Analysis and modeling of earthquake strong-motion site effects on Icelandic arrays for earthquake engineering applications

Saharalsadat Rahpeyma

Faculty of Civil and Environmental Engineering
University of Iceland
2018

Analysis and modeling of earthquake strong-motion site effects on Icelandic arrays for earthquake engineering applications

Saharalsadat Rahpeyma

Dissertation submitted in partial fulfillment of a
Philosophiae Doctor degree in Civil Engineering

Supervisor
Dr. Benedikt Halldórsson

PhD Committee
Dr. Benedikt Halldórsson
Dr. Sigurjón Jónsson
Dr. Russell A. Green
Dr. Birgir Hrafnkelsson

Opponents
Dr. John Douglas
Dr. Dominik H. Lang

Faculty of Civil and Environmental Engineering
School of Engineering and Natural Sciences
University of Iceland
Reykjavík, 2018

Analysis and modeling of earthquake strong-motion site effects on Icelandic arrays for earthquake engineering applications
Dissertation submitted in partial fulfilment of a *Philosophiae Doctor* degree in Civil Engineering

Faculty of Civil and Environmental Engineering
School of Engineering and Natural Sciences
University of Iceland
Tæknigarði, Dunhaga 5
107 Reykjavík
Iceland

Telephone: 525 4700

Bibliographic information:
Saharalsadat Rahpeyma, 2018, *Analysis and modeling of earthquake strong-motion site effects on Icelandic arrays for earthquake engineering applications*, PhD dissertation, Faculty of Civil and Environmental Engineering, University of Iceland, 215 pp.

ISBN 978-9935-9438-1-1
Author ORCID 0000-0001-6299-1514

Printing: Háskólaprent
Reykjavík, Iceland, November 2018

To my beloved family

for their unlimited love and constant encouragement

تقدیم به

پدر و مادر عزیز و مهربانم

که همواره یاوری دلسوز و فداکار و پشتیبانی محکم و مطمئن برایم بوده‌اند

Abstract

A comprehensive site-effect investigation framework was developed and applied to earthquake strong-motion data collected on small-aperture urban arrays, the ICEARRAY I in Hveragerði in the South Iceland Seismic Zone and ICEARRAY II in Húsavík in the Tjörnes Fracture Zone in the north, where considerable variations over short distances in ground motion amplitudes have routinely been observed. Through the Horizontal-to-Vertical Spectral Ratio and Standard Spectral Ratio methods the amplification levels were shown to remain relatively low on lava rock, but predominant frequencies of resonance were found that vary systematically and correlate with the local soil structure and geological units. For stations on lava-rock characterized by one or more velocity reversals due to softer sedimentary layers at depth, modeling the soil structure as a two-degree-of-freedom dynamic system captures the observed predominant frequencies. For non-reversal sites, an inversion procedure for the velocity profile based on Bayesian statistical theory was developed. Furthermore, a versatile Bayesian Hierarchical Model (BHM) was developed and applied to peak ground acceleration data, quantifying the contribution of earthquake source, wave propagation and local geological effects to PGA variations. This study thus improves the modeling of site effects in Iceland, quantifies the variabilities of physical parameters of the subsoil through a Bayesian inversion technique, and through the new BHM shows that for ICEARRAY I the earthquake effect dominates the variability while for ICEARRAY II, the site effects are dominating. The results of this study thus facilitate our understanding of local ground motions and have practical implications for urban planning and seismic hazard assessment.

Útdráttur

Í þessu verkefni voru staðbundin mögnunaráhrif efstu jarðlaga á jarðskjálftabylgjur greind með jarðskjálfta- og jarðsuðsgögnum frá ICEARRAY hröðunarmælafylkingunum í Hveragerði og á Húsavík, en þær einkennast af umtalsverðum breytileika á einkennum og umfangi yfirborðshreyfinga í jarðskjálftum yfir afar stuttar vegalengdir (tugi til hundruði metra). Hlutfallsgreining tíðnirófa sýndi að kerfisbundin mögnunaráhrif á tveimur sveiflutíðnum eru til staðar á þeim hluta Hveragerðis sem stendur á hrauni vegna viðsnúnings bylgjuhraða í mýkri setlögum undir tveimur hraunlögum. Greining á mögnun út frá hefðbundnu tveggja frígráðu sveiflukerfi sýnir að sveiflutíðnirnar eru bein afleiðing þessa viðsnúnings í hraða. Fyrir hefðbundin jarðlög var líkan af bylgjuútbreiðslu í lagskiptu efni notað til að meta efniseiginleika jarðlaganna og óvissu þeirra út frá andhverfuaðferðum og Bayesískri tölfræði. Mismunaáhrif jarðskjálftahreyfinga voru greind með nýju stigskiptu Bayesísku tölfræðilíkani sem ákvarðar framlag jarðskjálftaupptaka, útbreiðsluáhrifa, og staðbundinna áhrifa á jarðskjálftahreyfingarnar. Niðurstöðurnar sýna að hve miklu leyti fjölbreytt jarðfræði og jarðskjálftaupptök hafa áhrif á breytileika jarðskjálftahreyfinganna, og óvissugreiningin sýnir hversu marktækur munurinn er. Í jarðskjálftunum á Suðurlandi árið 2008 er breytileikinn í yfirborðshreyfingum í Hveragerði rakinn til jarðskjálftaupptakanna fyrst og fremst. Mögnunin og breytileiki hennar í Hveragerði voru bæði minni en á Húsavík í jarðskjálftunum undan Norðurlandi 2012-2013, en þar sýndu niðurstöður að staðbundnar jarðfræðilegar aðstæður höfðu afgerandi og meiri áhrif. Í þessari rannsókn hefur nýjum eðlisfræðilegum og tölfræðilegum líkönum verið beitt við greiningu á jarðskjálftahreyfingum og breytileika þeirra yfir stuttar vegalengdir. Niðurstöðurnar hafa hagnýta þýðingu því þær sýna hvaða líkön henta best eftir tegund jarðlaga og greiningaraðferðirnar skilgreina óvissu betur en áður hefur verið gert. Slíkt er forsenda bætts mats á jarðskjálftavá í byggð á Íslandi, sem hefur bein áhrif á mat á áhættu af völdum jarðskjálfta, bætingu jarðskjálftahönnunar og hagkvæmni byggðarskipulags.

List of Publications

This dissertation is based on the work presented in papers published in or submitted to International Scientific Indexing (ISI) journals and peer-reviewed proceedings of major conferences on Earthquake Engineering and Engineering Seismology as following:

1. Rahpeyma S., Halldorsson B., Olivera C., Green R. A., and Jónsson S. (2016). Detailed site effect estimation in the presence of strong velocity reversals within a small-aperture strong-motion array in Iceland. *Soil Dynamics and Earthquake Engineering, 89,* 136-151. doi:10.1016/j.soildyn.2016.07.001
2. Rahpeyma S., Halldorsson B., Hrafnkelsson B., and Jónsson S. (2018). Bayesian hierarchical model for variations in earthquake peak ground acceleration within small-aperture arrays. *Environmetrics,* 29(3). doi:10.1002/env.2497
3. Rahpeyma S., Halldorsson B., Hrafnkelsson B., Green R. A., and Jónsson S. (2018a). Site effect estimation on two Icelandic strong-motion arrays using a Bayesian hierarchical model of spatial distribution of earthquake peak ground acceleration. *Soil Dynamics and Earthquake Engineering, (Accepted, in press)*.
4. Rahpeyma S., Halldorsson B., Hrafnkelsson B., Green R. A., Polat O., and Jónsson S. (2018b). Estimate of shear-wave velocity profile using microseismic Horizontal-to-Vertical Spectral Ratios inversion. *Soil Dynamics and Earthquake Engineering, (Manuscript completed)*.
5. Halldorsson B., Olivera C., Rahpeyma S., Ólafsson S., Green R. A., and Snæbjörnsson J. T. (2016). On the HVSR estimation at Icelandic strong-motion stations. In *Proceedings of the Nordic Geotechnical Meeting*. Reykjavík, Iceland: Icelandic Geotechnical Society.
6. Snæbjörnsson J. T., Rahpeyma S., Halldorsson B., and Sigtryggsdóttir F. G. (2016). Effects of soil-structure interaction on the excitation and response of RC buildings subjected to strong-motion. In *Proceedings of the Nordic Geotechnical Meeting*. Reykjavík, Iceland: Icelandic Geotechnical Society.
7. Rahpeyma S., Halldorsson B., and Green R. A. (2017). On the distribution of earthquake strong-motion amplitudes and site effects across the Icelandic strong-motion arrays. In *16th World Conference on Earthquake Engineering (16WCEE)* (Paper no. 2762.). Santiago, Chile.
8. Rahpeyma S., Halldorsson B., and Hrafnkelsson B. (2017). Bayesian hierarchical model of peak ground acceleration for the Icelandic strong-motion arrays. In *International Conference on Earthquake Engineering and Structural Dynamics (ICESD)*. Reykjavík, Iceland.
9. Rahpeyma S., Halldorsson B., Hrafnkelsson B., and Polat O. (2018c). Shear-wave velocity modeling by inversion of microseismic Horizontal-To-Vertical Spectral Ratio. In *16th European Conference on Earthquake Engineering (16ECEE)* (Paper no. 11620). Thessaloniki, Greece, 18-21 June 2018.

Table of Contents

List of Figures

List of Tables

Abbreviations and Parametrizations

Parameter, Variable and Abbreviation	Explanation
SSR	Standard Spectral Ratio
HVSR	Horizontal-to-Vertical Spectral Ratio
SISZ	South Iceland Seismic Zone
TFZ	Tjörnes Fracture Zone
GMM	Ground Motion Model
V_S, β	Shear-wave velocity (m/s)
$V_{S,30}$	the average value of shear-wave velocity in the upper 30 meters of the subsoil structure
μ	shear modulus of the subsoil material
ρ	Density of the subsoil
k	Stiffness of the sedimentary layers
m	Mass of lava layers
μ_{es}, $\boldsymbol{\mu}$	Predicted median ground motion
δB_e, $\boldsymbol{\delta B}$	Event effect or inter-event residuals
δS_s, $\boldsymbol{\delta S}$	Station effect or inter station residuals
δWS_{es}	Event-station term or event-station residuals
δE_{es}	Unmodeled effects
τ	Inter-event standard deviation
ϕ	Intra-event standard deviation
ϕ_{S2S}	Inter-station standard deviation
ϕ_{SS}	Event-station standard deviation
ϕ_R	Unmodeled effects standard deviation
σ	Total standard deviation
Δ_{S2S}	Range parameter of Matérn covariance function of station terms
ν_{S2S}	Smoothness factor of Matérn covariance function of station terms
Δ_{SS}	Range parameter of Matérn covariance function of event-station terms
ν_{SS}	Smoothness factor of Matérn covariance function of event-station terms
$\boldsymbol{\Sigma}_{S2S}$	Covariance matrix of station effects
$\boldsymbol{\Sigma}_{SS}$	Covariance matrix of event-station effects

Acknowledgement

First and foremost, I would like to express my most sincere gratitude to my supervisor Dr. Benedikt Halldórsson for his continuous support through this four-year journey. You have been a tremendous mentor for me and without your guidance, motivation, and immense knowledge it would have been impossible to complete this dissertation. I would like to thank you for your trust, encouraging my research and for allowing me to grow as a research scientist.

I would like to extend my sincere gratitude to my doctoral committee members: Professor Sigurjón Jónsson for his constructive comments which incentivized me to widen my thoughts as well as my research from various perspectives, Professor Russell A. Green for his never-ending support and valuable feedbacks, also Professor Birgir Hrafnkelsson for his significant influence and invaluable contribution to this work.

I dedicate many special thanks to Professor Apostolos S. Papageorgiou from University of Patras and Dr. Steven Gibbons from NORSAR, who provided me with the opportunity of spending traineeship programs in their institutes to extend my experience in different aspects of my future career. I wish to thank Professor Orhan Polat from Dokuz Eylül University and Professor Brady Cox from the University of Texas at Austin for providing me with the data and valuable contribution.

The Earthquake Engineering Research Centre (EERC) in Selfoss is a memorable and unique place for me where I had my first experience of living and working abroad. My special thanks to Elínborg who is been always helpful and supportive. I wish to thank Símon, Sólveig, Rajesh and Puja for making the Centre feel like warm and bright home. Without you I could not survive dark and cold Icelandic winter.

Some special words of gratitude go to my amazing fellows, Tim and Milad for the stimulating discussions and for the sleepless nights at the EERC and VRII when we were working together before deadlines and for all the fun we have had during the last four years. I wish you the best of luck in your professional and personal life. To all my dear friends, in particular, Narges who has always been a major source of support when things would get a bit discouraging. I love and appreciate all of you so very much.

This work was made possible by the Icelandic Centre for Research through a Research Grant of Excellence no. 141261-051/52/53. Additionally, a Doctoral grant from the Eimskip Fund of the University of Iceland, and a grant from the University of Iceland Research Fund enabled this work. The basis of this work is the ICEARRAY I strong-motion dataset, recorded on instruments funded by a Marie Curie Reintegration Grant in 2008, and the ICEARRAY II strong-motion dataset, funded through a Rannís Equipment Grant in 2012 with their deployment funded by the Icelandic Catastrophe Insurance (Grant no. S112-2013). Additionally, the inhabitants of Hveragerði and Húsavík, and their municipalities, are acknowledged for their generosity and assistance in housing the recording equipment and for their dedication and support to the ICEARRAY project.

Last, but by no means least, I would like to thank my family, my lovely parents, Amir and Farzaneh, and my wonderful sisters, Sara and Asal, for encouraging me to follow my instincts and supporting me spiritually throughout this wonderful experience and my life in general. You have been always the true source of love, strength and motivation to get things done. I love you and could not have imagined any of this without you.

1 Introduction

1.1 The Purpose Statements

Earthquakes are one type of natural disaster that pose a great risk to the built environment and life safety. Although it is impossible to either predict or prevent earthquakes, it is possible to mitigate their deleterious effects and, thus, reduce the number of deaths, injuries, and damage. In this regard, a fundamental challenge for earthquake engineers, seismologists, and geologists is enhancing their knowledge to be better able estimate of the level and variability of ground motion from future earthquakes.

Essentially, the main factors that influence earthquake strong ground motion can be divided into source, path, and site effects. Site effects, defined as the variation of ground shaking in space, amplitude, frequency content and duration (Pitilakis 2004) due to differences in the geologic profiles, have received considerable attention during the last decades. In this regard, the research efforts in Japan, where as early as the 1930s site effects have been well recognized through pioneering works by Sezawa and Ishimoto (Kawase and Aki 1989). Evidence from past earthquakes such as 1985 Mexico City earthquake (Mexico); 1989 Loma Prieta earthquake (Northern California); 1995 Kobe earthquake (Japan); and 1999 Chi-Chi earthquake (Taiwan) illustrates that the distribution of seismic wave propagation can vary extremely due to heterogenous subsoil structures, even over relatively small areas. In other words, local geological and geostructural conditions strongly account for modifications of seismic wave amplitudes and propagation patterns and, therefore, the damage distribution in a relatively limited region (Aki 1988, 1993; Steidl 1993; Steidl et al. 1996; Field 2000; Pitilakis 2004). Accordingly, the estimation of the influence of site effects is one of the important steps in any reliable seismic hazard assessment and an important criterion in engineering applications. Figure 1.1 shows a multi-level and inter-disciplinary schema proposed by the Pacific Earthquake Engineering Research (PEER) center that links researchers in different field to mitigate earthquake disaster.

In current practice, site response analysis is a fundamental part of assessing seismic hazard in earthquake prone areas. In Iceland, however, the earthquake resistant design of structures generally oversimplifies the site amplification effects and considers uniform site response of rock for their applications. This is contrary to the complex geological features in Iceland consisting of repeating sedimentary and lava-rock layers resulting from glaciation and deglaciation together with sea level fluctuations. In 2002 however, Bessason and Kaynia illustrated considerable variation in site amplification due to different subsoil structure beneath the east and west abutments of the 80-meter long Thjorsa Bridge in South Iceland. The findings of their research emphasized the importance of the site effect investigations even over relatively short distances (Bessason and Kaynia 2002). Unfortunately, the lack of dense and local recordings of earthquake ground motions in Iceland has largely prevented researchers from establishing a reliable assessment of site amplification effects and a consistent model for strong ground motion variation across limited areas. Both of these are of paramount importance to accurately assessing the seismic hazard and properly apply the building code.

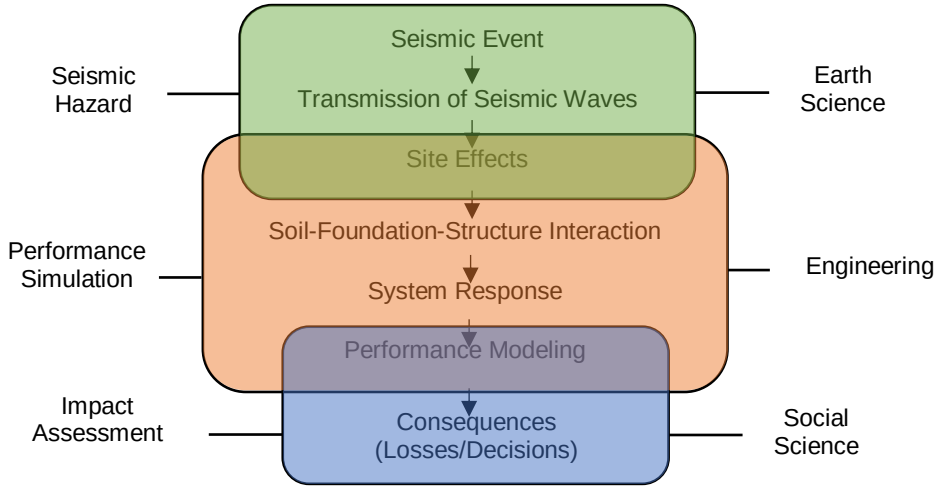

Figure 1.1 Performance-based seismic assessment proposed by Pacific Earthquake Engineering Research Centre (PEER).

Hence, the main goal of the Ph.D. research presented herein is to rigorously analyze and model site-specific site response characteristics of engineering significance using the unique dataset recently recorded by Icelandic strong-motion arrays and to tie these site response characteristics to the corresponding profile's geological features. In addition, a statistical model for describing the distribution of the strong ground motion parameters associated with the site characteristics has been developed. The outputs of the research presented herein are vital for performing reliable seismic hazard assessments, as well as microzonation studies, that provide inputs for urban planning and for evaluating the vulnerability of the structures.

1.2 Research Objectives

In order to setup a practical framework to analyze and model the characteristics of the earthquake strong-motion site effects, the current research fulfils the following main objectives:

Research objective #1: *To outline a detailed and reliable scheme for site effect estimation across two strong-motion arrays in Iceland*

The first objective of this research is to develop a framework for estimating site response effects that can be readily incorporated in the assessment of seismic hazard in earthquake prone areas. As introduced briefly in Section 2.3, there are several empirical techniques that can be used to evaluate site response characteristics, and the choice of method is mainly based on the importance and the nature of the project. In this research, two different practical and widely used techniques, Standard Spectral Ratio (SSR) and Horizontal to Vertical Spectral Ratio (HVSR) methods, are used to characterize site-specific site amplifications. For the purpose of evaluating the stability and reliability of the results, microseismic recordings collected at strong-motion stations are analyzed using the HVSR method. Applying different approaches and using different datasets helps us to explore the systematic variation in amplification, to include predominant frequencies, across the arrays and link this variability to the geological conditions. The results of site effect estimation across two Icelandic strong-motion arrays (ICEARRAY I and II) are discussed in Section 4.1 and the obtained results for ICEARRAY I is published in Rahpeyma et al. (2016).

Research objective #2: *To estimate shear-wave velocity models of geologic profiles by inverting experimental microseismic HVSR using a Bayesian framework.*

As a result of the research objective #1, there is a strong correlation between site amplification and geological features beneath the station. The evaluation of the site effects quantitatively requires one-, two-, or three-dimensional shear-wave velocity models of subsoil down to bedrock. The V_S profile is of paramount interest in many engineering applications. In particular, seismic building codes such as Eurocode 8 (European Committee for Standardization 2003), American Society of Civil Engineers (ASCE 2007, 2010), and also National Earthquake Hazard Reduction Program as well as site classification scheme such as NEHRP (Building Seismic Safety Council, BSSC 2003; Boore 2004a) consider $V_{S,30}$ (i.e., the average shear-wave velocity down to 30 m depth) to classify sites for evaluating the dynamic behavior of the near-surface geologic profile.

In this thesis, estimation of the soil properties is the main topic of research objective #2. The V_S profiling has been accomplished using a variety of either invasive (e.g., down-hole or cross-holes seismic surveys) or non-invasive (e.g., surface-wave or body-wave approaches, and refraction-reflection analyzes) processing tools as well as a various types of laboratory tests. The classical in-situ methods usually use borehole drilling to investigate the key soil properties and consequently approximate the V_S profile (Kramer 1996). However, although the in-situ material testing can provide the most detailed and accurate information with reasonable resolution between closely spaced boreholes, it is relatively expensive and time consuming to amass the measurements to cover the whole area under study; therefore, these methods are primarily recommended in projects of relative importance. A general introduction to the different methods to obtain V_S profile is presented in Section 2.5.

The objective #2 of this thesis is focused on modeling the subsoil V_S profile by inverting the experimental microseismic HVSR. The inversion problem is established based on Bayesian framework to estimate better understanding of the model parametrization and quantitative information about the associated variabilities. Two test/nominated stations are selected to model the V_S profile (Rahpeyma et al. 2018b).

Research objective #3: *To decompose ground motion parameters and quantify the spatial variability of strong ground motion amplitude*

The primarily analysis of strong-motion recordings of ICEARRAY I and II revealed that despite the small inter-station distances and assumed uniform site conditions there is a noticeable variation in strong-motion amplitudes across both arrays. Research objective #3 aims to model the spatial distribution of peak ground acceleration (PGA) and also to quantify the associated variabilities across both Icelandic arrays. A multi-level Bayesian Hierarchical Model (BHM) is proposed to determine the contribution of source, propagation path, and site characteristics and associated uncertainties.

It has long been known that the reliability of the estimation of the uncertainty in ground motion amplitude has a significant influence on the precision of the computed seismic hazard. Commonly, the random variability of ground motions is divided into two components: (1) aleatory variability (i.e., natural randomness in a process) and (2) epistemic uncertainty (i.e., limited knowledge or data of the system). Since reduction of the standard deviation is a critical issue in seismic hazard mitigation studies, especially at low probability

levels (Atkinson 2006), many researchers focus their efforts on decreasing uncertainty by improving physically justified Ground Motion Models (GMMs). However, in spite of the different and more complicated functional forms of GMMs (Abrahamson and Shedlock 1997; Abrahamson et al. 2008; Douglas and Aochi 2008; Strasser et al. 2009), total uncertainty in GMMs has not changed considerably over the last 50 years. Accordingly, the investigation of the main sources of variabilities and underlying mechanics associated with attenuation relationships is a challenging research topic that has important implications for critical structures (Bommer et al. 2004; Bommer and Abrahamson 2006; Atkinson 2006; Kowsari et al. 2017). In recent decades, many researchers have tried to split up the variabilities into independent terms to understand the calculated variability (Joyner and Boore 1981; Fukushima and Tanaka 1990; Abrahamson and Youngs 1992; Douglas and Gehl 2008; Kuehn and Scherbaum 2015).

In this thesis, therefore, the total variability is split into inter-event and intra-event variabilities. In particular, intra-event variability is split into inter-station, event-station, and measurement and model error variabilities. The comparison between the results from the two arrays illustrates the influence of the complexity of subsoil structure and geological effects. Establishing quantitative estimates of strong-motion spatial variability will develop the understanding of the key factors that affect the variation of seismic ground motions across even a relatively small area. This estimation is critical for detailed microzonation and decision making for urban planning. The proposed methodology in this thesis can be also applied to different datasets with similar constraints (Rahpeyma et al. 2018 & 2018a).

1.3 Dissertation Organization

This dissertation is divided into two parts. Part I contains the following chapters:

- Chapter 1: Introduction
- Chapter 2: Background
- Chapter 3: Aim
- Chapter 4: Applications and Results
- Chapter 5: Conclusion and future perspectives

Part II (Appendix) presents scientific papers based on this work. Three ISI papers have been accepted/published based on the research presented in the current dissertation. The manuscript of the fourth paper is completed. The papers are presented in full as appendices.

2 Background

2.1 Historical Evidence of Site Effects

Significant enhancement in knowledge of source mechanism, seismic waves propagation, and localized geological effects have been consequences of improvement in quantity and quality of the observations. A practical approach to reach this goal is essentially installation of closely spaced networks or seismic arrays in regions with high potential of seismicity. The first such array that became operational is SMART-1 (1980-1991) consists of 37 accelerographs installed in a circular aperture of $D = 2$ km in a seismic region of Taiwan with the aim of collecting data for earthquake engineering applications and seismological studies of the near-field events (Bolt et al. 1982; Abrahamson 1988). During deployment of the SMART-1 from 1980 to 1991, it has recorded considerable number of accelerogram traces from 48 earthquakes ranging in local magnitude from 3.6 to 7.0 and epicentral distances ranged from 3-200 km from the array center, and focal depths from shallow to 100 km. The recorded earthquakes had both reverse and strike- slip focal mechanisms associated with the subduction zone and transform faults. Peak ground accelerations have been recorded up to 0.33 g and 0.34 g on the horizontal and vertical components, respectively.

In accordance with the main geological feature of the region, almost all SMART-1 stations are located on a uniform site condition of an alluvial plain of the Lanyang River; however, initial analyses of the strong-motion recordings highlighted dissimilarities in ground motion intensities and site amplifications between stations. The differences between site responses across the array and the availability of high-quality data recorded during the deployment of the SMART-1 attracted many researchers to investigate the main source of the site response variations. In particular, the dynamic properties of the site conditions and spatial variability of earthquake motions were analyzed using the SMART-1 dataset (Bolt et al. 1982; Loh et al. 1983; Loh 1985; Harichandran and Vanmarcke 1986; Harichandran 1991; Chiu et al. 1995; Theodulidis and Bard 1995; Kiureghian 1996; Dimitriu et al. 2000). The analysis revealed that the variabilities in site amplification are closely related to the local subsoil geological construction (Theodulidis and Bard 1995; Beresnev et al. 1995; Dimitriu et al. 2000).

Although the site amplification variation between SMART-1 stations due to the geological features of the region attracted the attention of many researchers, the vital role of the spatial variability of ground motion intensities was exclusively highlighted for the first time after 1985 Mexico City (Michoacan) earthquake. This large ($M_s = 8.1$) subduction zone earthquake with epicenter located more than 300 km away from Mexico City resulted in severe structural collapse, injuries, and deaths. Enormous variations in ground shaking and distribution of the structural damages were observed in different parts of the city (Celebi et al. 1987; Seed et al. 1988). In particular, while in the southwest of the city ground motion intensities were moderate and structural damage was not considerable, in the northwest part of the city, destructive damage and building collapse were observed due to the severe ground shaking (Roullé and Chávez-García 2006). It is noteworthy that this variation in structural damage has been observed in previous earthquakes during the past decades; however, the

differences resulting from the 1985 Mexico City earthquake were somewhat more accentuated and noticeable. Furthermore, digital strong motion instruments were in operation in different parts of Mexico City and, therefore, recorded the ground-motions across the city during and after the event. The recorded data have provided a unique opportunity to explore the effects of source (Houston and Kanamori 1986; Eissler et al. 1986; Singh et al. 1988), site conditions (Ordaz and Singh 1992; Cárdenas et al. 1997; Cárdenas-Soto and Chávez-García 2003), intensities of shaking and amplification measurements as well as the associated effects on structural damage (Seed et al. 1988).

The geological and geostructural features of the Mexico City can be divided into three main parts for microzonation studies. The city is largely built on an ancient lake bed sediment; a relatively firm layer of gravel and fill at the surface ($5-10$ m thick), an underlying soft clay layer ($25-35$ m thick), and beneath that, a firm sand layer ($2-5$ m thick). The western part of the city sits on a hill characterized by a surface layer of lava flows or volcanic tuffs. Also, there is a transition zone composed of alluvial sandy and silty deposit layers between the hill and the lake bed zones (Beck and Hall 1986). Evidently, during the 1985 Mexico City earthquake, the severe structural failures mainly occurred on the lake bed deposits while the rest of the city did not show noticeable failures. Within the heavily damage area, the intensity of structural damage varied depending on the height of the structures, presumably reflecting the impact of geological structure, the intensity and frequency characteristics of the ground motions, the dynamic response characteristics of the structures and the criteria controlling the design of the structures. Investigations into site effects revealed that on the ancient lake bed the intensity of shaking was much greater than that to which the buildings were designed, due to resonance phenomenon (Beck and Hall 1986). This resonance effect resulted in the amplification of long period seismic waves that were prominent in the base-rock motions due to the long site-to-source distance and was the main cause of high-rise building collapse. In addition, long duration of the ground motion during 1985 Mexico City earthquake led to increase the amount of damage (Ansal et al. 2004).

Hence, the Mexico City earthquake and other recent destructive earthquakes (e.g., Armenia 1988, Loma Prieta 1989, Philippines 1990, Northridge 1994, Kobe 1995, Izmit 1999, Athens 1999, Chi-Chi Taiwan 1999, Haiti 2010, Christchurch 2011) have shed light on the importance of site effects. These destructive earthquakes have repeatedly demonstrated the influence of subsoil conditions on the destructive potential of ground shaking in terms of fatalities, economic losses, and functional disruption. In this context, reliable and accurate evaluation of site response in seismic regions, particularly urban areas, represents an important target in the framework of seismic-risk mitigation strategies, engineering applications, and lifeline network design.

2.2 Basic Physical Concept

As can be seen in Figure 2.1, the physical process of an earthquake ground motion recorded at a station can be divided into three main factors directly associated with causative fault (source effects), source-to-site travel paths (path effects), and local geological conditions (site effects) (Chin and Aki 1991; Boore 2004b; Yoshida 2015). Theoretically, the earthquake strong-motion observed at a station can be modeled as a function of the source, path, and site effects within the context of the well-known stochastic modeling approach (Boore 1983, 2003). Thus, the discrete Fourier amplitude spectrum of ground motion or response associated with an earthquake ($i = 1, \dots, N_E$) of seismic magnitude M_{0i} observed

at a receiver s ($j = 1, \dots, N_S$) with source-to-site distance of R_{ij}, and at frequency, f_k ($k = 1, \dots, N_F$), can be written in the form of the convolution of the main factors as Eq. (2.1):

$$Y_{ij}\left(M_{0i}, R_{ij}, s_j, f_k\right) = S_i(M_{0i}, f_k) \otimes P_{ij}(R_{ij}, f_k) \otimes G_j(s_j, f_k) \otimes I(f_k) \quad (2.1)$$

Where $\otimes$ indicates convolution, $S_i(M_{0i}, f_k)$ is the earthquake source spectrum, $P_{ij}(R_{ij}, f_k)$ is path effects, $G_j(s_i, f_k)$ accounts for site-specific site response and a frequency transfer function $I(f_k)$ for the type of motion (displacement, velocity, acceleration, harmonic response of a single-degree-freedom oscillator). The Source and path effects can be largely determined using either different models or numerical estimations and the site effect is defined as the main topic of the current Ph.D. research.

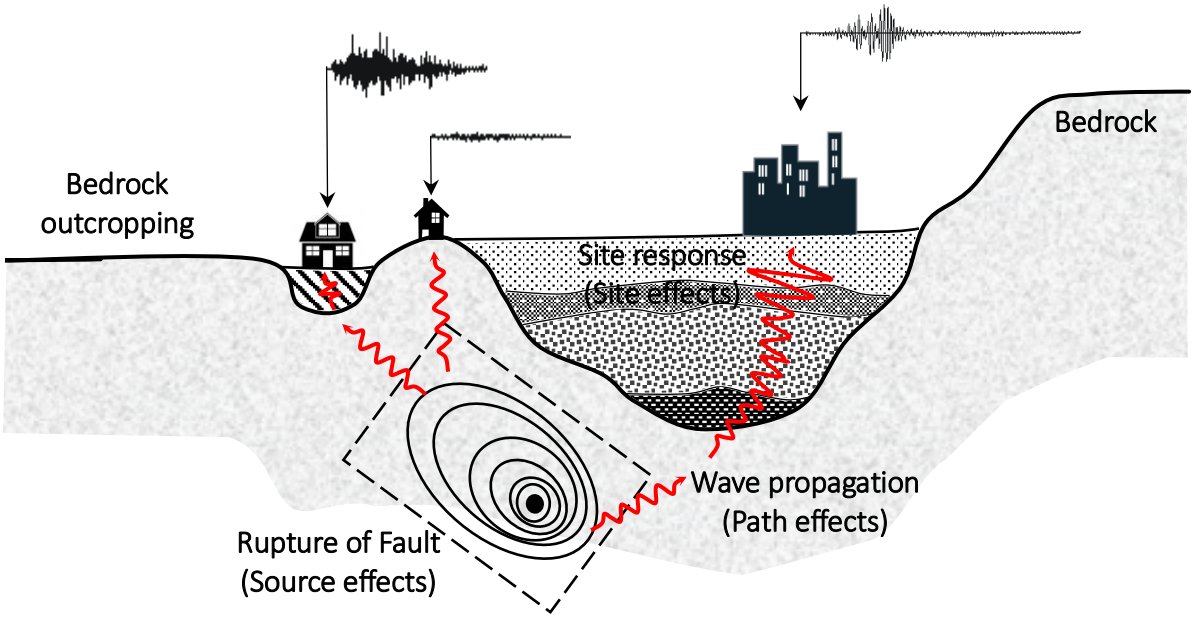

Figure 2.1 Schematic illustration of the wave propagation from fault to ground surface.

The main factors that characterize a site are the geometry of the soil stratigraphy (i.e., thickness and lateral discontinuities of the various soil layers), the mechanical and dynamic properties of the soil and rock materials, and the topographical conditions of the surface. In this regard, the term "*site effects*" accounts for the effects of local geology and topography in the modulation of seismic wavefield at a recording station, and via site effect estimation we try to investigate the main characteristics of the subsoil as well as their consequences on ground response at a site.

In general, site effects manifest in the near surface geology above the impedance contrast between soil deposits and the underlying bedrock or stiffer stratum, where the soil deposits act as a filter to incoming seismic waves and amplify motions having certain frequencies (Boore 2004b). Ground motions recorded on sites classified as "soil" are often larger in amplitude relative to those recorded on "rock" sites (Gutenberg 1957; Boore and Joyner 1997). The leading cause is that the deposits that form low-velocity layers near the Earth's surface trap energy, amplify or de-amplify some ranges of frequencies due to the decrease or increase in seismic impedance, and preferentially amplify resonant frequencies. Several researchers have shown that for layers of given thickness, the relative shaking response will be greatest where the shallowest geologic units have the lowest impedance values and where

the impedance contrast between the surface layer and the underlying one is the greatest (Ansal et al. 2004).

In addition to the geological effects, site amplification due to topography has been identified in theoretical as well as empirical studies. The top of isolated hills, elongated crests, edges of plateaus and cliffs are usually zones of amplification due to diffraction and focusing. The main results are that the topographic amplification is maximum at the top of the hill and is maximum at the frequency at which one shear wavelength equals the width of the hill base. Motions on the hillsides are not amplified much, and motions around the base of the hill are usually de-amplified with respect to motions far from the hill.

2.3 Methods to Estimate Site Effects

Due to improvement in seismic instrumentation and consequently developing high-quality datasets, the physical concept and importance of site amplification is more and more understood and quantified. Although the importance of site effects is clearly investigated, there is not yet a common framework as regards to what is the best method for site effect estimation. Over recent years, therefore, a large number of different methods and guidelines for ground response analysis with varying degree of simplification and accuracy have been proposed by various investigators. For instance, the European seismological study Site EffectS assessment using AMbient Excitations (SESAME) (Bard and SESAME-Team 2005) and the North Atlantic Treaty Organization (NATO) engineering study SfP Project 980857. The desired method is principally selected according to the importance and the main goal of the project. Generally, the site effect estimation methods are grouped in five main categories (Pitilakis 2004): (1) experimental-empirical methods that apply recordings of ground motion or ambient noise to estimate the basic site-specific characteristics usually in frequency domain; (2) Empirical methods that evaluate parameters of earthquake motions such as acceleration, velocity and response spectra based on site classification, average S-wave velocity, topography, earthquake magnitude and existing amplification relationships; (3) semi-empirical methods that compute time histories of earthquake motions by combining recorded earthquake motions of smaller earthquakes as element motions (i.e. Green's functions); (4) theoretical (numerical and analytical) methods that use analytical and more often numerical 1D, 2D or 3D wave propagation model; (5) hybrid methods that compute time histories of earthquake motions by coupling a longer period component determined by a theoretical seismic fault model with a computational seismic wave propagation model having a shorter period component determined by a semi-empirical method. In this section, some of the most commonly used methods that are implemented in this Ph.D. thesis are introduced:

2.3.1 Experimental-Empirical Methods

The majority of the site effect estimation methods use seismic ground motions to determine the main dynamic response characteristics of the site in the frequency domain. As shown schematically in Figure 2.1 and mathematically in Eq. (2.1), the seismic ground motions recorded at a station can be transferred in the frequency domain as the product of Fourier spectra of the source effect, the path effect and the site effects. In order to estimate the site effects, the source and path effects should be removed from the observed ground motions at the selected station. With the aim of determining site effects in frequency domain, experimental-empirical methods are built up in two major categories based on using a

"*reference*" motion. The "*reference site*" is defined as a close station free of any site-specific characteristics (e.g., sediments or topography), mainly located on rock, with the simultaneous recordings at the close by soil sites (Steidl et al. 1996). In other words, the upward component of the propagating seismic waves at reference site (i.e., surface-rock-site) can be considered the same as those at the base of the subsoil strata which theoretically has a flat transfer function with amplitude of one.

In the current research and specifically the first paper presented in the second part of this Ph.D. dissertation (Rahpeyma et al. 2016), two different and widely used experimental methods have been applied: (1) The Standard Spectral Ratio (SSR) as a reference site method and (2) Horizontal-to-Vertical Spectral ratio (HVSR) as a non-reference site method, both of these methods are described in the following sub-sections:

Standard Spectral Ratio Technique (SSR)

One of the most common and widely used methods to characterize site effects is Standard Spectral Ratio (SSR) method which is defined as the spectral ratio of a sedimentary site with respect to a bedrock site (i.e., reference site) from the same earthquake and component of motion (Borcherdt 1970). The SSR technique was initially introduced by Borcherdt (1970) and is still one of the most commonly used and reliable approaches to investigate site response. This technique is applicable only to cases that the data are derived from dense local arrays with at least one station on rock outcropping conditions defined as reference station. The SSR amplification curve as a function of frequency can be obtained using Eq. (2.2):

$$SSR(f) = \frac{A(f)_{soil}}{A(f)_{rock}} \tag{2.2}$$

Where $A(f)$ denotes Fourier amplitude spectra for horizontal motions recorded on the soil and rock sites, with the rock site being the reference site. The result of the SSR method is a site-specific "amplification curve" which is a function of frequency and reveals both the "*predominant frequency*" or f_0 of horizontal vibrations of the site, corresponding to the peak in the ratio, and its amplitude, A_0.

The reference site should fulfill at least two main following conditions: (1) First, it should be located close enough to the examined stations to ensure that differences between nominated stations are only due to site response characteristics and not differences in source radiation or travel path. This is the main hypothesis in this approach that the travel path through the Earth's crust is essentially the same for both sites and that the reference record is equivalent to the input motion at the base of the soil profile. Thus, the ratio of the Fourier amplitude spectra expresses only the effect of the local soil conditions at the specific site. (2) Second, the reference site also should not be influenced by any kind of site effects (e.g., geological or topographical). It should be emphasized that the choice of the reference site is critical because the rock sites can have a site response of their own, which can lead to an underestimation of seismic hazard when these sites are used as reference sites (Steidl et al. 1996). However, despite some criticisms and drawbacks for this approach, the SSR technique can still be used to estimate amplification of ground motions as a function of frequency in different geological conditions (e.g., Cranswick 1988; Boore and Joyner 1997; Raptakis et al. 1998; Guéguen et al. 2000; Atakan 2009; Cultrera et al. 2014; Rahpeyma et al. 2016). The SSR site effect estimate is relatively stable even if records are noisy (Field et al. 1992; Steidl 1993).

Horizontal to Vertical Spectral Ratio Technique (HVSR)

Although SSR method is known as a common and reliable approach to retrieve information about the shallow subsoil seismic properties which are of engineering interest, one of the main drawbacks is that in many cases we do not have access to the simultaneous recordings at a rock and soil sites in closely spaced array or local networks. Nakamura (1989) proposed a non-reference site approach which entails using the spectral ratio of the horizontal to vertical components of ground motion (see Eq. (2.3)). The Horizontal-to-Vertical Spectral Ratio (HVSR or H/V) also known as Nakamura's method was first suggested by Nogoshi and Igarashi based on initial studies of Kanai and Tanake (Kanai and Tanaka 1961; Nogoshi and Igarashi 1970, 1971). This technique has been widely used in seismic exploration as a practical tool to detect and evaluate of seismic amplification effects at a station.

From the experimental point of view, the HVSR method requires a three-component ground-motion acquisition and consists in performing the ratio between its horizontal and vertical Fourier spectrum, properly averaged on an adequate sample. The ratio can be computed by using ground-motion acceleration, velocity, and even displacement spectra. The HVSR amplification curve as a function of frequency can be obtained using Eq. (2.3):

$$HVSR(f) = \frac{A(f)_H}{A(f)_V} \tag{2.3}$$

where $A(f)$ is Fourier amplitude spectra from horizontal and vertical components recorded at a specific site. Today, due to the adequate accuracy, easy and fast implementation of the HVSR technique, it is known as the most practical tool for site effect estimations and has attracted the attention of numerous researchers who applied the method to estimate the soil amplification properties and the fundamental frequency (Konno and Ohmachi 1998; Mucciarelli and Gallipoli 2001; Di Giacomo et al. 2005; Sylvette et al. 2006; Herak 2008; D'Amico et al. 2008; Bonnefoy-Claudet et al. 2009; Gallipoli and Mucciarelli 2009; Rahpeyma et al. 2016, 2017). In addition, many researchers have investigated the reliability of using the HVSR technique, both numerically and experimentally, for the quantifications of site effects (Bard 1998; Di Giacomo et al. 2005; Sylvette et al. 2006; D'Amico et al. 2008; Pilz et al. 2009; Rahpeyma et al. 2016).

It has been well investigated that the shape of the amplification curve is firmly correlated to the subsoil characteristics and the frequency associated to the maximum amplitude (i.e. fundamental frequency, f_0) of the HVSR curve can be a representative factor of the velocity contrast between soil layers (Nakamura 2000, 2008; Bard and SESAME-Team 2005). The underlying premise of the HVSR technique is that the vertical component of the ground motion in cases where the soil stratigraphy is flat and horizontal is assumed to be free of any kind of influence related to the site conditions at the recording site. Nevertheless, there are different physical interpretations for the fundamental concepts of the Nakamura's technique (Arai and Tokimatsu 2004; Parolai et al. 2005; Arai and Tokimatsu 2005; Picozzi and Albarello 2007; Herak 2008; D'Amico et al. 2008; Albarello and Lunedei 2010; Sánchez-Sesma et al. 2011).

The HVSR Origins

The critical debate over the underlying theory of Nakamura's method focuses on the hypothesis that the obtained spectral ratio is chiefly determined by body waves that are vertically incident with the surface (Herak 2008; Nakamura 2008), or surface-waves,

Rayleigh and Love waves, with relevant higher modes (Arai and Tokimatsu 2004, 2005; Lunedei and Albarello 2010).

The fact that the HVSR spectrum can be described in term of body-waves travelling along particular patterns only is not well investigated. Although the composition of ambient vibrations in term of the different seismic phases it is not well defined, the majority of researchers believe that the vibrations are composed of all seismic phases travelling in the subsoil (Lunedei and Malischewsky 2015). In fact, different experimental and numerical results indicate that the content in different seismic phases can considerably change in dependence on the subsoil stratigraphy and on source characteristics as well as in different frequency ranges.

In a pioneer research study by Nogoshi and Igarashi (1970) the HVSR curves from ambient vibrations were compared with the ellipticity pattern of Rayleigh fundamental-mode, reckoning from the possibility of this comparison that this seismic phase plays the main role in the ambient vibrations. Afterwards, other researchers (Lachet and Bard 1994; Tokimatsu 1997; Fäh et al. 2001; Sylvette et al. 2006; Bonnefoy-Claudet et al. 2008) have emphasized the influence of the surface-waves on HVSR spectral ratio. Likewise, Arai et al. (1996) and Tokimatsu (1997) also showed that the HVSR amplification curves obtained from ambient vibrations are in association with the ellipticity of the first mode of Rayleigh waves. In fact, theoretical models (Tokimatsu 1997) and numerical simulations (Lachet and Bard 1994; Fäh et al. 2001; Sylvette et al. 2006; Bonnefoy-Claudet et al. 2008) highlighted the important role of surface waves, especially at the frequencies larger than the resonance frequency (f_0) of the subsoil. In this assumption, surface waves play a major role in the definition of the obtained HVSR curves (Konno and Ohmachi 1998; Fäh et al. 2001; Scherbaum et al. 2003; Arai and Tokimatsu 2004, 2005).

In contrast, Nakamura (2000) presented some arguments supporting the idea that the HVSR is in association with body-waves and directly represents the response function for S waves ($F_s(f_0)$) at the top of a sedimentary layer overlying a hard and rigid bedrock and the effects of Rayleigh waves is eliminated. In the case that both soil and bedrock are characterized by a weakly dissipative behavior, complex response function ($F_c(f)$) (where f is the frequency) relative to the body waves phase c (S or P) vertically propagating from depth can be computed numerically (Tsai 1970). According to Nakamura (2000) the HVSR at the fundamental frequency can be presented as Eq. (2.4):

$$HVSR(f_0) = |F_s(f_0)| \tag{2.4}$$

Recently, Herak (2008) generalized the Nakamura's assumption by considering that the ambient vibrations are constituted by body waves moving vertically and surface waves are considered to play a negligible role in the observed ambient vibrations. Within his assumption, the amplitudes of P and S phases respectively control vertical and horizontal ground motion components. If one also assumes that impinging P and S phases have the same amplitude in the bedrock, the HVSR at the surface is determined by the respective amplifications of these phases induced by seismic properties of sedimentary layers overlying the bedrock. In this position, the HVSR curve can be modeled as the ratio between transfer functions relative to S waves (horizontal components) and P waves (vertical component) as Eq. (2.5)

$$HVSR(f) = \frac{F_S(f)}{F_P(f)} \tag{2.5}$$

where $F_\beta(f)$ is the transfer function for S- and P-waves (β refers to body wave phases, i.e. S or P). The theoretical transfer function of a set of horizontally stratified, linearly elastic layers overlying a uniform half-space and excited by vertically incident (Figure 2.2), transient plane waves can be modeled analytically based on the fast-recursive algorithm proposed by (Tsai 1970), modified considering frequency dependent attenuation and body wave dispersion (Herak 2008) as Eq. (2.6)

$$F_\beta(f) = \left\{\cos\left[2\pi f \frac{H}{V_{c,a}(2\pi f)}\right] + i\frac{\rho_a V_{c,a}(2\pi f)}{\rho_a V_{c,b}(2\pi f)}\sin\left[2\pi f \frac{H}{V_{c,a}(2\pi f)}\right]\right\}^{-1} \tag{2.6}$$

where f is the frequency of the wave, H is the soft layer thickness, V_c is the complex velocity of phase c (P- or S-wave) and indices a and b respectively refer to the soft layer and bedrock for a single-layer model of soil overlaying the bedrock (Albarello and Lunedei 2010).

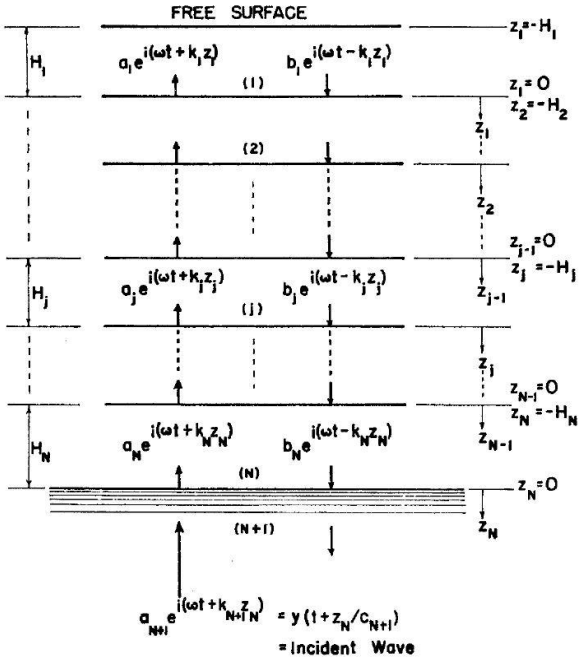

Figure 2.2 An N-Layer system under steady-state excitation (Tsai 1970).

The complex velocity, V_c, accounts for anelastic properties and can be defined as Eq. (2.7):

$$V_c(2\pi f) = \frac{V_c^e}{1 - \frac{1}{\pi Q_c}\log\left(\frac{f}{f_{ref}}\right)}\left(1 + \frac{1}{2Q_c}i\right) \tag{2.7}$$

where V_c^e is the elastic velocity of the body-waves and f_{ref} is a reference frequency (considered to be 1.0 Hz in this study), and Q_c is the quality factor used to account for material damping. It should be highlighted that the transfer function in Eq. (2.7) characterizes linear estimation of the amplification as it does not consider non-linear behavior of soil for large seismic vibrations (Herak 2008).

Comparison of HVSR and SSR Techniques

Among all site effect estimation tools SSR and HVSR techniques are well-known as the most practical and commonly used methods to determine the main site amplification characteristics. There are several literature references on comparative results on their reliability and relevance in different geological conditions using different datasets. The results of comparative investigations between SSR and HVSR techniques show good agreement between them (Stephenson 2007; Lozano et al. 2009; Akyol et al. 2013; Panzera et al. 2015; Rahpeyma et al. 2016; Stanko et al. 2017). Moreover, applying SSR and HVSR methods using weak and strong motion records (i.e., ambient noise measurements and earthquake strong-motion data) led to good agreement between techniques which validates the use of ambient noise in the absence of seismic strong-ground motion recordings (Rahpeyma et al. 2016).

In conclusion, both SSR and HVSR techniques are reliable in estimating the fundamental frequency of the soil profile. However, the amplification amplitude is comparable only when the soil layering is horizontal and there are not lateral geometrical variations. In those cases, due the presence of inward propagating surface waves, it is expected that part of them will affect the vertical component and hence the amplitude of the HVSR. For this reason, in cases where the stratigraphy is not flat and horizontal, which is pertinent in many real site conditions, the use of HVSR technique should be applied with caution, at least for the derivation of the amplification factor at the fundamental frequency.

2.4 Engineering Applications of Local Site Effects

Comprehensive understanding, as well as reliable evaluation, of local site effects have an enormous influence on the precise estimate of both level and distribution pattern of ground motions in a seismic region. Site effect can be considered as the process for estimating the response of soil layers under earthquake excitations and thus the variation of earthquake ground motion characteristics on the ground surface. Therefore, site factors and main soil characteristics are required in many engineering applications such as (1) ground motion models (GMMs) as practical tools to estimate the level of ground shaking, (2) building design codes and structural design criteria, (3) liquefaction studies, (4) microzonation studies, and (5) soil classification. In this section some of the most important engineering applications that are directly influenced by localized site conditions are discussed.

2.4.1 Ground Motion Models

Reliable estimate of the expected ground motion at a site of interest is essentially one of the dominant parameters for seismic hazard analysis (both deterministic and probabilistic) and earthquake risk mitigation. In the engineering practice, GMMs are essential tools to estimate ground motion parameters (e.g., PGA, peak ground velocity, PGV, peak ground displacement, PGD, and 5% damped pseudo-spectral acceleration, PSA). Modern GMMs utilize mathematical-based expressions that relate strong-motion parameter of ground shaking to different seismic parameters and quantitatively characterize the earthquake source, path, and site (Lee et al. 2002).

An empirical GMM generally has the form of

$$\ln Y_{es} = f(X_{es}, \boldsymbol{\theta}) + \Delta \tag{2.8}$$

where Y_{es} is the observed ground motion parameters for event e and station s, $f(X_{es}, \theta)$ represents the GMM, X_{es} is the vector of explanatory parameters (e.g., magnitude, distance, style of faulting, site conditions). $\boldsymbol{\theta}$ is vector of model coefficients, and Δ is a random variable describing the total variability of the ground motion. Δ can be decomposed into inter-event variability, δB_e, and intra-event variability, δW_{es}, which are zero-mean, independent, normally distributed random variables with standard deviations τ and ϕ, respectively (Al Atik et al. 2010):

$$\ln Y_{es} = f(X_{es}, \boldsymbol{\theta}) + \delta B_e + \delta W_{es} \tag{2.9}$$

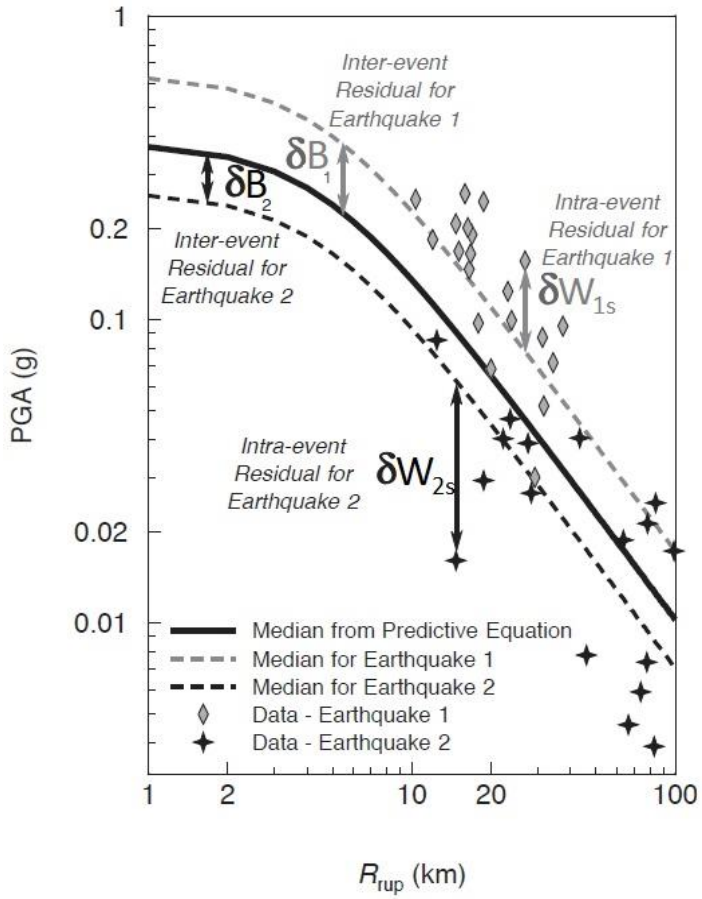

Figure 2.3 Inter-event and intra-event components of ground-motion variability (after Strasser et al. 2009).

As illustrated in Figure 2.3, the inter-event residuals (also known as between-event residual, or event term), δB_e, denotes the average shift of the observed ground motion from an individual earthquake, e, from the population median predicted by the GMM. The inter-events residual represents average source effects (averaged over all directions) and reflects the influence of factors such as stress drop and variation of slip in space and time that are not captured by the inclusion of magnitude, style of faulting, and source depth. The intra-event residual (also called within-event residual), δW_{es}, is the misfit between an individual observation at station s from the earthquake specific median prediction, which is defined as the median prediction of the model plus the inter-event term for earthquake e. The intra-event residual represents azimuthal variations in source, path, and site effects reflecting the

influence of those factors such as crustal heterogeneity, deeper geological structure, and near-surface layering that are not captured by a distance metric and a site-classification based on the average shear-wave velocity.

The GMM, $f(X_{es}, \boldsymbol{\theta})$, is a mathematical equation that relates a given strong-motion parameter to one or more parameters of the earthquake source, wave propagation path and local site conditions, collectively referred to as seismological parameters. Eq. (2.10) represents the most fundamental and commonly-used mathematical form of a GMM in logarithm form:

$$f(X_{es}, \boldsymbol{\theta}) \sim \ln \bar{Y}_{es} = c_1 + c_2 M_e - c_3 \ln R_{es} + c_4 F_e + c_5 S_s \tag{2.10}$$

where $\ln \bar{Y}_{es}$ is the natural logarithm of the strong-motion parameter of interest (e.g., PGA, PGV, and PSA), M_e is earthquake magnitude, R_{es} is a measure of source-to-site distance, F_e is a parameter to characterize type of faulting, S_s is a parameter to characterize the local site effects, and c_is are constant coefficients which can be obtained by regression analysis. During recent decades, many researchers have focused on physically quantifying the properties of the earthquake source, propagation path, and local site responses to develop the new generation of GMM which are used to predict the expected peak ground parameters along with the model variabilities. However, during recent decades, more complicated models such as NGA-West1 (Power et al. 2008) and NGA-West2 (Stewart et al. 2015) have been proposed that more complex GMMs, with the aim of increasing the accuracy and reliability of the prediction. These mathematical expressions can be physically explained in earthquake seismology (Lay and Wallace 1995).

In GMMs, localized site effects or site amplifications mainly reflect the type of deposits that lie under station and are commonly designated in GMMs in terms of surface and near-surface geology, shear-wave velocity (V_S), and sediment depth (Lee et al. 2002). The shear-wave velocity and sediment depth are often used because they can directly influence the dynamic response of the underlying subsoil structure subjected to vertically propagating body waves or horizontally propagating surface waves.

2.4.2 Building Design Codes

The specific-site site amplification is also of paramount importance in structural design criteria. The primary objective of seismic resistant design is to ensure life safety of a structure's occupants during and after a seismic event (i.e., collapse prevention). Building and seismic design codes, such as the 1997 Uniform Building Code (UBC), the 1997 National Earthquake Hazard Reduction Program (NEHRP) Recommended Provisions for Seismic Regulations for New Buildings and Other Structures (Building Seismic Safety Council, BSSC 2003; Boore 2004a), and the Eurocode 8 (EC8) (European Committee for Standardization 2003) take site condition into account. In all the above-mentioned building codes V_{S30} parameter, i.e., the average shear-wave velocity in the upper 30 m of a site profile is considered as the primary proxy for grouping sites into different classes for the purposes of incorporating local site conditions in the estimation of strong ground motion (Borcherdt 1992, 1994; Castellaro et al. 2008). As an example, NEHRP presents the minimum recommended requirements necessary for the design and construction of new buildings and other structures to resist earthquake ground motions throughout the United States. As can be seen in Table 2.1, NEHRP proposes six site classes which are a function of V_{S30}.

Table 2.1 Site Classes Defined in the NEHRP

Site Class	Soil Profile name	V_{S30} (m/s)	N_{SPT}	Undrained Shear Strength
A	Hard Rock	> 1500		
B	Rock	$760 - 1500$		
C	Very Dense Soil and Soft Rock	$360 - 760$	$> 50\ bpf^*$	$> 100\ kPa$
D	Stiff Soil	$180 - 360$	$15 - 50\ bpf$	$50 - 100\ kPa$
E	Soft Soil	< 180	$< 15\ bpf$	$< 50\ kPa$
F	Soil Requiring Site-Specific Evaluation			

* blows per foot (bpf)

2.4.3 Liquefaction Evaluation

The importance of site effect is also highlighted in liquefaction studies. Soil liquefaction is a natural phenomenon in which a saturated sandy soil behaves like a liquid in response to applied force due to losing contact between soil particles. Ground vibrations caused by earthquake or other sudden change in stress condition lead to soil liquefaction mainly in saturated, low density, un-compacted, and sandy soils. The extreme influences of liquefaction were highlighted after the 1964 Niigata earthquake, 1964 Alaska earthquake, and 2010-2011 Canterbury earthquake sequence (CES). During the 1989 Loma Prieta earthquake, soil liquefaction resulted in severe damage across San Francisco's Marina District. Hence, evaluation of the liquefaction resistance of soils has become an important and vital requirement in many geotechnical investigations and building codes in earthquake-prone regions.

Although there are various correlations have been developed relating soil resistance to liquefaction triggering and in situ test metrics (e.g., normalized standard penetration test, SPT, blow count; normalized cone penetration test, CPT, tip resistance; equivalent normalized SPT blow counted determined using a Becker hammer test, BHT; and normalized small strain shear-wave velocity, V_S), only shear-wave velocity test provides a metric that is a fundamental soil property and not an index metric. Also, it is less sensitive to problems of soil compression and reduced penetration resistance when soil fines are present, compared with SPT and CPT penetration methods (Kayen et al. 2013). In liquefaction assessments, V_S is considered an important indicator of soil capacity to resist permanent deformations and the rise of elevated pore fluid pressures.

2.4.4 Microzonation Maps

Key to microzonation for earthquake risk mitigation is the multidisciplinary contributions from the fields of geology, seismology, geophysics, geotechnical and structural engineering. Different zones can be delineated with respect to selected parameters to provide city planners with some guidelines for specifying population and building density, and more specifically, building characteristics. All of these analyses have to be considered within a probabilistic framework in order to account for all possibilities that may arise due different earthquake source mechanisms, which will have relevant exceedance probabilities (risk) that are suitable for the purpose. Geological formations, local site classification, equivalent shear wave velocity, spectral acceleration, spectral amplification and their variation are some of the parameters studied during seismic microzonation. A consistent approach has to be

implemented to assess each parameter with respect to all other parameters. Even though seismic microzonation contains important information for city and urban planning, considering different structures with different functions, site-specific studies need to be performed at each site to evaluate the effects of local soil conditions (Ansal et al. 2004).

2.4.5 Correlation between Shear-wave Velocity and HVSR

Over the last decade, the agreement between site classification and HVSR results have been explored. Zhao et al. (2006) investigated the correlation between HVSR curves and V_S profiles for the classification of seismic stations; they introduced an empirical site-classification method based on the mean HVSR amplitudes across all periods for strong-motion stations in Japan. Later, Fukushima et al. (2007) examined the impact of the site classification on empirical ground-motion models using the similar approach based on HVSR. Finally, Sokolov et al. (2007) studied a more complex behavior of sites grouped under the NEHRP B class by means of HVSRs measured at sites of the Taiwan strong-motion network observing several cases of amplification.

2.5 Estimates of Shear-wave Velocity (V_S)

Shear-wave velocity (V_S) is known as a practical indicator of the dynamic properties of the soil which can be related to the soil properties using Eq. (2.11)

$$G_{max} = \rho . V_s \tag{2.11}$$

where G_{max} is the small-strain shear modulus and can be measured in the laboratory, ρ is the total unit weight of the soil divided by acceleration due to gravity (9.81 m/sec^2) (Kramer 1996). The most common approach to quantitatively classify site conditions is by V_{S30} (m/s), which is calculated as Eq. (2.12)

$$V_{S30} = \frac{\sum_{i=1}^{N} h_i}{\sum_{i=1}^{N} \frac{h_i}{V_{s_i}}} \tag{2.12}$$

where h_i and V_{s_i} denote the thickness and V_S for the i-th formation or layer, and N is the total number of formations or layers in the upper 30 m of the profile. The summation in the numerator must equal 30 meters. In a pioneering study, Borcherdt (1994) characterized site conditions for a large number of strong-motion station sites in terms of V_{S30}. Although there are strong debates of practicality and reliability of the V_{S30} as a proxy to predict local site amplification (Castellaro et al. 2008; Kokusho and Sato 2008; Lee and Trifunac 2010; Héloïse et al. 2012), it has become a standard and widely used proxy to characterize site response in many of ground motion models and site classification schemes. Thus, the accuracy of shear-wave velocity estimation has become an important topic of research. Estimating the main properties (e.g., thickness and shear-wave velocity of layers) of the subsoil properties is one of the main objectives of this Ph.D. research and the topic of the fourth paper in this dissertation (Rahpeyma et al. 2018b).

Approximating the shear-wave velocity structure can be accomplished using a variety of either invasive or non-invasive processing tools. The classical in-situ methods usually

require more time and relatively expensive tools such as borehole drilling to investigate the key soil properties and consequently approximate the profile (Kramer 1996). Although the in-situ material testing can provide the most detailed and accurate information with reasonable resolution between closely spaced boreholes, it would be rather expensive and time-consuming to amass the measurements to cover the whole area under study; therefore, these methods are primarily recommended for projects of relative importance. Contrary to the in-situ measurements, non-invasive techniques have long been recognized as functional and cost-efficient alternatives to obtain shear-wave velocity profiles. Sections 2.5.1 and 2.5.2 describe invasive and non-invasive methods, respectively.

