake-bed zone. Field et al. (1990) gives a good review of this topic and investigate the SSRn methods in New York City. They found two amplification peaks at 0.8 Hz and 2.4 Hz that they attribute to the site effect after comparing their results with those from the theoretical one-dimensional transfer function. Again amplification factors were unrealistically high so the authors concluded that for urban environments the assumption that the SSRn is independent on the noise sources is not always strictly valid. Lermo and Chávez-García (1994) makes the comparison between four spectral ratio methods (including SSR, SSRn and HVSR) on 3 different sites. They concluded that the SSRn was not able to predict the amplitude contrary to the HVSR but their results seem to show that the SSRn was sometimes quite reliable to reproduce the general shape of the amplification function for higher modes in the frequency range 0.3-5 Hz. In a different way, Dravinski et al. (1996) investigated theoretically the potential of the HVSR and the SSRn to assess the response of a semicircular and semi-spherical valley to incident plane harmonic Rayleigh waves. They found better estimation of the resonance frequencies (f_0 and higher modes) and amplification with the SSRn than the HVSR.

Some studies made a qualitative comparison of the SSRn with the geology of the site (Okada et al. 1991; Ferritto 1995; Koyama et al. 1996; Milana et al. 1996; Haile et al. 1997; Ibs-von Seht and Wohlenberg 1999; Burjánek et al. 2012) and/or with the degree of damage after a major earthquake (Ohmachi et al. 1991; Koyama et al. 1996; Haile et al. 1997; Seo et al. 2000). While several other studies found mixed results regarding the comparison between

SSR and SSRn curves, depending mostly on the experimental conditions (type of soil, level and proximity of noise sources, distance between stations and the reference) and on processing (instrument used, criteria for the selection of the noise windows, stacking, frequency bands evaluated), some of them obtain at least a rough estimation of the SSR site transfer function using the SSRn approach, generally at low frequencies (Rovelli et al. 1991; Gutierrez and Singh 1992; Yamanaka et al. 1993; Gitterman et al. 1996; Zhao et al. 1998; Horike et al. 2001; Theodoulidis 2006). Some authors succeeded in achieving only f_0 (Gaull et al. 1995; Seekins et al. 1996; Atakan 1997), while few others did not find any convincing agreement between SSRn and SSR (Seo 1992; Field et al. 1995; Field 1996). A majority of these studies concluded that the SSRn method was reliable to predict, at least, f_0 but unable to estimate the amplification factor of the site response on the whole frequency band because of a greater influence of the source component on the SSRn than the site one. Because of this incapacity to assess the ground motion amplification and of the ramp up of the HVSR method, SSRn has been gradually abandoned. However, only few studies explicitly paid attention to the origin and to the distribution of the noise source, to the site-reference distance and/or to the geological configuration of the surrounding area. Aki (1988) showed that as the microtremors period decreases, their dependence on the local sources increases. Horike et al. (2001) tested the SSRn on a large (5 km diameter) and a small (0.5 km diameter) arrays between 1 and 10 Hz. The small array was located on a volcanic ash terrace as well as the reference while the majority of stations of the large array were located on a sand plain. They found that SSRn was able to achieve the amplification factor up to 5–8 Hz for the smallest area. At the larger area they didn't succeed in finding the amplitude even if the general shape of the SSRn transfer function was quite similar to those from the SSR. They explained this discrepancy by the presence of close anthropic sources and concluded that the incoming microtremors are the same only within limited areas of a few hundred meters in diameter. However, they did not discuss the possible influence of the differences in the geological conditions between both areas. Ferritto (1996) tested the repeatability of microseisms SSRn results from several arrays both with rock and soil reference site. He observed a good stability of the general shapes of the relative transfer functions estimate in both cases. He concluded that the SSRn referring to soil site is a good tool for microzonation.

4.4.4. SSRh

We introduce here the SSRh (hybrid SSR) approach that combines the SSRn based on ambient noise recording with the classical SSR perform on earthquakes recording. The SSR is used to assess the rock relative site response at one location inside the basin, and thus required only two long term or permanent stations. In a second time, the spatial variability of the basin response is estimated through the SSRn perform on several short time recording stations inside the basin, and relatively to the soil station where the rock relative SSR transfer function is known (Figure 47).



$$\mathrm{SSR}_{i/r} = \langle \frac{\{|\bar{U}_i|\}}{\{|\bar{U}_r|\}} \rangle$$

$$\mathrm{SSRn}_{i/r} = \left\{ \langle \frac{|\bar{u}_i|}{|\bar{u}_r|} \rangle \right\}$$

$$\mathrm{SSRh}_{i/r} = \langle \frac{\{|\bar{U}_s|\}}{\{|\bar{U}_r|\}} \rangle \cdot \left\{ \langle \frac{|\bar{u}_i|}{|\bar{u}_s|} \rangle \right\}$$

With

$|\bar{U}|$ = seismic spectrum
$|\bar{u}|$ = noise spectrum
$\{\cdot\}$ = smoothing
$\langle\cdot\rangle$ = geometrical mean

Figure 47. Illustration of the SSR, SSRn, and SSRh methods for the case of a sedimentary basin.

The SSRh can thus be expressed as:

$$SSRh_{i/r}(w) = SSR_{s/r}(w) \cdot SSRn_{i/s}(w) \qquad 35$$

where $SSRh_{i/r}$ is the SSRh at site i relatively to the rock reference site r, $\mathrm{SSRn}_{i/s}$ is the SSRn evaluated from noise at site i relatively to the soil reference site s and $\mathrm{SSR}_{s/r}$ is the classical SSR at the soil reference site s located inside the basin relatively to the rock references site r.

Eq. (32), (34) and (35) gives finally:

$$SSRh_{i/r}(w) = \langle \frac{\{|U_s(w)|\}}{\{|U_r(w)|\}} \rangle_{log} \cdot \langle \frac{\{|u_i(w)|\}}{\{|u_s(w)|\}} \rangle_{log} \qquad 36$$

where $|U_s(w)|$ and $|u_s(w)|$ are the FAS evaluated at the soil reference site s from the earthquake signal and from the microtremors respectively.

The main idea behind the SSRh is to insure that the SSRn reference site is in the same geological conditions (inside the basin) and closer to the studied sites. The SSRh approach aim to capture the site response spatial variability inside sedimentary basin from microtremors. A first suggestion of using a reference site inside the basin for the SSRn was done by Ferritto (1996). He assessed the SSRn relatively to a soil reference station co-located with a borehole where boring logs have been realized. He used the SSRn to describe

the spatial variability of the site response and he proposed using the borehole information to deduce the rock-relative site response.

4.4.5. Cross-correlation or coherency

The spatial correlation of the ambient noise has been shown both theoretically and experimentally to yield to the imaginary parts of the Green's function under some conditions. The Green's function between two points A and B is the record that would be obtained in B if an impulsive source (Dirac) was placed in A. Following Prieto et al. (2011) notation, the prewhitening processing formulation is the mathematical equivalent of the coherency but expressed in the time domain:

$$\gamma_{AB}(w) = \left\{ \left| \langle \frac{u_A(w) u_B^*(w)}{|u_A(w)||u_B(w)|} \rangle_{lin} \right| \right\} \Leftrightarrow \tilde{C}_{AB}(\tau) \qquad 37$$

where $\gamma_{AB}(w)$ is the coherency and $\tilde{C}_{AB}(\tau)$ is the prewhitened correlation between site A and B. $\Leftrightarrow$ represent the transition between the frequency domain and the time domain through the Fourier transform. $|u_A(w)|$ and $|u_B(w)|$ are the FAS evaluated from the ambient noise at site A and B respectively and $\langle\cdot\rangle_{lin}$ stand for the ensemble arithmetic mean.

The reconstruction of the Green's function through the correlation process requires correlating a diffuse wavefield that respects the equipartition condition (Sánchez-Sesma and Campillo 2006; Sánchez-Sesma et al. 2006, 2008; Weaver et al. 2009) that supposes that all modes of the medium are excited with the same level of energy and with a random phase. This condition is realized (1) when random noise sources are distributed homogeneously in the medium, or (2) when wavefields emitted by discrete sources propagate in a diffusive medium. The former is permitted by averaging many noise windows that randomizes the source distribution (Medeiros et al. 2015). The latter is generally the assumption made when using the earthquake latter arrivals (Coda waves); The seismic wave undergoes multiple scattering by the heterogeneity of the earth that randomized the signal and cancel the initial anisotropy and polarization coming from the source (Hennino et al. 2001). However the noise source distribution is rarely even and the medium is often not fully diffusive in the required manner (Weaver et al. 2009; Mulargia 2012; Hillers et al. 2013; Yoritomo and Weaver 2016). That implies that the correlation is generally not identical to the Green's function (e.g., Hanasoge 2013) and the spatial heterogeneity of the ambient noise field impacts directly the evaluation of the Green's function depending of the orientation of the station pairs (Stehly et al. 2006, 2008; Froment et al. 2010; Chen et al. 2016). However, different techniques have been proposed to take into account this ambient noise field anisotropy (Stehly et al. 2008; Prieto et al. 2009, 2011; Seats et al. 2012; Lin et al. 2012b; Weaver 2012; Zhang and Yang 2013; Ermert et al. 2016; Sadeghisorkhani et al. 2016). For instance, the prewhitening processing help reducing the contamination by strong transient noise (Bensen et al. 2007). Moreover, numerous studies have demonstrated the interest of the results obtained from the noise correlation even without an exact Green's function reconstruction (e.g., Weaver 2013).

The vast majority of them were conducted at the regional scale and at low frequency (<1–2 Hz) because the noise is then mainly composed by coherent Rayleigh waves that are emitted from identified oceanic sources.

About half a century ago, Claerbout (1968) introduced the potential of the spatial correlations to retrieve the Green's function of the medium. Few decades later, the cross-correlations $(\mathrm{u_A(w)u_B^*(w)})$ between ambient noise recordings at two points of the medium has been demonstrated to allow for the emergence of the Green's function in helioseismology and in acoustic (Weaver and Lobkis 2001; Lobkis and Weaver 2001; Derode et al. 2003). Pioneer applications in seismology consisted in considering the phase of the cross-correlation for imaging the surface waves velocity in the geological structure (tomography) purpose. Campillo and Paul (2003) estimate the deterministic surface-wave Green's function from the correlation on the seismic Coda of regional earthquakes. Soon after, Shapiro and Campillo (2004) recognized surface waves on the ambient noise correlation at low frequencies and assessed dispersion curves which agreed with global earth model. Many authors since then used the phase of the cross-correlation to image the velocity structure of Earth's crust at the regional scale (e.g., Shapiro et al. 2005; Stehly et al. 2007) and more recently at the local scale (e.g., Lin et al. 2012a).

While the phase of the ambient noise correlation has demonstrated its aptitude to achieve tomographic images of geological units, the Green's function amplitude has received less but growing attention in the last decade. The first works on the amplitude were done with the coherency (see Prieto et al. 2011). The coherency gives the phase similarity at every frequency between ground motions recorded at two stations spatially separated by an inter-station distance Δd. However, it may also contain information on the amplitude of the reconstruct waves along their propagations. Indeed, after the first understanding by Aki (1957) who established that the spatial correlation corresponds to a Bessel function, numerous publications supported that the coherency may be interpreted as the attenuation along the path: Tsai (2011), Cupillard et al. (2011) and Liu and Ben-Zion (2013) propose a theoretical framework to understand the amplitude of the coherency that takes into account the attenuation and the distribution of the noise source; Various authors found numerical evidence that the coherency contains the geometrical spreading and the attenuation information (Cupillard and Capdeville 2010; Prieto et al. 2011; Cupillard et al. 2011; Weaver 2011, 2013, Lawrence et al. 2012, 2013; Weemstra et al. 2013; Liu and Ben-Zion 2013). Several other authors retrieve the attenuation associated to geological structure from empirical evaluation of the noise coherency mainly at low frequency (Matzel 2007, 2008; Prieto et al. 2009; Lawrence and Prieto 2011; Lin et al. 2011; Lawrence et al. 2012; Zhang and Yang 2013; Hillers et al. 2013; Bowden et al. 2015). Most of these studies found a direct relation between coherency and attenuation even if the dependence of the coherency to the source distribution appears generally decisive (Tsai 2011; Weaver 2012; Hanasoge 2013; Hillers et al. 2013; Menon et al. 2014).

4.4.6. ANIRF

At the same time with the coherency, few studies tried to simulate the earthquake propagation including both the attenuation along the path and the site effect near the surface from the interferometry by deconvolution. This approach employs both the amplitude and the phase of the ambient noise to access to the ambient noise impulse response function (ANIRF) of the medium. The $ANIRF_{A\rightarrow B}$ is the record that would be obtained in B, if an impulse source were generated in A. The amplitude information, such as the site amplification, is conserved by making the spectral normalization only with the station where the Green's function is expected to be generated (reference station):

$$ANIRF_{A\rightarrow B}(w) = \left\{ \left| \langle \frac{u_A(w) u_B^*(w)}{|u_A(w)|^2} \rangle_{lin} \right| \right\} \qquad 38$$

The principal drawback of using the amplitude information is the impossibility of performing temporal normalization and prewhitening. Thus, transient noise and seismic perturbation can significantly bias the results (Yoritomo and Weaver 2016). Moreover, the amplitude dependence to the uniformity of the source distribution is widely suspected (Tsai 2011; Hanasoge 2013). However, the ANIRF approach found the support of several studies that show promising results on simulating the propagation of earthquakes waves and estimating the spatial variability of the site amplification in a sedimentary basin.

Prieto and Beroza (2008) first made the comparison between recording of a real earthquake and a virtual earthquake simulated by the ANIRF. They evaluated the impulse response between a station located nearby the epicenter of a moderate earthquake and several stations located inside and outside the Los Angeles basin and found strong similarity in the waveform relative amplitude and phase between 4 and 10 seconds of periods. Only the relative amplification between sites is retrieved with this approach since the amplitude of the ANIRF is always calibrated to the earthquake records by a scalar value different between components but common to all station pairs (Viens et al. 2014). Some differences remain but could be easily explained by the discrepancy between the impulse source emerging from the noise correlation and the double-couple supposed to be generated at the earthquake source (Maruyama 1963). Moreover, the localization difference between the 10 km deep earthquake hypocenter and the virtual source at the surface is expected to limit the comparison. However, the amplification factor obtained from the ANIRF at the stations located inside the basin relatively to the one obtained at the rock sites were very similar to those observed during the earthquake. Three years later, Prieto et al. (2011) made complementary observations from another earthquake on the same array. Again, the ANIRF was reliable to predict, in the long period range [4-10s], the waveform observed at several stations during the earthquake. They also found limited impact of the seasonal variations and of the choice of the virtual source on the amplification response of the basin. The source discrepancy was treated later by Denolle et al. (2013) who applied a source depth and double-couple focal mechanism corrections to the ambient noise impulse source. In the same way as the two

previous works, they compared at long periods [4-10s] in various position of Southern California the ground motion predicted by the corrected ANIRF to the one observed from four regional earthquake of moderate magnitude. They found firstly that even without the corrections, the ANIRF was able to capture the wave propagation including both the geometrical spreading and the site effect amplification and duration extension. In a second step, the corrected ANIRF gave them a better agreement to the real earthquake records both for the amplitude and for the phase. The improvement was greater for the two strongest earthquakes whose focal mechanisms are the farthest to a surface-impulse source. Soon after, Denolle et al. (2014a) extended this method to a finite rupture by deploying seismometers along a fault of concern. Thus, they simulate the ground motion induced by various rupture scenario of a strong magnitude 7 earthquake on the San Andreas fault and compare the results to prediction made by numerical simulation from the same scenario. They found substantial difference in the shaking patterns across the Los Angeles Basin between both approaches. The ANIRF generally provide signals with higher amplitude than the one predicted by the numerical simulation, especially for receivers in the basin. However it allowed the observation of the influence of the source directivity and of a geological waveguide phenomenon. At the same time, Viens et al. (2014) proposed a hybrid schemes combining the ANIRF at low frequency (<0.25 Hz) and a non-stationary stochastic method at higher frequency to simulate the broadband ground motion. At low frequency, the waveforms simulated by the ANIRF were very similar to those observed from moderate earthquakes and authors achieve to retrieve the geometrical spreading and local site amplifications. This good result was obtained without correction of the source mechanism and location. They also tested the extrapolation from low-to-moderate magnitude earthquake to larger magnitude earthquake by applying a scaling law to the calibration factor obtains between the amplitude of the ANIRF and the real low-to-moderate earthquakes recording. They found that such a scaling to larger magnitude process give satisfactory results to target moderate earthquakes up to magnitude 6. Viens et al. (2015) extended the applicability of the ANIRF to the case of subduction earthquakes. They successfully simulate a moderate subduction earthquake in Japan in the 5 to 8 s period range after applying the source correction of Denolle et al. (2013).

Recently, several studies tried to evaluate specifically the site response at the basin scale from the ANIRF. Denolle et al. (2014b) evaluated the relative amplification from a dense array of 296 receivers located in the Kanto Basin in Japan and stimulated by 375 virtual sources generated on distant stations (up to 300 km) outside the basin. This evaluation was done at low frequency only (2-10s) but included the f_0 of this very deep basin (~7 s) and the surface wave geometrical spreading was corrected by a $1/\sqrt{r}$ term with r the distance between stations to cancel out propagation effects. They were able to qualitatively reproduce the spatial variation of the basin amplification and found a linear relation between the amplitude of the ANIRF peak and the basin thickness. Bowden et al. (2015) evaluate high-resolution amplification maps in the Californian Long Beach basin at three frequencies (0.67, 1.0 and 2.0 Hz) from a dense array composed by 5200 geophones. They applied the wavefront tracking method developed few years before by Lin et al. (2012b). This approach proposed a 2D Rayleigh wavefield model that separates the source, the site amplification,

the attenuation and scattering by studying both directionally-dependent incoming and outgoing amplitude variation of the noise correlation. It used the wavefield gradient across the array in order to be less dependent on the evenness of the noise sources distribution but it requires a large number of stations uniformly distributed. They finally found a good agreement between the amplification maps and the amplification that could be expected regarding the previous phase velocity observations and the geometry and geology of the basin. Similarly to Denolle et al. (2014b), Viens et al. (2016) simulate wave propagation across the Kanto basin from the ANIRF between 1 and 10 s of period. They successfully simulate the waveform and amplification pattern observed across the basin during two moderate shallow earthquakes. Moreover, in addition to the observation of fundamental mode of Rayleigh waves, authors achieve to identify S-wave around 10 s of periods in the ANIRF. The ambient noise Green's function composition in the Kanto basin was studied more in detail by Boué et al. (2016). They demonstrate the potential of the noise correlation to capture the 3D wave propagation in the basin including the trapping and focusing effect and the complex wave reflection and conversion at the edges of the basin.

4.5. Data and processing

After this review, we now focus on the data presentation and processing. Two arrays located across sedimentary basins are studied, using both earthquakes and ambient noise recordings.

4.5.1. Provence site array

The studied industrial area is located in Provence in the southwestern Alpine foreland (Southeastern France). This region is relatively active compared to other regions of mainland France, even though, the seismic activity of this area remains low-to-moderate (Guéguen et al. 2007b; Sanchez et al. 2010). P1 and P2 are located on outcropping massive Cretaceous limestone. P3 to P7 are located within a rather small paleovalley (few hundred meters wide, 50 to 150 m deep) filled with stiff Miocene sands and sandstones and Quaternary softer deposits. Based on geophysical measurements, V_{S30} could be evaluated at 2100, 1800, 440 and 720 m/s for P1, P2, P3 and P4 respectively. P1 and P2 stations thus belong to the "hard rock" class, whereas P3 and P4 belong to "very dense soil" class according to the NEHRP classification.

Figure 48 presents the geological map and the velocity profiles measurements at the rock site P1 and the soil site P4 from two invasive techniques: cross-hole and PS Suspension Logging (PSSL). Site P1 also had non-invasive investigation, as it was one the studied site for the interPACIFIC project (Garofalo et al. 2016). Sensors are Güralp CMG6-TD broadband velocimeters with a sampling rate of 100 Hz. They have been continuously recording from few months to more than two years depending on the station (Perron et al. 2015). The main database was built with data recorded continuously between February 2012 and June 2014

at every site. More than 500 weak motion earthquakes were recorded, mostly low magnitude (M_L from 1 to 2.5) regional earthquakes.

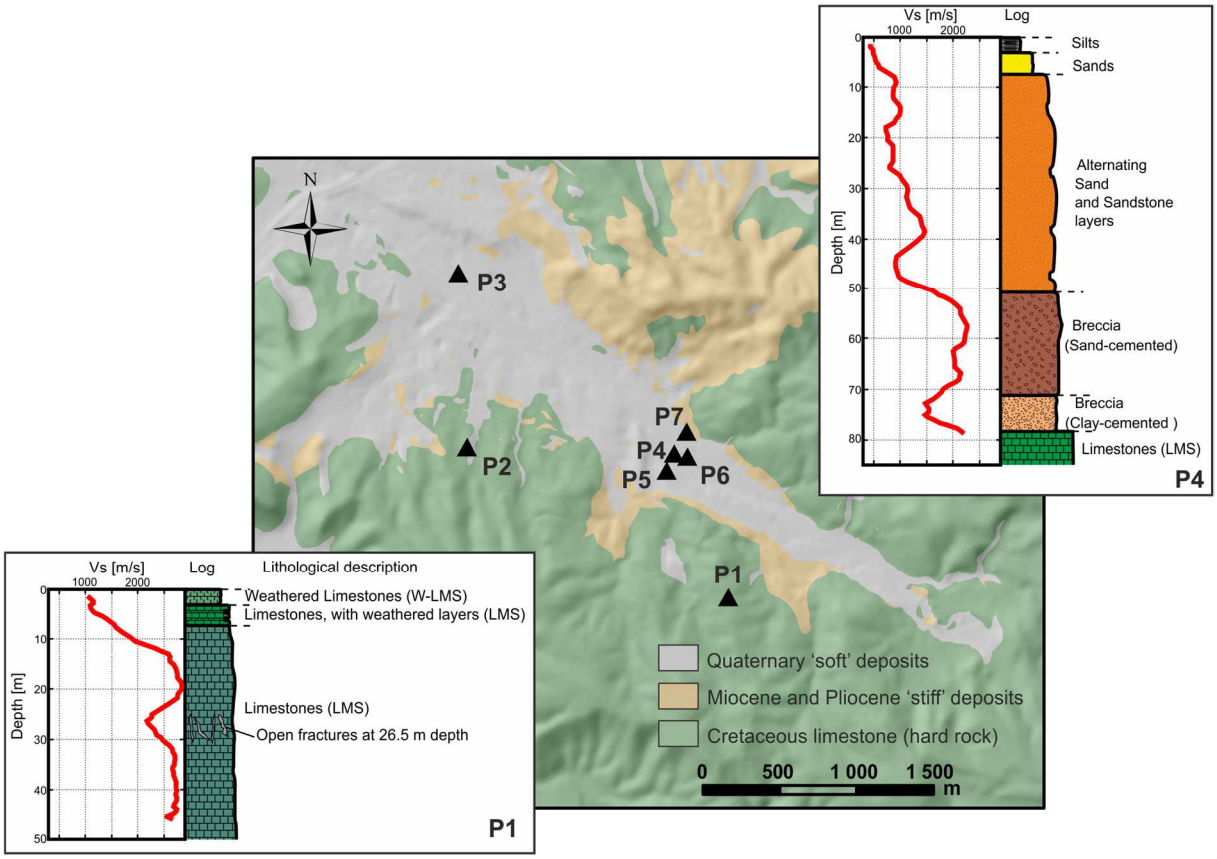


Figure 48. Geological map of the recording area in Provence. P1 and P2 are the rock sites and P3 to P7 are the 5 stiff soil sites. At stations P1 and P4, 3 boreholes allowed to determine the shear wave velocity profile with depth ($V_S(z)$) from two different techniques: cross-hole and PS Suspension Logging (PSSL).

4.5.2. Argostoli array

The Cephalonia Island (Western Greece, Ionian island) area is located at the north-western end of the Aegean subduction frontal thrust which is linked to the dextral Cephalonia-Levkas transform fault located at the west of Cephalonia. The seismotectonic context is one of the most active in Europe and the potential for destructive earthquakes is important, such as the 1953 M7.3 earthquake. The Plio-Quaternary Koutavos and Pliocene basins are located on a NE striking syncline bounded to the west by a thrusted asymmetric anticline (Figure 49). The western flank of this anticline is faulted by two east-dipping reverse faults (White Rock Fault and Argostoli Fault).

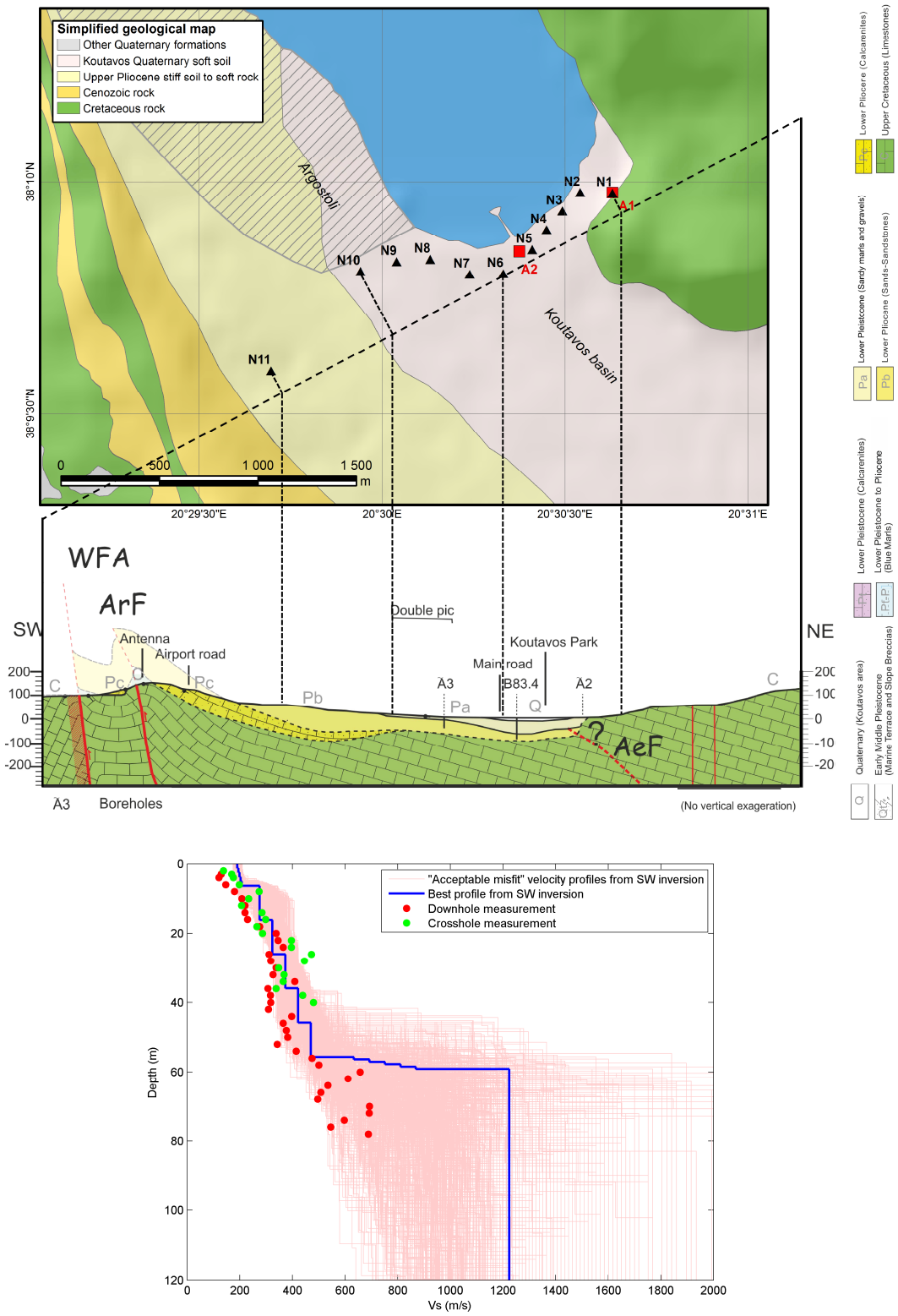


Figure 49. Geological map and section of the recording area near Argostoli. N1 is the reference rock site located on limestone while N2 to N11 are the soil sites across the basin. WFA is the White Rock Fault, while ArF correspond to the Argostoli fault. At site A2 (red square) three boreholes allowed

to determine the shear wave velocity profile with depth (Vs(z)) from two invasive techniques: cross-hole and down-hole. A temporary network at the surface also gave the opportunity to assess the velocity profile from the surface waves dispersion curve inversion. The blue line indicates the best estimate produce by the inversion.

A high resolution experiment took place from September 2011 to April 2012 in the framework of the NERA project (http://www.nera-eu.org/, www.orfeus-eu.org/organization/projects/NERA/Deliverables/). This experiment has been conducted by four institutes (Institut des Sciences de la Terre, Grenoble, France, ISTERRE ; Instituto Nationale di Geofisicale Vulcanologia, Rome, Italy, INGV; GeoForschungsZentrum, Potsdam, Germany, GFZ; Institute of Engineering Seismology and Earthquake Engineering, ITSAK). Hereafter we present results from a limited part of the network. Only records from 11 velocimeters installed by INGV have been used. All those velocimeters are 5s Lennartz with 100 Hz sample rate. This sub-array is arranged along a profile across the sedimentary basin as shown by the Figure 49. N1 is the rock reference site while N6 has been chosen as the sedimentary reference site. Hundreds of regional weak motion earthquakes have been recorded (Cultrera et al. 2014). The shear wave velocity profile at the center of the sedimentary basin is given at the bottom of Figure 49. Both the geological map and the velocity profile have been determined afterwards by the Sinaps@ project (http://www.institut-seism.fr/projets/sinaps/).The “Argonet” permanent vertical network was installed in 2015 at the surface and depth of the borehole and was used to determine the velocity profile together with ambient noise array measurements (Hollender et al. 2015; Cushing et al. 2016). However, this database is not used in the present work.

4.5.3. Processing

For the SSR method evaluated in Provence, FAS in Eq. (32) are processed from the full signal (i.e., P + S + coda waves) of tens to hundreds weak motion earthquakes depending of the site. SSR have been estimated also from the P-, S- and coda waves alone leading to similar results (an example is given latter Figure 53). For each frequency, the geometric mean is estimated from earthquakes satisfying SNR >3. The SSR assessment in Argostoli has been realized in a similar way in a precedent work by Cultrera et al. (2014).

For methods based on the ambient noise (HVSR, SSRn, SSRh, coherency and ANIRF) the quantity of input data available is huge allowing us to pick it through rigorous selective criteria. The main idea is to reduce transient noise through an appropriate window selection to allow the emergence of Green’s function (Baig et al. 2009; Seats et al. 2012; Melo et al. 2013; Bowden et al. 2015; Yoritomo and Weaver 2016). Indeed, transient noise may perturb the stationarity of the ambient noise field: at low frequency transient noise is composed of teleseism events, storms, etc. and at higher frequencies it corresponds principally to the anthropic activity (e.g. footsteps, close traffic). Because hours to years of noise are usually considered, instrumental issues (like time shift) can be easily introduced as well (Stehly et al. 2007). Bensen et al. (2007) gives in detail an example of ambient noise data processing allowing the emergence of the Green’s function.

In Provence, 150 files of one hour of noise are selected through a complete checking from many randomly picked one hour file of the continuous data. Each one-hour file is windowed in 60 windows of one minute. Every one-minute window, at the site or at the reference, presenting peaks exceeding 10 times the standard deviation evaluated from the full hour data is removed following Denolle et al. (2013). If more than 70% of the one-minute windows have been removed then the full hour window is rejected. Otherwise, the average spectrum and associated variation coefficient are evaluated from the selected one-minute windows. If the average variation coefficient on the whole frequency band is upper than 100% or if the maximal variation coefficient between 0.2 and 15 Hz is exceeding 150% then the full hour is rejected.

In Argostoli a similar procedure has been followed but the continuous data are recorded in 24h files and the ambient noise level is higher than in Provence leading us to select 8 hours at night from 25 files. Finally, the same procedure as in Provence is employed to extract one minute windows from each 8-hours record.

At the end of the selection process for both sites, thousands of one-minute windows have been selected. Those selected one-minute windows recorded simultaneously for each component at the site and at the reference are then used to assess HVSR, SSRn, SSRh, coherency and ANIRF approaches through Eq. (33), (34), (36), (37) and (38) respectively. The processing of the FAS is identical for both the ambient noise and earthquake-based approaches. Since the distribution of SSR is log-normal (Perron et al. 2015), the geometrical mean is considered for every approach relying on spectral ratio (SSRn, SSRh, HVSR). On the contrary, approach based on the noise correlation technique (ANIRF and coherency) requires the arithmetic mean of the complex spectra in order to keep the information of the phase at the numerator. The module and smoothing procedure are then applied after on the mean value. The smoothing is done by the Konno and Ohmachi (1998) procedure with a b-value of 30 and the frequency vector is resampled on a logarithmic scale.

Some studies propose to smooth or to impose a water level to the denominator to avoid division by zero. After having tested such processing, we do not find convincing results and prefer keeping the raw formulation. A simple test to evaluate the dependence of the results to transient noise is to compare evaluation made from noise recorded at days and at night. We did it for the SSRn and for the ANIRF and found very limited differences that gives us confidence in the stability of our results (not represented here).

4.6. Microtremors processing with amplitude-based methods (HVSR, SSRn and SSRh)

The comparison between the site transfer function estimated from the HVSR (brown), the SSRn (green) and the SSRh (blue) with the SSR (red) is made on Figure 50 and Figure 51 for the Provence and Argostoli network respectively. The rock reference stations are P1 and

N1, and the soil reference station for the SSRh approach are P4 and N6 for the Provence and the Argostoli array respectively. Figures display the quadratic mean of the horizontal component of the site transfer function. For each station the distance to the reference (Δd) for both referenced microtremor approaches (SSRn and SSRh) is indicated as well as f_0 of each site where the HVSR and SSR peak was clear enough to be picked (black dotted lines). Accordingly to the SSR and HVSR curves, the lower f_0 (f_{0min}) are approximately 3–4 Hz for the Provence basin and 1.5 Hz for the Argostoli one. SSR estimates in Provence between two similar rock sites (P2/P1) indicate a graduate amplification from 6 to 30 Hz at site P2. This illustrates the difficulty of selecting of reference site presenting a negligible site response. In Argostoli at sites N7 to N11 f_0 is less clear and the amplification factor is lower than for site N2 to N6. Those observations agree with the complex geological structure of the South-West part of the basin (Figure 49) which indicates variations in the thickness of superficial layers and stiffer soil (lower Pleistocene and lower Pliocene).

4.6.1. HVSR

The HVSR on both networks and in agreement with the literature shows a peak at the fundamental resonance frequency (f_0) for some sites (e.g., P4, P6, N4, N5, N6). When the f_0 peak is not sharp on the SSR, due to 2D or 3D effects, neither it is on the HVSR. To that respect, picking accurately f_0 can be difficult in case of broadband HVSR shape (e.g., sites P3, N7 to N11). The resonance frequencies for higher modes (above f_0) are, however, never visible and the amplification factors are always underestimated on the whole frequency band with the HVSR. This approach is, thus, not able to reproduce the site transfer function obtained from the SSR.

4.6.2. SSRn

For every site in Provence, the SSRn shows a good agreement when compared to the SSR up to approximately 4 Hz. In Argostoli the agreement is also good for every site up to approximately 1.5 Hz. These frequencies are very similar to the f_{0min} observed for each basin. At higher frequencies, the amplitudes of the SSRn are widely overestimated even if the general shapes of the transfer functions are roughly similar when compared to the classical SSR. This overestimation of the SSRn at high frequency has been widely pointed out in the literature.

4.6.3. SSRh

The SSRh results are very similar to the SSR ones up to approximately 12 Hz in Provence for sites located in the basin but not for P2 that is a rock site located outside the basin. In the Argostoli basin, the agreement is visible for the full-frequency range available (up to 20 Hz) and for most of the sites. Indeed, slight discrepancy can be observed for sites N10 and N11 that are located on stiffer soils than the other sites of the array. Taking an SSRn reference site inside the basin (SSRh) improved then greatly the similarity of the results with the SSR ones when compared to using SSRn referring directly to the rock site. This is the first time to

our knowledge that a microtremor-based approach is able to predict accurately the SSR transfer function in such a broadband frequency range. The SSRh approach seems reliable to predict f_0 as well as higher modes and the amplification factor up to high frequencies.

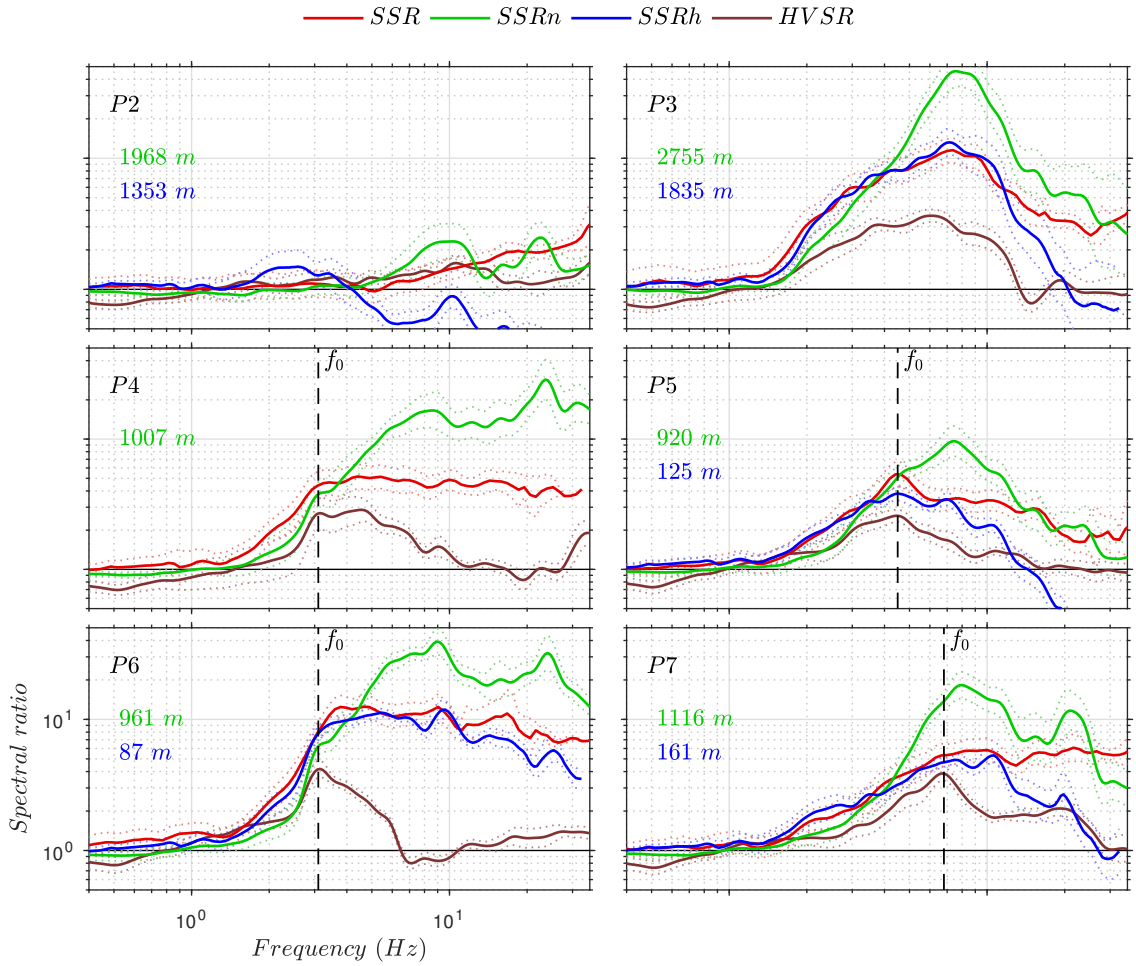


Figure 50. Comparison between the geometric mean of the horizontal mean component of the SSR (red), the SSRn (green), the SSRh (blue) and the HVSR (brown) in Provence at stations P2 to P7 with the rock station P1 has reference. P4 is the soil reference used for the SSRh technic. The Δd distances to station P1 (green) and to station P4 (blue) are indicated for each panel while f_0 (black dashed lines) is given for sites providing a noticeable peak in the HVSR and/or SSR. The curves corresponding to the geometric mean plus and minus standard deviation are represented with dotted lines.

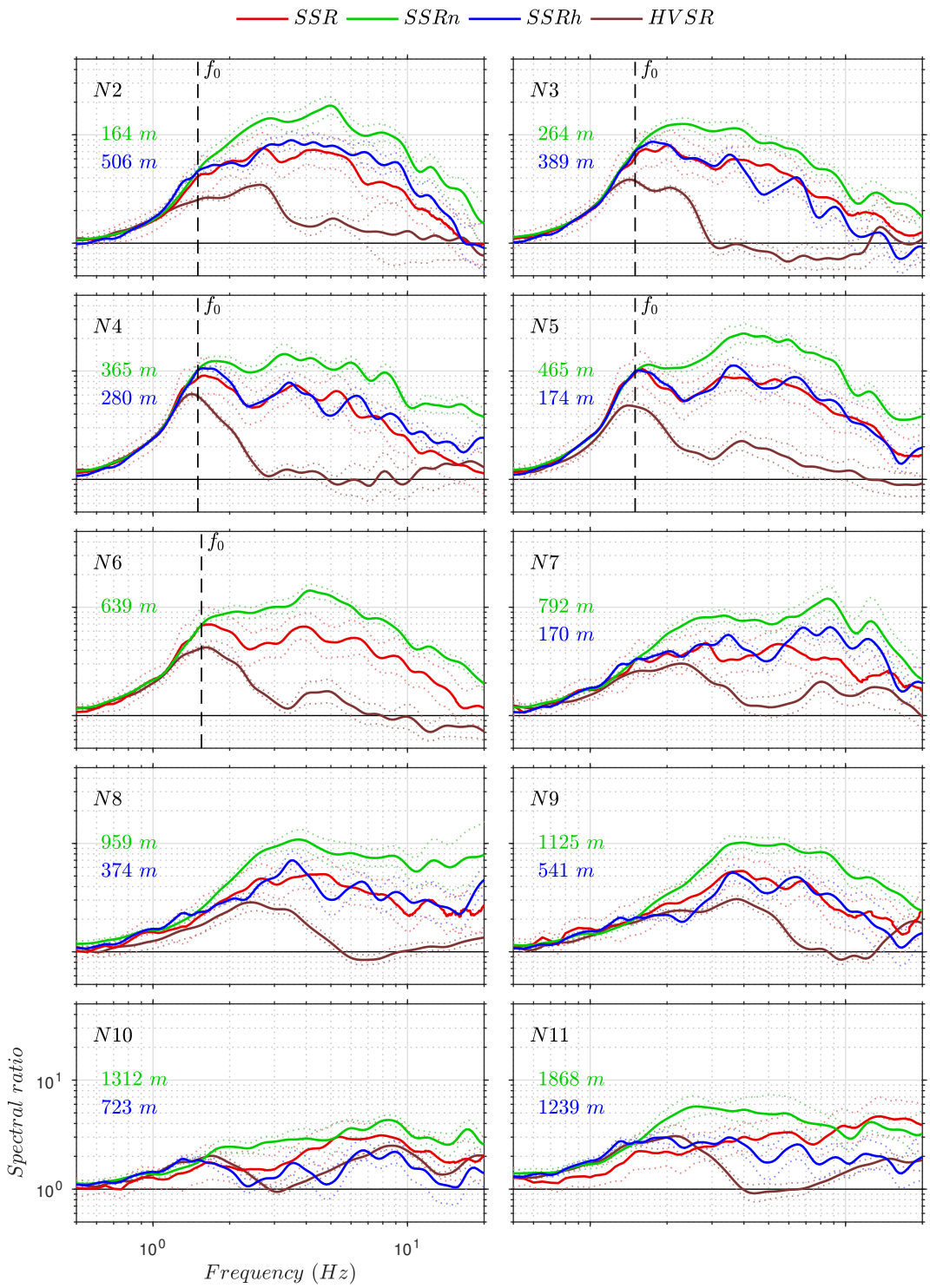


Figure 51. Identical to Figure 50 but for stations N2 to N11 in Argostoli.

4.6.4. A_0

The previous observations show that the amplification factor estimated by the SSRn approach can be reliable up to f_{0min} only. Thus, the SSRn approach gives the opportunity to have an estimation of the amplification factor at f_{0min} fully evaluated from microtremors. In a first step, the f_{0min} could be determined by the HVSR approach when it is possible and in a second step A_0 can be read on the SSRn at that frequency. Figure 52 gives an example of A_0 evaluation at site P6 relatively to P1 in Provence and N4 relatively to N1 in Argostoli.

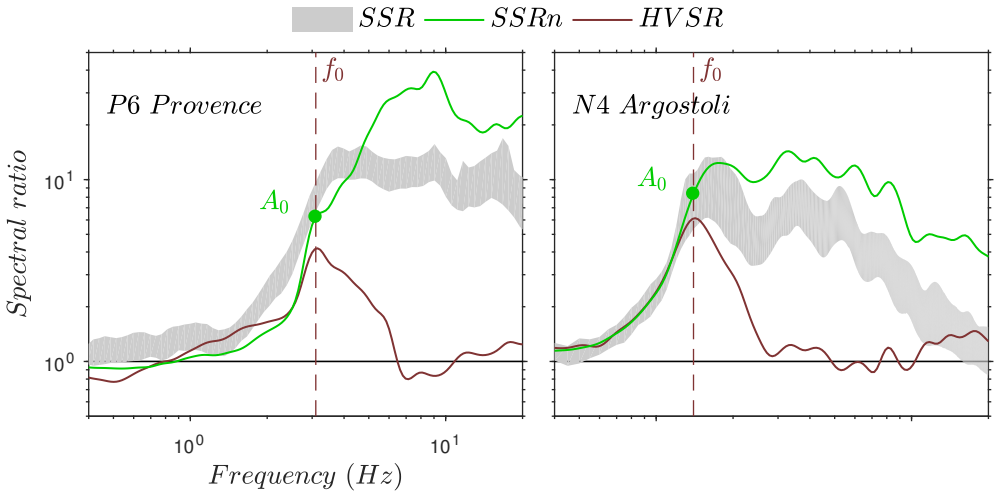


Figure 52. Example of A_0 evaluation from the mean HVSR (brown) and SSRn (green) curves at site P6 relatively to P1 in Provence (left panel) and at site N4 relatively to N1 in Argostoli (right panel). The SSR mean plus/minus standard deviation (grey) is represented as well for the comparison.

Table 3: Comparison of the amplification factor (A_0) at the minimal fundamental resonance frequency of the basin (f_{0min}) evaluate from the SSRn and from the SSR with the corresponding relative difference. The coefficient of variation (CV) of the SSRn and the SSR at f_{0min} are also indicated.

Network	Provence			Argostoli				
Site	P4	P5	P6	N2	N3	N4	N5	N6
A_0 from SSRn	3,8	5,1	6,2	4,7	7,2	10,2	9,8	6,9
A_0 from SSR	4.4	5,4	8,1	4,1	6,2	8,5	9,6	6,8
Relative diff. (%)	-15	-5	-23	14	16	20	2	2
SSRn CV at f_{0min} (%)	24	21	17	15	9	12	8	7
SSR CV at f_{0min} (%)	28	24	20	42	41	41	42	47

The estimation of A_0 from the SSRn is reliable to assess the one we would obtain from the SSR with fewer than 20% of relative difference for almost every site of Provence and Argostoli (Table 3) that is under the uncertainty associate to the SSR. This evaluation of A_0 gives a good first approximation of the maximal amplification factor of the site. It has the

advantage to be completely independent of the recording of earthquake. This approach could give a suitable solution for low seismicity areas where no earthquake can be recorded in a sufficiently short period of time.

4.6.5. Discussion

- On the observed agreement between microtremor and earthquakes-based approaches

Because body waves emitted by earthquakes are supposed to come with a quasi-vertical incidence whereas microtremors are mainly composed of surface waves propagating horizontally, the potential of microtremors utilization for site effect evaluation was long since debated. The theoretical point of view is not discussed here, but we address this question through the experimental comparison between the SSR and the SSRn for two sites located inside the basin. Figure 53 shows the comparison between the geometric mean of SSRn estimated between two sites located inside the basin at almost 2 km of distance (P3/P4 in Provence) and the geometric mean SSR estimated from various phase of the signals (P-, S-, coda-, full signal) recorded from more than one hundred earthquakes. The relative site responses estimated from the ambient noise and from the different phases of the signal are very close together up to 12 Hz for the mean horizontal component. Indeed, below 12 Hz, the geometric mean of SSRn estimated is within the inter-event uncertainty interval of the SSR for most of the considered phases. This result is also true for the vertical component even if the discrepancy of the geometric mean of SSRn above f_0 is slightly higher. This demonstrates that the satisfactory results observed previously with the SSRh can be entirely explained by the reliability of the SSRn to assess the response of the site to weak motion solicitations when the reference station is located inside the basin.

Some authors proposed explanations for some apparent agreement between spectral ratio from microtremors and from earthquakes recording. Field and Jacob (1993) computed the theoretical response of a horizontally stratified sedimentary layer to ambient noise sources and found that the expected horizontal-component spectrum of ambient noise contains its most prominent peak at the f_0 predicted for incident shear waves. Horike et al. (2001) observed a good agreement between SSRn and SSR for site located on similar geological condition that they explained, based on the work of Satoh et al. (2001), by the possible presence of surface waves in the S-waves portion of the seismic signal that may be generated by near-site inhomogeneities. Aki (1957) already notice that “it is well known that the characteristics of the ground are reflected more or less in its vibration whatever the origin of the vibration may be”. Our results seem to support this latter interpretation, especially in sedimentary basins where the strong trapping phenomena might be weakly dependent on the incoming wavefield (earthquake or noise) and where trapped waves may dominate the wavefield inside the basin.

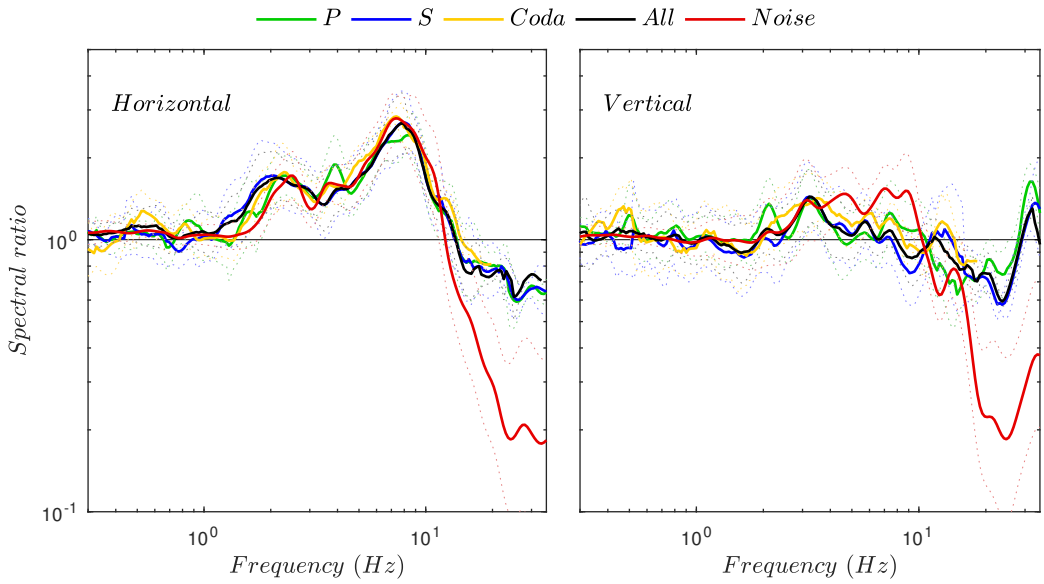


Figure 53. Comparison between the geometric mean SSRn (red) estimate from ambient noise between two stations located inside the basin in Provence (P3/P4) with the geometric mean SSR estimate from various phase of the earthquake signal: the P-wave (green), the S-wave (blue), the coda wave (yellow) and the full signal (black). The geometric mean of the horizontal (left) and vertical (right) components are estimated from more than an hundred earthquakes for the SSR and from thousands of selected noise windows for the SSRn. The curves corresponding to the mean plus and minus standard deviation are represented with dotted lines.

❖ On a proposition of interpretation for the rock-referenced SSRn over-estimation

Contrary to the HVSR, the SSRn discrepancy with the SSR has been poorly discussed in the literature. The majority of the previous studies observed an overestimation of the amplification factor obtain at high frequencies by the SSRn referring the rock sites but only few investigate its origin. Numerous authors consider that the SSRn limitation was due to disturbance of the ambient noise wavefield by local sources. Indeed, the source of discrepancy between SSR and SSRn is generally imputed to the concentration of human activities mainly into the basin that makes the average level of anthropogenic noise higher for the soil site than for the rock sites that are usually located farther in a quieter area.

Here, we propose that the SSRn incapacity to assess the amplification at high frequency might be primarily due to geological differences between the study site and the reference site. Indeed, based on the observations that the SSRn overestimate the amplification for frequency above the smallest f_0 of the basin (f_{0min}), we proposed that basins could be a natural barrier for the surface waves with wavelength smaller than the basin size that forbid a part of them to propagate outside the basin. Thus, the SSRn rock reference station located nearby the basin might see only a part of the ambient noise wavefield for frequency above f_{0min}, while the station located at the study site inside the basin see the full wavefield at every frequency. Leading thus to an overestimation of the amplification above f_{0min} in the SSRn spectral ratio between the soil site and the rock-reference site. In the following, we

referred to this suspected phenomenon as “source hiding”. The source hiding is expected to concerns only the surface waves that cross the basin along their propagation from the source to the recording site. For a given rock recording site, this is mostly the case when the ambient noise is generated inside or behind (according to the site) the basin. Differently for the classical SSR, all the body waves emitted by the earthquake source are received both at the study and at the rock reference site as the basin is unlikely crossed by the body waves arriving at the rock reference site with an incidence angle close to the vertical.

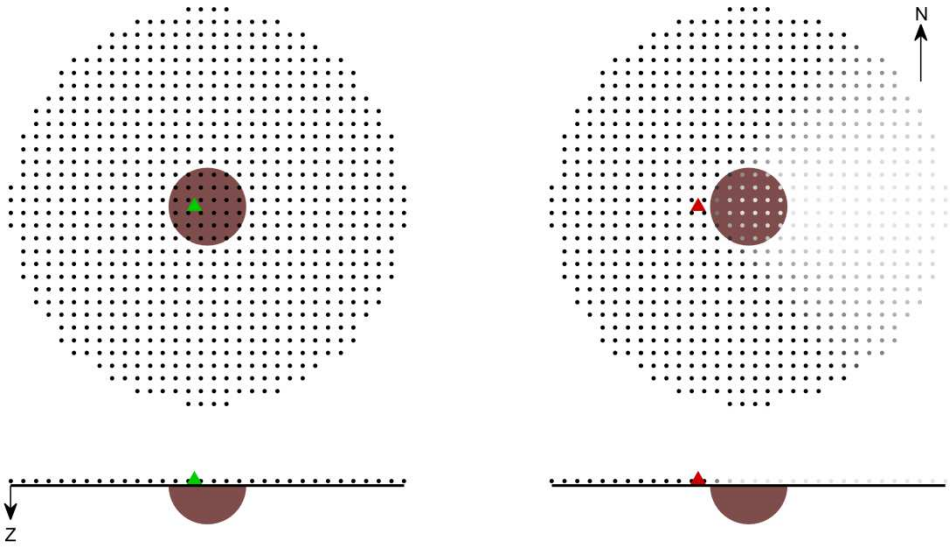


Figure 54. Illustration of the microtremor source hiding phenomenon. Plan view (at the top) and Est-West cross-section (at the bottom) of a sedimentary basin (brown half-ball) surrounding by a higher velocity medium (in white) and submitted to multiple microtremors source uniformly distributed at the surface (grid of grey to black points). The lighter is the color, the lower is the frequency above which waves are trapped during the propagation from the source to the receiver (triangle). The case of a receiver located inside the basin is illustrated on the left (green) while the one outside is given on the right (red). The wave is assumed to travel along a straight line between the source and the receiver.

The ambient noise source hiding is illustrated on Figure 54 by a very simple and idealistic example. A sedimentary basin is represented by a half-ball (in brown) surrounded by a stiffer medium (in white). The wave velocity in the basin is lower than the one in the surrounding medium. Multiple microtremor sources are represented with a uniform distribution at the free surface (grid of points) and are assumed to generate deterministic surface waves at every frequencies. The plan view is given at the top while the cross-section is represented at the bottom. Two cases are illustrated to show which sources are seen by a receiver depending on it position. First, a receiver is located inside the basin (green triangle – left panel) whereas secondly a receiver is located at the West of the basin (red triangle – right panel). On this illustrative example, every source is seen at every frequency by the receiver located inside the basin while waves emitted by sources located at the East of the basin are trapped from the view point of the receiver located at the West of the basin. The frequency above which

the waves might be trapped inside the basin is expected to corresponds to the lowest f_0 cross on the path from the noise sources to the rock reference site. Thus, the color (from white to black) of each point of the grid indicates the frequency above which the wavefield emitted by this source is expected to be trapped on the path to the receiver and thus is not seen by it. Lighter is the color lower is the frequency above which the waves are trapped. Black points correspond to points without any trapping. The source hiding phenomenon is then expected to start at f_{0min} and to increase slightly with increasing frequency. It gives a suitable explanation for the overestimation observed on the SSRn beyond f_{0min} here and at high frequency in the literature when compared to the classical SSR from earthquakes.

In terms of random noise wavefield, the source hiding phenomena could induce a difference in the number and in the spatial distribution of sources seen by the station located inside the basin and the reference station located outside.

4.7. Microtremors processing with coherency and ANIRF methods

4.7.1. Phase of the reconstructed wavefields

One of the most common and reliable way to assess ambient noise phase information is the prewhitened cross-correlation (Eq. (37)). It is equivalent to the coherency but expressed in the time domain. Because the virtual source and recorded signal are both three-component vectors, approaches based on noise correlation provide rank-two tensors with nine components. To limit the number of plots and because the vertical component is the simplest to process and interpret as it mainly contains Rayleigh waves, only the VV component of the tensor is displayed in the time series representation. In the same way only the results extracted from the Argostoli network are given since this network provides a better source-receiver distance distribution than the Provence site.

The mean of the causal and anti-causal parts of the VV component of the whitened correlation recorded between 2 and 4 Hz are given with N1 and with N11 taken as the virtual source at Figure 55. Taking N11 as the virtual source seems to improve the reconstruction of the green function when compared to taking N1. The amplitude of the reconstructed wavefield is much higher with the virtual source N11 and one group velocity may be identified from the correlation. This velocity of approximately 630 m/s is in good agreement with the velocity evaluated in the borehole for the S-wave (Figure 49). Indeed, at 2 Hz the penetration depth for the fundamental mode of the Rayleigh waves is approximately 100 m leading to a mean S-wave velocity that is close to 600 m/s. For the shorter Δd distance (< 1000 m) the causal and anti-causal parts are superposed close to $t = 0\,s$. Then no first arrival can be detected. Moreover, the reconstruct waves seem to arrive before at rock site N1 than at site N2 even if the former is located further. This might be partly explained by the noticeable parts

of limestone with high velocity in the path between the virtual source N11 and the recording site N1 (Figure 49).

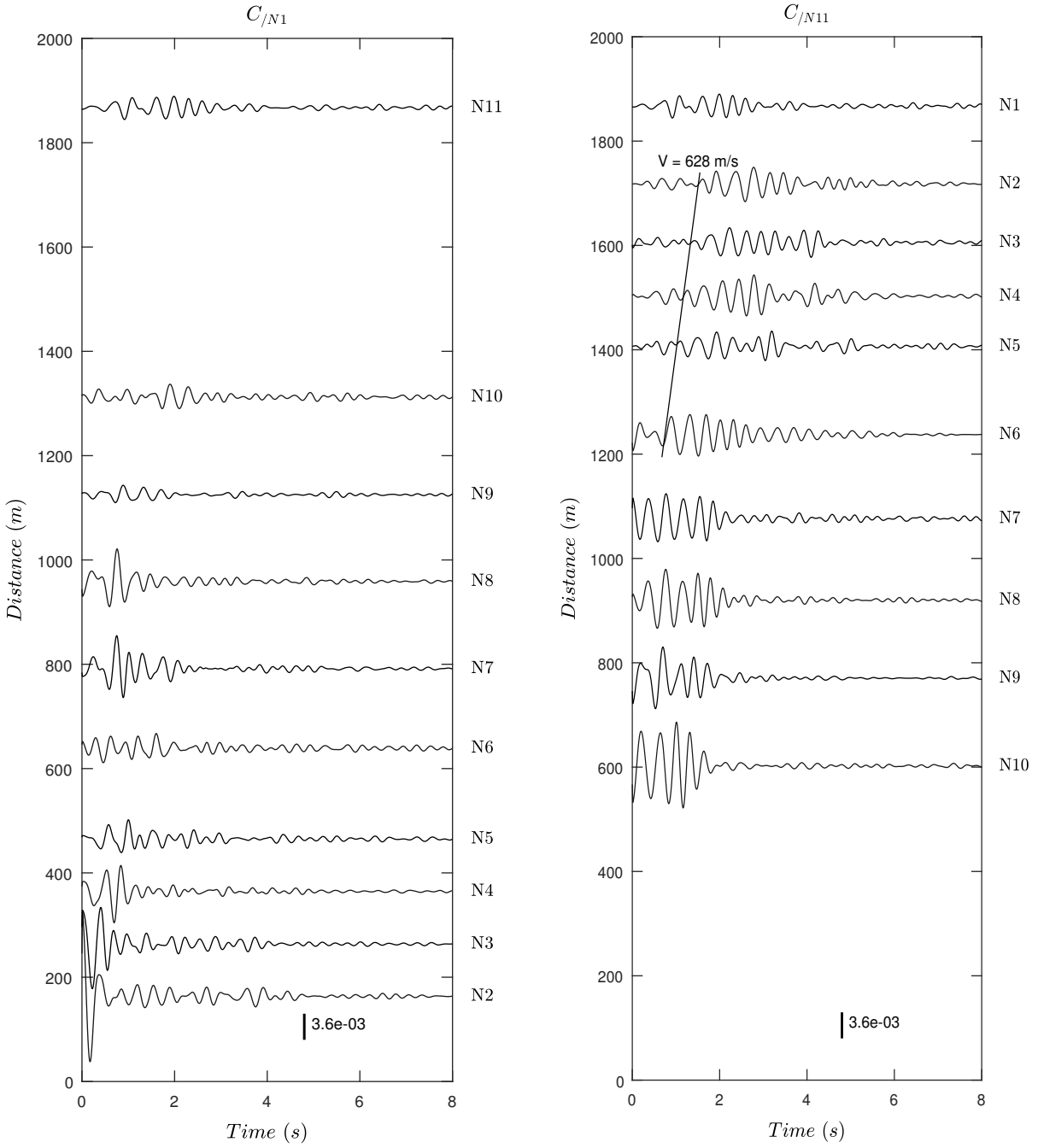


Figure 55. Mean of the causal and anti-causal parts of the VV component of the prewhitened correlation (C) between 2 and 4 Hz in Argostoli at sites N1 to N11 and with N1 (left panel) and N11 (right panel) taken as the virtual source. The time series are represented according to the distance between the receiver and the virtual source (Δd) and with a common amplitude scale.