In either invasive or non-invasive estimation methods there is uncertainty associated with the resulting V_S profiles. There are two different types of uncertainties related to site response estimates: (1) aleatory variability, and (2) epistemic uncertainty. In terms of V_S, the aleatory variability results from the variability (randomness) associated with the subsurface layering and stiffness across the subsoil strata. Thus, aleatory variability is linked to the horizontal and vertical spatial variability of V_S. If limited V_S data exists spatially, it is hard to realistically quantify aleatory variability and assumptions must be made. On the other hand, the epistemic uncertainty results from data uncertainty and/or a lack of scientific knowledge. Thus, even for a single location, epistemic uncertainty in V_S exists due to factors such as vertical sampling interval and method of data analysis/interpretation. Currently, epistemic uncertainty in V_S is rarely quantified by those performing either invasive or non-invasive testing. Rather, a single, deterministic V_S profile is typically provided for a single testing location without consideration of data uncertainty. This results in the need to make assumptions about the epistemic uncertainty in V_S.

2.5.1 Invasive techniques

The invasive (in-hole or in-situ) techniques, used to obtain subsoil properties (e.g., P-wave velocity, S-wave velocity, density, and thickness of layers) in different depths, are based on performing local drilling and in-situ sampling in several depths inside the medium. The most commonly used in-situ methods to measure shear-wave velocity are cross-hole logging, down-hole logging, suspension logging, and seismic CPT (SCPT). The SCPT is a modified down-hole measurement in conjunction with the conventional CPT. The SCPT has become more common and preferable in recent years because it is a relatively rapid and cost-effective method of measuring shear-wave velocity in soils. However, it is worth mentioning that the laboratory tests following in-situ sampling (e.g., the small-strain shear modulus obtained via penetration-based methods) can be strongly affected by disturbance induced by sampling and reconsolidation in laboratory.

2.5.2 Non-invasive techniques

Contrary to the invasive methods that require the placement of receivers into the medium, for non-invasive techniques the instruments can be located on the ground surface without drilling boreholes. Hence during recent decades, implementing different theoretical and numerical algorithms based on wave propagation have been dramatically developed to infer subsoil properties (Bard 1998; Garofalo et al. 2016a, b). A lot of non-invasive techniques use microtremors or ambient vibrations to infer subsoil properties. Microtremors, i.e., high frequency vibrations (> 1 Hz), are mainly associated to human activities. The availability of the ambient noise measurements leads to a large number numerical and experimental

studies to obtain shear-wave velocity models. Principally, the microtremor array method (Aki 1957; Asten and Henstridge 1984) implements simultaneous recordings of ambient noise to determine phase velocity dispersion of Rayleigh surface waves travelling across an array of seismic sensors.

Dispersion Curve based Methods

The use of surface-waves to characterize subsoil structure and obtain near-surface properties was firstly introduced by inverting surface-wave phase velocity dispersion data (Ewing et al. 1957; Dorman et al. 1960; Brune and Dorman 1963; Knopoff 1972). There are many different algorithms that have been utilized to perform surface-wave inversion namely Multilayer dispersion computation, Least square curve fitting, Knopoff's method, Direct search algorithm, High frequency Rayleigh wave inversion, Refraction microtremor method.

The first surface wave method (SWM) with the aim of site-specific characteristics estimate was developed in Germany in 1930s. In this technique the deformed shape of ground surface was measured during an induced vibration using a vertical harmonic excitation. After data collection the dispersion curve versus frequency was calculated and inverted to determine the V_S profile. Afterwards, in a pioneering experiment, Brune and Dorman (1963) used a "two-station" technique in which the phase velocity was computed from the phase differences up to an arbitrary integer for two stations located at distance (Brune and Dorman 1963; Knopoff 1972). However, a significant progress in the surface-wave methods occurred in the early 1980s, with the introduction of the Spectral Analysis of Surface Wave (SASW) method (Nazarian et al. 1983; Stokoe and Nazarian 1983). In the SASW technique the main focus is on analyzing the surface-wave (predominantly Rayleigh waves) dispersion relationship to produce near-surface V_S profiles. In this method a pair of vertical low-frequency (i.e., 1-2 Hz) geophones is coupled to the ground surface for recording the surface vibrations generated by either an impulsive source (e.g., hammer or drop weight) or a vibratory source (e.g., portable electromechanical shaker), located on the extension of the straight line defined by the two receivers and at a specified distance from the first receiver (Pelekis and Athanasopoulos 2011). After proposing the SASW method, the next major progress in SWM framework was introduced in the late 1990s and early 2000s and called as Multichannel Analysis of Surface Waves (MASW) (Park et al. 1999). In the MASW technique a group of geophones is installed for measuring either active or passive surface vibrations in a limited region. Instead of measuring phase differences of surface-wave arrivals at the receivers in SASW method, in the MASW method a dispersion image is constructed by transforming the time-space domain into a different domain (e.g., frequency-wavenumber domain or frequency-phase velocity domain) (Pelekis and Athanasopoulos 2011). In this technique the inversion of dispersion data is typically based on the fundamental mode of wave propagation, although it is possible to jointly invert two or more modes.

The SWMs are defined based on the analysis of the geometric dispersion of surface-waves (dispersion-based methods). The dispersion curve depends strongly on the S-wave velocity profile of the subsurface; hence V_S profiles may be estimated by inverting measured dispersion curves. The procedure consists of estimating the dispersive characteristics at a receiver, by means of acquisition and processing of seismic data, and finally inverting these data to estimate the subsoil properties. In addition to cost efficiency of SWMs there are many advantages of using surface waves to image the subsurface structures. For instance, surface-wave inversion more capably images low-velocity zones while refraction methods cannot see low-velocity zones because such a zone would bend the traversing wave deeper instead

of towards the surface. The disadvantages of SWMs are intensive computation of data processing and inversion of the experimental data required, as well as the influential a priori assumptions that need to be made (Scherbaum et al. 2003; Molnar et al. 2010; Garofalo et al. 2016a, b). In addition, surface-wave inversion is characterized as non-linear and ill-posed problem which can be strongly affected by solution non-uniqueness (Luke et al. 2003; Scherbaum et al. 2003; Foti et al. 2009; Teague and Cox 2016). In other words, several different S-wave velocity profiles can be found having an equally good agreement with the experimental model. Consequently, this significant ambiguity in the final V_S structure using dispersion curve inversion methods.

Non-dispersion Curve based Methods

Many researchers have investigated the reliability of using ambient noise, both numerically and experimentally, for the quantifications of site effects (Bard 1998; Di Giacomo et al. 2005; Sylvette et al. 2006; D'Amico et al. 2008; Pilz et al. 2009; Rahpeyma et al. 2016). It is well investigated that in the absence of earthquake strong-motion recordings, microseismic data can be easily obtained and provide additional constraints and spatial resolution of site effects via HVSR method. Since the shape and form of amplification curves are associated with the subsoil characteristics, the frequency associated with the maximum amplitude of the HVSR curve is representative of the velocity contrast between soil layers (Nakamura 2000; Bard and SESAME-Team 2005; Nakamura 2008). Hence, despite a wide debate over basic physical interpretations, by and large the HVSR method is considered a reliable and practical tool to obtain V_S profiles (see section 2.3.1). In 2008, Herak introduced a new methodology to retrieve S-wave velocity model without engaging the dispersion curves (Herak 2008). The proposed technique is based on using Monte Carlo search in model space (i.e., subsoil properties such as thickness, h, velocity of propagation of the body wave, V_P and V_S, density, ρ, frequency dependent Q-factor for P- and S-wave, Q_P and Q_S) to minimize defined misfit function to invert the experimental ambient noise HVSRs. In this study, computing theoretical transfer functions of layered soil models are based on the fast-recursive algorithm proposed by Tsai (1970) modified to take frequency-dependent attenuation and body-wave dispersion into account (Herak 2008).

2.5.3 Inversion Strategies to Estimate V_S Profiles

S-wave velocity estimates from inversion of microtremor dispersion data must be shown to be reliable and their uncertainties understood to be used with confidence for seismic design purposes (Molnar et al. 2010). Therefore, different approaches have been proposed and are commonly used to invert the experimental dispersion curve. Most of these methods are defined based on deterministic inversion techniques (i.e., linearized inversion techniques), with some differences in the data concerned, the model parameters, the computation of the partial derivatives, the inversion algorithms, the use of constraints, and etc. (Socco and Boiero 2008). However, a major criticism of SWMs is that the inversion of the dispersion curve is a non-linear problem and the solution is non-unique, with a wide range of V_S models that can reasonably fit the data. This can lead to a certain level of ambiguity in the final shear-wave velocity profile, which is then used in modeling the site seismic response.

Although many researchers proposed different methods to invert microseismic array dispersion data by minimizing the data misfit, only a few of them consider qualitative uncertainty analysis. On the one hand, local-search (LS) techniques are defined based on local derivative approximations (Menke 1989) evaluated by the best data fitting model, and

hence they become less accurate as the data–model relationship becomes strongly non-linear. Shapiro (1996) showed that the solutions obtained from classical surface-wave inversion schemes (e.g., least-square, LS, algorithms) are too restrictive and uncertainties are not correctly estimated. The V_S profile selected by LS methods is only one of the possible solutions. The result is very sensitive to the initial model, and the inversion process can easily be biased by wrong choices in terms of model parameterization that lead the solution into local minima (Luke et al. 2003; Wathelet et al. 2004). On the other hand, many of the recent studies have included plotting all models considered in the misfit-minimization procedure colored according to misfit (Giulio et al. 2006; Roten and Fäh 2007), plotting a subset of the models based on an arbitrary misfit threshold (Wathelet 2008; Foti et al. 2009), plotting a subset of models which achieve a misfit within an arbitrary level (e.g., 10%) of the best-fit model (Parolai et al. 2007), and plotting the lowest misfit models from multiple inversions of the same data together with their average (Kind et al. 2005). However, it should be noted that quantitative uncertainty estimation needs not only a non-linear inversion approach that draws models proportional to their probability, but also rigorous estimation of the data error statistics and an appropriate model parametrization. In 2006, the 3rd International Symposium on Effects of Surface Geology on Seismic Motion investigated the ability of the microtremor array methods to determine the subsoil structure by conducting a comprehensive noise-blind test. They found that fine-layering and basement depth and velocity were almost never retrieved (Cornou 2006). Hence, a critical issue identified for improvement of microtremor inversion was quantitative and meaningful evaluation of confidence intervals on V_S profiles.

To overcome the inversion problems many researchers have tried to propose more constrained and reliable outlines in order to overcome the ill-posed problem of inversions. In recent years, the Bayesian approach has gained favor as the advantages of its greater power are recognized in many geological applications. As an example, Molnar et al. (2010) implemented non-linear Bayesian inversion with evaluation of model uncertainties and model parametrizations to produce the most probable model of the subsurface V_S profile together with quantitative uncertainty estimates from microtremor array dispersion data. Applying Bayesian inversion of microtremor array data at two nominated stations in British Columbia with high seismic hazard in Canada confirmed the practicality of the Bayesian inversion to estimate the most probable family of V_S profiles.

Bayesian Inference

Bayesian statistical modeling presents a well-defined framework based on the Bayes theorem which attempts to statistically update observed data and make inference in the light of the observations (Gelman and Rubin 1992; Diggle et al. 1998; Berger 2013; Congdon 2014; Gelman et al. 2014). The Bayesian methodology principally differs from the classical frequentist methods in that all of the unknown parameters in the underlying probability model are treated as random variables, in contrast to unknown constants. In addition, Bayesian inversion maps the distributions of data uncertainty into parameter distributions and therefore the solution of the Bayesian inversion can be obtained and presented in terms of properties of the posterior probability density of the unknown model parameters that represents optimal parameter estimates (e.g., the mode of the posterior probability distribution).

A preferable model incorporates a priori knowledge about the model parameters through prior distribution. As can be seen in Eq. (2.13), the basic concepts of the Bayesian inference

can be described in mathematical terms. The inference for the model parameter vector $\boldsymbol{\theta}$ is based on the data $\boldsymbol{y}$ which contains information about $\boldsymbol{\theta}$ through the sampling distribution, denoted by $\pi(\boldsymbol{y}|\boldsymbol{\theta})$, also known as likelihood function. The prior density, denoted by $\pi(\boldsymbol{\theta})$, describes probabilistically assumptions about $\boldsymbol{\theta}$ and is another important component in the inference scheme. The posterior density, denoted by $\pi(\boldsymbol{\theta}|\boldsymbol{y})$, represents the knowledge about theta after seeing the data, and can be thought of as an update of the prior density. It is given by

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{\pi(\boldsymbol{\theta})\pi(\boldsymbol{y}|\boldsymbol{\theta})}{\pi(\boldsymbol{y})} \tag{2.13}$$

where $\pi(\boldsymbol{y})$ represents the marginal density function of the data, $\boldsymbol{y}$, which is independent of parameters $\boldsymbol{\theta}$:

$$\pi(\boldsymbol{y}) = \int \pi(\boldsymbol{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\theta \tag{2.14}$$

Since we are often unable to evaluate the integral of Eq. (2.14) analytically, we mainly consider a numerical approximation method instead and Eq. (2.13) can be expressed as Eq. (2.15):

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) \propto \pi(\boldsymbol{\theta})\pi(\boldsymbol{y}|\boldsymbol{\theta}) \tag{2.15}$$

In general, computing these properties requires optimizing and integrating the posterior probability density, which must be carried out numerically for nonlinear problems (Molnar and Cassidy 2006).

Markov Chain Monte Carlo (MCMC) Algorithm

There are several different numerical techniques that have been proposed to construct and sample from arbitrary posterior distributions and can be applicable to almost any Bayesian problem. Markov chain simulations, also called Markov chain Monte Carlo (MCMC) methods, are used to simulate samples, $\boldsymbol{\theta}$, from a posterior distribution with the desired (true) posterior distribution, $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ (see e.g., Brooks 1998; Gilks 2005; Gelman et al. 2014). The key motivation behind the MCMC is that they perform an intelligent search within a high dimensional space and thus Bayesian Models in high dimensions become tractable. The Markov chain sampling is performed in a sequence based on the distribution of the sampled draws and depending on the last value drawn. In this regard, Markov chain is a sequence of random variables $\boldsymbol{\theta}^1, \boldsymbol{\theta}^2, \ldots$, where for any k, the distribution of $\boldsymbol{\theta}^k$ given all previous $\boldsymbol{\theta}$'s depends only on the most recent value $\boldsymbol{\theta}^{k-1}$. This procedure is often carried out by the use of the Gibbs sampling algorithm (Geman and Geman 1984; Casella and George 1992) and the Metropolis-Hastings algorithm or Metropolis algorithm (Metropolis et al. 1953; Hastings 1970), which are both outlined in the subsequent sections.

Gibbs Sampling

The Gibbs sampler, also known as alternating conditional sampling, is a particular MCMC sampling algorithm considered to obtain samples from a joint distribution of multidimensional random variable (Casella and George 1992; Gelman et al. 2014). The Gibbs sampler is the simplest of the Markov chain simulation algorithms, and it is our first choice for conditionally conjugate models, where we can directly sample from each

conditional posterior distribution. Since the Gibbs sampler can be defined in terms of sub-vectors, it is suited for obtaining samples from the posterior distribution of a given multi-parameter Bayesian model. The mathematical formulation of the Gibbs sampler, as designed for Bayesian inference, is as follows. Assume $\boldsymbol{\theta}$ denotes a vector of unknown model parameters, $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_j)$, of a given multi-parameter Bayesian model that can be divided into J sub-vectors or components; therefore, there are J steps in iteration k (for total K iterations). The corresponding joint posterior density function is represented by $\pi(\boldsymbol{\theta}|\boldsymbol{y})$. At each iteration, the Gibbs sampler cycles through the sub-vectors of $\boldsymbol{\theta}$ and draws samples from conditional posterior distribution of the sub-vectors of $\boldsymbol{\theta}$ conditioned on the latest values of other sub-vectors of $\boldsymbol{\theta}$. This scheme generates a Markov chain consisting of samples of $\boldsymbol{\theta}$ obtained in every iteration, that can be shown to converge to the target posterior density $\pi(\boldsymbol{\theta}|\boldsymbol{y})$. The standard Gibbs sampler based on total K iterations with target posterior density of $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ can be mathematically formulated as following steps (Gelman et al. 2014):

```
Initialize the Markov chain, by choosing an arbitrary
starting value θ^0 such that π(θ^0|y) > 0
for k = 1, ..., K − 1
    for j = 1, ..., J
        sample θ_j^(k+1) from π(θ_j | y, θ_1^(k+1), ..., θ_(j−1)^(k+1), θ_(j+1)^k, ..., θ_J^k)
    end
end
```

Metropolis and Metropolis-Hasting Algorithm

The *Metropolis-Hastings algorithm* is one of the most practical and commonly used Markov chain Monte Carlo simulation methods for obtaining a sequence of random samples from a probability distribution. (Metropolis et al. 1953; Hastings 1970; Gelman et al. 2014). It also can be considered as an extension to the Gibbs sampler which is highly practical for sampling from multi-dimensional distributions. An appropriate selection of a proposal distribution for MCMC methods, for example for the Metropolis-Hastings algorithm, is well known to be a crucial factor for the convergence of the algorithm.

The *Metropolis algorithm* is an adaptation of a random walk with an acceptance/rejection rule to converge to the specified target distribution. The algorithm proceeds as follows (Gelman et al. 2014)

1- Draw a starting point of vector parameter $\boldsymbol{\theta}^0$, for which $\pi(\boldsymbol{\theta}^0|\boldsymbol{y}) > 0$

2- At step k

a) Sample a proposal value $\boldsymbol{\theta}^*$ from a proposal distribution conditioned on the previous iteration i.e. $q(\boldsymbol{\theta}^*|\boldsymbol{\theta}^{k-1})$. For the Metropolis algorithm (but not the Metropolis-Hastings algorithm), the jumping distribution must be *symmetric*, satisfying the condition that $q(\boldsymbol{\theta}^*|\boldsymbol{\theta}^k) = q(\boldsymbol{\theta}^k|\boldsymbol{\theta}^*)$,

b) Calculate the ratio of the densities,

$$r = \min\left\{1, \frac{\pi(\boldsymbol{\theta}^* \quad |\boldsymbol{y})}{\pi(\boldsymbol{\theta}^{k-1}|\boldsymbol{y})}\right\} \tag{2.16}$$

c) Sample u_k from uniform density on [0,1]. Accept or reject the proposed values of model parameters according to:

$$\boldsymbol{\theta}^k = \begin{cases} \boldsymbol{\theta}^{k-1} & if \quad r \leq u_k \\ \boldsymbol{\theta}^* & if \quad r > u_k \end{cases} \tag{2.17}$$

However, it should be noted that the MCMC estimate of model parameters using Metropolis steps can lead to inefficient and unreliable estimate due to (1) the jumps are short that cause the simulation moves very slowly through the target distribution, and (2) the jumps are almost all into low-probability regions of the target density, causing the Markov chains to remain steady and not progressive. However, it is always possible to improve the mixing simulation by properly adjusting the jumping distribution.

Computational Efficiency Assessment

Although the MCMC algorithms essentially ensure the convergence of simulation process to the target density, it is absolutely necessary to check the convergence of the simulated sequences once the simulation algorithm has been implemented and the simulations drawn. In general, fast convergence and low dependence between successive samples lead to higher quality of MCMC chains. There are several diagnostic techniques to qualify the MCMC chains and obtain the rate of acceptance. The convergence diagnostics tools which are normally used for assessing computational efficiency of MCMC chains are the following:

Trace plots: A trace plot is a plot of an MCMC simulation, in which the value of the MCMC chain is plotted as a function of iterations. By visual inspection of the trace plots, it is possible to identify if and where the MCMC chain gets stuck in the same value for many consecutive iterations. If the MCMC chain does get stuck, that indicates low computational efficiency.

Gelman-Rubin statistic: Gelman and Rubin (1992) proposed a metric for assessing convergence of iterative MCMC simulations. The Gelman–Rubin statistic is evaluated from the m simulated MCMC chains, which have different initial values and have been simulated independently of each other. The algorithm for calculating the Gelman–Rubin statistic is thoroughly outlined in Brooks and Gelman (1998). The Gelman–Rubin statistic can be interpreted as follows. A Gelman-Rubin statistic close to 1.00 suggests that the MCMC simulations are close to the target distribution. In most practical cases, values below 1.05 are acceptable. However large values of the Gelman–Rubin statistic, typically greater than 1.10, indicate that the simulations have not converged to the target density. Gelman-Rubin plots are plots where the Gelman-Rubin statistic is plotted as a function of iteration. These plots can be used as a visual tool for assessing the rate of convergence of the given MCMC chain.

Autocorrelation: The dependence between successive samples of the Markov chain is evaluated with its autocorrelation which is estimated with the sample correlation. The *j*-th lag autocorrelation, ρ_j, is defined as the correlation between every *j* successive draws. The *j*-th lag autocorrelation of a MCMC chain $\{\theta_k\}_{k=1}^K$ can be estimated with Eq. (2.18):

$$\hat{\rho}_j = \frac{\sum_{k=1}^{K-j}(\theta_k - \bar{\theta})(\theta_{k+j} - \bar{\theta})}{\sum_{k=1}^{K}(\theta_k - \bar{\theta})^2} \tag{2.18}$$

where $\bar{\theta} = K^{-1}\sum_{k-1}^{K}\theta_k$. How the *j*-th lag autocorrelation decreases as function of lag k yields insight into the computational efficiency of the MCMC sampler. That is, the *j*-th lag autocorrelation decreases rapidly if the MCMC algorithm is computationally efficient. However, high *j*-th lag autocorrelation for relatively high values of j indicates poor computational efficiency. Autocorrelation plots, which are plots showing the *j*-th lag autocorrelation as a function of lag *j*, are useful visual diagnostics tools for assessing the behavior of the autocorrelation.

Effective sample size: The quality of a simulated MCMC chain can be assessed through its effective sample size, which is an estimate of the equivalent number of independent iterations in the simulated MCMC chain.

Proposal Densities and Efficiency

The choice of proposal density is highly effective for MCMC algorithms. Applying an inappropriate density function can lead to bad mixing and slowly converging MCMC chains. Hence, designing a competent proposal density is vital when implementing the MCMC algorithm in practice. Many researchers offer practical guidelines for designing an efficient proposal density. In general, the efficiency and applicability of the proposal density function can often be controlled through the acceptance probability, r, or the expected acceptance ratio $\mathbb{E}[r]$ of the MCMC chain. In this regard, various performance of the acceptance rate is expected for optimum computational efficiency is expected due to the class of MCMC algorithms. This in turn, leads to guidelines for tuning the proposal densities for optimal efficiency.

Random walk proposal: If the MCMC chains show high acceptance ratio, the proposed values tend to be close to the current value of the chain. In other words, the walking space is too narrow, and the model parameter cannot properly explore the model space of the posterior density. Accordingly, in each iteration the Markov chain will take too small steps (i.e. jumps). This phenomenon can lead to high autocorrelation MCMC chain, and therefore reduces computational efficiency. Despite of high autocorrelation chains, when the acceptance rate is too low, the proposed draws take large steps from the current positions in the chain but are frequently rejected by the MCMC algorithm. This results in the chain getting stuck in the same state for many iterations, which reduces computational efficiency. Roberts et al. (1997) confirmed that the optimal acceptance rate for random walk proposals is 44% when the θ is a scalar and 23% when it is a multidimensional.

In practice, random walk proposal densities are tuned to achieve the desired acceptance rate for computational efficiency. An example of a random walk proposal for a target density $\pi(\boldsymbol{\theta}|\boldsymbol{y})$, which is tuned for computational efficiency, is given in Roberts et al. (1997). That is, a proposal density based on the Gaussian distribution centered on the last draw of θ with a covariance matrix $c(-\boldsymbol{H})^{-1}$ where $\boldsymbol{H}$ is the Hessian matrix of $\log\pi(\boldsymbol{\theta}|\boldsymbol{y})$ evaluated at the mode $\boldsymbol{\theta}_0$ and $c = 2.38^2/dim(\boldsymbol{\theta})$. The Hessian matrix is defined as a square matrix of second-order partial derivatives of a scalar-valued function and can be presented here as Eq. (2.19) that is:

$$\boldsymbol{H} = \nabla^2 \log \pi(\boldsymbol{\theta}|\boldsymbol{y})|_{\boldsymbol{\theta}=\boldsymbol{\theta_0}} \tag{2.19}$$

where ∇^2 is the second derivative operator for a multivariable function. Consequently, the resulting proposal density is

$$q(\boldsymbol{\theta}^*|\boldsymbol{\theta}^{k-1}) = \mathrm{N}(\boldsymbol{\theta}^*|\boldsymbol{\theta}^{k-1}, c(-\boldsymbol{H})^{-1}). \tag{2.20}$$

The scaling parameter *c* can be shown to yield optimal acceptance rates in a particular large dimension scenario (Roberts et al. 1997).

2.6 Quantifying Site Effects

Many seismic loss problems, e.g., disruption of spatially distributed infrastructure, are highly dependent upon the regional distribution of ground motion intensities. As discussed in section 2.4.1, the GMMs are the simplest and most common tools in engineering practice to quantify the systematic dependence of the frequency dependent earthquake ground motion amplitudes. Current needs of the site-specific hazard and risk analysis for critical facilities in conjunction with the accumulation of the observed records provide the impetus to develop site-specific GMMs in which earthquake ground motion amplitude is decomposed. In general, empirical GMMs are constructed by fitting a regression formula (e.g., by means of least-square method) of a particular functional form to available observations of peak-parameters evaluated from recordings of earthquake ground motion. The variability of the recorded earthquake ground motion amplitude across a region arises from different sources, such as earthquake rupture, seismic wave propagation path, and local site effects. Due to the vital effects of variability of GMMs on seismic hazard and seismic risk assessment, many researchers decompose the observed ground motion parameters as a reliable approach to estimate the associated variabilities.

2.6.1 Variabilities in Ground Motion Models

It has long been known that the reliability of the variability of ground motion amplitude plays an important role in a precise seismic hazard assessment. Commonly, the random variability of ground motions is divided into two components: (1) aleatory variability (i.e., natural randomness in in a process) and (2) epistemic uncertainty (i.e., limited knowledge or data of the system). The aleatory variability in ground motion prediction accounts for the apparent randomness in observed motions with respect to the predictive model and is interpreted as being inherent variability that cannot be reduced without changing the predictive model. This is contrasted with epistemic uncertainty, which is the component of the ground motion prediction that results from incomplete knowledge of the earthquake process and which can therefore, in theory, be reduced through the acquisition of additional and better data (Bommer and Crowley 2006). Since reduction of standard deviation is a critical issue in seismic hazard mitigation studies, especially at low probability levels (Atkinson 2006), many researchers put their efforts to decrease the uncertainty by improving physically justified GMMs. However, use of even more complicated functional forms of GMMs over the last 50 years (Abrahamson and Shedlock 1997; Abrahamson et al. 2008; Douglas and Aochi 2008; Strasser et al. 2009), has not changed the total uncertainty considerably. On this note, investigation of the main sources of variabilities associated with attenuation relationships and any logical physical and theoretical interpretation that could influence the possible variability are a challenging research topic, with important

implications for critical structures in engineering seismology (Bommer et al. 2004; Bommer and Abrahamson 2006; Atkinson 2006; Kowsari et al. 2017). During recent decades, many researchers have tried to split up the variabilities into independent terms and quantify the relative contributions of each to the overall variability (Joyner and Boore 1981; Fukushima and Tanaka 1990; Abrahamson and Youngs 1992; Bommer and Crowley 2006; Douglas and Gehl 2008; Kuehn and Scherbaum 2015). The objective is to separate the variations in amplitudes in order to find more reliable inference of source, station, and path effects.

The basic partitioning of the source (inter-event) effects from the site and path (intra-event) effects can be considered in relation to a median, reference GMM (see Figure 2.3). The inter-event and intra-event standard deviations of the ground-motion model represent the earthquake-to-earthquake variability and record-to-record variability, respectively. The inter-events and intra-event residuals are uncorrelated, so the total standard deviation of the ground-motion model, σ, can be written as Eq. (2.21):

$$\sigma = \sqrt{\tau^2 + \phi^2} \tag{2.21}$$

It has been recognized that the variance can be further partitioned to account for repeatable source, path, and site effects (Anderson and Brune 1999; Al Atik et al. 2010). Hence the total variability becomes:

$$\sigma = \sqrt{\tau^2 + \phi_{S2S}^2 + \phi_{SS}^2} \tag{2.22}$$

in which τ, ϕ_{S2S}, ϕ_{SS} are the standard deviation of event effects, station effects and event-station effects, respectively. τ encodes differences of a particular event from the mean of all events (such as a deviating stress drop), whereas ϕ_{S2S} and ϕ_{SS} are due to differences in the site and path related aspects, respectively.

The unknown effects and associated variabilities can be approximated using parameterized empirical models (e.g. Chin and Aki 1991; Boatwright et al. 1991; Moya et al. 2000; Shabestari et al. 2004; Ortiz-Alemán et al. 2017) or empirical techniques (e.g. Aki 1957; Borcherdt 1970; Nakamura 1989; Bard 1998).

2.6.2 Classical techniques

The empirical equations for predicting strong ground motion are classically fit to the strong-motion data set by the method of ordinary least squares (LS). Campbell (1981) used weighted least-squares (WLS) in an attempt to compensate for the nonuniform distribution of data with respect to distance. Fukushima and Tanaka (1990) introduced a two-stage regression method designed to decompose the determination of the magnitude dependence from the determination of the distance dependence. They implemented the proposed method on the Japanese peak horizontal acceleration data set and compared results with those from one-stage ordinary least squares. They showed that the one-stage ordinary least-squares results were seriously in error. They attributed the error to the strong correlation between magnitude and distance and the resulting trade-off between magnitude dependence and distance dependence. Nonetheless, Joyner and Boore (1993) found both one-stage and two-stage methods have comparable variabilities. Later, Boore et al. (1997) used the same two-stage regression technique to decouple the components. The correct distance dependence, given by the two-stage method and verified by analyzing individual earthquakes separately,

showed a much stronger decay of peak acceleration with distance than the one-stage ordinary least-squares method, which had been used previously.

In 1992, Masuda and Ohtake proposed a weighting matrix for the second-stage regression different from any used earlier (Masuda and Ohtake 1992). They revealed that off-diagonal terms need to be included in the weighting matrix, because the amplitude factors that are the dependent variables in the second-stage regression are mutually correlated as a consequence of the fact that they were determined in the first-stage regression along with the parameters that control the distance dependence. Brillinger and Preisler (1984, 1985) proposed what they called the random effects model, which incorporated an explicit earthquake-to-earthquake component of variance in addition to the record-to-record component. They described one-stage maximum-likelihood methods for evaluating the parameters in the prediction equation.

More recent, Abrahamson and Youngs (1992) introduced an alternative algorithm, which they considered more stable though less efficient. The concept of an earthquake-to-earthquake component of variance is implicit in the two-stage regression methods. The two-stage methods are not, however, exactly equivalent to the one-stage maximum-likelihood methods, and the relationship of one to the other is not obvious. Both the one-stage and two-stage methods are based on maximum likelihood. In the one-stage methods, the parameters are all determined simultaneously by maximizing the likelihood of the set of observations. In addition to the classical approaches to determine the ground motion model along with variability estimation, many researchers (e.g., Wang and Takada (2009); Kuehn and Scherbaum (2015, 2016)) implemented new statistical tools such as the Bayesian theory in order to decompose earthquake ground motion parameters and associated variabilities into source, path, and site terms.

2.6.3 Bayesian Hierarchical Model (BHM)

Many statistical applications involve several parameters that can be regarded as related or connected in some way by the structure of the problem, implying that a joint probability model for these parameters should reflect their dependence (Gelman et al. 2014). A key point of such applications is that the observed data can be used to estimate aspects of the population distribution even though the values of are not themselves observed. It is natural to model such a problem hierarchically, with observable outcomes modeled conditionally on certain parameters, which themselves are given a probabilistic specification in terms of further parameters, known as *hyperparameters*, with possible available *prior* distribution. Such hierarchical thinking helps in understanding multiparameter problems and also plays an important role in developing computational strategies. The hierarchy of the Bayesian hierarchical modeling adhered to in this thesis, can be idealized in the following three levels.

Data level: A data density is chosen for the observations conditioned on the latent processes and other potential parameters.

Latent level: A probability model is constructed for the latent processes conditioned on other potential parameters. This is attained by selecting prior distribution which should ideally incorporate a priori knowledge on the latent processes.

Hyper level: Prior distributions for the parameters of the latent processes are chosen.

BHM Formulation for Ground Motion Model

A GMM can be formulated using a three-level BHM consisting of what is referred to as the data-level, latent-level, and hyper-level (for details the reader is referred to Rahpeyma et al. 2018). At the latent-level which mainly consists of model parameters, the probability models are constructed using hyperparameters and other potential parameters. At this hyper-level, the probability models include the prior distribution of the model variabilities and other parameters defined in the latent-level. In the formulation convention of Bayesian statistics, the BHM model consists of the following independent terms as Eq. (2.23)

$$Y_{es} = \mu_{es}(M_e, R_{es}, D_e) + \delta B_e + \delta S_s + \delta WS_{es} + \delta E_{es}\,, \qquad (2.23)$$

$$for \quad e = 1, \dots, N\,, \quad s = 1, \dots, Q$$

where μ_{es} (i.e., GMM) provides median ground motions in terms of independent variables (magnitude, hypocentral distance, depth of the origin, and constant coefficients) for event e and station s, δB_e is the event effect (also called inter-event residual or event term) which denotes the over-all effect of event e in addition to the predicted median ground motion μ_{es}, δS_s is the station effect for station s, δWS_{es} is a spatially correlated event-station effect, and δR_{es} is an independent error term representing unmodeled effects or other factors that are not accounted for. The terms δWS_{es} and δR_{es} are assumed to follow Gaussian distributions, thus, conditioned on μ_{es}, δB_e and δS_s then Y_{es} also follows a Gaussian distribution.

The event effects are combined in the vector $\boldsymbol{\delta B} = (\delta B_1, \dots, \delta B_N)$, which we assume a priori to be normally distributed with a mean of zero, standard deviation of τ and independent of each other. The station effects in the vector $\boldsymbol{\delta S} = (\delta S_1, \dots, \delta S_Q)$ are assumed to stem from a zero mean Gaussian field governed by a Matérn covariance function with marginal variance ϕ_{S2S}^2 that describes the station-to-station variability. The spatially correlated event-station effects, δWS_{es}, are modeled as a zero mean Gaussian field for each event e that is also governed by a Matérn covariance function with marginal variance ϕ_{SS}^2 that describes within event variability. The Gaussian fields of any two events are independent. So, the vector of all δWS_{es} terms stacked by subvectors containing terms from the same event has a mean of zero and a block diagonal covariance matrix. Finally, unmodeled effects term, δR_{es}, is zero mean Gaussian with standard deviation of ϕ_R and the δR_{es} terms are independent of each other. The current values of the smoothness of the Mátern covariance functions were selected after trying different values for the smoothness parameters for station effects and event-station effects covariance functions.

The total variance of Y_{es} can be calculated as the sum of the inter-event variance τ^2, the inter-station variance ϕ_{S2S}^2 (i.e., station-to-station variability), the event-station variance ϕ_{SS}^2 (i.e., variability between station within an event), and the variance of the unmodeled effects and other unaccounted factors, ϕ_R:

$$\sigma^2 = \tau^2 + \phi_{S2S}^2 + \phi_{SS}^2 + \phi_R^2 \qquad (2.24)$$

The inter-event variance (τ^2) quantifies the variations between events after taking the effects of earthquake magnitude, M_e, source-to-site distance, R_{es}, and depth, D_e, of each event into account. The inter-station variance (ϕ_{S2S}^2) quantifies the variability in the station effects that stems from the varied local geological conditions. In other words, the inter-station variance represents the systematic deviation of the observed amplification at the specific station from

the median amplification predicted by the model. The event-station variance (ϕ_{SS}^2), on the other hand, is a measure of the spatial variability in PGA between stations within events after taking into account the overall effect of the event and the average effect of each station. And finally, the variance parameter ϕ_R^2 quantifies jointly the variability in the unmodeled effects and deviations that are not accounted for with other terms in the model. A detailed explanation of the BHM model is presented in the Appendix of Paper III.

3 Aim

The overall objective of this dissertation is to enhance our understanding about site-effects and shed new light on site-specific characteristics in engineering applications. Hence, the main goals are summarized in three main steps: Firstly, establish a comprehensive site-effects estimate framework by applying different empirical-experimental techniques and using both strong ground motion data set and microseismic measurements to quantify the main characteristics of the localized site effects. Secondly, propose a robust non-invasive approach which provides reliable inference on the main properties of the near-surface soil structure and the associated variabilities. Finally, quantitatively decouple recorded ground motions into the source, path, and site effects and estimate their associated variabilities. These goals are addressed in the four papers presented in this thesis. The following summarizes the outline presented in each paper.

Paper 1: The main goal of the research presented in paper I is to outline a comprehensive and detailed site-effect estimation in order to highlight the influence of site-specific characteristics on ground motion response even over relatively short distances using different site effect investigation methods as well as different data sets (i.e., weak and strong ground motions) recorded on the first small-aperture strong-motion array (ICEARRAY I) in south Iceland.

Paper 2: A novel Bayesian hierarchical spatial model (BHM) for characterizing variations in earthquake ground motion parameters is proposed in paper II. The proposed BHM model quantitatively accounts for an event effect, a station effect, an event-station effect, and an error term that jointly takes into account measurement error and model error, respectively. The BHM quantifies the local variabilities in the small region of the array stations and shows to what extent the source and site contribute to that variability.

Paper 3: In this paper the proposed BHM in paper II is used to quantitively estimate the site effects characteristics across two Icelandic strong motion arrays with different geological structures. Establishing quantitative estimates of strong-motion spatial variability will develop our understanding of the key factors which affect the variation of seismic ground motions even across a relatively small area.

Paper 4: The main goal of this paper is to propose a non-invasive technique in order to obtain the subsoil structure based on inversion of the observed microseismic Horizontal-to-Vertical spectrum. The inversion scheme is set up in the context of the Bayesian framework using Markov chain Monte Carlo technique with Metropolis steps in order to explore the space of model parameters and find the best fitting family of subsoil properties along with their associated uncertainties. A blind test is conducted over the number of layers to consistently investigate the best resolution of model parametrization.

4 Applications and Results

The main discussion in this chapter cycles through the proposed methodologies and statistical modeling developed in the thesis and is in accordance with the objectives outlined in Chapter 1, with appropriate references to the four papers prepared from this PhD plan.

4.1 Site Effect Estimation

4.1.1 Historical site effect estimation in Iceland

Iceland, the most seismicity active region in northern Europe, is lying across the Mid-Atlantic Ridge where the North American and Eurasian crustal plates are diverging at an average rate of approximately 2 cm/y (DeMets 1990; DeMets et al. 1994; Einarsson 2008). Passing across Iceland from southwest to north, the onshore part of the plate boundary shifts eastward, resulting in two main transform fault zones (cf. Figure 4.1): (1) the completely onshore South Iceland Seismic Zone (SISZ) with an approximately $80\ km$ long by $20\ km$ wide in the south Iceland, (2) the mostly offshore Tjörnes Fracture Zone (TFZ) with around 120 km long and as much as 70 km wide in the north coast of Iceland (Stefánsson et al. 1993; Guðmundsson et al. 1993). The seismic potential of the SISZ has been well investigated and characterized by a network of N-S right lateral strike-slip faults with potential to produce destructive earthquakes either as strong single or sequences of magnitude $7.0 - 8.0$ events over a period lasting from weeks to years (Einarsson et al. 1981; Einarsson 1991; Bjarnason et al. 1993; Stefánsson et al. 1993; Bellou et al. 2005). The TFZ is known as a tectonically complex triangular area which is primarily composed of three NW-SE lineaments: Dalvík lineament, Grimsey lineament, and Húsavík Flatey Fault (HFF). Seismic activity in TFZ is mainly associated with swarms with similar waveforms and frequently offshore (Hensch et al. 2008). Contrary to the SISZ, since TFZ is mostly offshore, historical information and geological observations are not well-documented and essentially limited so the earthquake catalogue of the TFZ is less complete and less comprehensive (Saemundsson 1974; Einarsson 1991; Guðmundsson et al. 1993; Stefansson et al. 2008).

The soil structure in Iceland in many places consists of repeating sedimentary and lava-rock layers that was mostly formed during and after the last Ice Age as a result of recurring glaciation and deglaciation, sea level fluctuations along with volcanic eruptions, (Einarsson and Douglas 1994). Such repeating structure of inter changing relatively soft (sediment) and stiff (rock or lava-rock) layers leads to several strong velocity reversals within depth. Furthermore, the surface geology is further complicated by fractures, fissures and faults of tectonic origin (Clifton and Einarsson 2005; Angelier et al. 2008). This is especially true for the south Iceland lowlands (SIL) known as geologically young and the most populous agricultural region in Iceland. The bedrock geology of the SISZ was formed during the Upper Pliocene and Pleistocene which mostly covered by staked layers of postglacial lava flows and Quaternary sediments of fluvial, glacial, and glaciofluvial origin (Sæmundsson 1979). On the other hand, the majority of the geological studies conducted in the TFZ have focused on the Tjörnes peninsula, particularly, the town of Húsavík (Sæmundsson and Karson 2006; Gudmundsson 2007). The oldest present rock in the Tjörnes peninsula is a

8.5 – 10 Ma old basalt and can be found north of the HFF as the basalt are cut at the fault and are not known to occur south of it (Sæmundsson and Karson 2006). Most of the lowest strata is of young Quaternary basalt, originating from a series of eruptions from the nearby Grjothals shield volcano; however, at relatively low elevations, the lava rock is superseded by hyalochlastite, indicating that the Grjothals lava had flowed into the sea (Sæmundsson 1974). The Tjörnes peninsula is also quite unique as it is one of the few places in Iceland with 500 m thick sedimentary layers, known as Tjörnes beds (Gudmundsson 2007).

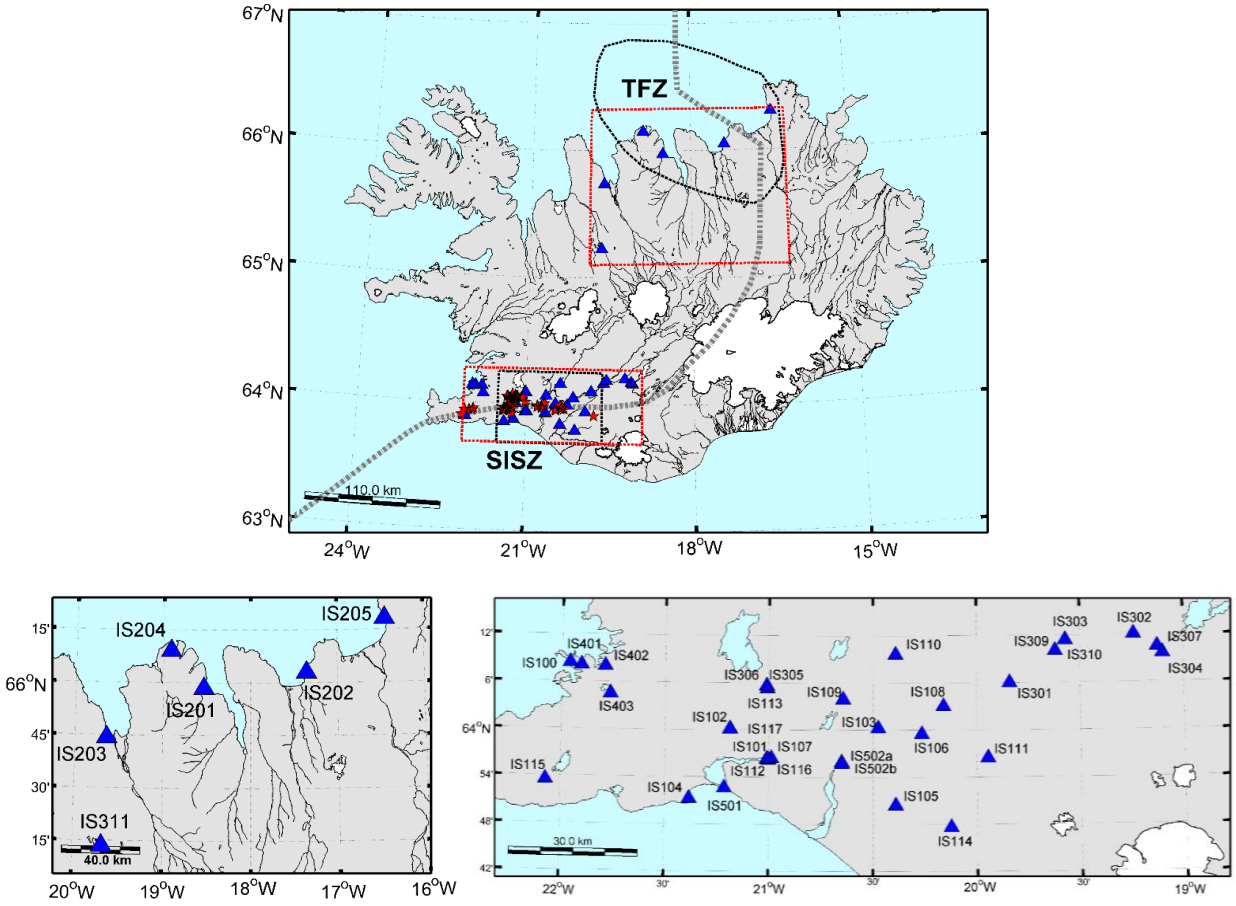

Figure 4.1 The ISMN strong-motion stations (represented by triangle symbols) across SISZ and TFZ in the south and north of Iceland. The picture at the bottom shows the cut area in the top picture with red dashed rectangular along with ISMN strong-motion station.

As a pioneering site effect study in Iceland, Atakan et al. (1997) investigated site amplification characteristics in an area of $\sim 400\ \mathrm{km}^2$ in westernmost part of SISZ by implementing two experimental methods of site effect estimation i.e., Single Station Spectral Ratio (SSSR, same as HVSR technique) and Standard Spectral Ratio (SSR) and using 15 earthquake strong-motion recordings including a magnitude 3.1M_L event as well as ambient noise measurements. They showed that there is good agreement between different data sets and different techniques. They argued that most of the spectral amplifications observed, should be related to the unconsolidated sediments of fluvial origin. However, the possible effects of underlying partly-consolidated, to compact, sedimentary deposits (i.e., hyaloclastites) are not fully known. In addition, they showed that the unconsolidated sediments pose a major concern in areas which are sensitive to damage by amplification of the earthquake signals. However, they claimed that the lack of digital data from earlier

strong-motions recorded on the sediments, makes it difficult to address this problem. They emphasized that detailed studies on the correlation between the results from this study with damage patterns from historical large earthquakes can give important clues in this respect.

In general, seismic wave amplification in Iceland due to localized site conditions is largely considered to be insignificant in earthquake engineering practice due to the easily exposed older bedrock or more recent volcanic lava-rock. However, recent site effect investigations have shed new light on the strong variation in earthquake ground motion amplitudes due to variations in site effects, originating in a complex subsoil structure, even over relatively short distances (Bessason and Kaynia 2002).

4.1.2 Icelandic Strong-motion Network and Arrays

Since 1984, earthquake strong-motion in Iceland has been monitored and recorded by the Icelandic strong-motion network (ISMN) in order to collect data for engineering objectives and seismic analysis (Sigbjörnsson 1990). The ISMN stations are mainly located in the SISZ and around the margin of the TFZ where the population, industrial centers, and lifeline networks (e.g., hydroelectric and geothermal power stations, hospitals, bridges, dams etc.) are largely concentrated. Around 40 permanent ISMN stations (Figure 4.1) with approximately $5-15$ km inter-station distance provide data for earthquake engineering objectives (Sigbjörnsson et al. 2004). However, the general lack of dense recordings of earthquake ground motion, specifically in urban areas, hampers the reliable assessment of variable strong ground motions over small areas which is required for accurate seismic hazard mitigation and decisive applications in building codes (Zerva and Zhang 1997; Zerva and Zervas 2002).

In 2007, the first small-aperture Icelandic strong-motion array (ICEARRAY I, top left Figure 4.2) was deployed in the town of Hveragerði in South Iceland with the aim of (1) monitoring and recording strong events in the region, (2) quantifying spatial variability of strong-motion over short distances, (3) shedding lights on earthquake source processes (Halldorsson et al. 2009). The array covers around 1.23 km^2 and consists of 13 accelerometric stations with inter-station distances ranging from $50-1900$ m (Halldorsson et al. 2009). Afterwards in 2012, the second strong-motion array (ICEARRAY II) was deployed in the TFZ in North Iceland, specifically in the town of Húsavík which is located effectively on top of the Húsavík Flatey Fault (HFF), the largest transform faults in the country (top right Figure 4.2). The ICEARRAY II consists of 6 free-field stations and one structural monitoring system in the regional hospital building. The free-field stations are three-component CUSP-3C instruments of Canterbury Seismic Instruments and the structural system is a CUSP-3D3 unit, with one ground floor instrument, and two three-component sensors on the fifth floor (Halldorsson et al. 2012). The identification ID, name, and location of the ICEARRAY I and II are listed in Table 4.1 and Table 4.2, respectively.

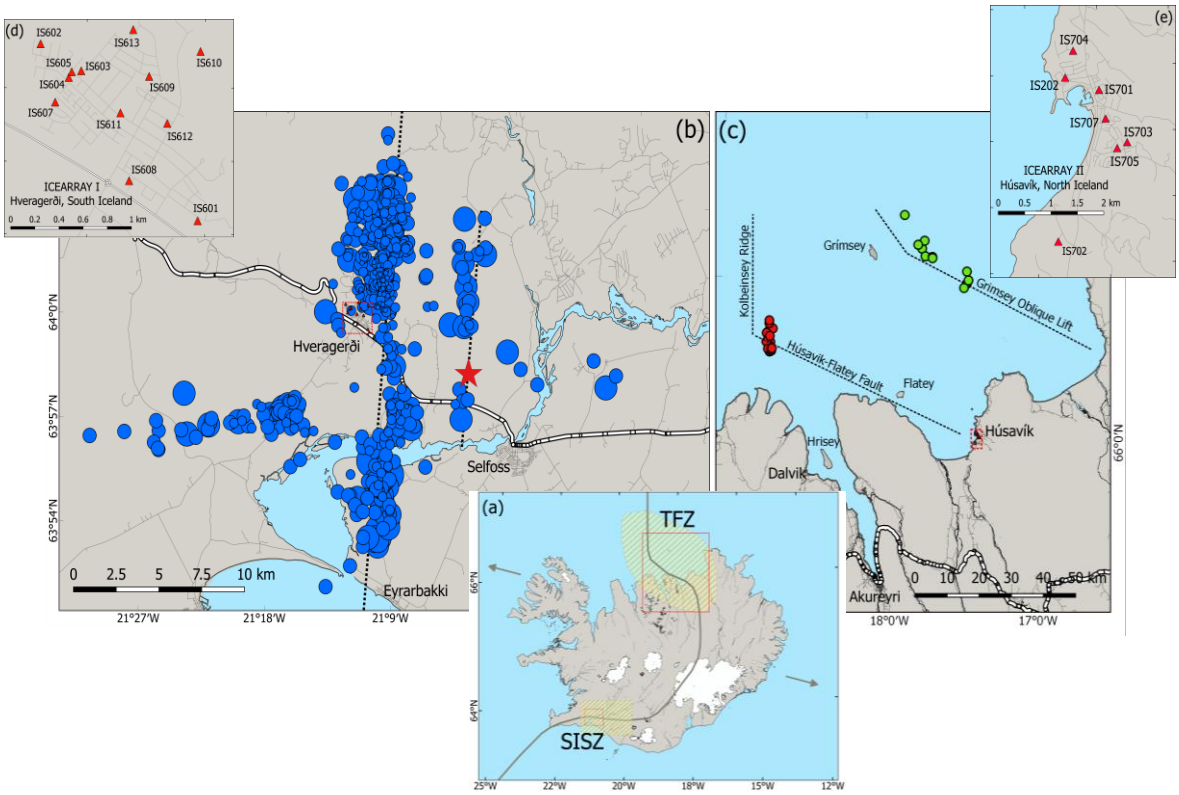

Figure 4.2 (a) Map of Iceland with the approximate location of the Mid-Atlantic Ridge (grey line) and the South Iceland Seismic Zone (SISZ) and the Tjörnes Fracture Zone (TFZ) marked by hatched areas. Red rectangles indicate the areas shown in (b) and (c) with more details. (b) The aftershock distribution (blue circles) from 29 May 2008 Ölfus earthquake in southwest Iceland outlining the two causative earthquake faults (dotted lines). The twelve ICEARRAY I stations (red triangles shown in (d) along with station ID-codes) are located within the town of Hveragerði (red dashed rectangle shown in (b)). (c) Locations of the main events during the October 2012 (red circles) and April 2013 (green circles) earthquake swarms in the TFZ in addition to the main seismic lineaments of the TFZ as black dotted lines. The seven ICEARRAY II stations (red triangular shown in (e) along with station ID-codes) are located within the town of Húsavík (red dashed rectangle shown in (c)).

Table 4.1 ICEARRAY I strong-motion array station locations.

Station ID	Station name	Latitude (˚N)	Longitude (˚E)
IS601	Heidarbrun 51	63.9927	-21.1776
IS602	Kambahraun 39	64.0047	-21.2043
IS603	Dynskogar 3	64.0029	-21.1974
IS604	Borgarhraun 12	64.0024	-21.1995
IS605	Borgarhraun 8	64.0028	-21.1990
IS607	Arnarheidi 26	64.0007	-21.2018
IS608	Sunnumork 2	63.9954	-21.1893
IS609	Dvalarheimilid As	64.0025	-21.1859
IS610	Reykir	64.0042	-21.1772
IS611	Heidmork 31	64.0000	-21.1908
IS612	Reykjamork 17	63.9993	-21.1828
IS613	Laufskogar 39	64.0057	-21.1886

Table 4.2 ICEARRAY II strong-motion array station locations.

Station ID	Station name	Latitude (°N)	Longitude (°E)
IS202	Husavik-Slokkvistod	66.0490	-17.3550
IS701	Husavik-Ketilsbraut 9	66.0468	-17.3411
IS702	Husavik-Kaldbakur	66.0242	-17.3614
IS703	Husavik-Litlagerdi 2	66.0386	-17.3305
IS704	Husavik-Hofdi	66.0530	-17.3511
IS705	Husavik-Holl	66.0378	-17.3348
IS707	Husavik-Gardarsbraut 39	66.0424	-17.3390

Strong-motion Recordings

The deployment of the Icelandic strong-motion arrays in the SISZ and TFZ has started a new era of earthquake strong-motion recordings in Iceland and specifically since then the size of the Icelandic strong-motion data set has been dramatically increased. On close scrutiny, the ICEARRAY I was in the extreme near-fault region of the $M_w 6.3$ Ölfus earthquake on 29 May 2008. The ICEARRAY I stations recorded the main shock and more than 1700 aftershocks during a year of seismicity after the main shock. Despite the relatively small inter-station distances across the array and fairly uniform local site conditions, ground motions of which were characterized by intense ground accelerations (PGA of $38 - 89\%$ g) of relatively short durations ($\sim 5 - 6$ seconds) and large amplitude near-fault velocity pulses due to simultaneous rupture directivity and permanent tectonic displacements. On the other hand, ICEARRAY II has recorded much less data, the far-field ground motions of total of 26 small-to-moderate size earthquakes, during the largest seismic sequence over the last 30 years in North Iceland during $2012 - 2013$ (Halldorsson et al. 2012; Olivera et al. 2014; Rahpeyma et al. 2017) (see Figure 4.2).

The earthquake parametric information was obtained from the SIL seismic network of the Icelandic Meteorological Office which monitors the seismicity of Iceland (Böðvarsson et al. 1999), and the temporary LOKI seismograph network which was deployed in the seismic region within two days of the main shock to provide more accurate hypocentral locations of aftershocks (Brandsdóttir et al. 2010). Unfortunately, however, only 700 of the aftershocks recorded by the ICEARRAY I were found to match with events reported by either SIL or LOKI networks. For the other aftershocks, the source-site distances were estimated from P- and S-phase arrival times and their local magnitudes estimated using an empirical relationship for peak ground velocity (Pétursson and Vogfjörð, Kristín 2009). Figure 4.3 shows the characteristics of the ICEARRAY I and II databases in terms of seismic parameters.

Microseismic data

Microseismic noise is defined as low amplitude vibrations from surface sources randomly distributed in space and time (e.g., vibrations of natural origin such as wind and sea tides, or of manmade origin such as traffic, industrial machinery, etc.). During recent decades, many researchers have investigated the reliability of using ambient noise, both numerically and experimentally, for the quantifications of site effects (Bard 1998; Triantafyllidis et al. 2006; Pilz et al. 2009). In the absence of earthquake strong-motion recordings, microseismic data

is easily obtained and can provide additional constraints and spatial resolution of site effects. Continuous ambient noise recordings of a minimum one-hour duration were performed at the ICEARRAY I and II stations using REF TEK 130-01 Broadband Seismic Recorders and Lennartz LE-3D/5s three-component sensors for which gain level and sampling rate were configured to unity and 100 Hz, respectively (Halldorsson et al. 2009).

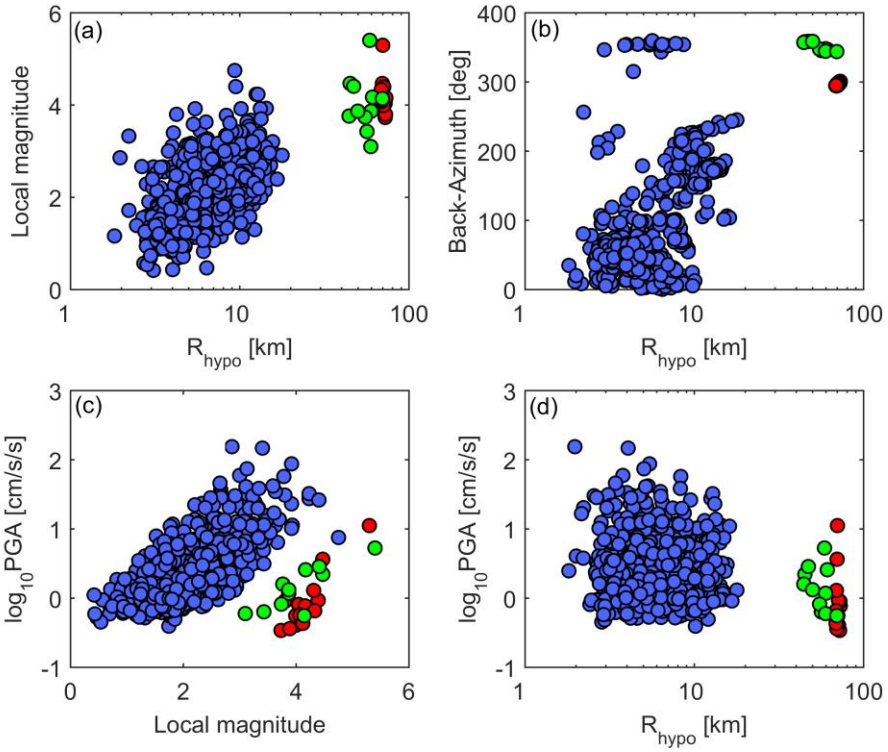

Figure 4.3 Distribution of different earthquake parameters with (a) local magnitude and (b) back-azimuth plotted versus hypocentral distance (R_{hypo}), and then log_{10} PGA plotted versus (c) local magnitude and (d) hypocentral distance as recorded by ICEARRAY I (blue circles) and ICEARRAY II (red circles for the swarm in 2012 and green circles for the 2013 swarm).

4.1.3 Geological settings across ICEARRAY I & II

The main subsoil geological characteristics across ICEARRAY I and II are entirely different (cf. Figure 4.4). Across ICEARRAY I, the uppermost lava layer (~5,000 year-old) lies on top of a softer sedimentary layer, which in turn lies on top of another lava rock layer (~10,000 year-old) resulting in a velocity reversal. We expect this situation to be the case for Hveragerði because based on geological and borehole information the uppermost lava rock layer (~5,000 year-old,) lies on top of a softer sedimentary layer, which in turn lies on top of another lava rock layer (~10,000 year-old) resulting in a velocity reversal. As can be seen in the geological map of Hveragerði, the majority of the ICEARRAY I strong-motion stations sit on stiff lava rock layer. Hence, due to the uniform site condition of the ICEARRAY I and short inter-station distances we expect to see small variability in ground motion amplitudes using ICEARRAY I recordings. However, significant variations in PGA and PGV were observed during the recorded ground motions of the main-shock and aftershocks data set (Halldorsson and Sigbjörnsson 2009; Douglas and Halldorsson 2010). In contrast, local soil conditions across ICEARRAY II are much more complicated in comparison to the ICEARRAY I (see Figure 4.4). Húsavík itself is characterized by several

soft sedimentary layers, which generally overlay the Grjóthváls lava rock and hyaloclastite, and vary in softness as well as depth from site to site in the town (Gudmundsson 2007). The subsoil structure beneath the ICEARRAY II can be clustered into four main geological units: (1) the northernmost part of the town sits on relatively hard Tillite, (2) while the geology along the shoreline towards south is characterized by horizontally layered fluvial sediments. (3) On top of the horizontal sediments lies a delta formation of glacial deposits. (4) The oldest sediments are glacial deposits which have over time been altered to solid Tillite rock which underlay parts of Húsavík (Saemundsson and Karson 2006).

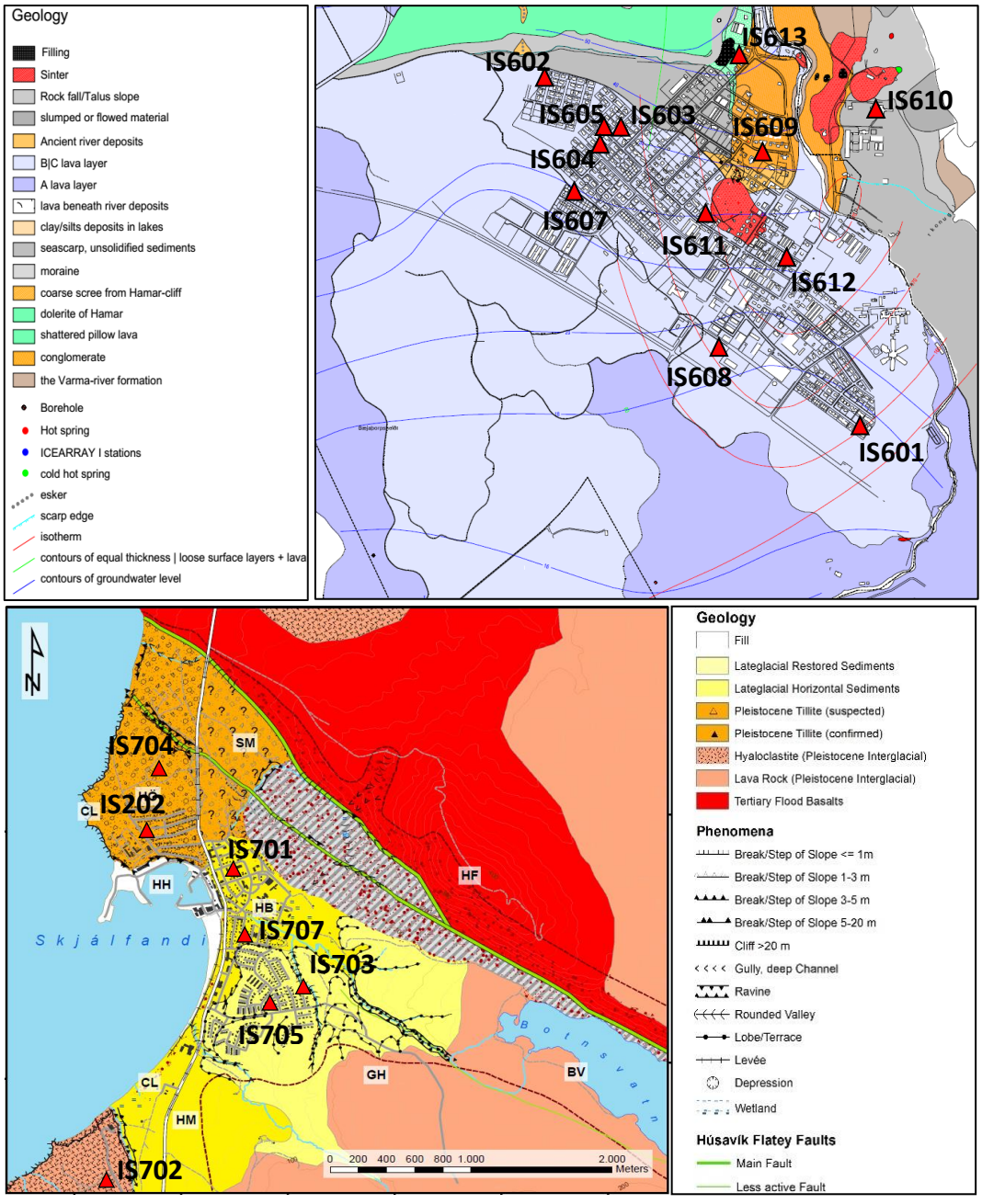

Figure 4.4 Geological map of (top) ICEARRAY I area adopted from Sæmundsson and Kristinsson (2005) and (bottom) ICEARRAY II area adopted from Waltl (2013). The station locations are marked as red triangles.

4.1.4 Results

The extensive ICEARRAY I near-field strong-motion aftershocks recorded after the Ölfus earthquake and ICEARRAY II far-field strong-motion data set provided a great opportunity not only to quantify the local site effects at the array stations on such soil structure, but also the relative differences over short distances. In this study, we applied the methodology chiefly recommended by Site EffectS assessment using AMbient Excitations (SESAME) research project for the task in HVSR analysis (Bard and SESAME-Team 2005). Figure 4.5 and Figure 4.6 represent the HVSR results for each station across ICEARRAY I and II, respectively.

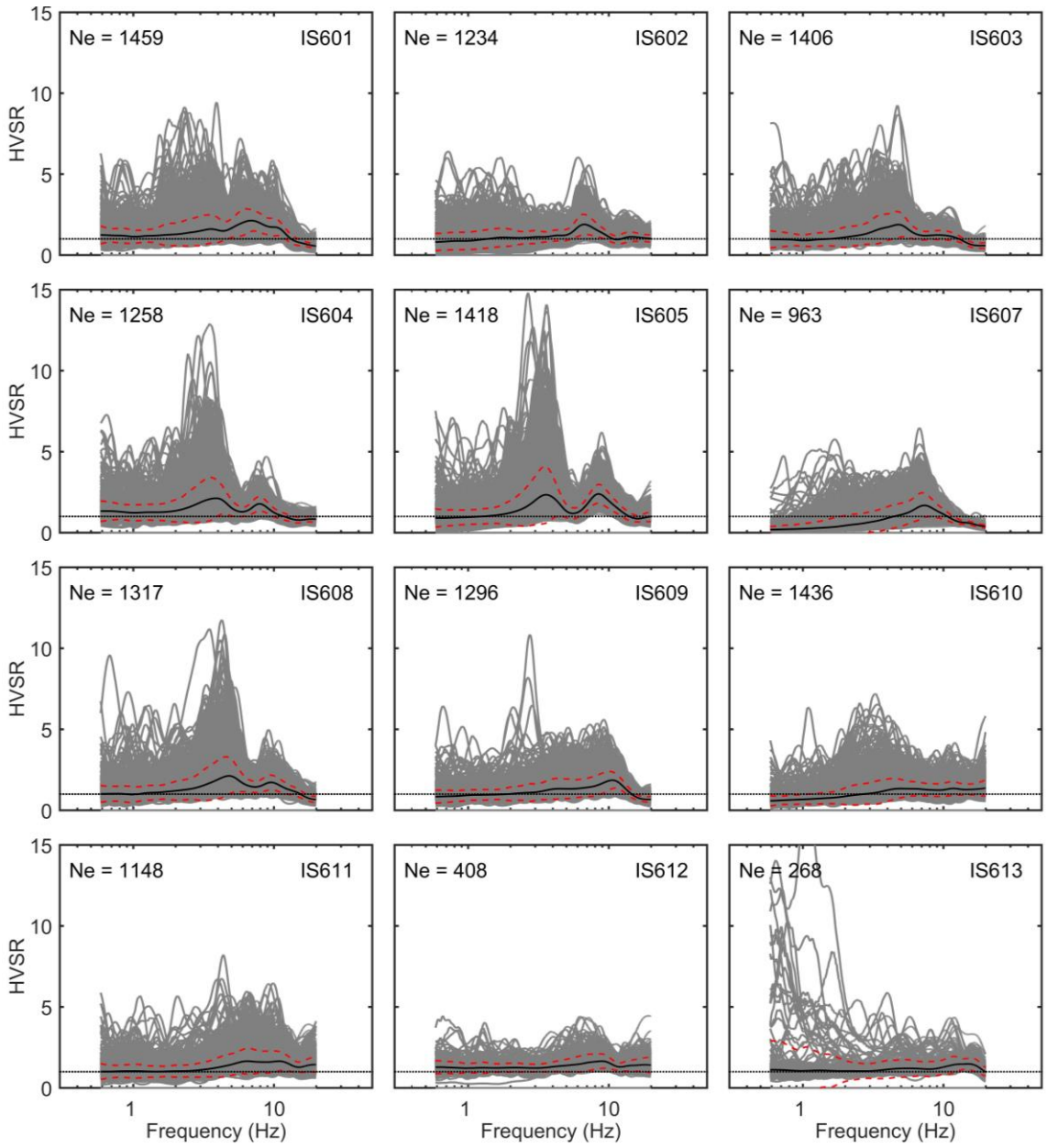

Figure 4.5 The earthquake recordings (gray lines) as a function of frequency for each of the twelve ICEARRAY I strong-motion stations. The geometric mean HVSR (black solid lines) and their corresponding $\pm 1\sigma$ (red dashed lines) are also shown, with Ne the number of available earthquake events used to derive the mean.

To obtain the HVSR, we calculated the absolute Fourier spectrum of each of the three components over the duration of the record after applying several sensitivity analyses for selecting input parameters (e.g. smoothing coefficient, averaging method, etc.). The spectral amplitudes were smoothed using the Konno and Ohmachi with a smoothing coefficient of $B = 20$ (Konno and Ohmachi 1998) for the selected bandwidth. A single smoothed spectrum representing horizontal ground motions was obtained by calculating the geometric mean of the two smoothed horizontal spectra. Dividing the spectrum for the horizontal motions by the spectrum of the vertical component produced the HVSR curve as a function of frequency for each event-station pair. Finally, the average HVSR and the corresponding standard deviation as a function of frequency, have been calculated using geometric mean of HVSR from the recordings.

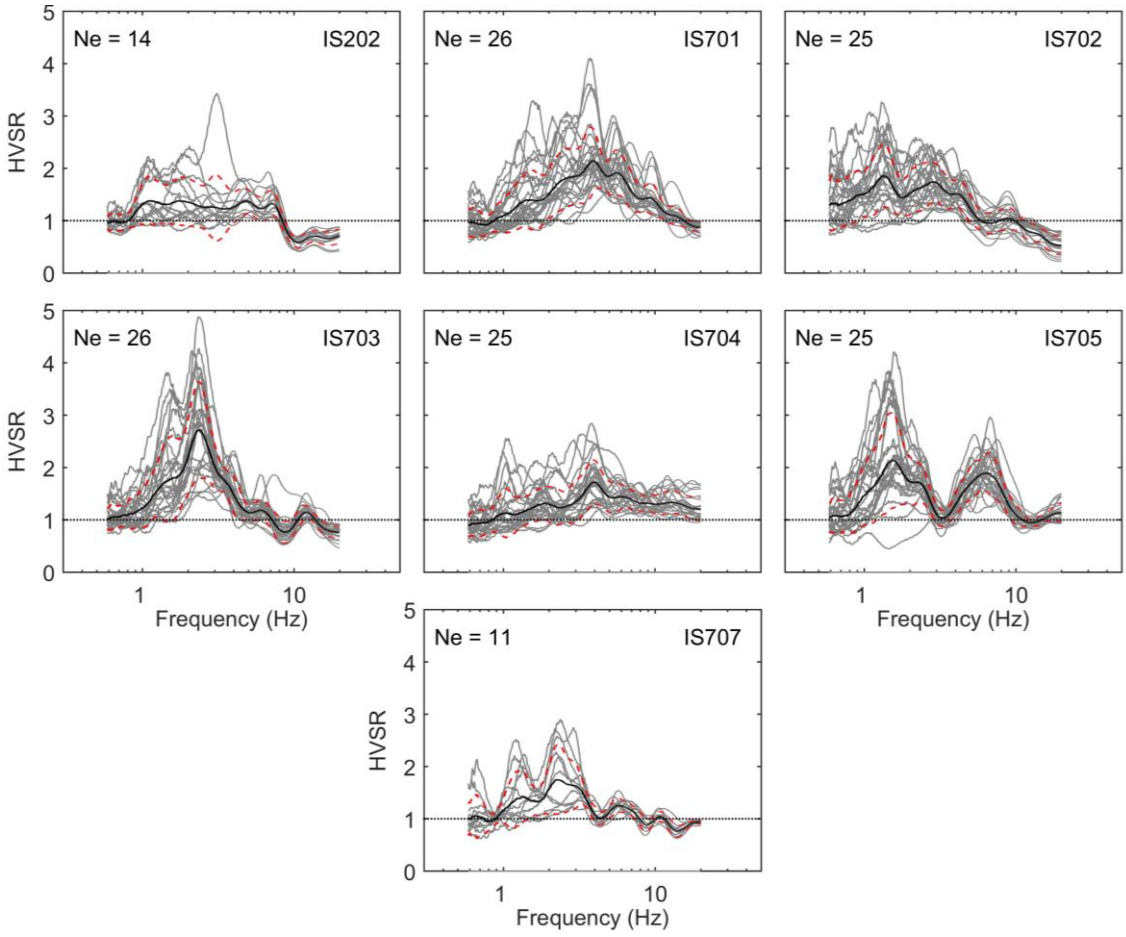

Figure 4.6 The earthquake recordings (gray lines) as a function of frequency for each of the seven ICEARRAY II strong-motion stations. The geometric mean HVSR (black solid lines) and their corresponding $\pm 1\sigma$ (red dashed lines) are also shown, with Ne the number of available earthquake events used to derive the mean.

Sensitivity analysis

Since the microseismic noise is not necessarily random in time and place due to man-made disturbances in the town and environmental factors (e.g., temporary weather/storm vibrations), we tested the sensitivity of the HVSR results in different aspects. First, the smoothing coefficient value was determined on the basis of sensitivity analysis to ensure that sufficient detail was preserved in the resulting spectral ratio. We also compared the

HVSRs obtained for different time windows using both the window length from S-wave arrival time to the end of the coda waves and the entire length of each recording. No significant differences were observed, and as a result, the spectral ratios presented in this study were calculated over the duration of each record. In calculating spectral ratios, we used the arithmetic, geometric, and the quadratic/squared mean methods for combining the two horizontal components. Comparison of the spectral ratios showed insignificant differences, and therefore, the geometric mean method was applied in this study. After reviewing the data and removing segments containing spurious spikes the HVSR was calculated for each station for different times of the day, night and week, and for various durations ranging from several minutes to several hours.

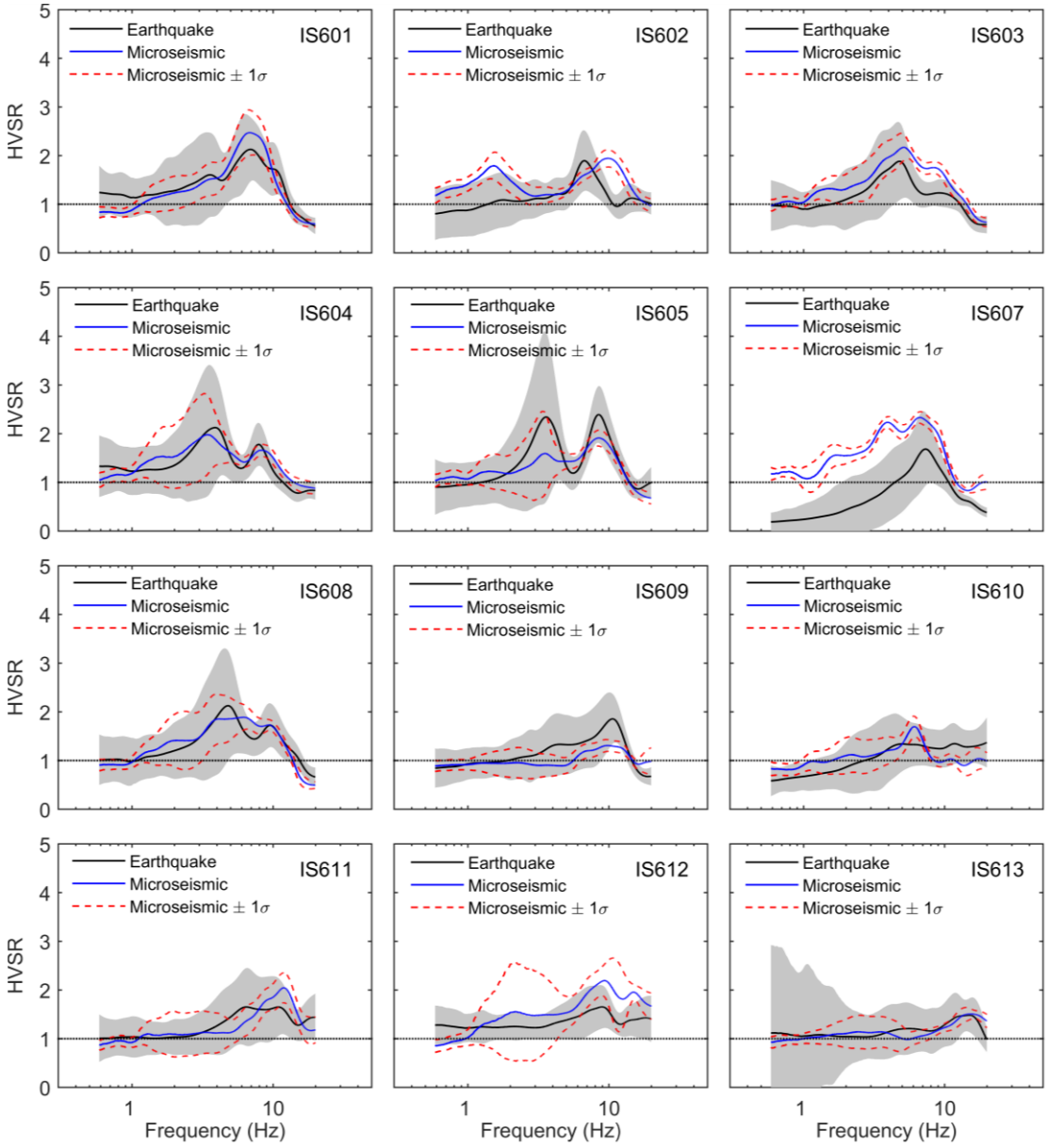

Figure 4.7 Comparison of the mean HVSR estimated from earthquake (black lines) and microseismic (blue lines) data for ICEARRAY I stations. Standard deviations $\pm 1\sigma$ are shown with gray shaded areas (earthquake HVSR) and red dashed lines (microseismic HVSR).

The analysis showed that a stable HVSR at each station was obtained for the optimal window length of 20-minutes. Therefore, for sites where long recordings were available the most stable ones (insofar as being relatively free of spurious signals, traffic and obvious man-made temporary disturbances) were split up into multiple and unique parts of 20 minutes.

In order to evaluate the reliability of the HVSR results and at the same time the applicability of HVSR results from microseismic measurements, the same HVSR procedure was applied to the ambient noise measurements. Figure 4.7 and Figure 4.8 compare the mean HVSR results from earthquake recordings with the mean HVSR results from microseismic data for ICEARRAY I and II stations, respectively. The agreement in terms of overall shape of the amplification curves and their amplitudes is remarkable at almost all stations. The comparison seems to confirm results reported by many studies in other regions (e.g., D'Amico et al. 2008; Pilz et al. 2009) that microseismic data and HVSR analysis may be used with confidence to map the overall amplification characteristics of ICEARRAY I.

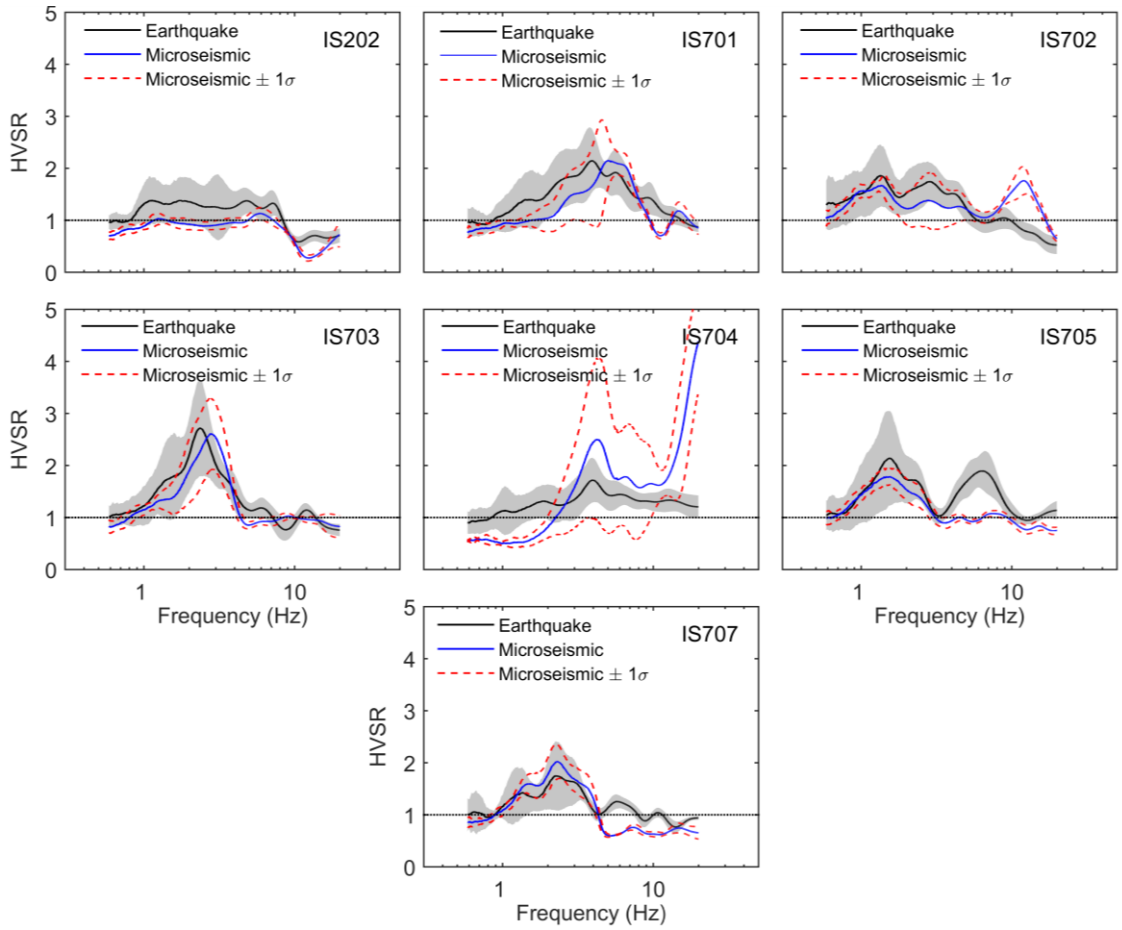

Figure 4.8 Comparison of the mean HVSR estimated from earthquake (black lines) and microseismic (blue lines) data for ICEARRAY II stations. Standard deviations $\pm 1\sigma$ are shown with gray shaded areas (earthquake HVSR) and red dashed lines (microseismic HVSR).

To implement the SSR method the ratio of the geometric mean of the Fourier amplitude spectra for the horizontal components of motions recorded at a site of interest to that of a reference site was computed. As can be seen in the geological map (cf. Figure 4.4), all the ICEARRAY I stations sit on lava-rock. Therefore, all sites were viewed as potential reference sites in the context of the SSR method. On the basis of the HVSR results, relatively

low and approximately uniform (across the frequency range considered) amplifications were observed at stations IS609-IS613, making them potential candidates for a reference site. Station IS612 was excluded since it is located on a relatively young lava-rock, while stations IS609, IS610 and IS613 are located on a considerably older, and presumably more stable, bedrock. Stations IS610 (located on a hillside) and IS613 (relatively few data and unstable HVSR) were excluded, leaving station IS609 as the selected reference station for the SSR method; this station has been in operation since 1999 as a permanent station of the Icelandic strong-motion network. The results of the SSR method using IS609 as the reference station are shown in Figure 4.9, along with the results from the HVSR method for comparison. In general, the results from the two methods are in good agreement.

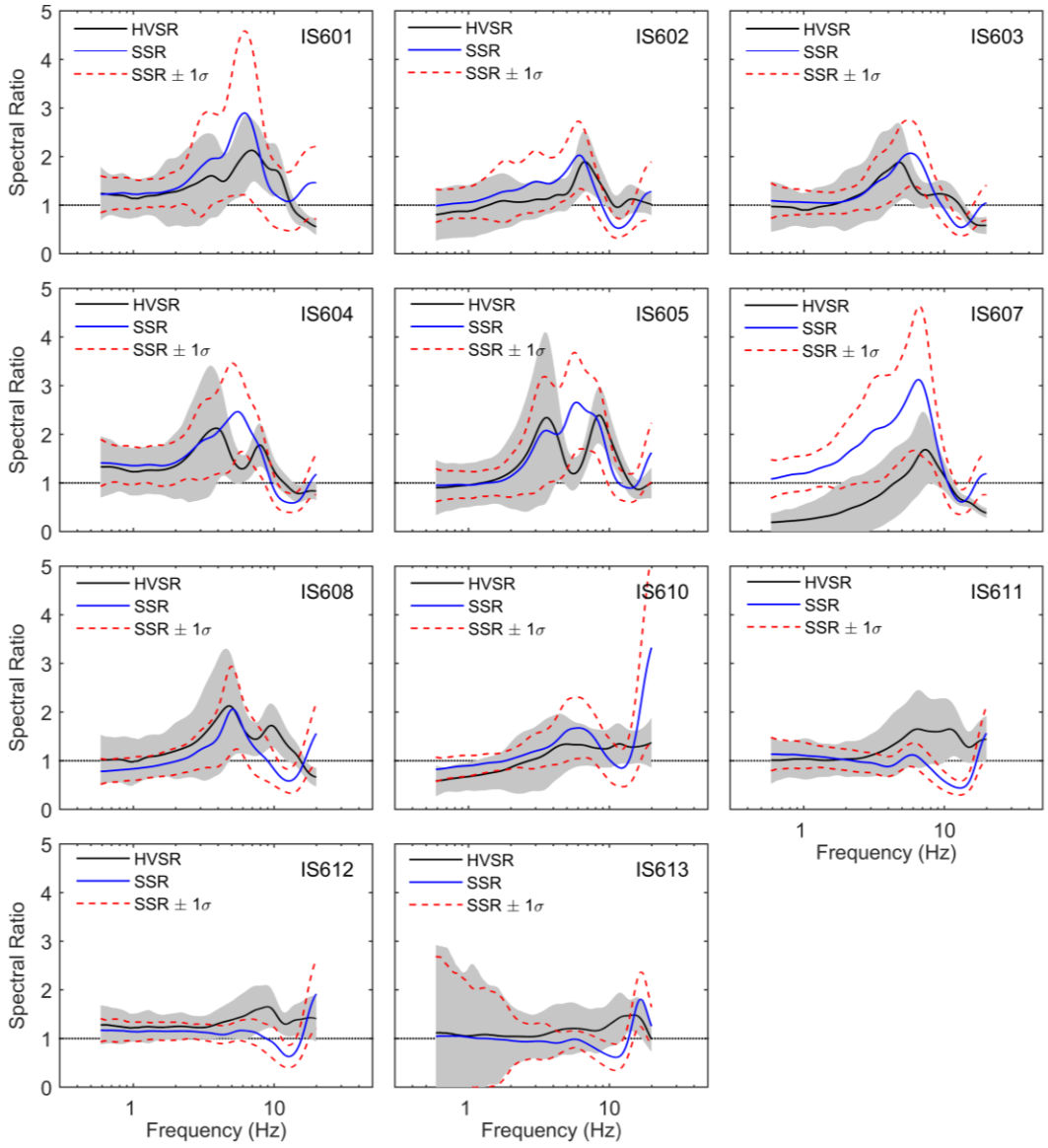

Figure 4.9 Comparison of the mean HVSR (black lines) and mean SSR (blue lines) determined from the earthquake data for each of the ICEARRAY I strong-motion stations, using station IS609 as a reference site. Standard deviations of the means are indicated by red dashed lines for SSR and with gray shaded areas for HVSR.

Focusing on the distribution of predominant frequency in ICEARRAY-II shows that there is a general north-south trend of decreasing peak predominant frequency that is in high

correlation with geological settings of the area. As it is abovementioned, the northernmost part of the town lies on Pleistocene Tillite hard rock and the rest of town is mostly located on sedimentary layer. Figure 4.6 shows that motions from stations located on sedimentary layer in central part of the city are characterized by lower predominant frequency with broad or multiple amplification curve and higher strong-motion amplitude. Therefore, we nominated station IS704 as reference site to obtain the SSR amplification curves. The comparing results for ICEARRAY II stations are presented in Figure 4.10.

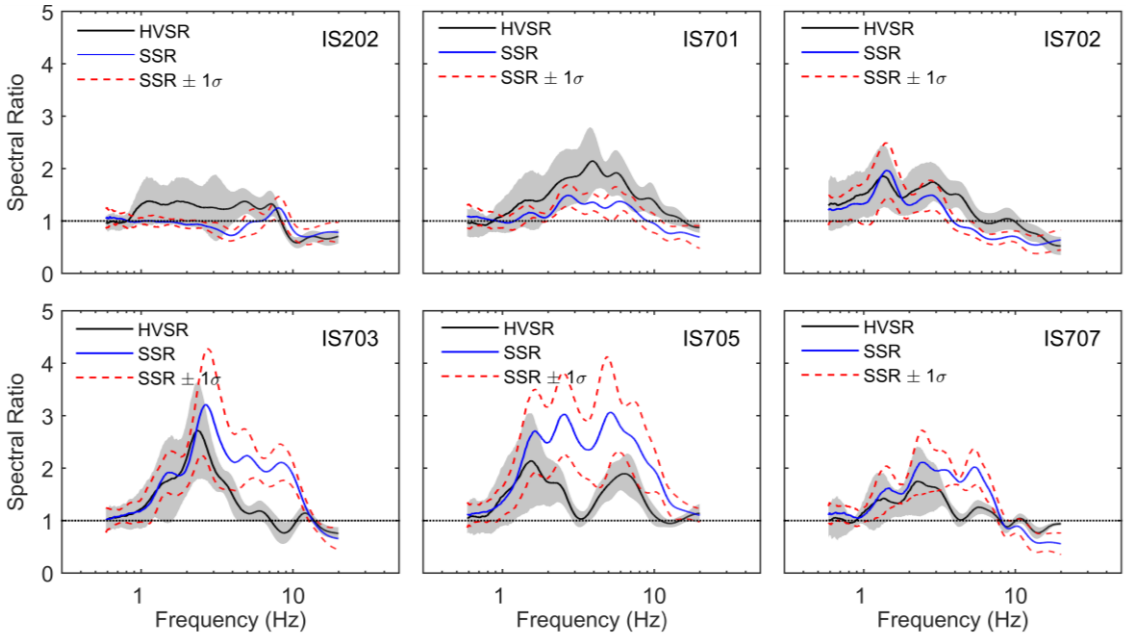

Figure 4.10 Comparison of the mean HVSR (black lines) and mean SSR (blue lines) determined from the earthquake data for each of the ICEARRAY II strong-motion stations, using station IS704 as a reference site. Standard deviations of the means are indicated by red dashed lines for SSR and with gray shaded areas for HVSR.

Topology of the HVSR and SSR results

Despite the small aperture of the Icelandic arrays, Figure 4.7 to Figure 4.10 highlight significant variations in the site effect characterizations. On close scrutiny, we can cluster the results into four following groups:

- Bimodal amplification curves (ICEARRAY I: IS604, IS605, IS608 and ICEARRAY II: IS705)
- Single narrow-band peak amplification curves (ICEARRAY I: IS602 and ICEARRAY II: IS703)
- Broad amplification curves over a wide frequency range (ICEARRAY I: IS601, IS603 and ICEARRAY II: IS701, IS707)
- very low and uniform amplification curves across the frequency range (ICEARRAY I: IS609-613 and ICEARRAY II: IS202, IS704)

This variation in the HVSR highlights the existence of the significant site effects due to complex and varying geostructural settings within the town of Hveragerði and Húsavík.

With the exceptions of stations IS609, IS610, and IS613, which are located on a very old bedrock, the ICEARRAY I strong-motion stations are located on lava-rock layer overlying sedimentary layer (see Figure 4.4). In Hveragerði, the uppermost lava layer is relatively young basaltic lava of varying thickness (10-30 meters according to boring log information). The lava layer lies on top of a sedimentary layer (of similarly varying thickness) which in turn overlies an older lava layer. No information on the thickness of the lower layer exists, but it flowed from the same volcanic fissure and based on the spatial extent of the lava it is most likely of similar thickness as the younger lava layer. From the typical layered structure of young geological formations in south Iceland, it is extremely likely that the older lava layer is underlain by another sedimentary layer. The relatively low and uniform amplification in the northeast of the ICEARRAY I reveals that the geologic profile does not exhibit any sharp impedance contrast (e.g. IS609-613). In contrast, theoretical and numerical investigations for station IS605 in this study imply that the bimodal amplification occurs because of two large impedance contrasts, one deep and the other shallow. The same trend can be observed at IS604 which is located 54 meters southwest from IS605, indicating similar substructures. When compared with the amplification curve of station IS603, however, the difference indicates that the velocity contrast under the station is much less abrupt compared to IS603 and especially IS605. These amplification curves are in stark contrast with those at the bedrock stations, which are characterized by high frequency peaks (~10 Hz) of very low amplitudes.

In contrast to Hveragerði, the geological setting of Húsavík is generally characterized by several Pleistocene and Holocene sedimentary layers, which mainly overlie the Grjótháls lava and breccia. The HVSR of approximately unity across the frequency range of IS704 and IS202 is representative of the hard layer underlying the sites, since compact sediments and hard rock sites generally tend to experience no significant ground amplification. The southwest of Húsavík, where station IS702 is located, is characterized as being underlain by hyaloclastites (Pleistocene Interglacial), and the stations in the central part of the town (i.e., IS701, IS703, IS705, and IS707) sit on the Lateglacial restored/horizontal sediments. It should be noted that there is considerable variation in subsurface topography within the town and such geologic profiles enhance the variation in site effects and ground-motion amplitudes.

4.1.5 Seismic parameters vs. HVSR characteristics

It has been well investigated that a sharp peak in HVSR curve is indicative of high impedance contrast between a softer layer overlying a stiffer layer (Bard and SESAME-Team 2005; Di Giacomo et al. 2005; Nakamura 2008). Hence, we interpret the clear bimodal amplification curve from earthquake data at station IS605 (ICEARRAY I) and station IS705 (ICEARRAY II) as coming from two considerable velocity contrasts within depth, indicating a complex structure of repeated hard rock (e.g. lava, Tillite, stiff sediments)-soft sediment strata. However, the explanations for bimodal amplification curves for stations IS605 and IS705 are likely entirely different. It should be emphasized that due to small epicentral distances (less than 20 km) and high apparent velocity over the ICEARRAY I, the waves consists exclusively of body waves with near vertical incidence angles. In contrast, the waves in ICEARRAY II are assumed to be comprised mostly of surface waves as a result of large epicentral distance ($\sim 40 - 80$ km) and likely lower velocity over the array.

A close scrutiny on Figure 4.7 reveals that the HVSR in station IS605 is characterized by a greater variability in the lower-frequency peak ($\sim 3-4$ Hz) but less in the higher-frequency peak ($\sim 8-9$ Hz) for both earthquake and microseismic data, with the level of mean amplification from the strong-motion recordings for both peaks being almost the same. Moreover, the variability of the lower amplification peak obtained from microseismic data is explicitly higher than the second peak. Hence, in Figure 4.11 we try to perform a detailed investigation to characterize the possible pattern of the bimodal amplification curve versus different seismic parameters by clustering the aftershocks into two groups, those associated with lower (gray, Figure 4.11(b)) and higher (cyan, Figure 4.11(c)) predominant frequencies. Figure 4.11(a) reveals that the two groups are comprised of aftershocks that occurred nearly equally on the two different fault structures. Furthermore, there is no azimuthal dependency of predominant frequency (Figure 4.11(g)). Similarly, both groups of earthquakes have approximately similar coverage of hypocentral distances (Figure 4.11(f)). Nevertheless, there is a noticeable correlation between earthquake intensity and the associated predominant frequency (Figure 4.11(d-e)). This means that events with higher energy generally appear to excite the lower-frequency peak. In other words, the lower predominant frequency peak is clearly observed for earthquakes of large magnitude that mainly are associated with maximum HVSR amplitude and much greater scatter in amplitude.

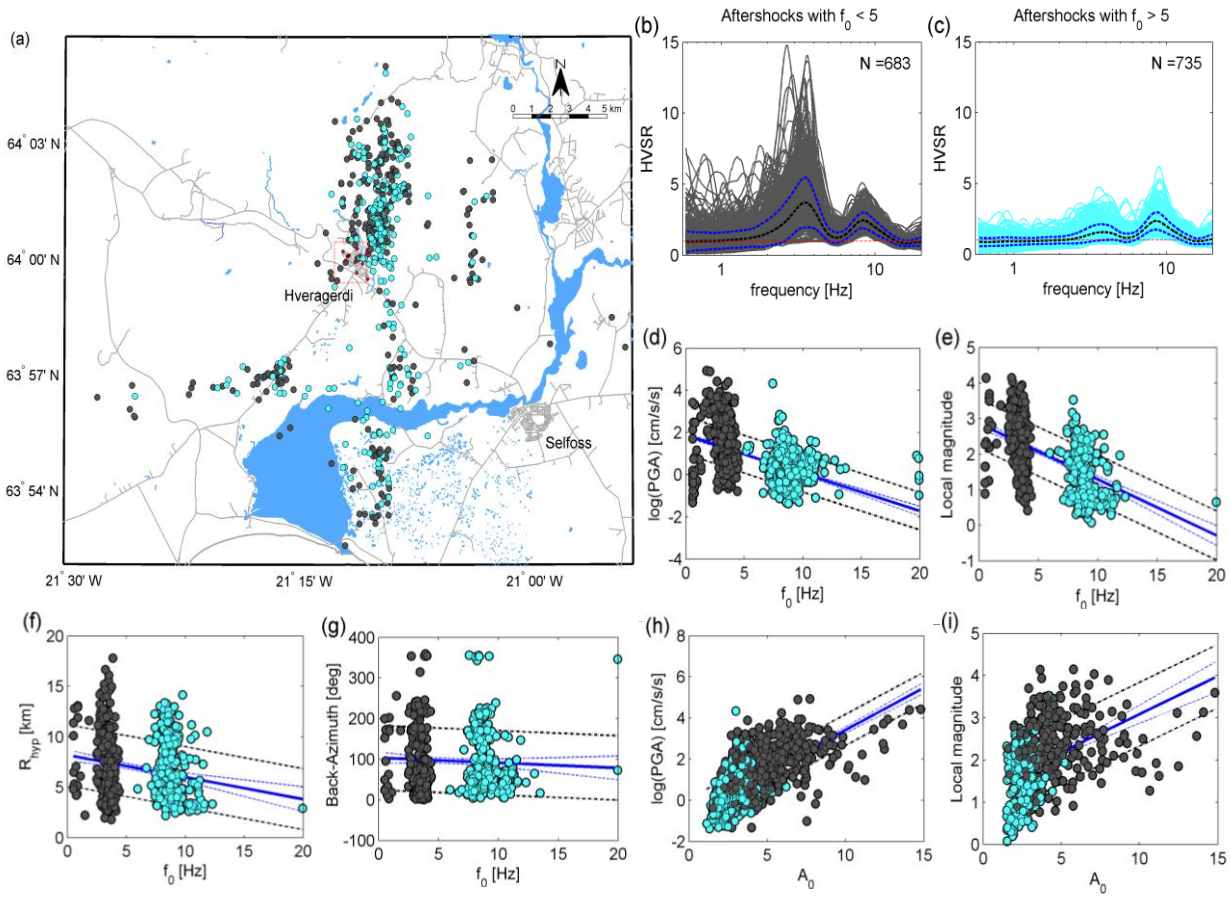

Figure 4.11 (a) distribution of aftershock locations recorded by ICEARRAY I grouped according to the f_0 range above (cyan) and below (gray) 5 Hz, using station IS605. The HVSR curves for two groups of aftershocks is shown in (b) and (c) as well as the mean HVSR $\pm 1\sigma$, with N the number of available earthquakes. Also shown the (d) PGA, (e) local magnitude, (f) hypocentral distance, and (g) back-azimuth versus the predominant frequency f_0. (h) PGA and (i) event magnitude versus Amplification factor A_0 (Gray: $f_0 \leq 5$ Hz and cyan: $f_0 > 5$ Hz).

As can be seen in Figure 4.11(h-i), smaller magnitude earthquakes with lower energy mainly associated to relatively constant and low peak amplification at higher frequency (cyan symbols with $f_0 \sim 8-9$ Hz). In other words, almost all earthquakes that are associated with the amplification peak at higher frequency have amplitudes less than ~3. This observation specifies that the peak in the HVSR amplification curve at lower predominant frequency is largely associated with earthquakes of relatively larger magnitudes, and in the vast majority of the cases, they are associated with maximum HVSR amplitudes higher than ~3 and with much greater variability in amplitudes.

In contrast to the ICEARRAY I that we could not capture any directional dependency, the predominant frequencies of bimodal amplification in ICEARRAY II are highly correlated to the sources. As can be seen in Figure 4.8, microseismic measurements are not able to capture the shallower layer in soil structure and the soil layering filters out the high frequencies. It is clear that low-frequency peaks are mainly associated with the 2012 swarm occurred on HFF, while the 2013 swarm that took place on the Grimsey Oblique Rift is linked to the high-frequency peaks (Figure 4.12(a)). The significant dependency to the direction of incoming waves is clear in Figure 4.12(g). It is also conspicuous that events associated with lower-frequency peaks have larger hypocentral distance (Figure 4.12(f)) and events with larger magnitude are associated with lower-frequency peaks and larger amplification.

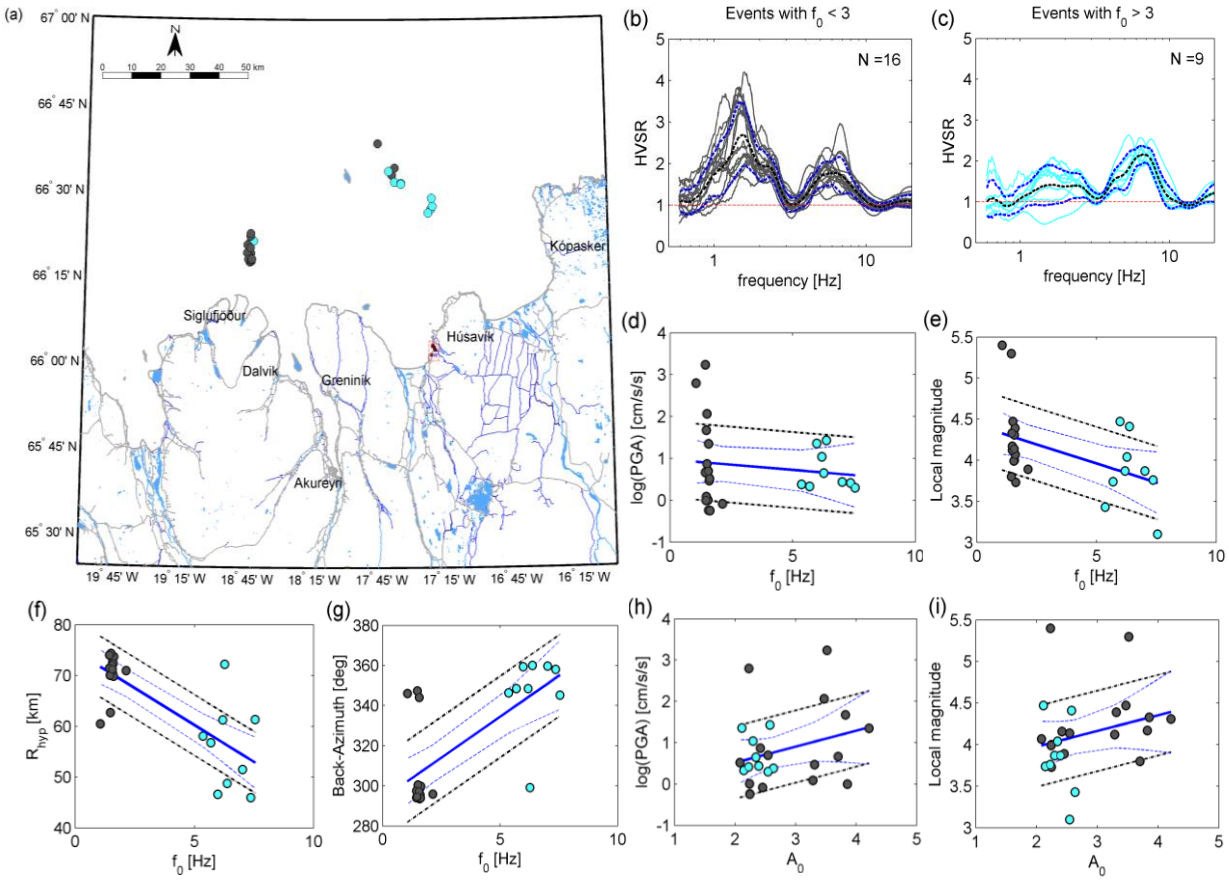

Figure 4.12 (a) distribution of 26 events locations recorded by ICEARRAY II grouped according to the f_0 range above (cyan) and below (gray) 3 Hz, using station IS705 located in the town of Húsavík (red dashed rectangle). The HVSR for the two groups of events is shown in (b) and (c) as well as the mean HVSR $\pm 1\sigma$, with N the number of available earthquakes. Also shown the (d) PGA, (e) local magnitude, (f) hypocentral distance, (g) back-azimuth versus the predominant frequency f_0. (h) PGA, (i) event magnitude versus Amplification factor A_0 (Gray: $f_0 \leq 3$ Hz and cyan: $f_0 > 3$ Hz).

4.1.6 Modeling the bimodal response

It has been well investigated that the shape of the HVSR amplification curves are in correlation with the subsoil structure (Nakamura 2008; Bonnefoy-Claudet et al. 2009). We physically interpret the two peaks of the amplification curve from earthquake data at station IS605 as coming from two considerable and abrupt velocity contrasts at depth, indicating that a multilayer subsoil model is needed for modeling. During recent decades, many researchers tried to apply the HVSR technique as a practical tool to determine the V_s profile (e.g., Arai and Tokimatsu 2004, 2005; Herak 2008). It should be mentioned that the inversion of physical parameters of the subsoil structure as a function of depth on the basis of the HVSR using the body-wave approximation provides better results around the resonance frequency (Tsai 1970; Albarello and Lunedei 2010).

In this study, we apply the theoretical approach presented in section 2.3.1 using the body-wave approximation. However, our attempts to reproduce the bimodal amplification curve at station IS605 using this approach were largely unsuccessful. As can be seen in Figure 4.13, for reasonable values of the density and shear wave velocity of the lava and sedimentary layers, and using layer thicknesses from borehole logs, the method fails to reproduce two physical peaks at the observed predominant frequencies of station IS605. Figure 4.13(a) shows that adding two layers of hard rock to the top and in the middle of the soft soil stratum (70 m of sedimentary layer) does not significantly change the predominant frequency or the amplification curve. On the contrary, Figure 4.13(b) shows that inserting two soil layers with lower impedance ratios dramatically affects the respective amplification. The amplification curve is thus dominated by the deepest velocity contrast (the underlying bedrock assumed as half-space and an overlaying sedimentary layer) and the reverse velocity contrast between the hard layers and the intermediate stiff layers is essentially invisible to the method. As a result, we find that the body wave assumption in interpreting the HVSR results from data recorded on a profile with velocity reversals may lead to ambiguous interpretation.

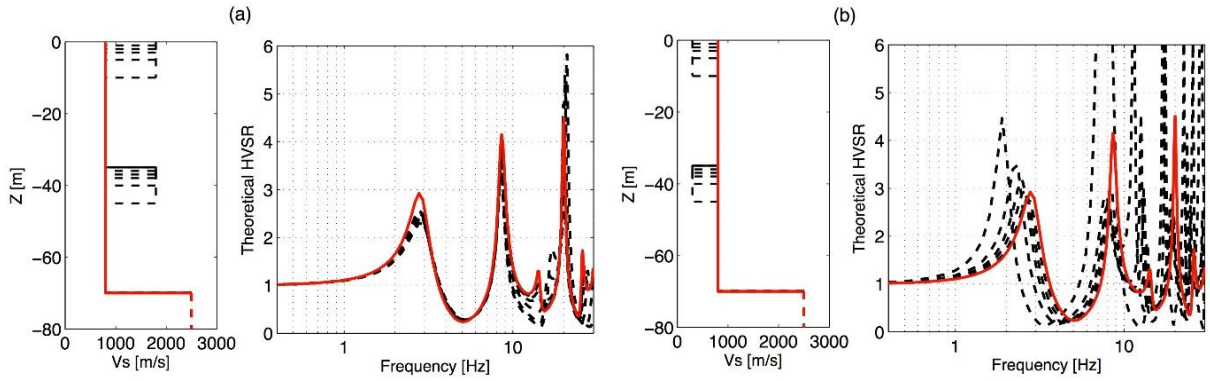

Figure 4.13 The effects of velocity reversals on wave propagation and soil amplification. Inserting two layers with (a) higher and (b) lower velocities.

Alternatively, due to the obvious mechanical similarities to that of a dynamic structural system, we model the geologic profile as a classically damped dynamic system subjected to a base excitation (Chopra 1981; Yoshida 2015). Available geological information and borehole records show the existence of two Holocene lava layers under IS605 and for that reason a two-degree of freedom (2DOF) system was assumed to model the bimodal HVSR amplification curve at station IS605.

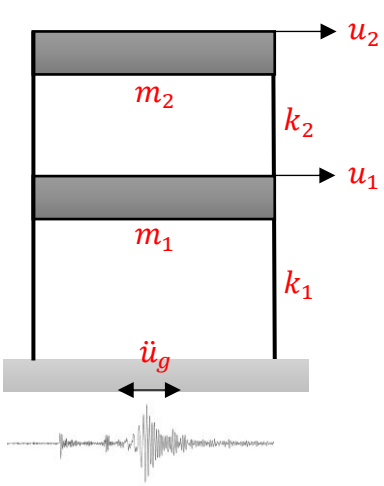

Figure 4.14 Mechanical analogue of a 2DOF undamped system subjected to earthquake excitation at the base.

In Figure 4.14, two-degree-of-freedom undamped system subjected to earthquake excitation at the base is illustrated. In this above mechanical analogue, we consider the lava-rock layers rigid (contributing only in the masses). The sedimentary layers are modeled as shear columns. Their mass is lumped in the masses m_1 and m_2. Specifically, half of the mass of layer 1 is lumped in mass m_1 and the other half in mass m_2.

The equation of motion and system matrices for the system in Figure 4.14 is

$$\begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \begin{Bmatrix} \ddot{u}_1 \\ \ddot{u}_2 \end{Bmatrix} + \begin{bmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_2 \end{bmatrix} \begin{Bmatrix} u_1 \\ u_2 \end{Bmatrix} = -\begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} \cdot \ddot{u}_g(t) \tag{4.1}$$

Eq. (4.1) can be presented in the general format as Eq. (4.2)

$$\boldsymbol{m\ddot{u}} + \boldsymbol{ku} = -\boldsymbol{m\iota} \cdot \ddot{u}_g(t) \tag{4.2}$$

where $\boldsymbol{k}$ and $\boldsymbol{m}$ are the stiffness and mass matrices, respectively, and $\boldsymbol{\iota}$ is the influence vector representing the mass displacements from static application of a unit-ground displacement. We will solve the Eq. (4.1) and Eq. (4.2) using modal decomposition, assuming that classical modes exist (i.e., that the damping matrix $\mathbf{c}$ of the system is diagonalizable by the normal modes of the undamped system). Admittedly, this is a strong assumption for our system (alternating hard magma layers and soft sediment layers) and most probably is not valid. However, assuming the more general case of non-classical damping would require more involved computations and defeats the purpose of this exercise. As presented in Eq. (4.3), the nodal displacement vector $\boldsymbol{u}(t)$ of the system can be expressed in terms of modal coordinates by using the expansion theorem for multi-degree-of-freedom (MDOF) systems (modal superposition).

$$\begin{aligned} \boldsymbol{u}(t) = \sum_{n=1}^{N} \boldsymbol{u}_n(t) &= \sum_{n=1}^{N} \boldsymbol{\phi}_n q_n(t), \qquad n = 1,2 \\ &= \begin{Bmatrix} \phi_{11} \\ \phi_{21} \end{Bmatrix} q_1(t) + \begin{Bmatrix} \phi_{21} \\ \phi_{22} \end{Bmatrix} q_2(t) \end{aligned} \tag{4.3}$$

where $q_n(t)$ are the modal coordinates, and $\boldsymbol{\phi}_n$ are the natural mode shapes. The undamped modal frequencies ω_n and modes $\boldsymbol{\phi}_n$ can be obtained by solving the eigenvalue problem $((\boldsymbol{k} - \omega_n^2 \boldsymbol{m})\boldsymbol{\phi}_n = \mathbf{0})$.

Using the normal modes, the equations of motion are decoupled, and we obtain the uncoupled modal equations:

$$\ddot{q}_n(t) + 2\zeta_n \omega_n \dot{q}_n(t) + \omega_n^2 q_n(t) = -\Gamma_n \ddot{u}_g(t), \quad n = 1,2 \tag{4.4}$$

Where the mode participation factor Γ_n, the factor converting ground acceleration to force at each degree of freedom L_n, and modal masses M_n are

$$\Gamma_n = \frac{L_n}{M_n} \quad , \quad L_n = \boldsymbol{\phi}_n^T \mathbf{m}\boldsymbol{\iota} \quad , \quad M_n = \boldsymbol{\phi}_n^T \mathbf{m} \boldsymbol{\phi}_n$$

In general, the equation for a damped system can be written as

$$\boldsymbol{M}_n \ddot{\boldsymbol{q}}_n(t) + \boldsymbol{C}_n \dot{\boldsymbol{q}}_n(t) + \boldsymbol{K}_n \boldsymbol{q}_n(t) = -\boldsymbol{m}\boldsymbol{\iota} \ddot{u}_g(t) \tag{4.5}$$

where $\boldsymbol{M}_n$, $\boldsymbol{C}_n$, and $\boldsymbol{K}_n$ are the generalized mass, damping and stiffness matrices, respectively.

We are looking for steady-state solutions of the modal equations to an excitation of the form

$$\ddot{u}_g(t) = \ddot{u}_{go}(\omega) e^{i\omega t} \tag{4.6}$$

where $\ddot{u}_{go}(\omega)$ is the frequency response of the excitation and ω is the forcing frequency; the same functional form for the modal coordinates is also assumed. For such an excitation, the steady-state solution of the modal coordinates is of the form

$$q_n(t) = q_{no}(\omega) e^{i\omega t} \tag{4.7}$$

Substituting these expressions into the modal Eq. (4.4) we obtain:

$$[-\omega^2 + 2\xi_n \omega_n (i\omega) + \omega_n^2] \cdot q_{no}(\omega) e^{i\omega t} = -\Gamma_n \ddot{u}_{go}(\omega) e^{i\omega t} \tag{4.8}$$

Or equivalently,

$$q_{no}(\omega) = -\frac{\Gamma_n \ddot{u}_{go}(\omega)}{\omega^2 \cdot \left[\left(\left(\frac{\omega_n}{\omega}\right)^2 - 1\right) + i2\xi_n \left(\frac{\omega_n}{\omega}\right)\right]} \quad , \quad (n = 1,2) \tag{4.9}$$

Upon insertion into Eq. (4.5) the relative displacement frequency response of interest

$$u_2^R(t) = \phi_{21} q_1(t) + \phi_{22} q_2(t) \tag{4.10}$$

the relative displacement response of the second degree of freedom that is of interest to us (response of the free surface) is obtained as (i.e., top of the soil structure, denoted as $u_{2o}^R(\omega)$ in harmonic from):

$$u_{2o}^R(\omega) = -\left(\frac{\phi_{21}\Gamma_1}{\omega^2 \left[\left(\left(\frac{\omega_1}{\omega}\right)^2 - 1\right) + i2\xi_1 \left(\frac{\omega_1}{\omega}\right)\right]} + \frac{\phi_{22}\Gamma_2}{\omega^2 \left[\left(\left(\frac{\omega_2}{\omega}\right)^2 - 1\right) + i2\xi_2 \left(\frac{\omega_2}{\omega}\right)\right]} \right) \ddot{u}_{go}(\omega) \tag{4.11}$$

where for the *n*-th mode of the 2DOF system ξ_n is the damping ratio, ω_n is the undamped natural circular frequency, ϕ_{2n} is natural mode shapes, and Γ_n is the modal participation factor. Eq. (4.11) can be written as $u_{2o}^R(\omega) = H(\omega)\ddot{u}_{go}(\omega)$, where $H(\omega)$ is the displacement transfer function.

$$H(\omega) = -\left(\frac{\phi_{21} \cdot \Gamma_1}{\omega^2 \cdot \left[\left(\left(\frac{\omega_1}{\omega} \right)^2 - 1 \right) + i2\xi_1 \left(\frac{\omega_1}{\omega} \right) \right]} + \frac{\phi_{22} \cdot \Gamma_2}{\omega^2 \cdot \left[\left(\left(\frac{\omega_2}{\omega} \right)^2 - 1 \right) + i2\xi_2 \left(\frac{\omega_2}{\omega} \right) \right]} \right) \tag{4.12}$$

We record the absolute motion at the free surface. Therefore, the total displacement on the free surface (ground) denotes as

$$u_{2o}^T(t) = u_{2o}^R(t) + u_{go}(t) \tag{4.13}$$

Which when written in harmonic form gives

$$\begin{aligned} u_{2o}^T(\omega)e^{i\omega t} &= u_{2o}^R(\omega)e^{i\omega t} + u_{go}(\omega)e^{i\omega t} \\ &= H(\omega)\ddot{u}_{go}(\omega)e^{i\omega t} + u_{go}(\omega)e^{i\omega t} \end{aligned} \tag{4.14}$$

Subsequently, we can obtain the ratio of the total displacement frequency response at the free surface, $u_{2o}^T(\omega)$, to the input displacement at the bottom, $u_{go}(\omega)$, in the harmonic form as Eq. (4.15).

$$\frac{u_{2o}^T(\omega)}{u_{go}(\omega)} = H(\omega)\frac{\ddot{u}_{go}(\omega)}{u_{go}(\omega)} + 1 \tag{4.15}$$

From

$$\ddot{u}_{go}(\omega) = (i\omega)^2 u_{go}(\omega) \tag{4.16}$$

we obtain the ratio of the total displacement frequency response at the surface $u_{2o}^T(\omega)$ to the input displacement at the bottom $u_{go}(\omega)$ as

$$\frac{u_{2o}^T(\omega)}{u_g(\omega)} = -\omega^2 \cdot H(\omega) + 1 \tag{4.17}$$

In order to numerically compute the relative and absolute response from Eq. (4.11) and Eq. (4.13) we need to estimate the input parameters for the 2DOF dynamic system. In our modeling, the lava layers are assumed to be rigid masses m_1 and m_2 (where the subscript 1 refers to the lower and 2 to the upper lava layer in Figure 4.4) and the sedimentary layers are assumed to be massless lateral springs having stiffnesses k_1 and k_2. The masses of the sedimentary layers are lumped into m_1 and m_2 (i.e., one half of the mass of each sedimentary layer above and below the lava layer is lumped with the mass of the lava layer). Considering a unit-area vertical column of the soil profile the mass is calculated from $m = \rho H$ where ρ

is density and H is the layer thickness. The shear stiffness is calculated by taking the advantage of the relation of shear modulus of the material (μ), its density (ρ) and shear wave velocity (β) or $k = \mu/H$ where $\mu = \rho\beta^2$. From shallow borehole logs in Hveragerði the thickness of the upper lava-rock layer ($B|C$) is seen to vary between $14 - 22$ m, and the thickness of the sedimentary layer comprised of loose alluvial and marine sand and gravel is between approximately $10 - 22$ m. However, there is no information about other material characteristics (ρ or β) or the layering of deeper layers. As a result, we relied on material properties from other studies in South Iceland on similar geology. By solving the eigenvalue problem numerically, we estimate two natural frequencies of oscillation at $f_1 \approx 3.5$ Hz and $f_2 \approx 8.5$ Hz which match almost exactly the two mean predominant frequencies observed in the HVSR amplification curves.

Table 4.3 summarizes the parameters of the final model which is schematically shown in Figure 4.15. The resulting soil structure is shown schematically in Figure 4.15(a), which from the surface to the top of the bedrock, consists of a 15 m thick lava layer (top), 22 m sedimentary layer, 12 m lava layer (bottom) and a 12 m thick sedimentary layer. The corresponding ratio of the absolute displacement frequency response at the free surface to the input ground displacement is shown in Figure 4.15(b). Figure 4.15(b) clearly indicates that there are two fundamental frequencies of oscillations.

Table 4.3 Soil and 2DOF model properties.

Layer	Soil				2DOF system	
	$H\ [m]$	$\rho\ [g/cm^3]$	$V_s\ [m/s]$	$G_s\ [N/m^2]$	$k\ [N/m]$	$m\ [kg]$
Lava (L_1)	15	2.2	1800	7.13×10^9		51.7×10^3
Sediment (L_2)	22	1.7	750	0.95×10^9	43.5×10^6	
Lava (L_3)	12	2.2	1800	7.12×10^9		55.9×10^3
Sediment (L_4)	12	1.8	800	1.15×10^9	96.0×10^6	

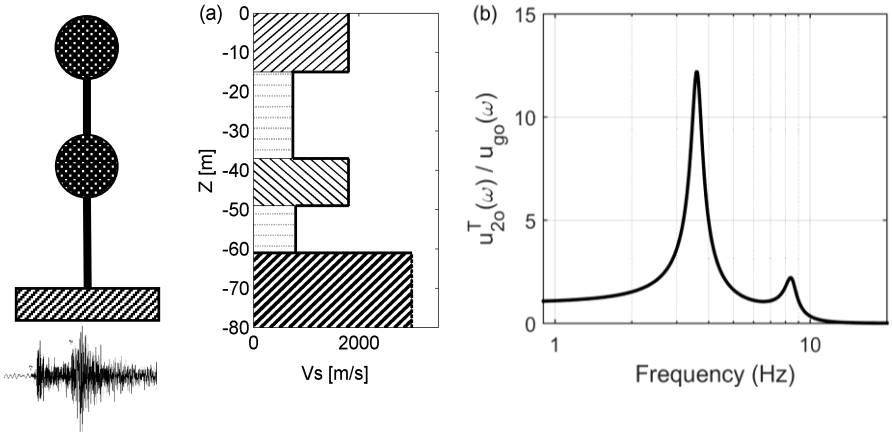

Figure 4.15 Shear-wave velocity profile obtained by modal analysis (hatched and dotted areas denote lava and sedimentary layers, respectively, with the bottom layer indicating bedrock). (b) The total displacement transfer function corresponding to the soil structure in (a), exhibiting two fundamental modes at the predominant frequencies.

These results, along with geological mapping and borehole logs, imply that the characteristics of the soil structure underneath IS605 could, at least as a first approximation,

be assumed to apply to the part of Hveragerði that lies on lava rock. However, while the HVSR results of the nearest stations IS604 and IS603 (50 and 70 m away from IS605, respectively) show all the same HVSR characteristics for both earthquake and microseismic data, they do not exhibit the clear bimodal HVSR amplification curve from earthquake data. The same can be said for station IS608 which is the closest station to the borehole locations. Other stations on lava rock show different patterns of amplification. Therefore, the results indicate that the primary characteristics of the soil structure such as layer thicknesses and the impedance contrasts of the velocity reversals, may vary considerably under Hveragerði, contributing to the observed variations in ground motion amplitudes, even over short distances. We also note that the predominant frequency of the HVSR from earthquake recordings on lava is considerably lower than for stations on older bedrock. From the results for station IS605, the lava-sedimentary layer that is responsible for this peak lies between $\sim 40 - 60$ m deep, which in turn raises questions about the applicability of the average shear-wave velocity in the uppermost 30 meters (V_{S30}) as the parameter on which to base site characterization for earthquake resistant design for this type of profile.

In order to confirm the reliability of applying dynamic system modeling for stations characterized with velocity reversals, we test the only other available numerical modeling of site response on a lava-sedimentary soil structure in Iceland. Bessason and Kaynia (2002) compared strong-motion site effects on the west and east abutments of the base-isolated and instrumented 80 m long Thjorsa-Bridge during two $M_w 6.5$ and 6.4 earthquakes in South Iceland on 17 and 21 June 2000, respectively, and a several of their aftershocks. As can be seen in Figure 4.16, the borehole information show that the west abutment of the bridge is located on the lava-rock and the east abutment sites on the bedrock (i.e., dolerite).

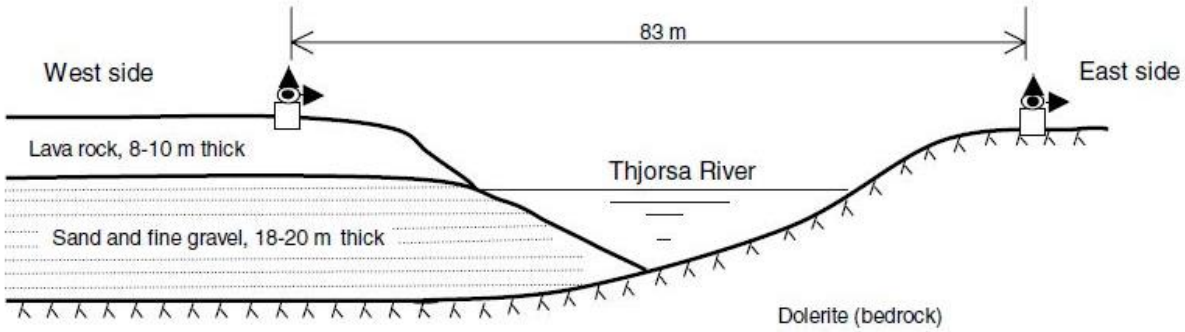

Figure 4.16 The soil profile at the Thjorsa-Bridge (Bessason and Kaynia 2002).

Numerical modeling of the site response revealed considerable and consistent site amplification on the west side relative to the east side which manifests itself in a considerable peak at around $2 - 5$ Hz in the response spectral ratio. They attribute the relative differences in amplification to the differences in site conditions on each side of the bridge. Namely, based on geotechnical surveys the soil profile on the west side of the bridge consists of an $8 - 10$ m thick lava-rock underlain by an $18 - 20$ m thick sedimentary layer of loose sand and gravel on top of on bedrock, while the east site is located directly on bedrock. After testing several parameter values (i.e., shear-wave velocity, density, and Poisson's ratio) by trial and error, they were able to produce a peak in the simulated response spectra at 3 Hz for the soil structure using the geological information presented in Figure 4.17 using the computer program SHAKE.

In the context of this study however, we model the soil structure as a single-degree-of-freedom-system (SDOF) since it consists of a single lava/sediments stack above the bedrock. The corresponding natural frequency of oscillation is

$$f_0 = \frac{1}{2\pi}\sqrt{\frac{K}{M}} \tag{4.18}$$

where K and M represent the stiffness and mass of the system, respectively. For the range of shear wave velocities assumed for the sedimentary layer we find that values of the natural (i.e., predominant) frequency of the site lies in the range of $3.2 - 4.8$ Hz, which is in excellent agreement with observations.