4.7.2. Amplitude of the reconstructed wavefields

Figure 56 shows the time series reconstructed from the prewhitened correlation (black) and the ANIRF (red) at sites N10 to N1 with N11 taken as the virtual source for three frequency range (1–2 Hz, 2–4 Hz, and 4–8 Hz). Phases of the signal obtained with the coherence and with ANIRF approaches are very close, in agreement with equations (37) and (38). The amplitude of the prewhitened correlation shows a decay according to Δd that is found to be close to the one predicted by the geometrical spreading for surface waves ($\sqrt{\Delta d}$). This observation is in agreement with several studies that support that the coherency might contain the attenuation information along the path. Compared to the prewhitened correlation, the ANIRF also predicts an amplification of seismic waves for sites located inside the softer parts of the basin (N2 to N6), as well as an de-amplification for the rock site N1 (relatively to site N11 used as the virtual source) for the three frequency bands. In agreement with the site response obtained from the SSR (Figure 51), the ANIRF also predicts that the site amplification inside the basin is greater above 2 Hz than between 1 and 2 Hz. Between 4 and 8 Hz, the reconstruct wavefield does not display any clear arrival for farthest sites (N8 to N1), especially with the ANIRF.

4.7.3. Site transfer function assessment from the ANIRF

A better quantification of the ANIRF potential for the site effect assessment can be done by direct comparison with the SSR results in the frequency domain. To that purpose, the coherent parts of the time series reconstruct from the ANIRF are used for the SSR computation. Identically to the SSR computation based on earthquake records, the distance of the site and reference under study to the virtual source has to be much greater than the distance between the site and the reference. We test this approach in Argostoli by using as a virtual source N11 that is the farthest station from N2 and N1, respectively the target sedimentary site and the rock reference site (Figure 57). The distance between N11 and N2 or N1 is more than ten times the one between N2 and N1. Moreover, similarly to the SNR>3 criteria for the SSR approach, only the most coherent reconstructed wavefields (between 0 and 7 s in the time series, see Figure 56) which present frequencies having a SNR>3 both at the site and at the reference are considered. The horizontal mean component of the ANIRF is then computed as the quadratic mean of the spectrum for the EE and NN components.

Figure 58 presents the comparison between the $\mathrm{SSR}_{\mathrm{N2/N1}}$, the $\mathrm{SSRn}_{\mathrm{N2/N1}}$ the $\mathrm{SSRn}_{\mathrm{N2/N11}}/\mathrm{SSRn}_{\mathrm{N1/N11}}$, and $\mathrm{ANIRF}_{\mathrm{N11\rightarrow N2}}/\mathrm{ANIRF}_{\mathrm{N11\rightarrow N1}}$, that is the SSR computed on the coherent part of the signal virtually propagating from N11, as detailed previously. Below 4 Hz, the ANIRF results computed here using only one virtual source predicts well the site response that would have been obtained for a single earthquake as $\mathrm{ANIRF}_{\mathrm{N11\rightarrow N2}}/\mathrm{ANIRF}_{\mathrm{N11\rightarrow N1}}$ is mostly within the uncertainty interval of the SSR. Above 4 Hz, the $\mathrm{ANIRF}_{\mathrm{N11\rightarrow N2}}/\mathrm{ANIRF}_{\mathrm{N11\rightarrow N1}}$ provides results that are closer to $\mathrm{SSRn}_{\mathrm{N2/N1}}$ than to the $\mathrm{SSR}_{\mathrm{N2/N1}}$. This is probably due to the bad wavefield reconstruction already observed in Figure 56 at these frequencies.

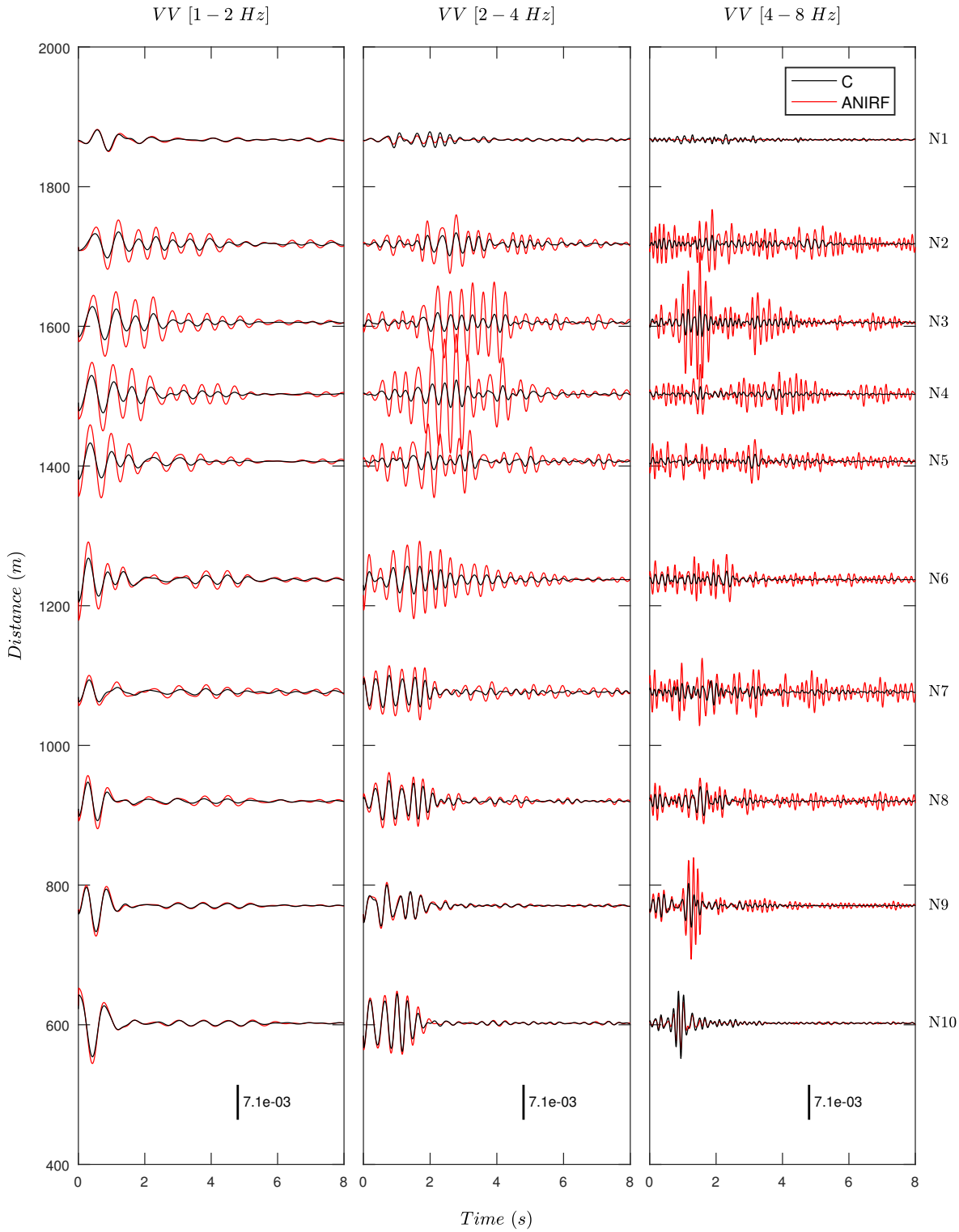


Figure 56. Mean of the causal and anti-causal parts of the VV component of the prewhitened correlation (black), and of the ANIRF (red) for three frequency ranges at sites N2 to N10 in Argostoli with N11 taken as the virtual source. The time series are represented according to the distance between the receiver and the virtual source (Δd) and with a common amplitude scale.

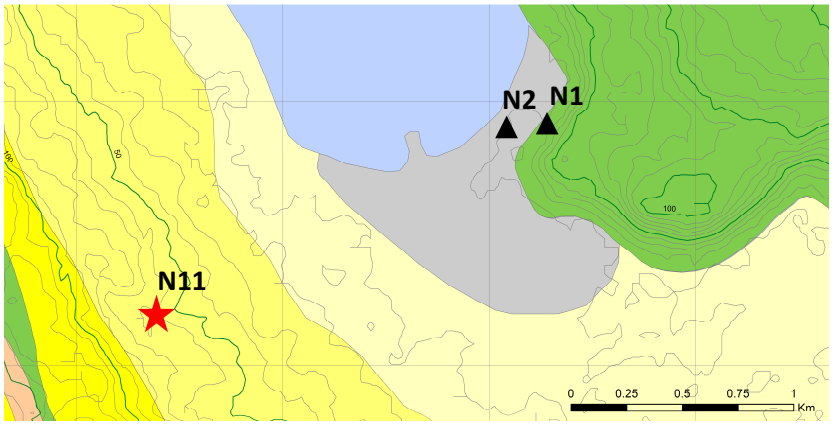


Figure 57. Illustration of the utilization of the ANIRF for site effect assessment. Site N11 is taken as a virtual source that is used to realize the traditional SSR between the soil site N2 and the rock reference site N1.

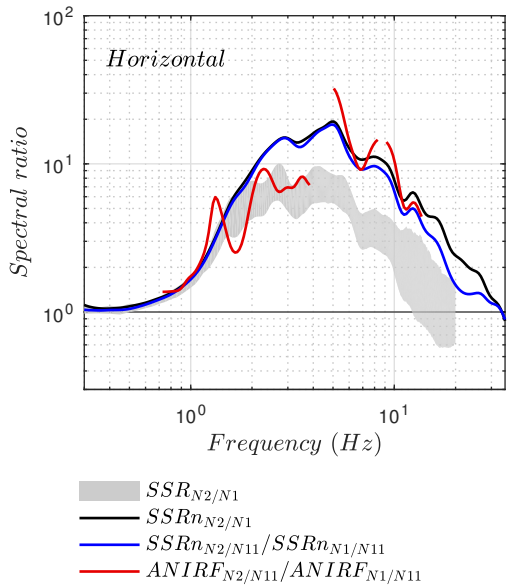


Figure 58. Comparison in Argostoli between the direct evaluation of the SSR (grey) and the SSRn (black and blue curves) between sites N2 and N1 with the evaluation of the ANIRF (red curve) when considering the distant station N11 as a virtual source. Only the horizontal mean component is represented.

This result is in agreement with the observation already made from the SSRh that the ambient noise is able to provide the site transfer function evaluated from earthquake recordings, even though the wavefield composition is not the same. As already observed before (Figure 51), the SSRn overestimate the SSR amplitude at high frequencies. This discrepancy is also observed when using N11 as an intermediate site ($\mathrm{SSRn_{N2/N11}}$/ $\mathrm{SSRn_{N1/N11}}$). Differently to the rock-referred SSRn, the ANIRF does not overestimate the amplification above f_0. Then, using only the coherent part of the ambient noise wavefield allow improving the agreement of the results obtained from the microtremors-based

approach. If we suppose the source hiding interpretation as true, the coherent ambient noise wavefield between the station inside and outside the basin might correspond only to the parts of the wavefield that is transmitted through the geological interface and thus seen by both stations. Using multiple ANIRF virtual sources to assess the site transfer function is then very promising. However, for most of the case, the targeted soil site is located at least hundreds of meters from the rock reference site. Then, respecting the ten times distance between these sites for selecting the site station taken as the virtual source leads to pluri-kilometers inter-station distance for the ANIRF computation. Thus, assessing the site transfer function from the ANIRF approach is mostly limited to the deep basins presenting low f_0 or to site arrays presenting very close rock reference sites as the quality of the reconstructed wavefield decrease rapidly at high frequencies when the inter-station increase.

4.8. Investigating the parameters that control the site response assessment from microtremors

It is of major interest to understand what controls the capability or incapability of microtremors approaches to predict the ground motion that would be produced by an earthquake. Indeed, limiting the number of stations that need recording earthquakes is crucial to predict the ground motion in large areas with a limited cost and time especially in low seismicity area where few earthquakes only are recordable in weeks or even months.

We evaluate the reliability of the SSRh by representing its variation with respect to the SSR (SSRh/SSR) and compare it to various parameters in Provence (Figure 59) and in Argostoli (Figure 60) in order to discuss what may control the SSRh reliability. These parameters are:

- frequencies corresponding to the wavelengths $\lambda = \Delta d$, $\lambda = \Delta d/2$ and $\lambda = \Delta d/3$ (vertical black lines) where Δd is the distance between the site and the soil reference site (SSRn parts). These frequencies are computed considering an average shear wave velocity derived from the velocity profiles in P4 and P1 for Provence (Figure 48) and in the center of the Argostoli basin (Figure 49). In Provence, the proportion of rock and sediment in the path between the site and the microtremor reference site is taken into account for the estimation of the average velocity while in Argostoli the rock portion is negligible. Hereafter we refer to these frequencies as $fh_{\Delta d}$, $fh_{\Delta d/2}$ and $fh_{\Delta d/3}$.

- The geometric (logarithmic) standard deviation of the SSRn used for computing the SSRh (σ_{SSRh} - lighted blue dotted curve) between evaluation of one-hour spectral ratios. It is representative for the time variability of the SSRn. The SSR inter-event variability (σ_{SSR} - red dotted curve) is represented as well for the comparison.

- The horizontal coherency (Eq. (37)) that expresses the similarity of the phase of the ambient noise wavefield between the studied site and the soil reference site (purple

curve). The average coherency is represented here and encompasses the two horizontal components. A coherency equal to 1 indicates identical phases while low coherency (e.g., below 0.5) corresponds to incoherent microtremors. A more detailed discussion on the coherency physical meaning is given section 4.4.5.

To facilitate the reading of all the curves but the coherency, three horizontal grey bands stand for the relative variation of 10, 30 and 70% from the darkest to the lightest bands respectively.

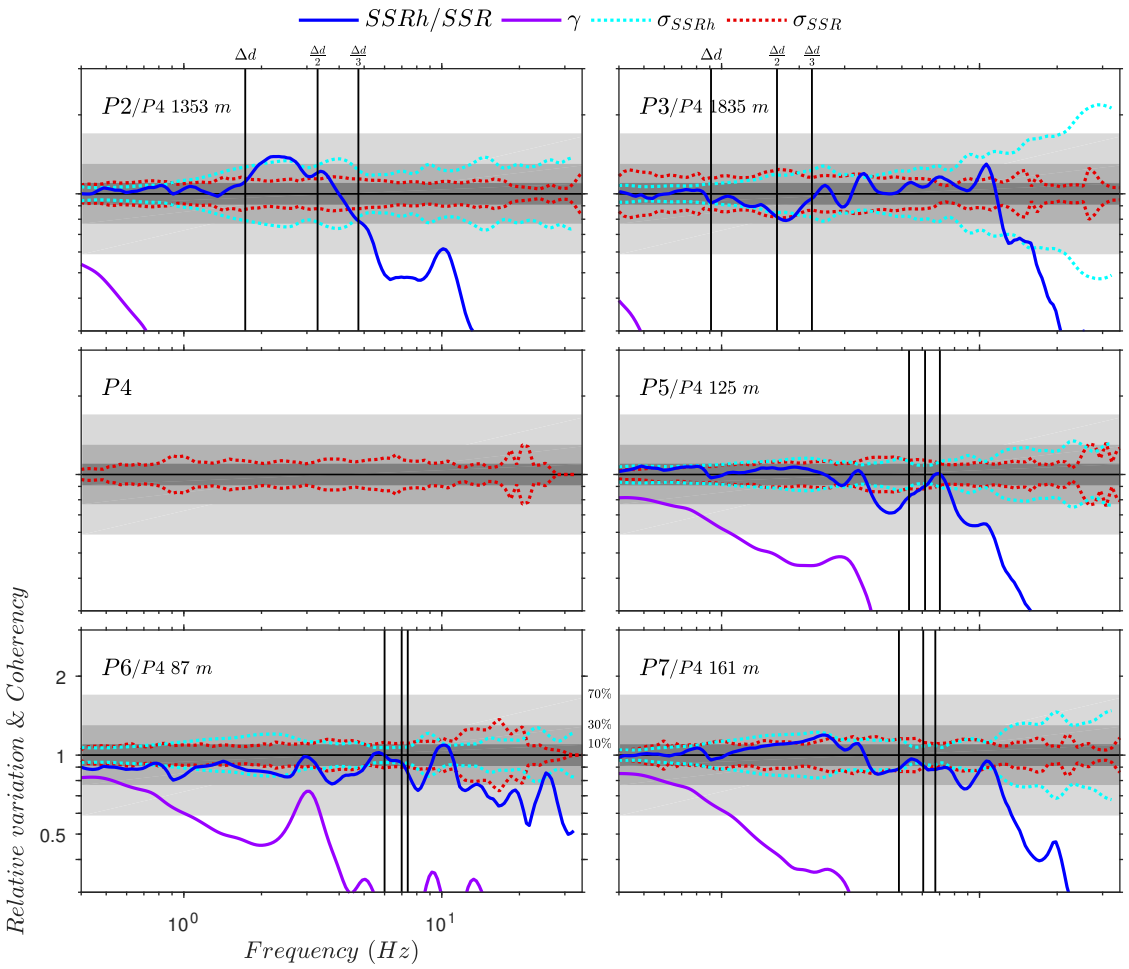


Figure 59. Comparison between the relative variation of the SSRh (blue curves) with respect to the SSR in Provence at site P2 to P7 according to the soil reference site P4. The time variability of the SSRh is given by the geometrical standard deviation σ_{SSRh} (blue dotted curves) and the inter-event variability of the SSR is represented by σ_{SSR} (red dotted curves). The horizontal mean coherency that vary between 0 and 1 is also represented (purple curves). The site-reference distances (Δd) is written on the top left corner and the frequency corresponding to the wavelength equal to Δd, $\Delta d/2$ and $\Delta d/3$ are represented as well (black vertical lines).

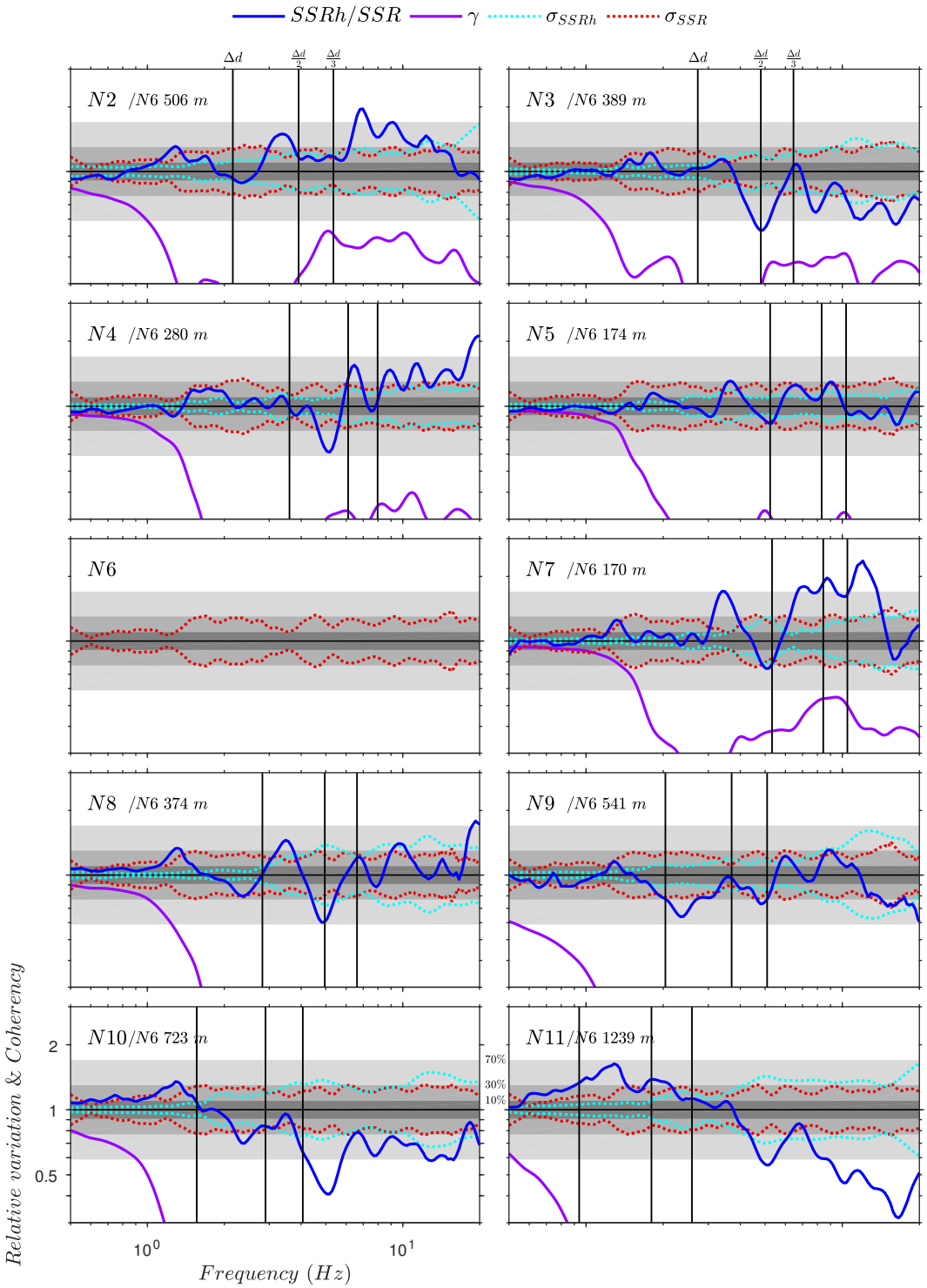


Figure 60. Identical to Figure 59 but in Argostoli at sites N2 to N11 and with N6 the soil reference site.

The time variability of the SSRh is mostly of the same order of magnitude than the inter-event variability of the SSR approach (slightly higher in Provence and lower in Argostoli). However, the SSR variability between earthquakes has different origins (Perron et al. 2015). The mean SSRh appear to be within the SSR uncertainty up to approximately 10 Hz for every site in Provence and 3 Hz in Argostoli. On the whole, σ_{SSRh} tends to increase with increasing frequencies. This increase of the time variability with frequency coincides with the increase of the discrepancies between SSRh and SSR. However, σ_{SSRh} does not always show strong variations when important discrepancy appears between the SSRh and the SSR approaches (e.g., P5 at 7 Hz, N2 at 6 Hz, N7 at 3 Hz).

The time variability of approaches based on the cross-correlation is often attributed to variations in the time of the source distribution. Many authors conclude that the phase and amplitude of the ambient noise cross-correlation are strongly dependent on the evenness of the source distribution (e.g. Stehly et al. 2006; Froment et al. 2010 ; Hanasoge 2013; Hillers et al. 2013; Chen et al. 2016). In the same way, the dependence of the SSRn to the source distribution was often conjectured to be responsible for the unsatisfactory results provided by this approach (see section 4.4.3). This interpretation might give a suitable explanation for the simultaneous increase of σ_{SSRh} and of the discrepancy of the SSRh with increasing frequency as the temporal variability (for short time windows) of the source spatial distribution generally increase with increasing frequency. However, the uniformity of the noise source distribution has not been evaluated in this study.

On both networks, the coherency vanished quickly at low frequency while the amplitude of the SSRh is reliable up to higher frequency. For instance, at site P3 in Provence the SSRh agreed with the SSR up to 12 Hz while the noise is incoherent at every frequency. Thus, it seems that if the phase is coherent then the SSRh amplitude is reliable while the opposite is not true. This could be related to the strong basin response that leads the incoming seismic wavefield to be amplified whatever its origin but this amplification does not correspond to an amplification that occurs in-phase along the path between the two stations.

Because the rock outcrop is generally located far from the site of interest, the site-reference distance criteria is often considered as a critical parameter for the SSRn approach depending on the frequency band considered. Scarce are studies that discuss explicitly SSRn results according to the site-reference distance. Based on circular arrays of different diameters, Horike et al. (2001) conclude that the incoming microtremors are the same only within limited areas of a few hundred meters in diameter. Concerning the SSRh a good agreement with the SSR is observable up to approximately $fh_{\Delta d/3}$ in Provence for site P5, P6 and P7 that are the closest one to the soil reference site. However, in P3 the SSRh is reliable up to approximately 10 Hz while $fh_{\Delta d/3}$ is around 2 Hz. In Argostoli this criterion also give variable results. The discrepancy (>30%) between SSR and SSRh starts approximately between $fh_{\Delta d}$ and $fh_{\Delta d/2}$ on every site but N7 where it starts at lower frequency and N5 where it starts at higher frequency. Our results with the SSRh approach (for sites with similar geological conditions) indicates that microtremors can sometimes deliver reliable information up to high frequency (>10 Hz) and for site-reference distance that is equal to almost two kilometers

(e.g., P3 in Provence). The dependence of the SSRh to Δd is not obvious contrarily to the one of the coherency.

4.9. General Conclusions

The use of the ambient noise has received a growing attention since almost a century and more specifically since the 1980s with an exploding number of studies that proposed and/or tested different technics to assess site effects. The majority of these studies did it at low frequency only (often around the first microseism peak ~7 s) where the ambient noise is mostly composed by coherent Rayleigh wave generates by known and/or distant oceanic sources. Evaluating the site transfer function that is obtained from the SSR for a broadband frequency range is very challenging and has never been successfully realized in our knowledge. This experimental study was conducted in two sedimentary basins (Provence and Argostoli) where velocimeters recorded both earthquakes and ambient noise. We tested different approaches to process the ambient noise in order to assess the relative site transfer function in a broad frequency range. We compared the results to those obtained from the SSR processing of earthquakes recordings.

We did not discuss the theoretical point of view but we made several observations which show that the mean of the relative site response measured inside both sedimentary basin does not depend on the input wave, whatever if it corresponds to earthquakes or to ambient noise, even up to frequencies higher than 10 Hz. These observations support the weak dependence of the site effects to the nature of the excitation, especially for sedimentary basin where the site response dominates the ambient noise wavefield.

The different tested approaches led to different results and interpretations:

- The HVSR is found here, and in agreement with the literature, unable to infer more than f_0. Moreover, it does not provide a reliable estimation of f_0 when 2D or 3D site effects induce broadband amplifications.

- Again in agreement with previous studies, the SSRn shows a wide overestimation (comparatively to the SSR) of the relative amplification factors at high frequency when referring to the rock site located outside the basin. However, the results of the SSRn are similar to the SSR one for frequencies up to f_{0min}. None of the explanations commonly referred to seems to justify the discrepancy between the SSRn and the SSR above f_{0min} here. Indeed, no clear evidence of the influence of the source distribution through the time variability of the SSRn has been found, while the impact of the inter-station distance is not obvious up to the maximal dimension of our two networks (~2 km). Finally, based on the observations that the SSRn with both stations inside the basin gives much better results above f_{0min} than taking the reference station outside it leads us to propose an interpretation for this apparent geological dependence of the microtremors based approaches. We propose that the ambient noise wavefield generated inside and behind (according to the rock-

reference site) the basin at frequency above f_{0min} is reflected in parts by the basin interface due to the velocity contrast and so is not completely seen by the station located outside the basin, at the rock-reference site. This so-called “source hiding” suspected phenomena does not arise with the SSR because earthquake waves arrive generally at the site with a vertical incidence. It gives a suitable explanation for the overestimation of the amplification by the SSRn approaches. This source hiding interpretation may allow to considered new approaches to assess the site effect from the ambient noise if it is confirmed (e.g., A_0 and SSRh).

- The SSRh aim at assessing the spatial variability inside the basin from the SSRn with the reference station located inside the basin. Then, the relative transfer function can be corrected by the SSR estimated between the SSRn reference and the rock reference site located outside the basin or by any other estimation of the site response at the SSRn local reference station. This hybrid approach provides very reliable site transfer function estimation up to high frequency (12 Hz in Provence and 20 Hz in Argostoli).
- The ANIRF approach allows obtaining qualitative results in the time domain that correspond to the propagation expected across the basin. To extract the site component in the frequency domain from the ANIRF, we used a distance station as a virtual source and apply the SSR technique on the generated signal recorded at the rock reference site and at the closest soil site in Argostoli. It provided comparable result with the SSR up to 4 Hz even if f_{0min} = 1.5 Hz. This seems to indicate that considering only the coherent part of the ambient noise wavefield prevents being biased by the source hiding phenomena or any other phenomena explaining the overestimation observed above f_{0min} with the SSRn. However, contrary to the SSRh that provides some good results up to very high frequency (~20 Hz) between distant stations, the coherency decreases rapidly at high frequencies for increasing inter-station distance.

To conclude, we consider that theoretical investigations on the presented approaches and results are necessary. Numerical modeling would be useful to understand the importance of the equipartition assumption, of the evenness of the source distribution and of the source hiding phenomena on the SSRn, SSRh and the ANIRF results for site effect assessment. Further research should also consider other sites to understand to what extent and under which conditions these results may be applicable to and up to which frequency.

4.10. Acknowledgments

We are thankful to all participants of the Cashima, NERA and Sinaps@ projects for providing the Provence and Argostoli resource necessary for this work. A special thanks must be given to Cédric Guyonnet-Benaize, Cécile Cornou, Elise Delavaud, Marc Nicolas, Olivier Sèbe,

Bruno Hernandez, Afifa Imtiaz, Michel Campillo, Philippe Roux, Philippe Guéguen, Hiroshi Kawase and Emmanuel Chaljub for their comments or/and participation to this work.

Conclusions et Perspectives

Évaluer l'aléa sismique sur un site donné consiste à prédire les mouvements du sol associés à un/des scénarios de séismes locaux ou régionaux probables pour une période de temps donnée. Ces mouvements dépendent de la magnitude de la source sismique, de l'atténuation des ondes lors de leur propagation en profondeur et des effets de site localement à la surface. Ces effets de site peuvent induire de fortes variabilités spatiales et inter-évènements qui sont fonction de la géométrie et des propriétés physiques des structures géologiques superficielles, et de la fréquence.

Les enregistrements sismiques sont essentiels pour les études d'aléa sismique, en raison de l'information qu'ils contiennent sur les sources et les propagations d'ondes en profondeur et à la surface. Cette information est traitée de façon statistique au travers des équations de prédiction du mouvement du sol (GMPEs) qui sont établies empiriquement à partir d'une grande variété de séismes enregistrés sur une grande diversité de sites. Lorsque pour une région donnée, les bases de données disponibles ne sont pas suffisantes pour réaliser une GMPEs, il convient alors d'adapter celles établies ailleurs à l'aide des ajustements host-to-target. Ces ajustements sont calculés en pratique à partir de quelques proxys de site tels que V_{S30} et κ_0. Si V_{S30} peut-être évaluée à partir de mesures géophysiques ou géotechniques *in situ*, ce n'est pas le cas de κ_0 qui est traditionnellement mesuré sur les spectres en accélération de séismes modérés à forts (M>3). De plus, ces proxys ne sont pas représentatifs des effets de site associés aux géologies à géométrie 2D-3D (p. ex. bassin sédimentaire, vallée alluviale, etc.), qui complexifient très fortement la propagation des ondes dans les couches superficielles. Dans ce cas, une évaluation spécifique des effets de site est nécessaire.