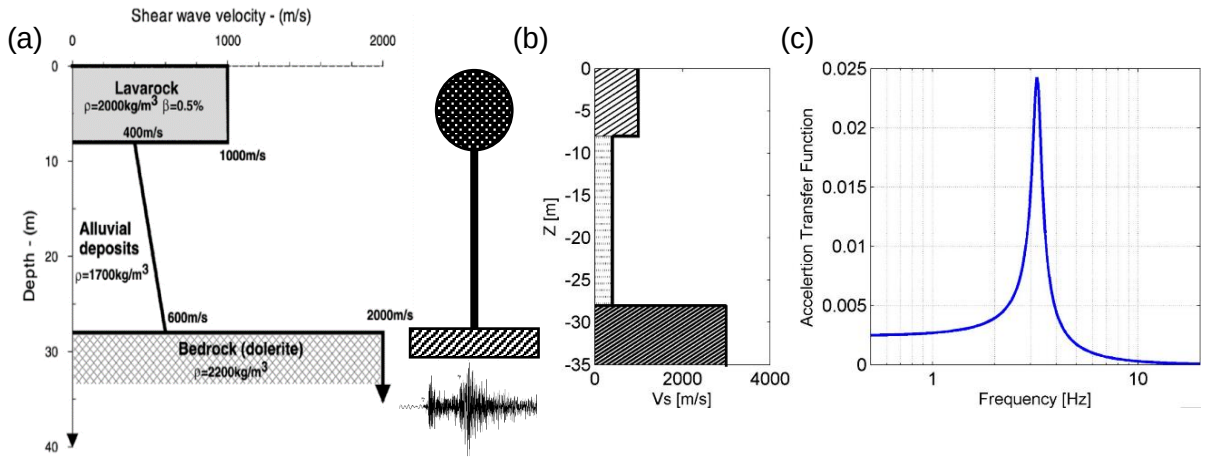

Figure 4.17 (a) soil profile used in SHAKE analyses (Bessason and Kaynia 2002). (b) Shear-wave velocity profile obtained by modal analysis (hatched and dotted areas denote lava and sedimentary layers, respectively, with the bottom layer indicating bedrock). (c) Acceleration transfer function corresponding to the soil structure in (b), exhibiting one fundamental mode at the predominant frequencies.

Thus, our results confirm that modeling soil structure composed of a lava layer on top of a sedimentary layer, therefore introducing velocity reversals, as either a SDOF system (e.g., under the west abutment of the Thjorsa brigdge) or as a 2DOF system (as in Hveragerði) is a robust way of estimating the predominant frequency(ies) of site amplification. This is especially the case when details of layer thicknesses are known, but alternatively, in the absence of such information the dynamic model may be used to infer the geologic structure underneath the site. The results of this section related to the site effect estimation across ICEARRAY I has been published in **Paper I**.

Source: Rahpeyma S., Halldorsson B., Olivera C., Green R. A., and Jónsson S. (2016) Detailed site effect estimation in the presence of strong velocity reversals within a small-aperture strong-motion array in Iceland, Soil Dynamics and Earthquake Engineering 89:136–151.

4.2 Shear-wave Velocity Profiling

4.2.1 Bayesian statistical inference

We propose a non-invasive method to estimate the shear-wave velocity model based on inversion of the HVSR spectrum of microseismic measurements. Due to incapacity of the theoretical approaches to determine the H/V spectrum where velocity reversals exist, we take advantage of the Bayesian inversion approach of the microseismic HVSR to approximation the most probable V_s profile (a set of optimal V_s models) in test stations in two different regions with high potential seismic risk. The test stations in this study are located in Izmir (BYN) in Turkey and Mirandola (MIR) in Italy ("soft soil"). We implement Bayesian approach with the aim of providing a constructive framework for making inference on different soil properties in the light of the observations. In the context of the Bayes theorem, the unknown model parameters are assumed to be random variables and assigned prior probability distribution logically defined based on available information or a priori subjective beliefs (Congdon 2014; Gelman et al. 2014). As it is shown in Eq. (4.19), the prior information about the model parameters will be updated by conditioning on the observed data with respect to the underlying probability model.

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) \propto \pi(\boldsymbol{\theta})\pi(\boldsymbol{y}|\boldsymbol{\theta}) \tag{4.19}$$

where, $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ is the joint posterior distribution of the model parameters, $\boldsymbol{\theta}$ (which represents soil properties), given the derived theoretical transfer function of the subsoil as data, $\boldsymbol{y}$, requires information about the sampling distribution $\pi(\boldsymbol{y}|\boldsymbol{\theta})$ and also a sensible assumption about the prior distribution $\pi(\boldsymbol{\theta})$ if it exists. In this regard, the obtained posterior distribution integrates updated knowledge on model parameters considering knowledge found from the observed data. In order to numerically approximate posterior density function of model parameters, a Markov Chain Monte Carlo (MCMC) algorithm, is employed (Gelman and Rubin 1992; Smith and Roberts 1993; Gilks 2005). MCMC is basically applicable to almost any Bayesian modeling and is a general algorithm for simulating independent Markov chains which has a desired target density. This procedure is mainly carried out using the Gibbs sampling framework (Geman and Geman 1984; Casella and George 1992) and the Metropolis algorithm (Metropolis et al. 1953) as an updating strategy which tracks adaptation of a random walk in parameters space to define the acceptance or rejection of the samples to converge to the specified target distribution (cf. section 2.5.3).

Bayesian convergence diagnostics

In this study, we use three different convergence diagnostics to assess the convergence of multiple MCMC chains. First, visual inspection can expose bad mixing of Markov chains or chaotic behavior of separate chains. Secondly, the Gelman-Rubin statistics (Gelman and Rubin 1992) by relying on the within-chain variance to the between-chain variance tests whether the chains all converge to the same posterior distribution. Large values (> 1.00) of Gelman-Rubin test indicates that simulated chains have not converged to the target density. Finally, the autocorrelation plots evaluate the exist dependency between successive samples within each Markov chain.

Inversion strategy

We implement the theoretical HVSR obtained based on the transfer functions of a set of horizontally stratified, linearly elastic layers overlaying half-space excited by vertically incident proposed by Tsai (1970) (see sections 2.3.1 and 4.1.6 for more details). The parametrization of the theoretical HVSR is based on the assumptions of 1-D layered models consisting of a stack of homogenous linear elastic layers over a half-space. Hence, subsoil physical properties such as thickness (H), density (ρ), shear-wave velocity (V_s), compressional velocity (V_p), and elastic properties for S- and P-waves (Q_s and Q_p) are considered as model parameters.

Trial inversion using all model parameters reveled that due to the non-uniqueness results and large trade-off between parameters the model cannot converge reliably. We observed that thickness and S-wave velocity are the most influential and correlated variables. Therefore, we set model parameter $\boldsymbol{\theta} = (H, V_s)$ and fix the rest of parameters with the aim at better convergence. It has been also proven that the theoretical transfer function chiefly depends on S-wave velocity and depth of the subsoil and negligibly on the other soil properties (Foti et al. 2009; Molnar et al. 2010). Therefore, in this study, $\boldsymbol{\theta} = (H, V_s)$ are assumed to be unknown and the rest of parameters are defined as fixed parameters and their values can be approximated on the basis of available geological information.

At each iteration k of the MCMC process, the unknown variables are drawn, as input for theoretical HVSR, from a normal distribution centered at an adaptive mean (i.e., the latest accepted value) value and pre-defined standard deviation for each layer over all chains as can be seen in Eq. (4.20):

$$\theta_{p,l}^{k} \sim \mathrm{N}\left(\theta_{p,l}^{k-1},\ \sigma_{\theta_{p,l}}^{2}\right) \tag{4.20}$$

where subscripts $p = 1,2$ and $l = 1, \dots, L$ indicate the model parameters indicator $\boldsymbol{\theta} = (H, V_s)$ and the layer, respectively. The standard deviation, $\sigma_{\theta_{p,l}}$, is defined based on 5% of the mean value of the model parameters. For each model parameter at each layer, the lower ($\theta_{p,L}$) and upper ($\theta_{p,U}$) bounds are chosen reasonably to avoid the inversion stick into a wrong convergence track due to the trade-off between model parameters; however, the boundaries should be wide enough to allow the data to determine the S-wave velocity profile parameters ($\theta_{p,L} \leq \theta_{p,l} < \theta_{p,U}$). The initial values of the model parameters, which are used to produce the initial theoretical HVSR, are instinctive approximations.

As can be seen in Eq. (4.19), the Bayesian framework requires specifying prior probabilities for all model parameters and a likelihood function. Due to the lack of precise information about the subsoil structure, the prior probability density function of each parameter is chosen as a uniform probability density functions on the interval $\theta_{p,L} \leq \theta_{p,l} \leq \theta_{p,U}$, such that

$$\pi\left(\theta_{p,l}\right) = \begin{cases} 1 & , \quad \theta_{p,L} \leq \theta_{p,l} \leq \theta_{p,U} \\ 0 & , \quad otherwise \end{cases} \tag{4.21}$$

And the join prior probability density function for all the model parameters in $\boldsymbol{\theta}$ is the product of the individual prior densities. It is assumed that the probability density function of the spectral amplitudes $HVSR_i$ (i.e., y_i) in each frequency bin f_i with $i = 1, \dots, I$ (I is the total number of frequency bins) is lognormal with parameters $\mu_{HV}(f_i)$ and $\sigma_{HV}^2(f_i)$, i.e., the

expected value and the variance of $\log(HVSR_i^{obs})$. So, the probability density function for each y_i within frequency bin i is given by Eq. (4.22):

$$\pi(y_i|\theta_i) \sim \mathrm{LN}\left(y_i \mid \mu_{HV}(f_i), \sigma_{HV}^2(f_i)\right) \tag{4.22}$$

A practical issue influencing convergence to an unbiased estimate include deleting early samples of the Markov chain, commonly referred to as "burn-in" (a burn-in length of at least 25% of total samples is applied here). We run many sets of combination with different number of layers, various prior assumptions and initial values, number of iterations, number of chains, and burn-in sample size to find the most consistence results. All chains would be analyzed together after simulating the desired number of iterations by removing burn-in part.

4.2.2 Application to IzmirNet, Turkey

Geological setting

In 2008, a small aperture local seismic network, IzmirNet, consisting of 16 stations (see Figure 4.18) was established across the Izmir Bay (Polat et al. 2009). Historically, this region has a prominent seismic risk due to its large and growing population and key infrastructure which are surrounded by active faults. Figure 4.18 shows the distribution of IzmirNet stations across the Izmir Gulf (Polat et al. 2009; Gok and Polat 2014). As can be seen in Figure 4.18, the majority of the settlements (industrial and populated areas) are collocated on top of Quaternary alluvial deposits around the Gulf of Izmir. The unconsolidated deposits in the Izmir basin can significantly change the propagation of ground motions to the surface; hence, the assessment of seismic hazard for the Izmir region is an imperative issue.

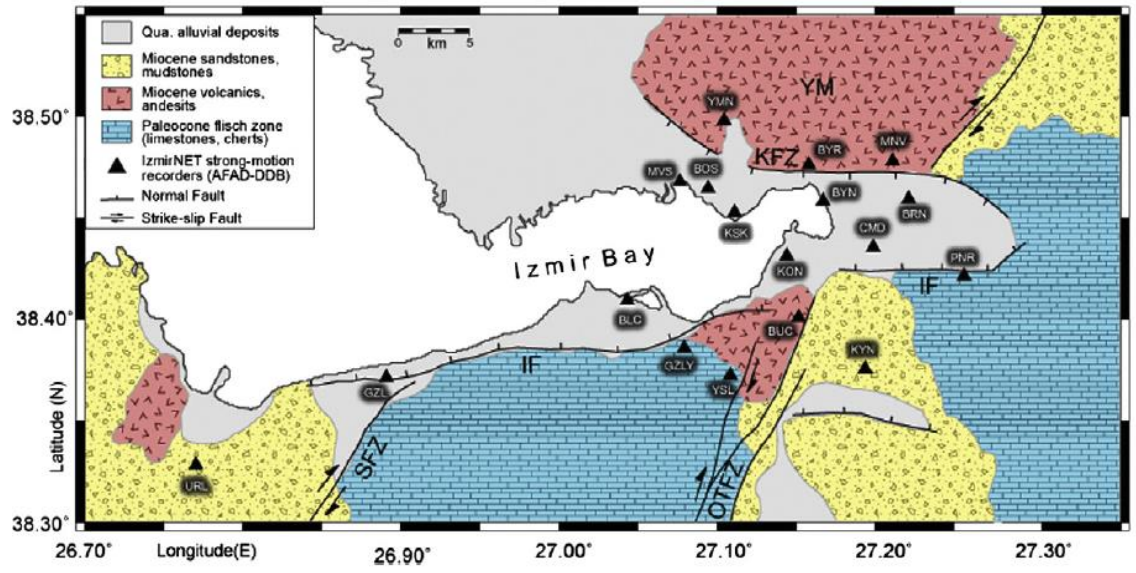

Figure 4.18 Location of IzmirNet array (filled triangles) on geology of Izmir and simplified geological features.

Results

The microseismic HVSR results for IzmirNet stations are presented in Figure 4.19. In this study, we selected station BYN located on soft soil in the eastern part of Izmir Bay with a clear fundamental frequency peak at $0.7 - 0.8$. The HVSR characteristic for BYN station

suggests that the soil column acts as a single layer on top of a high impedance contrast between layers, where strong amplification and frequency dependent resonance are known to occur.

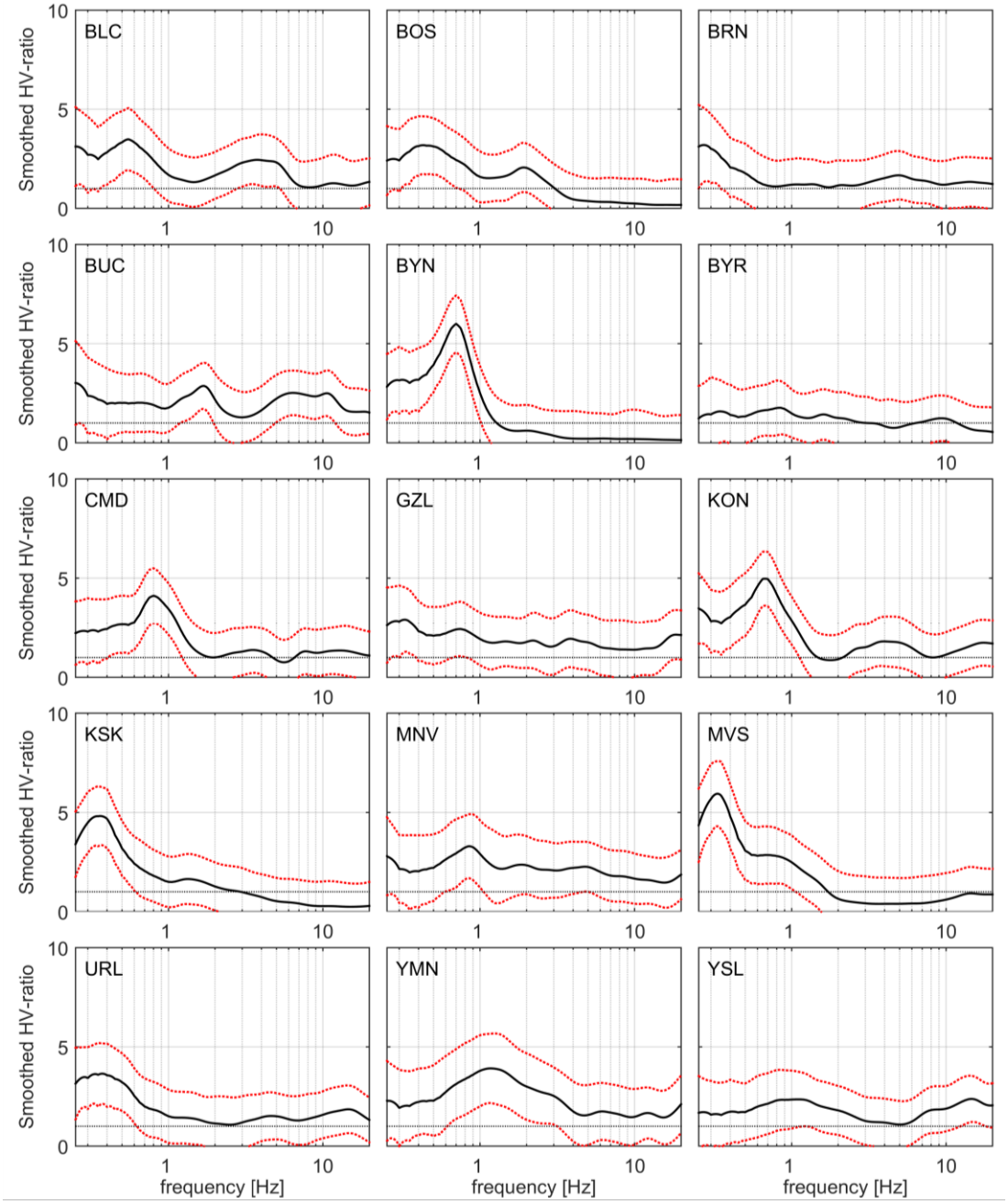

Figure 4.19 Mean microseismic HVSR $\pm 1\sigma$ for IzmirNet stations in Izmir, Turkey.

A grid search of the MCMC initiated with a starting model whose parameters are randomly perturbed within the bounds defined $[\theta_L, \theta_U]$ results in posterior probability distribution of the model parameters. Figure 4.20 shows the trace plots of the model parameters for a single layer soil profile overlaying the bed rock. Figure 4.20 represents the visual inspection of the successive MCMC samples of the model parameters. As can be seen in the Figure 4.20, trace plots show good mixing in successive samples to estimate the model parameters. The posterior histograms of posterior samples after burn-in period represent unimodal posterior distribution for thickness and S-wave velocity.

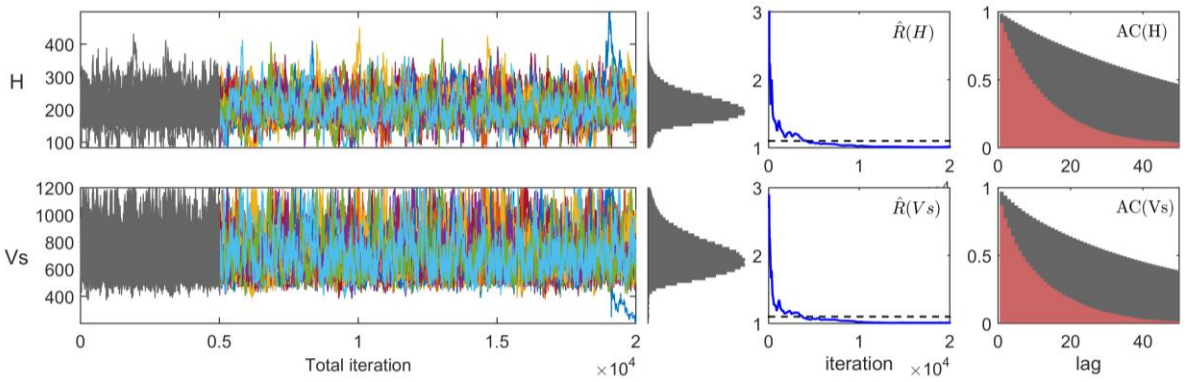

Figure 4.20 Convergence diagnostics: trace plots of all sampling chains ($N_c = 20$) with total iteration ($N_T = 20{,}000$) and considering burn-in period ($N_B = 5000$) in gray based on the MCMC simulations; model parameters posterior histogram; Gelman-Rubin plots; autocorrelation plots based on the MCMC simulations until lag 50 for a single-layer soil structure (gray bars), autocorrelation function after thinning of scale 5 (red bars).

Figure 4.21(a) shows the observed and the initial theoretical HVSR model obtained based on the initial assumption of V_s profile. It should be noted that we invert the observed HVSR over the nominated range of frequencies defined around the fundamental frequency (grey area in Figure 4.21(a)). It may be argued that at relatively high frequency, no HVSR peak associated to a shallow stratigraphic horizon can be observed; thus, the basic requirement for the proposed procedure to concentrate around the fundamental mode could be satisfied. Furthermore, in the seismic microzonation practice, attention has generally only been paid to the main resonance frequency, which is the largest HVSR peak, while other stable humps and troughs in the curve were not considered. As can be seen in Figure 4.21(a), the initial model is considered to be different from the observed HVSR. Furthermore, we assumed a wide range of model parameters space with small jumping steps (0.05 times of model parameters values at each iteration) which let the synthetic models converge to the highest probability ratio obtained by sampled model parameters. According to the available geological data, BYN station sits directly on top of an alluvial deposit layer of around $180 - 200$ m depth. It is explicit in Figure 4.21(b) that for a single-layer model the mean posterior of sedimentary layer thickness over the half-space is estimated around 200 m with ~ 50 m of standard deviation that is in very good agreement with available information.

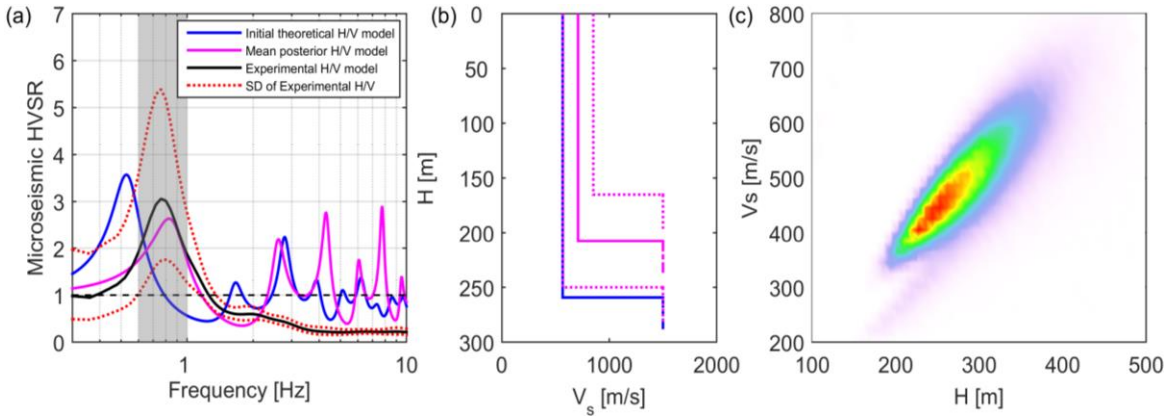

Figure 4.21 (a) The observed (black) and standard deviation (dotted red) HVSR of microseismic recordings, initial (blue), and determined theoretical HVSR from Bayesian MCMC inversion (magenta) for a single-layer subsoil structure at BYN station. (b) initial and final S-wave velocity model (c) joint marginal probability distribution from inversion of the microseismic data.

We found that increasing the number of parallel chains can reasonably lead to better convergence due to producing more combination of the input parameters and generating more synthetic HVSR curves. In this regard, within the MCMC grid search all perturbations start around different initial parameters values. Hence, the determined V_s profiles with this approach and the associated uncertainties would be reliable. Although with a single-layer subsoil structure we could quantitatively estimate model parameters' characteristics, the key question is that to what extent a detailed V_s profile can be extracted from the recorded data? It is noteworthy that due to complexity of the subsoil structure, non-linearity of the model, and non-uniqueness solution for the inversion process, it is likely that a single-layer model cannot precisely capture the subsoil structure. In soil properties inversion problems, defining enough parameters (e.g. number of layers, fixed or dynamic model parameters) is essentially critical to estimate a proper resolution for layering and parametrization.

We conduct a blind test over the number of layers to consistently investigate the best resolution of model parametrization. Therefore, we continue adding the number of layers as far as the posterior distributions of model parameters do not provide any consistent information. The effect of increasing the number of layers is shown in Figure 4.22 the marginal posterior probability distributions of a five-layer model over bedrock parameters for H and V_s by illustrating.

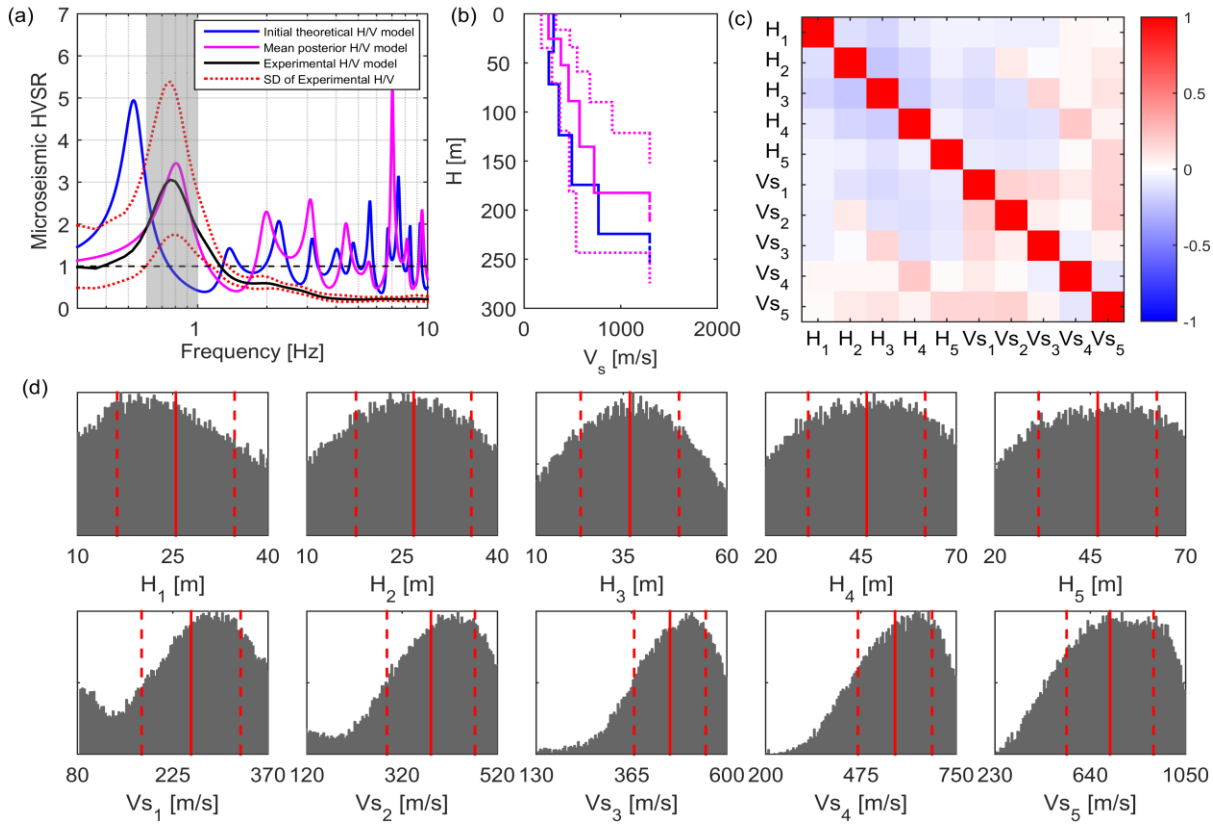

Figure 4.22 (a) The observed, initial, and posterior theoretical HVSR model obtained from Bayesian MCMC inversion for a 5-layer subsoil structure at BYN station, Turkey; (b) S-wave velocity prrofile for the initial (blue) and posterior (magenta) model; (c) Correlation matrix of posterior samples of model parameters; (d) posterior histograms for thickness (H) and S-wave velocity (Vs).

As can be seen in Figure 4.22(a-b), although the initial V_s profile and HVSR model behaves differently, the posterior model fits the observed HVSR very well and the estimated depth

of the sedimentary layer over the half-space is around $180-190$ meter. Figure 4.22(c) illustrates the correlation matrix and evaluate the correlation between posterior model parameters. Contrary to the relatively high correlation for a single-layer subsoil structure (cf. $\sim 80\%$ in Figure 4.22(c)), the correlation between model parameters of a multi-level subsoil structure is not large ($\sim \pm 30-40\%$). The negative and positive correlation can be observed between thicknesses and S-wave velocities of layers. The posterior histograms of thickness and S-wave velocity for each layer are shown in Figure 4.22(d). As can be seen the obtained posterior distributions of specifically thickness are weakly informative with large uncertainties for deeper layers comparing to the shallower layers. In order to confirm the efficiency of the inversion results is finding the predominant frequency from posterior simulations of model parameters we calculate the natural frequency of the soil, f_n, using the harmonic average defined in Eq. (4.23):

$$f_n = {(2n-1)} \Big/ {4\sum_{i=1}^{L}\left(\frac{H_i}{V_{s_i}}\right)} \tag{4.23}$$

where n is the mode number, H_i is the thickness and V_{s_i} is the shear-wave velocity for the i^{th} soil layer, and L refers to the number of layers overlaying the half-space. The estimated mean posterior predominant frequency (f_0) for a five-layer subsoil structure over half-space estimated in Figure 3 for station BYN is 0.63 Hz with the posterior 95% interval of $0.58-0.72$ Hz which is in approximately good agreement with the experimental fundamental frequency of 0.70 Hz.

4.2.3 Application to Mirandola, Italy

Geological setting

Mirandola city sits on the Po river plain in North Italy. As can be seen in Figure 4.23, the subsoil structure of the region is mainly characterized by layers of alluvial deposits with sandy horizons and silty-clayey layers overlaying a stiff rock layer of marine and transitional deposits of lower-middle Pleistocene age (Garofalo et al. 2016a, b; Cox and Teague 2016; Tarabusi and Caputo 2017). As can be seen in Figure 4.23 The city is located near the epicenter of the recent Emilia seismic sequence in 2012 (Anzidei et al. 2012). The invasive tests indicate that V_s profile is rather plain and smooth with no abrupt velocity contrasts until the top soft sedimentary layer is reached the stiff soil at a depth between 110 and 120 m (Cox and Teague 2016).

The continuous microseismic recordings of a minimum 60-minute duration were collected at Mirandola station (MIR). The signals at three components were sampled at 200 Hz sampling frequency with continues GPS synchronization. The same data processing routine as applied for BYN station was implemented for MIR station. MIR station is located on flat ground and it is rather quiet and at the boarder of a residential area, without noisy facilities. As can be seen in the insert figure on top right of Figure 4.23, there is a clear and sharp peak at fundamental frequency at $\sim 0.7-0.8$ Hz which indicates a strong velocity contrast within depth.

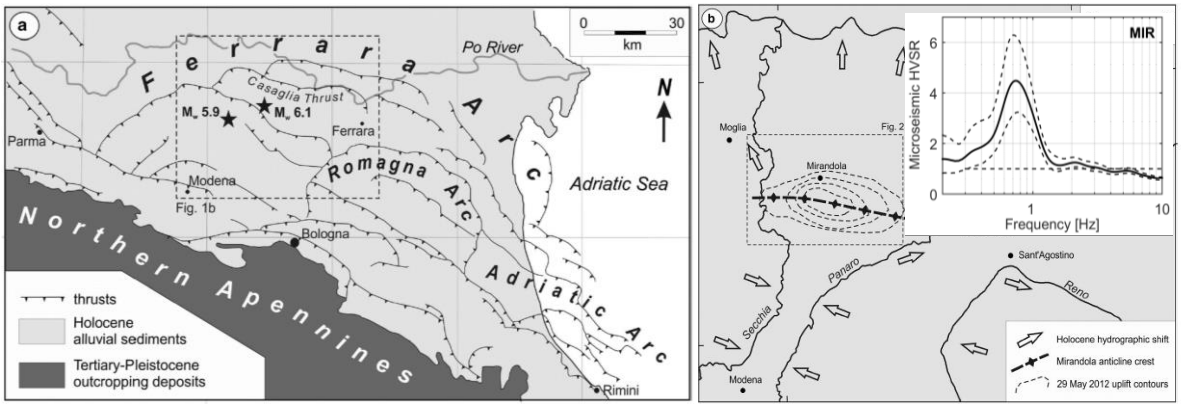

Figure 4.23 Geological and tectonic sketch map of the buried Northern Apennines fold-and-thrust belt. Stars represent the epicenters of 20 May (M_w6.1) and 9 May (M_w5.9) 2012 earthquakes (Tarabusi and Caputo 2017). The insert figure on right top shows the observed HVSR form microseismic measurements with a Konno and Ohmachi smoothing coefficient $B = 20$ for MIR station.

Results

The same inversion framework was applied for the Mirandola station. The sharp peak of the HVSR curve at MIR station can be modeled by a column of soft soil sits over the hard rock. We run minimum $N_C = 30$ chains of a total length of $N_T = 30{,}000$ iterations by considering $N_B = 5{,}000$ samples as burn-in period for different model sets. The initial results from a single-layer soil structure in addition to the convergence diagnostics results for MIR station is shown in Figure 4.24.

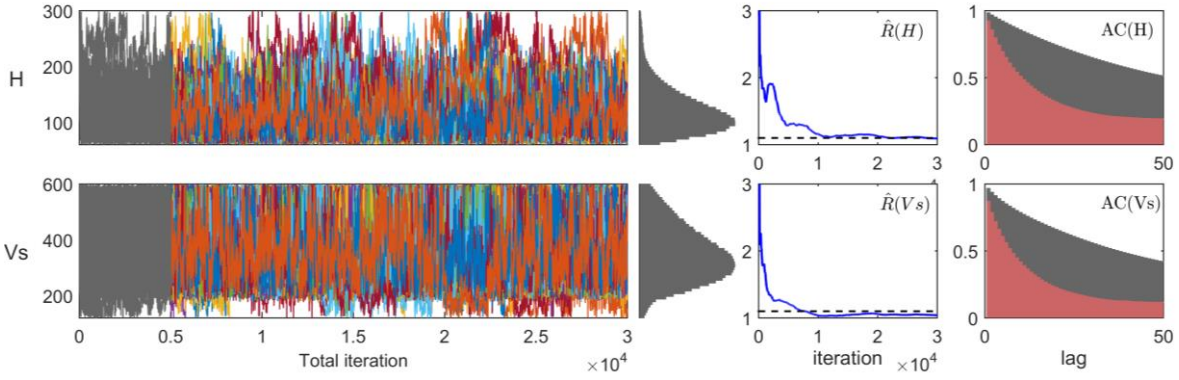

Figure 4.24 Convergence diagnostics: trace plots of all sampling chains ($N_C = 30$) with total iteration ($N_T = 30{,}000$) and considering burn-in period ($N_B = 5000$) in gray based on the MCMC simulations; model parameters posterior histogram; Gelman-Rubin plots, the blue line denotes the median of Gelman-Rubin statistics as a function of iterations; autocorrelation plots based on the MCMC simulations until lag 50 for a single-layer soil structure (gray bars), autocorrelation function after thinning of scale 5 (red bars).

Figure 4.25(a-b) represent the initial and the obtained HVSR and V_s profile using mean posterior of model parameters. Figure 4.25(c) highlights the strong trade-off between model parameters using marginal probability distribution of the model parameters. We gradually increase the number of layers as far as the posterior distributions of model parameters do not provide any consistent information for the layers. Figure 4.26 shows a three layers subsoil structure for MIR station.

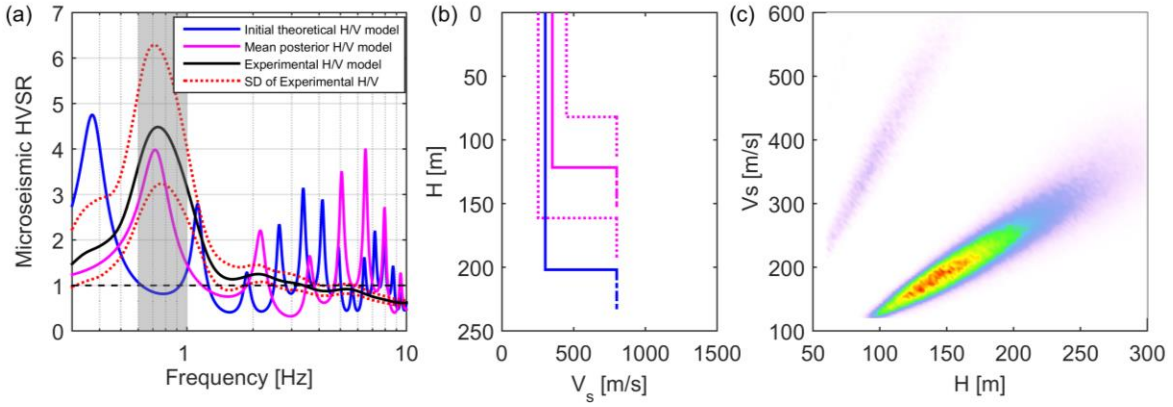

Figure 4.25. (a) The observed (black) and standard deviation (dotted red) HVSR of microseismic recordings, initial (blue), and determined theoretical HVSR from Bayesian MCMC inversion (magenta) for a single-layer subsoil structure at BYN station. (b) initial and final S-wave velocity model (c) joint marginal probability distribution from inversion of the microseismic data.

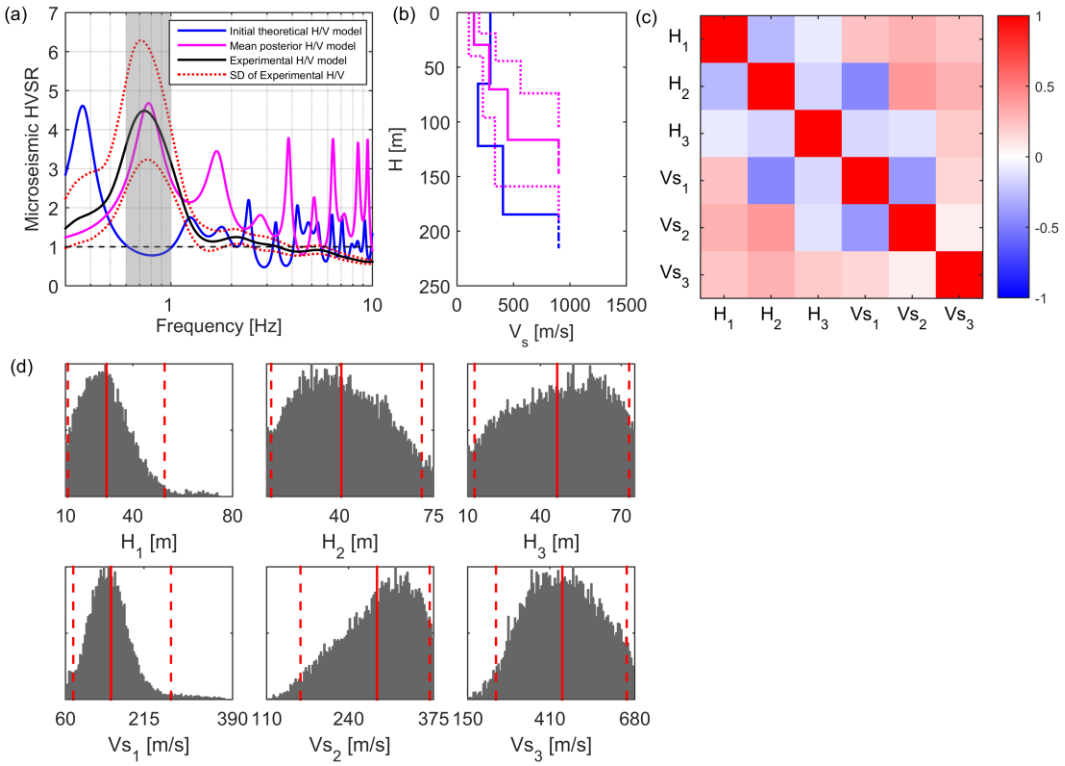

Figure 4.26 (a) The observed, initial, and posterior HVSR obtained from Bayesian MCMC inversion for a 3-layer subsoil structure at MIR station, Italy; (b) S-wave velocity profile for the initial (blue) and posterior (magenta) model; (c) Correlation matrix of posterior samples of model parameters; (d) posterior histograms for thickness (H) and S-wave velocity (V_S). The red thick lines and red dashed lines show posterior mean and posterior 95% percentile.

The results of this section have been presented in ***Paper IV***.

Source: Rahpeyma S., Halldorsson B., Hrafnkelsson B., Green R. A., Polat O., and Jónsson S. (2018b). Estimate of shear-wave velocity profile using microseismic Horizontal-to-Vertical Spectral Ratios inversion. Soil Dynamics and Earthquake Engineering, (manuscript completed)

4.3 Spatial Model of Ground Motion Amplitudes

Due to uniform site condition across ICEARRAY I insignificant variability in ground motion amplitudes was expected; however, As can be seen in considerable variations in PGA and peak-ground velocity (PGV) were observed during the recorded ground motions of the main-shock and aftershocks (Halldorsson and Sigbjörnsson 2009; Douglas and Halldorsson 2010; Rahpeyma et al. 2016). It is noteworthy that no azimuthal dependency was captured for ICEARRAY I recordings (Rahpeyma et al. 2016). Rationally, this variation is more highlighted across ICEARRAY II as a result of the variability of subsoil structure. It is well acknowledged that localized geological structure can significantly vary within short distances and influence site responses (Bessason and Kaynia 2002; Di Giacomo et al. 2005; Rahpeyma et al. 2016). Consequently, quantifying the origins of strong-motion variability over the relatively small area can have important effects on structural applications.

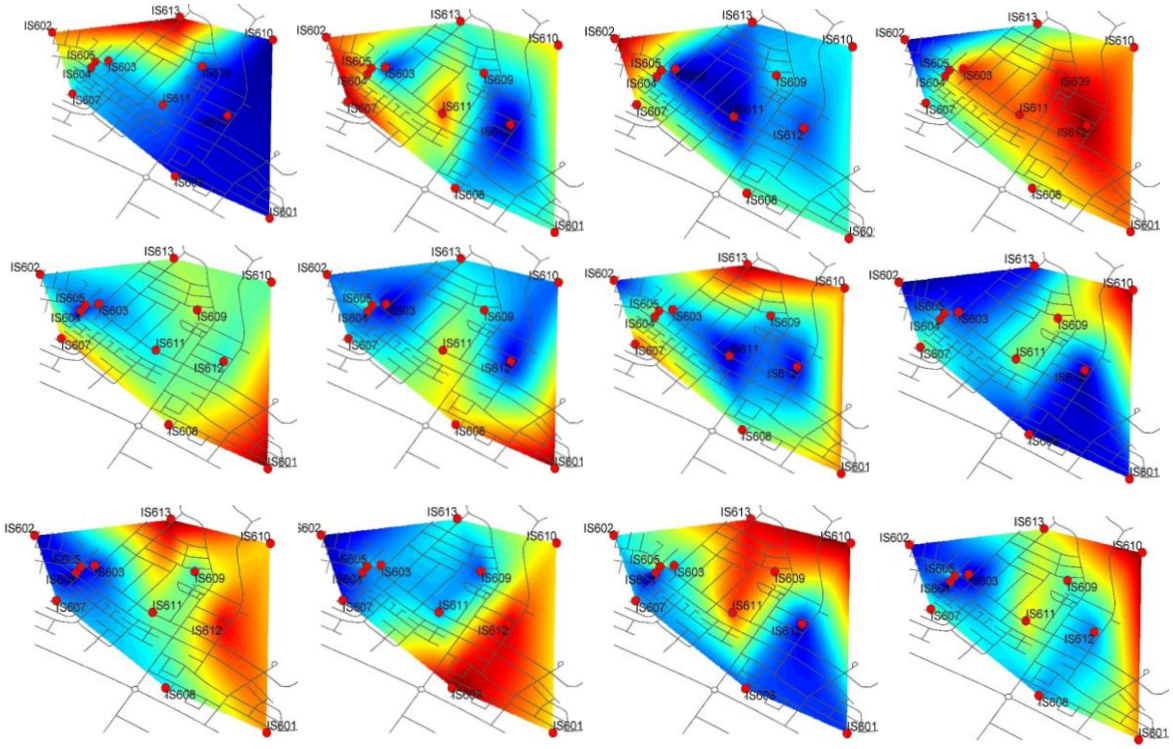

Figure 4.27 Spatial distribution of the normalized PGA across ICEARRAY I using aftershocks from 29 May 2008 M_w6.3 Ölfus earthquake.

In this section a spatial distribution model for peak ground accelerations recorded by ICEARRAY I and II strong-motion stations will be presented with the aim of decomposing source, propagation path, and localized site effects. The proposed method is on the basis of the Bayesian statistical framework to provide better understanding of distribution of strong-motion amplitudes as well as quantifying the corresponding variabilities. The proposed model, the so-called Bayesian Hierarchical Model (BHM), offers a flexible probabilistic framework for multi-level modeling of earthquake ground motion parameters. A BHM usually consists of three levels, referred to as the data level, the latent level, and the hyper-level. The first level (data), describes the distributional model for the observation conditioned on the model parameters. The latent level describes the distribution of the model variables which are conditioned on the parameters in the third level (hyper), the so-called

hyperparameters (i.e. mean and standard deviation of model parameters distribution). The third level consists of prior distributions for the model parameters at the data and latent levels.

The observed $\log_{10} PGA_{es}$ is modelled with a normal distribution at the data level of the BHM. Let Y_{es} represent the base 10 logarithm of PGA for event e at station s. The proposed model, consists of independent terms, can be presented as Eq. 4.25:

$$Y_{es} = \mu(M_e, R_{es}, D_e, \boldsymbol{\beta}) + \delta B_e + \delta S_s + \delta WS_{es} + \delta R_{es}$$
$$e = 1, \dots, N\,, \quad s = 1, \dots, Q \tag{4.24}$$

where $\mu_{es} = \mu(M_e, R_{es}, D_e, \boldsymbol{\beta})$ is the predicted median ground motion that is a function of independent variables (magnitude, hypocentral distance and depth of the origin) for event e and station s, δB_e is the event effect which denotes the over-all effect of event e in addition to the predicted median ground motion μ_{es}, δS_s is the station effect for station s, δWS_{es} is a spatially correlated event-station effect, and δR_{es} is an independent error term representing unmodeled effects or other factors that are not accounted for. The terms δWS_{es} and δR_{es} are assumed to follow Gaussian distributions, thus, conditioned on μ_{es}, δB_e and δS_s then Y_{es} also follows a Gaussian distribution.

The term μ_{es} is a commonly-used linear predictive function for median ground motion. Although there are different functional forms, in this study, we nominate the most commonly-used and less complicated model. The nominated predictive model links PGA to the local magnitude of the e-th earthquake, M_e, the hypocentral distance from e-th event to s-th station, R_{es}, and the depth of the origin of the e-th earthquake, D_e:

$$\mu_{es} = \beta_1 + \beta_2 M_e + \beta_3 \log_{10}(R_{es}) + \beta_4 D_e \tag{4.25}$$

The coefficients $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)$ correspond to the particular area and volume under study in which the seismic waves travel on their way to the surface, reflecting peculiarities of the local seismic regime and geological structure. The station terms, δS_s, are modeled as a mean zero Gaussian distribution with an exponential form of covariance function from the Matérn family with marginal variance ϕ_{S2S}^2 which describes the inter-station variability and Δ_{S2S} as the range parameter (read Paper III for more details). The event-station terms, δWS_{es}, defined as spatially correlated variables from a zero mean Gaussian field governed by a covariance function from the Matérn class with marginal covariance ϕ_{SS}^2 and range parameter Δ_{SS}.

The total variance of Y_{es}, when conditioning on M_e, R_{es}, and D_e, can be calculated as the sum of the inter-event variance τ^2, the inter-station variance ϕ_{S2S}^2, the event-station variance ϕ_{SS}^2, and the variance of the unmodeled effects and other unaccounted factors, ϕ_R:

$$\sigma_T = \sqrt{\tau^2 + \phi_{S2S}^2 + \phi_{SS}^2 + \phi_R^2} \tag{4.26}$$

The inter-event variance (τ^2) quantifies the variations between events after taking the effects of M_e, R_{es}, and D_e of each event into account. The inter-station variance (ϕ_{S2S}^2) quantifies the variability in the station effects that stems from the varied local geological conditions. In

other words, the inter-station variance represents the systematic deviation of the observed amplification at the specific station from the median amplification predicted by the model. The event-station variance (ϕ_{SS}^2), on the other hand, is a measure of the spatial variability in PGA between stations within events after taking into account the overall effect of the event and the average effect of each station. Finally, the variance parameter ϕ_R^2 quantifies jointly the variability in the unmodeled effects that are not accounted for with other terms.

In the proposed BHM, the vectors of $\vec{\boldsymbol{\eta}} = (\vec{\boldsymbol{\beta}}, \vec{\boldsymbol{v}}, \vec{\boldsymbol{\alpha}})$ and $\vec{\boldsymbol{\theta}} = (\sigma_\varphi, \sigma_v, \sigma_\varepsilon, \phi_\varepsilon, \sigma_\alpha)$ were defined as latent parameters in the second level and hyperparameters in the third level, respectively. In order to infer $\boldsymbol{\eta}$ and $\boldsymbol{\theta}$ we used a Gibbs sampler (Geman and Geman 1984; Casella and George 1992), an iterative Markov Chain Monte Carlo (MCMC) sampling algorithm, with Metropolis step in each iteration (Metropolis et al. 1953; Hastings 1970). We analyzed 610 earthquake events recorded in 10 ICEARRAY I strong-motion stations. Likewise, for ICEARRAY II we used 14 and 11 events from the October 2012 and April 2013 sequences, respectively, recorded on 6 strong-motion stations. To eliminate the path effects from ICEARRAY II recordings, we split the data set into two groups (i.e. October 2012 and April 2013) and set up the BHM separately for each data set.

Results

We used four parallel Markov chains of total length of 8,000 samples and considering the first 25% of iterations (2,000) as burn-in period when analyzing the ICEARRAY I data. In the case of ICEARRAY II however, more samples were required to get the convergence due to lack of data points compared with the ICEARRAY I dataset. Thus, all 4 chains used for the BHM proposed for ICEARRAY II were sampled with a total length of 20,000 iterations where the first 25% iterations (5,000) were used as burn-in period. The accuracy of the obtained model was diagnosed both visually in addition to the Gelman-Rubin Statistics and autocorrelation function (Gelman and Rubin 1992).

Hyperparameters inference

The posterior mean, standard deviation and 95% percentiles of the hyperparameters for ICEARRAY I and II are presented in Table 4.4. A key point to evaluate how well the model fits to the data is the variation of the parameters in addition to its posterior distribution. A close scrutiny on the posterior distribution of the hyperparameters in Figure 4.28 indicates that the ICEARRAY I results are well obtained with low variation and narrow with distribution in comparison to the ICEARRAY II results. Undoubtedly the extensive dataset recorded on the ICEARRAY I allow a robust and more reliable estimation of the hyperparameters of the BHM. Nevertheless, we cannot exclude also contributing effects of greater variability of local geology and topography of the ICEARRAY II.

Table 4.4 Posterior mean, standard deviation and 95 percentiles of $\boldsymbol{\theta}$ for ICEARRAY I and II

θ	ICEARRAY I			ICEARRAY II (swarm 2012)			ICEARRAY II (swarm 2013)		
	mean	SD	(2.5, 97.5) %	mean	SD	(2.5, 97.5) %	mean	SD	(2.5, 97.5) %
τ	0.198	0.007	(0.185, 0.211)	0.222	0.059	(0.129, 0.359)	0.149	0.065	(0.043, 0.306)
ϕ_{S2S}	0.092	0.024	(0.057, 0.151)	0.197	0.073	(0.099, 0.383)	0.218	0.082	(0.112, 0.428)
ϕ_{SS}	0.116	0.003	(0.114, 0.123)	0.052	0.032	(0.009, 0.135)	0.098	0.035	(0.032, 0.179)
ϕ_R	0.058	0.002	(0.056, 0.061)	0.053	0.007	(0.040, 0.065)	0.065	0.019	(0.023, 0.098)
Δ_{SS}	0.282	0.018	(0.250, 0.319)	1.851	1.216	(0.180, 4.767)	1.062	0.797	(0.204, 3.026)

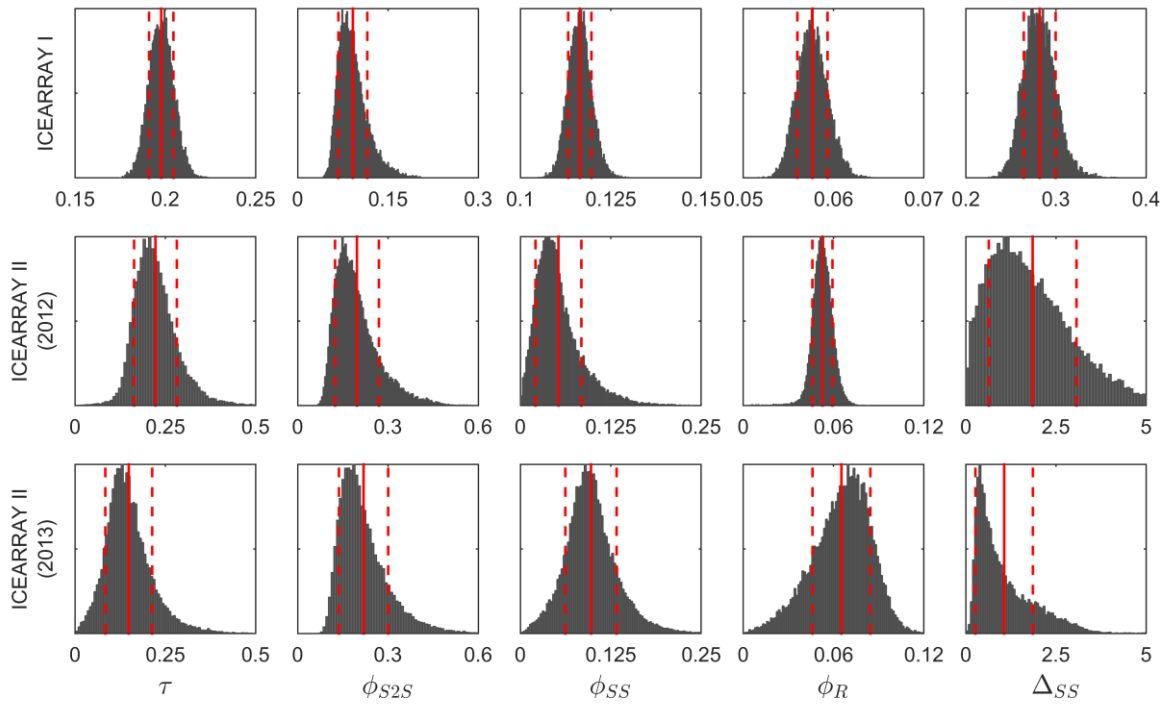

Figure 4.28 Posterior histograms of hyperparameters, $\boldsymbol{\theta}$, for ICEARRAY I (1st row) and ICEARRAY II (2nd and 3rd rows for swarm 2012 and swarm 2013, respectively). The red solid and dashed lines show posterior mean and posterior mean ± 1 posterior standard deviation, respectively.

Latent parameters inference

As can be seen in Table 4.5 the inference of βs, i.e., the constant coefficients of nominated GMM (Eq. (4.25)) in the BHM formulation, resulted in well-simulated posterior samples for ICEARRAY I. However, the obtained results for ICEARRAY II datasets, indicate extremely large standard deviations, indicating unreliable results, most likely due to the lack of data.

Table 4.5 Posterior mean, standard deviation and 95 percentiles of $\boldsymbol{\beta}$ for ICEARRAY I and II

β	ICEARRAY I			ICEARRAY II (swarm 2012)			ICEARRAY II (swarm 2013)		
	mean	SD	(2.5, 97.5) %	mean	SD	(2.5, 97.5) %	mean	SD	(2.5, 97.5) %
β_1	0.88	0.05	(0.79, 0.98)	-10.89	8.59	(-27.43, 6.10)	-1.25	1.57	(-4.32, 1.93)
β_2	0.71	0.02	(0.68, 0.74)	0.96	0.16	(0.63, 1.28)	0.47	0.08	(0.31, 0.63)
β_3	-2.86	0.08	(-3.01, -2.72)	3.61	4.57	(-5.43, 12.45)	-0.34	0.91	(-2.18, 1.46)
β_4	0.09	0.01	(0.08, 0.11)	0.02	0.03	(-0.05, 0.08)	0.00	0.01	(-0.02, 0.03)

Despite the high correlation between station terms, the posterior density of the inter-station variance ϕ_{S2S}^2 is indeed different from zero, confirming that the amplitude of the spatial field $\delta S(\cdot)$ is different from zero. This strongly indicates that site effects influence the PGA parameter. On closer scrutiny, the station terms for stations IS601, IS604, and IS611 are significantly affecting (either positively i.e., amplifying, or negatively i.e., de-amplifying, relative to the mean) the earthquake strong-motion intensities (see Table 4.6), based on the 95% marginal posterior intervals for the individual δS's across ICEARRAY I. In particular, the impact of station term value for station IS601 (95% confidence interval is [0.07; 0.20]) is estimated to intensify the earthquake ground-motions, whereas δS values for stations

IS604 (95% confidence interval is [-0.15; -0.02]), and IS611 (95% posterior interval is [-0.15; -0.02]) are expected to diminish the seismic ground motions. On the other hand, the rest of the station terms are not statistically significant from one another. For instance, a 95% confidence interval for station IS609 is [-0.05; 0.08] with an amplifying mean value of 0.01.

Table 4.6 Posterior mean, standard deviation and 95 percentiles of $\boldsymbol{\delta S}$ *for ICEARRAY I*

δS	ICEARRAY I			
	Station	mean	SD	(2.5, 97.5) %
δS_1	IS601	0.14	0.03	(0.07, 0.20)
δS_2	IS602	-0.04	0.03	(-0.11, 0.02)
δS_3	IS603	-0.06	0.03	(-0.13, 0.00)
δS_4	IS604	-0.08	0.03	(-0.15, -0.02)
δS_5	IS605	0.06	0.03	(-0.01, 0.12)
δS_6	IS607	0.05	0.03	(-0.01, 0.11)
δS_7	IS608	0.04	0.03	(-0.03, 0.10)
δS_8	IS609	0.01	0.03	(-0.05, 0.08)
δS_9	IS611	-0.08	0.03	(-0.15, -0.02)
δS_{10}	IS612	-0.04	0.03	(-0.01, 0.02)

As can be seen in Table 4.7, the same statistical analysis on ICEARRAY II's station terms with regards to their contribution to overall amplification or de-amplification, relative to the mean PGA across the array, found no significance (i.e., negative 2.5% percentile and positive 97.5% percentile), primarily due to the large standard deviations of the station term posterior distributions. As more data is being recorded however, it is expected that significant site-effects will be revealed at ICEARRAY II.

Table 4.7 Posterior mean, standard deviation and 95 percentiles of $\boldsymbol{\delta S}$ *for ICEARRAY II*

δS	ICEARRAY II (swarm 2012)				ICEARRAY II (swarm 2013)		
	Station	mean	SD	(2.5, 97.5) %	mean	SD	(2.5, 97.5) %
δS_1	IS202	-0.07	0.11	(-0.30, 0.16)	--	--	--
δS_2	IS701	-0.01	0.11	(-0.24, 0.21)	-0.03	0.13	(-0.29, 0.23)
δS_3	IS702	-0.17	0.11	(-0.40, 0.06)	-0.18	0.13	(-0.44, 0.08)
δS_4	IS703	0.24	0.11	(0.01, 0.47)	0.09	0.13	(-0.17, 0.35)
δS_5	IS704	-0.05	0.11	(-0.28, 0.18)	0.03	0.13	(-0.23, 0.29)
δS_6	IS705	0.23	0.11	(0.00, 0.46)	0.28	0.13	(0.02, 0.55)
δS_7	IS707	--	--	--	0.01	0.13	(-0.25, 0.27)

Figure 4.29 shows the comparison of the obtained station terms from ICEARRAY I and II. It is noteworthy that posterior distributions of station terms specify to what relative extent PGA can be expected to be either higher or lower than the mean PGA over the array. Contrary to the ICEARRAY I, the ICEARRAY II station terms are not well constrained as they have considerably large standard. It is largely due to the complexity of station conditions and limited data.

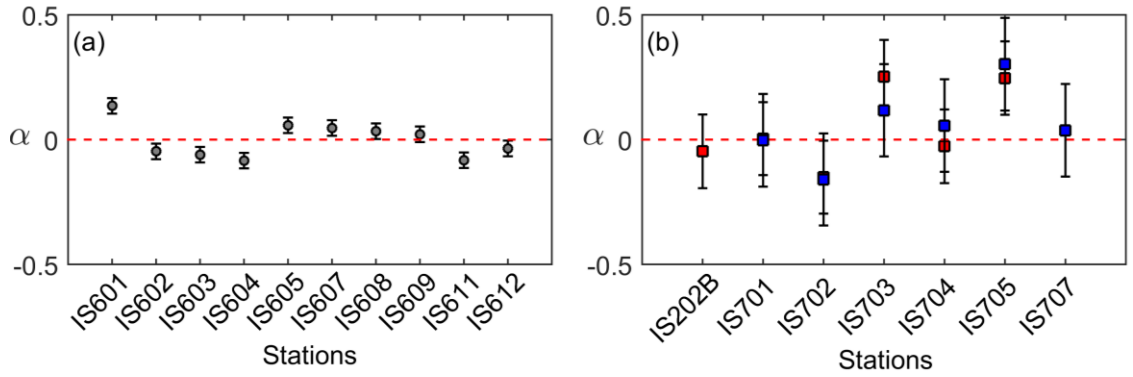

Figure 4.29 Posterior mean and standard deviation of the station term for (a) ICEARRAY I (grey symbols), and (b) ICEARRAY II (red symbols for swarm 2012, and blue symbols for swarm 2013).

Even though large station term variabilities for ICEARRAY II, mapping the spatial distribution of mean posterior of the station terms and mean $\log_{10}$ PGA confirms good agreement between spatial distribution patterns across both arrays (see Figure 4.30 and Figure 4.31). As can be seen in Figure 4.29(b), the station terms follow almost the same trend for both data sets across ICEARRAY II; however, the obtained mean values (except IS703) and standard deviations are slightly larger for the 2013 event. It is noteworthy that although stations IS703 and IS705 with interstation distance of 210 m in the western part of the array are the closest stations, their obtained station terms are noticeably different. The only convincing explanation for such difference can be related to the station locations in which station IS703 sits on the edge of the southern terraces of the town and different wave direction can largely affect the station term while station IS705 is placed in the middle of the terraces.

Variabilities

The inference of the hyperparameters for ICEARRAY I shows that inter-event standard deviation is larger than intra-event standard deviation which consists of both inter-station and inter-record variabilities (see Table 4.4). In general, the inter-event variability in empirical GMMs is reported to be smaller than intra-event variability (Strasser et al. 2009; Kuehn and Scherbaum 2015); however, the inter-event variability obtained for ICEARRAY I is larger than the intra-event variability. This is mainly due to the relatively uniform site conditions and also similar wave propagation across the region. In contrast, from the ICEARRAY II results intra-event variability appears higher than the inter-event variability (Table 4.4). Contrary to the ICEARRAY I, the strong-motion recorded by ICEARRAY II are characterized by small spatial dimensions of source regions compared to large hypocentral distances of ~50-80 km (cf. Figure 4.2). Especially for the 2012 events which are located in the same small source region on the HFF, while the 2013 event locations are distributed over the GOR lineament as a result have a wider azimuthal range. But, an important factor is the relative size differences of the datasets between ICEARRAY II compared to ICEARRAY I.

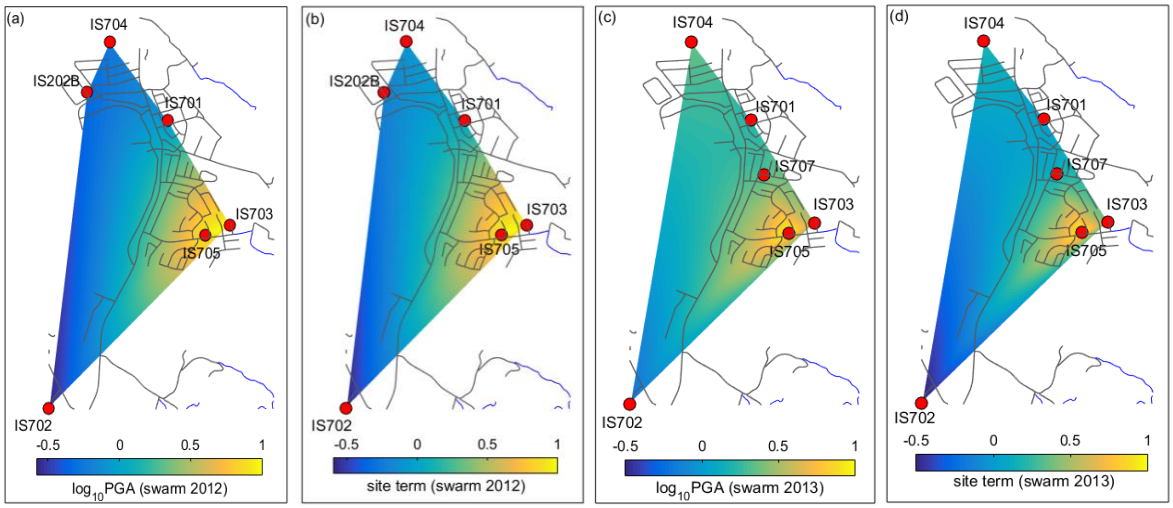

Figure 4.30 Mapping the spatial distribution of normalized (a) and (c) mean of log_{10}PGA, (b) and (d) posterior mean of station terms for ICEARRAY II stations obtained from swarm 2012 and 2013.