L'évaluation des effets de site 2D-3D peut être réalisée par simulations numériques, mais ces dernières sont limitées aux basses fréquences (<2-4 Hz) par la résolution des connaissances géotechniques et géophysiques du milieu géologique. À l'inverse, les approches empiriques, telles que la méthode SSR, permettent une évaluation large-bande fiable de la réponse relative des sites. L'enregistrement des tremblements de Terre dans les régions à sismicité faible à modérée est donc un enjeu majeur pour l'amélioration des évaluations de l'aléa sismique, et ce particulièrement pour les sites à géométrie 2D-3D.

Ce travail de thèse apporte des éléments sur l'applicabilité d'une mesure fiable de κ_0 et des méthodes empiriques d'évaluation des effets de site dans les régions de sismicité faible à modérée.

Les séismes de magnitude modérée à forte (M>3) ont une période de retour de l'ordre de plusieurs mois à plusieurs années dans les régions faiblement actives. Pour les séismes les plus forts (M>6), celle-ci peut même être de plusieurs siècles. Dans ces régions, il convient donc d'utiliser les enregistrements de séismes locaux de faible voire de très faible magnitude (M<2) qui ont des périodes de retour beaucoup plus courtes comme le décrit la loi de Gutenberg-Richter, ainsi que la quinzaine de téléséismes (R_e>3000 km et M>7) qui sont répertoriés et enregistrables partout dans le monde chaque année en moyenne. Ces évènements génèrent des mouvements de très faible amplitude au niveau du site qui ne sont détectables que si le niveau de bruit local est bas et les instruments suffisamment sensibles. En aléa sismique, historiquement, les accéléromètres sont plébiscités du fait qu'ils ne saturent pas et permettent ainsi d'enregistrer les mouvements forts induits par les évènements locaux de forte magnitude. En effet, ce sont ces derniers qui intéressent le plus les ingénieurs pour les études de tenue des infrastructures. Cependant, les accéléromètres paramétré pour enregistrer les mouvements forts sont peu sensibles et ne permettent pas l'enregistrement des mouvements les plus faibles. Nos résultats montrent, par exemple, que l'utilisation de vélocimètres large-bande permet l'enregistrement de bonne qualité (SNR>3) de 10 à 50 fois plus de séismes à basse fréquence (<5 Hz) que celle d'accéléromètres. Ainsi, sur un site industriel du sud-est de la France (Provence) pour lequel le niveau de bruit local est plutôt bas, ce sont plusieurs centaines de séismes qui ont pu être extraits (avec SNR>3 à au moins une fréquence) en l'espace de deux ans et demi grâce à une instrumentation vélocimétrique enregistrant en continu. Bien que cette base de données soit sans commune mesure en quantité et en qualité d'enregistrements avec celle constituée en un temps plus court dans le contexte de forte activité post-sismique à Argostoli, elle est suffisante pour réaliser les évaluations empiriques des effets de site qui sont classiquement effectuées dans les zones de sismicité modérée à forte.

L'applicabilité des approches empiriques en France nous permet de proposer une méthodologie d'évaluation de l'aléa spécifique aux sites situés dans les zones de sismicité faible à modérée qui tienne compte de façon plus fidèle des effets de site. Celle-ci se décompose en trois étapes qui correspondent aux chapitres 2 à 4 dont les principaux résultats sont rappelés ici :

- Les GMPEs ne peuvent pas être calculées jusqu'aux fortes magnitudes dans les zones de sismicité faible à modérée où les évènements majeurs (M>6) ont une période de récurrence généralement supérieure au siècle. Les ajustements host-to-target sont donc requis et réclament l'évaluation empirique de κ_0 qui retranscrit alors le terme d'atténuation propre au site (atténuation considérée comme indépendante de la fréquence). La physique sous-jacente n'est cependant pas claire et les influences de la source, de l'amplification de site ainsi que de la part dépendante en fréquence de l'atténuation sont probables. κ_0 est classiquement mesuré à haute fréquence sur les

spectres en accélération (κ_{0_AS}), ce qui exige l'enregistrement de nombreux séismes locaux de magnitude modérée à forte (M>3). La mesure de κ_0 à plus basse fréquence sur les spectres en déplacement (κ_{0_DS}) s'applique, à l'inverse, sur des séismes de très petites magnitudes (M<1,5). Localement (R_e<150 km), la base de données de Provence est principalement constituée d'évènements de magnitude intermédiaire à ces deux approches (1,5<M<3), qui sont ceux présents dans les catalogues sismiques. Ceci a pour conséquence que la bande de fréquence disponible pour mesurer κ_{r} est plus étroite dans les zones à sismicité faible à modérée. Ainsi, la mesure de κ_{r} est plus incertaine encore, car plus sensible à de petites modifications du spectre induites par la source ou par l'amplification de site. De plus, l'hypothèse d'indépendance en fréquence de κ semble incorrecte au vu de nos résultats, bien que son influence sur κ_0 ne soit pas clairement quantifiable. La composante de source de κ contrôle très distinctement la dispersion des valeurs de κ_{r} sans modifier notablement les valeurs moyennes de κ_0. Finalement, nos résultats montrent qu'une évaluation fiable de κ_0 n'est pas possible pour les sites sédimentaires à cause de l'interaction avec l'amplification de site. Pour les sites au rocher, les effets de site interfèrent beaucoup moins, même si une influence est encore visible. Cette influence résiduelle est moins pénalisante pour κ_{0_DS} à basse fréquence où les effets de site au rocher sont moins forts et où les profils de vitesse générique peuvent être utilisés pour en corriger l'effet. Pour les deux sites au rocher et suivant les deux approches testées, nos résultats sont en accord (κ_0~0,03 s) ce qui confirme qu'il est possible de réaliser une mesure empirique de ce paramètre même en zone de sismicité faible à modérée. Dans ces zones, les corrélations très incertaines de κ_0 avec V_{S30} sont souvent utilisées en l'absence de mesure de κ_0. Celles-ci sont à proscrire et devraient être substituées par l'approche κ_{DS} qui est peu sensible à la source et aux effets de site quand elle est réalisée sur des sites au rocher. Plus généralement, nous encourageons très fortement le développement des GMPEs et l'application des ajustements host-to-target exclusivement au rocher où les effets de site sont plus faibles et plutôt à géométrie 1D. Ainsi, ils sont représentés de façon beaucoup plus réaliste par les proxys V_{S30} et κ_0 et la mesure de κ_0 est bien plus fiable. Ceci ne peut se faire qu'avec un développement simultané des réseaux accéléromètriques au rocher dans les zones fortement sismiques où ils manquent encore cruellement. La première étape de notre méthodologie consiste donc à réaliser les ajustements host-to-target avec V_{S30} et κ_{0_DS} à partir d'une GMPEs si possible développée au rocher vers un site au rocher situé à proximité du site sédimentaire que l'on cherche à étudier.

- Il convient ensuite de transposer la GMPEs ajustée au rocher vers les sites sédimentaires dont on veut connaitre l'aléa. Pour ce faire la méthode SSR est tout à fait adaptée en ce sens qu'elle fournit la fonction de transfert empirique relative entre le site de référence au rocher où la GMPEs a été ajustée et les sites sédimentaires considérés. Cette fonction de transfert a l'avantage de pourvoir aussi bien la réponse relative moyenne des sites, que l'incertitude associée, et ce, quelle que soit la géométrie contrôlant l'effet de site que l'on cherche à caractériser (1D, 2D ou 3D). Comme pour la mesure du paramètre κ_0, cette approche nécessite pour chaque site de mesure

l'enregistrement d'un nombre statistiquement représentatif d'évènements. Si 8 évènements semblent suffisants pour caractériser la distribution log-normale des fonctions de transfert SSR, il en faut un nombre beaucoup plus important pour quantifier complètement l'incertitude inter-évènement qui leurs est associées, en particulier dans le cas de structures géologiques présentant des géométries 2D ou 3D. Il faudrait alors idéalement une couverture spatiale des sources (back-azimut, profondeur, et distance épicentrale) complète et homogène. Effectivement, nos résultats montrent que l'éclairage qui est donné au bassin par la source influence fortement sa réponse (jusqu'à plus de 100 %). Cette incertitude est épistémique et pourrait être retranchée en tenant compte de la position relative du séisme vis-à-vis du bassin dans les évaluations de la réponse du site. Il ne resterait alors que la variabilité réellement aléatoire qui semble être de l'ordre de 20 % seulement. À l'inverse, si la distribution spatiale des séismes utilisés pour estimer l'incertitude totale n'est pas homogène alors celle-ci est très certainement sous-estimée lorsque la géométrie des structures est 2D ou 3D. Cela est pourtant rarement pris en compte dans les évaluations empiriques des effets de site. La simulation numérique peut alors être requise pour combler ce manque, au moins à basse fréquence. Similairement, seule la simulation numérique permet de prédire la réponse à de fortes sollicitations des sols mous aux comportements non-linéaires. Les fonctions de transfert linéaires empiriques restent cependant un prérequis incontournable pour caler les modèles numériques.

- Lorsque la zone d'étude est étendue (complexe industriel, ville, etc.) et/ou que l'effet de site est très variable spatialement, alors son évaluation empirique doit être réalisée selon un maillage adapté. Il n'est pas envisageable d'employer la méthode SSR pour faire du microzonage, sauf à déployer un très grand nombre de stations pendant un temps relativement long (de plusieurs mois à plusieurs décennies). À l'inverse, l'utilisation du bruit ambiant permet la réalisation de campagnes de mesures relativement courtes (de l'ordre de quelques jours à quelques semaines). Cependant, les méthodes utilisant le bruit ambiant n'ont, jusqu'à présent, pas été en mesure de fournir la fonction de transfert des sites jusqu'aux hautes fréquences (>1-4 Hz). Nos résultats, en accord avec la littérature, montrent que la méthode HVSR ne permet pas d'identifier plus que f_0, tandis que la méthode SSRn surestime grandement l'amplification de site à haute fréquence lorsqu'elle utilise le rocher à l'extérieur du bassin comme référence. Étonnamment, cette dernière donne des résultats très similaires à ceux du SSR jusqu'à haute fréquence (12-20 Hz) lorsque le site de référence est pris dans le bassin, donc dans des conditions de site similaires au site d'étude. Nous pensons que ceci peut être dû à la réflexion d'une partie des ondes de surface du champ sismique ambiant par le contraste d'impédance marqué entre le sédiment et le rocher. Ainsi, une partie de l'énergie générée par les sources de bruit n'atteint pas le rocher avoisinant (« source hiding »). Cette interprétation et nos observations nous incitent à proposer la méthode hybride, SSRh, qui consiste à se référer au rocher avec le SSR puis à cartographier la variabilité spatiale de l'effet de site à partir du SSRn. Cette méthode est très prometteuse pour le microzonage des effets de site à l'intérieur des bassins. Cependant, pour certains sites ne bénéficiant pas

d'enregistrement de séismes, des solutions doivent être envisagées pour établir la fonction de transfert vers le rocher exclusivement à partir du bruit ambiant. L'approche ANIRF, qui utilise les corrélations de bruit, montre des résultats très encourageants à basse fréquence (<4 Hz) lorsque les sources virtuelles sont utilisées en remplacement des séismes dans le calcul de la fonction de transfert SSR. L'ensemble de ces résultats dénote le grand potentiel des méthodes de bruit ambiant pour l'évaluation des effets de site et le microzonage dans les années à venir.

Finalement, cette méthodologie permet d'envisager la réalisation d'évaluations d'aléa sismique spécifique à un site, qui tiennent compte de toute la complexité et la variabilité des effets de site pour l'ensemble du bassin. Celle-ci requiert simplement l'enregistrement de séismes sur une longue période pour un minimum absolu de deux stations seulement (un site de référence au rocher et un site sédimentaire), ainsi que la réalisation d'une campagne temporaire d'enregistrements du bruit ambiant si nécessaire. Il reste que pour avoir une information complète sur la variabilité du phénomène pour tous les back-azimuts de source et tenant compte des effets non-linéaires, des simulations 3D et/ou non-linéaires doivent être réalisées, mais toujours sur la base de résultat empirique linéaire afin d'être ancré au plus près de la réalité du phénomène.

Dans les années à venir, compte tenu de la nécessité absolue d'arrimer les résultats de la simulation numérique sur les mesures instrumentales, l'amélioration des prédictions de mouvement du sol passera nécessairement par le déploiement de nouveaux instruments. Tout d'abord d'accéléromètres sur les sites au rocher et rocher très dur dans les zones sismiquement actives pour améliorer la représentativité des GMPEs pour ces classes de sites. Dans le même temps, la mise en place d'une instrumentation vélocimétrique large-bande enregistrant en continu conjointement avec l'instrumentation accéléromètrique sur les sites d'intérêt (centrale nucléaire, barrage, etc.) dans les régions faiblement actives est d'une nécessité absolue pour envisager l'application des méthodes empiriques d'évaluation des effets de site. Ces méthodes empiriques doivent encore être améliorées pour envisager un jour de s'acquitter de l'enregistrement de séismes, en particulier par la méthode ANIRF qui est très prometteuse. De même, l'information contenue dans la coda des enregistrements sismiques devrait faire l'objet d'une recherche plus approfondie pour permettre l'évaluation relative des effets de site entre des sites distants les uns des autres (Sèbe 2004). Ceci est essentiel pour les sites d'études ne bénéficiant pas d'affleurements rocheux à proximité pouvant servir de référence. De plus, ce champ d'étude est également très prometteur pour l'amélioration des inversions généralisées, permettant d'envisager une séparation plus fiable des composantes de source, de trajet et de site du mouvement sismique.

Bibliographie

Aki K (1957) Space and time spectra of stationary stochastic waves, with special reference to microtremors. Bull Earthq Res Inst 35:415–456

Aki K (1993) Local site effects on weak and strong ground motion. Tectonophysics 218:93–111

Aki K (1967) Scaling law of seismic spectrum. J Geophys Res 72:1217–1231 . doi: 10.1029/JZ072i004p01217

Aki K (1987) Magnitude-frequency relation for small earthquakes: A clue to the origin of fmax of large earthquakes. J Geophys Res 92:1355

Aki K (1980) Scattering and attenuation of shear waves in the lithosphere. J Geophys Res Solid Earth 85:6496–6504 . doi: 10.1029/JB085iB11p06496

Aki K (1988) Local site effects on ground motion. Earthq Eng Soil Dyn II-Recent Adv Ground Motion Eval Geotech Spec Publ 20:103–155

Aki K, Chouet B (1975) Origin of coda waves: source, attenuation, and scattering effects. J Geophys Res 80:3322–3342

Aki K, Richards PG (1980) Quantitative Seismology

Ameri G, Hollender F, Perron V, Martin C (2017) Site-specific partially nonergodic PSHA for a hard-rock critical site in southern France: adjustment of ground motion prediction equations and sensitivity analysis. Bull Earthq Eng. doi: 10.1007/s10518-017-0118-6

Anderson JG (2007) Physical processes that control strong ground motion. In: Treatise on geophysics, earthquake seismology. pp 513–565

Anderson JG (1986) Implication of Attenuation for Studies of the Earthquake Source. In: Das S, Boatwright J, Scholz CH (eds) Earthquake Source Mechanics. American Geophysical Union, pp 311–318

Anderson JG, Hough SE (1984) A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies. Bull Seismol Soc Am 74:1969–1993

Anderson JG, Humphrey JR (1991) A least squares method for objective determination of earthquake source parameters. Seismol Res Lett 62:201–209

Anderson JG, Lee Y, Zeng Y, Day S (1996) Control of strong motion by the upper 30 meters. Bull Seismol Soc Am 86:1749–1759

ASN (2001) Règle fondamentale de sûreté n°2001-01 relative à la détermination du risque sismique pour la sûreté des installations nucléaires de base de surface

ASN (1984) RÈGLE N° I.3.b relative à l'Instrumentation sismique

Atakan K (1997) Empirical site response studies in Central America: Present status. In: Proceedings, Seminar on Evaluation and Mitigation of Seismic Risk in the Central American Area, Universidad Centroamericana "Jose Simeón Cañas" San Salvador, El Salvador. pp 22–26

Atkinson GM (2006) Single-Station Sigma. Bull Seismol Soc Am 96:446–455 . doi: 10.1785/0120050137

Baer M, Kradolfer U (1987) An automatic phase picker for local and teleseismic events. Bull Seismol Soc Am 77:1437–1445

Baig AM, Campillo M, Brenguier F (2009) Denoising seismic noise cross correlations. J Geophys Res Solid Earth 114:B08310 . doi: 10.1029/2008JB006085

Bard P-Y (1999) Microtremor Measurements: A tool for site effect estimation? In: ESG98 symposium. (ed. Irikura, Kudo, Okada and Sasatani), pp 1251–1279

Bensen GD, Ritzwoller MH, Barmin MP, et al (2007) Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. Geophys J Int 169:1239–1260 . doi: 10.1111/j.1365-246X.2007.03374.x

Beresnev IA, Atkinson GM (1997) Modeling finite-fault radiation from the $ømega$n spectrum. Bull Seismol Soc Am 87:67–84

Berge-Thierry C, Svay A, Laurendeau A, et al (2016) Toward an integrated seismic risk assessment for nuclear safety improving current french methodologies through the SINAPS@ research project. Nucl Eng Des. doi: 10.1016/j.nucengdes.2016.07.004

Biasi GP, Smith KD (2001) Site effects for seismic monitoring stations in the vicinity of Yucca Mountain. Nevada

Bindi D, Parolai S, Cara F, et al (2009) Site Amplifications Observed in the Gubbio Basin, Central Italy: Hints for Lateral Propagation Effects. Bull Seismol Soc Am 99:741–760 . doi: 10.1785/0120080238

Bonilla LF, Steidl JH, Lindley GT, et al (1997) Site amplification in the San Fernando Valley, California: variability of site-effect estimation using the S-wave, coda, and H/V methods. Bull Seismol Soc Am 87:710–730

Bonnefoy-Claudet S, Cotton F, Bard P-Y (2006) The nature of noise wavefield and its applications for site effects studies. Earth-Sci Rev 79:205–227 . doi: 10.1016/j.earscirev.2006.07.004

Boore DM (2006) Orientation-Independent Measures of Ground Motion. Bull Seismol Soc Am 96:1502–1511 . doi: 10.1785/0120050209

Boore DM (2004) Can site response be predicted? J Earthq Eng 8:1–41

Boore DM (1986) Short-period P-and S-wave radiation from large earthquakes: implications for spectral scaling relations. Bull Seismol Soc Am 76:43–64

Boore DM (2003) Simulation of ground motion using the stochastic method. Pure Appl Geophys 160:635–676

Boore DM, Campbell KW (2017) Adjusting Central and Eastern North America Ground-Motion Intensity Measures between Sites with Different Reference-Rock Site Conditions. Bull Seismol Soc Am 107:132–148 . doi: 10.1785/0120160208

Boore DM, Gibbs JF, Joyner WB, et al (2003) Estimated Ground Motion From the 1994 Northridge, California, Earthquake at the Site of the Interstate 10 and La Cienega Boulevard Bridge Collapse, West Los Angeles, California. Bull Seismol Soc Am 93:2737–2751 . doi: 10.1785/0120020197

Borcherdt RD (1970) Effects of local geology on ground motion near San Francisco Bay. Bull Seismol Soc Am 60:29–61

Boué P, Denolle M, Hirata N, et al (2016) Beyond basin resonance: characterizing wave propagation using a dense array and the ambient seismic field. Geophys J Int 206:1261–1272 . doi: 10.1093/gji/ggw205

Bowden DC, Tsai VC, Lin FC (2015) Site amplification, attenuation, and scattering from noise correlation amplitudes across a dense array in Long Beach, CA. Geophys Res Lett 42:1360–1367 . doi: 10.1002/2014GL062662

Brune JN (1970) Tectonic stress and the spectra of seismic shear waves from earthquakes. J Geophys Res 75:4997–5009

Burjánek J, Moore JR, Yugsi Molina FX, Fäh D (2012) Instrumental evidence of normal mode rock slope vibration: Evidence of normal mode rock slope vibration. Geophys J Int 188:559–569 . doi: 10.1111/j.1365-246X.2011.05272.x

Cadet H, Bard P-Y, Duval A-M, Bertrand E (2012) Site effect assessment using KiK-net data: part 2—site amplification prediction equation based on f0 and Vsz. Bull Earthq Eng 10:451–489 . doi: 10.1007/s10518-011-9298-7

Calvet M, Sylvander M, Margerin L, Villaseñor A (2013) Spatial variations of seismic attenuation and heterogeneity in the Pyrenees: Coda Q and peak delay time analysis. Tectonophysics 608:428–439 . doi: 10.1016/j.tecto.2013.08.045

Campbell KW (2009) Estimates of Shear-Wave Q and 0 for Unconsolidated and Semiconsolidated Sediments in Eastern North America. Bull Seismol Soc Am 99:2365–2392 . doi: 10.1785/0120080116

Campbell KW (2003) Prediction of strong ground motion using the hybrid empirical method and its use in the development of ground-motion (attenuation) relations in eastern North America. Bull Seismol Soc Am 93:1012–1033

Campbell KW (2004) Erratum: Prediction of Strong Ground Motion Using the Hybrid Empirical Method and Its Use in the Development of Ground-Motion (Attenuation) Relations in Eastern North America. Bull Seismol Soc Am 94:2418–2418 . doi: 10.1785/0120040148

Campillo M, Paul A (2003) Long-Range Correlations in the Diffuse Seismic Coda. Science 299:547–549 . doi: 10.1126/science.1078551

Celebi M, Dietel C, Prince J, et al (1987) Site amplification in Mexico City (determined from 19 September 1985 strong-motion records and from recordings of weak motions). Elsevier, Amsterdam

Chaljub E, Maufroy E, Moczo P, et al (2015) 3-D numerical simulations of earthquake ground motion in sedimentary basins: testing accuracy through stringent models. Geophys J Int 201:90–111 . doi: 10.1093/gji/ggu472

Chaljub E, Moczo P, Tsuno S, et al (2010) Quantitative Comparison of Four Numerical Predictions of 3D Ground Motion in the Grenoble Valley, France. Bull Seismol Soc Am 100:1427–1455 . doi: 10.1785/0120090052

Chen Z, Gerstoft P, Bromirski PD (2016) Microseism source direction from noise cross-correlation. Geophys J Int 205:810–818 . doi: 10.1093/gji/ggw055

Claerbout J (1968) Synthesis of a layered medium from its acoustic transmission response. GEOPHYSICS 33:264–269 . doi: 10.1190/1.1439927

Cormier VF (1982) The effect of attenuation on seismic body waves. Bull Seismol Soc Am 72:S169–S200

Cornou C, Bard P-Y (2003) Site-to-bedrock over 1D transfer function ratio: An indicator of the proportion of edge-generated surface waves? Geophys Res Lett 30: . doi: 10.1029/2002GL016593

Cotton F, Scherbaum F, Bommer JJ, Bungum H (2006) Criteria for Selecting and Adjusting Ground-Motion Models for Specific Target Regions: Application to Central Europe and Rock Sites. J Seismol 10:137–156 . doi: 10.1007/s10950-005-9006-7

Cultrera G, De Rubeis V, Theodoulidis N, et al (2014) Statistical correlation of earthquake and ambient noise spectral ratios. Bull Earthq Eng. doi: 10.1007/s10518-013-9576-7

Cupillard P, Capdeville Y (2010) On the amplitude of surface waves obtained by noise correlation and the capability to recover the attenuation: a numerical approach. Geophys J Int 181:1687–1700 . doi: 10.1111/j.1365-246X.2010.04586.x

Cupillard P, Stehly L, Romanowicz B (2011) The one-bit noise correlation: a theory based on the concepts of coherent and incoherent noise. Geophys J Int 184:1397–1414 . doi: 10.1111/j.1365-246X.2010.04923.x

Cushing EM, Hollender F, Guyonnet-Benaize C, et al (2016) Close to the lair of Odysseus Cyclops : the SINAPS@ postseismic campaign and accelerometric network installation on Kefalonia island – Site effect characterization experiment. In: Active Tectonics and Archeoseismology (PATA). Crestone, Colorado, USA

Dainty AM (1981) A scattering model to explain seismic Q observations in the lithosphere between 1 and 30 Hz. Geophys Res Lett 8:1126–1128

Davis PM, Rubinstein JL, Liu KH, et al (2000) Northridge Earthquake Damage Caused by Geologic Focusing of Seismic Waves. Science 289:1746–1750 . doi: 10.1126/science.289.5485.1746

Delavaud E, Cotton F, Akkar S, et al (2012) Toward a ground-motion logic tree for probabilistic seismic hazard assessment in Europe. J Seismol 16:451–473 . doi: 10.1007/s10950-012-9281-z

Denolle MA, Dunham EM, Prieto GA, Beroza GC (2013) Ground motion prediction of realistic earthquake sources using the ambient seismic field. J Geophys Res Solid Earth 118:2102–2118 . doi: 10.1029/2012JB009603

Denolle MA, Dunham EM, Prieto GA, Beroza GC (2014a) Strong Ground Motion Prediction Using Virtual Earthquakes. Science 343:399–403 . doi: 10.1126/science.1245678

Denolle MA, Miyake H, Nakagawa S, et al (2014b) Long-period seismic amplification in the Kanto Basin from the ambient seismic field. Geophys Res Lett 41:2319–2325 . doi: 10.1002/2014GL059425

Derode A, Larose E, Tanter M, et al (2003) Recovering the Green's function from field-field correlations in an open scattering medium (L). J Acoust Soc Am 113:2973–2976 . doi: 10.1121/1.1570436

Douglas J, Gehl P, Bonilla LF, Gélis C (2010) A κ model for mainland France. Pure Appl Geophys 167:1303–1315 . doi: 10.1007/s00024-010-0146-5

Dravinski M, Ding G, Wen K-L (1996) Analysis of spectral ratios for estimating ground motion in deep basins. Bull Seismol Soc Am 86:646–654

Drouet S (2006) Analyse des données accélérométriques pour la caractérisation de l'aléa sismique en France métropolitaine

Drouet S, Cotton F, Guéguen P (2010) Vs30, κ, regional attenuation and Mw from accelerograms: application to magnitude 3-5 French earthquakes: vS30, κ, attenuation and Mw from accelerograms. Geophys J Int 182:880–898 . doi: 10.1111/j.1365-246X.2010.04626.x

Duval A-M, Bertrand E, Vidal S, Delgado J (2013) Détection des effets de site sismiques : mise au point de méthodes expérimentales et application à Nice. Bull Lab Ponts Chaussées 3–20

Edwards B, Fäh D, Giardini D (2011) Attenuation of seismic shear wave energy in Switzerland: Seismic attenuation in Switzerland. Geophys J Int 185:967–984 . doi: 10.1111/j.1365-246X.2011.04987.x

Edwards B, Ktenidou O-J, Cotton F, et al (2015) Epistemic uncertainty and limitations of the κ_0 model for near-surface attenuation at hard rock sites. Geophys J Int 202:1627–1645 . doi: 10.1093/gji/ggv222

Ermert L, Villaseñor A, Fichtner A (2016) Cross-correlation imaging of ambient noise sources. Geophys J Int 204:347–364 . doi: 10.1093/gji/ggv460

Eva C, Cattaneo M, Augliera P, Pasta M (1991) Regional coda Q variations in the western Alps (northern Italy). Phys Earth Planet Inter 76–86

Faccioli E, Vanini M (2003) Complex" site effects in earthquake ground motion, including topography. Earthq Eng Struct Dyn 5:223–238

Ferritto J (1995) Ground Motion Amplification Using Microseisms. Int Conf Recent Adv Geotech Earthq Eng Soil Dyn

Ferritto JM (1996) Repeatability of microseism measurements in Port Hueneme case study. Bull Seismol Soc Am 86:428–435

Field EH (1996) Spectral amplification in a sediment-filled valley exhibiting clear basin-edge-induced waves. Bull Seismol Soc Am 86:991–1005

Field EH, Clement AC, Jacob KH, et al (1995) Earthquake site-response study in Giumri (formerly Leninakan), Armenia, using ambient noise observations. Bull Seismol Soc Am 85:349–353

Field EH, Hough SE, Jacob KH (1990) Using microtremors to assess potential earthquake site response: a case study in Flushing Meadows, New York City. Bull Seismol Soc Am 80:1456–1480

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

Field EH, Jacob KH (1993) The theoretical response of sedimentary layers to ambient seismic noise. Geophys Res Lett 20:2925–2928

Froment B, Campillo M, Roux P, et al (2010) Estimation of the effect of nonisotropically distributed energy on the apparent arrival time in correlations. GEOPHYSICS 75:SA85-SA93 . doi: 10.1190/1.3483102

Futterman WI (1962) Dispersive body waves. J Geophys Res 67:5279–5291 . doi: 10.1029/JZ067i013p05279

Gariel J-C, Campillo M (1989) The influence of the source on the high-frequency behavior of the near-Field acceleration spectrum: A numerical study. Geophys Res Lett 16:279–282

Garofalo F, Foti S, Hollender F, et al (2016) InterPACIFIC project: Comparison of invasive and non-invasive methods for seismic site characterization. Part II: Inter-comparison between surface-wave and borehole methods. Soil Dyn Earthq Eng 82:241–254 . doi: 10.1016/j.soildyn.2015.12.009

Gaull BA, Kagami H, Taniguchi H (1995) The Microzonation of Perth, Western Australia, Using Microtremor Spectral Ratios. Earthq Spectra 11:173–191 . doi: 10.1193/1.1585810

Gitterman Y, Zaslavsky Y, Shapira A, Shtivelman V (1996) Empirical site response evaluations: case studies in Israel. Soil Dyn Earthq Eng 15:447–463 . doi: 0267-7261/96/$15.00

Graves RW, Pitarka A (2010) Broadband Ground-Motion Simulation Using a Hybrid Approach. Bull Seismol Soc Am 100:2095–2123 . doi: 10.1785/0120100057

Guéguen P, Cornou C, Garambois S, Banton J (2007a) On the Limitation of the H/V Spectral Ratio Using Seismic Noise as an Exploration Tool: Application to the Grenoble Valley (France), a Small Apex Ratio Basin. Pure Appl Geophys 164:115–134 . doi: 10.1007/s00024-006-0151-x

Guéguen P, Michel C, LeCorre L (2007b) A simplified approach for vulnerability assessment in moderate-to-low seismic hazard regions: application to Grenoble (France). Bull Earthq Eng 5:467–490 . doi: 10.1007/s10518-007-9036-3

Gutenberg B (1927) Grundlagen der erdbebenkunde. Gebrüder Borntraeger

Gutierrez C, Singh SK (1992) A site effect study in Acapulco, Guerrero, Mexico: comparison of results from strong-motion and microtremor data. Bull Seismol Soc Am 82:642–659

Haghshenas E, Bard P-Y, Theodulidis N, Team SW (2008) Empirical evaluation of microtremor H/V spectral ratio. Bull Earthq Eng 6:75–108 . doi: 10.1007/s10518-007-9058-x

Haile M, Seo K, Kurita K, et al (1997) Study of Site Effects in Kobe Area Using Microtremors. J Phys Earth 45:121–133

Hanasoge SM (2013) The influence of noise sources on cross-correlation amplitudes. Geophys J Int 192:295–309 . doi: 10.1093/gji/ggs015