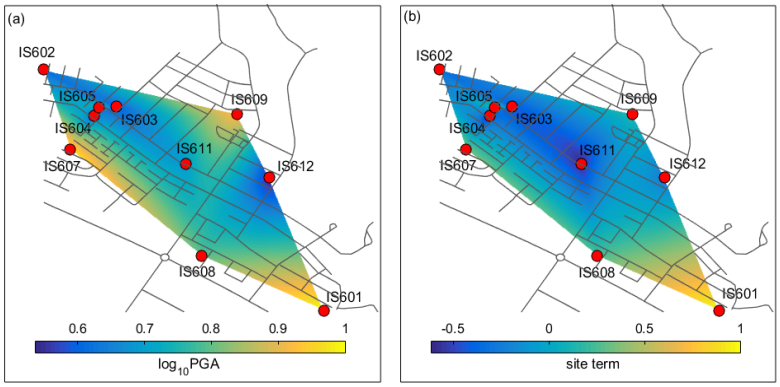

Figure 4.31 Mapping the spatial distribution of normalized (a) mean of log_{10}PGA and (b) posterior mean of station terms for ICEARRAY I stations obtained from aftershocks of 2008 Ölfus earthquake.

The proposed BHM quantifies the observed variabilities and the relative contributions of the earthquake and path effects, localized station effects, and their interdependent effects, respectively, to those variabilities of array peak ground accelerations. The inference of the posterior distribution of the latent parameters as well as the hyperparameters obtained for ICEARRAY I and II indicates significant station-effects with respect to local geology and topography. The large difference in confidence limit in the results depending on the size of the dataset emphasizes the need for more detailed geological measurements and analysis of the ground motion characteristics across relatively small areas where relative differences of ground motion intensities are of interest and importance, such as in urban areas near active faults. More detailed results of this section have been presented in ***Paper II*** and ***III***.

Source:

a) *Rahpeyma S., Halldorsson B., Hrafnkelsson B., and Jónsson S. (2018) Bayesian hierarchical model for earthquake peak ground acceleration recorded on a small-aperture array, Environmetrics, 29(3).*
b) *Rahpeyma S., Halldorsson B., Hrafnkelsson B., Green R. A., and Jónsson S. (2018a) Site effect estimation on two Icelandic strong-motion arrays using a Bayesian hierarchical model of spatial variation of earthquake peak ground acceleration, Soil Dynamics and Earthquake Engineering, (accepted, in press)*

5 Conclusion and Future Perspectives

5.1 Summary

The South Iceland and Tjörnes Fracture Zones are the regions in Iceland that possess the greatest potential for large and destructive earthquake occurrence. In Iceland, standard earthquake engineering practice generally assumes that the site conditions across small distances are identical. Hence, for an urban area of relatively uniform building stock subjected to strong earthquake ground motions, a uniform damage distribution can be expected. This condition would have been the case for the South Iceland due to relatively uniform site conditions (i.e., surficial lava-rock or old exposed bedrock) across the region. However, the preliminary analysis of the strong-motion data recorded by ICEARRAY I after 2008 Ölfus earthquake showed strong variation in the site responses across short distances ($\sim$1.23 km^2). The same variations have been observed by analyzing 2012 and 2013 swarms recorded by ICEARRAY II in the North Iceland with more complex and sedimentary origins subsoil structures. Therefore, in this thesis various aspects, methods, and data sets of site effects investigation with focus on earthquake engineering applications have been explored.

A comprehensive site effect analysis has been employed across ICEARRAY I (Hveragerði, South Iceland) and ICEARRAY II (Húsavík, North Iceland) by applying Standard Spectral Ratio (SSR) and Horizontal-to-Vertical Spectral Ratio (HVSR) methods using microseismic measurements as well as earthquake strong-motion recordings. The results between the methods are consistent and show that while the amplification levels remain relatively low, the predominant frequency varies systematically between stations and is found to correlate with the geological units. In particular, for stations on lava-rock the underlying geologic structure is characterized by repeated lava-soil stratigraphy characterized by reversals in the shear wave velocity with depth.

In order to estimate the shear-wave velocity model as the most common and practical indicator of site conditions in geological and earthquake engineering applications, two different frameworks have been proposed for different site conditions (1) regular velocity profiles i.e., monotonically increasing shear-wave velocity and density values with depth and (2) velocity profiles with reversals i.e., where a stiff lava layer lies on top of softer sediment, leading to a sudden decrease in the shear wave velocity with depth. The latter is common subsoil feature in geological younger parts of Iceland. For the first subsoil structure we proposed a practical method based on the inversion of the observed microseismic HVSR spectrum. The inversion scheme is set up in the context of the Bayesian framework using Markov chain Monte Carlo (MCMC) technique with Metropolis steps in order to explore the space of model parameters and find the best fitting family of subsoil properties along with their associated variabilities. The theoretical HVSR spectrum are calculated through body-wave approximation as a reliable estimate for subsoil structure of the sedimentary layers overlaying the half-space. The proposed technique has been assessed using microseismic recordings from two nominated station from IzmirNet (Turkey) and Mirandola (Italy). Nevertheless, we investigated that standard modeling of the HVSR using vertically incident body waves is not applicable for regions characterized by velocity reversals.

Instead, modeling the soil structure as a two-degree-of-freedom dynamic system is found to capture the observed predominant frequencies of site amplification.

Finally, we developed a ground motions model for peak ground accelerations from ground motion databases recorded by the Icelandic arrays using a Bayesian Hierarchical Model (BHM). The BHM allows the partitioning of the model into event, station, and event-station terms, which in turn allows the relative contributions of source, path, and site effects to be quantified. The results indicate that site effects influenced the motions across both ICEARRAY I and II. Although the site conditions across ICEARRAY I have been classified as uniform (i.e., "rock" with relatively flat topography), site effects contribute ~13% to the total variability in the amplitudes of predicted ground motions across the array. In contrast to ICEARRAY I, the variation of the geologic profiles and topography is much greater across ICEARRAY II. As a result, the site effects contribute ~37%-55% (using different data sets) to the of the total variability in the amplitudes of predicted ground motions across the array, with the contributions being less constrained for ICEARRAY II than ICEARRAY I due to the relative sizes of the recorded ground motion databases. These results facilitate our understanding of the key factors that affect the variation of seismic ground motions across a relatively small area. Such a detailed microzonation is of great importance for earthquake hazard assessment on a local scale and has practical implications for engineering decision making.

5.2 Future Work

Following the main objective of the first paper in this thesis and take advantage of the proposed forward modeling of the soil structure as a two-degree-of freedom dynamic system to capture the observed predominant frequencies of site amplification, an inversion framework can be set to estimate the most probable subsoil structure in the presence of the strong velocity reversals. Theoretically, an inversion scheme can be set up in the context of the Bayesian statistical framework using the Markov Chain Monte Carlo (MCMC) technique along with the Metropolis step algorithm in order to explore the space of model parameters and find the best fitting family of subsoil properties.

The characterization of the spatially correlated ground motion parameters is of great value and importance and extension the current research would be treasured. While the peak ground acceleration was the ground motion parameter studied in this thesis, the efficiency of the proposed BHM in this study becomes particularly beneficial and other ground motion parameters can be obtained using the same methodology, such as peak ground velocity, the pseudo-spectral accelerations/velocities/displacements, ground motion duration, etc. Therefore, the next logical step of the current study would be integration the proposed model to estimate the spatial correlation of peak ground motions and response spectra. As a result, the correlation coefficients of different ground motion parameters between the response spectrum at different frequencies can be calculated. It is noteworthy that the correlation coefficients between ground motion parameters are highly important for the estimation of joint hazard or the conditional mean spectrum which has not been implemented for Iceland yet. In this approach, contrary to the previous studies, we will show how one can estimate jointly a ground motion model and the correlations.

Despite of using dense recordings of the Icelandic arrays to determine the spatial distribution of the ground motion amplitudes in local scale, the ISMN recordings can be used in order to

decompose ground motion parameters into source, path, and site terms in particular to quantitatively estimate the regional effects in macroscale. Subsequently, the proposed modeling will provide the opportunity for the first time in Iceland to develop partially (e.g., single station) or fully non-ergodic GMMs to quantitatively determine source, path and site effects and their associated variabilities which are highly essential for seismic hazard analysis in Iceland.

The proposed BHM in this research project enables us to quantitatively estimate the level of site terms without engaging the V_{S30}. To date, the site term for almost all the GMMs is defined as a function of V_{S30} namely the most common indicator to determine the site. Therefore, implementing the proposed BHM can help us to quantitatively investigate how reliable is this indicator for different site classes. The performance of various site-condition proxies in commonly used GMMs in Iceland will be examined under the Bayesian framework. Different site proxies such as V_{S30} and slope, which are generally considered a priori as more relevant for short period ground motions, and f_0 (predominant frequency) along with H_{800} (the depth beyond which V_s exceeds 800 m/s) that should in principle be more suitable for long periods as well as different combination of site proxies will be examined. The results will pioneeringly investigate the velocity reversals effects on site-effects proxies.

Moreover, the proposed BHM modeling can be very helpful when we do not have reliable information about the geological structure of the stations. In general, in the author's opinion, implementation of the proposed statistical BHM model for spatial observations to other areas in the world along with future improvements of the modeling and components, presents an interesting and vivid area of future research.

Such analyses facilitate the understanding of the key factors that affect the variation of seismic ground motions across a relatively small area. The resulting detailed microzonation is of great importance for earthquake hazard assessment on a local and macro-scale, with practical implications for engineering decision making e.g., for urban planning and microzonation studies.

Finally, array data is especially practical when applying array processing techniques which are powerful tools in quantifying the geophysical and geological characteristics of the array area. Hence, the available data recorded by ICEARRAY I and II provides unique opportunity to analyze and model seismic wave propagation across Hveragerði and Húsavík using array processing techniques. In this basis, one of the main objectives of our future work will be two-dimensional shear-wave velocity structure estimate using seismic travel time inversion. The inversion scheme will be defined based on ray tracing in the context of the Bayesian statistical framework using the Markov Chain Monte Carlo (MCMC) technique along with the Metropolis step algorithm. On the one hand, the obtained distinct subsoil structures and their associated variabilities will be used to construct a two-dimensional model of the subsoil overburden underneath Hveragerði and Húsavík implementing a geostatistical interpolation technique known as *Kriging* method. As the result, the spatial distribution of the lava flow and sedimentary strata will be estimated.

Bibliography

Abrahamson NA (1988) Statistical properties of peak ground accelerations recorded by the SMART 1 array. Bulletin of the Seismological Society of America 78:26–41.

Abrahamson NA, Atkinson G, Boore DM, Bozorgnia Y, Campbell K, Chiou B, Idriss IM, Silva W, and Youngs R (2008) Comparisons of the NGA ground-motion relations. Earthquake Spectra 24:45.

Abrahamson NA, Shedlock KM (1997) Overview. Seismological Research Letters 68:9–23.

Abrahamson NA, Youngs RR (1992) A stable algorithm for regression analyses using the random effects model. Bulletin of the Seismological Society of America 82:505–510.

Aki K (1957) Space and Time Spectra of Stationary Stochastic Waves, with Special Reference to Microtremors. Bulletin of Earthquake Research Institute 35:415-456.

Aki K (1988) Local site effects on ground motion. Earthquake Engineering and Soil Dynamics II - Recent Advances in Ground Motion Evaluation, Geotechnical Special Pudlication 20:103–155

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

Akyol Ni, Kurtulmuş TÖ, Çamyildiz M, Güngör T (2013) Spectral Ratio Estimates for Site Effects on the Horst–Graben System in West Turkey. Pure Appl Geophys 170:2107–2125.

Al Atik L, Abrahamson N, Bommer JJ, et al (2010) The variability of ground-motion prediction models and its components. Seismological Research Letters 81:794–801.

Albarello D, Lunedei E (2010) Alternative interpretations of horizontal to vertical spectral ratios of ambient vibrations: new insights from theoretical modeling. Bulletin of Earthquake Engineering 8:519–534.

Anderson JG, Brune JN (1999) Probabilistic seismic hazard analysis without the ergodic assumption. Seismological Research Letters 70:19–28.

Angelier J, Bergerat F, Stefansson R, Bellou M (2008) Seismotectonics of a newly formed transform zone near a hotspot: Earthquake mechanisms and regional stress in the South Iceland Seismic Zone. Tectonophysics 447:95–116.

Ansal A, Laue J, Buchheister J, Erdik M, Springman SM, Studer J and Köksal D (2004) Site characterization and site amplification for a seismic microzonation study in Turkey. In: Proceedings of the 11th International Conference on Soil Dynamic and Earthquake Engineering (11th ICSDEE) and the 3rd International Conference on

Earthquake Geotechnical Engineering (3rd ICEGE), January 7–9, Berkeley. pp 53–60.

Anzidei M, Maramai A, Montone P (2012) Preface. Annals of Geophysics 55.

Arai H, Tokimatsu K (2004) S-Wave Velocity Profiling by Inversion of Microtremor H/V Spectrum. Bulletin of the Seismological Society of America 94:53–63.

Arai H, Tokimatsu K (2005) S-Wave Velocity Profiling by Joint Inversion of Microtremor Dispersion Curve and Horizontal-to-Vertical (H/V) Spectrum. Bulletin of the Seismological Society of America 95:1766–1778.

Arai H, Tokimatsu K, Abe A (1996) Comparison of local amplifications estimated from microtremor F–K spectrum analysis with earthquake records. In: 11th World Conference on Earthquake Engineering. Acapulco, Mexico.

ASCE (2007) Seismic rehabilitation of existing buildings. Reston, VA.

ASCE (2010) Minimum Design Loads for Buildings and Other Structures. Reston, VA.

Asten M, Henstridge J (1984) Array estimators and the use of microseisms for reconnaissance of sedimentary basins. Geophysics 49:1828–1837.

Atakan K (2009) The Need for Standardized Approach for Estimating the Local Site Effects Based on Ambient Noise Recordings. In: Mucciarelli M, Herak M, Cassidy J (eds) Increasing Seismic Safety by Combining Engineering Technologies and Seismological Data. Springer Netherlands, pp 3–15

Atakan K, Brandsdóttir B, Halldórsson P, Friðleifsson G. (1997) Site Response as a Function of Near-Surface Geology in the South Iceland Seismic Zone. Natural Hazards 15:139–164.

Atkinson GM (2006) Single-Station Sigma. Bulletin of the Seismological Society of America 96:446–455.

Bard PY (1998) Microtremor measurements: A tool for site effect estimation? In: Proceeding of the Second International Symposium on the Effects of Surface Geology on Seismic Motion. Yokohama, Japan. pp 1251–1279

Bard PY, SESAME-Team (2005) Guidelines for the implementation of the H/V spectral ratio technique on ambient vibrations: measurements, processing, and interpretations. SESAME European research project. SESAME European research project.

Beck JL, Hall JF (1986) Factors contributing to the catastrophe in Mexico City during the earthquake of September 19, 1985. Geophysical Research Letters 13:593–596.

Bellou M, Bergerat F, Angelier J, Homberg C (2005) Geometry and segmentation mechanisms of the surface traces associated with the 1912 Selsund Earthquake, Southern Iceland. Tectonophysics 404:133–149.

Beresnev IA, Wen K-L, Yeh YT (1995) Nonlinear soil amplification: Its corroboration in Taiwan. Bulletin of the Seismological Society of America 85:496–515.

Berger JO (2013) Statistical Decision Theory and Bayesian Analysis. Springer Science & Business Media.

Bessason B, Kaynia AM (2002) Site amplification in lava rock on soft sediments. Soil Dynamic and Earthquake Engineering 22:525–540.

Bjarnason IT, Cowie P, Anders MH, Seeber L, and Scholz CH (1993) The 1912 Iceland Earthquake Rupture: Growth and Development of a Nascent Transform System. Bulletin of the Seismological Society of America 83:416–435.

Boatwright J, Fletcher JB, Fumal TE (1991) A general inversion scheme for source, site, and propagation characteristics using multiply recorded sets of moderate-sized earthquakes. Bulletin of the Seismological Society of America 81:1754–1782.

Böðvarsson R, Rögnvaldsson ST, Slunga R, Kjartansson E (1999) The SIL data acquisition system—at present and beyond year 2000. Physics of the earth and planetary interiors 113:89–101.

Bolt BA, Tsai YB, Yeh K, Hsu MK (1982) Earthquake strong motions recorded by a large near-source array of digital seismographs. Earthquake Engineering and Structural Dynamics 10:561–573.

Bommer JJ, Abrahamson NA (2006) Why do modern probabilistic seismic-hazard analyses often lead to increased hazard estimates? Bulletin of the Seismological Society of America 96:1967–1977.

Bommer JJ, Abrahamson NA, Strasser FO, Pecker A, Bard PY, Bungum H, Cotton F, Fäh D, Sabetta F, Scherbaum F and Studer J (2004) The Challenge of Defining Upper Bounds on Earthquake Ground Motions. Seismological Research Letters 75:82–95.

Bommer JJ, Crowley H (2006) The influence of ground-motion variability in earthquake loss modelling. Bulletin of Earthquake Engineering 4:231–248.

Bonnefoy-Claudet S, Baize S, Bonilla LF, et al (2009) Site effect evaluation in the basin of Santiago de Chile using ambient noise measurements. Geophysical Journal International 176:925–937.

Bonnefoy-Claudet S, Cécile C, Pierre-Yves B, Fabrice C, Peter M, Jozef K, and Fäh D (2006) H/V ratio: a tool for site effects evaluation. Results from 1-D noise simulations. Geophysical Journal International 167:827–837.

Bonnefoy-Claudet S, Köhler A, Cornou C, Wathelet M, and Bard PY (2008) Effects of Love Waves on Microtremor H/V RatioEffects of Love Waves on Microtremor H/V Ratio. Bulletin of the Seismological Society of America 98:288–300.

Boore DM (2004a) Estimating Vs30 (or NEHRP site classes) from shallow velocity models (depths < 30 m). Bulletin of the Seismological Society of America 94:591–597.

Boore DM (2004b) Can site response be predicted? Journal of Earthquake Engineering 8:1–41.

Boore DM (1983) Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra. Bulletin of the Seismological Society of America 73:1865–1894.

Boore DM (2003) Simulation of Ground Motion Using the Stochastic Method. Pure and Applied Geophysics 160:635–676.

Boore DM, Joyner WB (1997) Site amplifications for generic rock sites. Bulletin of the Seismological Society of America 87:327–341.

Boore DM, Joyner WB, Fumal T (1997) Equations for estimating horizontal response spectra and peak acceleration from North American earthquakes: A summary of recent work. Seismological Research Letters 687:128–154.

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

Borcherdt RD (1992) Simplified site classes and empirical amplification factors for site-dependent code provisions. In: NCEER, SEAOC, BSSC Workshop on Site Response during Earthquakes and Seismic Code Provisions. University of Southern California, Los Angeles, California.

Borcherdt RD (1994) Estimates of site-dependent response spectra for design (methodology and justification). Earthquake spectra 10:617–653.

Brandsdóttir B, Parsons M, White RS, Guðmundsson O, Drew J, and Thorbjarnadóttir B (2010) The May 29th 2008 earthquake aftershock sequence within the South Iceland Seismic Zone: Fault locations and source parameters of aftershocks. Jokull: journal of the glaciological and geological societies of Iceland 60:23–46.

Brillinger DR, Preisler HK (1984) An exploratory analysis of the Joyner-Boore attenuation data. Bulletin of the Seismological Society of America 74:1441–1450.

Brillinger DR, Preisler HK (1985) Further analysis of the Joyner-Boore attenuation data. Bulletin of the Seismological Society of America 75:611–614.

Brooks SP (1998) Markov chain Monte Carlo method and its application. The statistician 69–100.

Brooks SP, Gelman A (1998) General methods for monitoring convergence of iterative simulations. Journal of computational and graphical statistics 7:434–455.

Brune J, Dorman J (1963) Seismic waves and earth structure in the Canadian shield. Bulletin of the Seismological Society of America 53:167–209.

Building Seismic Safety Council, BSSC (2003) NEHRP recommended provisions for seismic regulations for new buildings and other structures, Washington.

Campbell KW (1981) Near-source attenuation of peak horizontal acceleration. Bulletin of the Seismological Society of America 71:2039–2070.

Cárdenas M, Chávez-García FJ, Gusev A (1997) Regional amplification of ground motion in central Mexico. Results from coda-length magnitude data and preliminary modeling. Journal of Seismology 1:341–355.

Cárdenas-Soto M, Chávez-García FJ (2003) Regional Path Effects on Seismic Wave Propagation in Central Mexico. Bulletin of the Seismological Society of America 93:973–985.

Casella G, George EI (1992) Explaining the Gibbs Sampler. The American Statistician 46:167–174.

Castellaro S, Mulargia F, Rossi PL (2008) Vs30: Proxy for Seismic Amplification? Seismological Research Letters 79:540–543.

Celebi M, Prince J, Dietel C, Onate M, and Chavez G (1987) The Culprit in Mexico City—Amplification of Motions. Earthquake Spectra 3:315–328.

Chin B-H, Aki K (1991) Simultaneous study of the source, path, and site effects on strong ground motion during the 1989 Loma Prieta earthquake: a preliminary result on pervasive nonlinear site effects. Bulletin of the Seismological Society of America 81:1859–1884.

Chiu H-C, Amirbekian RV, Bolt BA (1995) Transferability of strong ground-motion coherency between the SMART1 and SMART2 arrays. Bulletin of the Seismological Society of America 85:342–348.

Chopra AK (1981) Dynamics of structures, a primer. Earthquake Engineering Research Institute, Berkeley, California.

Clifton A, Einarsson P (2005) Styles of surface rupture accompanying the June 17 and 21, 2000 earthquakes in the South Iceland Seismic Zone. Tectonophysics 396:141–159.

Congdon P (2014) Applied Bayesian modelling. John Wiley & Sons.

Cornou C (2006) Derivation of Structural Models from Ambient Vibration Array Recordings: Result from an Interpretational Blind Test. Proc 3rd Int Symp on the Effects of Surface Geology on Seismic Motion, Grenoble, France, 29 August-1 September.

Cox BR, Teague DP (2016) Layering ratios: a systematic approach to the inversion of surface wave data in the absence of a priori information. Geophysical Journal International 207:422–438.

Cranswick E (1988) The information content of high-frequency seismograms and the near-surface geologic structure of “hard rock” recording sites. In Scattering and Attenuations of Seismic Waves, Part I 128:333–363.

Cultrera G, Rubeis VD, Theodoulidis N, et al (2014) Statistical correlation of earthquake and ambient noise spectral ratios. Bulletin of Earthquake Engineering 12:1493–1514.

D'Amico V, Picozzi M, Baliva F, Albarello D (2008) Ambient noise measurements for preliminary site-effects characterization in the urban area of Florence, Italy. Bulletin of the Seismological Society of America 98:1373–1388.

DeMets C (1990) Current plate motions. Geophysical Journal International 101:425–478.

DeMets C, Gordon RG, Angus DF, Stein S (1994) Effect of recent revisions to the geomagnetic reversal time scale on estimates of current plate motions. Geophysical Research Letters 21:2191–2194.

Di Giacomo D, Maria Rosaria G, Mucciarelli, M, Parolai S, and Richwalski SM (2005) Analysis and Modeling of HVSR in the Presence of a Velocity Inversion: The Case of Venosa, Italy. Bulletin of the Seismological Society of America 95:2364–2372.

Diggle PJ, Tawn JA, Moyeed RA (1998) Model-based geostatistics. Journal of the Royal Statistical Society: Series C (Applied Statistics) 47:299–350.

Dimitriu P, Theodulidis N, Bard PY (2000) Evidence of nonlinear site response in HVSR from SMART1 (Taiwan) data. Soil Dynamics and Earthquake Engineering 20:155–165.

Dorman J, Ewing M, Oliver J (1960) Study of shear-velocity distribution in the upper mantle by mantle Rayleigh waves. Bulletin of the Seismological Society of America 50:87–115.

Douglas J, Aochi H (2008) A Survey of Techniques for Predicting Earthquake Ground Motions for Engineering Purposes. Surveys in Geophysics 29:187–220.

Douglas J, Gehl P (2008) Investigating strong ground-motion variability using analysis of variance and two-way-fit plots. Bulletin of Earthquake Engineering 6:389–405.

Douglas J, Halldorsson B (2010) On the use of aftershocks when deriving ground-motion prediction equations. In: 9th US National and 10th Canadian Conference on Earthquake Engineering (9USN/10CCEE). Toronto, Canada, Paper no. 220.

Einarsson P (1991) Earthquakes and present-day tectonism in Iceland. Tectonophysics 189:261–279.

Einarsson P (2008) Plate boundaries, rifts and transforms in Iceland. Jökull 58:35–58.

Einarsson P, Björnsson S, Foulger G, Stefánsson R, and Skaftadóttir T (1981) Seismicity pattern in the South Iceland seismic zone. Earthquake Prediction 141–151.

Einarsson Þ, Douglas G (1994) Geology of Iceland: rocks and landscape. Mál og menning, Reykjavík, Iceland.

Eissler H, Astiz L, Kanamori H (1986) Tectonic Setting and Source Parameters of the September 19, 1985 Michoacan, Mexico Earthquake. Geophysical Research Letters 13:569–572.

European Committee for Standardization (2003) Eurocode 8: Design of Structures for earthquake resistance - Part1: General rules, seismic actions and rules for buildings. European Standard.

Ewing WM, Jardetzky WS, Press F, Beiser A (1957) Elastic Waves in Layered Media. Physics Today 10:27.

Fäh D, Kind F, Giardini D (2001) A theoretical investigation of average H/V ratios. Geophysical Journal International 145:535–549.

Field EH (2000) Accounting for Site Effects in Probabilistic Seismic Hazard Analyses of Southern California: Overview of the SCEC Phase III Report. Bulletin of the Seismological Society of America 90:S1–S31.

Field EH, Jacob KH, Hough SE (1992) Earthquake site response estimation: A weak-motion case study. Bulletin of the Seismological Society of America 82:2283–2307.

Foti S, Comina C, Boiero D, Socco LV (2009) Non-uniqueness in surface-wave inversion and consequences on seismic site response analyses. Soil Dynamics and Earthquake Engineering 29:982–993.

Fukushima Y, Bonilla LF, Scotti O, Douglas J (2007) Site Classification Using Horizontal-to-vertical Response Spectral Ratios and its Impact when Deriving Empirical Ground-motion Prediction Equations. Journal of Earthquake Engineering 11:712–724.

Fukushima Y, Tanaka T (1990) A new attenuation relation for peak horizontal acceleration of strong earthquake ground motion in Japan. Bulletin of the Seismological Society of America 80:757–783.

Gallipoli MR, Mucciarelli M (2009) Comparison of Site Classification from VS30, VS10, and HVSR in Italy. Bulletin of the Seismological Society of America 99:340–351.

Garofalo F, Foti S, Hollender F, et al (2016a) InterPACIFIC project: Comparison of invasive and non-invasive methods for seismic site characterization. Part II: Inter-comparison between surface-wave and borehole methods. Soil Dynamics and Earthquake Engineering 82:241–254.

Garofalo F, Foti S, Hollender F, Bard PY, Cornou C, Cox BR, Ohrnberger M, Sicilia D, Asten M, Di Giulio G, and Forbriger T (2016b) InterPACIFIC project: Comparison of invasive and non-invasive methods for seismic site characterization. Part I: Intra-comparison of surface wave methods. Soil Dynamics and Earthquake Engineering 82:222–240.

Gelman A, Carlin J., Stern HS, Rubin DB (2014) Bayesian data analysis. Chapman & Hall/CRC, London.

Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. Statistical Science 457–472.

Geman S, Geman D (1984) Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. IEEE Transactions on Pattern Analysis and Machine Intelligence 6:721–741.

Gilks WR (2005) Markov Chain Monte Carlo. In: Encyclopedia of Biostatistics. John Wiley & Sons, Ltd.

Giulio GD, Cornou C, Ohrnberger M, Wathelet M, and Rovelli A (2006) Deriving Wavefield Characteristics and Shear-Velocity Profiles from Two- Dimensional Small-Aperture Arrays Analysis of Ambient Vibrations in a Small-Size Alluvial Basin, Colfiorito, Italy. Bulletin of the Seismological Society of America 96:1915–1933.

Gok E, Polat O (2014) An assessment of the microseismic activity and focal mechanisms of the Izmir (Smyrna) area from a new local network (IzmirNET). Tectonophysics 635:154–164.

Gudmundsson A (2007) Infrastructure and evolution of ocean-ridge discontinuities in Iceland. Journal of Geodynamics 43:6–29.

Guðmundsson Á, Brynjolfsson S, Jónsson MT (1993) Structural analysis of a transform fault-rift zone junction in North Iceland. Tectonophysics 220:205–221.

Guéguen P, Chatelain JL, Guillier B, Yepes H (2000) Technical note: An indication of soil topmost layer response in Quito (Ecuador) using noise H/V spectral ratio. Soil Dynamics and Earthquake Engineering 19:127–133.

Gutenberg B (1957) Effects of ground on earthquake motion. Bulletin of the Seismological Society of America 47:221–250.

Halldorsson B, Jónsson S, Papageorgiou AS, Green RA, Kalogeras I, Avery H, Ólafsson S, Oliveira CS, Polat O, and Sigurbjörnsson R (2012) ICEARRAY II: A new multidisciplinary strong-motion array in North Iceland. In: 15th World Conference on Earthquake Engineering (15WCEE). Lisbon, Portugal, Paper no. 2567.

Halldorsson B, Sigbjörnsson R (2009) The Mw6. 3 Ölfus earthquake at 15: 45 UTC on 29 May 2008 in South Iceland: ICEARRAY strong-motion recordings. Soil Dynamics and Earthquake Engineering 29:1073–1083.

Halldorsson B, Sigbjörnsson R, Schweitzer J (2009) ICEARRAY: the first small-aperture, strong-motion array in Iceland. Journal of Seismology 13:173–178.

Harichandran RS (1991) Estimating the spatial variation of earthquake ground motion from dense array recordings. Structural Safety 10:219–233.

Harichandran RS, Vanmarcke EH (1986) Stochastic Variation of Earthquake Ground Motion in Space and Time. Journal of Engineering Mechanics 112:154–174.

Hastings WK (1970) Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57:97–109.

Héloïse C, 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. Bulletin of Earthquake Engineering 10:451–489.

Hensch M, Riedel C, Reinhardt J, Dahm T, and N.I.C.E. (2008) Hypocenter migration of fluid-induced earthquake swarms in the Tjörnes Fracture Zone (North Iceland). Tectonophysics 447:80–94.

Herak M (2008) ModelHVSR-A Matlabs tool to model horizontal-to-vertical spectral ratio of ambient noise. Computers & Geosciences 34:1514–1526.

Houston H, Kanamori H (1986) Source characteristics of the 1985 Michoacan, Mexico Earthquake at periods of 1 to 30 seconds - Houston - 1986 - Geophysical Research Letters - Wiley Online Library. Geophysical Research Letters 13:597–600.

Joyner WB, Boore DM (1981) Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake. Bulletin of the Seismological Society of America 71:2011–2038.

Joyner WB, Boore DM (1993) Methods for regression analysis of strong-motion data. Bulletin of the Seismological Society of America 83:469–487.

Kanai K, Tanaka T (1961) On Microtremor VIII. Bulletin of the Earthquake Research Institute, Tokyo University 39:97–114.

Kawase H, Aki K (1989) A study on the response of a soft basin for incident S, P, and Rayleigh waves with special reference to the long duration observed in Mexico City. Bulletin of the Seismological Society of America 79:1361–1382.

Kayen R, Moss RES, Thompson EM, et al (2013) Shear-Wave Velocity–Based Probabilistic and Deterministic Assessment of Seismic Soil Liquefaction Potential. Journal of Geotechnical and Geoenvironmental Engineering 139:407–419.

Kind F, Fäh D, Giardini D (2005) Array measurements of S-wave velocities from ambient vibrations. Geophysical Journal International 160:114–126.

Kiureghian AD (1996) A Coherency Model for Spatially Varying Ground Motions. Earthquake Engineering and Structural Dynamics 25:99–111.

Knopoff L (1972) Observation and inversion of surface-wave dispersion. Tectonophysics 13:497–519.

Kokusho T, Sato K (2008) Surface-to-base amplification evaluated from KiK-net vertical array strong motion records. Soil Dynamics and Earthquake Engineering 28:707–716.

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

Kowsari M, Halldorsson B, Hrafnkelsson B, et al (2017) On the Selection of Ground-Motion Prediction Equations for Seismic Hazard Assessment in the South Iceland Seismic Zone. In: 16th World Conference on Earthquake Engineering (16WCEE). Santiago, Chile, Paper no. 2809.

Kramer SL (1996) Geotechnical earthquake engineering. Prentice Hall, USA: New Jersey

Kuehn NM, Scherbaum F (2015) Ground-motion prediction model building: a multilevel approach. Bulletin of Earthquake Engineering 13:2481–2491.

Kuehn NM, Scherbaum F (2016) A partially non-ergodic ground-motion prediction equation for Europe and the Middle East. Bulletin of Earthquake Engineering 1–14.

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

Lay T, Wallace TC (1995) Modern Global Seismology. Elsevier

Lee VW, Trifunac MD (2010) Should average shear-wave velocity in the top 30m of soil be used to describe seismic amplification? Soil Dynamics and Earthquake Engineering 30:1250–1258.

Lee WH, Jennings P, Kisslinger C, Kanamori H (2002) International Handbook of Earthquake & Engineering Seismology. Academic Press

Loh C-H (1985) Analysis of the spatial variation of seismic waves and ground movements from smart-1 array data. Earthquake engineering & structural dynamics 13:561–581.

Loh CH, Ang AH-S, Wen YK (1983) Spatial Correlation Study of Strong Motion Array Data with Application to Lifeline Earthquake Engineering. University of Illinois Engineering Experiment Station. College of Engineering. University of Illinois at Urbana-Champaign.

Lozano L, Herraiz M, Singh SK (2009) Site effect study in central Mexico using H/V and SSR techniques: Independence of seismic site effects on source characteristics. Soil Dynamics and Earthquake Engineering 29:504–516.

Luke B, Calderón-Macías C, Stone R, Huynh M (2003) Non-Uniqueness in Inversion of Seismic Surface-Wave Data. In: Symposium on the Application of Geophysics to Engineering and Environmental Problems 2003. Environment and Engineering Geophysical Society, pp 1342–1347.

Lunedei E, Albarello D (2010) Theoretical HVSR curves from full wavefield modelling of ambient vibrations in a weakly dissipative layered Earth. Geophysical Journal International 181:1093–1108.

Lunedei E, Malischewsky P (2015) A Review and Some New Issues on the Theory of the H/V Technique for Ambient Vibrations. In: Perspectives on European Earthquake Engineering and Seismology. Springer, Cham, pp 371–394.

Masuda T, Ohtake M (1992) Comment on "a new attenuation relation for peak horizontal acceleration of strong earthquake ground motion in Japan" by Y. Fukushima and T. Tanaka. Bulletin of the Seismological Society of America 82:521–522.

Menke W (1989) Geophysical data analysis: Discrete inverse theory.

Metropolis N, Rosenbluth AW, Rosenbluth MN, et al (1953) Equation of state calculations by fast computing machines. The journal of chemical physics 21:1087–1092.

Molnar S, Cassidy JF (2006) A comparison of site response techniques using weak-motion earthquakes and microtremors. Earthquake Spectra 22:169.

Molnar S, Dosso SE, Cassidy JF (2010) Bayesian inversion of microtremor array dispersion data in southwestern British Columbia. Geophysical Journal International 183:923–940.

Moya A, Aguirre J, Irikura K (2000) Inversion of Source Parameters and Site Effects from Strong Ground Motion Records using Genetic Algorithms. Bulletin of the Seismological Society of America 90:977–992.

Mucciarelli and Gallipoli (2001) A critical review of 10 years of microtremor HVSR technique.pdf. Bollettino di Geofisica Teorica ed Applicata 42:255–266.

Nakamura Y (1989) A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface. Quarterly Report of Railway Technical Research Institute 30:25–33.

Nakamura Y (2000) Clear identification of fundamental idea of Nakamura's technique and its applications. In: Proceedings of the 12th World Conference on Earthquake Engineering. p 2656.

Nakamura Y (2008) On the H/V spectrum. In: Proceedings of the 14th World Conference on Earthquake Engineering. Beijing, China.

Nazarian S, Ii S, H K, Hudson WR (1983) Use of Spectral analysis of surface waves method for determination of moduli and thicknesses of pavement system. Transportation Research Record.

Nogoshi M, Igarashi T (1970) On the propagation characteristics estimations of subsurface using microtremors on the ground surface. Journal of Seismological Society of Japan 23:264–280.

Nogoshi M, Igarashi T (1971) On the amplitude characteristics of microtremor (Part 2). Journal of Seismological Society of Japan 24:26–40.

Olivera CI, Halldorsson B, Green RA, Sigbjörnsson R (2014) Site effects estimation using ambient noise and earthquake data in Húsavík, north Iceland. In: Proceedings of the

2nd European Conference on Earthquake and Engineering Seismology (2ECEES). Istanbul, Turkey.

Ordaz M, Singh SK (1992) Source spectra and spectral attenuation of seismic waves from Mexican earthquakes, and evidence of amplification in the hill zone of Mexico City. Bulletin of the Seismological Society of America 82:24–43.

Ortiz-Alemán C, Reyes-Olvera M, Iglesias-Mendoza A, Orozco-Del-Castillo MG, and Hernández-Gómez JJ (2017) Estimation of source; path and site effects at MASE array stations: a comprehensive study. Geofísica Internacional 56:71-85.

Panzera F, Lombardo G, Monaco C, Stefano AD (2015) Seismic site effects observed on sediments and basaltic lavas outcropping in a test site of Catania, Italy. Natural Hazards 79:1–27.

Park C, Miller R, Xia J (1999) Multichannel analysis of surface waves. GEOPHYSICS 64:800–808.

Parolai S, Mucciarelli M, Gallipoli MR, Richwalski SM, and Strollo A (2007) Comparison of Empirical and Numerical Site Responses at the Tito Test Site, Southern Italy. Bulletin of the Seismological Society of America 97:1413–1431.

Parolai S, Picozzi M, Richwalski SM, Milkereit C (2005) Joint inversion of phase velocity dispersion and H/V ratio curves from seismic noise recordings using a genetic algorithm, considering higher modes. Geophysical Research Letters 32:L01303.

Pelekis PC, Athanasopoulos GA (2011) An overview of surface wave methods and a reliability study of a simplified inversion technique. Soil Dynamics and Earthquake Engineering 31:1654–1668.

Pétursson GG, Vogfjörð, Kristín (2009) Attenuation relations for near- and far-field peak ground motion (PGV, PGA) and new magnitude estimates for large earthquakes in SW-Iceland. Icelandic Meteorological Office, Reykjavík, Iceland.

Picozzi M, Albarello D (2007) Combining genetic and linearized algorithms for a two-step joint inversion of Rayleigh wave dispersion and H/V spectral ratio curves. Geophysical Journal International 169:189–200.

Pilz M, Parolai S, Leyton F, Campos J, and Zschau J (2009) A comparison of site response techniques using earthquake data and ambient seismic noise analysis in the large urban areas of Santiago de Chile. Geophysical Research Letters 178:713–728.

Pitilakis K (2004) Site Effects. In: Recent Advances in Earthquake Geotechnical Engineering and Microzonation. Springer, Dordrecht, pp 139–197.

Polat O, Ceken U, Uran T, Gok E, Yilmaz N, Beyhan M, Koc N, Arslan B, Yilmaz D, and Utku M (2009) IzmirNet: A strong-motion Network in Metropolitan Izmir, Western Anatolia, Turkey. Seismological Research Letters 80:831–838.

Power M, Chiou BS-J, Abrahamson NA, Bozorgnia Y, Shantz T, and Roblee, C (2008) An Overview of the NGA Project. Earthquake Spectra 24:3–21.

Rahpeyma S, Halldorsson B, Olivera C, Green, RA, and Jonsson, S (2016) Detailed site effect estimation in the presence of strong velocity reversals within a small-aperture strong-motion array in Iceland. Soil Dynamics and Earthquake Engineering 89:136–151.

Rahpeyma S, Halldorsson B, Green RA (2017) On the Distribution of Earthquake Strong-motion Amplitudes and Site Effects Across the Icelandic Strong-motion Arrays. In: 16th World Conference on Earthquake Engineering (16WCEE). Santiago, Chile, p Paper no. 2762.

Rahpeyma S, Halldorsson B, Hrafnkelsson B, Jónsson S (2018) Bayesian hierarchical model for variations in earthquake peak ground acceleration within small-aperture arrays. Environmetrics, 29:e2497.

Rahpeyma S, Halldorsson B, Hrafnkelsson B, Green RA, and Jónsson S (2018a) Site effect estimation on two Icelandic strong-motion arrays using a Bayesian hierarchical model for the spatial distribution of earthquake peak ground acceleration. Soil Dynamics and Earthquake Engineering (Accepted, in press).

Rahpeyma S, Halldorsson B, Hrafnkelsson B, Green RA, Polat O, and Jónsson S (2018b) Estimate of shear-wave velocity profile using microseismic Horizontal-to-Vertical Spectral Ratios inversion. (Manuscript completed).

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. Earthquake Spectra 14:203–224.

Roberts GO, Gelman A, Gilks WR (1997) Weak convergence and optimal scaling of random walk Metropolis algorithms. The annals of applied probability 7:110–120.

Roten D, Fäh D (2007) A combined inversion of Rayleigh wave dispersion and 2- D resonance frequencies. Geophysical Journal International 168:1261–1275.

Roullé A, Chávez-García FJ (2006) The strong ground motion in Mexico City: Analysis of data recorded by a 3D array. Soil Dynamics and Earthquake Engineering 26:71–89.

Sæmundsson K (1974) Evolution of the axial rifting zone in northern Iceland and the Tjornes fracture zone. Bulletin of the Geological Society of America 85:495.

Sæmundsson K (1979) Outline of the geology of Iceland.

Sæmundsson K, Karson JA (2006) Stratigraphy and tectonics of the Húsavík-Western Tjörnes Area. Iceland Geosurvey.

Sæmundsson K, Kristinsson S (2005) Hveragerði. Soil temperature measurements and faults. Iceland GeoSurvey (ÍSOR), Reykjavík, Iceland (in Icelandic).

Sánchez-Sesma FJ, Rodríguez M, Iturrar´an-Viveros U, Luzón F, Campillo M, Margerin L, García-Jerez A, Suarez M, Santoyo MA, and Rodríguez-Castellanos A (2011) A theory for microtremor H/V spectral ratio: application for a layered medium. Geophysical Journal International 186:221–225.

Scherbaum F, Hinzen K-G, Ohrnberger M (2003) Determination of shallow shear wave velocity profiles in the Cologne, Germany area using ambient vibrations. Geophysical Journal International 152:597–612.

Seed HB, Romo MP, Sun JI, Jaime A, and Lysmer J (1988) The Mexico Earthquake of September 19, 1985—Relationships Between Soil Conditions and Earthquake Ground Motions. Earthquake Spectra 4:687–729.

Shabestari KT, Yamazaki F, Saita J, Matsuoka M (2004) Estimation of the spatial distribution of ground motion parameters for two recent earthquakes in Japan. Tectonophysics 390:193–204.

Shapiro N (1996) Etude de l'intéraction des ondes sismiques guidées régionales avecles hétérogénéités delacroute. PhD thesis Universite´ Joseph Fourier.

Sigbjörnsson R (1990) Strong motion measurements in Iceland and seismic risk assessment. In: 9th European Conference on Earthquake Engineering. pp 213–222.

Sigbjörnsson R, Ólafsson S, Thórarinsson Ó (2004) Strong-motion recordings in Iceland. In: Proceedings of the 13th World Conference on Earthquake Engineering. Mira, Vancouver, BC, Canada, Paper no. 2370.

Singh S, Mena E, Castro R (1988) Some aspects of the source characteristics and ground motion amplifications in and near Mexico City from acceleration data of the September 1985, Michoacan, Mexico earthquakes. Bulletin of the Seismological Society of America 78:451–477.

Smith AFM, Roberts GO (1993) Bayesian Computation Via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods. Journal of the Royal Statistical Society Series B (Methodological) 55:3–23.

Socco LV, Boiero D (2008) Improved Monte Carlo inversion of surface wave data. Geophysical Prospecting 56:357–371.

Sokolov VY, Loh C-H, Jean W-Y (2007) Application of horizontal-to-vertical (H/V) Fourier spectral ratio for analysis of site effect on rock (NEHRP-class B) sites in Taiwan. Soil Dynamics and Earthquake Engineering 27:314–323.

Stanko D, Markušić S, Strelec S, Gazdek M (2017) HVSR analysis of seismic site effects and soil-structure resonance in Varaždin city (North Croatia). Soil Dynamics and Earthquake Engineering 92:666–677.

Stefánsson R, Böðvarsson R, Slunga R, Einarsson P, Jakobsdottir S, Bungum H, Gregersen S, Havskov J, Hjelme J and Korhonen H (1993) Earthquake prediction research in the South Iceland Seismic Zone and the SIL project. Bulletin of the Seismological Society of America 83:696–716.

Stefansson R, Gudmundsson GB, Halldorsson P (2008) Tjörnes fracture zone. New and old seismic evidences for the link between the North Iceland rift zone and the Mid-Atlantic ridge. Tectonophysics 447:117–126.

Steidl JH (1993) Variation of Site Response at the UCSB Dense Array of Portable Accelerometers. Earthquake Spectra 9:289–302.

Steidl JH, Tumarkin AG, Archuleta RJ (1996) What Is a Reference Site? Bulletin of the Seismological Society of America 86:1733–1748.

Stephenson WR (2007) HVSR, SSR, and VS30 Values at Three New Zealand Basin Arrays. Journal of Earthquake Engineering 11:987–1009.

Stewart JP, Boore DM, Seyhan E, Atkinson GM (2015) NGA-West2 Equations for Predicting Vertical-Component PGA, PGV, and 5%-Damped PSA from Shallow Crustal Earthquakes. Earthquake Spectra 32:1005–1031.

Stokoe KH, Nazarian S (1983) Effectiveness of ground improvement from spectral analysis of surface waves. In: Proceedings of the European Conference on Soil Mechanics and Foundation Engineering.

Strasser FO, Abrahamson NA, Bommer JJ (2009) Sigma: Issues, Insights, and Challenges. Seismological Research Letters 80:40–56.

Suzuki T (1933) Amplitude of Rayleigh on the surface of a stratified medium. Bulletin of the Earthquake Research Institute, Tokyo Imperial University.

Tarabusi G, Caputo R (2017) The use of HVSR measurements for investigating buried tectonic structures: the Mirandola anticline, Northern Italy, as a case study. Int J Earth Sci (Geol Rundsch) 106:341–353.

Teague DP, Cox BR (2016) Site response implications associated with using non-unique Vs profiles from surface wave inversion in comparison with other commonly used methods of accounting for Vs uncertainty. Soil Dynamics and Earthquake Engineering.

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

Tokimatsu K (1997) Geotechnical site characterization using surface waves. Proc lst Inter Conf on Earthquake Geotechnical Engineering, Tokyo, 1997 3:1333–1368.

Triantafyllidis P, Theodulidis N, Savvaidis A, Papaioanou C, and Dimitriu P (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.

Tsai NC (1970) A Note on the Steady-State response of an Elastic Half-Space. Bulletin of the Seismological Society of America 60:795–808.

Waltl P (2013) Geomorphology and Building Stock of Húsavík, North Iceland: A Uniform GIS Database for Application in Hazard and Risk Modeling. Master's Thesis, BOKU University of Natural Resources and Life Sciences. Vienna, Austria.

Wang M, Takada T (2009) A Bayesian Framework for Prediction of Seismic Ground Motion. Bulletin of the Seismological Society of America 99:2348–2364.

Wathelet M (2008) An improved neighborhood algorithm: Parameter conditions and dynamic scaling. Geophysical Research Letters 35:L09301.

Wathelet M, Jongmans D, Ohrnberger M (2004) Surface-wave inversion using a direct search algorithm and its application to ambient vibration measurements. Near Surface Geophysics 2:211–221.

Yoshida N (2015) Seismic Ground Response Analysis. Springer Netherlands, Dordrecht.

Youd TL, Idriss IM (2001) Liquefaction Resistance of Soils: Summary Report from the 1996 NCEER and 1998 NCEER/NSF Workshops on Evaluation of Liquefaction Resistance of Soils. Journal of Geotechnical and Geoenvironmental Engineering 127:297–313.

Zerva A, Zervas V (2002) Spatial variation of seismic ground motions: An overview. Applied Mechanics Reviews 55:271.

Zerva A, Zhang O (1997) Correlation patterns in characteristics of spatially variable seismic ground motions. Earthquake engineering & structural dynamics 26:19–39.

Zhao JX, Irikura K, Zhang J, Fukushima Y, Somerville PG, Asano A, Ohno Y, Oouchi T, Takahashi T, and Ogawa H (2006) An Empirical Site-Classification Method for Strong-Motion Stations in Japan Using h/v Response Spectral Ratio. Bulletin of the Seismological Society of America 96:914–925.

Appendix A
Papers

Paper I

Detailed site effect estimation in the presence of strong velocity reversals within a small-aperture strong-motion array in Iceland

Sahar Rahpeyma, Benedikt Halldorsson, Christian Olivera,
Russel A. Green and Sigurjon Jónsson

Rahpeyma S, Halldorsson B., Olivera C, Green R.A, and Jónsson S (2016) Detailed site effect estimation in the presence of strong velocity reversals within a small-aperture strong-motion array in Iceland, *Soil Dynamics and Earthquake Engineering* **89**:136–151. doi:10.1016/j.soildyn.2016.07.001

Soil Dynamics and Earthquake Engineering 89 (2016) 136–151

Contents lists available at ScienceDirect

Soil Dynamics and Earthquake Engineering

journal homepage: www.elsevier.com/locate/soildyn

Detailed site effect estimation in the presence of strong velocity reversals within a small-aperture strong-motion array in Iceland

Sahar Rahpeyma [a], Benedikt Halldorsson [a,*], Christian Olivera [b], Russell A. Green [c], Sigurjón Jónsson [d]

[a] Earthquake Engineering Research Centre & Faculty of Civil and Environmental Engineering, School of Engineering and Natural Sciences, University of Iceland, Selfoss, Iceland
[b] ECS Mid-Atlantic, LLC, 6710 Oxon Hill Road, Suite 101, National Harbor, MS 20745, USA
[c] Department of Civil & Environmental Engineering, Virginia Tech, Blacksburg, VA, USA
[d] King Abdullah University of Science and Technology (KAUST), Physical Science and Engineering Division (PSE), Thuwal 23955-6900, Saudi Arabia

ARTICLE INFO

Article history:
Received 11 March 2016
Received in revised form
5 July 2016
Accepted 7 July 2016

Keywords:
Site effect
HVSR
SSR
Velocity reversal
Array
Iceland

ABSTRACT

The rock site characterization for earthquake engineering applications in Iceland is common due to the easily exposed older bedrock and more recent volcanic lava rock. The corresponding site amplification is generally assumed to be low but has not been comprehensively quantified, especially for volcanic rock. The earthquake strong-motion of the M_w6.3 Ölfus earthquake on 29 May 2008 and 1705 of its aftershocks recorded on the first small-aperture strong-motion array (ICEARRAY I) in Iceland showed consistent and significant variations in ground motion amplitudes over short distances (<2 km) in an urban area located mostly on lava rock. This study analyses the aftershock recordings to quantify the local site effects using the Horizontal to Vertical Spectral Ratio (HVSR) and Standard Spectral Ratio (SSR) methods. Additionally, microseismic data has been collected at array stations and analyzed using the HVSR method. The results between the methods are consistent and show that while the amplification levels remain relatively low, the predominant frequency varies systematically between stations and is found to correlate with the geological units. In particular, for stations on lava rock the underlying geologic structure is characterized by repeated lava-soil stratigraphy characterized by reversals in the shear wave velocity with depth. As a result, standard modeling of HVSR using vertically incident body waves does not apply. Instead, modeling the soil structure as a two-degree-of-freedom dynamic system is found to capture the observed predominant frequencies of site amplification. The results have important implications for earthquake resistant design of structures on rock sites characterized by velocity reversals.

1. Introduction

Iceland is the largest subaerial part of the Mid-Atlantic Ridge where the North American and Eurasian crustal plates are drifting apart with an average rate of approximately 2 cm/year (Fig. 1) [1–3]. Passing across Iceland from south to north, the onshore part of the plate boundary is shifted eastward, resulting in two transform zones: the South Iceland Seismic Zone (SISZ) in the south and the Tjörnes Fracture Zone (TFZ) in the north. The largest and most populous agricultural region in Iceland is located in the SISZ for which the seismic potential and characteristics has been well documented on the basis of historical seismicity. It is known as a region in which destructive earthquakes occur, either as strong single earthquakes or in earthquake sequences of magnitude 6–7 events over a period lasting from weeks to years. The causative faults of strong earthquakes in the SISZ occur as parallel and near vertical north-south striking faults, which is perpendicular to the underlying east-west trending plate boundary [4–13].

Earthquake strong-motion in Iceland has been monitored over the last three decades by the Icelandic strong-motion network (ISMN) which is owned and operated by the Earthquake Engineering Research Centre of the University of Iceland. At present, the network consists of 40 free-field stations that are primarily located in the SISZ and the TFZ, along with several key strong-motion stations in urban centers and key infrastructures such as hydroelectric and geothermal powerplants, dams, hospitals, bridges etc. Additionally, the first Icelandic strong-motion array (ICEARRAY I) was deployed in 2007 in the town of Hveragerði in the SISZ. The ICEARRAY I consists of 13 strong-motion stations with interstation distances of only 50–1900 m [8], in contrast to the more typical ISMN interstation distances of 5–10 km in the

* Corresponding author.
E-mail addresses: sahar@hi.is (S. Rahpeyma), skykkur@hi.is (B. Halldorsson), rugreen@vt.edu (R.A. Green), sigurjon.jonsson@kaust.edu.sa (S. Jónsson).

http://dx.doi.org/10.1016/j.soildyn.2016.07.001
0267-7261/

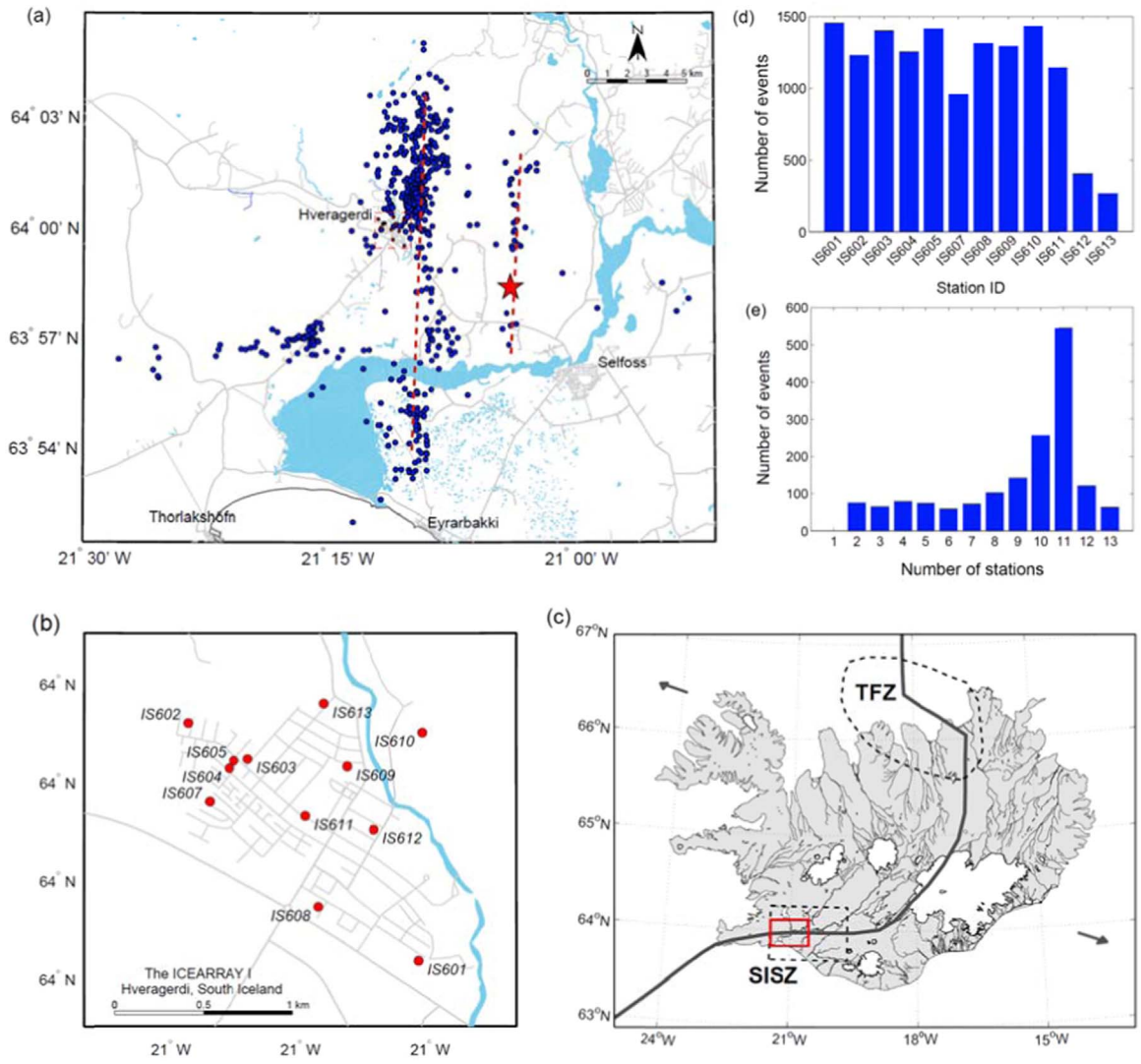

Fig. 1. (a) North–south trending alignments of aftershock locations (blue circles) recorded by ICEARRAY I after the 29 May 2008 Ölfus earthquake (red star) in southwest Iceland indicate the location of the causative faults (approximated by the red dashed lines), The ICEARRAY I stations (red dots) are located within the town of Hveragerði (red dashed rectangle). (b) Map of Hveragerði showing the locations of the twelve ICEARRAY I strong motion stations (red circles) (c) Map of Iceland showing the rough location of the present-day Mid-Atlantic plate boundary (dark solid line), the Tjörnes Fracture Zone (TFZ) and South Iceland Seismic Zone (SISZ). Histograms show (d) the number of recordings at each ICEARRAY I station and (e) the number of events recorded by a given number of stations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

SISZ [14]. During the 29 May 2008 M_w 6. 3 Ölfus earthquake and its aftershock sequence, the ICEARRAY I recorded the strong-motions of the mainshock and 1705 of its aftershocks [8,15]. The main shock recordings were characterized by intense ground accelerations of relatively short durations (5–6 s) and large amplitude near-fault velocity pulses. Despite the relatively small interstation distances of the array considerable variations of earthquake ground motion amplitudes and frequency content were observed. The geometric mean of the horizontal peak ground acceleration (PGA) varied from about 44% to 88% of the acceleration of gravity (g) and peak ground velocity (PGV) from 26 to 62 cm/s [8]. Similar variations of relative amplitudes of the recorded aftershocks have also been reported but not yet fully investigated [15].

The spatial variation in amplitude and frequency content of earthquake ground motions can be attributed to wave propagation effects and localized site effects. During recent decades, it has been recognized that propagation of seismic waves may vary significantly due to local geological and geostructural settings, even over relatively small distances [16,17]. In general, motions recorded on sites classified as "soil" are larger in amplitude relative to those recorded on "rock" sites [18,19]. This is due to impedance contrast where soil deposits acting as filters to incoming seismic waves and amplifying motions at certain frequencies. Consequently, site effects is a major aspect of geotechnical earthquake engineering and has a major influence on seismic hazard [e.g., 20,21–25]. It is noteworthy that in earthquake engineering practice in Iceland, site effects are generally not considered to be a key factor, presumably due to the relatively thin topsoil which is in most cases is easily removed from the uppermost competent rock (e.g., lava rock, hyaloclastite, dolerite, etc.). However, for lava rock the presence of pronounced site effects has been reported [17]. Namely, in geologically younger parts of Iceland the interplay of repeated glaciation/deglaciation and fluctuating sea levels with the primary basaltic volcanism has resulted in the geological profiles consisting of recurring layers of basaltic lavas, as well as tuff layers, often with intermediate layers of sediments or alluvium [26]. This is especially true in the SISZ where the topography is approximately flat and of low elevation, and largely covered with postglacial lava flows underlain by Quaternary sediments of mainly fluvial, glacial, and glaciofluvial origin [27]. In such cases

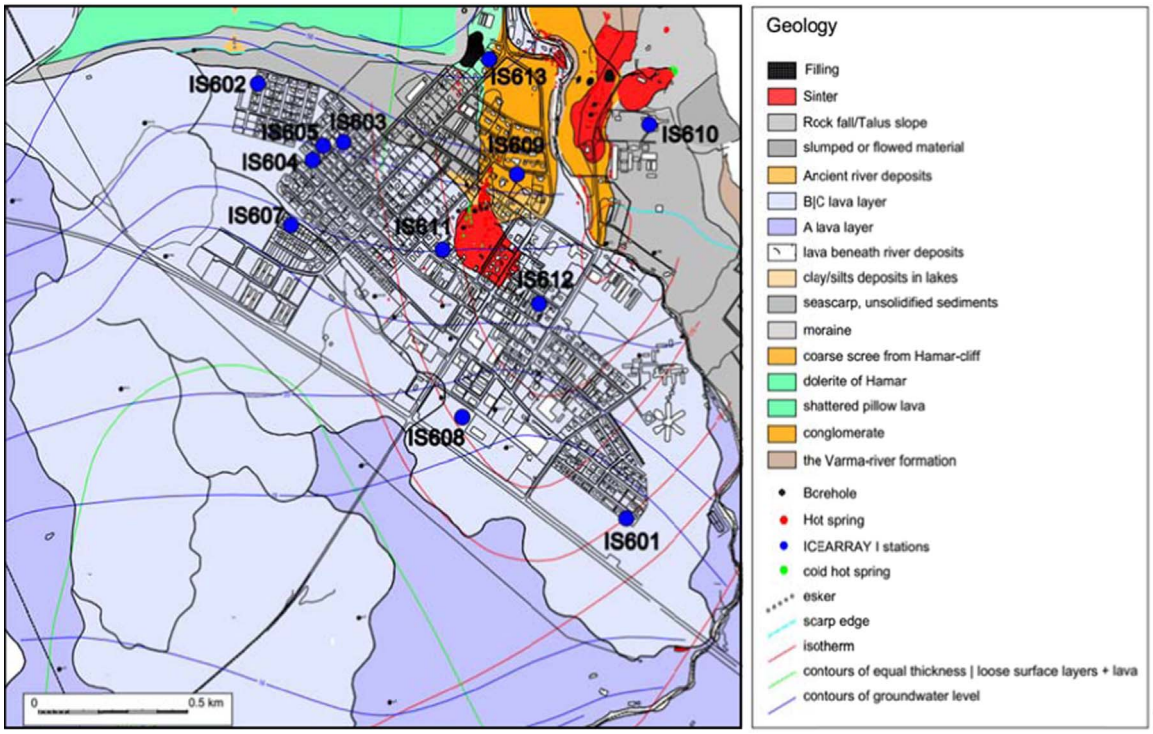

Fig. 2. Geological map of Hveragerði, adopted from [66], showing that most of the town is located on lava. The ICEARRAY I station locations are marked as blue circles. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the lack of consolidation of the sedimentary deposits between the layers of lava rock causes reversal in the velocity profile with depth, leading to significant differences in site response relative to the older bedrock and should therefore not be ignored [17]. We expect this situation to be the case for Hveragerði because based on geological and borehole information the uppermost lava rock layer (B'C in Fig. 2, ~5000 year-old,) lies on top of a softer sedimentary layer, which in turn lies on top of another lava rock layer (A in Fig. 2, ~10,000 year-old) resulting in a velocity reversal [14]. The extensive ICEARRAY I strong-motion dataset from the aftershocks of the Ölfus earthquake now provide the opportunity not only to quantify the local site effects at the array stations on such soil structure, but also the relative differences over short distances.

There are several techniques that can be used to evaluate and quantify site effects [e.g., 21,28] and the choice of the method is usually based on the importance and nature of the project. One of the most popular and widely used techniques to characterize site amplification is the standard spectral ratio (SSR) method [29]. The SSR is defined as the ratio of the Fourier amplitude spectra of an earthquake motion recorded on a soil site to that recorded on a selected reference site. Essentially, the SSR method for estimating site response is based on the comparison of ground motions at soil sites of interest to nearby rock site that is considered a reference motion [29,30]. The result of the SSR method is a site-specific "amplification curve" which is a function of frequency and reveals both the "predominant frequency" of horizontal vibrations of the site, corresponding to the peak in the ratio, and its amplitude. Pragmatically, the SSR technique can only be used for cases where data are available from dense local arrays, to include a station on a reference site that has negligible site response and is located in close proximity to the soil sites of interest. Careful examination of the reference site is required for acceptable amplification estimation at the sedimentary sites [30]. However, finding a proper reference site can be challenging [30–32], and large spatial separation of the soil and rock sites may require correcting the recordings for path and finite-source effects [33]. Therefore, the horizontal to vertical spectral ratio (HVSR) method, which was introduced by Nakamura (1989), is also used herein, along with the SSR method. The HVSR method does not require recordings on a reference site, making it a more convenient approach to estimate site response. The HVSR method entails using the spectral ratio of the horizontal to the vertical component of ground motion [34]. The method, also called Nakamura´s technique, was first introduced by Nogoshi and Igarashi based on the initial studies of Kanai and Tanaka [35–37]. Since then, many investigators worldwide used this relatively easy to implement approach. The fundamental assumption of the HVSR method is that when the soil stratigraphy is comprised of horizontal layers and that the vertical component of the ground motion is free of any kind of influence related to the soil conditions at the recording site. The primary parameter obtained from the HVSR method is the predominant frequency of the soil profile, which corresponds to a peak in the HVSR. The estimate of the predominant frequency is deemed more reliable than the amplitude of the site amplification and is an indicator of a significant velocity contrast at some depth beneath the station that results in the amplification of horizontal ground motion relative to the vertical motion [24,38,39]. Although seismologists still debate the HVSR method's physical and theoretical bases, the approach has attracted the attention of many researchers [40]. Noticeably, the SSR and HVSR methods yield similar estimates of the predominant frequencies of soil profiles, but the amplification values determined by HVSR method are generally less than those determined using the SSR method.

In the present study the strong-motions of the aftershocks of the Ölfus earthquake recorded on the ICEARRAY I as well as recordings of microseismic noise at the array stations have been analyzed using the HVSR method. Additionally, the earthquake recordings were analyzed using the SSR method. The results are presented as frequency dependent amplification curves from which the predominant frequency and the amplitude of site amplification are estimated. The results of the different methods are compared and interpreted in terms of local geological conditions,

ground motion amplitudes, source-site distances and azimuth, and earthquake parameters such as magnitude and depth. Finally, we interpret the HVSR results by considering that a vertical column of a soil structure consisting of a lava rock layer on top of a softer sedimentary layer, representing a velocity reversal with depth, can be modeled as a simple dynamic system.

2. Data

2.1. Array strong-motion data

The 29 May 2008 M_w6.3 Ölfus earthquake occurred on two parallel, near vertical north-south striking right-lateral strike slip faults that are approximately 4.5 km apart (see Fig. 1a). While the epicenter was located on the eastern fault, ~6.5 km S-E of Hveragerði, the majority of the aftershocks occurred on the western fault that lies only 1–2 km from the town and ruptured ~2 s after the eastern fault [11,41]. The ICEARRAY I recorded the strong-motions during the main shock and those from 1705 aftershocks in the region (Fig. 1a) on up to 13 three-component CUSP-3Clp accelerograph instruments [8,15,42]. The recordings of 12 stations are used in the present study because two stations (IS608 and IS688) are collocated and thus provide essentially identical results [8]. Since the array instruments run in triggered mode, the complete strong-motion dataset has been organized and reviewed to ensure the quality of the data, resulting in a dataset of approximately 1600 aftershocks after removing unusable records due to either spurious signals or events recorded on only a few array station. The dataset is characterized by earthquake strong-motion with geometric mean of the horizontal PGA in the range of 3.5~38%g (where g is the coefficient of acceleration due to gravity) from earthquakes of magnitudes in the range 0.42~4.75 at epicentral distances of ~1.6–15 km. The aftershocks occurred between 29 May 2008 and 10 May 2009, after which seismicity in the region has been relatively low.

The earthquake parametric information was obtained from the SIL seismic network of the Icelandic Meteorological Office which monitors the seismicity of Iceland [43], and the temporary LOKI seismograph network which was deployed in the macroseismic region within two days of the main shock to provide more accurate hypocentral locations of aftershocks [44]. Unfortunately, however, only 700 of the aftershocks recorded by the ICEARRAY I were found to match with events reported by either SIL or LOKI networks. For the other aftershocks, the source-site distances were estimated from P- and S-phase arrival times and their local magnitudes estimated using an empirical relationship for peak ground velocity [45]. Fig. 3 shows the ICEARRAY I database characteristics from parametric earthquake data and the identification number, name, and location of the array stations shown in Fig. 1 are listed in Table 1.

2.2. Microseismic data

The spectral analysis of microseismic (ambient) vibrations via the HVSR method is an alternative approach to estimate the site response in urban environment. Microseismic noise is defined as

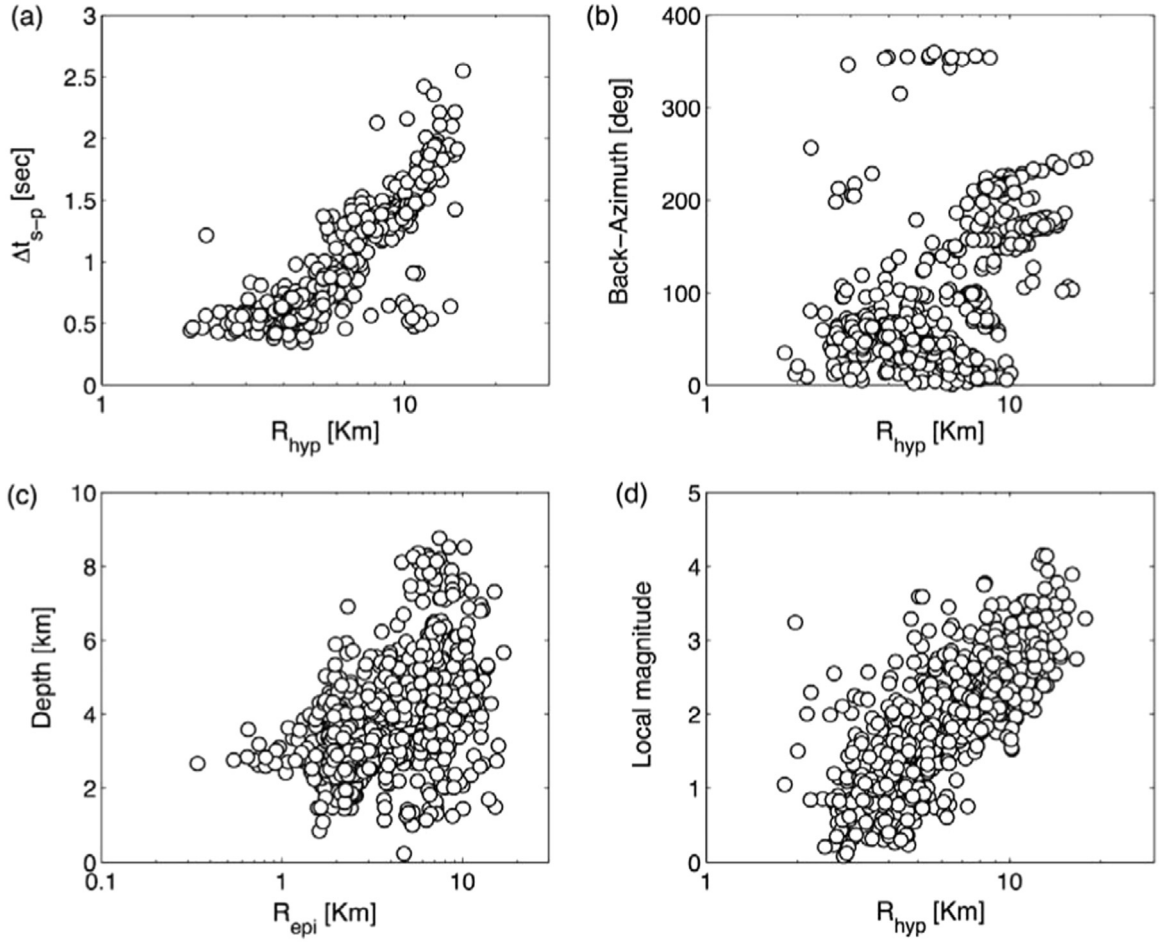

Fig. 3. Different earthquake parameters plotted vs. source-site distance from strong-motion station IS605: (a) difference between P- and S-wave arrival time, (b) back azimuth (in degrees clockwise from north), (c) depth, and (d) local magnitude.

Table 1
ICEARRAY I strong-motion array station locations.

	Station ID	Station name	Latitude (°N)	Longitude (°E)
ICEARRAY I	IS601	Heidarbrun 51	63.9927	−21.1776
	IS602	Kambahraun 39	64.0047	−21.2043
	IS603	Dynskogar 3	64.0029	−21.1974
	IS604	Borgarhraun 12	64.0024	−21.1995
	IS605	Borgarhraun 8	64.0028	−21.1990
	IS607	Arnarheidi 26	64.0007	−21.2018
	IS608	Sunnumork 2	63.9954	−21.1893
	IS609	Dvalarheimilid As	64.0025	−21.1859
	IS610	Reykir	64.0042	−21.1772
	IS611	Heidmork 31	64.0000	−21.1908
	IS612	Reykjamork 17	63.9993	−21.1828
	IS613	Laufskogar 39	64.0057	−21.1886

low amplitude vibrations from surface sources randomly distributed in space and time (e.g., vibrations of natural origin such as wind and sea tides, or of manmade origin such as traffic, industrial machinery, etc.). During recent decades, many researchers have investigated the reliability of using ambient noise, both numerically and experimentally, for the quantifications of site effects [46–48]. In the absence of earthquake strong-motion recordings, microseismic data is easily obtained and can provide additional constraints and spatial resolution of site effects via the HVSR method.

Continuous ambient noise recordings of a minimum one-hour duration were performed at the ICEARRAY I sites using REF TEK 130-01 Broadband Seismic Recorders and Lennartz LE-3D/5s three-component sensors for which gain level and sampling rate were configured to unity and 100 Hz, respectively. The microseismic measurements across ICEARRAY I stations were carried out during October and December 2012 in several subarray formations, recording at three to six stations simultaneously for several hours to several days.

2.3. Data analysis

Each earthquake strong-motion record from a given station is of limited duration centered on the actual earthquake strong-motion, with a pre-and post-event buffer of 30–60 s typically. The array recordings of each event in the dataset have been trimmed appropriately to limit the extent of noise. As a result, each array recording consists of 10–12 stations, with the first sample of each component record having the same absolute time and duration. For each event-station pair, we calculated the absolute Fourier spectrum of each of the three components over the duration of the record. The spectral amplitudes were smoothed using the Konno and Ohmachi with a smoothing coefficient of $B = 20$ [38] for the selected bandwidth. A single smoothed spectrum representing horizontal ground motions was obtained by calculating the geometric mean of the two smoothed horizontal spectra. Dividing the spectrum for the horizontal motions by the spectrum of the vertical component produced the HVSR curve as a function of frequency for each event-station pair. As we have multiple earthquake recordings at the same station, we calculated the average HVSR and the corresponding standard deviation as a function of frequency, thus producing the station-representative HVSR from the earthquake strong-motion records. This approach is also the recommended method by SESAME and is the most commonly used method for HVSR analyses [see e.g. 20,28]. We tested the sensitivity of the HVSR results in several ways. First, the smoothing coefficient value was determined on the basis of sensitivity analysis to ensure that sufficient detail was preserved in the resulting spectral ratio. We also compared the HVSRs obtained for different time windows using both the window length from S wave arrival time to the end of the coda waves and the entire length of each recording. No significant differences were observed, and as a result, the spectral ratios presented in this study were calculated over the duration of each record. In calculating spectral ratios we used the arithmetic, geometric, and the quadratic/squared mean methods for combining the two horizontal components. Comparison of the spectral ratios showed insignificant differences, and therefore, the geometric mean method was applied in this study.

3. Results

Despite small aperture of ICEARRAY I there are noticeable differences in the HVSR results across the array. The HVSR amplification curves for all earthquake recordings for each of the twelve ICEARRAY I strong-motion stations were calculated over the frequency range of interest to this study (0.5–20 Hz), along with the geometric mean HVSR and the associated $\pm 1\sigma$ (Fig. 4). Some stations exhibit bimodal amplification curves with one mode being more dominant and of relatively larger amplitude than the other (e.g., IS605, IS604, IS608) while other stations have a single narrow-band peak of relatively low amplitudes (e.g., IS602) or amplification curves of relatively high amplification over a wide frequency range (e.g., IS601, IS603), or even very low and uniform amplification curves across the frequency range (IS609 through IS613). One station (IS607) exhibits unreliable results at lower frequencies, believed to be contaminated by a possibly faulty vertical component because it contains anomalously large low-frequency content. These results indicate that significant site response variation exist on a spatial scale of less than 2 km within the town of Hveragerði.

To implement the SSR method the ratio of the geometric mean of the Fourier amplitude spectra for the horizontal components of motions recorded at a site of interest to that of a reference site was computed. While the geology of the town of Hveragerði is somewhat complex (see Fig. 2), all of the stations lie on what is classified as “rock” [8]. Therefore, all sites were viewed as potential reference sites in the context of the SSR method. On the basis of the HVSR results (see Fig. 4), relatively low and approximately uniform (across the frequency range considered) amplifications were observed at stations IS609-IS613, making them potential candidates for a reference site. Station IS612 was excluded since it is located on a relatively young lava-rock, while stations IS609, IS610 and IS613 are located on a considerably older, and presumably more stable, bedrock. Stations IS610 (located on a hillside) and IS613 (relatively few data and unstable HVSR) were excluded, leaving station IS609 as the selected reference station for the SSR method; this station has been in operation since 1999 as a permanent station of the Icelandic strong-motion network.

The results of the SSR method using IS609 as the reference station are shown in Fig. 5, along with the results from the HVSR method for comparison. In general, the results from the two methods are in good agreement. At stations IS601-IS607 the SSR results not only reproduce the fundamental frequencies of the HVSR results, but the HVSR amplification curves are generally within the 68 percentile limits of the SSR results across the frequency range considered. We note that while the bimodal HVSR amplification curve at station IS605 is not reproduced by the SSR curve, the latter absolutely confirms the overall amplification effects by exhibiting a broad amplification curve covering the peaks at both predominant frequencies. We suspect that slight variations in the vertical spectrum at the reference station is likely the cause of the peaks not being reproduced. We note that while the HVSR and SSR results appear to diverge at around 10 Hz for stations IS611-IS613 the amplifications are low (near unity) and those discrepancies are associated with the apparent “signature” peak of

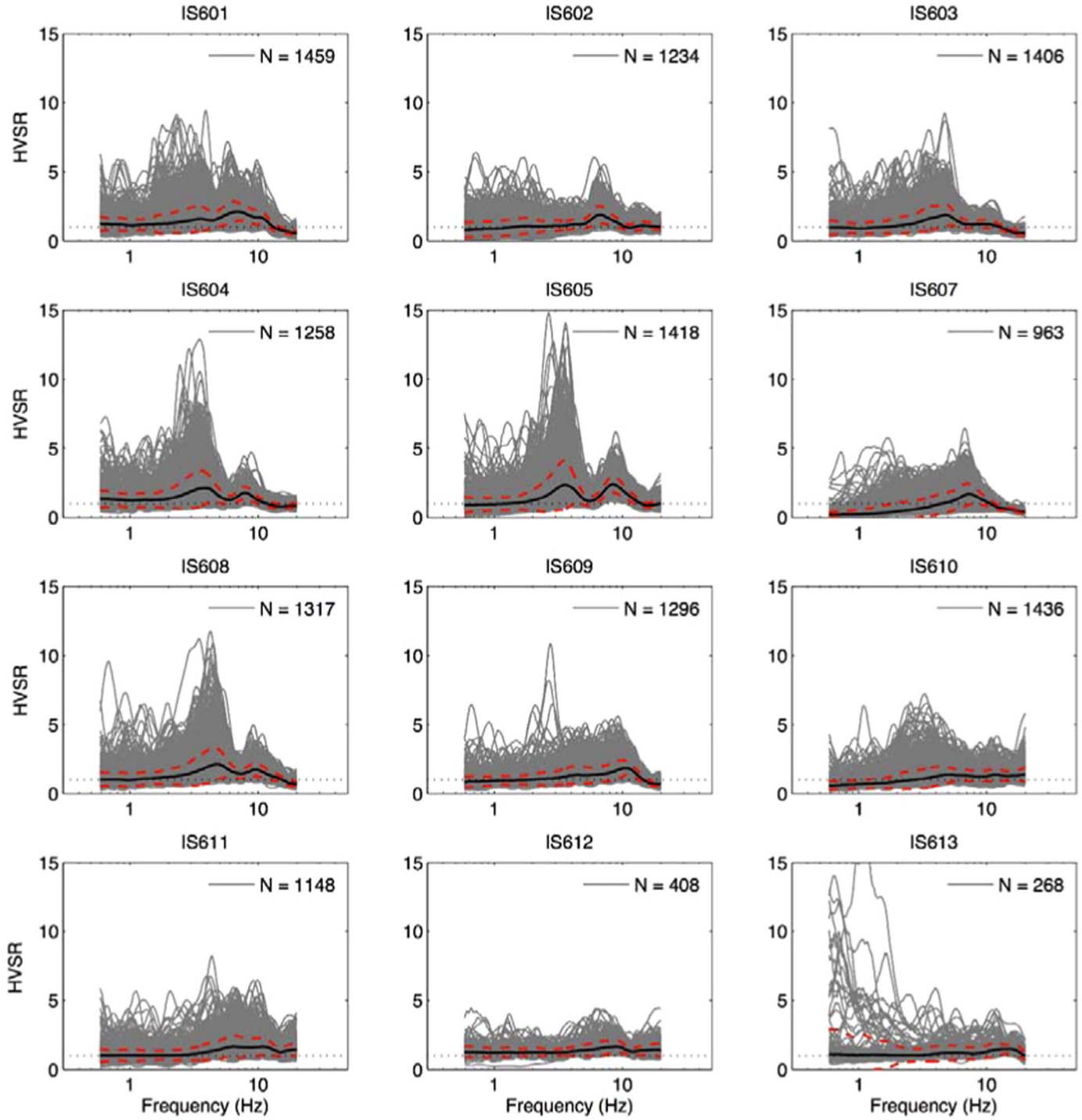

Fig. 4. The earthquake recordings (gray lines) as a function of frequency (Konno and Ohmachi smoothing, coefficient B = 20) for each of the twelve ICEARRAY I strong-motion stations. The geometric mean HVSR (black solid lines) and their corresponding $\pm 1\sigma$ (standard deviation, red dashed lines) are also shown, with N the number of available earthquake events used to derive the mean. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the IS609 HVSR results at that frequency. Finally, the SSR results for station IS607 essentially confirm our suspicion that a faulty vertical component is affecting the HVSR results, and reveal a clear peak between 6 and 7 Hz.

In order to empirically assess the reliability of the HVSR results and at the same time the applicability of HVSR results from microseismic measurements, the same HVSR procedure was applied to the ambient noise measurements. First however, since the microseismic noise is not necessarily random in time and place due to man-made disturbances in the town and environmental factors (e.g., temporary weather/storm vibrations), an additional sensitivity study was carried out. After reviewing the data and removing segments containing spurious spikes the HVSR was calculated for each station for different times of the day, night and week, and for various durations ranging from several minutes to several hours. The analysis showed that a stable HVSR at each stations was obtained for the optimal window length of 20-min [49]. Therefore, for sites where long recordings were available the most stable ones (insofar as being relatively free of spurious signals, traffic and obvious man-made temporary disturbances) were

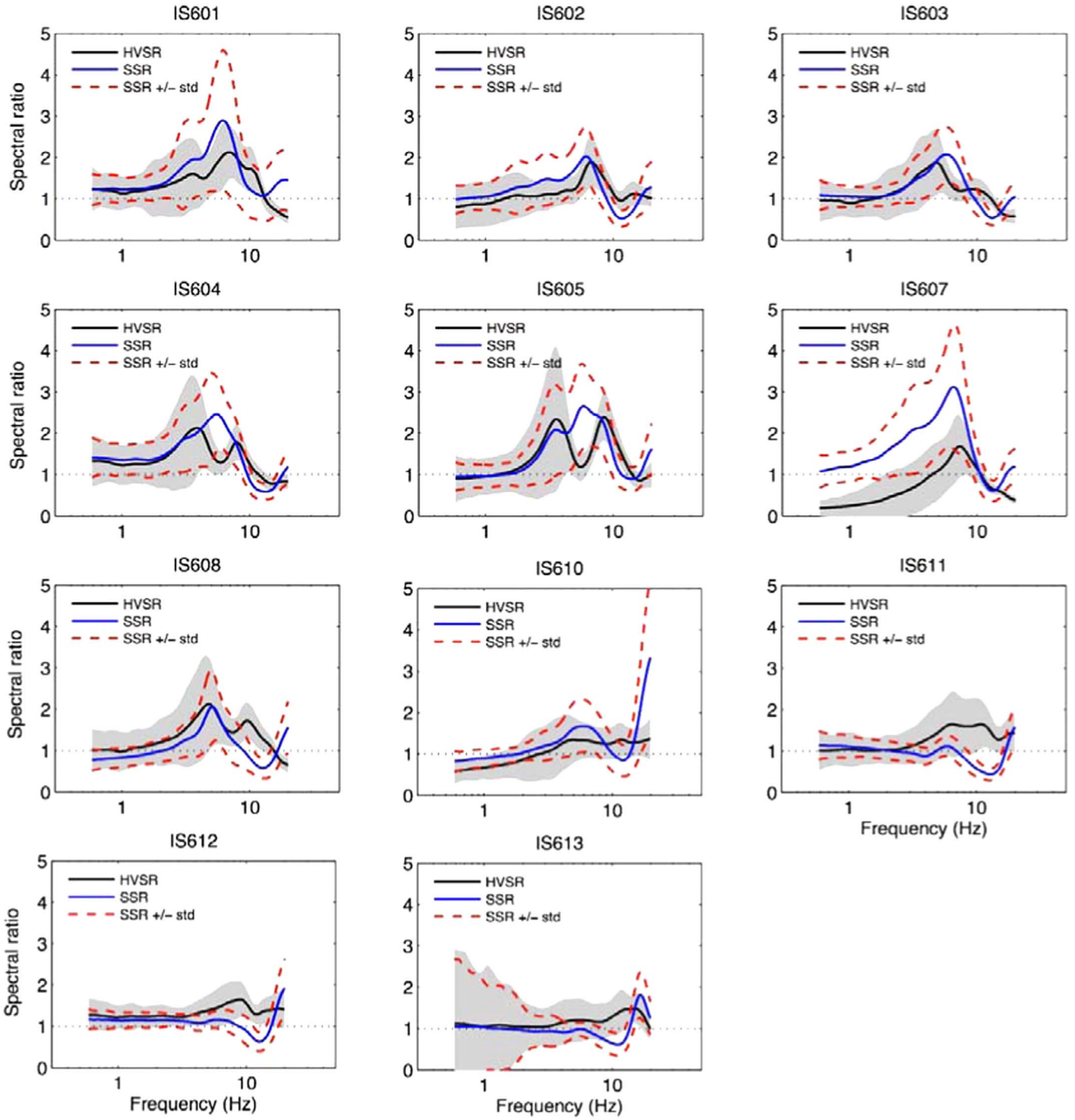

Fig. 5. Comparison of the mean HVSR (black lines) and mean SSR (blue lines) determined from the earthquake data for each of the ICEARRAY I strong-motion stations, using station IS609 as a reference site. Standard deviations of the means are indicated by red dashed lines for SSR and with gray shaded areas for HVSR. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

split up into multiple and unique parts of 20 min. Subsequently, the HVSR was calculated for each segment, and the mean HVSR and corresponding standard deviation were calculated as a function of frequency, producing the station-representative HVSR from microseismic data.

A comparison is shown in Fig. 6 of the mean HVSR results from earthquake recordings with the mean HVSR results from microseismic data. The agreement in terms of overall shape of the amplification curves and their amplitudes is remarkable at almost all stations. The comparison seems to confirm results reported by many studies in other regions [e.g., 50,51] that microseismic data and HVSR analysis may be used with confidence to map the overall amplification characteristics of ICEARRAY I.

4. Discussion

4.1. Spatial distribution of HVSR characteristics

To provide further insight into the HVSR results, the spatial

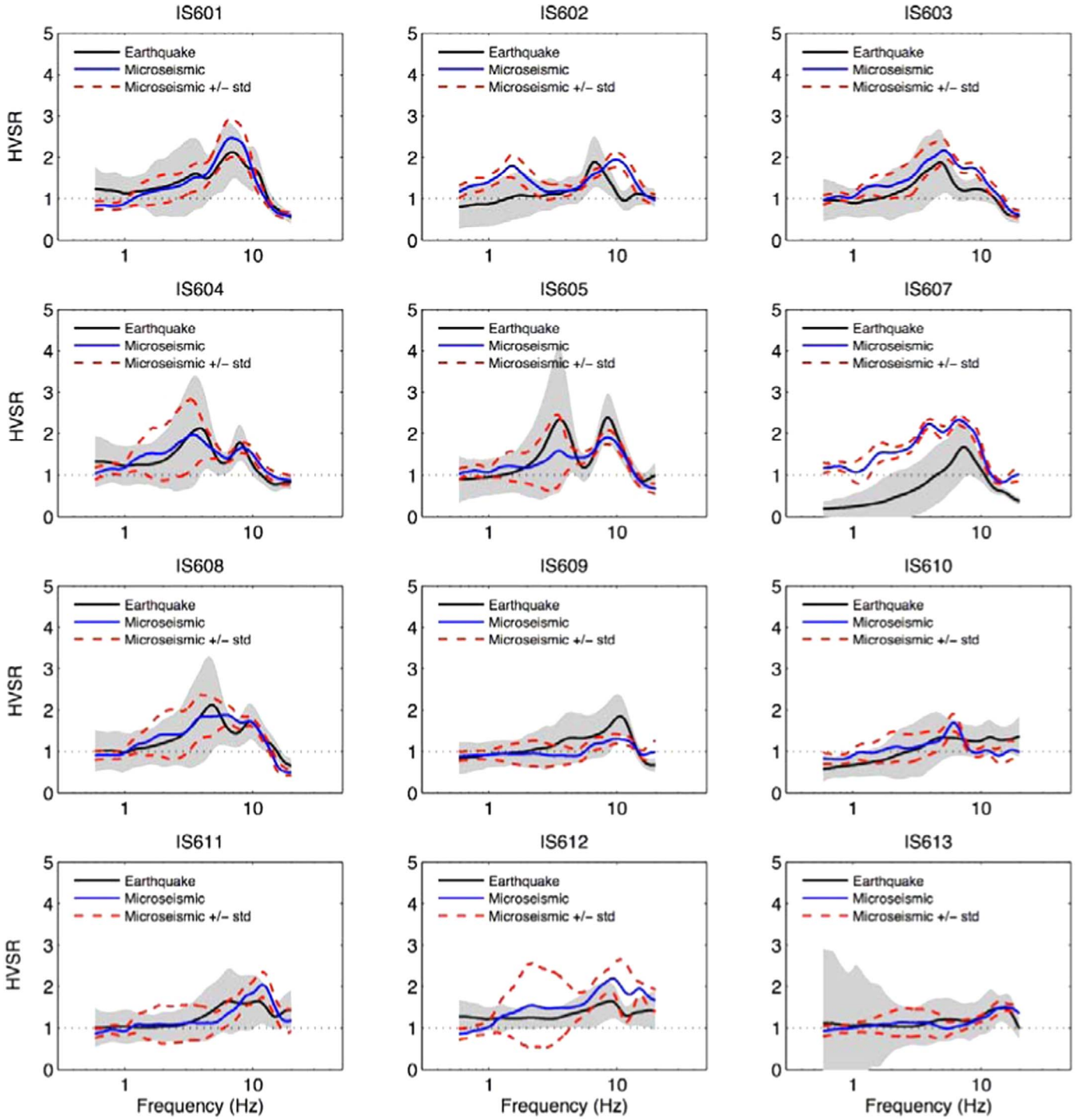

Fig. 6. Comparison of the mean HVSR estimated from earthquake (black lines) and microseismic (blue lines) data for the twelve ICEARRAY I stations. Standard deviations $\pm 1\sigma$ are shown with gray shaded areas (earthquake HVSR) and red dashed lines (microseismic HVSR). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

distribution of fundamental frequencies and maximum amplitudes of the amplification curves based on earthquake and microseismic data, respectively, are shown as surface plots in Fig. 7 over the area confined by the array stations. While the slight differences between the HVSR characteristics between stations may become more pronounced when presented spatially in this way, both methods show the same general trend that from north to south amplification tends to increase and predominant frequency decrease. When focusing on the distribution of predominant frequency as revealed from earthquake data, the results show a general northeast-southwest trend of decreasing peak predominant frequency. [Parenthetically we note from microseismic data that the difference between the spatial distribution of predominant frequency is mainly due to station IS611 having a significantly higher (above 10 Hz) and IS610 having lower predominant frequency]. The almost linear NW-SE pattern shown in Fig. 7 is meaningful because it is in accord with a major geological transition in the area. Namely, stations IS609, IS613, and IS610 are located on older bedrock (as mentioned previously) while the remaining stations are located on a post-glacial lava rock layer on

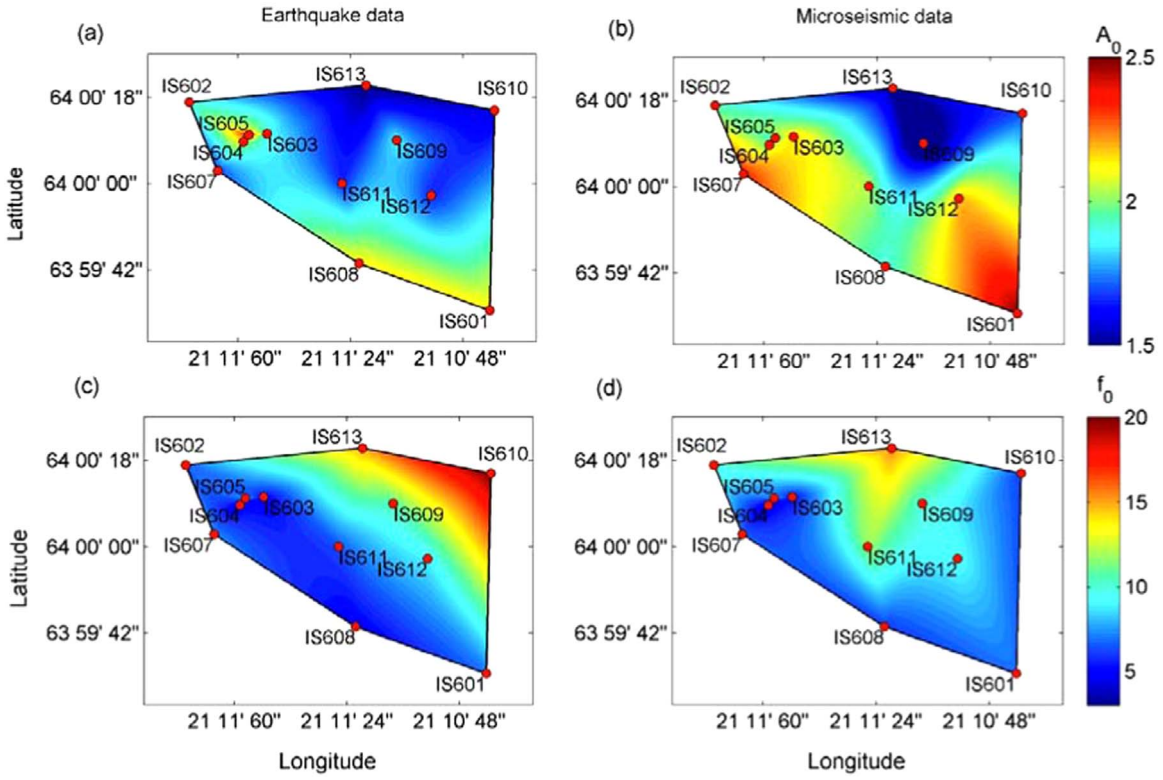

Fig. 7. Spatial variability of local site effects in Hveragerði with the ICEARRAY I stations marked as red dots. Amplification factor A_0 estimated from (a) earthquake recording and (b) microseismic data, and fundamental frequency f_0 from (c) earthquake recordings and (d) microseismic data. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

top of a sedimentary layer, according to logs of shallow boreholes in the town. The latter geologic profile is a typical layering encountered primarily in geologically younger parts of Iceland in the vicinity of the volcanic zones. In Hveragerði, the uppermost lava layer (BlC, see Fig. 2) is relatively young basaltic lava of varying thickness (on the order of 10–30 m based on borehole data). Lava BlC lies on top of a sedimentary layer (of similarly varying thickness), which in turn lies on top of an older lava layer (A). No information about the thickness of the lower layer exists, but it flowed from the same volcanic fissure, and based on the spatial extent of the lava it is likely of similar thickness as Lava B=C. From the typical layered structure of young geological formations in Iceland, is likely that it is underlain by sediments, introducing the second velocity reversal with depth, above the bedrock.

The results of this study show that for stations located on the younger formation the amplification curves are characterized by lower predominant frequencies and larger relative amplitudes than those on older bedrock. Moreover, the bimodal amplification curves (seen particularly in earthquake data) are only seen at stations on this formation and clearly represent the complex structure of repeated lava-sediment layers. This is especially apparent for station IS605 where the amplification curve is characterized by two clear predominant frequencies and relatively large amplitudes, which clearly suggests the presence of large and sharp velocity contrasts underneath the station. The same characteristic, albeit slightly less pronounced, can be observed at IS604 which is located 54 m southwest from IS605, indicating that the substructure is similar under the two stations. When compared with the amplification curve of station IS603 (located 80 m from IS605 and 114 m from IS604, due east), however, the difference indicates that the velocity contrast under the station is much less abrupt compared to IS603 and especially IS605. These amplification curves are in stark contrast with those at the bedrock stations, which are characterized by high frequency peaks (~10 Hz) of very low amplitudes.

4.2. Earthquake parameters vs. HVSR characteristics

The standard deviation of the HVSR amplification curves from earthquake data in Fig. 5 are similar at most stations and relatively constant over the frequency range. Notable exceptions from this are amplification curves at stations IS604, IS605, and IS608 which exhibit bimodal amplification curves. For station IS605 for example, there is considerably greater scatter in the bimodal amplification curve for the lower-frequency peak but less for the higher-frequency peak. This scatter is in turn reflected in the associated standard deviation of amplification around ~3–4 Hz (dashed lines in Fig. 5). It is of interest that while both peaks have very similar mean amplification, individual amplifications can be much higher at the lower-frequency peak, while the higher-frequency peak at ~8–9 Hz is associated with much less scatter.

Moreover, in many cases individual amplification curves exhibited different trends in the amplification, warranting further inspection of the HVSR curves for each event for azimuthal, distance, and magnitude dependence. Using the bimodal character of the amplification curve of station IS605 as an example, we start by grouping the aftershocks shown in Fig. 8a into two classes, those associated with the peak amplification at frequencies less than (gray, see Fig. 8b) and higher than (blue, see Fig. 8c) 5 Hz.

The spatial distribution of aftershocks outlines the two N-S trending causative earthquake faults and a third (non-causative) fault segment trending E-W. From Fig. 8a it appears that the two groups of aftershocks are distributed more or less evenly between the fault structures. In more quantitative terms, as shown in Fig. 8g, it is clear that no azimuthal dependence of predominant frequency exists (in other words, the null hypothesis of the aftershock groups being two independent random samples is rejected at the 5% significance level). Similarly, no correlation was found between predominant frequency and earthquake depth (not shown), and both groups of earthquakes have similar coverage of

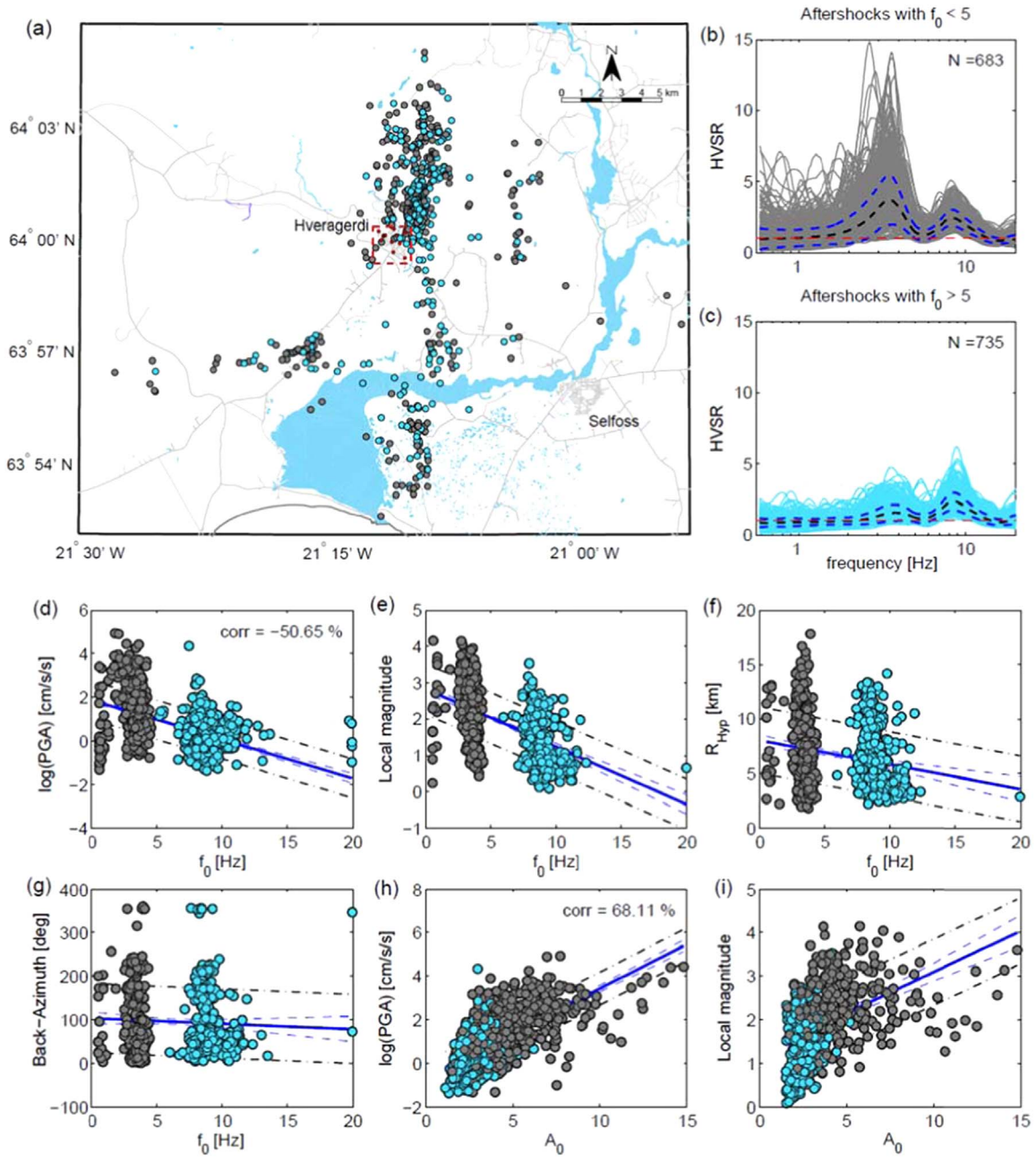

Fig. 8. (a) Map of ~630 aftershock locations recorded by ICEARRAY I grouped according to the predominant frequency range above (blue) and below (gray) 5 Hz, using station IS605 located in the town of Hveragerði (red dashed rectangle). The HVSR for the two groups of aftershocks is shown in (b) and (c) as well as the mean HVSR $\pm 1\sigma$, with N the number of available earthquake. Also shown the (d) peak ground acceleration (PGA), (e) local magnitude, (f) hypocentral distance and (g) back-azimuth vs. the predominant frequency f_0 (Gray: $f_0 \leq 5Hz$ and cyan: $f_0 > 5Hz$), and (h) PGA and (i) event magnitude vs. Amplification factor A_0. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

hypocentral distances (Fig. 8f). On the other hand, there is clearly a strong correlation between earthquake magnitude and the associated predominant frequency (Fig. 8e); this is in turn reflected in a strong correlation with peak ground acceleration. Thus, as can be seen when comparing Fig. 8d and e, larger magnitude earthquakes appear to produce a bimodal amplification curve at station IS605, where the peak at (lower) 3–4 Hz predominant frequency dominates the peak at (higher) 8–9 Hz predominant frequency.

On the contrary, smaller magnitude earthquakes produce relatively constant and low peak amplification at higher frequency (8–9 Hz) (Fig. 8h and i). Interestingly, almost all earthquakes that are associated with the amplification peak at higher frequency have amplitudes less than ~3 (Fig. 8h and i). In other words, the peak in the HVSR amplification curve at lower predominant frequency is only clearly observed for earthquakes of relatively larger magnitudes, and in the vast majority of the cases, they are associated with maximum HVSR amplitudes higher than 3 and with much greater variability in amplitudes.

In the frequency domain the total spectrum of earthquake ground motion can be considered as the result of the combined effects of the earthquake source spectrum, propagation path effects, and local site effects. The results indicate therefore that earthquake source parameters that scale with earthquake size are responsible for the correlation of the lower-frequency peak with relatively larger magnitude earthquakes compared with the higher-frequency peak. To a first order approximation we assume

that an earthquake having a seismic moment $M_\circ$ may be modeled as a symmetric circular crack of diameter $2\rho_0$ on which a local stress drop $\Delta\sigma$ takes place as it ruptures radially and radiates seismic waves. The earthquake source acceleration spectrum [52] of the radiation may for simplicity be expressed by the "ω-square" spectrum [53] with a corner frequency f_0 that scales as

$$M_\circ(f)=(2\pi f)^2\frac{M_\circ}{1+\left(\frac{f}{f_0}\right)^2} \tag{1}$$

where

$$M_\circ=\frac{16}{7}\Delta\sigma\rho_0^3 \tag{2}$$

and f_0 is related to the crack radius ρ_0 by

$$f_0=\frac{C_s\beta}{2\pi\rho_0} \tag{3}$$

where β is the shear wave velocity in the vicinity of the source, and C_s is an increasing function of the ratio v/β where v is the rupture velocity inside the crack [54,55]. The spectral amplitudes of acceleration are highest (and constant) above f_0, while below f_0 they decrease rapidly with decreasing frequency. For a constant stress drop the physical dimensions of the crack ($2\rho_0$) increases with seismic moment, which in turn results in lower corner frequency. For example, assuming a realistic value of the stress drop of 130 bars [9], $v/\beta = 0.75$ and $\beta = 3.2$ km/s, the corner frequency of a magnitude 3.5 earthquake is 4.7 Hz, while that for a magnitude 2.5 earthquake is 14.9 Hz. An earthquake larger by one-magnitude unit, which is approximately the difference in mean magnitudes between the two groups (Fig. 8e), has on the average three-times larger spectral amplitudes at frequencies of 3–4 Hz than the smaller earthquake. Therefore, we can expect the larger earthquakes to have significantly higher seismic energy at those frequencies which may be enough to induce horizontal vibrations, something that the smaller events, due to higher corner frequency, are much less likely to induce. The observed scatter of the HVSR curves at the lower predominant frequency is most likely a manifestation of this disparity of spectral amplitudes. Finally, the range of magnitudes in the two groups shown in Fig. 8e is most likely attributed to the natural variation of the main earthquake source parameters ($\Delta\sigma$) and the inherent uncertainty in magnitude determination.

4.3. HVSR from earthquake vs. microseismic data at station IS605

It is of particular interest that when comparing the HVSR results from earthquake data with microseismic data at station IS605 that two distinct features stand out: (1) anomalously large scatter at the lower predominant frequency is observed in both the earthquake and microseismic data, and (2) only the amplification peak at the higher predominant frequency is captured by the microseismic data. Due to small epicentral distances and high apparent velocity over the array, the earthquake waves consists exclusively of body waves with near vertical incidence angles while the microseismic noise is generally assumed to be comprised mostly of ambient surface waves [34,56,57]. However, recent studies have shown that modeling the complete seismic wavefield gives numerical results that are in better overall agreement with empirical HVSR results over the frequency range of interest, as compared to exclusively using the surface or body wave approximation. The body wave approximation was reported to give better estimates around the predominant frequency of the sedimentary layer while the surface wave approximation gave better results at higher frequencies [58].

We note however that most of the results in the current study are obtained on sites for which the site profile is characterized by velocity reversals. In such cases researchers have reported significant differences between the HVSR results from earthquake vs. microseismic data [59] (albeit for a site characterized by a thin, shallow and dense anthropogenic layer on top of very thick clay layer).

In our case, the velocity reversal is due to softer sedimentary layers sandwiched between the lava layers of similar thicknesses. Such conditions have been shown to produce different HVSR results from earthquake data compared to those recorded on bedrock [17]. It is therefore likely that for station IS605 the velocity reversal is considerable and abrupt and that the earthquake waves are sampling deeper parts of the subsoil layers which microseismic wave field generated on or near the surface is not able to reach (except for the part of the wave field that is comprised of body waves). In general, the HVSR results from earthquake data for the ICEARRAY I stations on lava rock are more or less reproduced by microseismic data, both in overall amplitude and predominant frequency, which may actually indicate that the velocity contrasts (and reversals) in general do not vary laterally. Additionally, according to borehole data lava and sedimentary layer thicknesses vary considerably over the study area, which is most likely the cause of the relative differences in HVSR results among the stations. It appears, however, likely that the failure of the amplification curve from microseismic data to capture the predominant frequency and amplification at station IS605 (and to a similar but lesser extent IS604 and IS608) is due to an anomalously sharp onset of the velocity reversal in the stack of lava and sedimentary layers at that site.

4.4. Modeling the bimodal site response at station IS605

The shape of HVSR curves are connected to the impedance contrast between layers [16] and a sharp peak is indicative of high impedance contrast between a softer layer overlaying a harder layer [57,59]. For this reason, we physically interpret the two peaks of the amplification curve from earthquake data at station IS605 as coming from two considerable and abrupt velocity contrasts at depth, indicating that a multilayer subsoil model is needed for modeling.

In general, inversion of physical parameters of the subsoil profile on the basis of HVSR is carried out in the context of the body wave approximation [58,60]. Essentially, the body wave approximation is based on computing the transient response of a set of horizontally stratified, linearly elastic layers overlying a uniform half-space and excited by vertically incident, transient plane waves. In this study we apply the theoretical transfer function of the layered soil model based on the fast recursive algorithm proposed by Tsai [61], modified considering frequency dependent attenuation and body wave dispersion [60], thus modeling the theoretical HVSR at the surface as the ratio between transfer functions for the vertically propagating incident S-waves (horizontal components) and P-waves (vertical component) [58,60]:

$$HVSR(f)=\frac{F_S(f)}{F_P(f)} \tag{4}$$

The transfer functions $F_c(f)$ for S- and P-waves (c refers to body wave phase, i.e. S or P), used to calculate the theoretical HVSR, can be modeled analytically as [58].

$$F_c(f)=\left\{\cos\left[2\pi f\frac{H}{V_{c,a}(2\pi f)}\right]+i\frac{\rho_a V_{c,a}(2\pi f)}{\rho_b V_{c,b}(2\pi f)}\sin\left[2\pi f\frac{H}{V_{c,a}(2\pi f)}\right]\right\}^{-1} \tag{5}$$

where f is the frequency of the wave, H is the soft layer thickness, V_c is the complex velocity of phase c and indices s and r respectively refer to the soft layer and bedrock (for one layer model of

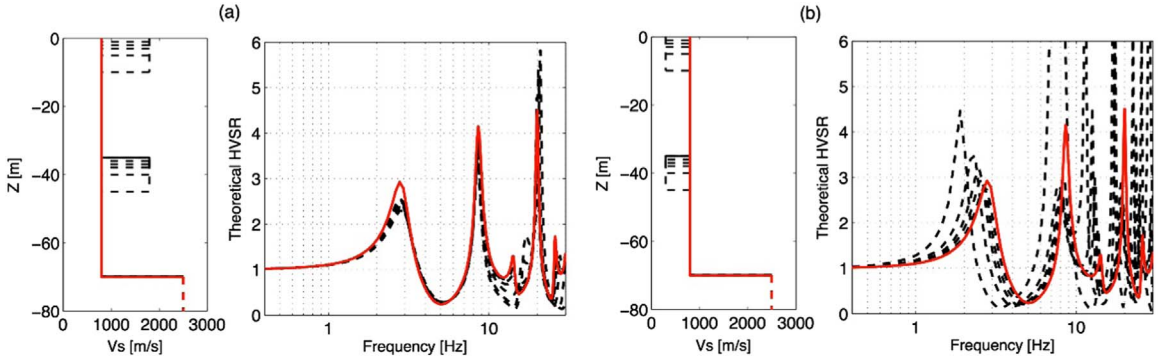

Fig. 9. The effects of velocity reversals on wave propagation and soil amplification. Inserting two layers with (a) higher velocities (b) lower velocities. Sub-soil profile system over the half-space is also shown.

soil overlaying the bedrock). The quality factor (Q_c) is used to account for damping:

$$V_c(2\pi f)=\frac{V_c^e}{1-\frac{1}{\pi Q_c}\log\left(\frac{f}{f_{ref}}\right)}\left(1+i\frac{1}{2Q_c}\right) \tag{6}$$

where V_c^e is the elastic velocity of the body waves, and f_{ref} is a reference frequency (considered to be unity in this study) [55]. However, attempts at reproducing the bimodal amplification curve at station IS605 using this approach were largely unsuccessful. For reasonable values of the density and shear wave velocity of the lava and sedimentary layers, and using layer thicknesses from borehole logs, the method fails to reproduce two physical peaks at the observed predominant frequencies. In fact, adding two layers of hard rock to the top and in the middle of the soft soil stratum (70 m of sedimentary layer) does not significantly change the predominant frequency or the amplification curve (see Fig. 9a). On the contrary, Fig. 9b shows that inserting two soil layers with lower impedance ratios dramatically affects the respective amplification. The amplification curve is thus dominated by the deepest velocity contrast (the underlying bedrock assumed as half-space and an overlaying sedimentary layer) and the reverse velocity contrast between the hard layers and the intermediate stiff layers is essentially invisible to the method. As a result, we find that the body wave assumption in interpreting the HVSR results from data recorded on a profile with velocity reversals may lead to ambiguous interpretation. Instead, we model the dynamic response of such a profile as a classically damped linear oscillator subjected to a base excitation [62–64]. For the bimodal amplification curve at station IS605, a two-degree of freedom (2DOF) linear oscillator may be used. The nodal displacement vector $\boldsymbol{u}(t)$ of the system can be expressed in terms of modal coordinates by using the expansion theorem for multi-degree-of-freedom (MDOF) systems (modal superposition)

$$\boldsymbol{u}(t)=\sum_{n=1}^{N}\boldsymbol{u}_n(t)=\sum_{n=1}^{N}\boldsymbol{\Phi}_n q_n(t),\ N=2 \tag{7}$$

where $q_n(t)$ are the modal coordinates, and $\boldsymbol{\Phi}_n$ are the natural mode shapes. The undamped modal frequencies ω_n can be obtained by solving the eigenvalue problem $\left(\boldsymbol{K}-\omega_n^2\boldsymbol{M}\right)\boldsymbol{\Phi}_n=0$) where $\boldsymbol{K}$ and $\boldsymbol{M}$ are the stiffness and mass matrices, respectively [62]. In general, the equation for a damped system can be written as

$$\boldsymbol{M_n}\ddot{\boldsymbol{q}}_{\boldsymbol{n}}(\boldsymbol{t})+\boldsymbol{C_n}\dot{\boldsymbol{q}}_{\boldsymbol{n}}(\boldsymbol{t})+\boldsymbol{K_n}\boldsymbol{q}_{\boldsymbol{n}}(\boldsymbol{t})=-m\iota\ddot{u}_g(t) \tag{8}$$

where $\boldsymbol{M}_n$, $\boldsymbol{C}_n$, and $\boldsymbol{K}_n$ are the generalized mass, damping and stiffness matrices, respectively, and ι is the influence vector representing the mass displacements from static application of a unit-ground displacement. We assume that the basement rock excitation is of the form $\ddot{u}_g(t)=\ddot{u}_{go}(\omega)e^{i\omega t}$, where $\ddot{u}_{go}(\omega)$ is the frequency response of the excitation and ω is the forcing frequency; the same functional form for the modal coordinates is also assumed. Upon insertion into Eq. (8) the displacement frequency response of interest (i.e., top of the soil structure, denoted as $u_{2o}(\omega)$) becomes

$$u_{2o}(\omega)=-\left(\frac{\Gamma_1}{\omega^2\left[\left(\left(\frac{\omega_1}{\omega}\right)^2-1\right)+i2\xi_1\left(\frac{\omega_1}{\omega}\right)\right]}+\frac{\Gamma_2}{\omega^2\left[\left(\left(\frac{\omega_2}{\omega}\right)^2-1\right)+i2\xi_2\left(\frac{\omega_2}{\omega}\right)\right]}\right)\ddot{u}_{go}(\omega) \tag{9}$$

where for the n-th mode of the 2DOF system ξ_n is the damping ratio, ω_n is the undamped natural circular frequency, and Γ_n is the modal participation factor. Eq. (9) can be written as $u_{2o}(\omega)=H(\omega)\ddot{u}_{go}(\omega)$, where $H(\omega)$ is the displacement transfer function.

In order to calculate the acceleration response $\ddot{u}_{2o}(\omega)=\omega^2 u_{2o}(\omega)$ from Eq. (9) we have to estimate the parameters of the system. The lava layers are assumed to be rigid masses m_1 and m_2 (where the subscript 1 refers to the lower (A) and 2 to the upper (B=C) lava layer) and the sedimentary layers are assumed to be massless lateral springs having stiffnesses k_1 and k_2. The masses of the sedimentary layers are lumped into m_1 and m_2 (i.e., one half of the mass of each sedimentary layer above and below the lava layer is lumped with the mass of the lava layer).

Considering a unit-area vertical column of the soil profile the mass is calculated from $m=\rho H$ where ρ is density and H is the layer thickness. The shear stiffness is calculated by taking the advantage of the relation of shear modulus of the material (μ), its density (ρ) and shear wave velocity (β) or $k=\mu/H$ where $\mu=\rho\beta^2$. From shallow borehole logs in Hveragerði the thickness of the upper lava-rock layer (B=C) is seen to vary between 14–22 m, and the thickness of the sedimentary layer comprised of loose alluvial and marine sand and gravel is between approximately 10–22 m. However, there is no information about other material characteristics (ρ or β) or the layering of deeper layers. As a result, we relied on material properties from other studies in South Iceland on similar geology [33]. By solving the eigenvalue problem numerically we obtained two natural frequencies of oscillation at $f_1\approx 3.5$ Hz and $f_2\approx 8.5$ Hz which match almost exactly the two mean predominant frequencies observed in the HVSR amplification curves.

This result was obtained for the parameters shown in Table 2 which are based on values that are fully consistent with borehole logs and independent observations [17] and required only minimal

Table 2
Properties of soil structure and 2DOF system representative of station IS605.

Layer	Soil				2DOF system	
	H [m]	ρ [g/cm^3]	β [m/s]	μ [N/m^2]	k [N/m]	m [kg]
L_1 (Lava B=C)	15	2.2	1800	7.13×10^9		51.7×10^3
L_2 (Sediment)	22	1.7	750	0.95×10^9	43.5×10^6	
L_3 (Lava A)	12	2.2	1800	7.12×10^9		55.9×10^3
L_4 (Sediment)	12	1.8	800	1.15×10^9	96.0×10^6	

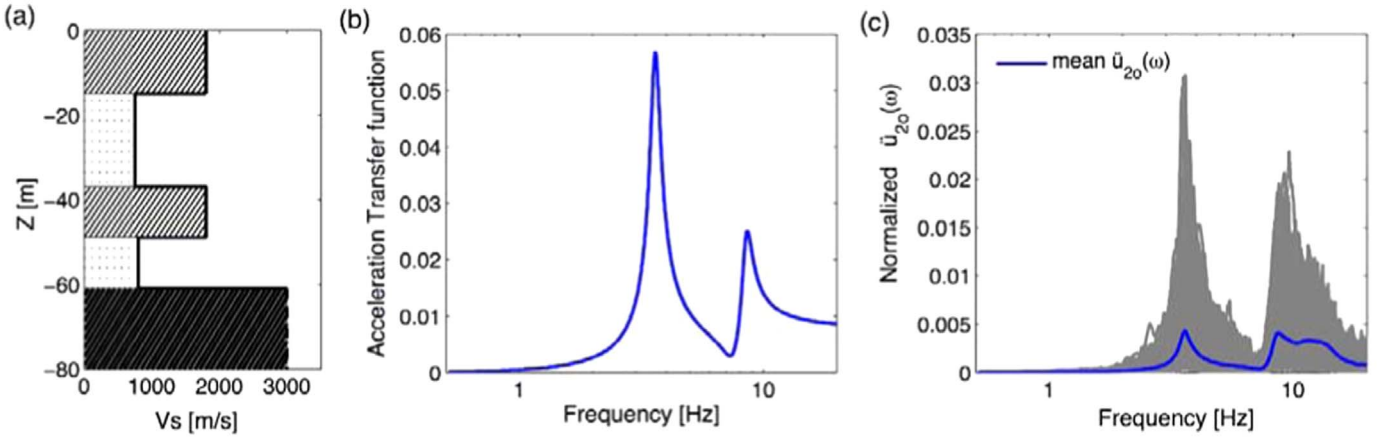

Fig. 10. (a) Shear-wave velocity profile obtained by modal analysis (hatched and dotted areas denote lava and sedimentary layers, respectively, with the bottom layer indicating bedrock). (b) The absolute acceleration transfer function corresponding to the soil structure in (a). (c) Absolute acceleration frequency response of horizontal motion at the free-surface for the 2DOF system (gray curves). The gray curves were calculated using motions at the reference station IS609 and the response of each event was normalized by the maximum frequency response of the vertical ground motion. The blue curve indicates the mean response. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

adjustment. The resulting soil structure is shown schematically in Fig. 10a, which from the surface to the top of the bedrock, consists of a 15 m thick lava layer (B=C), 22 m sedimentary layer, 12 m lava layer (A) and a 12 m thick sedimentary layer. The corresponding absolute acceleration transfer function is shown in Fig. 10b. For illustration purposes we also plot the absolute acceleration frequency response at the free surface (top of the soil structure) in Fig. 10c for each event (gray lines) assuming damping ratios (ξ_1 and ξ_2) for both layers as 5%. Because the bedrock excitation is unavailable we used for simplicity the recorded strong-motion at the reference station IS609 multiplied by 0.5 to remove free surface effect and normalized the absolute horizontal frequency response with the vertical response to reduce scatter. The response in Fig. 10c clearly shows the two fundamental frequencies of oscillations. The response associated with the larger frequency is broader than indicated by the transfer function because the response is affected by the slight peak in the amplification curve of IS609 with predominant frequency at 10 Hz and above.

These results, along with geological mapping and borehole logs, imply that the characteristics of the soil structure underneath IS605 could, at least as a first approximation, be assumed to apply to the part of Hveragerði that lies on lava rock. However, while the HVSR results of the nearest stations IS604 and IS603 (50 and 70 m away from IS605, respectively) show all the same HVSR characteristics for both earthquake and microseismic data, they do not exhibit the clear bimodal HVSR amplification curve from earthquake data. The same can be said for station IS608 which is the closest station to the borehole locations. Other stations on lava rock show different patterns of amplification. Therefore, the results indicate that the primary characteristics of the soil structure such as layer thicknesses and the impedance contrasts of the velocity reversals, may vary considerably under Hveragerði, contributing to the observed variations in ground motion amplitudes, even over short distances. Additionally, while in Fig. 10 the reference site recordings were used to highlight the variations in acceleration response of the 2DOF soil structure, the acceleration response scales with the bedrock acceleration response according to Eq. (9). Accordingly, even though no borehole instruments exist in ICEARRAY I the acceleration response in the bedrock could theoretically be recovered via Eq. (9) from the recorded ground motions at station IS605. We note that the predominant frequency of HVSR from earthquake recordings on lava is considerably lower than for stations on older bedrock. From the results for station IS605, the lava-sedimentary layer that is responsible for this peak lies between ~40 and 60 m deep, which in turn raises questions about the applicability of the average shear-wave velocity in the uppermost 30 m ($V_{S,30}$) as the parameter on which to base site characterization for earthquake resistant design for this type of profile.

The results suggest that the only other available numerical modeling of site response on a lava-sedimentary soil structure in Iceland may now be interpreted in terms of the proposed model of this study. Bessason and Kaynia [17] compared strong-motion site effects on the west and east abutments of the base-isolated and instrumented 80 m long Thjorsa-Bridge during two M_w6.5 and 6.4 earthquakes in South Iceland on 17 and 21 June 2000, respectively, and a several of their aftershocks. They showed that considerable and consistent site amplification is observed on the west side relative to the east side which manifests itself in a considerable peak at around 2–5 Hz in the response spectral ratio. They attribute the relative differences in amplification to the differences in site conditions on each side of the bridge. Namely,

Table 3
Properties of soil structure and SDOF system representative of west abutment of Thjorsa-Bridge.

Layer	Soil [17]			SDOF system	
	H [m]	ρ [g/cm^3]	β [m/s]	k [N/m]	m [kg]
Lava-rock	8	2.0	1000		33×10^3
Sediment	20	1.7	400 – 600	$13.6 - 30.6 \times 10^6$	

based on geotechnical surveys the soil profile on the west side of the bridge consists of an 8–10 m thick lava-rock underlain by an 18–20 m thick sedimentary layer of loose sand and gravel on top of on bedrock, while the east site is located directly on bedrock. After testing several parameter values (i.e., shear-wave velocity, density, and Poisson's ratio) by trial and error, they were able to produce a peak in the simulated response spectra at 3 Hz for the soil structure shown in Table 3 using the computer program SHAKE. In the context of this study however, we model the soil structure as a single-degree-of-freedom-system (SDOF) since it consists of a single lava/sediments stack above the bedrock. The corresponding natural frequency of oscillation is $f_0=\frac{1}{2\pi}\sqrt{\frac{K}{M}}$ where K and M represent the stiffness and mass of the system, respectively. For the range of shear wave velocities assumed ([17], and Table 3) for the sedimentary layer we find that values of the natural (i.e., predominant) frequency of the site lies in the range of 3.2–4.8 Hz, which is in excellent agreement with observations. Thus, our results confirm that modeling soil structure composed of a lava layer on top of a sedimentary layer, therefore introducing velocity reversals, as either a SDOF system (e.g., under the west abutment of the Thjorsa brigdge) or as a 2DOF system (as in Hveragerði) is a robust way of estimating the predominant frequency/ies of site amplification. This is especially the case when details of layer thicknesses are known, but alternatively, in the absence of such information the dynamic model may be used to infer the geologic structure underneath the site.

5. Summary and conclusions

The earthquake strong-motion during the M_w6.3 Ölfus earthquake on 29 May 2008 and 1705 aftershocks were recorded on the first small-aperture urban strong-motion array in Iceland, the ICEARRAY I. The array consists of 12 stations in the town of Hveragerði, located in the western part of the South Iceland Seismic Zone, and makes possible detailed studies of many aspects of engineering seismology ranging from source to site effects. In this vein, the first comprehensive study of localized site effects in an urban area in Iceland was presented herein. The frequency dependent horizontal-to-vertical spectral ratio (HVSR) amplification curves are determined for each strong-motion station using both earthquake data and recordings of ambient noise from an independent microseismic campaign. The station-specific results are compared and interpreted on the basis of geology and borehole information.

The earthquake data was analyzed using both HVSR and standard-spectral-ratio (SSR) methods, while the microseismic data was analyzed using only HVSR. A comparison of the results from the different methods and datasets shows very good agreement among the mean amplification curves. In particular, the agreement between earthquake and microseismic HVSR results indicates that in the absence of earthquake data a microseismic campaign and HVSR analysis is a useful tool to map the spatial distribution of predominant frequencies and amplification for microzonation purposes in Iceland. In fact, the results show that the majority of stations located in the town of Hveragerði present amplification functions of relatively low amplitudes, consistent with the "stiff rock" that characterizes the surface geology of much of Hveragerði. Mapping the spatial distribution of the predominant frequency of amplification shows that the array stations can be divided into two groups, one to the south/southwest having relatively lower frequencies ($\sim$5 Hz), and one to the north/northeast of relatively higher frequencies ($>$10 Hz). A comparison with the surface geology shows that most of the stations to the south/southwest lie on top of a relatively young basaltic lava layer ($\sim$5000 years in age) for which the HVSR amplification is relatively high and characterized by either broad single peak or multiple peaks at relatively low frequencies. In contrast, the stations to the north are located on much older bedrock. In fact, borehole data shows the existence of a surface lava layer on top of sediments, which in turn lie on top of another lava layer, indicating a reversal in the velocities at depth in the profile. Such repeating structure of interchanging lava and sedimentary layers is characteristic of relatively young geology in Iceland.