Hanks TC (1982) fmax. Bull Seismol Soc Am 72:1867–1879

Hanks TC, Kanamori H (1979) A moment magnitude scale. J Geophys Res B 84:2348–2350

Hennino R, Trégourès N, Shapiro NM, et al (2001) Observation of Equipartition of Seismic Waves. Phys Rev Lett 86:3447–3450 . doi: 10.1103/PhysRevLett.86.3447

Hillers G, Ben-Zion Y, Landès M, Campillo M (2013) Interaction of microseisms with crustal heterogeneity: A case study from the San Jacinto fault zone area. Geochem Geophys Geosystems 14:2182–2197 . doi: 10.1002/ggge.20140

Hollender F, Perron V, Imtiaz A, et al (2015) Close to the lair of Odysseus Cyclops: the SINAPS@ post-seismic campaign and accelerometric network installation on Kefalonia island. In: 9ème Colloque National AFPS. Marne-la-Vallée

Horike M, Zhao B, Kawase H (2001) Comparison of site response characteristics inferred from microtremors and earthquake shear waves. Bull Seismol Soc Am 91:1526–1536

Hough SE, Anderson JG (1988) High-frequency spectra observed at Anza, California: implications for Q structure. Bull Seismol Soc Am 78:692–707

Hough SE, Anderson JG, Brune J, et al (1988) Attenuation near Anza, California. Bull Seismol Soc Am 78:672–691

Hough SE, Lees JM, Monastero F (1999) Attenuation and source properties at the Coso Geothermal Area, California. Bull Seismol Soc Am 89:1606–1619

Ibs-von Seht M, Wohlenberg J (1999) Microtremor measurements used to map thickness of soft sediments. Bull Seismol Soc Am 89:250–259

Imtiaz A (2015) Champ d'ondes, variabilité spatiale et cohérence des mouvements sismiques: effets en champ proche et en vallée alluviale. Doctoral thesis, Université Grenoble Alpes

Imtiaz A, Perron V, Svay A, et al (2017) Wavefield characteristics and coherency of seismic ground motion from a rock site at Argostoli, Greece. In: 16th World Conference on Earthquake Engineering (16WCEE)

Irikura K, Kawanaka T (1980) Characteristics of microtremors on ground with discontinuous underground structure. Bull Disas Prev Res Inst 30:81–96

Ishimoto M (1937) Observations sur des secousses d'une petite amplitude. Bull Earthq Res Inst 15:697–705

Jarpe SP, Cramer CH, Tucker BE, Shakal AF (1988) A comparison of observations of ground response to weak and strong ground motion at Coalinga, California. Bull Seismol Soc Am 78:421–435

Jongmans D, Campillo M (1993) The response of the Ubaye Valley (France) for incident SH and SV waves: Comparison between measurements and modeling. Bull Seismol Soc Am 83:907–924

Kagami H, Duke CM, Liang GC, Ohta Y (1982) Observation of 1-to 5-second microtremors and their application to earthquake engineering. Part II. Evaluation of site effect upon seismic wave amplification due to extremely deep soil deposits. Bull Seismol Soc Am 72:987–998

Kagami H, Okada S, Shiono K, et al (1986) Observation of 1-to 5-second microtremors and their application to earthquake engineering. Part III. A two-dimensional study of site effects in the San Fernando Valley. Bull Seismol Soc Am 76:1801–1812

Kanai K (1983) Engineering seismology. Tokyo

Kanai K, Suzuki T (1955) Relation between the Property of Building Vibration and the Nature of Ground. Bull Earthq Res Inst 33:109–120

Kanai K, Tanaka T (1961) On Microtremors 8. Bull Earthq Res Inst 39:97–114

Kanai K, Tanaka T, Osada K (1954) Measurement of the Micro-tremor. I. Bull Earthq Res Inst 32:199–209

Kennett BLN (1974) Reflections, rays, and reverberations. Bull Seismol Soc Am 64:1685–1696

Kilb D, Biasi G, Anderson J, et al (2012) A Comparison of Spectral Parameter Kappa from Small and Moderate Earthquakes Using Southern California ANZA Seismic Network Data. Bull Seismol Soc Am 102:284–300 . doi: 10.1785/0120100309

King JL, Tucker BE (1984) Observed variations of earthquake motion across a sediment-filled valley. Bull Seismol Soc Am 74:137–151

Knopoff L (1964) Q Rev. Rev Geophys Space Phys 2:625–660

Konno K, Ohmachi T (1998) Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor. Bull Seismol Soc Am 88:228–241

Kottke AR (2017) VS30-κ0 relationship implied by ground motion models? In: 16th World Conference on Earthquake Engineering (16WCEE). Santiago, Chile

Koyama S, Okama I, Kashima T, Kitagawa Y (1996) Characteristics of ground motion in Kobé area. Elsevier Science Ltd, Acapulco

Kramer SL (1996) Geotechnical earthquake engineering. Pearson Education India

Ktenidou O-J, Abrahamson NA (2016) Empirical Estimation of High-Frequency Ground Motion on Hard Rock. Seismol Res Lett. doi: 10.1785/0220160075

Ktenidou O-J, Abrahamson NA, Drouet S, Cotton F (2015) Understanding the physics of kappa (κ): insights from a downhole array. Geophys J Int 203:678–691 . doi: 10.1093/gji/ggv315

Ktenidou O-J, Chavez-Garcia FJ, Pitilakis KD (2011) Variance Reduction and Signal-to-Noise Ratio: Reducing Uncertainty in Spectral Ratios. Bull Seismol Soc Am 101:619–634 . doi: 10.1785/0120100036

Ktenidou O-J, Chávez-García F-J, Raptakis D, Pitilakis KD (2016) Directional dependence of site effects observed near a basin edge at Aegion, Greece. Bull Earthq Eng 14:623–645 . doi: 10.1007/s10518-015-9843-x

Ktenidou O-J, Cotton F, Abrahamson NA, Anderson JG (2014) Taxonomy of Kappa : a review of definitions and estimation approaches targeted to applications. Seismol Res Lett 85:135–146 . doi: 10.1785/0220130027

Ktenidou O-J, Gelis C, Bonilla L-F (2013) A Study on the Variability of Kappa in a Borehole: Implications of the Computation Process. Bull Seismol Soc Am 103:1048–1068 . doi: 10.1785/0120120093

Ktenidou O-J, Silva WJ, Darragh RB, et al (2017) Squeezing Kappa (κ) Out of the Transportable Array: A Strategy for Using Bandlimited Data in Regions of Sparse Seismicity. Bull Seismol Soc Am 107:256–275 . doi: 10.1785/0120150301

Kudo K (1995) Practical estimates of site response. State-of-the-art report. In: Proceedings of the fifth International Conference on Seismic Zonation. Nice, France

Küperkoch L, Meier T, Lee J, et al (2010) Automated determination of P-phase arrival times at regional and local distances using higher order statistics. Geophys J Int 181:1159–1170 . doi: 10.1111/j.1365-246X.2010.04570.x

Lachet C, Bard P-Y (1994) Numerical and Theoretical Investigations on the Possibilities and Limitations of Nakamura's Technique. J Phys Earth 42:377–397

Langston CA (1979) Structure under Mount Rainier, Washington, inferred from teleseismic body waves. J Geophys Res Solid Earth 84:4749–4762

Laurendeau A, Bard P-Y, Hollender F, et al (2017) Derivation of consistent hard rock (1000 < VS < 3000 m/s) GMPEs from surface and down-hole recordings: analysis of KiK-net data. Bull Earthq Eng 1–32 . doi: 10.1007/s10518-017-0142-6

Laurendeau A, Cotton F, Ktenidou O-J, et al (2013) Rock and stiff-soil site amplification: dependency on VS30 and Kappa (K0). Bull Seismol Soc Am 103:3131–3148 . doi: 10.1785/0120130020

Lawrence JF, Denolle M, Seats KJ, Prieto GA (2013) A numeric evaluation of attenuation from ambient noise correlation functions. J Geophys Res Solid Earth 118:6134–6145 . doi: 10.1002/2012JB009513

Lawrence JF, Prieto G, Denolle M, Seats KJ (2012) Attenuation - The Ugly Stepsister of Velocity in the Noise Correlation Family. AGU Fall Meet Abstr 51:

Lawrence JF, Prieto GA (2011) Attenuation tomography of the western United States from ambient seismic noise. J Geophys Res Solid Earth 116:B06302 . doi: 10.1029/2010JB007836

Legifrance (2010) Arrêté du 22 octobre 2010 relatif à la classification et aux règles de construction parasismique applicables aux bâtiments de la classe dite « à risque normal »

Lermo J, Chávez-García FJ (1994) Are microtremors useful in site response evaluation? Bull Seismol Soc Am 84:1350–1364

Lermo J, Chávez-García FJ (1993) Site effect evaluation using spectral ratios with only one station. Bull Seismol Soc Am 83:1574–1594

Lin F-C, Li D, Clayton RW (2012a) Interferometry with a dense 3D dataset

Lin F-C, Ritzwoller MH, Shen W (2011) On the reliability of attenuation measurements from ambient noise cross-correlations. Geophys Res Lett 38:L11303 . doi: 10.1029/2011GL047366

Lin F-C, Tsai VC, Ritzwoller MH (2012b) The local amplification of surface waves: A new observable to constrain elastic velocities, density, and anelastic attenuation. J Geophys Res Solid Earth 117:B06302 . doi: 10.1029/2012JB009208

Liu X, Ben-Zion Y (2013) Theoretical and numerical results on effects of attenuation on correlation functions of ambient seismic noise. Geophys J Int 194:1966–1983 . doi: 10.1093/gji/ggt215

Lobkis OI, Weaver RL (2001) On the emergence of the Green's function in the correlations of a diffuse field. J Acoust Soc Am 110:3011–3017 . doi: 10.1121/1.1417528

Macmurdo J (1824) XXI. Papers relating to the earthquake which occurred in India in 1819. Philos Mag Ser 1 63:105–119 . doi: 10.1080/14786442408644477

Malagnini L, Rovelli A, Hough SE, Seeber L (1993) Site amplification estimates in the Garigliano Valley, Central Italy, based on dense array measurements of ambient noise. Bull Seismol Soc Am 83:1744–1755

Mallet R (1862) Great Neapolitan Earthquake of 1857: The First Principles of Observational Seismology as Developed in the Report to the Royal Society of London of the Expedition Made by Command of the Society Into the Interior of the Kingdom of Naples, to Investigate the Circumstances of the Great Earthquake of Demember 1857. Chapman and Hall

Maruyama T (1963) On the Force Equivalents of Dynamical Elastic Dislocations with Reference to the Earthquake Mechanism. 弾性的転位に等価な力の模型と発震機構の力の模型

Matzel E (2007) Imaging Seismic Attenuation in the Crust and Upper Mantle by Ambient Noise Correlation. AGU Fall Meet Abstr 33:

Matzel E (2008) Attenuation tomography using ambient noise correlation. AGU Fall Meet Abstr 24:

Maufroy E, Chaljub E, Hollender F, et al (2016) 3D numerical simulation and ground motion prediction: Verification, validation and beyond – Lessons from the E2VP project. Soil Dyn Earthq Eng 91:53–71 . doi: 10.1016/j.soildyn.2016.09.047

Maufroy E, Chaljub E, Hollender F, et al (2015a) Earthquake Ground Motion in the Mygdonian Basin, Greece: The E2VP Verification and Validation of 3D Numerical Simulation up to 4 Hz. Bull Seismol Soc Am 105:1398–1418 . doi: 10.1785/0120140228

Maufroy E, Chaljub E, Hollender F, et al (2015b) Validating the numerical simulation approach for ground motion prediction: General framework and latest lessons from the E2VP project. In: vited theme lecture, 6ICEGE (Proceedings of the 6th international conference on earthquake geotechnical engineering). Christchurch, New-Zealand

Maufroy E, Chaljub E, Theodoulidis NP, et al (2017a) Source-related variability of site response in the Mygdonian basin (Greece) from accelerometric recordings and 3-D numerical simulations. Bull Seismol Soc Am 107:

Maufroy E, Chaljub E, Theodoulidis NP, et al (2017b) Source-Related Variability of Site Response in the Mygdonian Basin (Greece) from Accelerometric Recordings and 3D Numerical Simulations. Bull Seismol Soc Am 107:787–808 . doi: 10.1785/0120160107

Maufroy E, Cruz-Atienza VM, Cotton F, Gaffet S (2015c) Frequency-Scaled Curvature as a Proxy for Topographic Site-Effect Amplification and Ground-Motion Variability. Bull Seismol Soc Am 105:354–367 . doi: 10.1785/0120140089

Mayor J, Calvet M, Margerin L, et al (2016) Crustal structure of the Alps as seen by attenuation tomography. Earth Planet Sci Lett 439:71–80 . doi: 10.1016/j.epsl.2016.01.025

Medeiros WE, Schimmel M, Nascimento AF do (2015) How much averaging is necessary to cancel out cross-terms in noise correlation studies? Geophys J Int 203:1096–1100 . doi: 10.1093/gji/ggv336

Melo G, Malcolm A, Mikesell D, Wijk K van (2013) Using SVD for improved interferometric Green's function retrieval. Geophys J Int 194:1596–1612 . doi: 10.1093/gji/ggt172

Menon R, Gerstoft P, Hodgkiss WS (2014) On the apparent attenuation in the spatial coherence estimated from seismic arrays. J Geophys Res Solid Earth 119:3115–3132 . doi: 10.1002/2013JB010835

Milana G, Barba S, Pezzo ED, Zambonelli E (1996) Site response from ambient noise measurements: New perspectives from an array study in Central Italy. Bull Seismol Soc Am 86:320–328

Milne J (1908) Seismology. K. Paul, Trench, Trübner

Mucciarelli M, Gallipoli MR (2004) The HVSR technique from microtremor to strong motion: empirical and statistical considerations. In: Proc. of 13th World Conference of Earthquake Engineering, Vancouver, BC, Canada, Paper. Vancouver

Mulargia F (2012) The seismic noise wavefield is not diffuse. J Acoust Soc Am 131:2853–2858 . doi: 10.1121/1.3689551

Nakamura Y (1989) A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface. Railw Tech Res Inst Q Rep 30:

Nogoshi M, Igarashi T (1971) On the amplitude characteristics of microtremor (part 2). Jour Seism Soc Jpn 24:26–40

Ohmachi T, Nakamura Y, Toshinawa T (1991) Ground Motion Characteristics of the San Francisco Bay Area Detected by Microtremor Measurements. Int Conf Recent Adv Geotech Earthq Eng Soil Dyn

Ohta Y, Kagami H, Goto N, Kudo K (1978) Observation of 1-to 5-second microtremors and their application to earthquake engineering. Part I: Comparison with long-period accelerations at the Tokachi-oki earthquake of 1968. Bull Seismol Soc Am 68:767–779

Okada S, Muraki HO, Kagami H (1991) Two-dimensional analysis on site effects in the Sapporo Urban Region, Japan, based on observation of 1 to 10 second microtremors. In: Fourth International Conference on Seismic Zonation

Oth A, Parolai S, Bindi D, Wenzel F (2009) Source Spectra and Site Response from S Waves of Intermediate-Depth Vrancea, Romania, Earthquakes. Bull Seismol Soc Am 99:235–254 . doi: 10.1785/0120080059

Papageorgiou AS (1988) On two characteristic frequencies of acceleration spectra: patch corner frequency and fmax. Bull Seismol Soc Am 78:509–529

Papageorgiou AS (2003) The barrier model and strong ground motion. Pure Appl Geophys 160:603–634

Papageorgiou AS, Aki K (1983) A specific barrier model for the quantitative description of inhomogeneous faulting and the prediction of strong ground motion. Part II. Applications of the model. Bull Seismol Soc Am 73:953–978

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

Parolai S, Bindi D, Baumbach M, et al (2004) Comparison of different site response estimation techniques using aftershocks of the 1999 Izmit earthquake. Bull Seismol Soc Am 94:1096–1108

Parolai S, Bindi D, Pilz M (2015) κ0: The role of Intrinsic and Scattering Attenuation. Bull Seismol Soc Am 105:1049–1052 . doi: 10.1785/0120140305

Perron V, Hollender F, Bard P-Y, et al (2015) Site instrumentation usefulness: implementation recommendations and application examples of empirical site effect estimation in low to moderate seismicity context. In: 9ème Colloque National AFPS. Marne-la-Vallée

Perron V, Hollender F, Bard P-Y, et al (2017a) Robustness of kappa (κ) measurement in low-to-moderate seismicity areas: insight from a site-specific study in Provence, France. Bull Seismol Soc Am 107:2272–2292 . doi: 10.1785/0120160374

Perron V, Laurendeau A, Hollender F, et al (2017b) Selecting time windows of seismic phases and noise for engineering seismology applications: a versatile methodology and algorithm. Bull Earthq Eng. doi: 10.1007/s10518-017-0131-9

Pilz M, Fäh D (2017) The contribution of scattering to near-surface attenuation. J Seismol 1–19 . doi: 10.1007/s10950-017-9638-4

Prieto GA, Beroza GC (2008) Earthquake ground motion prediction using the ambient seismic field. Geophys Res Lett 35: . doi: 10.1029/2008GL034428

Prieto GA, Denolle M, Lawrence JF, Beroza GC (2011) On amplitude information carried by the ambient seismic field. Comptes Rendus Geosci 343:600–614

Prieto GA, Lawrence JF, Beroza GC (2009) Anelastic Earth structure from the coherency of the ambient seismic field. J Geophys Res Solid Earth 114:B07303 . doi: 10.1029/2008JB006067

Protopapa E, Papastamatiou D, Michaelidas O, Gazetas G (1998) The Ionianet Accelerometer Array: Early Results and Analysis. In: Proceedings of the 11th European Conference on Earthquake Engineering

Purvance MD, Anderson JG (2003) A Comprehensive Study of the Observed Spectral Decay in Strong-Motion Accelerations Recorded in Guerrero, Mexico. Bull Seismol Soc Am 93:600–611

Raptakis D, Theodulidis N, Pitilakis K (1998) Data Analysis of the Euroseistest Strong Motion Array in Volvi (Greece): Standard and Horizontal-to-Vertical Spectral Ratio Techniques. Earthq Spectra 14:203–224 . doi: 10.1193/1.1585996

Régnier J, Bonilla L-F, Bard P-Y, et al (2016) International Benchmark on Numerical Simulations for 1D, Nonlinear Site Response (PRENOLIN): Verification Phase Based on Canonical Cases. Bull Seismol Soc Am 106:2112–2135 . doi: 10.1785/0120150284

Reid HF (1910) The mechanics of the earthquake. Carnegie institution of Washington

Reitherman RK (2012) Earthquakes and Engineers: An International History, Civil Engineering. Reston, Virginia

Riepl J, Bard P-Y, Hatzfeld D, et al (1998) Detailed evaluation of site-response estimation methods across and along the sedimentary valley of Volvi (EURO-SEISTEST). Bull Seismol Soc Am 88:488–502

Rovelli A (1982) On the frequency dependence of Q in friuli from short period digital records. Bull Seismol Soc Am 72:2369–2372

Rovelli A, Singh SK, Malagnini L, et al (1991) Feasibility of the Use of Microtremors in Estimating Site Response During Earthquakes: Some Test Cases in Italy. Earthq Spectra 7:551–561 . doi: 10.1193/1.1585643

Sadeghisorkhani H, Gudmundsson Ó, Roberts R, Tryggvason A (2016) Mapping the source distribution of microseisms using noise covariogram envelopes. Geophys J Int 205:1473–1491 . doi: 10.1093/gji/ggw092

Sanchez G, Rolland Y, Schreiber D, et al (2010) The active fault system of SW Alps. J Geodyn 49:296–302

Sánchez-Sesma FJ, Campillo M (2006) Retrieval of the Green's Function from Cross Correlation: The Canonical Elastic Problem. Bull Seismol Soc Am 96:1182–1191 . doi: 10.1785/0120050181

Sánchez-Sesma FJ, Pérez-Ruiz JA, Campillo M, Luzón F (2006) Elastodynamic 2D Green function retrieval from cross-correlation: Canonical inclusion problem. Geophys Res Lett 33:L13305 . doi: 10.1029/2006GL026454

Sánchez-Sesma FJ, Pérez-Ruiz JA, Luzón F, et al (2008) Diffuse fields in dynamic elasticity. Wave Motion 45:641–654 . doi: 10.1016/j.wavemoti.2007.07.005

Satoh T, Kawase H, Matsushima S ’ichi (2001) Differences between site characteristics obtained from microtremors, S-waves, P-waves, and codas. Bull Seismol Soc Am 91:313–334

Satoh T, Kawase H, Matsushima S (1998) Difference of empirical site characteristics obtained from microtremors, S-wave, P-wave, and coda, and their Theoretical Interpretation. J Seismol Soc Jpn 51:

Sbaa S, Perron V, Hollender F, et al (2016) Analysis of rotation sensor data from the SINAPS@ Kefalonia post-seismic experiment. In: 5th IASPEI/IAEE (ESG5). Taipen

Seats KJ, Lawrence JF, Prieto GA (2012) Improved ambient noise correlation functions using Welch’s method. Geophys J Int 188:513–523 . doi: 10.1111/j.1365-246X.2011.05263.x

Sèbe O (2004) Déconvolution aveugle et sismologie: estimation de la source sismique et des effets de site. Université Joseph-Fourier-Grenoble I

Seekins LC, Wennerberg L, Margheriti L, Liu H-P (1996) Site amplification at five locations in San Francisco, California: A comparison of S waves, codas, and microtremors. Bull Seismol Soc Am 86:627–635

Seo K (1992) A joint work for measurements of microtremors in the Ashigara valley. In: Int. Symp. Effects of Surf. Geol. on Seismic Motion. ESG Japan, pp 43–52

Seo K, YAMANAKA H, Kurita K, et al (2000) A joint research on microtremors in Fukui basin, Japan– For site effects evaluation during the 1948 Fukui (Japan) earthquake. In: Proceedings

SESAME team (2004) Guidelines for the implementation of the H/V spectral ratio technique on ambient vibrations: measurements, processing and interpretation. SESAME European research project

Shapiro NM, Campillo M (2004) Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise. Geophys Res Lett 31:n/a-n/a . doi: 10.1029/2004GL019491

Shapiro NM, Campillo M, Stehly L, Ritzwoller MH (2005) High-Resolution Surface-Wave Tomography from Ambient Seismic Noise. Science 307:1615–1618 . doi: 10.1126/science.1107895

Stehly L, Campillo M, Froment B, Weaver RL (2008) Reconstructing Green’s function by correlation of the coda of the correlation (C3) of ambient seismic noise. J Geophys Res Solid Earth 113:B11306 . doi: 10.1029/2008JB005693

Stehly L, Campillo M, Shapiro NM (2006) A study of the seismic noise from its long-range correlation properties. J Geophys Res 111: . doi: 10.1029/2005JB004237

Stehly L, Campillo M, Shapiro NM (2007) Traveltime measurements from noise correlation: stability and detection of instrumental time-shifts. Geophys J Int 171:223–230 . doi: 10.1111/j.1365-246X.2007.03492.x

Steidl JH, Tumarkin AG, Archuleta RJ (1996) What Is a Reference Site? Bull Seismol Soc Am 86:1733–1748

Svay A, Perron V, Imtiaz A, et al (2017) Spatial coherency analysis of seismic ground motions from a rock site dense array implemented during the Kefalonia 2014 aftershock sequence. Earthq Eng Struct Dyn. doi: DOI: 10.1002/eqe.2881

Takemura M, Motosaka M, Yamanaka H (1995) Strong Motion Seismology in Japan. J Phys Earth 43:211–257 . doi: 10.4294/jpe1952.43.211

Theodoulidis N (2006) Site characterization using strong motion and ambient noise data: Euroseistest (N. Greece). In: 3th IASPEI/IAEE (ESG3). Grenoble, pp 1–10

Thompson EM, Baise LG, Tanaka Y, Kayen RE (2012) A taxonomy of site response complexity. Soil Dyn Earthq Eng 41:32–43 . doi: 10.1016/j.soildyn.2012.04.005

Thouvenot F (1983) Frequency dependence of the quality factor in the upper crust: a deep seismic sounding approach. Geophys J Int 73:427–447

Triantafyllidis P, Theodulidis N, Savvaidis A, et al (2006) Site effects estimation using earthquake and ambient noise data: The case of Lefkas town (W. Greece). In: Procceedings of the 1st European Conference on Earthquake and Engineering Seismology (1ECEES), Geneva, Switzerland. pp 3–8

Trifunac MD, Todorovska MI (2001) Evolution of accelerographs, data processing, strong motion arrays and amplitude and spatial resolution in recording strong earthquake motion. Soil Dyn Earthq Eng 21:537–555 . doi: 10.1016/S0267-7261(01)00013-6

Tsai C-CP, Chen K-C (2000) A model for the high-cut process of strong-motion accelerations in terms of distance, magnitude, and site condition: An example from the SMART 1 array, Lotung, Taiwan. Bull Seismol Soc Am 90:1535–1542

Tsai VC (2011) Understanding the amplitudes of noise correlation measurements. J Geophys Res Solid Earth 116:B09311 . doi: 10.1029/2011JB008483

Udwadia FE, Trifunac MD (1973) Comparison of earthquake and microtremor ground motions in El Centro, California. Bull Seismol Soc Am 63:1227–1253

Van Houtte C, Drouet S, Cotton F (2011) Analysis of the Origins of K (Kappa) to Compute Hard Rock to Rock Adjustment Factors for GMPEs. Bull Seismol Soc Am 101:2926–2941 . doi: 10.1785/0120100345

Van Houtte C, Ktenidou O-J, Larkin T, Holden C (2014) Hard-Site 0 (Kappa) Calculations for Christchurch, New Zealand, and Comparison with Local Ground-Motion Prediction Models. Bull Seismol Soc Am 104:1899–1913 . doi: 10.1785/0120130271

Viens L, Koketsu K, Miyake H, et al (2016) Basin-scale Green's functions from the ambient seismic field recorded by MeSO-net stations. J Geophys Res Solid Earth 121:2016JB012796 . doi: 10.1002/2016JB012796

Viens L, Laurendeau A, Bonilla LF, Shapiro NM (2014) Broad-band acceleration time histories synthesis by coupling low-frequency ambient seismic field and high-frequency stochastic modelling. Geophys J Int 199:1784–1797 . doi: 10.1093/gji/ggu362

Viens L, Miyake H, Koketsu K (2015) Long-period ground motion simulation of a subduction earthquake using the offshore-onshore ambient seismic field. Geophys Res Lett 42:5282–5289 . doi: 10.1002/2015GL064265

Warren N (1972) Q and structure. The moon 4:430–441 . doi: 10.1007/BF00562009

Weaver R, Froment B, Campillo M (2009) On the correlation of non-isotropically distributed ballistic scalar diffuse waves. J Acoust Soc Am 126:1817–1826 . doi: 10.1121/1.3203359

Weaver RL (2012) On the retrieval of attenuation from the azimuthally averaged coherency of a diffuse field. ArXiv12066513 Phys

Weaver RL (2013) On the retrieval of attenuation and site amplifications from ambient noise on linear arrays: further numerical simulations. Geophys J Int ggt063 . doi: 10.1093/gji/ggt063

Weaver RL (2011) On the amplitudes of correlations and the inference of attenuations, specific intensities and site factors from ambient noise. Comptes Rendus Geosci 343:615–622 . doi: 10.1016/j.crte.2011.07.001

Weaver RL, Lobkis OI (2001) Ultrasonics without a Source: Thermal Fluctuation Correlations at MHz Frequencies. Phys Rev Lett 87:134301 . doi: 10.1103/PhysRevLett.87.134301

Weemstra C, Boschi L, Goertz A, Artman B (2013) Seismic attenuation from recordings of ambient noise. GEOPHYSICS 78:Q1–Q14 . doi: 10.1190/geo2012-0132.1

Wen J, Chen X (2012) Variations in fmax along the Ruptured Fault during the Mw 7.9 Wenchuan Earthquake of 12 May 2008. Bull Seismol Soc Am 102:991–998 . doi: 10.1785/0120110105

Wood HO (1908) Distribution of apparent intensity in San Francisco. Calif Earthq April 18:220–245

Yamanaka H, Dravinski M, Kagami H (1993) Continuous measurements of microtremors on sediments and basement in Los Angeles, California. Bull Seismol Soc Am 83:1595–1609

Yoritomo JY, Weaver RL (2016) Fluctuations in the cross-correlation for fields lacking full diffusivity: The statistics of spurious features. J Acoust Soc Am 140:702–713 . doi: 10.1121/1.4959002

Yu J, Haines J (2003) The Choice of Reference Sites for Seismic Ground Amplification Analyses: Case Study at Parkway, New Zealand. Bull Seismol Soc Am 93:713–723 . doi: 10.1785/0120010289

Zhang H, Thurber C, Rowe C (2003) Automatic P-wave arrival detection and picking with multiscale wavelet analysis for single-component recordings. Bull Seismol Soc Am 93:1904–1912 . doi: 10.1785/0120020241

Zhang J, Yang X (2013) Extracting surface wave attenuation from seismic noise using correlation of the coda of correlation. J Geophys Res Solid Earth 118:2191–2205 . doi: 10.1002/jgrb.50186

Zhao BM, Horike M, Takeuchi Y (1998) Reliability of estimation of seismic ground characteristic by microtremor observation. In: Proceedings of the XIth European conference on earthquake engineering, Paris September. pp 6–11

Annexe 1

Selecting time windows of seismic phases and noise for engineering seismology applications: a versatile methodology and algorithm

Bull Earthquake Eng
DOI 10.1007/s10518-017-0131-9

ORIGINAL RESEARCH PAPER

Selecting time windows of seismic phases and noise for engineering seismology applications: a versatile methodology and algorithm

Vincent Perron[1,2,3] · Aurore Laurendeau[4,7] · Fabrice Hollender[1,3] · Pierre-Yves Bard[3] · Céline Gélis[2] · Paola Traversa[5] · Stéphane Drouet[6]

Received: 14 October 2016 / Accepted: 28 March 2017
© The Author(s) 2017. This article is an open access publication

Abstract Seismic signal windowing is the preliminary step for many analysis procedures in engineering seismology (standard spectral ratio, quality factor, general inversion techniques, etc.). Moreover a noise window is often necessary for the data quality control through the signal-to-noise verification. Selecting the noise window can be challenging when large heterogeneous datasets are considered, especially when they include short pre-event noise signals. This study proposes a fully automatic and configurable (i.e., with default parameters that can also be user-defined) algorithm to windowing the noise and the P, S, coda and full signal once the P-wave (T_P) and S-wave (T_S) first arrivals are known. An application example is given on a KiK-net dataset. A Matlab language implementation of this algorithm is proposed as an online resource.

Keywords Seismic phase windowing · Noise windowing · Signal-to-noise ratio · Matlab algorithm

Electronic supplementary material The online version of this article (doi:10.1007/s10518-017-0131-9) contains supplementary material, which is available to authorized users.

✉ Vincent Perron
vincent.perron.mail@gmail.com

1 French Alternative Energies and Atomic Energy Commission (CEA), Cadarache, 13108 Saint-Paul-lez-Durance Cedex, France

2 IRSN, PRP-DGE/SCAN/BERSSIN, BP 17, 92262 Fontenay-aux-Roses, France

3 Univ. Grenoble Alpes, IFSTTAR, ISTerre, CS 40700, 38058 Grenoble, France

4 CEA, DAM, DIF, 91297 Arpajon, France

5 EDF, CEIDRE/TEGG/SGG, 905 av. du Camp de Menthe, 13097 Aix-en-Provence, France

6 GEOTER, Fugro Group, Auriol, France

7 Present Address: Instituto Geofisico Escuela Politecnica Nacional, Quito, Ecuador

Published online: 17 April 2017

1 Introduction

Selecting specific signal phases (i.e., P, S, or coda waves) is required for diverse applications in seismology. For instance, the early part of shear-wave phase is often used for site effects assessment (e.g., Borcherdt 1970; Jongmans and Campillo 1993; Horike et al. 2001; Satoh et al. 2001) and is the basis of the evaluation of the kappa parameter (Anderson and Hough 1984; Ktenidou et al. 2014), while the quality factor (Q) related to the attenuation is regularly estimated from the later coda arrivals (e.g., Aki and Chouet 1975; Mayor et al. 2016). Moreover, to estimate the quality of a signal and its frequency range of validity, the signal-to-noise ratio (SNR) is computed from the ratio between the Fourier amplitude spectrum (FAS) estimated on parts of the signal and the FAS generally evaluated for a noise window of the same duration that is often selected before the event.

While numerous studies have proposed automatic picking algorithms to determine the P-wave (T_P) and/or S-wave (T_S) first arrivals (e.g., Baer and Kradolfer 1987; Sleeman and van Eck 1999; Zhao and Takano 1999; Zhang et al. 2003; Stefano et al. 2006; Wong et al. 2009; Küperkoch et al. 2010; Tan and He 2016), there have only been a few that have offered a solution for windowing the different phases of the earthquake signal. Most studies have considered a constant window duration for every event, without taking into account the earthquake rupture duration or the expansion of the signal duration as waves are being propagated to larger distances (Phillips and Aki 1986; Bonilla et al. 1997; Drouet et al. 2010; Douglas et al. 2010). Recently, some studies have proposed more complex approaches based on the signal analysis (e.g., Maggi et al. 2009) or based on a model using the information extracted from seismic bulletins (e.g., Kishida et al. 2016).

When working with a large and heterogeneous dataset, once the T_P and T_S first arrivals have been picked, defining a specific window can be complicated. This complexity increases when a noise window has to be assessed for SNR computation. Indeed, time-series extracted from triggered instruments and/or generated automatically from regional or national networks, often present short and variable pre-event noise durations. When a large dataset has to be processed, as for generalized inversion techniques (GIT—e.g., Drouet et al. 2008) or ground motion prediction equations (e.g., Laurendeau et al. 2013), then a complex automatic procedure has to be used to avoid the introduction of poor quality data into the processing and to minimize the number of rejected data due to difficult window selection.

The motivation behind the present study was to provide a suitable windowing tool for spectral estimation on different phases with due account to signal-to-noise ratio issues. A method to select the phase windows of any dataset for which the T_P and T_S first arrivals are known is proposed, and a suitable solution to estimate the noise window from heterogeneous datasets with variable noise level and duration is provided. An automatic Matlab algorithm was developed and tested on a KiK-net Japanese dataset that is composed of more than 2000 manually picked events with short and variable durations of pre-event noise (Laurendeau et al. 2013). The records are accelerograms from local to regional events that are used between 0.25 and 30 Hz mostly. This study was initially developed for the application of GIT to a specific KiK-net subset (Foundotos et al. 2016—same issue), and for the correction of the KiK-net surface records for local site effects for prediction of hard-rock reference ground motion prediction (Laurendeau et al. 2016—same issue): more details on the dataset can be found in these papers.

After a short reminder on the relationships between window duration and the associated minimum frequency valid for the FAS, (f_{min}), we present the model proposed for the

seismic-phase windowing in a first step, and the methodology used for the noise window selection in a second step. In a final step, some examples are given to discuss the windowing obtained for local, regional and teleseimic events, as well as for complex examples including after- or fore-shocks.

2 Spectral resolution and sensitivity to time window duration

Many studies have tested the sensitivity of their data to signal window duration (e.g., Satoh et al. 2001; Ktenidou 2010; Douglas et al. 2010). These have mostly reported only limited dependence, provided that the same seismic phase is considered. These observations have sometimes been used to justify the choice of a constant window duration for every event.

In addition to the potential differences in the input seismic signal delimited by different windows, the resolution of the FAS differs as well. Indeed, the longer the selected time window, the higher the number of wavelengths considered for the FAS at each frequency. For instance, a 10-s-long window contains only one wavelength of a 10-s period, but 10 wavelengths of a 1-s period, and 100 wavelengths of a 0.1-s period, while a 1-s-long window contains one tenth of the wavelengths at each of these periods. If N is the number of wavelengths necessary to insure a good spectral resolution, then the minimal reliable frequency for the FAS is given by:

$$f_{min} = \frac{N}{D}, \tag{1}$$

where D is the duration of the window considered. The higher is N, the better is the spectral resolution. Based on our experience, taking $N = 10$ is enough to give the assurance of good spectral resolution. However, taking such a high N number is not always possible, as this depends on the seismic phase or noise duration available, and on the minimal frequency necessary for the application. When it is required, the N-value can be optimized by tested it values for different signals. Figure 1 shows the sensitivity of the FAS to the S-wave window duration (D_S). The colors from blue to yellow show the results for window lengths from 2.5 to 60 s. In this example and for various other KiK-net signals tested (not represented here), the minimal N-value for this dataset is around three. Indeed, clear discrepancies appear at low frequency for the shortest window, generally just below the f_{min} criteria (vertical lines) obtained with $N = 3$. Because the KiK-net dataset present very limited duration of noise for the analyses and tests with $N = 3$ appear satisfying, we keep this N-value in the following examples.

In agreement with the literature, we find a good stability of the FAS over the frequency range where the N-value criterion is satisfied (Fig. 1). Thus, the FAS seems weakly dependent on the duration of the signal considered, provided the most energetic part of the signal is common to every window. It means that small variations on the duration of the selected phase window would lead to negligible change on the FAS.

3 Seismic-phase windowing

Phase windowing consists of using the first arrival time of P waves and S waves to automatically select different windows: for P, S, coda, and full signals. The nomenclature for the phase intervals and the different times considered is given in Fig. 2. In addition to

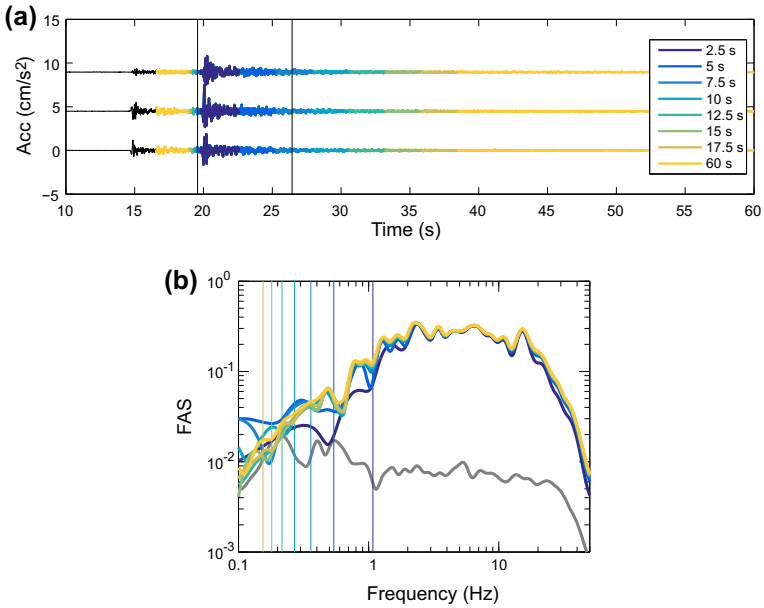


Fig. 1 S-wave window duration impact on the computation of the Fourier acceleration spectrum (FAS). **a** The time histories of the three components and the selected window are shown. The window is a little larger because the cosine taper is applied outside the limit interval of the defined S-wave window. **b** The corresponding FAS are shown. FAS in gray correspond to the noise spectrum. The *vertical lines* indicate the minimum frequency associated with the S-wave window and allowed to have at least three wavelengths (**N** = 3). In this example, it is necessary to have at least 10 s of signal to have a reliable spectrum at 0.3 Hz

T_P and T_S, the ending signal time (T_{end}) can be defined as well. This time is used as an upper limit for the duration of the S phase, the coda phase, and the full signal. We recommend selecting T_{end} directly from the spectrogram with precautions in the chosen color scale, to be able to detect the end of the coda waves at low frequency, as well as eventual strong noise or aftershocks at each frequency. If the T_{end} value is not provided, then it is automatically taken as the time corresponding to 95% of the cumulated energy evaluated on the three components between T_P and the end of the record. It is particularly useful to pick T_{end} for low SNR records and in the case of close aftershocks or strong transient noise, which can be included by the cumulated energy approach. Moreover, the cumulated energy approach presents the drawback that it depends on the duration and level of noise included in the record around the signal, especially when the latter is weak. The time of the initial sample of the record is denoted T_i, while the final sample is given by T_f. The time for the earthquake occurrence is noted as T_0. The method and its associated algorithm have been developed for FAS processing purposes. The window's edge must be tapered to satisfy the infinite signal assumption made for the Fourier transform on a finite window. Thus, the rate of tapering (tx) can be specified in the window selection process, to enlarge the windows and apply the tapering outside the accurate delimitation of the phase windows.

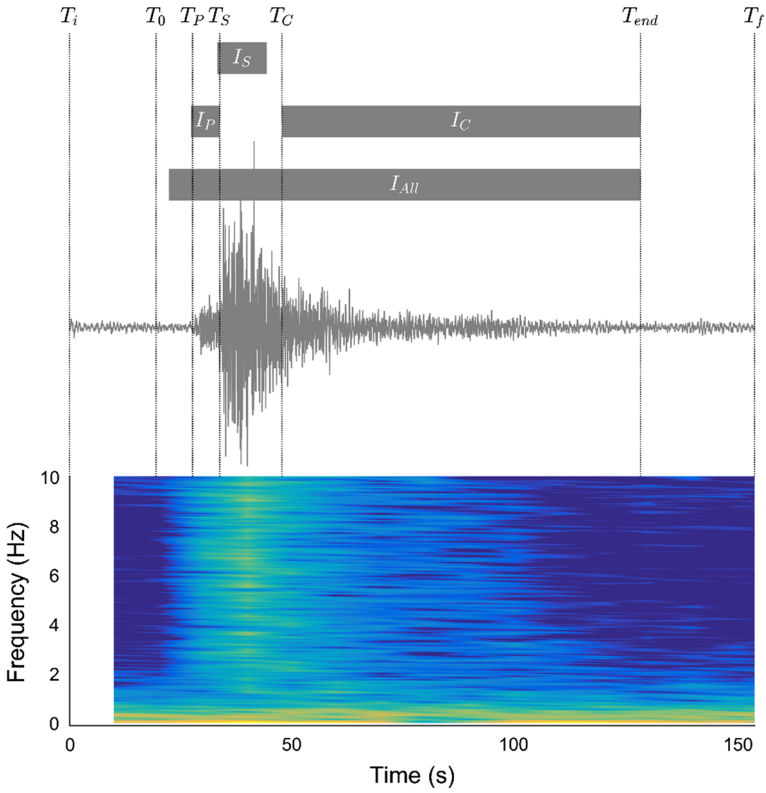


Fig. 2 East–West component record from a M_L 2.2 earthquake at 56 km epicentral distance and occurring at $\boldsymbol{T_0}$. The P, S, coda and all phases are represented by the gray bands indicating $\boldsymbol{I_P}$, $\boldsymbol{I_S}$, $\boldsymbol{I_C}$ and $\boldsymbol{I_{All}}$, respectively, with their corresponding durations ($\boldsymbol{D_P}$, $\boldsymbol{D_S}$, $\boldsymbol{D_C}$, $\boldsymbol{D_{All}}$) and considering a rate of tapering of 5% ($\boldsymbol{tx} = 0.05$). The first arrival times ($\boldsymbol{T_P}$, $\boldsymbol{T_S}$, $\boldsymbol{T_C}$), phases ending ($\boldsymbol{T_{end}}$) and beginning ($\boldsymbol{T_i}$) and ending ($\boldsymbol{T_f}$) of the record are also shown. *Bottom* spectrogram for the seismic energy as a function of the time and frequency

3.1 P-wave windowing

The P-wave window is the easiest phase to delimit as it starts at T_P and ends at T_S. The duration of the P-wave phase (D_P) can be written as:

$$D_P = \frac{T_S - T_P}{1 - tx}, \tag{2}$$

where tx takes into account the single edge enlargement in the noise before the P-wave onset, to apply the tapering on this pre-event noise and thus to avoid losing some parts of the P-wave phase in the tapering process for the FAS evaluation. Finally, the P-wave window interval is defined as:

$$I_P = [T_P - D_P \times tx; T_S]. \tag{3}$$

3.2 S-wave windowing

The S-wave window duration is given according to a source term through the inverse of the corner frequency (f_c—Brune 1970), and a propagation term taking into account the difference time between the P-wave and S-wave first arrivals (T_S–T_P):

$$D_S = max\left\{\begin{matrix} D_{Smin} \\ \frac{1}{f_c} + (T_S - T_P) \end{matrix}\right\} \Big/ (1 - 2tx). \tag{4}$$

Here, the window is being enlarged by a factor tx for both edges. Thus, the enlargement of the first edge includes a small portion of P waves in the S window, although P waves are already included in the S-wave phase anyway. A minimal duration (D_{smin}) can be defined for the question of spectral resolution at low frequency. f_c is estimated directly by the Brune (1970) relationship [Eq. (4)] from the seismic moment (M_0), considering a stress drop ($\Delta\sigma$) of 10 bars and a mean shear-wave velocity (β_S) for the crust of 3500 m/s:

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

$\Delta\sigma$ and β_S can be, however, easily adapt to the target region if needed. The seismic moment can be deduced from the moment magnitude (M_W) according to Eq. (6) (Hanks and Kanamori 1979):

$$M_0 = 10^{1.5M_w + 9.1}. \tag{6}$$

If the moment magnitude is not available, M_W can be approximated by the local magnitude (M_L) extracted from the seismic catalog. This estimation of the source duration is anyway approximate, but is supported by the observed stability in the spectrum evaluated from windows with different D_S as shown in Fig. 1. Kishida et al. (2016) proposed a similar formulation for D_S, with a part related to the source with different durations defined empirically according to the magnitude, and a part related to the propagation defined as one tenth of the hypocentral distance ($0.1R_h$).

First, for the source term, $f_c$is high for small and moderate earthquakes (M < 5), and thus D_S is only controlled by the propagation term. Then, the source duration can be neglected for M < 5, making Eq. (4) usable without the need for parameters other than T_S and T_P. The use of Eq. (4) for earthquakes with M > 7.5 can lead to very large source durations (see Kishida et al. 2016). A maximal S-wave window duration (D_{smax}) can be chosen in this context. Secondly, for the propagation term, (T_S–T_P) is widely accepted to be close to a 1/8 of the hypocentral distance given in kilometers, making both expressions similar. However, the formulation in Eq. (3) has the advantage of being independent of uncertainties in the source localization (especially the depth) given by the seismic catalog.

The S-wave interval is finally defined as:

$$I_S = \left[T_S - D_S \times tx; \min\left\{\begin{matrix} T_S + D_S(1 - tx) \\ T_{end} \end{matrix}\right\}\right]. \tag{7}$$

Figure 3 shows the variation of source and path component duration of D_S as defined by Eq. (4) with the magnitude and rupture distance for the KiK-net dataset example. The maximum duration due to the source is around 17 s, and that due to the propagation is

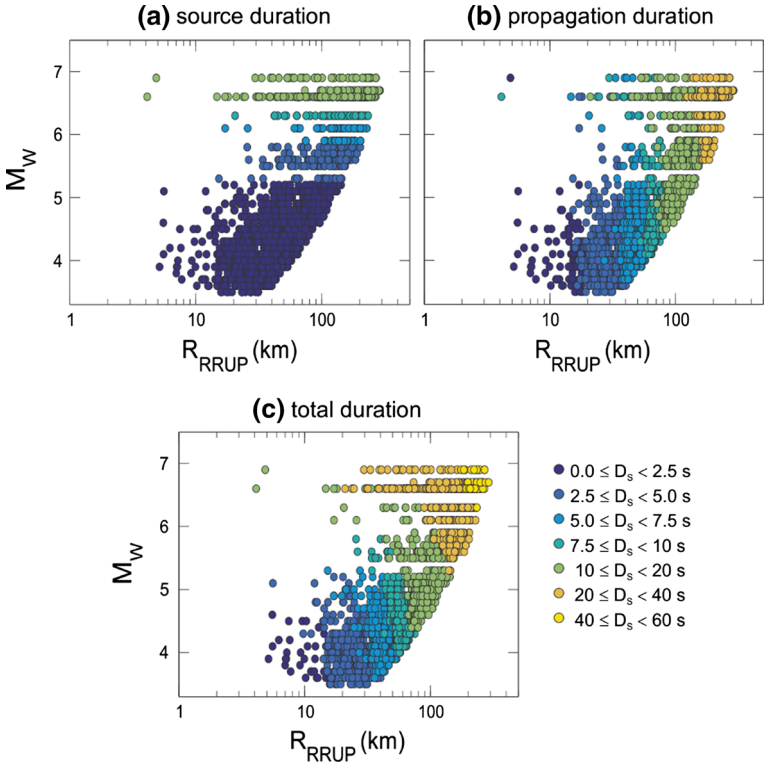


Fig. 3 Duration of the S-wave windows (**c**) with respect to the term duration of the source (**a**; M_W effect) and the path (**b**; R_{RRUP} effect)

around 38 s. The total duration has a minimum of 1.2 s and a maximum of 55 s. However, in the following, we consider a target minimal frequency of 0.3 Hz and a minimum of three wavelengths included inside the window (N = 3). Equation (1) finally gives a minimal duration $D_{Smin} = 10$ s for applications on the KiK-net dataset.

3.3 Coda-wave windowing

Aki (1969) defined the beginning of the coda phase as twice the S-wave travel time ($2(T_S - T_0)$) after earthquake occurrence T_0. To be independent of the information extracted from the seismic bulletin, we propose a formulation equivalent to the Aki (1969) definition, but based only on the T_P and T_S parameters. Using the approximation commonly accepted that $R_h \propto 8(T_S - T_P)$ and that $\beta_S \approx 3.5\,\mathrm{km/s}$ and considering $R_h/\beta_S \approx (T_S - T_0)$, we easily find $2(T_S - T_0) \approx 4.6(T_S - T_P)$. The time of the first 'arrival' of the coda phase (T_C) is finally $T_C = 4.6(T_S - T_P) + T_0$, which is equivalent to $T_C = 2.3(T_S - T_P) + T_S$. This formulation is also empirically confirmed through the good coefficient of determination $R^2 = 0.98$ for the linear regression on the local events ($R_h < 500$ km) between Aki (1969) and our T_C expression. This definition of T_C has the advantage that it is independent of the uncertainty on the time origin of the earthquake. Finally, the coda-wave interval can be written as:

$$I_C = [3.3T_S - 2.3T_P; T_{end}]. \tag{8}$$

Its associate duration is simply:

$$D_C = T_{end} - (3.3T_S - 2.3T_P). \tag{9}$$

A minimal coda wave window duration (D_{Cmin}) can be defined. If $D_C < D_{Cmin}$, then no coda wave window is returned by the algorithm since the signal is too weak and the amplitude of the coda waves falls rapidly below the noise level. A coda signal is generally available only for events with good SNR (i.e. >10) in a large frequency range.

3.4 Full-signal windowing

A full-signal interval (I_{All}) is also proposed that starts at T_P and finishes at T_{end}. This takes into account the enlargement on the first edge for the tapering, and it is defined as:

$$I_{All} = [T_P - D_{All} \times tx; T_{end}] \quad with \quad D_{All} = \frac{T_{end} - T_P}{1 - tx}. \tag{10}$$

This full signal window is particularly useful when no specific phase is mandatory, and to obtain long enough windows (D_{All}) to assess spectra with good resolution up to low frequencies.

4 Noise window selection

The noise window selection generally consists of taking the duration of the target window and reporting it before the first P-wave arrival. Here, a more complex scheme has been developed, to take into account the availability of data with short windows of pre-event noise. Figure 4 shows the pre-event noise duration that is available for the KiK-net dataset. This shows generally that these durations are short and variable, which makes the noise window selection difficult. Only one noise window of duration (D_N) is selected to assess the SNR of several seismic phase windows of different durations. The duration of the target noise window (D_t) has to be at least as long as the longest seismic signal window requested (D_P, D_S, D_C or D_{All}), or long enough to satisfy the f_{min} criterion. Thus, FAS estimated from these windows of different durations have to be normalized by the square root of the duration to obtain the Fourier amplitude spectrum density ($FASD = FAS/\sqrt{D}$) that is length-independent, for SNR purposes especially.

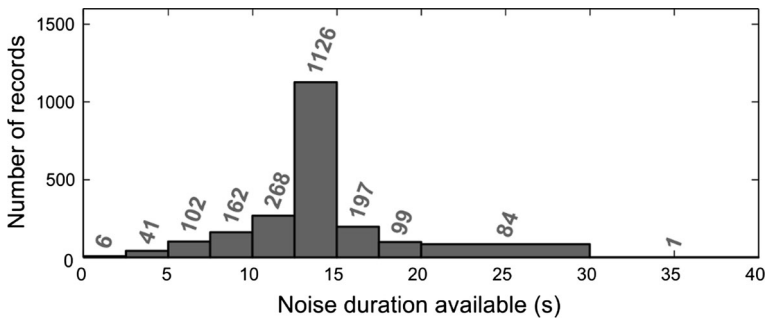


Fig. 4 The number of records versus the pre-event noise duration available

In the present example, we consider a minimal noise duration D_{min} of 10 s to be consistent with the minimal S-wave window duration taken previously. However, 10 s of noise is not available before the event for all of the records of the KiK-net dataset, as shown in Fig. 4. The dataset contains 311 records without 10 s of pre-event noise. Thus, different noise window definitions are tested, for which the energy was then compared. The idea was to select a noise window with sufficient duration, and also a window with a representative level of energy (without strong seismic signal included). To do this, one pre-event noise window (I_{N1}) of duration (D_{N1}) is tested as well as two post-events noise windows: a short window I_{N2} of duration D_{N2}, and a long window I_{N3} of duration D_{N3}.

$$I_{N1} = \left[max\left\{ \begin{matrix} (T_P - 0.1) - max\left\{ \begin{matrix} D_{min} \\ D_t \end{matrix} \right\} \\ T_i \end{matrix} \right\}; T_P - 0.1 \right]; \tag{11}$$

$$I_{N2} = \left[max\left\{ \begin{matrix} T_f - max\left\{ \begin{matrix} D_{min} \\ D_{N1} \end{matrix} \right\} \\ T_S + D_S \end{matrix} \right\}; T_f \right]; \tag{12}$$

$$I_{N3} = \left[max\left\{ \begin{matrix} T_f - max\left\{ \begin{matrix} D_{min} \\ D_{N1} \\ D_t \end{matrix} \right\} \\ T_S + D_S \end{matrix} \right\}; T_f \right]. \tag{13}$$

The pre-event window (I_{N1}) is the preferred one, as no earthquake signal can be introduced in the noise evaluation. However, if the pre-event noise that is available is too short, or if there is a fore-shock before the mainshock, then other windows have to be considered. The post-event window I_{N3} is longer than I_{N2} only when the target phase duration (D_t) is greater than D_{min} and D_{N1}. No S wave can be introduced in the post-event window, while the coda wave can be accepted to minimize the number of rejected signals due to a lack of noise available. An example is given in Fig. 5 that illustrates the noise window selecting process for the S-wave duration targeted on a record presenting a limited pre-event noise duration available. In this example, I_{N3} is preferable to the two other noise windows because it is the only one that has the same duration as the targeted S-wave window while it provides

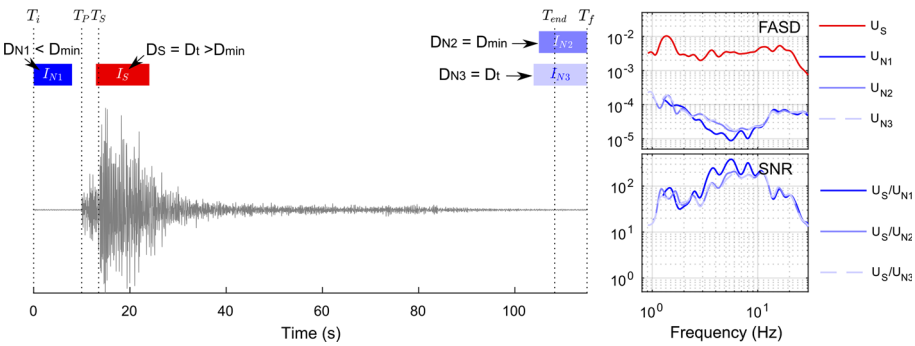


Fig. 5 Example of the noise window selection process when the S-wave window duration is targeted ($\boldsymbol{D_t = D_s}$) and considering a 10 s minimal window duration ($\boldsymbol{D_{min}}$ = 10 s). The S-wave window ($\boldsymbol{I_S}$), the pre-event noise window ($\boldsymbol{I_{N1}}$) and the short and long post-event noise window ($\boldsymbol{I_{N2}}$ and $\boldsymbol{I_{N3}}$) are represented as well as their corresponding amplitude spectral density (FASD) and deduce signal-to-noise ratio (SNR)

similar FASD than the two other noise windows even if some coda waves may be included inside I_{N3}. This latter verification is carried out on the comparison of the mean energy of the different windows. A minimum pre-event noise duration of 1 s is mandatory for these comparisons. To be able to compare just the representativeness of each noise window, the mean energy (E) is estimated for the three noise windows (E_{N1}, E_{N2}, E_{N3}) according to the following definition:

$$E = \frac{\sum_{i=N_{min}}^{Nf} FAS_{EW}(f_i)^2 + FAS_{NS}(f_i)^2 + FAS_Z(f_i)^2}{3 \times (Nf - N_{min} + 1)}, \tag{14}$$

where N_{min} is the index of the minimum frequency (f_{min}), Nf is the number of frequency samples, and FAS are the Fourier amplitude spectra computed for the three components.

A scheme can take into account the length of each window as well as their mean energy, to determine the best noise window for the noise FASD evaluation, as detailed in Fig. 6. The energy for the three noise window comparisons is weighted by some factors defined by the user (F_1, F_2, F_3, F_4). This allows I_{N1}, I_{N2}, or I_{N3} to be favored, depending on the duration available for each one, and the number of rejected records to be minimized due to lack of noise for the SNR evaluation. In addition to the noise and the P, S, coda and all signal windows, the algorithm returns a Flag value that indicates which noise window has been selected and how. For the KiK-net application the scheme of the noise window selection process is configured with: $D_{min} = 10\,\mathrm{s}, F_1 = 5, F_2 = 3, F_3 = 2$, and $F_4 = 0.67$.

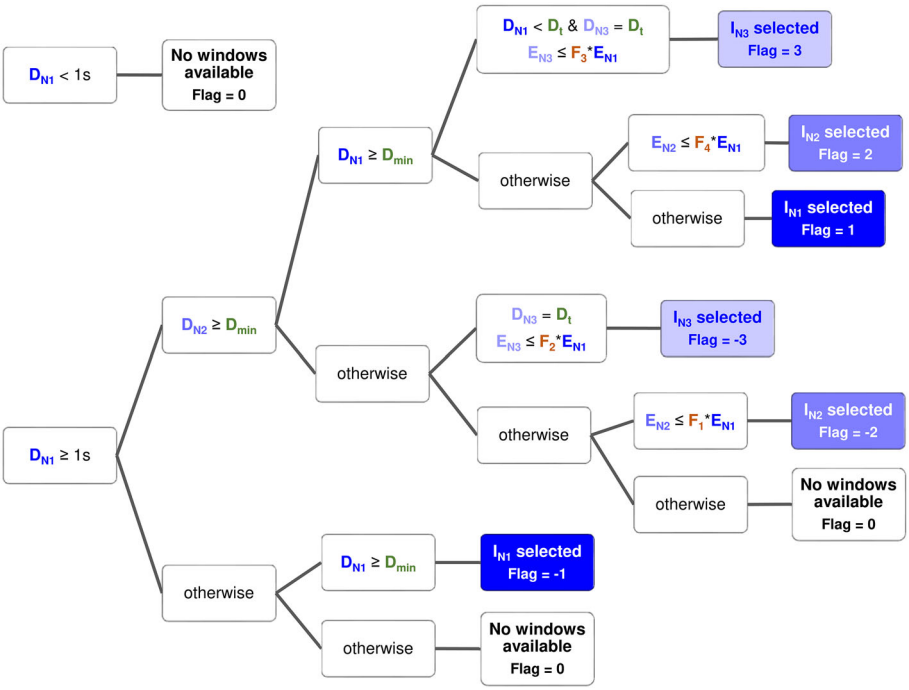


Fig. 6 Scheme of the noise window selection methodology. $\mathbf{D_{N1}}$, $\mathbf{D_{N2}}$ and $\mathbf{D_{N3}}$ are the pre-event and the short and the long post-event noise window duration respectively. The corresponding spectral energy of these noise windows are given by $\mathbf{E_{N1}}$, $\mathbf{E_{N2}}$ and $\mathbf{E_{N3}}$. Few other parameters can be easily adapted to each dataset: the minimal and target duration ($\mathbf{D_{min}}$ and $\mathbf{D_t}$) and the weighted factors $\mathbf{F_1}$–$\mathbf{F_4}$

Table 1 Number of records selected both at depth and at the surface from the 2119 events of the KiK-net dataset for each noise window definition

Flag	Definition	Records (n)	
		At depth	At the surface
−3	Long post-event noise window (I_{N3}) selected without the possibility to take a pre-event noise window	297	250
−2	Short post-event noise window (I_{N2}) selected without possibility to take a pre-event noise window	60	72
−1	Pre-event noise window (I_{N1}) selected without possibility to take a post-event noise window	31	33
0	No noise window selected	6	30
1	Pre-event noise window (I_{N1}) selected	1538	1547
2	Short post-event noise window (I_{N2}) selected	6	25
3	Long post-event noise window (I_{N3}) selected	181	162

The main idea is that the pre-event window is favored when both the pre-event and post-event windows are longer than D_{min} and D_t (Flag 1). However if the pre-event window is shorter than D_t, then the longest post-event window is preferred (Flag 3). In the same way, if I_{N1} is shorter than D_{min}, then the long post-event is preferred if not too much signal is included inside it (Flag-3), otherwise the short post-event window is chosen (Flag-2).

These Flags for the noise selection are indicated in Fig. 6 and are given in Table 1, with the corresponding number of events that were selected for the KiK-net dataset example. Most of the noise windows selected were pre-event windows. The post-event windows selected were mostly constituted by the long window. Only a few signals are removed from the dataset due to the absence of a noise window for the SNR evaluation. All of the parameters and factors given in this article can be adjusted as an input of the algorithm.