In addition to microzonation, the ability to estimate physical parameters of the subsurface layering using HVSR results would be of primary practical importance, particularly for the stations on lava. Accordingly, the bimodal HVSR results at station IS605 were analyzed further, focusing on the fact that the mode with lower predominant frequency at 3–4 Hz was not induced by all earthquakes, while the one with higher predominant frequency at 8–9 Hz was. The prominence of the lower frequency mode did not depend on azimuth or distance. However, there was a strong correlation between prominence of the lower frequency mode on earthquake strong-motion amplitudes (i.e., PGA) and magnitude, indicating that lower magnitude earthquakes did not excite the mode at lower frequencies. The results from microseismic HVSR analysis moreover showed that only the mode at higher frequencies was consistently excited, indicating that the microseismic noise, consisting largely of surface waves, either did not contain enough energy at the lower frequencies or penetrate deep enough to excite the mode at lower frequencies. Attempts to model the bimodal amplification curves using the body-wave approximation were unsuccessful because the results did not reproduce the predominant frequencies of the bimodal HVSR curve nor give realistic physical parameters for the layers below, in comparison with borehole data and other sources. In fact, the body-wave approximation was found not to apply to sites such as IS605, and for most of ICEARRAY I stations on lava rock under which velocity reversals exist. Thus, we considered the bimodal amplification at IS605 to be the response of a soil structure consisting of two sets of lava-sedimentary layers, one on top of the other, and modeled it as a classically damped 2DOF linear oscillator. We took advantage of analytical results of the acceleration response of the top mass (i.e., surface) of a 2DOF oscillator subjected to a base (bedrock) excitation and calculated the system properties. The mass and stiffness were derived from the physical properties of the lava and sedimentary layers from typical values from literature, considering the lava layer to be a rigid mass, and borehole records provided constraints on the layer thicknesses. In this way the dynamic response on the free surface exhibited two clear fundamental modes of oscillation, one at 3–4 Hz, the other at 8–9 Hz, matching the observations. The soil structure obtained by modal analysis of two lava-soil profile is in accordance with geologic mapping and borehole data. Finally, we have confirmed the applicability of our approach using the only other case of numerical simulation of site response on a lava/sedimentary soil structure in Iceland.

The results of this study indicate that for rock sites in Iceland characterized by considerable velocity reversals, due to one or more sets of a lava rock layer above a softer sedimentary layer,

higher site amplification over a relatively narrow frequency range can be expected relative to rock sites without velocity reversals. Such conditions, especially if they lie at considerable depth (several tens of meters) may not be captured by microzonation studies that rely solely on microseismic recordings. Moreover, in such cases site characterization based on the average shear wave velocity in the uppermost 30 m is not expected to apply. These results have important implications for earthquake resistant design considerations for structures on rock sites characterized by velocity reversals.

Acknowledgements

This work was supported by the Icelandic Centre for Research (Grant of Excellence no. 141261-051), the Icelandic Catastrophe Insurance (Grant no. S112-2013) and the University of Iceland Research Fund. Christian I. Olivera was supported by the Leifur Eiriksson Foundation and Virginia Tech, and the Earthquake Engineering Research Centre of the University of Iceland provided him material support. The seismographs used in this study were rented from the LOKI instrument bank at the Icelandic Meteorological Office. We thank Judith Marinière, Pierre Turquais, and Romain Millan, interns from EOST in France during for arranging and organizing the ICEARRAY I strong-motion dataset. We would also like to thank the National Energy Authority of Iceland for compiling and sharing the borehole data. Finally, we are grateful to the people of Hveragerði for housing the recording equipment and acknowledge the municipality of Hveragerði for their support.

References

[1] Einarsson P. Plate boundaries, rifts and transforms in Iceland. Jökull 2008;58:35–58.
[2] DeMets C. Current plate motions. Geophys J Int 1990;101:425–78.
[3] DeMets C, Gordon RG, Angus DF, Stein S. Effect of recent revisions to the geomagnetic reversal time scale on estimates of current plate motions. Geophys Res Lett 1994;21:2191–4.
[4] Bellou M, Bergerat F, Angelier J, Homberg C. Geometry and segmentation mechanisms of the surface traces associated with the 1912 Selsund Earthquake, Southern Iceland. Tectonophysics 2005;404:133–49.
[5] Bjarnason IT, Cowie P, Anders MH, Seeber L, Scholz CH. The 1912 Iceland earthquake rupture: growth and development of a nascent transform system. Bull Seism Soc Am 1993;83:416–35.
[6] Einarsson P, Björnsson S, Foulger G, Stefánsson R, Skaftadóttir T. Seismicity pattern in the South Iceland seismic zone. Earthq Predict 1981:141–51.
[7] Einarsson P. Earthquakes and present-day tectonism in Iceland. Tectonophysics 1991;189:261–79.
[8] Halldorsson B, Sigbjörnsson R. The M_w6.3 Ölfus earthquake at 15:45 UTC on 29 May 2008 in South Iceland: ICEARRAY strong-motion recordings. Soil Dyn Earthq Eng 2009;29:1073–83.
[9] Halldórsson B, Olafsson S, Sigbjornsson R. A fast and efficient simulation of the far-fault and near-fault earthquake ground motions associated with the June 17 and 21, 2000, earthquakes in South Iceland. J Earthq Eng 2007;11:343–70.
[10] Pagli C, Pedersen R, Sigmundsson F, Feigl KL. Triggered fault slip on June 17, 2000 on the Reykjanes Peninsula, SW-Iceland captured by radar interferometry. Geophys Res Lett 2003;30.
[11] Sigbjornsson R, Snæbjörnsson JT, Higgins S, Halldórsson B, Olafsson S. A note on the Mw 6.3 earthquake in Iceland on 29 May 2008 at 15:45 UTC. Bull Earthq Eng 2009;7:113–26.
[12] Stefánsson R, Halldórsson P. Strain release and strain build-up in the South Iceland seismic zone. Tectonophysics 1988;152:267–76.
[13] Vogfjord K. Triggered seismicity after the June 17, Mw=6.5 earthquake in the south Iceland seismic zone: the first five minutes. EGS-AGU-EUG Joint Assembly; 2003 1. p. 11251.
[14] Halldórsson B, Sigbjornsson R, Schweitzer J. ICEARRAY: the first small-aperture, strong-motion array in Iceland. J Seism 2009;13:173–8.
[15] Douglas J, Halldorsson B. On the use of aftershocks when deriving ground-motion prediction equations. 9th US national 10th canadian conference on earthquake engineering 9USN10CCEE, Toronto, Canada; 2010, Paper (no. 220).
[16] Bonnefoy-Claudet S, Baize S, Bonilla LF, Berge-Thierry C, Pasten C, Campos J, et al. Site effect evaluation in the basin of Santiago de Chile using ambient noise measurements. Geophys J Int 2009;176:925–37.
[17] Bessason B, Kaynia AM. Site amplification in lava rock on soft sediments. Soil Dyn Earthq Eng 2002;22:525–40.
[18] Fäh D. Microzonation of the city of Basel. J Seism 1997;1:87–102.
[19] Lermo J, Chávez-García FJ. Are microtremors useful in site response evaluation? Bull Seism Soc Am 1994;84:1350–64.
[20] Su F, Aki K, Teng T, Zeng Y, Koyanagi S, Mayeda K. The relation between site amplification factor and surficial geology in central California. Bull Seism Soc Am 1992;82:580–602.
[21] Aki K. Local site effects on weak and strong ground motion. Tectonophysics 1993;218:93–111.
[22] Ansal A, Seco P, Pedro S. Strong ground motions and site amplification. In: Proceedings of the second international conference on earthquake geotechnical engineering; 1999. p. 879–94.
[23] Ansal A, Laue J, Buchheister J, Erdik M, Springman SM, Studer J, et al. Site characterization and site amplification for a seismic microzonation study in Turkey. Proceedings of the 11th International Conference Soil Dynamics and Earthquake Engineering 11th ICSDEE, 3rd International Conference On Earthquake Geotechnical Engineering 3rd ICEGE, January 7–9 Berkeley; 2004. p. 53–60.
[24] Bard PY, SESAME-Team. Guidelines for the implementation of the H/V spectral ratio technique on ambient vibrations: measurements, processing, and interpretations. SESAME European research project. SESAME European Research Project; 2005.
[25] Barani S, Massa M, Lovati S, Spallarossa D. Effects of surface topography on ground shaking prediction: implications for seismic hazard analysis and recommendations for seismic design. Geophys J Int 2014.
[26] Einarsson P., Douglas G. Geology of Iceland: rocks and landscape. Mál Og Menn; 1994.
[27] Atakan K, Brandsdóttir B, Halldórsson P, Friðleifsson G. Site response as a function of near-surface geology in the south iceland seismic zone. Nat Hazards 1997;15:139–64.
[28] Boore DM. Can site responsebe predicted? J Earthq Eng 2004;8:1–41.
[29] Borcherdt RD. Effects of local geology on ground motion near San Francisco Bay. Bull Seism Soc Am 1970;60:29–61.
[30] Steidl JH, Tumarkin AG, Archuleta RJ. What is a reference site? Bull Seism Soc Am 1996;86:1733–48.
[31] Cranswick E. The information content of high-frequency seismograms and the near-surface geologic structure of "hard rock" recording sites. Pure Appl Geophys 1988;128:333–63.
[32] Boore DM, Joyner WB. Site amplifications for generic rock sites. Bull Seism Soc Am 1997;87:327–41.
[33] Field EH, Zeng Y, Johnson PA, Beresnev IA. Non-linear ground-motion amplification by sediments during the 1994 Northridge earthquake: observation and finite source simulation. J Geophys Res: Solid Earth 1998;103:26869–83.
[34] Nakamura Y. A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface. Q Rep Railw Tech Res Inst 1989;30:25–33.
[35] Kanai K, Tanaka T. On Microtremor VIII. Bull Earthq Res Inst Tokyo Univ 1961;39:97–114.
[36] Nogoshi M, Igarashi T. On the propagation characteristics estimations of subsurface using microtremors on the ground surface. J Seism Soc Jpn 1970;23:264–80.
[37] Nogoshi M, Igarashi T. On the amplitude characteristics of microtremor (Part 2). J Seism Soc Jpn 1971;24:26–40.
[38] Konno K, Ohmachi T. Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor. Bull Seism Soc Am 1998;88:228–41.
[39] Sivaram K, Mahesh P, Rai SS. Stability assessment and quantitative evaluation of H/V spectral ratios for site response studies in Kumaon Himalaya, India using ambient noise recorded by a broadband seismograph network. Pure Appl Geophys 2012;169:1801–20.
[40] Mucciarelli Marco, Gallipoli Maria Rosaria. A critical review of 10 years of microtremor HVSR technique. Boll Geofis Teor Appl 2001;42:255–66.
[41] Halldorsson B, Sigbjörnsson R, Rupakhety R, Chanerley AA. Extreme near-fault strong-motion of the M6.3 Ölfus earthquake of 29 May 2008 in South Iceland. 14th European Conference Earthquake Engineering 14ECEE, Ohrid, Macedonia; 2010, Paper no. 1640.
[42] Halldorsson B, Avery H. Converting strong-motion networks to arrays via common triggering. Seism Res Lett 2009;80:572–8.
[43] Böðvarsson R, Rögnvaldsson ST, Slunga R, Kjartansson E. The SIL data acquisition system-at present and beyond year 2000. Phys Earth Planet Inter 1999;113:89–101.
[44] Brandsdóttir B, Parsons M, White RS, Guðmundsson Ó, Drew J, Thorbjarnardóttir BS. The May 29th, 2008 earthquake aftershock sequence within the South Iceland Seismic Zone: fault locations and source parameters of aftershocks. Jokull J Glaciol Geol Soc Icel 2010;60:23–46.
[45] Pétursson GG, Vogfjörð Kristín. Attenuation Relations for near- and Far-Field Peak Ground Motion (PGV, PGA) and New Magnitude Estimates for Large Earthquakes in SW-Iceland. Reykjavík, Iceland: Icelandic Meteorological Office; 2009.
[46] Bard PY. Microtremor measurements: a tool for site effect estimation? Proc Second Int Symp Eff Surf Geol Seism Motion Yokohama Jpn 1998:1251–79.
[47] Triantafyllidis P, Theodulidis N, Savvaidis A, Papaioannou C, Dimitriu P. Site effects estimation using earthquake and ambient noise data: the case of Lefkas town (W. Greece); 2006. p. 3–8.
[48] Pilz M, Parolai S, Leyton F, Campos J, Zschau J. A comparison of site response techniques using earthquake data and ambient seismic noise analysis in the

large urban areas of Santiago de Chile. Geophys Res Lett 2009;178:713–28.
[49] Olivera CI, Halldorsson B, Ólafsson S, Green RA, Sigbjörnsson R. A first look at site effects at icelandic strong-motion stations using microseismic data. In: Proceedings of the 2nd european conference earthquake engineering seismology 2ECEES, Paper no. 2044, Istanbul, Turkey; 2014.
[50] Parolai S, Bormann P, Milkereit C. Assessment of the natural frequency of the sedimentary cover in the Cologne area (Germany) using noise measurements. J Earthq Eng 2001;5:541–64.
[51] D'Amico V, Picozzi M, Baliva F, Albarello D. Ambient noise measurements for preliminary site-effects characterization in the urban area of Florence, Italy. Bull Seism Soc Am 2008;98:1373–88.
[52] Sato T, Hirasawa T. Body wave spectra from propagating shear cracks. J Phys Earth 1973;21:415–31.
[53] Aki K. Scaling law of seismic spectrum. J Geophys Res 1967;72:1217–31.
[54] Sato T, Hirasawa T. Body wave spectra from propagating shear cracks. J Phys Earth 1973;21:415–31.
[55] Aki K, Richards PG. Quantitative seismology. theory and methods, vol. I, II. San Francisco, USA: W.H. Freeman and Company; 1980.
[56] Nakamura Y. Clear identification of fundamental idea of Nakamura's technique and its applications. In: Proceedings of the 12th World Conference on Earthquake Engineering, 2000, p. 2656.
[57] Nakamura Y. On the H/V spectrum. In: Proceedings of the 14th World Conference on Earthquake Engineering, Beijing, China; 2008.
[58] Albarello D, Lunedei E. Alternative interpretations of horizontal to vertical spectral ratios of ambient vibrations: new insights from theoretical modeling. Bull Earthq Eng 2010;8:519–34.
[59] Giacomo DD, Gallipoli MR, Mucciarelli M, Parolai S, Richwalski SM. Analysis and modeling of hvsr in the presence of a velocity inversion: the case of Venosa, Italy. Bull Seism Soc Am 2005;95:2364–72.
[60] Herak M. ModelHVSR—A Matlabs tool to model horizontal-to-vertical spectral ratio of ambient noise. Comput Geosci 2008;34:1514–26.
[61] Tsai NC. A note on the steady-state response of an elastic half-space. Bull Seism Soc Am 1970;60:795–808.
[62] Chopra AK. Dynamics of structures, vol. (3). New Jersey: Prentice Hall; 1995.
[63] Chopra AK. Modal analysis of linear dynamic systems: Physical interpretation. J Struct Eng 1996;122:517–27.
[64] Yoshida N. Seismic ground response analysis, vol. 36. Dordrecht: Springer Netherlands; 2015.
[66] Sæmundsson K, Kristinsson S. Hveragerði. Soil temperature measurements and faults. Reykjavík, Iceland: Iceland GeoSurvey (ÍSOR); 2005 (In Icelandic).

Further reading

[65] Schnabel PB, Lysmer J, Seed HB. SHAKE–A computer program for response analysis of horizontally layered sites United States: University of Califorina Berkeley; 1972 Report No. EERC 72-12.

Paper II

Bayesian hierarchical model of variations in earthquake peak ground acceleration within small-aperture arrays

Sahar Rahpeyma, Benedikt Halldorsson, Birgir Hrafnkelsson, and Sigurjon Jónsson

Rahpeyma S, Halldorsson B, Hrafnkelsson B, Jónsson S (2018) Bayesian hierarchical model of variations in earthquake peak ground acceleration within small-aperture arrays. *Environmetrics, 29(3)*, doi:10.1002/env.2497

Received: 2 November 2017 | Revised: 27 February 2018 | Accepted: 27 February 2018

DOI: 10.1002/env.2497

RESEARCH ARTICLE

WILEY

Bayesian hierarchical model for variations in earthquake peak ground acceleration within small-aperture arrays

Sahar Rahpeyma[1] | Benedikt Halldorsson[1,2] | Birgir Hrafnkelsson[3] | Sigurjón Jónsson[4]

[1]Earthquake Engineering Research Centre & Faculty of Civil and Environmental Engineering, School of Engineering and Natural Sciences, University of Iceland, Selfoss, Iceland

[2]Geoscience Research Group, Division of Processing and Research, Icelandic Meteorological Office, Reykjavík, Iceland

[3]Department of Mathematics, Faculty of Physical Sciences, School of Engineering and Natural Sciences, University of Iceland, Reykjavik, Iceland

[4]Physical Science and Engineering Division (PSE), King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia

Correspondence
Benedikt Halldorsson, Earthquake Engineering Research Centre & Faculty of Civil and Environmental Engineering, School of Engineering and Natural Sciences, University of Iceland, Selfoss, Iceland.
Email: skykkur@hi.is

Funding information
Icelandic Centre for Research, Grant/Award Number: 141261-051/52/53; Eimskip Fund of the University of Iceland; University of Iceland Research Fund; Marie Curie International Re-integration Grant

Knowledge of the characteristics of earthquake ground motion is fundamental for earthquake hazard assessments. Over small distances, relative to the source–site distance, where uniform site conditions are expected, the ground motion variability is also expected to be insignificant. However, despite being located on what has been characterized as a uniform lava-rock site condition, considerable peak ground acceleration (PGA) variations were observed on stations of a small-aperture array (covering approximately 1 km^2) of accelerographs in Southwest Iceland during the Ölfus earthquake of magnitude 6.3 on May 29, 2008 and its sequence of aftershocks. We propose a novel Bayesian hierarchical model for the PGA variations accounting separately for earthquake event effects, station effects, and event-station effects. An efficient posterior inference scheme based on Markov chain Monte Carlo (MCMC) simulations is proposed for the new model. The variance of the station effect is certainly different from zero according to the posterior density, indicating that individual station effects are different from one another. The Bayesian hierarchical model thus captures the observed PGA variations and quantifies to what extent the source and recording sites contribute to the overall variation in ground motions over relatively small distances on the lava-rock site condition.

KEYWORDS

Bayesian hierarchical model, ground motion variabilities, peak ground acceleration, posterior density function, strong-motion array

1 | INTRODUCTION

The spatial variability of strong ground motions can highly affect the response of lifeline systems such as pipelines, tunnels, bridges, and transmission systems. One of the important lessons learned from recent catastrophic earthquakes (e.g., Mexico City earthquake, 1985; Loma Prieta earthquake, 1989; Kobe earthquake, 1995; İzmit earthquake, 1996; and Chi-Chi earthquake, 1999) is that the distribution pattern of seismic waves propagation can significantly change due to variation in geological conditions even over relatively small areas. Quantitatively estimating the strong-motion properties will provide practical information of the key factors that affect the spatial variation of ground motion parameters. Accordingly, it is of paramount importance to obtain the contribution of the earthquake source, path, and site effects

to the overall variability of ground motions. This paper develops a statistical method to address the contribution of the earthquake source and localized station effects into the overall variability of strong ground motions in addition to quantifying the associated variabilities. This estimation is essential for microzonation studies and decision making for urban planning.

After the installation of the SMART-1 array, the first digital strong-motion circular array (37 accelerographs in an aperture of $D = 2$ km) in Taiwan in 1980, noticeable variations in seismic ground motions were observed over short distances (Abrahamson, 1988). The recorded data have provided a great opportunity to identify the properties of the earthquake source, recording station, and propagation path effects on distribution of ground motions across the array (Bolt, Tsai, Yeh, & Hsu, 1982; Chiu, Amirbekian, & Bolt, 1995; Dimitriu, Theodulidis, & Bard, 2000; Harichandran, 1991; Loh, 1985; Loh, Ang, & Wen, 1983; Theodulidis & Bard, 1995). Since then, attempts have been made to characterize the spatial variability of seismic ground motions by using closely spaced networks and seismic arrays due to the critical effects ground motion variations can have on reliable earthquake disaster mitigation and early-stage damage assessments. A large number of seismic arrays and networks, permanent or temporary, have been installed around the world to investigate the characteristics of seismic ground motions. For instance, Meremonte, Frankel, Cranswick, Carver, and Worley (1996) deployed portable digital seismographs in the San Fernando Valley to investigate relevant factors to seismic zonation in urban areas, such as site amplification, sedimentary basin effects, and the variability of ground motion over short distances. They observed high variation in ground motion parameters over a distance of 200 m. Bessason and Kaynia (2002) observed the variability of strong-motion recordings and response spectra on the west and east abutments of the base-isolated and instrumented 80-m-long Thjorsa Bridge in South Iceland. Shabestari, Yamazaki, Saita, and Matsuoka (2004) estimated the spatial distribution of ground motion parameters using the data recorded by K-NET and KiK-net (strong-motion seismograph networks) stations in Japan over a 1 km × 1 km area. Most recently, Rahpeyma, Halldorsson, Olivera, Green, and Jónsson (2016) investigated the variability across the small-aperture and urban strong-motion arrays in south and north of Iceland.

In general, the observed ground motions at a station can be expressed as a convolution of the primary terms of source, path, and site effects (Boore, 1983, 2003; Field & Jacob, 1995). The unknown source, path, and site effects are usually approximated using parameterized empirical models that are fitted to a given data set via regression analysis (e.g., Baltay, Hanks, & Abrahamson, 2017; Bindi, Castro, Franceschina, Luzi, & Pacor, 2004; Boatwright, Fletcher, & Fumal, 1991; Boore, Joyner, & Fumal, 1997; Chin & Aki, 1991; Moya, Aguirre, & Irikura, 2000; Ortiz-Alemán, Reyes-Olvera, Iglesias-Mendoza, Orozco-Del-Castillo, & Hernández-Gómez, 2017; Shabestari et al., 2004) or using empirical techniques (e.g., Aki, 1957; Bard, 1998; Borcherdt, 1970; Nakamura, 1989). However, the inversion problems are mostly nonlinear and can be influenced by non-uniqueness solution, which leads to high level of uncertainties. On the other hand, over the last couple of decades, many researchers have focused on physically quantifying the properties of the earthquake source, path, and site responses to develop the new generation of ground motion models (GMMs), which are used to predict the expected peak ground parameters along with the model variability (e.g., Boore, Stewart, Seyhan, & Atkinson, 2014; Campbell & Bozorgnia, 2014; Chiou & Youngs, 2014).

It has long been known that the reliability of ground motion amplitudes plays an important role in an accurate seismic hazard assessment, especially for near-fault events. The total aleatory variability (i.e., of natural origin) in empirical attenuation models can be split into two main independent components: inter-event variability and intra-event variability (e.g., Abrahamson & Youngs, 1992; Atik et al., 2010). The inter-event variability can be described as the ground motion variability related to the event-specific factors (e.g., randomness in the source e.g., possible source variation and associated directivity effects), whereas the intra-event variability can be interpreted as the ground motion variability related to the record-specific factors (e.g., randomness in site response) (Strasser, Abrahamson, & Bommer, 2009).

After the M_w 6.3 Ölfus earthquake on May 29, 2008 in South Iceland over 1,700 aftershocks were recorded on 12 strong-motion stations of the Icelandic strong-motion array in the town of Hveragerði (ICEARRAY I), deployed in 2007 and located in the extreme near-fault region of the mainshock (Halldorsson & Sigbjörnsson, 2009). While insignificant variability in ground motion amplitudes was expected due to uniform station condition across ICEARRAY I, the opposite was observed, and the recorded ground motions of the mainshock and aftershocks exhibited considerable variations in peak ground acceleration (PGA) and peak ground velocity (PGV) (Douglas & Halldorsson, 2010; Halldorsson & Sigbjörnsson, 2009; Rahpeyma et al., 2016). These data thus provide an ideal opportunity to statistically model and physically interpret the spatial distribution of ground motion intensities over relatively small area. The objective is to separate the variations in amplitudes in order to find more reliable inference of source, path, and site effects. The dense recordings at short interstation distances are expected to provide a more reliable assessment of ground motion amplitude distribution and its variability.

In this paper, we introduce a practical scheme based on the Bayesian statistical framework to give us a better understanding of the distribution of strong-motion amplitudes and to quantify the corresponding variabilities. For that purpose, we propose a new Bayesian hierarchical model (BHM) for the variability of PGA across a small-aperture array. The proposed model incorporates many of the commonly used seismic parameters and offers a flexible probabilistic framework for multilevel modeling of PGA that accounts for the effects of the earthquake source, propagation path effects and localized site effects, along with their respective variabilities. The proposed multilevel model is designed to explain the variability in the data with an event effect, a station effect, an event-station effect, and a term representing unexplained effects and other factors that are not accounted for (e.g., see Kuehn & Scherbaum, 2015, 2016; Landwehr, Kuehn, Scheffer, & Abrahamson, 2016; Sigurdarson & Hrafnkelsson, 2016). Our proposed BHM model makes it possible to separate the intra-event variability into two variability terms, that is, a station effect and an event-station effect. This is an improvement as the models of Kuehn and Scherbaum (2015, 2016) do not separate the event-station effect and the term representing the unexplained effects and unaccounted factors but model these two terms as one term. The Matérn covariance function is used to model the spatial fields corresponding to the station effect and the event-station effect. The Matérn covariance function has three parameters that control the spatial correlation, namely, the decay parameter, the smoothness parameter, and the amplitude of the spatial field (see, e.g., Cressie & Huang, 1999). We note that Jayaram and Baker (2009) have outlined some of the most common models that have been fitted to spatial data in various fields. Although the exponential model appeared to be the most appropriate model for general application, they found that common fitting procedures were not optimal for obtaining the model parameters. Hence, they employed a manual fitting approach in which the model parameters were "tuned" using visual judgment. In contrast, the proposed multilevel modeling approach in this study was found to be both appropriate and reliable to determine the model parameters and the associated variability.

The BHM in this study is thus useful for the prediction of ground motion parameters with location across the array and will improve our understanding of the ground motion distribution by separating it into contributions attributed to the source and localized site effects, respectively. The results indicate the importance of regionalization of the ground motion in the effort of understanding the underlying sources of the aleatory variability, which consequently affects the seismic hazard analysis. We note at the outset that the proposed BHM is a general model and could be applied to any ground motion parameter of interest.

2 | ARRAY STRONG-MOTION DATA

Iceland is seismically the most active region in Northern Europe and is located on the extensional plate boundary of the North American and Eurasian tectonic plates known as the Mid-Atlantic Ridge. The extensional spreading rate of approximately 2 cm/year manifests itself as a belt of volcanic and seismic zones across Iceland from southwest to northeast (DeMets, Gordon, Angus, & Stein, 1994; Einarsson, 1991, 2008; Sæmundsson, 1979). Due to an eastward ridge jump in Iceland, two transfer zones have formed, the onshore South Icelandic Seismic Zone (SISZ) and the largely offshore Tjörnes Fracture Zone (TFZ) in the north (Figure 1). Historically, the most destructive earthquakes have taken place in these two transfer zones, and the size of these earthquakes has been estimated between magnitude 6 and 7 (Sigbjörnsson, Sigurdsson, Snæbjörnsson, & Valsson, 2006; Stefánsson et al., 1993; Stefánsson, Guðmundsson, & Halldórsson, 2008). The SISZ is characterized by a series of parallel north–south right-lateral strike-slip faults on which the largest earthquakes take place, the latest one being the earthquake of magnitude 6.3 on May 29, 2008.

ICEARRAY I, the first strong-motion array in Iceland, was installed in the town of Hveragerði in the western part of the SISZ in 2007 (Halldorsson & Sigbjörnsson, 2009; Halldorsson, Sigbjörnsson, & Schweitzer, 2009). It consists of 12 strong-motion stations in an area of ~1.23 km^2 with interstation distance ranging from 50 to 1,900 m (Figure 1). As can be seen in a geological map of Hveragerði (Figure 2), most of the town lies directly on top of lava-rock (Sæmundsson & Kristinsson, 2005) considered to represent a uniform "rock" site condition within ICEARRAY I. Accordingly, a uniform site amplification of seismic waves would be expected at the ICEARRAY I stations. Nevertheless, based on local geological surveys and available log information for shallow boreholes across the Hveragerði, the rock-site condition can be separated into an old bedrock condition on one hand, and a young lava-rock condition on the other hand. For the latter, the uppermost lava-rock layer (B|C in Figure 2) is found to lie on top of a softer sedimentary layer, which in turn lies on top of another lava-rock layer (A in Figure 2). Below that is most likely another sedimentary layer on top of the old bedrock at 70- to 80-m depth, resulting in two contrasts in material density and velocity between rock and sedimentary layers, which characterize the site amplification on lava-rock in Hveragerði (Rahpeyma et al., 2016).

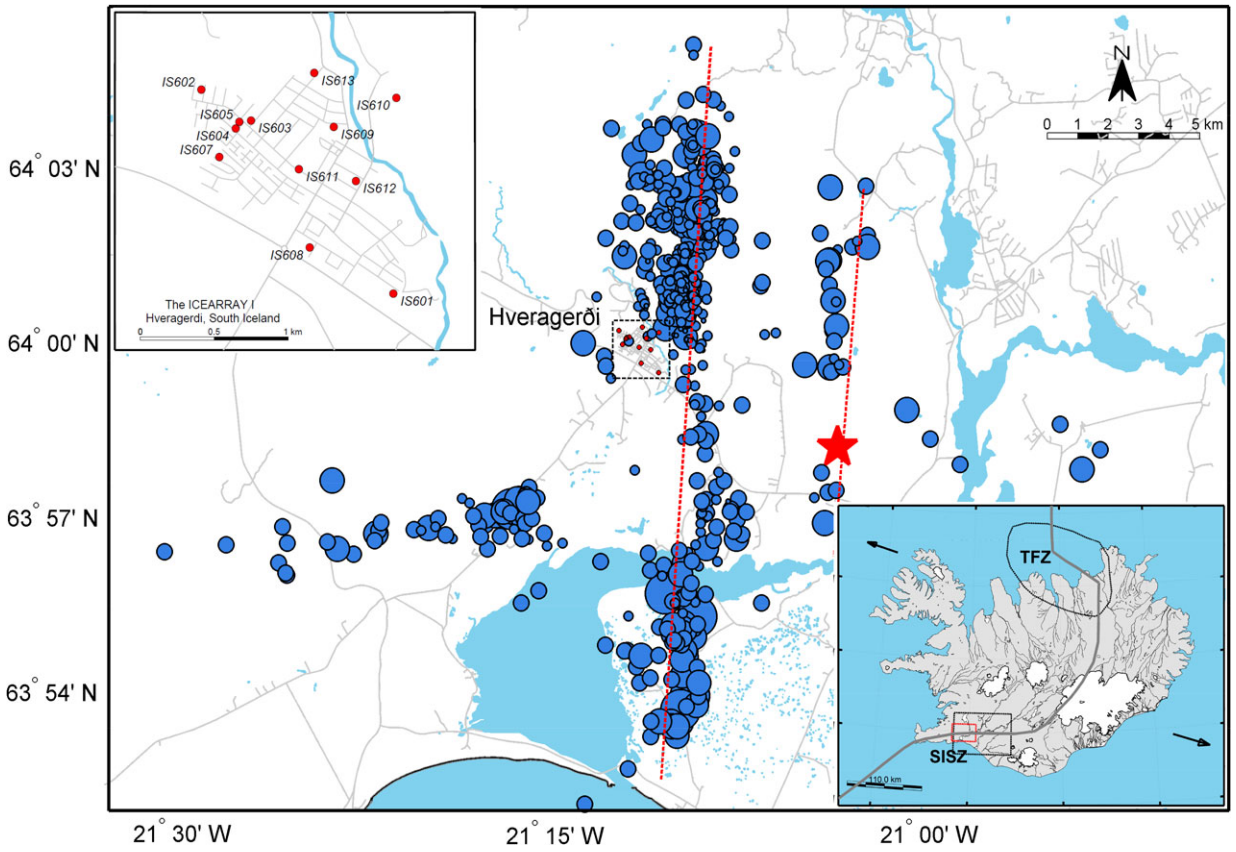

FIGURE 1 Aftershock distribution (blue circles) of the Ölfus earthquake on May 29, 2008 (red star shows the mainshock epicenter) in Southwest Iceland recorded by Icelandic strong-motion array in the town of Hveragerði (ICEARRAY I) network, with the location of the causative faults approximated by the red dashed lines. The top-left figure shows the 12 ICEARRAY I strong-motion stations (red circles) located within the town of Hveragerði (black dashed rectangle in the central figure). The bottom-right map of Iceland shows the approximate location of the Mid-Atlantic plate boundary (dark solid line), the Tjörnes Fracture Zone (TFZ), and South Iceland Seismic Zone (SISZ). The red solid rectangle marks the area shown in the central figure

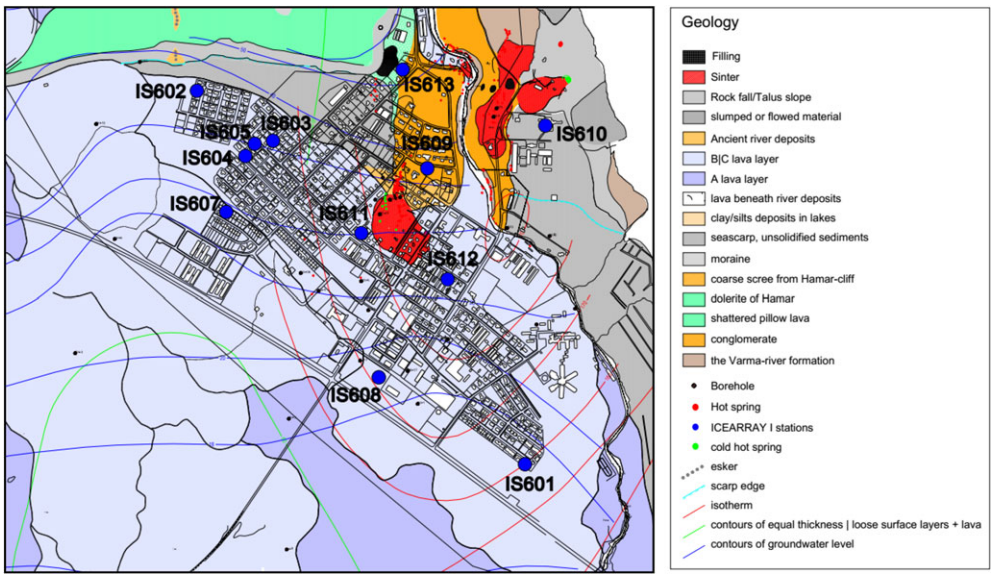

FIGURE 2 Geological map of Hveragerði, adopted from (Sæmundsson & Kristinsson, 2005), showing that most of the town is located on lava. The Icelandic strong-motion array in the town of Hveragerði (ICEARRAY I) station locations are marked as blue circles

Less than a year after ICEARRAY I was deployed, on May 29, 2008, the M_w 6.3 Ölfus earthquake took place, and the array recorded over 1,700 of its aftershocks (Figure 1). The primary analysis of the distribution of peak ground motion parameters (e.g., PGA, PGV, and response spectral acceleration (SA)) during the mainshock and a few aftershocks exhibited significant ground motion variation within the array (Douglas & Halldorsson, 2010; Halldorsson & Sigbjörnsson, 2009; Rahpeyma et al., 2016, 2017). As mentioned above, it is well recognized that localized geological structure can significantly vary within short distances and influence site responses (Bessason & Kaynia 2002; Di Giacomo, Gallipoli, Mucciarelli, Parolai, & Richwalski, 2005; Loh, 1985; Rahpeyma et al., 2016; Shabestari et al., 2004). Consequently, it is important to quantify such variation and its associated variability in order to evaluate its importance in the mitigation of seismic risk. The extensive ICEARRAY I strong-motion data set from M_w 6.3 Ölfus earthquake and its aftershocks therefore provides an ideal opportunity to model, parameterize, and interpret the distribution of ground motion across the array along with the associated variability.

The aftershock distribution of the Ölfus earthquake delineates two parallel and near-vertical north–south striking right-lateral strike-slip faults approximately 4.5 km apart (Figure 1). While the mainshock epicenter was recorded on the eastern fault, the majority of the aftershocks occurred on the western fault that is located only 1–2 km from Hveragerði and ruptured ~2 s after the eastern fault (Halldorsson, Sigbjörnsson, Rupakhety, & Chanerley, 2010; Sigbjörnsson, Snæbjörnsson, Higgins, Halldórsson, & Ólafsson, 2009). The mainshock ground motion was characterized by intense but short-duration ground acceleration (~5–6 s) and large amplitude and long-period near-fault velocity pulses. The geometric mean of the horizontal PGA of the mainshock varies in the range of ~44–88% of the acceleration of gravity (g) at distances ranging ~1.0–2.3 km from the causative faults (Halldorsson & Sigbjörnsson, 2009). Similar variations have been noted for the aftershocks, which range in local magnitudes of 0.42–4.75 and epicentral distances of 1.6–15 km (Figure 3), and they resulted in maximum geometric mean of horizontal PGA values in the range 0.03–38% g across stations (Douglas & Halldorsson, 2010; Rahpeyma et al., 2016).

To investigate the variability of strong motion recorded on the same geological unit, we used the aftershock recordings from 10 stations that are directly located on the lava-rock layer that underlies the main part of the town. We thus excluded station IS610 that sits on older bedrock on a hillside and station IS613 that is also on the old bedrock and for which only a few recordings were available. In addition, two stations, IS608 and IS688, are collocated and thus provide essentially identical results (Halldorsson & Sigbjörnsson, 2009). After reviewing the three-component acceleration time histories for

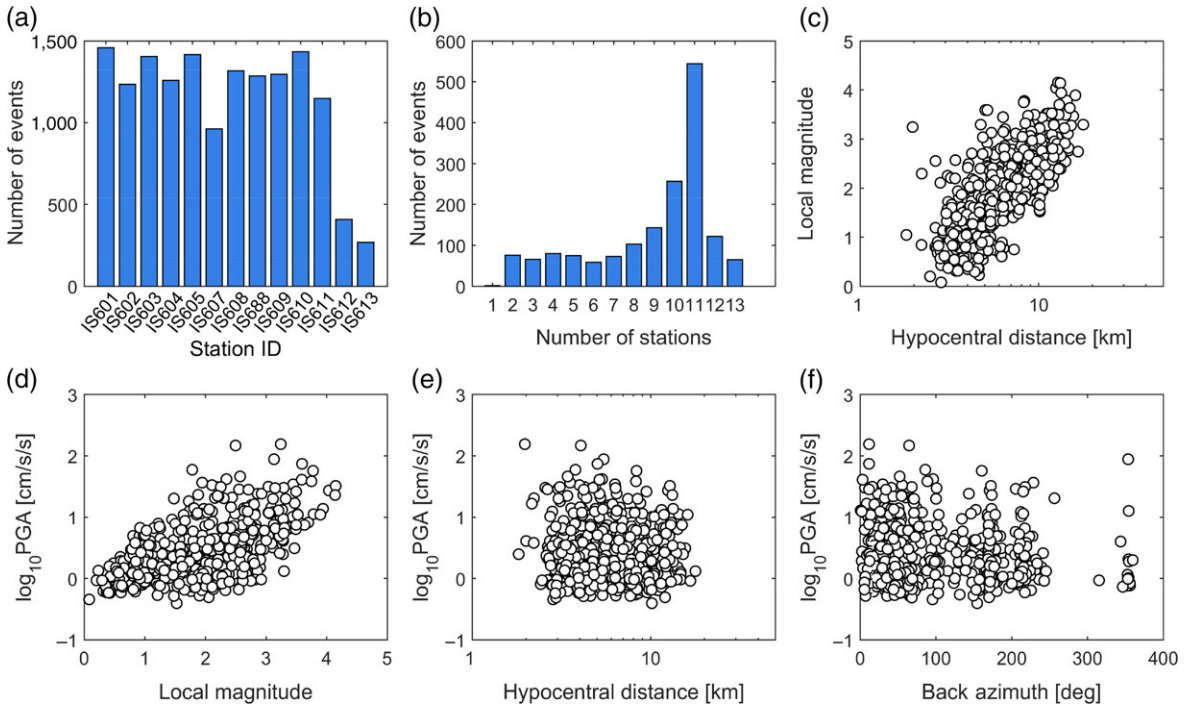

FIGURE 3 Parametric description of the aftershock recordings of the Icelandic strong-motion array in the town of Hveragerði (ICEARRAY I). (a) The number of aftershocks recorded by each station; (b) the number of events recorded by a given number of stations. Distribution of (c) local magnitude versus source-to-station distance; distribution of PGA recorded at station IS605 versus (d) local magnitude, (e) hypocentral distance, and (f) back-azimuth

all events and all stations for quality purposes, we selected 610 events recorded by 10 strong-motion stations to use in this study. After excluding missing data points (i.e., stations with no records), a total of 4,620 data points of PGA (geometric mean of PGA from two horizontal components) remained for the analysis. For all these records, reliable seismic parameters were available (e.g., local magnitude and hypocentral location) from other networks (Brandsdóttir et al., 2010; Stefánsson et al., 1993).

3 | STATISTICAL MODELING

3.1 | A BHM for PGA

The proposed statistical model for the spatially referenced PGA data is a BHM with a data level, a latent level, and a hyperparameter level. The observed log-scale PGA, $\log_{10} PGA$, is modeled at the data level of the BHM. For N events recorded at Q stations, let Y_{es} represent the base-10 logarithm of the observed PGA for event e at station s. The proposed model consists of five independent terms and can be presented as Equation (1), as follows:

$$\begin{aligned} Y_{es} &= \mu_{es}(M_e, R_{es}, D_e) + \delta B_e + \delta S_s + \delta WS_{es} + \delta R_{es}, \\ & e = 1, \ldots, N, \quad s = 1, \ldots, Q \end{aligned} \tag{1}$$

where μ_{es} provides median ground motion in terms of independent variables (magnitude, hypocentral distance, and depth of the origin) for event e and station s, δB_e is the event effect (also called inter-event residual or event term) that denotes the overall effect of event e in addition to the predicted median ground motion μ_{es}, δS_s is the station effect for station s, δWS_{es} is a spatially correlated event-station effect, and δR_{es} is an independent error term representing unexplained effects or other factors that are not accounted for. The terms δWS_{es} and δR_{es} are assumed to follow Gaussian distributions, thus conditioned on μ_{es}, δB_e, and δS_s; then, Y_{es} also follows a Gaussian distribution.

The term μ_{es} is a commonly used linear predictive function for median ground motion. Although GMMs come in a variety of different functional forms, in this study, we nominate a commonly used and parsimonious model that links PGA to the local magnitude of the eth earthquake, M_e; the hypocentral distance from the eth event to the sth station, R_{es}; and the depth of the origin of the eth earthquake, D_e, as follows:

$$\mu_{es} = \beta_1 + \beta_2 M_e + \beta_3 \log_{10}(R_{es}) + \beta_4 D_e. \tag{2}$$

The coefficients $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)$ correspond to the particular area and volume under study in which the seismic waves travel on their way to the surface, reflecting the peculiarities of the local seismic regime and geological structure. We assume a weakly informative prior distribution for the $\boldsymbol{\beta}$ coefficients, with the mean $\boldsymbol{\mu}_\beta$ and the diagonal covariance matrix $\boldsymbol{\Sigma}_\beta$ (see Appendix). The event effects are combined in the vector $\boldsymbol{\delta B} = (\delta B_1, \ldots, \delta B_N)$, which we assume a priori to be normally distributed with mean zero and standard deviation τ, independent of each other. The station effects in the vector $\boldsymbol{\delta S} = (\delta S_1, \ldots, \delta S_Q)$ are assumed to stem from a mean-zero Gaussian field governed by a Matérn covariance function with marginal variance ϕ^2_{S2S} that describes the station-to-station variability, smoothness parameter $\nu_{S2S} = 0.5$, (see Equation (3), i.e., exponential covariance function), and range parameter Δ_{S2S}. We note that the exponential covariance has been found to be an appropriate model to fit to the spatial data (Foulser-Piggot & Stafford 2012; Goda & Atkinson, 2009; Goda & Hong, 2008; Jayaram & Baker, 2009; Landwehr et al., 2016). The covariance matrix of the vector $\boldsymbol{\delta S}$ is denoted by $\boldsymbol{\Sigma}_{S2S}$, and it can be presented as

$$\{\boldsymbol{\Sigma}_{S2S}\}_{ij} = \left\{ \phi^2_{S2S} \exp\left(-\frac{d_{ij}}{\Delta_{S2S}}\right) \right\}_{ij}. \tag{3}$$

The spatially correlated event-station effects, δWS_{es}, are modeled as a mean-zero Gaussian field for each event e that is also governed by a Matérn covariance function with marginal variance ϕ^2_{SS} that describes within event variability, smoothness parameter $\nu_{SS} = 1.5$ (after testing other plausible values), and range parameter Δ_{SS}, as follows:

$$\{\boldsymbol{\Sigma}_{SS}\}_{ij} = \left\{ \phi^2_{SS} \left(1 + \frac{d_{ij}}{\Delta_{SS}}\right) \exp\left(-\frac{d_{ij}}{\Delta_{SS}}\right) \right\}_{ij}. \tag{4}$$

It has been shown that the correlation structure of the residuals from GMMs is not dependent on the direction and is therefore isotropic (e.g., Goda & Hong, 2008; Jayaram & Baker, 2009; Wang & Takada, 2005). In addition, our analysis

of residuals revealed that it is reasonable to assume that the spatial correlation functions for event-station effects are isotropic. The Gaussian fields of any two events are independent, and thus, the vector of all δWS_{es} terms stacked by subvectors containing terms from the same event has mean zero and a block-diagonal covariance matrix. Finally, the term representing unexplained effects, δR_{es}, is assumed to be Gaussian with zero mean and standard deviation of ϕ_R, and the δR_{es} terms are assumed to be independent of each other. Theoretically, it is difficult to infer the parameters of the spatial field representing the station effect because there is only one replicate of it in the proposed model (see Zhang, 2004). Thus, both the smoothness parameter and the range parameter are selected beforehand, and only the marginal variance is inferred. We assume that the station field is continuous but not its first derivative (in terms of expected mean square error). This is achieved by setting the smoothness parameter equal to 0.5. Further, it is assumed that the spatial correlation is small for distances of same scale as the smallest interstation distance, leading to a range parameter equal to 0.06 km. In the case of the event-station effect, there are many replicates of that effect, and thus, the marginal variance, the smoothness parameter, and the range parameter can be inferred. However, it can be hard to determine the smoothness parameter when an uncorrelated error term such as δR_{es} has to be taken into account and its variance to be estimated along with the other parameters. Thus, the smoothness parameter is selected under the assumption that the fields corresponding to δWS_{es} have a continuous first derivative (again in terms of expected mean square error), so the fields are rather smooth relative to fields that are only continuous. This is achieved by setting the value of the smoothness parameter equal to 1.5. A more detailed description of the BHM is given in the Appendix.

The total variance of Y_{es}, when conditioning on M_e, R_{es}, and D_e, can be calculated as the sum of the inter-event variance τ^2, the interstation variance ϕ^2_{S2S} (i.e., station-to-station variability), the event-station variance ϕ^2_{SS} (i.e., variability between stations for an event), and the variance of the unexplained effects and other unaccounted factors, ϕ^2_R, as follows:

$$\sigma^2 = \tau^2 + \phi^2_{S2S} + \phi^2_{SS} + \phi^2_R \,. \tag{5}$$

The inter-event variance (τ^2) quantifies the variations between events after taking the effects of M_e, R_{es}, and D_e of each event into account. The interstation variance (ϕ^2_{S2S}) quantifies the variability in the station effects that stems from the varied local geological conditions. In other words, the interstation variance represents the systematic deviation of the observed amplification at the specific station from the median amplification predicted by the model. The event-station variance $\left(\phi^2_{SS}\right)$, on the other hand, is a measure of the spatial variability in PGA between stations within events after taking into account the overall effect of the event and the average effect of each station. Finally, the variance parameter ϕ^2_R quantifies jointly the variability in the unexplained effects and deviations that are not accounted for by other terms in the model.

3.2 | Posterior inference

MCMC algorithm is used to sample from the posterior density of the proposed BHM. As explained in the previous section, we set $\Delta_{S2S} = 0.06$ (units in kilometer), which is effectively equal to the minimum interstation distance. A sensitivity analysis indicated that increasing the range parameter accounts for unreliable results mainly due to high correlation of station effects for short distances. Accordingly, the vector of hyperparameters is defined as $\boldsymbol{\theta} = (\tau, \phi_{S2S}, \phi_{SS}, \phi_R, \Delta_{SS})$, and the latent parameter vector is $\boldsymbol{\eta} = (\boldsymbol{\beta}, \boldsymbol{\delta B}, \boldsymbol{\delta S})$. To simplify the posterior sampling scheme, the hyperparameters are reparametrized by the logarithm function, that is, $\boldsymbol{\psi} = (\psi_1, \psi_2, \psi_3, \psi_4, \psi_5) = (\log(\tau), \log(\phi_{S2S}), \log(\phi_{SS}), \log(\phi_R), \log(\Delta_{SS}))$. In general, it is easier to sample from the posterior density of parameters that are defined on the real line, as in $\boldsymbol{\psi}$, compared with parameters that are strictly positive, as in $\boldsymbol{\theta}$.

The MCMC algorithm consists of two steps. In the first step, the hyperparameters are sampled jointly from the marginal posterior density $\pi(\boldsymbol{\psi}|\boldsymbol{y})$ using the Metropolis step of Roberts, Gelman, and Gilks (1997). In the second step, the latent parameters are sampled jointly from $\pi(\boldsymbol{\eta}|\boldsymbol{y}, \boldsymbol{\psi})$; see Appendix for further details. In this Metropolis algorithm, random samples $\boldsymbol{\psi}^*$ are drawn from a multivariate normal density centered at the previous draw $\boldsymbol{\psi}^{k-1}$ with a covariance matrix $c(-\boldsymbol{H})^{-1}$, where $\boldsymbol{H}$ is the Hessian matrix of $\log\pi(\boldsymbol{\psi}|\boldsymbol{y})$ evaluated at the mode $\boldsymbol{\psi}_0$ and with $c = 2.38^2/\ dim\ (\boldsymbol{\psi})$ (Roberts et al.,1997). The Hessian matrix is defined as a square matrix of second-order partial derivatives of a scalar-valued function and can be presented here as Equation (6), as follows:

$$\boldsymbol{H} = \nabla^2 \log\ \pi(\boldsymbol{\psi}|\boldsymbol{y})|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} \,, \tag{6}$$

where ∇^2 is the second derivative operator for a multivariable function. Consequently, the resulting proposal density is

$$q\left(\boldsymbol{\psi}^*|\boldsymbol{\psi}^{k-1}\right) = \mathrm{N}\left(\boldsymbol{\psi}^*|\boldsymbol{\psi}^{k-1}, c(-\boldsymbol{H})^{-1}\right). \tag{7}$$

Moreover, the scaling parameter c can be shown to yield optimal acceptance rates in a particular and large dimensional scenario (Roberts et al., 1997).

Although finding the local mode in linear and nonlinear problems is relatively straightforward by using classical local optimization methods, finding the global mode is more problematic especially in nonlinear problems. In general, in many nonlinear optimization problems, the objective function has multiple local minima and maxima. In this study, we use a numerical grid search optimization scheme that employs three steps of simple MCMC simulation to estimate the approximate values for the posterior mode of the log-transformed hyperparameters ($\boldsymbol{\psi}_0$). In the first step, we generate $N_g = 1,000$ uniform random values of log-transformed hyperparameters. Then, the mean and standard deviation of 5% of the samples closest to the maximum of $\log\pi(\boldsymbol{\psi}|\boldsymbol{y})$ are computed. The obtained mean and standard deviation are used to generate $N_g = 1,000$ random samples of $\boldsymbol{\psi}$ assuming a normal distribution for the second step. This last step can be repeated several times. Finally, the hyperparameters that produce the maximum $\log\pi(\boldsymbol{\psi}|\boldsymbol{y})$ are considered as the global mode.

In this study, we run $N_C = 4$ chains with a total length of $N_T = 10{,}000$ samples by considering the first 25% of the iterations ($N_B = 2{,}500$) as burn-in samples. The following scheme summarizes the sampling of the posterior density of $\boldsymbol{\psi}$:

1. Initialize the MCMC process with preliminary estimates of parameters.
   Set $\Delta_{S2S} = 0.06$. Evaluate the mode, $\boldsymbol{\psi}_0$, and the Hessian matrix, $\boldsymbol{H}$, of the marginal posterior density of the hyperparameters, $\boldsymbol{\psi}$ at $\boldsymbol{\psi}_0$. Generate the initial value for $\boldsymbol{\psi}$ from a normal density with mean $\boldsymbol{\psi}_0$ and covariance $c(-\boldsymbol{H})^{-1}$.
2. At step k, sample a proposal value $\boldsymbol{\psi}^*$ from a normal density with mean $\boldsymbol{\psi}^{k-1}$ and covariance $c(-\boldsymbol{H})^{-1}$. Sample a proposal value $\boldsymbol{\eta}^*$ from the conditional posterior density of $\boldsymbol{\eta} = (\boldsymbol{\beta}, \boldsymbol{\delta B}, \boldsymbol{\delta S})$, conditioned on $\boldsymbol{\psi} = \boldsymbol{\psi}^*$. Note that $\pi(\boldsymbol{\eta}|\,\boldsymbol{\psi}, \boldsymbol{y})$ is a normal density (see formulation in the Appendix).
3. Calculate the ratio

$$r = \min\left\{1, \frac{\pi\left(\boldsymbol{\psi}^* \quad |\boldsymbol{y}\right)}{\pi\left(\boldsymbol{\psi}^{k-1}|\boldsymbol{y}\right)}\right\}. \tag{8}$$

4. Sample u_k from uniform density on [0, 1]. Accept or reject the proposed values of $(\boldsymbol{\psi}, \boldsymbol{\eta})$ according to

$$\left(\boldsymbol{\psi}^k, \boldsymbol{\eta}^k\right) = \begin{cases} \left(\boldsymbol{\psi}^{k-1}, \boldsymbol{\eta}^{k-1}\right) & \text{if} \quad r \le u_k \\ \left(\boldsymbol{\psi}^* \quad , \boldsymbol{\eta}^*\right) & \text{if} \quad r > u_k \end{cases}, \tag{9}$$

   where $\boldsymbol{\eta}^*$ is the vector of latent parameters at iteration k. It is noteworthy that one can also sample first from the marginal density of $\boldsymbol{\psi}$ and store the results. Then, the next step is to sample $\boldsymbol{\eta}$ conditioned on the previously sampled $\boldsymbol{\psi}$.
5. Run the MCMC algorithm with the updated estimates and repeat steps 2–4.
6. Obtain posterior summaries for $\boldsymbol{\psi}$ and $\boldsymbol{\eta}$ using their posterior samples.

3.3 | Convergence diagnostics

We evaluated the convergence of the MCMC algorithm by visualizing trace plots and histograms of the parameters, which display the distribution of posterior samples. Furthermore, we evaluated the convergence by calculating the Gelman–Rubin ($\hat{R}$) statistic for all parameters (Gelman & Rubin, 1992). As a reference, the point estimates of the $\hat{R}$ statistics should be between 1.00 and 1.05 for all parameters (Brooks & Gelman, 1998; Gelman & Rubin, 1992). Moreover, we assessed the dependence between successive samples of the Markov chain using the sample autocorrelation from lag 1 to 50. The faster the autocorrelation decreases with lag, the better is the effective sample size (Gelman, Carlin, Stern, & Rubin, 2014).

4 | RESULTS AND DISCUSSION

The proposed BHM model was applied in the analysis of the ICEARRAY I PGA data set described above. The posterior sampling scheme described in the previous section was used to infer the hyperparameters and the latent parameters that resulted in well-determined, unimodal, and nearly symmetric marginal posterior distributions for each parameter. The estimates of the marginal posterior density of the hyperparameters along with the convergence diagnostic results (i.e., Gelman–Rubin statistics and autocorrelation plots) are graphically illustrated in Figure 4. The figure shows that the hyperparameters have nearly symmetric and unimodal densities, and their posterior mean, standard deviation and percentiles are presented in Table 1. The Gelman–Rubin diagnostic ($\hat{R}$) or the potential scale reduction is in

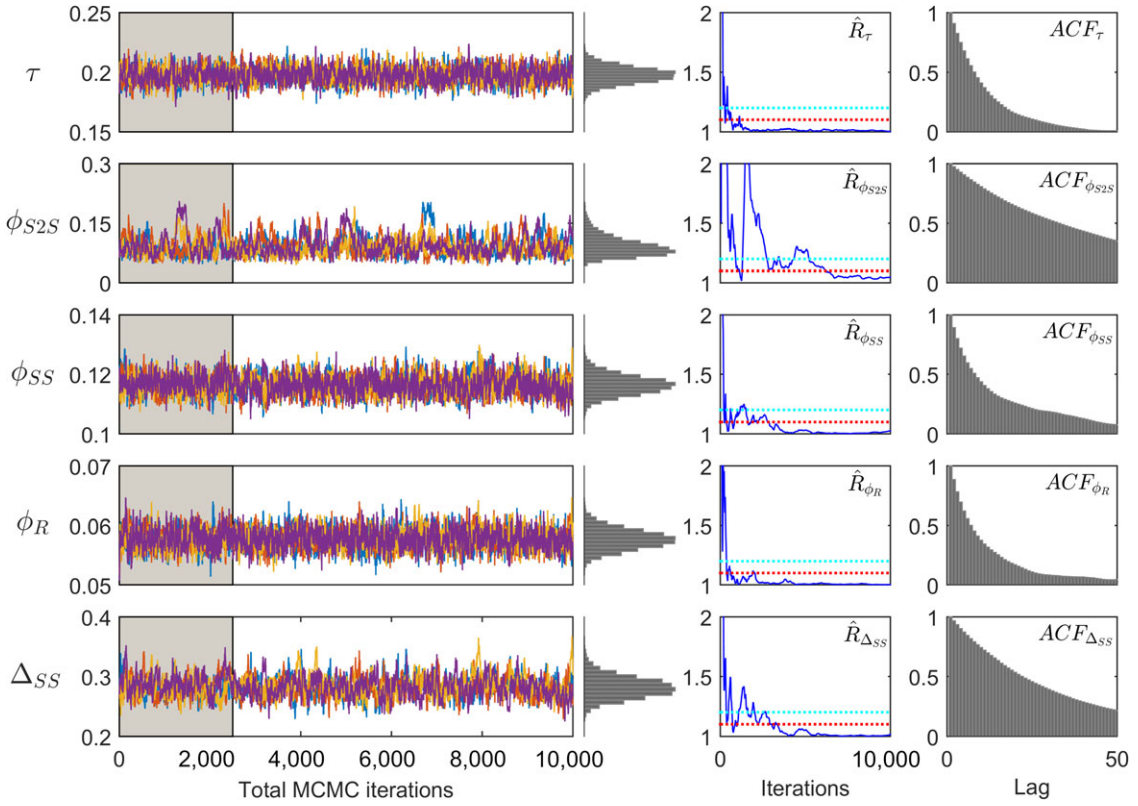

FIGURE 4 Trace plots of posterior samples of the hyperparameters in $\boldsymbol{\theta}$ using 10,000 iterations and 4 chains within Markov chain Monte Carlo (MCMC) simulations (left) with the initial 25% of iterations during the burn-in period indicated in gray and resulting histograms of the hyperparameter values. The Gelman–Rubin convergence statistics (middle) and the autocorrelation function plots (right) are also shown

TABLE 1 Posterior statistics for hyperparameters, $\boldsymbol{\theta}$

θ	Mean	Standard deviation	Percentiles %				
			2.5	25.0	50.0	75.0	97.5
τ	0.1977	0.0067	0.1847	0.1930	0.1977	0.2022	0.2112
ϕ_{S2S}	0.0915	0.0240	0.0571	0.0741	0.0873	0.1039	0.1510
ϕ_{SS}	0.1164	0.0032	0.1104	0.1142	0.1164	0.1185	0.1231
ϕ_R	0.0577	0.0017	0.0546	0.0565	0.0576	0.0588	0.0611
Δ_{SS}	0.2819	0.0177	0.2502	0.2697	0.2810	0.2932	0.3193

the range of 1.01–1.05, and the degree of autocorrelation is less than 0.10 at a lag of 50 for all the hyperparameters (Figure 4). In addition, the overall acceptance rate for hyperparameters is ~42%, which is within the recommended range (Roberts et al., 1997). A close scrutiny of the posterior statistics in Table 1 indicates that the results from our proposed BHM are robust and provide reliable estimates of the hyperparameters and that their uncertainties are low.

The posterior distributions for the latent parameters $\boldsymbol{\beta}$ also meet the convergence diagnostics, and the posteriors exhibit nearly symmetric and unimodal distributions (Figure 5 and Table 2). The posterior means and 95% posterior confidence intervals of each of the elements in the latent parameter $\boldsymbol{\delta B}$, the event effect for each event in addition to the GMM defined in Equation (2), are shown in Figure 6. The posterior means of the event effects are centered around zero as expected, and the mean value of posterior standard deviations of the event effects is around 0.07. Finally, the posterior estimates of the latent parameter $\boldsymbol{\delta S}$, representing the individual station effects across the array, show similar characteristics (Figure 7 and Table 3), with all the δS_s meeting the convergence diagnostics criteria, and their marginal posterior densities are nearly symmetric and unimodal with similar standard deviations but with different means.

As the model parameters are well determined, an accurate Bayesian prediction of the geometric mean PGA across the array can be achieved. A comparison of the model predictions with the data is captured by the total residual difference between the data and model predictions model, defined as $Y_{es} - \mu_{es}$, as is evident from Equation (2). Theoretically, all observations from a single event have the same inter-event effect that represents the unique effect occurring

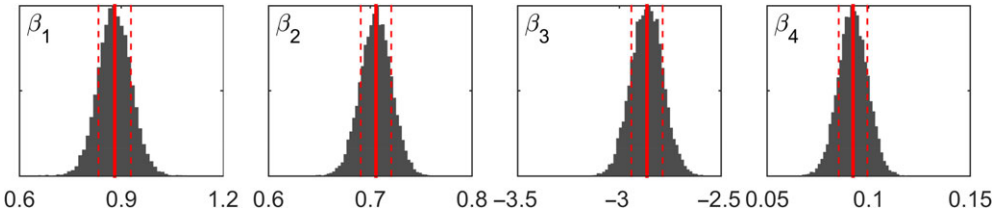

FIGURE 5 The resulting posterior histograms of latent ground motion model parameters in $\boldsymbol{\beta}$ after using four chains within Markov chain Monte Carlo simulations. The red thick and dashed lines show the posterior mean and posterior mean ± posterior standard deviation, respectively

TABLE 2 Posterior statistics for ground motion model coefficients, $\boldsymbol{\beta}$

$\boldsymbol{\beta}$	Mean	Standard deviation	Percentiles %				
			2.5	25.0	50.0	75.0	97.5
β_1	0.8807	0.0479	0.7866	0.8488	0.8802	0.9128	0.9760
β_2	0.7056	0.0150	0.6763	0.6955	0.7056	0.7157	0.7350
β_3	−2.8645	0.0769	−3.0129	−2.9168	−2.8638	−2.8126	−2.7117
β_4	0.0923	0.0070	0.0784	0.0876	0.0922	0.0971	0.1060

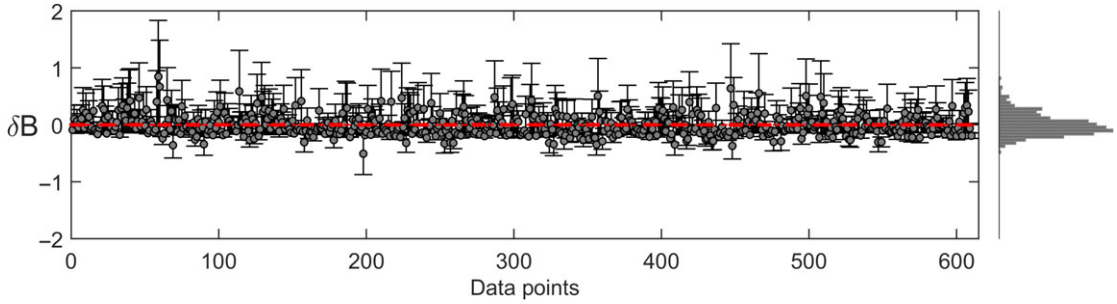

FIGURE 6 The error bars show posterior mean (gray dots) and 95% posterior intervals (error bars) of the overall effect of events, $\boldsymbol{\delta B}$, using four chains within Markov chain Monte Carlo simulations. The histogram represents the distribution of posterior means of the elements in $\boldsymbol{\delta B}$. The red dashed line represents the zero reference line

during that earthquake rupture and influences the corresponding ground shaking at all stations in the same manner. This event-specific influence is captured by the posterior mean of inter-event effects (δB_e).