5 Application examples and discussion

The advantage of our windowing formulation is illustrated in Fig. 7 by the comparison with two simple formulations: a 30- and a 10-s constant S-wave window. Three earthquakes recorded in Provence (Southeastern France) with increasing epicentral distance are presented. The influence of the S-wave window duration is visible on the FASD and the SNR. For the closest earthquake (7a), the 30-s window includes a lot of coda waves and underestimates greatly the FASD obtained with the 10-s window and our formulation. For the intermediate epicentral distance (7b), our formulation is in between the two constant window leading to a FASD that is close to the average between the three window definitions. Only the beginning of the S-wave window is included in the two constant windows for the regional earthquake (7c) while our formulation provides longer duration leading to a slightly lower FASD amplitude even if the three curves are very similar. Thus, a constant window duration is not adapted when working with datasets composed by both local and regional earthquakes while our formulation gives a suitable solution. Concerning the other phases of the signal, we used the classical P wave formulation that seems appropriate, while our coda wave window formulation begins very close to the one predicted by Aki (1969), as expected. The T_{end95} cumulative energy estimation is always later than the

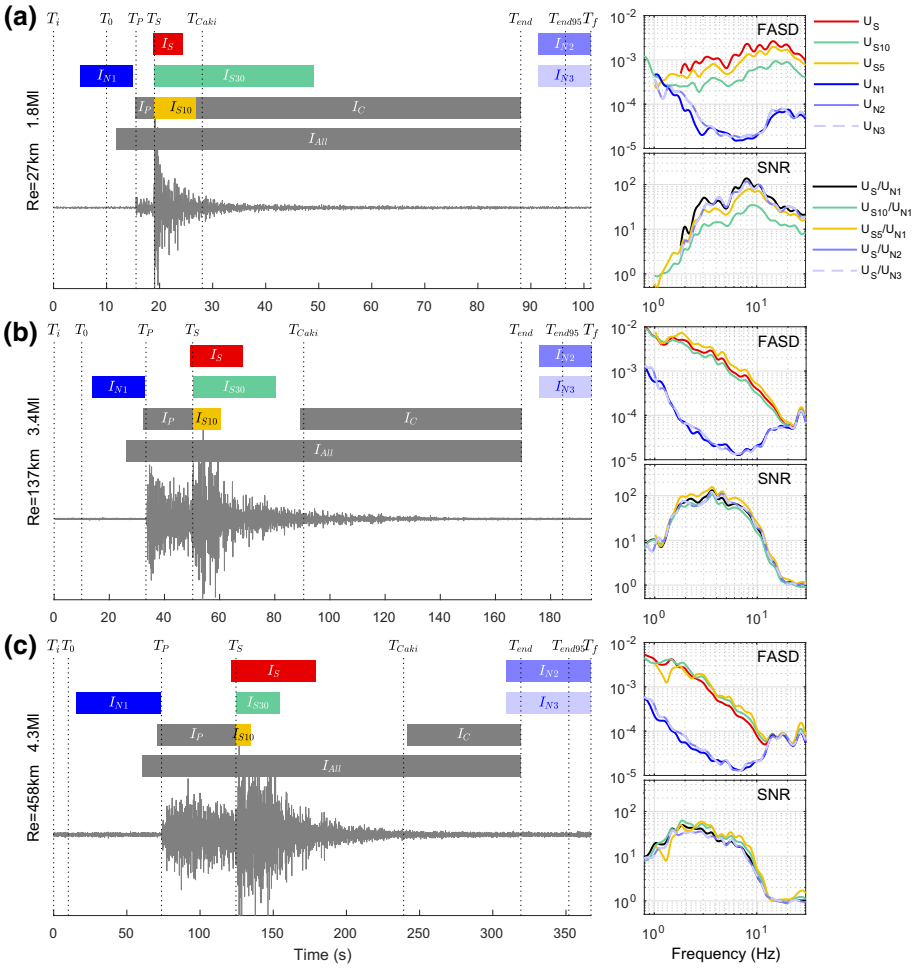


Fig. 7 Example of the signal phase and noise windowing for 3 earthquakes located at 27 **(a)**, 137 **(b)** and 458 km **(c)** from the recording site. The P-wave window ($\boldsymbol{I_P}$), the S-wave window ($\boldsymbol{I_S}$), the coda wave window ($\boldsymbol{I_C}$) and the full signal window ($\boldsymbol{I_{All}}$) are represented. To make the comparison between our formulation and simple constant window formulation, 30- and 10-s long windows ($\boldsymbol{I_{S30}}$ and $\boldsymbol{I_{S10}}$) are also represented as well as their corresponding amplitude spectral density (FASD) and associated signal-to-noise ratio (SNR). The picked phase ending ($\boldsymbol{T_{end}}$) and the one predicted by the 95% fractile of the cumulative energy ($\boldsymbol{T_{end95}}$) are represented as well as the Aki (1969) coda beginning formulation ($\boldsymbol{T_{Caki}}$) that can be compared to our formulation ($\boldsymbol{I_C}$). Identically to Fig. 5, the noise window selection process is also represented by the pre-event noise window ($\boldsymbol{I_{N1}}$) and the short and long post-event noise window ($\boldsymbol{I_{N2}}$ and $\boldsymbol{I_{N3}}$) and their associated FASD and SNR. Here, the pre-event noise duration is long enough to target $\boldsymbol{D_S}$ leading to the same duration for the three noise windows ($\boldsymbol{D_{N1} = D_{N2} = D_{N3} = D_S}$)

manually picked one, and this duration increase may be accentuated for longer records. The three noise windows have the same duration and present similar FASD. However, long pre-event noise is available and this may not be true for each record of every dataset.

Figure 8 presents two examples recorded during the 2014 Cephalonia seismic crisis in Greece, for which the selection of noise and phase windows is uneasy and might be unreliable. Here, the phase ending time has not been picked and is defined by T_{end95}. In the

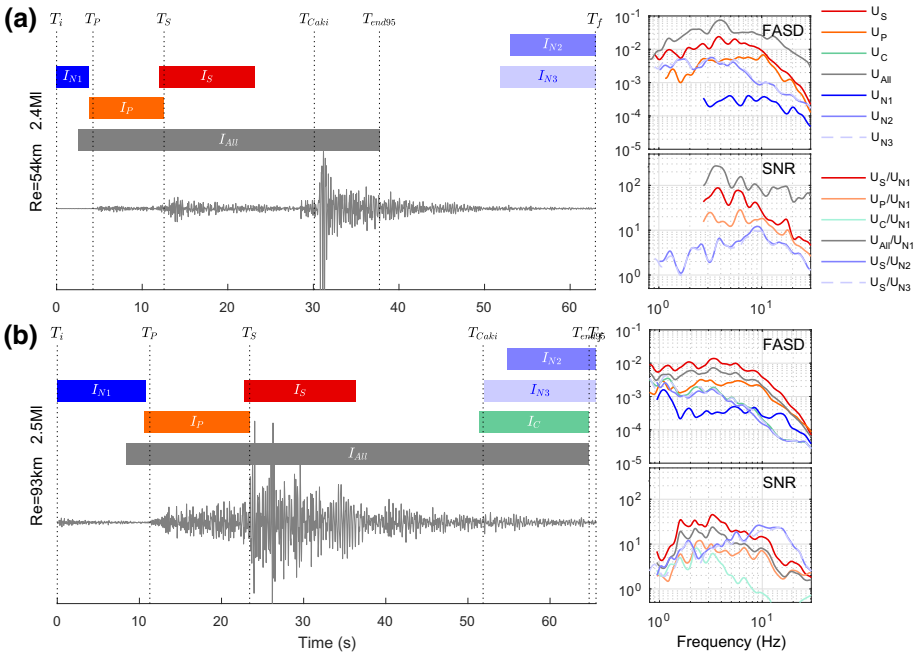


Fig. 8 Examples of uneasy signal phase and noise windowing when no phase ending ($\boldsymbol{T_{end}}$) has been picked and when either after—**(a)** or fore-shock **(b)** are present in the record. Similarly to Fig. 7, all the signal phases ($\boldsymbol{I_P}$, $\boldsymbol{I_S}$, $\boldsymbol{I_C}$ and $\boldsymbol{I_{All}}$) and noise windows ($\boldsymbol{I_{N1}}$, $\boldsymbol{I_{N2}}$ and $\boldsymbol{I_{N3}}$) are represented with their corresponding FASD and SNR. Here, the pre-event noise is limited, leading to test different post-event noise windows durations

first case (a), the target earthquake is followed by a stronger one that strongly biases the T_{end95} estimation and leads to include the P- and S-wave phases of the second earthquake in the coda and the full signal phase of the first event. However, the duration of the coda wave is not long enough here to satisfy the minimal coda wave duration criteria ($D_C < D_{Cmin} = 10$ s). Moreover, for this particular example, the noise selection leads to the rejection of this record (Flag = 0) since the pre-event noise window is too short and the FASD of the two post-event noise windows are too different from the pre-event one.

In the second example, a small fore-shock is present in the pre-event noise while the record is cut before the actual end of the signal. Here the signal phases are not biased but the noise selection is complex. Indeed, I_{N1} is enriched at high frequency compared to I_{N2} and I_{N3} while it is the opposite for low frequency due to the presence of long-period coda waves in the post-event noise window. When using our parameterization ($D_{min} = 10\,\text{s}, F_1 = 5, F_2 = 3, F_3 = 2,$ and $F_4 = 0.67$), the algorithm selects the pre-event noise window since it exhibits sufficient duration for good spectral resolution. The SNR is, however, significantly lower at higher frequency. To avoid such difficulty, we strongly recommend visualizing and flag such records when picking P- and S-wave first arrivals. Picking T_{end} or testing more accurate automatic procedures than the cumulative energy one is also required for an accurate coda and full signal phase window ending.

6 Conclusions

Seismic signal windowing is the preliminary step for many applications in seismology and for SNR verification. While the vast majority of previous studies have used very simple windowing formulations, such as constant duration, we propose a more complex method that takes into account source and propagation terms. This study provides a suitable solution for heterogeneous datasets where the P-wave and S-wave first arrivals have been picked beforehand. The earthquake signal phases are selected exclusively from the T_P and T_S parameters for the majority of the events, which makes the windowing independent of the uncertainties present in the information given by the seismic bulletin. For strong earthquakes (M > 6) that have source durations that are not negligible, a source term is estimated through the inverse of the corner frequency evaluated from the magnitude.

Selecting the noise window can be challenging when large heterogeneous datasets are considered, especially if for some events the duration of the available pre-event noise is short. The noise window has to be the most representative of the noise level, and long enough to allow SNR estimation with good resolution at low frequency. To get around this issue, we defined and tested three different windows: one pre-event and two post-event windows. A scheme is proposed for selecting of the best noise window in terms of duration (as long as possible) and mean energy (as low as possible), without including undesirable transients. This approach gave good results on the KiK-net dataset, with a very limited number of rejected signals.

A Matlab algorithm was developed and can be adapted to each dataset through a few parameters to be defined by the user. This algorithm is freely available as electronic online supplementary material of this paper and ready to be used for every windowing application.

Acknowledgements This work was conducted within the framework of the CASHIMA project (*Cadarache Seismic Hazard Integrated Multidisciplinary Assessment)*, jointly funded by the French Alternative Energies and Atomic Energy Commission (CEA), the Laue-Langevin Institute (ILL) and the International Thermonuclear Experimental Reactor (ITER) organization, and the SIGMA project (*Seismic Ground Motion Assessment*) funded by Electricité De France (EDF), CEA, AREVA and the Ente Nazionale per l'Energia Elettrica (ENEL). We are thankful to Christopher Berrie and to the two anonymous reviewers for their careful suggestions and corrections.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Aki K (1969) Analysis of the seismic coda of local earthquakes as scattered waves. J Geophys Res 74:615–631. doi:10.1029/JB074i002p00615

Aki K, Chouet B (1975) Origin of coda waves: source, attenuation, and scattering effects. J Geophys Res 80:3322–3342

Anderson JG, Hough SE (1984) A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies. Bull Seismol Soc Am 74:1969–1993

Baer M, Kradolfer U (1987) An automatic phase picker for local and teleseismic events. Bull Seismol Soc Am 77:1437–1445

Bonilla LF, Steidl JH, Lindley GT et al (1997) Site amplification in the San Fernando Valley, California: variability of site-effect estimation using the S-wave, coda, and H/V methods. Bull Seismol Soc Am 87:710–730

Borcherdt RD (1970) Effects of local geology on ground motion near San Francisco Bay. Bull Seismol Soc Am 60:29–61

Brune JN (1970) Tectonic stress and the spectra of seismic shear waves from earthquakes. J Geophys Res 75:4997–5009

Douglas J, Gehl P, Bonilla LF, Gélis C (2010) A κ model for Mainland France. Pure Appl Geophys 167:1303–1315. doi:10.1007/s00024-010-0146-5

Drouet S, Chevrot S, Cotton F, Souriau A (2008) Simultaneous inversion of source spectra, attenuation parameters, and site responses: application to the data of the French accelerometric network. Bull Seismol Soc Am 98:198–219. doi:10.1785/0120060215

Drouet S, Cotton F, Guéguen P (2010) Vs30, κ, regional attenuation and Mw from accelerograms: application to magnitude 3–5 French earthquakes: vS30, κ, attenuation and Mw from accelerograms. Geophys J Int 182:880–898. doi:10.1111/j.1365-246X.2010.04626.x

Foundotos L, Hollender F, Laurendeau A et al (2016) Assessing site effects at different scales using different approaches: learnings from generalized inversion of KiK-net data from stiff and rocky sites. Bull Earthq Eng (**to submitted**)

Hanks TC, Kanamori H (1979) A moment magnitude scale. J Geophys Res B 84:2348–2350

Horike M, Zhao B, Kawase H (2001) Comparison of site response characteristics inferred from microtremors and earthquake shear waves. Bull Seismol Soc Am 91:1526–1536

Jongmans D, Campillo M (1993) The response of the Ubaye Valley (France) for incident SH and SV waves: comparison between measurements and modeling. Bull Seismol Soc Am 83:907–924

Kishida T, Ktenidou O-J, Darragh RB, Walter S (2016) Semi-automated procedure for windowing time series and computing Fourier amplitude spectra (FAS) for the NGA-West2 database. Pacific Earthquake Engineering Research Center, Berkeley

Ktenidou O-J (2010) Experimental and theoretical study of seismic ground motion in the city of Aegion, Greece, focusing on the local site and topographic effects. Doctoral thesis, Aristotle University Thessaloniki

Ktenidou O-J, Cotton F, Abrahamson NA, Anderson JG (2014) Taxonomy of Kappa: a review of definitions and estimation approaches targeted to applications. Seismol Res Lett 85:135–146. doi:10.1785/0220130027

Küperkoch L, Meier T, Lee J et al (2010) Automated determination of P-phase arrival times at regional and local distances using higher order statistics. Geophys J Int 181:1159–1170. doi:10.1111/j.1365-246X.2010.04570.x

Laurendeau A, Cotton F, Ktenidou O-J et al (2013) Rock and stiff-soil site amplification: dependency on VS30 and Kappa (K0). Bull Seismol Soc Am 103:3131–3148. doi:10.1785/0120130020

Laurendeau A, Bard P-Y, Hollender F et al (2016) Derivation of consistent hard rock ($1000 < Vs < 3000$ m/s) GMPEs from surface and down-hole recordings: analysis of KiK-net data. Bull Earthq Eng (**to submitted**)

Maggi A, Tape C, Chen M et al (2009) An automated time-window selection algorithm for seismic tomography. Geophys J Int 178:257–281. doi:10.1111/j.1365-246X.2009.04099.x

Mayor J, Calvet M, Margerin L et al (2016) Crustal structure of the Alps as seen by attenuation tomography. Earth Planet Sc Lett 439:71–80. doi:10.1016/j.epsl.2016.01.025

Phillips WS, Aki K (1986) Site amplification of coda waves from local earthquakes in central California. Bull Seismol Soc Am 76:627–648

Satoh T, Kawase H, Matsushima S (2001) Differences between site characteristics obtained from microtremors, S-waves, P-waves, and codas. Bull Seismol Soc Am 91:313–334

Sleeman R, van Eck T (1999) Robust automatic P-phase picking: an on-line implementation in the analysis of broadband seismogram recordings. Phys Earth Planet Int 113:265–275. doi:10.1016/S0031-9201(99)00007-2

Stefano RD, Aldersons F, Kissling E et al (2006) Automatic seismic phase picking and consistent observation error assessment: application to the Italian seismicity. Geophys J Int 165:121–134. doi:10.1111/j.1365-246X.2005.02799.x

Tan Y, He C (2016) Improved methods for detection and arrival picking of microseismic events with low signal-to-noise ratios. Geophysics 81:KS93–KS111. doi:10.1190/geo2015-0213.1

Wong J, Han L, Bancroft J, Stewart R (2009) Automatic time-picking of first arrivals on noisy microseismic data. CSEG 0 02 04 06 08 1:1–4

Zhang H, Thurber C, Rowe C (2003) Automatic P-wave arrival detection and picking with multiscale wavelet analysis for single-component recordings. Bull Seismol Soc Am 93:1904–1912. doi:10.1785/0120020241

Zhao Y, Takano K (1999) An artificial neural network approach for broadband seismic phase picking. Bull Seismol Soc Am 89:670–680

Annexe 2

Accelerometer, velocimeter dense-array and rotation sensor datasets from the Sinaps@ post-seismic survey (2014-2015 aftershock sequence in Cephalonia, Greece)

Perron Vincent[(1),(2),(3)], Hollender Fabrice[(1),(2)], Mariscal Armand[(2)], Theodoulidis Nikolaos[(4)], Andreou Chrisostomos[(5)], Bard Pierre-Yves[(2)], Cornou Cécile[(2)], Cottereau Régis[(6)], Cushing Edward Marc[(3)], Frau Alberto[(7)], Hok Sébastien[(3)], Konidaris Agisilaos[(5)], Langlaude Philippe[(8)], Laurendeau Aurore[(9)], Savvaidis Alexandros[(4)], Svay Angkeara[(10)].

[(1)] CEA, DEN, F-13108 Saint-Paul-lez-Durance, France
[(2)] Univ. Grenoble Alpes, ISTerre, CNRS, IRD, IFSTTAR, F-38041 Grenoble, France
[(3)] IRSN, F-92260 Fontenay-aux-Roses, France
[(4)] ITSAK-EPPO, Institute of Engin. Seismology & Earthquake Engin. GR-55102 Thessaloniki, Greece

(5) TEI, Technological Educational Institute of Ionian Islands, GR-28100 Argostoli, Greece
(6) MSSMat, CNRS, CentraleSupélec, Université Paris-Saclay, F-92295 France
(7) CEA, DEN, F-91191 Gif-sur-Yvette, France
(8) CEREMA, F-06300 Nice, France
(9) CEA, DAM, DIF, F-91297 Arpajon, France (now at Instituto Geofisico Escuela Politecnica Nacional, Quito, Ecuador)
(10) EDF Lab. Paris-Saclay, F-91120 Palaiseau, France

1.1. Abstract

Cephalonia Island is one of the most active seismic zones of the Euro-Mediterranean area due to the activity of the Cephalonia transform fault connecting the Hellenic subduction to NW-Greece collision. Following the two M6+ earthquakes that occurred in early 2014, a post-seismic survey was organized within the framework of the Sinaps@ project near to the city of Argostoli. Different sensor types were deployed for different soil conditions: five accelerometers, a dense array of 21 broadband velocimeters, and one rotation sensor. The survey implementation and the event dataset preparation are presented here. The very rich aftershock sequence allowed high quality recording of thousands of events (signal-to-noise ratio >10). These are gathered into three main datasets: the accelerometric dataset (4147 events), the velocimetric dataset (1843 events), and the rotation dataset (1373 events). These events are associated to seismicity catalogs to compile source metadata. Each event benefited from a visual check and picking of P-wave and S-wave arrival times. The datasets are now open and may be used for various studies.

1.2. Introduction

Cephalonia Island is one of the most active seismic zones of the Euro-Mediterranean area. This high seismicity is mainly due to the activity of the right-lateral Cephalonia transform fault zone (Scordilis et al., 1985), which is located west of Cephalonia Island (Figure 1a, b), and plays an important role in the transition zone between the African subducting plate and the continental Apulian plate (see, e.g., Louvari et al., 1999). Cephalonia Island was struck by several destructive earthquakes during the 20th century, such as that of M7.2 that occurred on August 12, 1953, which raised the island by up to 70 cm (Stiros et al., 1994), killed 455 and injured 2412 people (Papazachos and Papazachou, 1997)

Near the town of Argostoli (the main city on Cephalonia Island), the Quaternary Koutavos sedimentary basin (Figure 1c) has already been the object of instrumentation and various studies. An array with three accelerometers was in operation from 1996 to 2000 (Protopapa et al., 1998; Theodoulidis and Bard, 2016). Then, during the European research project NERA a temporary network was deployed, from September 20, 2011, to April 17, 2012, which included 76 sensors of various types (Theodoulidis et al., 2017). This also included a

dense-array of 21 broadband velocimeters that were designed to study wavefield composition and spatial coherency, and were set up on soft-soil conditions (Imtiaz, 2015).

Given its seismicity and the previous studies that demonstrated the presence of a sedimentary basin with large site effects, this site was chosen for the implementation of a new test site within the framework of the French National Research Agency PIA Sinaps@ project (www.institut-seism.fr/projets/sinaps/) (Berge-Thierry et al., 2016). This test site consists of a vertical array with five accelerometers (one at the surface, four at depth), as well as another accelerometer at the surface on a rock reference site, at about four hundred meters distance northeast of the vertical array. This array has been in operation since July, 2015, and is known as 'ARGONET' (Hollender et al., 2015; Cushing et al., 2016). The long-term objective of this array is the validation of three-dimensional nonlinear computer simulation codes.

Before the installation of the permanent vertical accelerometric array, Cephalonia Island was struck by two significant earthquakes, on January 26, 2014, at 13:55:43 UTC and on February 3, 2014, at 03:08:45 UTC, with magnitudes of M_w=6.1 and M_w=6.0, respectively (Valkaniotis et al., 2014; Karakostas et al., 2015; Karastathis et al., 2015; Sakkas and Lagios, 2015; Theodoulidis et al., 2015; Mavroulis et al., 2017). These earthquakes included the very active subsequent aftershock sequence that triggered the organization of a post-seismic survey (Hollender et al., 2015). This survey involved many instruments (accelerometers, broad-band velocimeters, rotation sensors) that were deployed on different soil conditions. This instrumentation set was provided by the Sinaps@ partners who are involved in the post-seismic survey: CEA (French Alternative Energies and Atomic Energy Commission, France), IRSN (Institute for Radiological Protection and Nuclear Safety, France), CEREMA (Centre for Studies and Expertise on Risks, Environment, Mobility and Development, France), ISTerre (Institute of Earth Science, France), and ITSAK (Institute of Engineering Seismology and Earthquake Engineering, Greece).

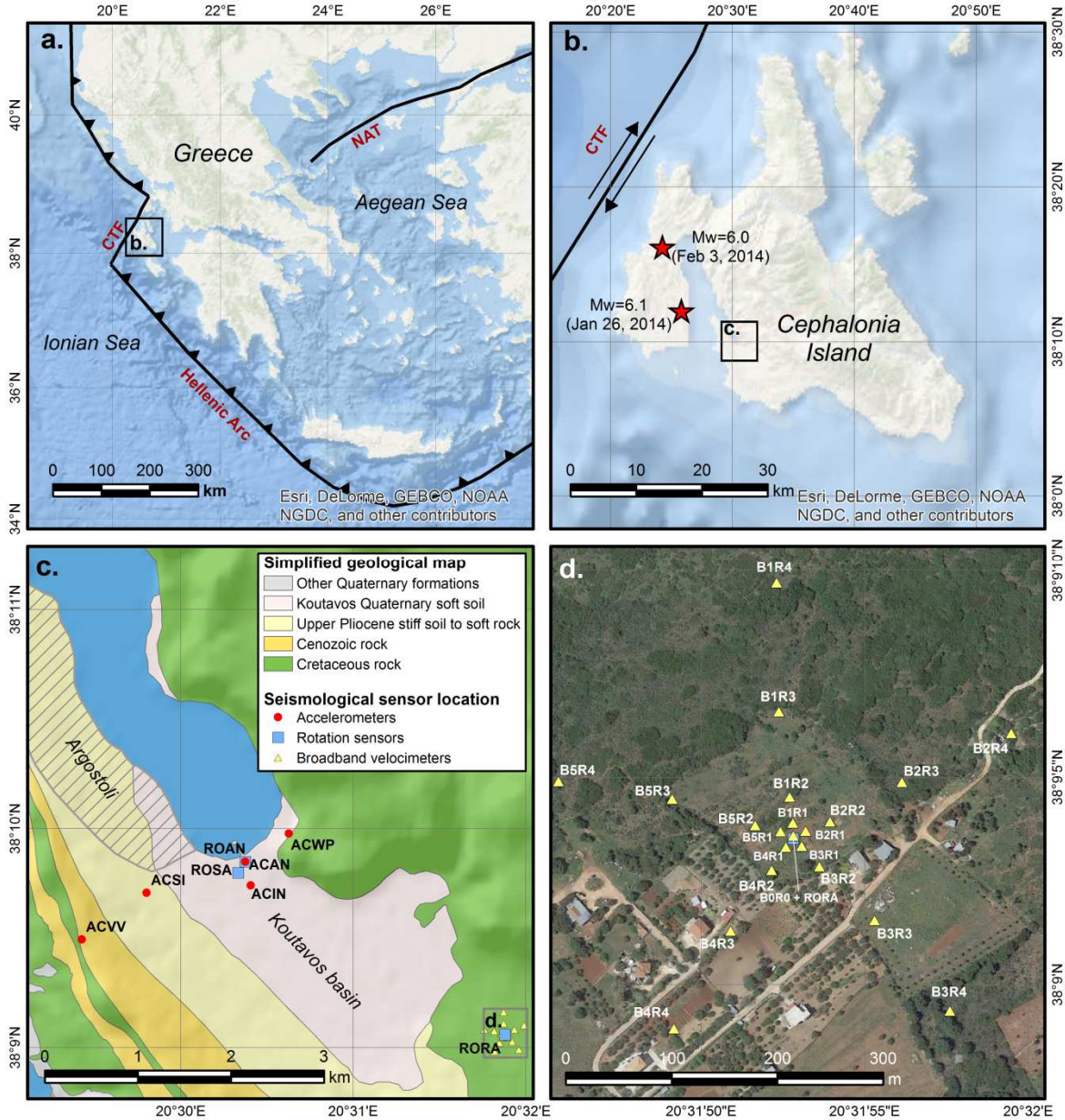


Figure 1: Location of the study site and sensors. (a) Cephalonia Island with respect to Greece, with the main tectonic features (from Karakostas et al., 2015). (b) The Argostoli area with respect to Kefalonia Island, with the position of the two main shocks of the Cephalonia 2014 seismic sequence (from Karakostas et al., 2015). (c) Geological map of the Argostoli–Koutavos area with location of the different sensors (from Cushing et al. 2016). (d) Zoom on the dense-array area (aerial photo: © Google Earth, date of satellite image: April 29, 2013).

The three main motives behind this experiment are associated to the three different datasets presented in this paper:

1. The accelerometers deployed were designed to record any strong motion associated to a possible major aftershock and to anticipate the installation of the permanent ARGONET

array and investigate different possible locations for it. Finally, the associated datasets led to 4147 well recorded events from February 2014 to July 2015, and thus might be interesting for statistical analysis of various ground-motion parameters.

2. The broadband velocimeters were deployed in a dense-array configuration designed to study the short-scale spatial coherency and wavefield composition, specifically for a rock site. This experiment was complementary to the former NERA dense-array that was deployed on a soft-soil site (Imtiaz, 2015) allowing thus the analysis of the effects of rock site conditions on spatial coherency and wavefield composition. The rock site dense-array was deployed from February to March 2014 and led to 1843 well recorded events.

3. Finally, this experiment was the opportunity to complement standard translational measurements by rotation measurements that allowed 'six degrees of freedom' recordings. The rotation measurements were performed for different sites, which also allowed a study of the soil conditions in terms of rotation motions. The rotation measurements led to the recordings of a total of 1373 events for three successive sites.

The present paper presents the data that were recorded during this experiment. These data are now open in the form of well documented datasets of seismic event recordings associated to 'flat-files' that include source metadata and the main features of the recorded signals for all of the events in the datasets.

1.3. Instrument Deployment and Details

For the accelerometric measurements, five Güralp CMG5 accelerometers were used (provided by the Sismob instrumentation pool), which were connected to Nanometrics Taurus digitizers (provided by CEA), and Nanometrics Centaur digitizers (provided by ISTerre). Unfortunately, some of the accelerometers malfunctioned in at least one component and they were only partly operational at the beginning of the experiment. In the event dataset provided, only data that were acquired during the good operation period are provided.

Four out of five accelerometric stations were deployed along a profile that was perpendicular to the Koutavos basin: the ACWP station (accelerometer located near the central water pump station of Argostoli), on a Cretaceous limestone rock site, which is considered as the 'reference station' for site effect studies; the ACAN station (at the position of the future ARGONET array) within Koutavos Park on Quaternary soft-soil; the ACSI station (within the Simatos farm) located on Pliocene stiff soil; and the ACVV station (within the Villa Vassiliki) located on an older Cenozoic rock site (Figure 1c). A fifth station (ACIN) was also installed on the Quaternary soft-soil of the Koutavos basin at the former 'IoniaNet' instrumented borehole site (Protopapa et al., 1998). Within the dataset provided, the ACWP and ACAN sensors recorded from February 6, 2014, to July 2, 2015, and were completed after this date by data that were produced by the ARGONET array (Cushing et al., 2016). The events

recorded at the other stations (i.e., ACVV, ACSI, ACIN) were from October 28, 2014, to July 2, 2015.

The rock site dense-array experiment was carried out using 21 Güralp CMG6TD broadband seismometers with built-in digitizers. These sensors were provided by CEA, IRSN, and CEREMA. The sensors were distributed five by five around four circles with radii of 10, 30, 90 and 180 m, plus a central station (Figure 1d). This geometry led to a star form with five branches (corresponding to five equally spaced azimuths (~72o) with four stations on each branch. Each station is defined by the ID of the branch and the ID of the radius (e.g., station B2R3, “branch #2, radius #3”; B0R0 was thus the central station). The dense-array was set up on Cretaceous limestone rock, about 2 km south-east of the Koutavos Park. The use of seismometers implied good instrumental sensitivity that provides good signal-to-noise ratio for low frequencies, but induced saturation for the strongest motions. The dense-array was operational from February 6, 2014, to March 10, 2014.

Only one Eentec R1 rotation sensor was used, which was connected to a Nanometrics Taurus digitizer (both provided by ISTerre). It was successively placed in three different locations that led to three recording phases. For the first, the rotation sensor was colocated in Koutavos Park, from February 6, 2014, to February 19, 2014, at the center of the former dense-array that was operated during the NERA project experiment, on soft-soil. The rotation sensor station for this period is called ROSA (for soft-soil array). Originally, it was placed in colocation with the accelerometer, but the north component of this accelerometer failed. Note that the NERA soft-soil dense-array was no longer in place for the Sinaps@ post-seismic survey, but the results of the rotation measurement could, however, be compared with the Imtiaz (2015) analysis.

The rotation sensor was then colocated with the broadband seismometer at the center of the rock site dense-array (i.e., B0R0), from February 19, 2014, to March 10, 2014 (i.e., for ~3 weeks). For this period the rotation station was called RORA (for rock array). This dataset allows comparisons between rotation and translation data, and also comparisons between rotation measurements with the dedicated rotation sensor and rotation derived from the ‘translation sensor array’. Note, however, that the strongest events cannot be analyzed for this phase because the seismometers were saturated.

At the end of the dense-array experiment, the rotation sensor was moved back in the Koutavos basin and was installed in colocation with the accelerometer ACAN, and operated from March 11, 2014, to July 2, 2015 (~16 months). The rotation station was called ROAN for this period (location of the then-future ARGONET array). This dataset allows comparisons between rotation and translation data, without any upper limit in terms of strong motion, which thus allows analysis over a wider range of ground motion amplitude levels and event magnitudes.

1.4. Overall Data preparation, Quality and Availability

All of the sensor signals were recorded in continuous mode, using a sampling rate of 200 Hz. To extract relevant time windows from these raw recordings that correspond to earthquake ground motions, two seismicity catalogs were used. The first was the National Observatory of Athens (NOA) catalog, which covers the whole recording period and includes earthquakes across the whole area of Greece and the surrounding region. The NOA catalog provides local magnitudes (ML) for all of the events, as well as moment magnitudes (MW) for the highest magnitude events. The second catalog used is based on the seismicity analyses of Karakostas et al. (2015). This catalog provides moment magnitudes, but it is limited to analysis of the Cephalonia aftershock sequence and runs to the end of June 2014. As this catalog is based on precise relocation work, it might provide more accurate source parameters. Information from both catalogs is included in the dataset provided. Each event of these datasets related to the accelerometers and the rotation sensor is at least linked to the source parameter characterization provided by the NOA catalog, and sometimes complemented by the Karakostas et al. (2015) catalog. As for the broadband seismometer dense-array dataset, a few events are linked only to the Karakostas et al. (2015) catalog, as they were not identified in the NOA catalog.

The extracted time windows were visually checked, to include only the well recorded signals that correspond to the source parameters from the seismicity catalogs. During these visual checks, the arrival times of both P-waves and S-waves were also picked for the ACWP time histories (for the accelerometer dataset, and for the rotation datasets linked to ROSA and ROAN) or for the B0R0 time histories (for the seismometer dense-array dataset, and for the rotation dataset linked to RORA). Note, however, that these pickings are only indicative and are designed to identify the different signal phases. It is not intended that they be used, for example, for precise relocation work.

For the data dissemination, these other selection criteria were applied as the final 'post-processing' event selection:

- accelerometer dataset: 1/ presence of records for at least both the ACWP and ACAN stations; 2/ signal-to-noise ratio >10 at ACWP and ACAN;
- broadband velocimeter dense-array dataset: 1/ inclusion of records for B0R0; 2/ inclusion of records for at least 19 (of the 21) sensors across the whole array; 3/ signal-to-noise ratio >10 at B0R0. Note that events that produced (or might produce) saturation were kept in the dataset; however, a 'flag' that informs of these possible saturation conditions is included within the corresponding flat-file;
- rotation sensor datasets: ROSA: signal-to-noise ratio >10; RORA: 1/ signal-to noise-ratio >10 at both B0R0 and RORA; 2/ presence of records for at least 20 seismometers across the whole array; 3/ exclusion of events with possible saturation on seismometers; ROAN: signal-to-noise ratio >10 at both ACAN and ROAN.

The signal-to-noise ratios used to select these events were estimated by measuring the peak values (acceleration, velocity, or rotation rate) on the seismic event time windows and noise windows before the event signals after applying a 1-10 Hz band-pass filter. Of note, this filter was just used for the signal-to-noise analysis, with the data being made available without filtering.

The data are provided in physical units (acceleration, velocity, or rotation rate). Acceleration and broadband velocimeter datasets are provided without phase corrections from the instrumental responses. However, within the frequency band of interest for earthquake engineering applications, these corrections are negligible and we prefer to leave the interested users who really need 'pole and zero' correction to do this themselves. This choice is also motivated on the basis that all of the accelerometers on the one hand, and all of the velocimeters on the other, are from the same devices and thus have the same instrumental response. Contrariwise, the data associated to the rotation sensors are fully corrected for instrumental responses, as the sensor used presents a relatively complex response, even within the earthquake engineering frequency band.

Table 1 summarizes the mean features of the different datasets. The accelerometer dataset led to 4127 events, with maximum peak ground acceleration of 3.72 m/s² recorded at ACAN. The broadband velocimeter dense-array recorded 1843 events, even if the period of operation was relatively short (5 weeks) due to high velocimeter sensitivity. Even if the potentially saturated events are kept in the dataset, they only represent around 35 events. As the overall instrumental noise of the rotation sensor is relatively high in comparison with the accelerometer instrumental noise (and hence much higher than the broadband velocimeter instrumental noise), this means that the rotation sensor datasets have fewer events in comparison to those involving only the other sensors. The three rotation sensor datasets led to 421, 118, and 834 events for ROSA, RORA, and ROAN, respectively. Note that for the ROAN dataset, 29 events were also included from the permanent ARGONET array best records (up to July 28, 2016), to complete the dataset with the strongest motions. The ROAN dataset included ground motion rotation up to 6.29 mrad/s associated to the event that reached a peak ground acceleration of 3.72 m/s² at ACAN.

Figure 2 shows the maps of the earthquakes contained in each dataset. Even if the Cephalonia aftershock sequence dominates all of the datasets, numerous other events are also included within the accelerometer, broadband velocimeter dense-array and ROAN datasets. This might allow these datasets to be used with selections of the events to avoid 'single path' or 'single azimuthal' effects. Figure 3 shows the distribution of each dataset in terms of magnitude, hypocentral distance, and peak values. Figure 4 gives an example of recordings (time histories) that correspond to one event (M_w=4.1) for the five accelerometers, which occurred on November 24, 2014, with a hypocentral distance of 25 km.

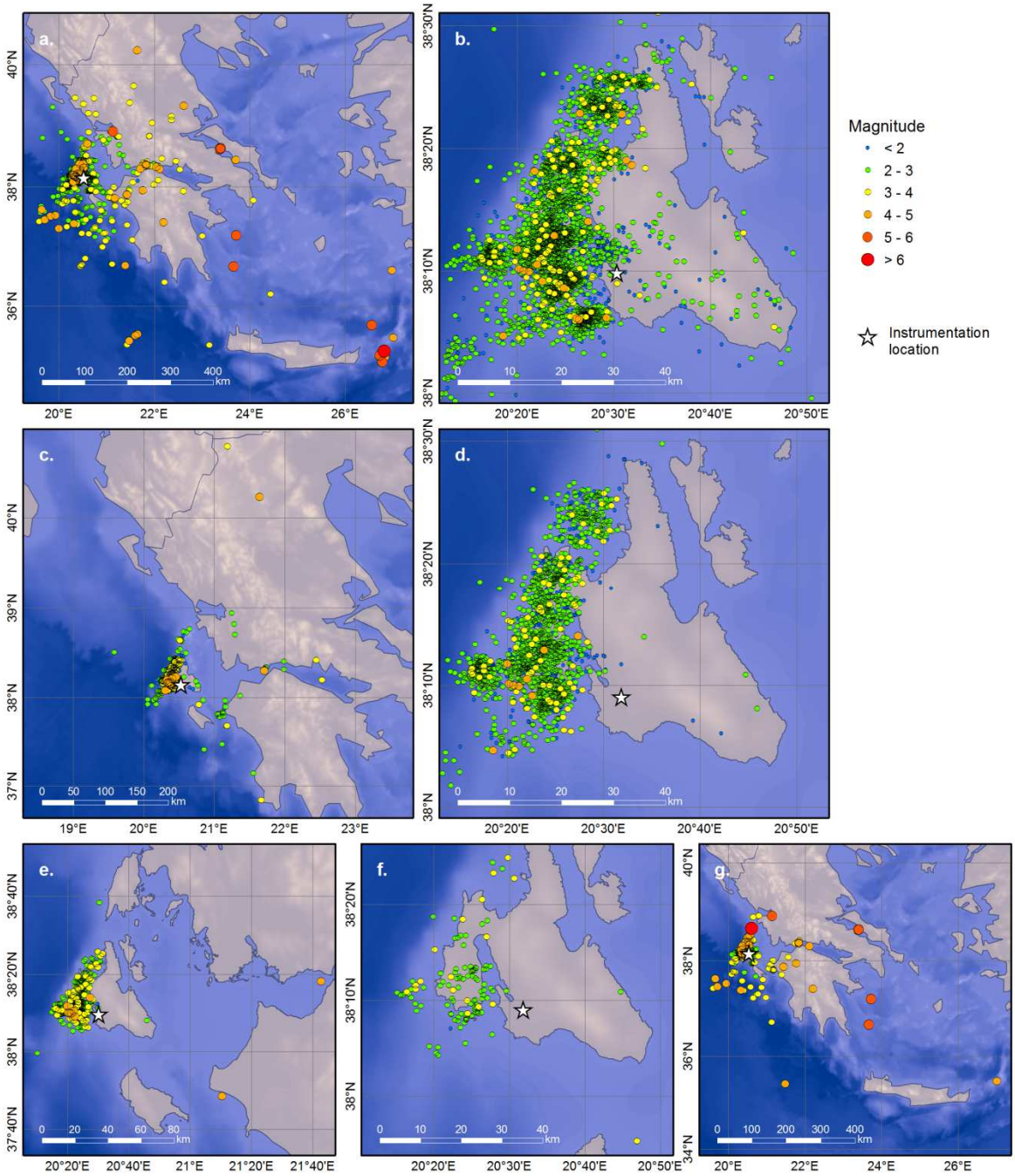


Figure 2: Event maps of the different datasets. (a) Accelerometer dataset – whole view. (b) Accelerometer dataset – zoom of Cephalonia Island. (c) Broadband velocimeter dense-array dataset – whole view. (d) Broadband velocimeter dense-array dataset – zoom of Cephalonia Island. (e) ROSA rotation sensor dataset. (f) RORA rotation sensor dataset. (g) ROAN rotation sensor dataset.

Table 1: Main features of the different event datasets from the Sinaps@ post-seismic survey.

	Accelerometer	Broadband velocimeter (rock site dense array)	Rotation sensor		
			ROSA (provided alone)	RORA (provided with velocimeter dense-array data)	ROAN (provided with colocated accelerometer data)
Number of events	4147	1843	421	118	834[(2)]
Date of the first event	Feb 6, 2014	Feb 6, 2014	Feb 6, 2014	Feb 19, 2014	Mar 11, 2014
Date of the last event	Jul 2, 2015	Mar 10, 2014	Feb 19, 2014	Mar 10, 2014	Jul 28, 2016[(2)]
Maximum peak value	3.72 m/s² (at ACAN)	3.4 mm/s[(1)]	2.06 mrad/s	0.23 mrad/s	6.29 mrad/s
Maximum magnitude (M_W)	6.0	4.8[(1)]	4.8	3.6	6.4[(2)]

[(1)] Maximum amplitude limited by velocimeter saturation. [(2)] Including the period extension using the ARGONET data.

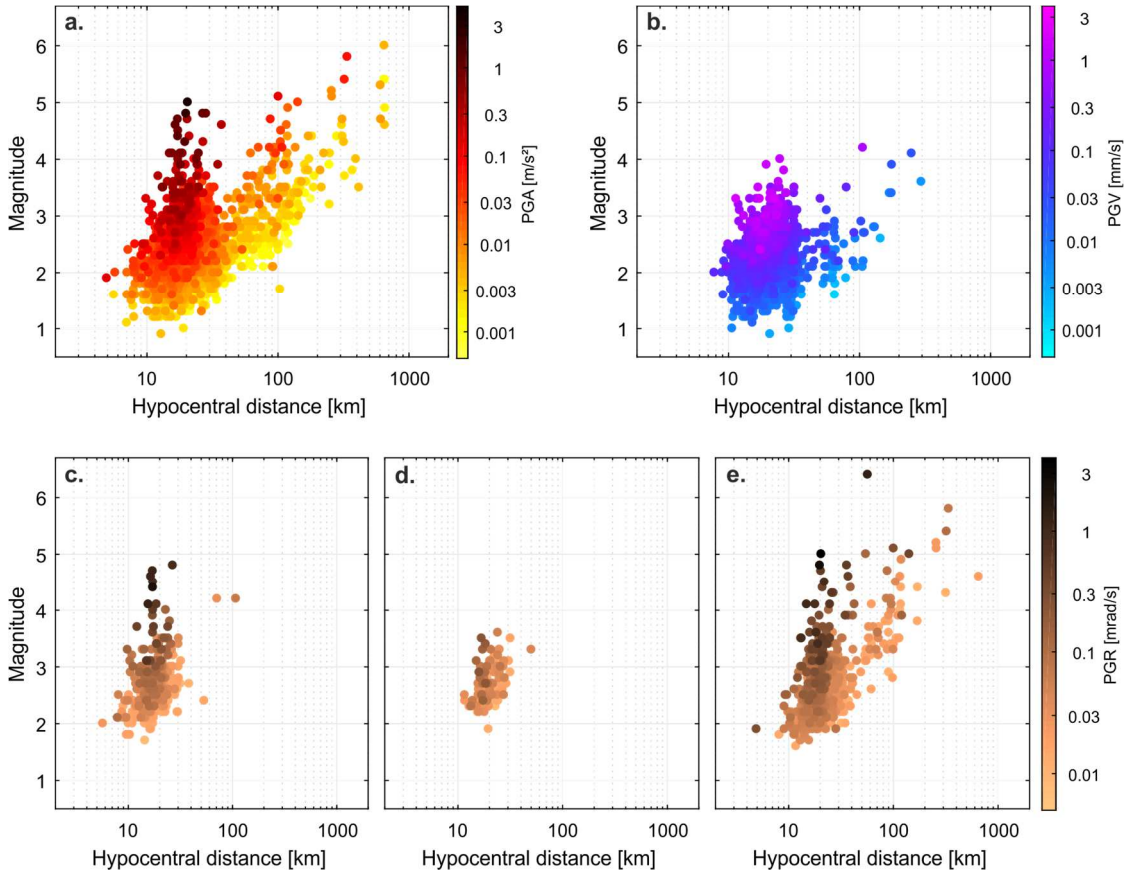


Figure 3: Magnitude versus hypocentral distance distributions of the different datasets ('peak values' are given through the color scales). (a) Accelerometer dataset. (b) Broadband velocimeter dense-array dataset. (c) ROSA rotation sensor dataset. (d) RORA rotation sensor dataset. (e) ROAN rotation sensor dataset.

The datasets are open and available on repository internet sites given in the 'data and resources' section. Each dataset is provided in SAC and ASCII, and in Matlab 'mat file' format. The datasets are provided with 'flat-files' that gather all of the important event metadata (i.e., location, magnitude, from NOA and Karakostas et al., 2015, catalogs), the information on records (i.e., peak value, signal-to-noise ratio, among others). Complete descriptions of the 'flat-file' content and the Matlab file structure are available on the repository websites. The accurate sensor location was determined using differential GNSS measurements, post-processed using data from the permanent NOA GNSS network (NOANET; Ganas et al., 2011).

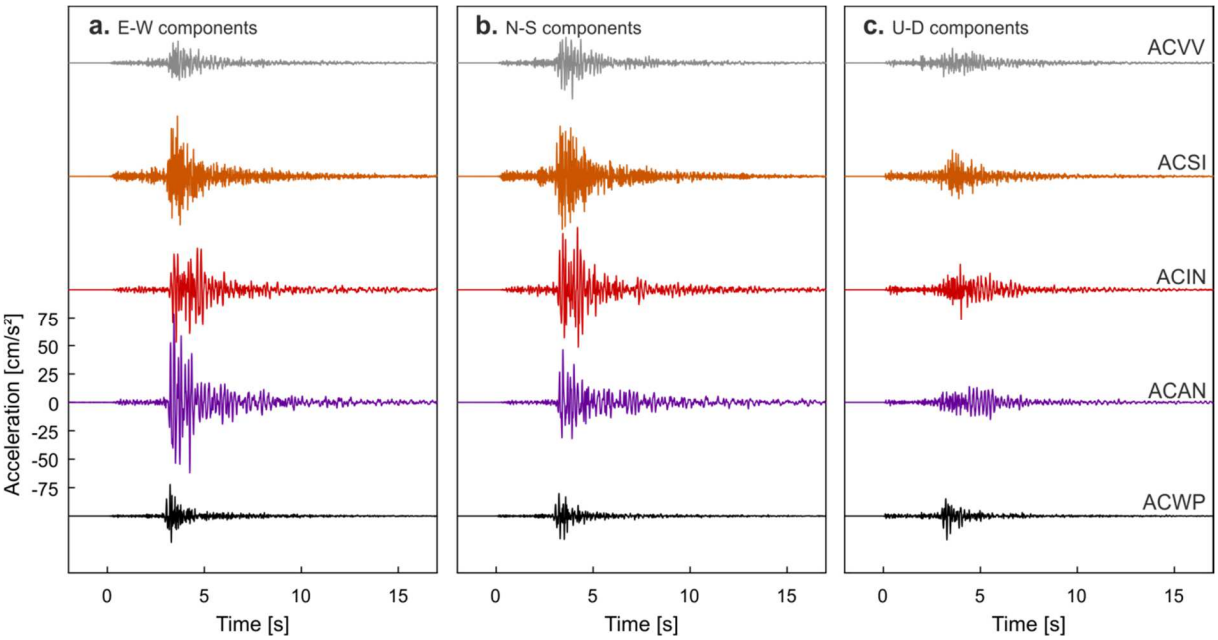


Figure 4: Example of time histories recorded with the five accelerometers (event occurred on November 24, 2014, M_W=4.1, hypocentral distance: 25 km). Amplification of signals with respect to the reference station ACWP is coherent with the known geology.

1.5. Initial observations and results

The datasets presented here have already been the subject of studies within the framework of the Sinaps@ project, especially on event subsets of the broadband velocimeter dense-array experiment. For instance, Svay et al. (2017) analyzed the plane-wave spatial coherency on 100 events and compared results to several widely used coherency models from the literature. Imtiaz et al. (2017) analyzed the lagged spatial coherency for about 40 events, as well as the wavefield composition (e.g., amount of Love and Rayleigh waves within the whole seismic energy) and they compared the results with those obtained within the NERA project soft-soil dense-array (Imtiaz, 2015). The rotation data were studied by Sbaa et al. (2016) and Sbaa et al. (2017) in a comparison of their results from the three datasets presented here and those in the literature. They highlighted the importance of soil

conditions on the rotation wavefield. Other studies are in progress, which are taking advantage of the high number of events within the accelerometer database to compute statistics on standard spectral ratio (SSR) (Borcherdt, 1970) analysis of site amplification using station ACWP as reference. As a preliminary result, Figure 5 shows the SSR analysis for ACVV, ACSI, ACIN, and ACAN with respect to ACWP, with the associated standard deviation. This analysis was performed using the Perron et al. (2017) algorithm to select seismic phase windows. Figure 5 also shows a simplified geological cross-section of the valley, to facilitate the interpretation of the SSR results. The ACIN and ACAN site amplification is large, as up to 10 for ACAN, with a fundamental frequency of around 1.6 Hz. The ACSI site amplification remains relatively low (~2) up to 10 Hz. The higher amplification at higher frequency might be attributed here to a shallow soft layer. These observations suggest that the main velocity contrast between the sedimentary layers occurred between the Quaternary soft-soil unit and the Pliocene stiff soil unit. ACVV station shows low broadband amplification associated to standard deviation that is globally larger than for other sites, possibly due to topographic site effects (Maufroy et al., 2015).

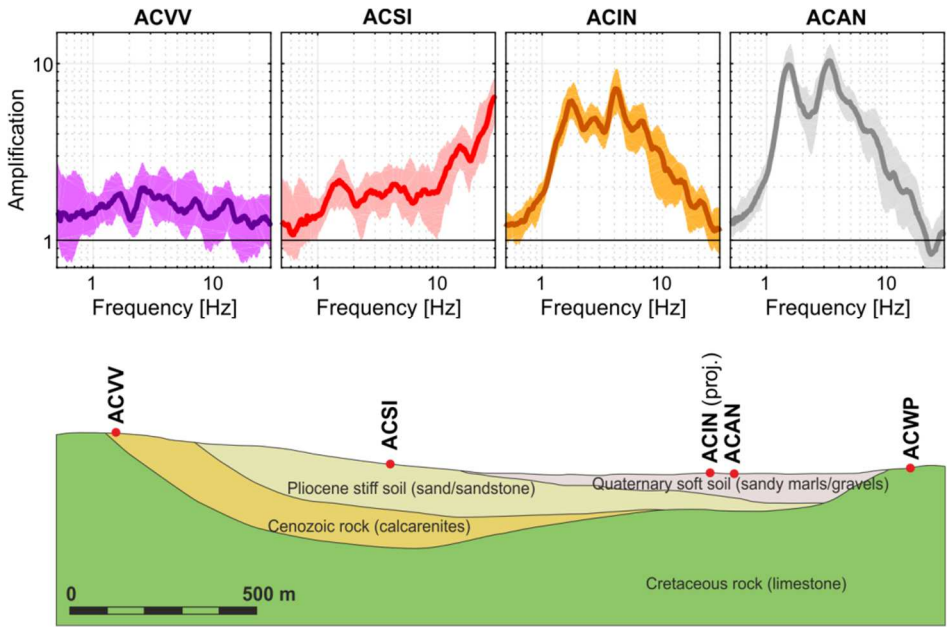


Figure 5: Top: Site amplification estimation using standard spectral ratio analysis for ACVV, ACSI, ACIN, and ACAN stations, with respect to ACWP station, with the associated standard deviation. Bottom: Simplified cross-section along the station profile.

1.6. SUMMARY

The Sinaps@ post-seismic survey started after two successive M6+ earthquakes that occurred in early 2014 on Cephalonia Island, Greece. It succeeded in producing thousands of well recorded events (signal-to-noise ratio >10) associated to seismicity catalogs that are distributed in several event datasets (accelerometer, broadband velocimeter dense-array,

rotation sensor) that are now open for dissemination. Part of this data has already been analyzed. However, they could surely be used in further research efforts. We hope that the preparation of these datasets (visual check, picking of P-wave and S-wave arrivals, 'flat-files') will help to minimize the time needed for data preparation for further studies.

1.7. DATA AND RESOURCES

The datasets presented in this paper can be uploaded from one of these three repository websites: www.isterre.fr/XXX; http://www.institut-seism.fr/projets/sinaps/XXX; http://argonet-kefalonia.org/data/XXX. Please cite the present paper if you use these datasets. If there is any need to use the raw data for a specific research project, please contact the authors.

The raw data associated to accelerometer sensors are also available from the EIDA web portal, using the temporary public network code "YP", which began in 2014. The numbering of the stations used in the present paper is different to the previous paper used for the EIDA site: (ACWP=AC01; ACAN=AC03; ACIN=AC09; ACSI=AC10; ACVV=AC11).

1.8. ACKNOWLEDGEMENTS

This study was carried out under the Sinaps@ project that receives French funding managed by the French National Research Agency under the program "Future Investments" (Sinaps@ reference: ANR-11-RSNR-0022). Sinaps@ is a "Seism Institute" project (http://www.institut-seism.fr/en/). The authors would like to thank all Sinaps@ partners who did not take part in the post-seismic survey itself but who helped us in its logistical organization: Catherine Berge-Thierry (CEA), Etienne Bertrand (CEREMA), Didier Clouteau (CentraleSupélec), Aline Dechamp (CEA), Fernando Lopez-Caballero (CentraleSupélec), Olivier Stéphanovitch (ISTO), Sandrine Roussel (ISTerre) and many others. We are also grateful to people who took part in the visual checks and the first arrival picking of the events: L. de Abreu Corrêa, M. Benjelloun, L. de Carvalho Paludo, F. Gatti, J.-B. Perroni and S. Sbaa. C. Berrie provided valuable improvement of the manuscript.

1.9. REFERENCES

Berge-Thierry, C., Svay, A., Laurendeau, A., Chartier, T., Perron, V., Guyonnet-Benaize, C., Kishta, E., Cottereau, R., Lopez-Caballero, F., Hollender, F., Richard, B., Ragueneau, F., Voldoire, F., Banci, F., Zentner, I., Moussallam, N., Lancieri, M., Bard, P.-Y., Grange, S., Erlicher, S., Kotronis, P., Le, M., Nicolas, M., Régnier, J., Bonilla, F. and Theodoulidis, N., 2016. Toward an integrated seismic risk assessment for nuclear safety improving current French methodologies through the SINAPS@ research project: Nuclear Engineering and Design,, p. In press.

Borcherdt, R.D., 1970. Effects of local geology on ground motion near San Francisco Bay: Bulletin of the Seismological Society of America, 60, p. 29–61.

Cushing, E.M., Hollender, F., Guyonnet-Benaize, C., Perron, V., Imtiaz, A., Svay, A., Bard, P.-Y., Cottereau, R., Lopez Caballero, F., Theodoulidis, N., Moiriat, D. and Gélis, C., 2016. Close to the lair of Odysseus Cyclops : the SINAPS@ postseismic campaign and accelerometric network installation on Kefalonia Island – site effect characterization experiment. In: Active Tectonics and Archeoseismology (PATA) Crestone, Colorado, USA.

Ganas, A., Chousianitis, K., Drakatos, G., Papanikolaou, M., Argyrakis, P., Kolligri, M., Petrou, P., Batsi, E. and Tsimi, C., 2011. NOANET: high-rate GPS network for seismology and geodynamics in Greece. In: Geophysical Research Abstracts, EGU General Assembly.

Hollender, F., Perron, V., Imtiaz, A., Svay, A., Mariscal, A., Bard, P.-Y., Cottereau, R., Lopez Caballero, F., Cushing, E.M., Theodoulidis, N.P. and Moiriat, D., 2015. Close to the lair of Odysseus Cyclops: the SINAPS@ post-seismic campaign and accelerometric network installation on Kefalonia Island. In: Ninth National AFPS Conference Marne-la-Vallée, France.

Imtiaz, A., 2015. Seismic wave field, spatial variability and coherency of ground motion over short distances: near source and alluvial valley effects. PhD thesis. Grenoble Alpes University, France.

Imtiaz, A., Perron, V., Svay, A., Hollender, F., Bard, P.-Y. and Theodoulidis, N.P., 2017. Wave-field characteristics and coherency of seismic ground motion from a rock site at Argostoli Greece (paper #1743). In: Proceedings of the 16th World Conference on Earthquake Engineering Santiago, Chile.

Karakostas, V., Papadimitriou, E., Mesimeri, M., Gkarlaouni, C. and Paradisopoulou, P., 2015. The 2014 Kefalonia Doublet (Mw6.1 and Mw6.0), Central Ionian Islands, Greece: Seismotectonic Implications along the Kefalonia Transform Fault Zone: Acta Geophysica, 63, p. 1–16, doi: 10.2478/s11600-014-0227-4.

Karastathis, V.K., Mouzakiotis, E., Ganas, A. and Papadopoulos, G.A., 2015. High-precision relocation of seismic sequences above a dipping Moho: the case of the January–February 2014 seismic sequence on Cephalonia island (Greece): Solid Earth, 6, p. 173–184, doi: 10.5194/se-6-173-2015.

Louvari, E., Kiratzi, A. and Papazachos, C., 1999. The Cephalonia Transform Fault and its extension to western Lefkada Island (Greece): Tectonophysics, 308, p. 223–236, doi: 10.1016/S0040-1951(99)00078-5.

Maufroy, E., Cruz-Atienza, V.M., Cotton, F. and Gaffet, S., 2015. Frequency-Scaled Curvature as a Proxy for Topographic Site-Effect Amplification and Ground-Motion Variability: Bulletin of the Seismological Society of America, 105, p. 354–367, doi: 10.1785/0120140089.

Mavroulis, S., Carydis, P., Alexoudi, V., Grambas, A. and Lekkas, E., 2017. The January-February 2014 Cephalonia (Ionian Sea, Western Greece) earthquakes: tectonic and seismological aspects. In: Proceedings of the 16th World Conference on Earthquake Engineering Santiago, Chile.

Papazachos, B. and Papazachou, C., 1997. The earthquakes of Greece: Editions ZITI, Thessaloniki.

Perron, V., Laurendeau, A., Hollender, F., Bard, P.-Y., Gélis, C., Traversa, P. and Drouet, S., 2017. Selecting time windows of seismic phases and noise for engineering seismology applications: a versatile methodology and algorithm: Bulletin of Earthquake Engineering,, doi: 10.1007/s10518-017-0131-9.

Protopapa, E., Papastamatiou, D., Michaelides, O. and Gazetas, G., 1998. The Ionianet accelerometer array: early results and analysis. In: Proceedings of the 11th European Conference on Earthquake Engineering Paris, France.

Sakkas, V. and Lagios, E., 2015. Fault modelling of the early-2014 ~M6 Earthquakes in Cephalonia Island (W. Greece) based on GPS measurements: Tectonophysics, 644–645, p. 184–196, doi: 10.1016/j.tecto.2015.01.010.

Sbaa, S., Hollender, F., Perron, V., Imtiaz, A., Bard, P.-Y., Mariscal, A., Cochard, A. and Dujardin, A., 2017. Analysis of rotation sensor data from the Sinaps@ Cephalonia (Greece) post-seismic experiment - link to surface geology and wave-field characteristics: Earth, Planets and Space (submitted to),.

Sbaa, S., Imtiaz, A., Hollender, F., Perron, V., Mariscal, A. and Bard, P.-Y., 2016. Analysis of rotation sensor data from the Sinaps@ Kefalonia post-seismic experiment. In: Proceedings of the 5th IASPEI / IAEE International Symposium Taipei, Taiwan.

Scordilis, E.M., Karakaisis, G.F., Karacostas, B.G., Panagiotopoulos, D.G., Comninakis, P.E. and Papazachos, B.C., 1985. Evidence for transform faulting in the Ionian Sea: The Cephalonia Island earthquake sequence of 1983: Pure and Applied Geophysics, 123, p. 388–397.

Stiros, S.C., Pirazzoli, P.A., Laborel, J. and Laborel-Deguen, F., 1994. The 1953 earthquake in Cephalonia (Western Hellenic Arc): coastal uplift and halotectonic faulting: Geophysical Journal International, 117, p. 834–849, doi: 10.1111/j.1365-246X.1994.tb02474.x.

Svay, A., Perron, V., Imtiaz, A., Zentner, I., Cottereau, R., Clouteau, D., Bard, P.-Y., Hollender, F. and Lopez-Caballero, F., 2017. Spatial coherency analysis of seismic ground motions from a rock site dense array implemented during the Kefalonia 2014 aftershock sequence: Earthquake Engineering & Structural Dynamics,.

Theodoulidis, N. and Bard, P.-Y., 2016. Site effects and non-linear soil behavior study using data from the IONIANET borehole array. In: Proceedings of the 35th European Seismological Commission General Assembly Trieste, Italy.

Theodoulidis, N., Cultrera, G., Cornou, C., Bard, P.-Y., Boxberger, T., Di Giulio, G., Imtiaz, A., Kementzetzidou, D., Makra, K. and NERA Team, 2017. Basin Effects on Ground Motion: The Case of a High-Resolution Experiment in Cephalonia (Greece): Bulletin of Earthquake Engineering (submitted to),.

Theodoulidis, N., Karakostas, C., Lekidis, V., Makra, K., Margaris, B., Morfidis, K., PAPAIOANNOU, C., Rovithis, E., Salonikios, T. and Savvaidis, A., 2015. The Cephalonia, Greece, January 26 (M6.1) and February 3, 2014 (M6.0) Earthquakes: Near-Fault Ground Motion and Effects on Soil & Structures: Bulletin of Earthquake Engineering,.

Valkaniotis, S., Ganas, A., Papathanassiou, G. and Papanikolaou, M., 2014. Field observations of geological effects triggered by the January–February 2014 Cephalonia (Ionian Sea, Greece) earthquakes: Tectonophysics, 630, p. 150–157, doi: 10.1016/j.tecto.2014.05.012.