The overall distribution of the inter-event effects in Figure 6 is plotted as a function of event index. However, a comparison of the inter-event effects with physical parameters such as local magnitude, distance, depth, and back-azimuthal angle is more meaningful, as is done in Figure 8a–d. The distribution of $\boldsymbol{\delta B}$ is close to being symmetric around zero, especially in the central range of magnitude, distance, and depth of the data. Near the ends of the ranges of the physical parameters, some slight deviations are seen (e.g., Figure 8a), which indicate that for this limited part of the data set, the GMM can be improved with parameters that aim at capturing these deviations. For that, a meaningful physical explanation of the deviation should ideally be sought after which the appropriate parameter or function could be added to the empirical GMM. That is, however, not the objective of this study, particularly as the BHM captures the overall data set quite well with the simple GMM, and the inter-event residuals are overall well behaved.

The earthquake-specific effects may now be removed from the total residuals as $\delta W_{es} = Y_{es} - (\mu_{es} + \delta B_e)$, which defines the intra-event residuals. They vary across the array and represent all other differences between ground motion at different locations that are not captured by the model (Abrahamson & Youngs, 1992). Figure 8e–h shows the intra-event residual behavior as a function of the various seismic parameters (independent variables) of the data set. The residuals behave as expected, that is, they have mean of zero and show absolutely no trends with local magnitude, distance, depth, and back-azimuthal angle.

We note that while the earthquake magnitude, the source–site distance, and depth are the independent variables of the GMM, the back-azimuth is not. It is the angle from ICEARRAY I to the epicenter, used here to effectively identify groups of events on different faults or parts of faults. The unbiased distribution of intra- and inter-event residuals in Figure 8

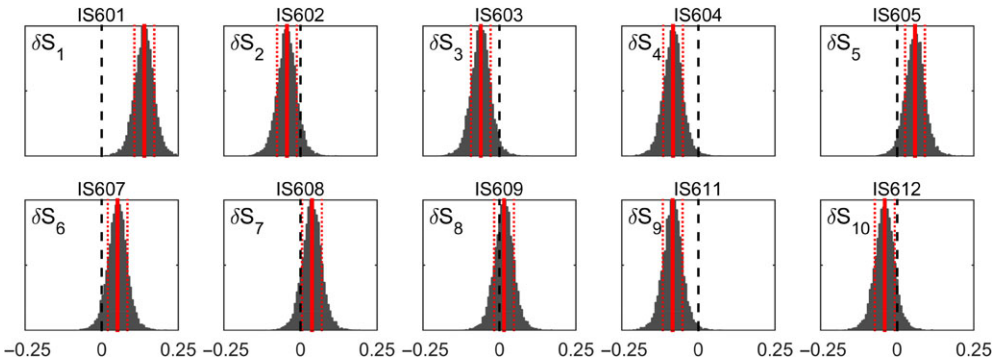

FIGURE 7 Posterior probability density functions of the station effects in $\boldsymbol{\delta S}$, based on Markov chain Monte Carlo simulations of four chains. The red thick and dashed lines show the posterior mean and posterior mean ± posterior standard deviation, respectively. The zero line is marked by a black dashed line

TABLE 3 Posterior statistics for station effects, $\boldsymbol{\delta S}$

Station effect	Station ID	Mean	Standard deviation	Percentiles %				
				2.5	25.0	50.0	75.0	97.5
δS_1	IS601	0.1393	0.0320	0.0754	0.1187	0.1395	0.1596	0.2035
δS_2	IS602	−0.0443	0.0323	−0.1090	−0.0645	−0.0443	−0.0239	0.0194
δS_3	IS603	−0.0610	0.0322	−0.1248	−0.0813	−0.0608	−0.0406	0.0023
δS_4	IS604	−0.0828	0.0322	−0.1471	−0.1029	−0.0825	−0.0627	−0.0190
δS_5	IS605	0.0580	0.0322	−0.0057	0.0377	0.0582	0.0783	0.1210
δS_6	IS607	0.0518	0.0324	−0.0126	0.0317	0.0520	0.0720	0.1154
δS_7	IS608	0.0373	0.0324	−0.0277	0.0171	0.0375	0.0580	0.1015
δS_8	IS609	0.0149	0.0324	−0.0502	−0.0053	0.0153	0.0354	0.0786
δS_9	IS611	−0.0841	0.0323	−0.1483	−0.1042	−0.0837	−0.0636	−0.0207
δS_{10}	IS612	−0.0409	0.0325	−0.1063	−0.0614	−0.0405	−0.0200	0.0244

indicates that the relative ground motions over the array are not dependent on the causative fault. That effectively confirms that this parameter indeed does not need to be included in the GMM. It should be noted that comparing the correlation of the residuals indicates that the correlations between inter-event residuals, v_e, are very low (max ∼20%), which is in agreement with the high variability of the aftershock PGAs ($\tau = 0.19$ and ∼60% of total variance), whereas the high level of ground motion correlations of intra-event residuals, δW_{es}, may be caused by the joint influence of path and complex site effects.

Similar to the earthquake-specific effects, there are station-specific contributions to the overall variability of the intra-event residuals in Figure 8e–h. They are captured by the station effect, $\boldsymbol{\delta S}$, the elements of which are shown individually in Figure 7 and have well-behaved posterior distributions, respectively. Comparing the spatial distribution of the posterior mean values of the station terms with the sample average of the intra-event residuals at each station indicates that the pattern and the values of the two spatial distributions are essentially identical across the array, in turn confirming that the posterior mean of the station effect is well determined and the remaining residuals are randomly distributed after the station effect has been accounted for. Further insight into the characteristics of the spatial distribution is provided by the posterior correlation matrix of the station effects, shown in Figure 9. Of particular note is that the station effects show very high positive correlation values (> 0.95). The correlation is even higher among stations IS602, IS603, IS604, IS605, and IS607, which are located close to each other in the northwest part of the array, particularly IS603–605, which have interstation distances of 50 –114 m (see Figure 2). Not surprisingly, they have the greatest positive correlation between the station effects. Despite the high correlation between station effects, the posterior density of the interstation variance ϕ^2_{S2S} clearly indicates that ϕ^2_{S2S} is different from zero, confirming that the amplitude of the spatial field $\delta S(\cdot)$ is different from zero. This strongly indicates that the station effects are different. As can be seen in Table 3, considering the 95% marginal posterior intervals for the individual δS across ICEARRAY I, the station effect only for stations IS601, IS604, and IS611 are significantly affecting (i.e., amplifying or de-amplifying) the earthquake strong-motion intensities. For instance, the impact of station effect for station IS601 (95% posterior interval is [0.0754, 0.2035]) is estimated to

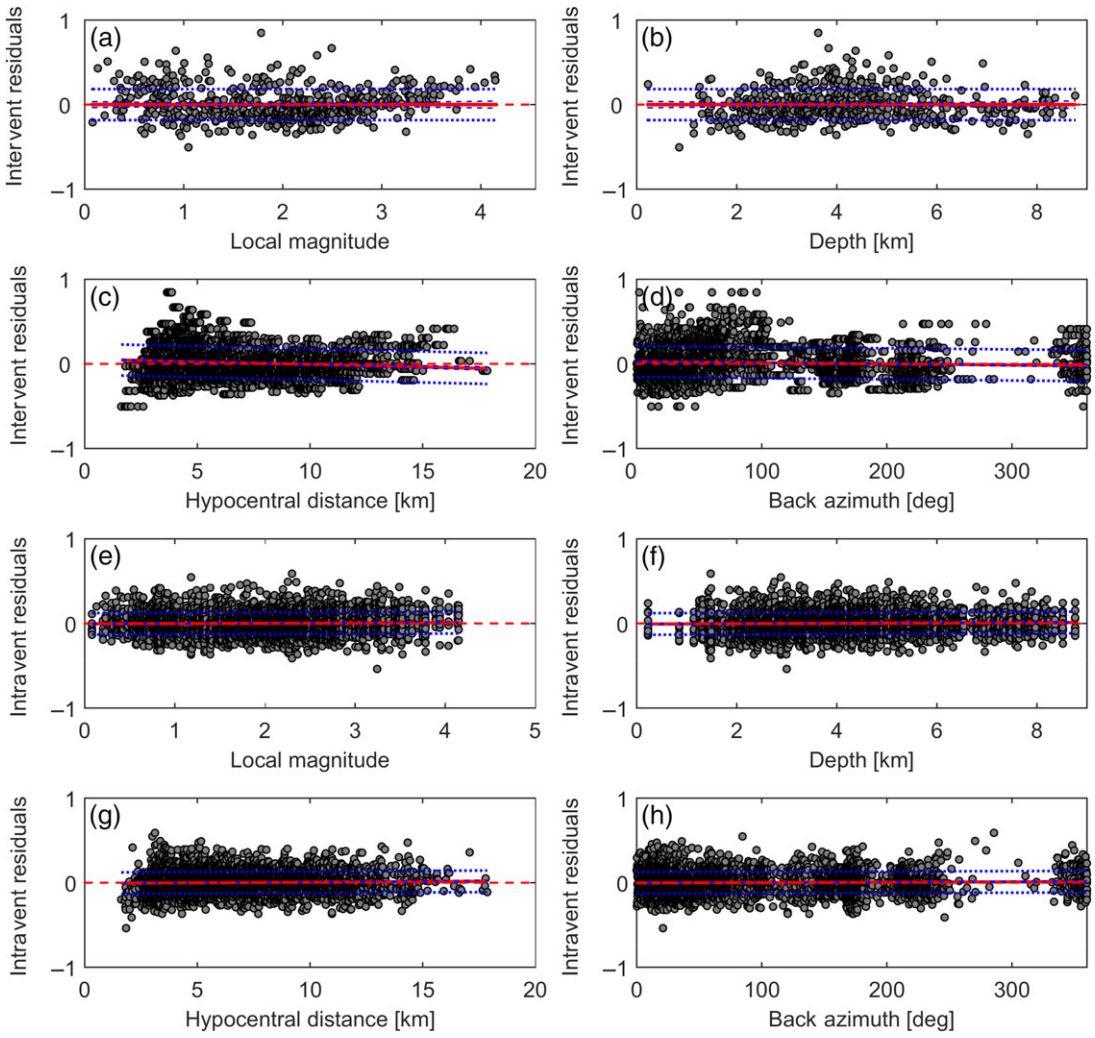

FIGURE 8 Inter-event residuals versus (a) moment magnitude, (b) depth, (c) hypocentral distance, and (d) back-azimuth. Intra-event residuals versus (e) moment magnitude, (f) depth, (g) hypocentral distance, and (h) back-azimuth. The red thick and blue dotted lines show the sample regression line and regression line ± one standard deviation, respectively

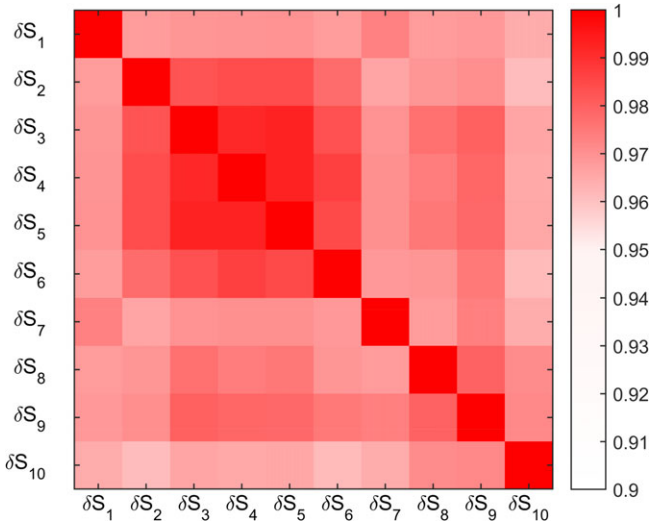

FIGURE 9 Station effect correlation matrix for the Icelandic strong-motion array in the town of Hveragerði

intensify the relative earthquake ground motions, whereas δS values for stations IS604 and IS609 (95% posterior intervals are [− 0.1471, − 0.0190] and [− 0.1483, − 0.0207], respectively) are expected to diminish the relative earthquake ground motions. In the case of the other seven stations, the ones with positive posterior means are more likely to intensify the earthquake ground motions, whereas the ones with negative posterior means are more likely to diminish the earthquake ground motions.

The overall high correlation shows that all stations record a similar level of ground motion, and the remaining and relative average differences of which are then captured by the station effects. The high correlation is thus a manifestation of the combination of the small-aperture array relative to the large (but very similar) source–site distances and the relatively uniform site condition. Taking this further, one expects that as the array aperture decreases, the overall correlation approaches unity, whereas if the array aperture is increased, the correlation should decrease rapidly due to the seismic waves traveling through increasingly different surface material and crustal paths, resulting in increasingly different relative source–site distances. Our results for ICEARRAY I now provide a baseline for the expected station correlations over a relatively uniform area and the individual station effects and thus highlight the overall importance of modeling the station effect.

When inspecting the hyperparameters, we note that generally, the inter-event variability reported for empirical GMMs is smaller than the intra-event variability (Strasser et al., 2009). On the contrary, our results show that the inter-event standard deviation ($\tau = 0.19$) is larger than the intra-event standard deviation (which is calculated from the square root of the sum of squares of the interstation, event station, and unexplained effects variances as $\phi = \sqrt{\phi_{S2S}^2 + \phi_{SS}^2 + \phi_R^2} = 0.16$). This is a manifestation of the similar characterization of the ground motions recorded on relatively uniform site conditions, which can be assumed to be affected by similar propagation path effects. In addition, it is reasonably comparable with previously reported values of PGA amplitude variability over the ICEARRAY I, but given the uniformity of soil structure under the array stations used in this study, the corresponding intra-event variability may suggest a lower bound on the standard deviations attainable for GMMs (Douglas & Halldorsson, 2010). The total standard deviation as measured by Equation (5) is $\sigma \simeq 0.25$, which is in the range of values for the total standard deviation that are commonly used for GMMs in hazard analyses in Iceland (0.20 − 0.30) (Ambraseys, Douglas, Sarma, & Smit, 2005; Olafsson & Sigbjörnsson, 2000). It is higher than the standard deviation of empirical models that have been calibrated to strong-motion data from South Iceland. The main differences between the previous studies and the current one are the relatively few events used, considerably larger source–site distances, and, most importantly, much larger earthquake magnitudes (average magnitude $\sim M_{\mathrm{w}}$ 6) compared with the data set of this study (Kowsari et al., 2017; Olafsson & Sigbjörnsson, 2000).

One of the important advantages of our proposed BHM is estimating the intrastation variability that represents the variation of ground motions recorded within a single station. Quantifying intrastation variability makes it possible to estimate single-station sigma (Atkinson, 2006). The single-station sigma was introduced to detect and avoid double counting those components of ground motion variability at a specific station that are repeatable from different sources (Atkinson, 2006; Kuehn & Scherbaum, 2015, 2016; Morikawa et al., 2008; Rodriguez-Marek et al., 2013; Rodriguez-Marek, Montalva, Cotton, & Bonilla, 2011). To obtain the single-station standard deviation, either we need a station-specific data set (Atkinson, 2006; Chen & Faccioli, 2013; Lin et al., 2011; Morikawa et al., 2008; Rodriguez-Marek et al., 2011; Zafarani & Soghrat, 2017) or we propose a nonergodic model (Kuehn & Scherbaum, 2015, 2016). With respect to the ergodic assumption, there is no difference in ground motion variability observed in a global data set and the one at a specific station source. The majority of GMMs have been obtained based on ergodic assumption (i.e., combining observations at different stations from different events); however, in this study, multilevel modeling allows splitting the variability and quantifying the interstation standard deviation. In our study, the exclusion of the interstation variability results in a decrease of standard deviations (single-station sigma $\sqrt{\tau^2 + \phi_{SS}^2 + \phi_R^2}$) by about 6.8%, which is likely much less than the variability of the hazard estimates themselves and thus unlikely to significantly influence the seismic hazard analysis. While this small value is mainly due to uniform and hard site conditions across ICEARRAY I, it is nevertheless important to quantify.

Finally, contrary to the classical frequentist methods (i.e., the different regression algorithms described in Boore et al., 1997, and the references therein) designed to decouple and provide point estimates of the relative source, propagation path, and local site effects in addition to estimating the associated variability, the proposed BHM methodology provides a more flexible framework that additionally obtains the probability distribution of each model parameter. On that basis, for example, the relative contribution of station effects on the observed ground motion amplitudes can be quantitatively estimated, the posterior probability density function of the station effects can be identified, and their posterior intervals can be used to determine whether a given station effect is positive or negative or whether its sign

cannot be determined. That is a clear advantage over the empirical–experimental methods that, for example, determine point estimates of site amplification (e.g., Bard, 1998; Theodulidis & Bard, 1995). Moreover, instead of the PGA parameter and the GMM used in this study, other (and more complex if required) GMMs can be introduced and applied to other strong-motion measures of interest, such as PGV, maximum response of dynamic systems, and ground motion duration. Effectively, therefore, one of the main benefits of the proposed BHM is to facilitate the robust spatial correlation analysis of such earthquake ground motions of interest, along with their uncertainties, which are key factors in local seismic hazard analyses with implications for estimating locally varying seismic risk (e.g., Goda & Atkinson, 2009, 2010; Goda & Hong, 2008; Jayaram & Baker 2009).

5 | CONCLUSIONS

We have introduced a novel BHM for characterizing variations in earthquake ground motion parameters, taken in this study as the geometric mean of the recorded PGA from two horizontal components of each station for each event. The proposed BHM model quantitatively accounts for an event effect, a station effect, an event-station effect, and an unexplained effect. The relative differences in observed PGA attributed to the earthquake source were captured in the BHM by an event effect, and those attributed to the localized site conditions were captured by a station effect. The variability between stations within an event was captured by the spatially varying event-station effect. Thus, the BHM quantifies the local variability in the small region of the array stations and shows to what extent the source and site contribute to that variability.

All model parameters were shown to be well determined and have unimodal and nearly symmetric posterior densities. The event effects were found to be relatively stable across the magnitude and distance ranges of the data and exhibit no clear trends that might justify a modification of the GMM. As a result, the unbiased residuals confirm that the BHM model performs well in modeling the distribution of PGA values across ICEARRAY I. The station effects show clear and systematic variation over the array with significant differences between several stations. However, the station effects are highly correlated as a result of the stations being located on the same geological structure and the array aperture being small compared with the source–site distances. The intrastation variability is an important part of the proposed model. While the extent of the intrastation variability was shown to be low for the area under study, along with the quantification of the station effects, it still has direct practical implications. The high resolution of the spatial distribution of earthquake ground motion and its variability as quantified in this study now forms the basis of a more accurate estimation of the earthquake hazard for this area.

ACKNOWLEDGEMENTS

This work was supported by the Icelandic Centre for Research (Rannís Grant of Excellence no. 141261-051/52/53), the Doctoral grants from Eimskip Fund of the University of Iceland, and the University of Iceland Research Fund. The instruments were funded through a Marie Curie International Re-integration Grant. The ICEARRAY I data set was initially arranged and organized by Judith Marinière, Pierre Turquais, and Romain Millan during their internship at the Earthquake Engineering Research Centre of the University of Iceland. The authors would like to express their gratitude to Hveragerði inhabitants and municipality for housing the recording equipment and for their dedication and support to the ICEARRAY project.

REFERENCES

Abrahamson, N. A. (1988). Statistical properties of peak ground accelerations recorded by the SMART 1 array. *Bulletin of the Seismological Society of America*, *78*(1), 26–41.

Abrahamson, N. A., & Youngs, R. R. (1992). A stable algorithm for regression analyses using the random effects model. *Bulletin of the Seismological Society of America*, *82*(1), 505–510.

Aki, K. (1957). Space and time spectra of stationary stochastic waves, with special reference to microtremors. *Bulletin of the Earthquake Research Institute*, *35*, 415–456.

Ambraseys, N. N., Douglas, J., Sarma, S. K., & Smit, P. M. (2005). Equations for the estimation of strong ground motions from shallow crustal earthquakes using data from Europe and the Middle East: Horizontal peak ground acceleration and spectral acceleration. *Bulletin of Earthquake Engineering*, *3*(1), 1–53.

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

Atkinson, G. M. (2006). Single-station sigma. *Bulletin of the Seismological Society of America*, *96*(2), 446–455.

Baltay, A. S., Hanks, T. C., & Abrahamson, N. A. (2017). Uncertainty, variability, and earthquake physics in ground-motion prediction equations. *Bulletin of the Seismological Society of America*, *107*(4), 1754–1772.

Bard, P. Y. (1998). Microtremor measurements: A tool for site effect estimation? Paper presented at: Proceeding of the Second International Symposium on the Effects of Surface Geology on Seismic Motion, Yokohama, Japan.

Bessason, B., & Kaynia, A. M. (2002). Site amplification in lava rock on soft sediments. *Soil Dynamics and Earthquake Engineering*, *22*(7), 525–540.

Bindi, D., Castro, R. R., Franceschina, G., Luzi, L., & Pacor, F. (2004). The 1997–1998 Umbria-Marche sequence (central Italy): Source, path, and site effects estimated from strong motion data recorded in the epicentral area. *Journal of Geophysical Research: Solid Earth*, *109*(B4).

Boatwright, J., Fletcher, J. B., & Fumal, T. E. (1991). A general inversion scheme for source, site, and propagation characteristics using multiply recorded sets of moderate-sized earthquakes. *Bulletin of the Seismological Society of America*, *81*(5), 1754–1782.

Bolt, B. A., Tsai, Y. B., Yeh, K., & Hsu, M. K. (1982). Earthquake strong motions recorded by a large near-source array of digital seismographs. *Earthquake Engineering and Structural Dynamics*, *10*(4), 561–573.

Boore, D. M. (1983). Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra. *Bulletin of the Seismological Society of America*, *73*(6A), 1865–1894.

Boore, D. M. (2003). Simulation of ground motion using the stochastic method. In *Seismic motion, lithospheric structures, earthquake and volcanic sources: The Keiiti Aki volume* (pp. 635–676). Basel, Switzerland: Birkhäuser.

Boore, D. M., Joyner, W. B., & Fumal, T. (1997). Equations for estimating horizontal response spectra and peak acceleration from North American earthquakes: A summary of recent work. *Seismological Research Letters*, *68*(1), 128–154.

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

Borcherdt, R. D. (1970). Effects of local geology on ground motion near San Francisco Bay. *Bulletin of the Seismological Society of America*, *60*(1), 29–61.

Brandsdóttir, B., Parsons, M., White, R. S., Guðmundsson, O., Drew, J., & Thorbjarnadóttir, B. (2010). The May 29 2008 earthquake aftershock sequence within the South Iceland Seismic Zone: Fault locations and source parameters of aftershocks. *Jökull: Journal of the Glaciological and Geological Societies of Iceland*, *60*, 23–46.

Brooks, S. P., & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, *7*(4), 434–455.

Campbell, K., & Bozorgnia, Y. (2014). NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra. *Earthquake Spectra*, *30*(3, (2014), 1087–1115.

Chen, L., & Faccioli, E. (2013). Single-station standard deviation analysis of 2010–2012 strong-motion data from the Canterbury region, New Zealand. *Bulletin of Earthquake Engineering*, *11*(5), 1617–1632.

Chin, B. H., & Aki, K. (1991). Simultaneous study of the source, path, and site effects on strong ground motion during the 1989 Loma Prieta earthquake: A preliminary result on pervasive nonlinear site effects. *Bulletin of the Seismological Society of America*, *81*(5), 1859–1884.

Chiou, B. S.-J., & Youngs, R. R. (2014). Update of the Chiou and Youngs NGA model for the average horizontal component of peak ground motion and response spectra. *Earthquake Spectra*, *30*(3), 1117–1153.

Chiu, H. C., Amirbekian, R. V., & Bolt, B. A. (1995). Transferability of strong ground-motion coherency between the SMART1 and SMART2 arrays. *Bulletin of the Seismological Society of America*, *85*(1), 342–348.

Cressie, N., & Huang, H.-C. (1999). Classes of nonseparable, spatio-temporal stationary covariance functions. *Journal of the American Statistical Association*, *94*(448), 1330–1339.

DeMets, C., Gordon, R. G., Angus, D. F., & Stein, S. (1994). Effect of recent revisions to the geomagnetic reversal time scale on estimates of current plate motions. *Geophysical Research Letters*, *21*(20), 2191–2194.

Di Giacomo, D., Gallipoli, M. R., Mucciarelli, M., Parolai, S., & Richwalski, S. M. (2005). Analysis and modeling of HVSR in the presence of a velocity inversion: the case of Venosa, Italy. *Bulletin of the Seismological Society of America*, *95*(6), 2364–2372.

Dimitriu, P., Theodulidis, N., & Bard, P. Y. (2000). Evidence of nonlinear site response in HVSR from SMART1 (Taiwan) data. *Soil Dynamics and Earthquake Engineering*, *20*(1–4), 155–165.

Douglas, J., & Halldorsson, B. (2010). On the use of aftershocks when deriving ground-motion prediction equations. Paper presented at: 9th US National and 10th Canadian Conference on Earthquake Engineering (9USN/10CCEE); Toronto, Canada.

Einarsson, P. (1991). Earthquakes and present-day tectonism in Iceland. *Tectonophysics*, *189*(1–4), 261–279.

Einarsson, P. (2008). Plate boundaries, rifts and transforms in Iceland. *Jökull*, *58*(12), 35–58.

Field, E. H., & Jacob, K. H. (1995). A comparison and test of various site-response estimation techniques, including three that are not reference-site dependent. *Bulletin of the Seismological Society of America*, *85*(4), 1127–1143.

Foulser-Piggott, R., & Stafford, P. J. (2012). A predictive model for Arias intensity at multiple sites and consideration of spatial correlations. *Earthquake Engineering & Structural Dynamics*, *41*(3), 431–451.

Gelman, A., Carlin, J., Stern, H. S., & Rubin, D. B. (2014). *Bayesian data analysis*. London, UK: Chapman & Hall/CRC.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457–472.

Goda, K., & Atkinson, G. M. (2009). Probabilistic characterization of spatially correlated response spectra for earthquakes in Japan. *Bulletin of the Seismological Society of America*, *99*(5), 3003–3020.

Goda, K., & Atkinson, G. M. (2010). Intraevent spatial correlation of ground-motion parameters using SK-net data. *Bulletin of the Seismological Society of America*, *100*(6), 3055–3067.

Goda, K., & Hong, H. P. (2008). Spatial correlation of peak ground motions and response spectraspatial correlation of peak ground motions and response spectra. *Bulletin of the Seismological Society of America*, *98*(1), 354–365.

Halldórsson, B., & Sigbjörnsson, R. (2009). The M_w6.3 Ölfus earthquake at 15:45 UTC on 29 May 2008 in South Iceland: ICEARRAY strong-motion recordings. *Soil Dynamics and Earthquake Engineering*, *29*(6), 1073–1083.

Halldórsson, B., Sigbjörnsson, R., Rupakhety, R., & Chanerley, A. A. (2010). Extreme near-fault strong-motion of the M6.3 Ölfus earthquake of 29 May 2008 in South Iceland. Paper presented at: 14th European Conference on Earthquake Engineering (14ECEE), Ohrid, Macedonia.

Halldórsson, B., Sigbjörnsson, R., & Schweitzer, J. (2009). ICEARRAY: The first small-aperture, strong-motion array in Iceland. *Journal of Seismology*, *13*(1), 173–178.

Harichandran, R. S. (1991). Estimating the spatial variation of earthquake ground motion from dense array recordings. *Structural Safety*, *10*(1–3), 219–233.

Jayaram, N., & Baker, J. W. (2009). Correlation model for spatially distributed ground-motion intensities. *Earthquake Engineering and Structural Dynamic*, *38*(15), 1687–1708.

Kowsari, M., Halldorsson, B., Hrafnkelsson, B., Snæbjörnsson, J. T., Ólafsson, S., & Rupakhety, R. (2017). On the selection of ground-motion prediction equations for seismic hazard assessment in the South Iceland seismic zone. Paper presented at: 16th World Conference on Earthquake Engineering (16WCEE), Santiago, Chile.

Kuehn, N. M., & Scherbaum, F. (2015). Ground-motion prediction model building: a multilevel approach. *Bulletin of Earthquake Engineering*, *13*(9), 2481–2491.

Kuehn, N. M., & Scherbaum, F. (2016). A partially non-ergodic ground-motion prediction equation for Europe and the Middle East. *Bulletin of Earthquake Engineering*, *14*(10), 2629–2642.

Landwehr, N., Kuehn, N. M., Scheffer, T., & Abrahamson, N. (2016). A nonergodic ground-motion model for California with spatially varying coefficients. *Bulletin of the Seismological Society of America*, *106*(6), 2574–2583.

Lin, P.-S., Chiou, B., Abrahamson, N., Walling, M., Lee, C.-T., & Cheng, C.-T. (2011). Repeatable source, site, and path effects on the standard deviation for empirical ground-motion prediction models. *Bulletin of the Seismological Society of America*, *101*(5), 2281–2295.

Loh, C.-H. (1985). Analysis of the spatial variation of seismic waves and ground movements from smart-1 array data. *Earthquake Engineering and Structural Dynamics*, *13*(5), 561–581.

Loh, C.-H., Ang, A. H.-S., & Wen, Y. K. (1983). Spatial correlation study of strong motion array data with application to lifeline earthquake engineering. Champaign, IL: University of Illinois at Urbana-Champaign.

Meremonte, M., Frankel, A., Cranswick, E., Carver, D., & Worley, D. (1996). Urban seismology—Northridge aftershocks recorded by multi-scale arrays of portable digital seismographs. *Bulletin of the Seismological Society of America*, *86*(5), 1350–1363.

Morikawa, N., Kanno, T., Narita, A., Fujiwara, H., Okumura, T., Fukushima, Y., & Guerpinar, A. (2008). Strong motion uncertainty determined from observed records by dense network in Japan. *Journal of Seismology*, *12*(4), 529–546.

Moya, A., Aguirre, J., & Irikura, K. (2000). Inversion of source parameters and site effects from strong ground motion records using genetic algorithms. *Bulletin of the Seismological Society of America*, *90*(4), 977–992.

Nakamura, Y. (1989). A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface. *Quarterly Report of Railway Technical Research Institute*, *30*(1), 25–33.

Olafsson, S., & Sigbjörnsson, R. (2000). Attenuation of strong ground acceleration: A study of the South Iceland earthquakes 2000. Manuscript submitted for publication.

Ortiz-Alemán, C., Reyes-Olvera, M., Iglesias-Mendoza, A., Orozco-Del-Castillo, M. G., & Hernández-Gómez, J. J. (2017). Estimation of source; path and site effects at MASE array stations: A comprehensive study. *Geofisica Internacional*, *56*(1), 71–85.

Rahpeyma, S., Halldorsson, B., & Green, R. A. (2017). On the distribution of earthquake strong-motion amplitudes and site effects across the icelandic strong-motion arrays. Paper presented at: 16th World Conference on Earthquake Engineering (16WCEE), Santiago, Chile.

Rahpeyma, S., Halldorsson, B., Olivera, C., Green, R. A., & Jónsson, S. (2016). Detailed site effect estimation in the presence of strong velocity reversals within a small-aperture strong-motion array in Iceland. *Soil Dynamics and Earthquake Engineering*, *89*, 136–151.

Roberts, G. O., Gelman, A., & Gilks, W. R. (1997). Weak convergence and optimal scaling of random walk Metropolis algorithms. *The Annals of Applied Probability*, *7*(1), 110–120.

Rodriguez-Marek, A., Cotton, F., Abrahamson, N. A., Akkar, S., Al Atik, L., Edwards, … Dawood, H. M. (2013). A model for single-station standard deviation using data from various tectonic regions. *Bulletin of the Seismological Society of America*, *103*(6), 3149–3163.

Rodriguez-Marek, A., Montalva, G. A., Cotton, F., & Bonilla, F. (2011). Analysis of single-station standard deviation using the KiK-net data. *Bulletin of the Seismological Society of America*, *101*(3), 1242–1258.

Sæmundsson, K. (1979). Outline of the geology of Iceland. *Jökull*, *29*, 7–28.

Sæmundsson, K., & Kristinsson, S. (2005). Hveragerði. Soil temperature measurements and faults. Reykjavík, Iceland: Iceland GeoSurvey (ÍSOR) (in Icelandic).

Shabestari, K. T., Yamazaki, F., Saita, J., & Matsuoka, M. (2004). Estimation of the spatial distribution of ground motion parameters for two recent earthquakes in Japan. *Tectonophysics*, *390*(1), 193–204.

Sigbjörnsson, R., Sigurdsson, T., Snæbjörnsson, J. T., & Valsson, G. (2006). Mapping of crustal strain rate tensor for Iceland with applications to seismic hazard assessment. Paper presented at: First European Conference on Earthquake Engineering and Seismology (1ECEES), Geneva, Switzerland.

Sigbjörnsson, R., Snæbjörnsson, J. T., Higgins, S. M., Halldórsson, B., & Ólafsson, S. (2009). A note on the M_w 6.3 earthquake in Iceland on 29 May 2008 at 15:45 UTC. *Bulletin of Earthquake Engineering*, *7*(1), 113–126.

Sigurdarson, A. N., & Hrafnkelsson, B. (2016). Bayesian prediction of monthly precipitation on a fine grid using covariates based on a regional meteorological model. *Environmetrics*, *27*(1), 27–41.

Simpson, D., Rue, H., Riebler, A., Martins, T. G., & Sørbye, S. H. (2017). Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science*, *32*(1), 1–28.

Stefánsson, R., Böðvarsson, R., Slunga, R., Einarsson, P., Jakobsdóttir, S., Bungum, H., … Korhonen, H. (1993). Earthquake prediction research in the South Iceland Seismic Zone and the SIL project. *Bulletin of the Seismological Society of America*, *83*(3), 696–716.

Stefánsson, R., Guðmundsson, G. B., & Halldórsson, P. (2008). Tjörnes fracture zone. New and old seismic evidences for the link between the North Iceland rift zone and the Mid-Atlantic ridge. *Tectonophysics*, *447*(1–4), 117–126.

Strasser, F. O., Abrahamson, N. A., & Bommer, J. J. (2009). Sigma: Issues, insights, and challenges. *Seismological Research Letters*, *80*(1), 40–56.

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

Wang, M., & Takada, T. (2005). Macrospatial correlation model of seismic ground motions. *Earthquake Spectra*, *21*(4), 1137–1156.

Zafarani, H., & Soghrat, M. R. (2017). Single-station Sigma for the Iranian strong motion stations. *Pure and Applied Geophysics*, *174*(11), 4077–4099.

Zhang, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *Journal of the American Statistical Association*, *99*(465), 250–261.

How to cite this article: Rahpeyma S, Halldorsson B, Hrafnkelsson B, Jónsson S. Bayesian hierarchical model for variations in earthquake peak ground acceleration within small-aperture arrays. *Environmetrics*. 2018;e2497. https://doi.org/10.1002/env.2497

APPENDIX

A.1 | Bayesian hierarchical model

The three levels of the spatial BHM proposed in Section 3 for the peak ground acceleration parameter of each station of an array are presented below in a detailed mathematical format.

1. Data level

At the data level, the conditional density of the observed data is given by

$$\pi\left(\boldsymbol{y}|\boldsymbol{\beta},\boldsymbol{\delta B},\boldsymbol{\delta S}\right)=\mathrm{N}\left(\boldsymbol{y}|\boldsymbol{X\beta}+\boldsymbol{Z}_1\boldsymbol{\delta B}+\boldsymbol{Z}_2\boldsymbol{\delta S},\boldsymbol{\Sigma}_y\right)=\mathrm{N}\left(\boldsymbol{y}|\,\mathbf{K}\boldsymbol{\eta},\boldsymbol{\Sigma}_y\right), \tag{A.1}$$

where $\boldsymbol{y}$ represents a vector containing all the observations Y_{es}, ordered by station number s for the first event, $e=1$, followed by order according to station number for the second event, $e=2$, and so on. The matrices $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ are index matrices. Assuming there are no missing observations at the nominated stations, $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ are then defined as

$$\boldsymbol{Z}_1=\boldsymbol{I}_N\otimes\mathbf{1}_Q, \tag{A.2}$$

$$\boldsymbol{Z}_2=\mathbf{1}_N\otimes\boldsymbol{I}_Q, \tag{A.3}$$

where $\otimes$ denotes the Kronecker product, $\mathbf{1}_n$ is an n-element vector of ones, and $\boldsymbol{I}_n$ is an $n\times n$ identity matrix. N and Q are the total number of events and the total number of stations, respectively. The matrix $\boldsymbol{X}$ is a $(NQ)\times p$ dimensional matrix containing the linear predictors for all combinations of stations and events. It is such that the first column is the constant vector, $\mathbf{1}_{NQ}$, and the next $(p-1)$ vectors contain the linear predictors. The vector $\boldsymbol{\eta}$ contains the vectors $\boldsymbol{\beta}$, $\boldsymbol{\delta B}$, and $\boldsymbol{\delta S}$. $\mathbf{K}$ is defined as a $(NQ)\times(p+N+Q)$ dimensional matrix of linear predictors and index matrices; see Equation (A.4). Furthermore, $\boldsymbol{\Sigma}_y$ is a $(NQ)\times(NQ)$ dimensional block-diagonal matrix composed as Equation (A.5), as follows:

$$\mathbf{K}=[\boldsymbol{X}\quad\boldsymbol{Z}_1\quad\boldsymbol{Z}_2], \tag{A.4}$$

$$\boldsymbol{\Sigma_y}=\phi_R^2\,\boldsymbol{I}_{NQ}+\boldsymbol{I}_N\otimes\boldsymbol{\Sigma}_{SS}=\begin{pmatrix}\phi_R^2\boldsymbol{I}_Q+\boldsymbol{\Sigma}_{SS} & \cdots & \mathbf{0}\\ \vdots & \ddots & \vdots\\ \mathbf{0} & \cdots & \phi_R^2\boldsymbol{I}_Q+\boldsymbol{\Sigma}_{SS}\end{pmatrix}, \tag{A.5}$$

where

$$\{\boldsymbol{\Sigma}_{SS}\}_{ij}=\left\{\phi_{SS}^2\left(1+\frac{d_{ij}}{\Delta_{SS}}\right)\exp\left(-\frac{d_{ij}}{\Delta_{SS}}\right)\right\}_{ij}, \tag{A.6}$$

d_{ij} is the interstation distance between stations i and j, and the covariance matrix, $\{\boldsymbol{\Sigma}_{SS}\}_{ij}$, is based on a Matérn covariance function with standard deviation ϕ_{SS}, smoothness parameter $\nu_{SS} = 1.5$, and range parameter, Δ_{SS}. It should be highlighted that rows and columns are removed from all the matrices and vectors according to missing data points.

1. *Latent level*

The vector $\boldsymbol{\eta}$ consists of the parameters $\boldsymbol{\beta}$, $\boldsymbol{\delta B}$, and $\boldsymbol{\delta S}$ at the latent level of the BHM. A priori $\boldsymbol{\eta}$ follows a normal distribution with mean $\boldsymbol{\mu}_\eta$ and covariance matrix $\boldsymbol{\Sigma}_\eta$ where

$$\boldsymbol{\eta} = \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\delta B} \\ \boldsymbol{\delta S} \end{pmatrix}, \quad \boldsymbol{\mu}_\eta = \begin{pmatrix} \boldsymbol{\mu}_\beta \\ \mathbf{0}_N \\ \mathbf{0}_Q \end{pmatrix}, \quad \boldsymbol{\Sigma}_\eta = \begin{pmatrix} \boldsymbol{\Sigma}_\beta & 0 & 0 \\ 0 & \tau^2 \boldsymbol{I}_N & 0 \\ 0 & 0 & \boldsymbol{\Sigma}_{S2S} \end{pmatrix}, \tag{A.7}$$

where $\mathbf{0}_D$ is a column vector of zeros of length D, $\boldsymbol{\mu}_\beta = \mathbf{0}_p$, $\boldsymbol{\Sigma}_\beta = 100^2 \boldsymbol{I}_p$. $\boldsymbol{\Sigma}_{S2S}$ is a covariance matrix based on the Matérn covariance function with standard deviation, ϕ_{S2S}, smoothness parameter $\nu_{S2S} = 0.5$, and range parameter, Δ_{S2S} and defined as Equation (A.8), as follows:

$$\{\boldsymbol{\Sigma}_{S2S}\}_{ij} = \left\{ \phi_{S2S}^2 \exp\left(-\frac{d_{ij}}{\Delta_{S2S}} \right) \right\}_{ij}. \tag{A.8}$$

1. *Hyperparameter level*

The vector $\boldsymbol{\theta} = (\tau, \phi_{S2S}, \phi_{SS}, \phi_R, \Delta_{SS})$ contains the unknown hyperparameters. The parameter Δ_{S2S} is difficult to infer. Thus, it is set equal to a fixed value, $\Delta_{S2S} = 0.06$ km. We assume a priori that the hyperparameters are independent and that each of them follows an exponential distribution as specified in Equation (A.9) based on Simpson, Rue, Riebler, Martins, and Sørbye (2017), as follows:

$$\begin{aligned} \tau \sim \text{Exp}(\lambda_\tau), \quad & \phi_{S2S} \sim \text{Exp}\left(\lambda_{\phi_{S2S}}\right), \quad \phi_R \sim \text{Exp}\left(\lambda_{\phi_R}\right), \\ \phi_{SS} \sim \text{Exp}\left(\lambda_{\phi_{SS}}\right), \quad & \Delta_{SS} \sim \text{Exp}\left(\lambda_{\Delta_{SS}}\right). \end{aligned} \tag{A.9}$$

The probability density function of an exponential random variable is defined as

$$\pi(\theta|\lambda) = \lambda e^{-\lambda\theta}, \tag{A.10}$$

where λ is the rate parameter and λ^{-1} is the mean of the exponential distribution. The range parameter for the event-site effect, Δ_{SS}, can reflect roughly half of the largest interstation distance. The total variance of the observed $\log_{10} PGA_{es}$ is around 0.2255. We assume a priori that the four independent variances τ^2, ϕ_{S2S}^2, ϕ_{SS}^2, and ϕ_R^2 contribute equally to the total variability of the data (cf. Equation (5)). This is formulated by assigning them identical prior medians. This results in a prior median equal to 0.0564. The median of the prior densities for the standard deviations τ, ϕ_{S2S}, ϕ_{SS}, and ϕ_R is 0.2374, and the rate parameter is 4.21 and is rounded up to 5.0. Hence, the rate parameters in the prior densities are set as succeeding

$$\lambda_\tau = 5.0, \quad \lambda_{\phi_{S2S}} = 5.0, \quad \lambda_{\phi_{SS}} = 5.0, \quad \lambda_{\phi_R} = 5.0, \quad \lambda_{\Delta_{SS}} = 0.7.$$

In this study, we implement the Metropolis algorithm on the transformed hyperparameters with the aim of avoiding sampling negative values of $\boldsymbol{\theta}$ and improving the efficiency of the sampling process. Hence, the log-transformed hyperparameters will be $\boldsymbol{\psi} = (\psi_1, \psi_2, \psi_3, \psi_4, \psi_5) = (\log(\tau), \log(\phi_{S2S}), \log(\phi_{SS}), \log(\phi_R), \log(\Delta_{SS}))$. The marginal prior density of ψ_i is

$$\pi(\psi_i|\lambda_i) = \lambda_i \exp(-\lambda_i \exp(\psi_i) + \psi_i), \tag{A.11}$$

where

$$\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5) = \left(\lambda_\tau, \lambda_{\phi_{S2S}}, \lambda_{\phi_{SS}}, \lambda_{\phi_R}, \lambda_{\Delta_{SS}}\right). \tag{A.12}$$

A.2 | Posterior inference

The joint posterior density is defined as Equation (A.13), as follows:

$$\begin{gathered}\pi\ (\boldsymbol{\eta},\boldsymbol{\psi}|\boldsymbol{y}) \propto \pi(\boldsymbol{y}|\boldsymbol{\eta},\boldsymbol{\psi})\ \pi(\boldsymbol{\eta}|\boldsymbol{\psi})\ \pi(\boldsymbol{\psi}) \\ \pi\ (\boldsymbol{\eta}|\boldsymbol{\psi}) \propto \pi(\boldsymbol{\beta})\ \pi(\boldsymbol{\delta B}|\psi_1)\ \pi(\boldsymbol{\delta S}|\psi_2,\Delta_{S2S}) \\ \pi\ (\boldsymbol{\psi}) \propto \prod_{k=1}^{5} \pi(\psi_k)\,.\end{gathered} \tag{A.13}$$

The posterior distribution of $\boldsymbol{\eta}$ conditioned on data and hyperparameters is given by Equation (A.14), as follows:

$$\begin{aligned}\pi\ (\boldsymbol{\eta}|\boldsymbol{y},\boldsymbol{\psi}) &\propto \pi(\boldsymbol{y}|\boldsymbol{\eta},\boldsymbol{\psi})\,\pi(\boldsymbol{\eta}|\boldsymbol{\psi}) \\ &\propto \mathrm{N}\big(\boldsymbol{y}|\mathbf{K}\,\boldsymbol{\eta},\boldsymbol{\Sigma}_y\big)\ \mathrm{N}\big(\boldsymbol{\eta}|\boldsymbol{\mu}_\eta,\boldsymbol{\Sigma}_\eta\big) \\ &\propto \exp\left(-\tfrac{1}{2}(\boldsymbol{y}-\boldsymbol{K\eta})^T\boldsymbol{\Sigma}_y^{-1}\,(\boldsymbol{y}-\boldsymbol{K\eta})\right)\exp\left(-\tfrac{1}{2}\big(\boldsymbol{\eta}-\boldsymbol{\mu}_\eta\big)^T\boldsymbol{\Sigma}_\eta^{-1}\,\big(\boldsymbol{\eta}\ \boldsymbol{\mu}_\eta\big)\right) \\ &= \mathrm{N}\left(\boldsymbol{\eta}|\boldsymbol{\mu}_{\eta,\mathrm{post}},\boldsymbol{\Sigma}_{\eta,\mathrm{post}}\right),\end{aligned} \tag{A.14}$$

where

$$\begin{aligned}\boldsymbol{\mu}_{\eta,\mathrm{post}} &= \left(\boldsymbol{\Sigma}_\eta^{-1}+\mathbf{K}^T\boldsymbol{\Sigma}_y^{-1}\mathbf{K}\right)^{-1}\left(\boldsymbol{\Sigma}_\eta^{-1}\ \boldsymbol{\mu}_\eta+\mathbf{K}^T\boldsymbol{\Sigma}_y^{-1}\boldsymbol{y}\right), \\ \boldsymbol{\Sigma}_{\eta,\mathrm{post}} &= \left(\boldsymbol{\Sigma}_\eta^{-1}+\mathbf{K}^T\boldsymbol{\Sigma}_y^{-1}\mathbf{K}\right)^{-1}.\end{aligned}$$

The marginal posterior distribution of the log transform of hyperparameters, $\boldsymbol{\psi}$, is defined as Equation (A.15), as follows:

$$\pi\ (\boldsymbol{\psi}|\boldsymbol{y}) \propto \pi(\boldsymbol{y}|\boldsymbol{\psi})\ \pi(\boldsymbol{\psi})\quad . \tag{A.15}$$

Consequently, $\boldsymbol{\Sigma}_\eta$ and $\boldsymbol{\Sigma}_y$ will be functions of ψ given by Equation (A.16) and Equation (A.17), respectively, as follows:

$$\boldsymbol{\Sigma}_\eta\,(\psi) = \begin{pmatrix} \boldsymbol{\Sigma}_\beta & 0 & 0 \\ 0 & \exp(2\psi_1)\,\boldsymbol{I}_N & 0 \\ 0 & 0 & \left\{\exp(2\psi_2)\exp\left(-\frac{d_{ij}}{\Delta_{S2S}}\right)\right\}_{ij} \end{pmatrix}, \tag{A.16}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{y}}\,(\psi) = \exp(2\psi_4)\,\boldsymbol{I}_{NQ} + \boldsymbol{I}_N \otimes \left\{\exp(2\psi_3)\left(1+\frac{d_{ij}}{\exp(\psi_5)}\right)\exp\left(-\frac{d_{ij}}{\exp(\psi_5)}\right)\right\}_{ij}. \tag{A.17}$$

It can be shown that the density $\pi(\boldsymbol{y}\,|\,\boldsymbol{\psi})$ is a normal density (see Equation (A.20)) with mean and covariance given by Equation (A.18) and Equation (A.19), as follows:

$$\mathrm{E}\,[\,\boldsymbol{y}] = \mathbf{K}\,\boldsymbol{\mu}_\eta, \tag{A.18}$$

$$\mathrm{cov}\ [\,\boldsymbol{y}] = \mathbf{K}\,\boldsymbol{\Sigma}_\eta\,\mathbf{K}^T + \boldsymbol{\Sigma}_y, \tag{A.19}$$

thus,

$$\pi(\boldsymbol{y}|\boldsymbol{\psi}) = \mathrm{N}\left(\boldsymbol{y}|\mathbf{K}\,\boldsymbol{\mu}_\eta,\mathbf{K}\,\boldsymbol{\Sigma}_\eta\mathbf{K}^T+\boldsymbol{\Sigma}_y\right). \tag{A.20}$$

The marginal posterior density of the hyperparameters can be obtained as Equation (A.21), as follows:

$$\begin{aligned}\pi(\boldsymbol{\psi}|\boldsymbol{y}) &\propto \mathrm{N}\big(\boldsymbol{y}|\mathbf{K}\,\boldsymbol{\mu}_\eta,\mathbf{K}\,\boldsymbol{\Sigma}_\eta\,\mathbf{K}^T+\boldsymbol{\Sigma}_y\big)\prod_{k=1}^{5}\exp(\psi_k|\lambda_k) \\ &\propto \left|\mathbf{K}\,\boldsymbol{\Sigma}_\eta\,\mathbf{K}^T+\boldsymbol{\Sigma}_y\right|^{-\frac{1}{2}}\exp\left(-\tfrac{1}{2}\left(\big(\boldsymbol{y}-\mathbf{K}\,\boldsymbol{\mu}_\eta\big)^T\left(\mathbf{K}\,\boldsymbol{\Sigma}_\eta\mathbf{K}^T+\boldsymbol{\Sigma}_y\right)^{-1}\big(\boldsymbol{y}-\mathbf{K}\,\boldsymbol{\mu}_\eta\big)\right)\right) \\ &\quad \prod_{k=1}^{5}\lambda_k\exp(-\lambda_k\exp(\psi_k)+\psi_k)\,.\end{aligned} \tag{A.21}$$

Then, the logarithm of the $\pi(\boldsymbol{\psi}|\boldsymbol{y})$ can be calculated as Equation (A.22), as follows:

$$\begin{gathered}\log\left(\pi\left(\boldsymbol{\psi}|\boldsymbol{y}\right)\right) = C-\tfrac{1}{2}\log\left|\mathbf{K}\,\boldsymbol{\Sigma}_\eta\,\mathbf{K}^T+\boldsymbol{\Sigma}_y\right| \\ -\tfrac{1}{2}\left[\big(\boldsymbol{y}-\mathbf{K}\,\boldsymbol{\mu}_\eta\big)^T\big(\mathbf{K}\,\boldsymbol{\Sigma}_\eta\mathbf{K}^T+\boldsymbol{\Sigma}_y\big)^{-1}\big(\boldsymbol{y}-\mathbf{K}\,\boldsymbol{\mu}_\eta\big)\right]+\sum_{k=1}^{5}\left\{\log\lambda_k-\lambda_k\exp(\psi_k)+\psi_k\right\}.\end{gathered} \tag{A.22}$$

Paper III

Site effect estimation on two Icelandic strong-motion arrays using a Bayesian hierarchical model of spatial distribution of earthquake peak ground acceleration

Sahar Rahpeyma, Benedikt Halldorsson, Birgir Hrafnkelsson, Russel A. Green, and Sigurjon Jónsson

Rahpeyma S, Halldorsson B, Hrafnkelsson B, Jónsson S (2018a) Site effect estimation on two Icelandic strong-motion arrays using a Bayesian hierarchical model of spatial variation of earthquake peak ground acceleration. *Soil Dynamics and Earthquake Engineering (Accepted, in press)*

Site effect estimation on two Icelandic strong-motion arrays using a Bayesian hierarchical model for the spatial distribution of earthquake peak ground acceleration

Sahar Rahpeyma[1], Benedikt Halldorsson[2*,3], Birgir Hrafnkelsson[4], Russell A. Green[5], and Sigurjón Jónsson[6]

Abstract: Earthquake recordings on two small-aperture (covering ~ 1.2 km^2 each) strong-motion arrays in Iceland (ICEARRAY I and II) exhibit considerable variations in the spatial distribution of ground-motion amplitudes. To better understand this spatial variability, we use a Bayesian Hierarchical Model (BHM) that incorporates ground motions models (GMMs) for peak ground accelerations (PGA) developed from ground motion databases recorded by the two arrays, respectively. The posterior distributions of the model parameters are then determined using Markov Chain Monte Carlo simulations in the context of Bayesian statistical methods. The BHM allows the partitioning of a GMM into event, station, and event-station terms, which in turn allows the relative contributions of source, path, and site effects to be quantified. The results indicate that site effects can dominate the spatial distribution of ground-motion parameters (e.g., PGA) observed across both ICEARRAY I and II. Although the site conditions across ICEARRAY I have been classified as uniform (i.e., "rock" with a relatively flat topography), station terms contribute ~ 13% to the total variability in the amplitudes of predicted ground motions across the array. In contrast to ICEARRAY I, the variation of the geologic profiles and topography is much greater across ICEARRAY II. As a result, the inter-station variability is shown to contribute up to ~ 57% of the total variability in the amplitudes of predicted ground motions across the array, with the contributions being less constrained for ICEARRAY II than ICEARRAY I due to the relative sizes of the recorded ground motion databases. These results facilitate our understanding of the key factors that affect the variation of seismic ground motions across a relatively small area. Such a detailed microzonation is of great importance for earthquake hazard assessment on a local scale and has practical implications for engineering decision making.

Keywords: *Peak Ground Acceleration, Bayesian Hierarchical Model, strong-motion array, parameter estimation, parameter uncertainty, site effects.*

[1] Earthquake Engineering Research Centre & Faculty of Civil and Environmental Engineering, School of Engineering and Natural Sciences, University of Iceland, Selfoss, Iceland (sahar@hi.is)

[2*] Earthquake Engineering Research Centre & Faculty of Civil and Environmental Engineering, School of Engineering and Natural Sciences, University of Iceland, Selfoss, Iceland (skykkur@hi.is) (corresponding author)

[3] Geoscience Research Group, Division of Processing and Research, Icelandic Meteorological Office, Reykjavík, Iceland

[4] Department of Mathematics, Faculty of Physical Sciences, School of Engineering and Natural Sciences, University of Iceland, Reykjavik, Iceland (birgirhr@hi.is)

[5] Department of Civil & Environmental Engineering, Virginia Tech, Blacksburg, VA, USA (rugreen@vt.edu)

[6] King Abdullah University of Science and Technology (KAUST), Physical Science and Engineering Division (PSE), Thuwal, 23955-6900, Saudi Arabia (sigurjon.jonsson@kaust.edu.sa)

1. Introduction

Considerable variation of earthquake strong ground motion over small distances can significantly influence the seismic response of lifelines such as bridges, pipelines, and other structures of considerable spatial extent [1]. The significance of site effects on the spatial variation of ground motion intensities was highlighted by the catastrophic 1985, M_w8.0 Mexico City earthquake. In this event, extreme structural collapses, resulting in numerous injuries and fatalities, occurred in some parts of the city while nearby adjacent areas of the city were relatively unaffected, with the differences in the amount of damage being largely attributed to differences in the local ground motions [2,3]. The spatial variation of earthquake ground motions results from the combination of wave-passage effects and the natural incoherence of ground motion and varies both with frequency and distance. Accordingly, during recent decades, many researchers have attempted to quantitatively estimate the contribution of the earthquake source, path, and site effects to the total variability of ground motions [4,5]. It is generally acknowledged that localized site effects on earthquake ground motion are systematic and repeatable for a given site. This is because site effects are a manifestation of the impact that the shallow geological structure can have on the amplification and frequency content of seismic waves. Accordingly, if the geologic profiles vary between locations, relative differences should be expected in site effects which contribute to the incoherency of strong-motion [e.g., 6–11].

Small-aperture seismic arrays are powerful tools that, among other things, enable a comprehensive analysis of the variation of earthquake ground motions at small spatial scales [12]. The analysis of the SMART1 dataset in Taiwan, recorded on the first dense digital small-aperture array of strong-motion seismographs, showed a conspicuous variation in ground motion intensities across the array. This variation in ground motion intensities among the SMART1 stations has led to multiple studies focused on the analysis and modeling of the spatial variation of ground motion [1,3,12–18]. The analysis of the SMART1 recordings shed light on the association of the ground motion variability over short distances with the variation in the underlying geological profiles. In addition to SMART1, attempts have been made to characterize variability of seismic ground motions by using other closely spaced networks and seismic arrays [3,6,11,19–21]. The findings of these studies have highlighted the important role of local dense measurements in making detailed estimates of site effects across relatively small areas as large networks or individual strong-motion seismographs cannot provide the required resolution.

The deployment of the first Icelandic strong-motion array (ICEARRAY I) in the town of Hveragerði in the South Island Seismic Zone (SISZ) in 2007 commenced a new era in earthquake strong-motion monitoring in Iceland. The array covers around 1.23 km^2 and consists of 12 accelerometric stations with inter-station distances ranging from 50-1900 meters [22] (see Fig. 1). A second strong-motion array (ICEARRAY II) was deployed in the Tjörnes Fracture Zone (TFZ) in Húsavík (North Iceland) in 2012, with the array located effectively on top of the Húsavík-Flatey Fault (HFF), the largest transform fault in the country. ICEARRAY II consists of 6 free-field stations and one structural monitoring system, three of the stations are collocated with a linear array of high-frequency sampling continuous GPS instruments [23,24]. In the decade since their deployment, the arrays have dramatically increased the size of the Icelandic strong-motion dataset. In particular, ICEARRAY I recorded the extreme near-fault motions of the M_w6.3 Ölfus earthquake on 29 May 2008 and more than 1700 of its aftershocks over the following year. ICEARRAY II has, on the other hand, recorded much less data despite capturing two of the largest seismic sequences in the last 30 years in North Iceland; the array recorded far-field motions from 26 small-to-moderate size earthquakes [25,26].

The surface geology across ICEARRAY I is characterized as "rock" (see Fig. 2 and [11] and references therein). As a result, one would expect small variability in ground motion amplitudes across the array. However, preliminary analysis showed consistent and considerable variations in peak-ground accelerations (PGA), peak-ground velocity (PGV), and pseudo-spectral acceleration (PSA) of the recorded motions. To better understand this, we previously carried out a detailed analysis of the frequency dependence of site effects across ICEARRAY I [11]. The results from that study revealed noticeable differences between the lava-rock structure vs. the old bedrock that underlies portions of the array [11,26]. Namely, most of Hveragerði is located on a young and competent lava-rock layer, which in turn is underlain by a layer of softer sedimentary deposits (see Fig. 2). This geologic structure is characterized by a significant shear wave velocity reversal and is repeated with depth until the old bedrock is reached at ~ 70 m. Thus, the site effects under this part of the town are expected to be affected by at least two velocity reversals. As a result, two separate and predominant frequencies were observed in the horizontal-to-vertical-spectral-ratios (HVSR) recorded in the portion of the town underlain by the lava rock profiles. These profiles act as a transfer function that amplify the surface ground motions at certain predominant frequencies, with the extent of the amplification depending on the frequency content of the

incident seismic motion [11,26]. There are considerable variations in the recorded peak ground motion amplitudes, even for motions recorded on seemingly similar layered geologic profiles. Although the spatial distribution of events may introduce an azimuthal effect on the amplitudes of ground motions [27,28], the recorded aftershocks covered almost the entire azimuthal range around ICEARRAY I and no azimuthal dependency on amplitudes was observed [11].

In contrast to ICEARRAY I, the topography and near surface geology in the town of Húsavík in North Iceland, where ICEARRAY II is located, is more variable and complex. The geologic profiles underlying ICEARRAY II are characterized by four different types of geological units of varying stiffness and thicknesses across the town [29]. The ground motions in Húsavík are thus expected to be more variable than in Hveragerði, and early results have shown this [26]. Therefore, the two ICEARRAY networks present two very different cases, i.e., one case where the geological conditions appear uniform and limited ground motion variability is expected and then another case where the variable geology and topography should lead to significantly larger differences in ground motion.

As stated above, complex earthquake source, wave propagation, and local geological effects all contribute to the variation of ground motion that can cause essentially identical structures to respond significantly different during the same event, even over a relatively small area. The simplest and most common approach in engineering practice to quantify the systematic dependence of the frequency dependent earthquake ground motion amplitudes is by using ground motion models (GMMs). In particular, empirical GMMs are developed by fitting a regression formula (e.g., by means of least-square method) having a particular functional form to a database of observed ground motion parameters. The residuals of the GMMs relative to the observations are generally expected to follow a zero-mean normal distribution for which a standard deviation (i.e., the random variability of ground motions) is estimated.

The uncertainty associated with predicting ground motions has a significant influence on computed seismic hazard, with a larger GMM standard deviation resulting in higher computed seismic hazard, especially at low probability levels [8]. As a result, many researchers have expanded efforts to develop more physically justified empirical GMMs, with the hopes of decreasing the associated standard deviation. However, using more complicated functional forms of GMMs over the last decades [30–34] has not led to the expected decrease in the total variability. On this note, the investigation of the main sources

and physical understanding of uncertainties associated with GMMs is a challenging research topic that has important implications for the seismic capacity of critical structures [8,35–37]. During recent decades, many researchers have tried to split up the variabilities into independent effects and quantify the relative contributions of each to the overall predicted ground motion variability [38–42]. The objective is to separate the variation of predicted ground motions in order to better quantify the contributions from source, station, and path effects. Commonly, terms for the effects can be approximated using parameterized empirical models (e.g., [43–47]) or empirical techniques [48–51]. However, the inversion problems are mostly nonlinear, and the solution can be affected by non-uniqueness which leads to high level of uncertainties.

In this study we use a practical Bayesian inference scheme that we recently developed [52] and gives us better understanding of the distribution of ground motion parameters, as well as quantifying the corresponding uncertainties. The proposed model, the so-called Bayesian Hierarchical Model (BHM), offers a flexible probabilistic framework for multi-level modeling of earthquake ground motion parameters, in which a collection of random variables can be decomposed into a series of conditional models. We point out that the ground motion parameter that we model in this study is PGA due to its prevalence in the literature and use in earthquake building standards, but also note that the method used herein can be applied to any ground motion parameter of interest. The BHM is a practical statistical tool written in multiple levels. The Bayes theorem yields the posterior distribution of the unknown model parameters and is used for statistical inference and provides posterior estimates and intervals for the parameters. In general, a BHM for the data is constructed hierarchically in three levels: data level, latent level, and hyperparameter level. The data level describes the distributional model for the observation conditioned on the model parameters. The latent level describes the distribution of the latent parameters, some of which are found in the distributional model of the observations. Furthermore, the joint distribution of the latent parameters is conditioned on the parameters in the hyperparameter level; these parameters are referred to as hyperparameters. A typical hyperparameter is a standard deviation of the distribution describing the corresponding latent parameter. The hyperparameter level includes prior distributions for the hyperparameters at the data and latent levels [53,54]. The structure of the BHM is such that different ground motion parameters of interest are split into subsets of event, station, event-station, and unexplained terms [42,55,56]. Since separating the residuals and variabilities into inter-event and intra-event, respectively, is both desired

from a physical (e.g., source studies [57]) and practical (e.g., in seismic hazard studies [4]) standpoint, we estimate both inter-event and intra-event standard deviations to account for inter-station variability within the two strong-motion arrays in Iceland, respectively [42]. In the following, we first present the ground motion datasets recorded on ICEARRAY I and II, followed by an overview of the BHM, a description of the results of implementing the model, and a discussion of the results.

2. Array Strong-motion Data

2.1. ICEARRAY I in Southwest Iceland

On 29 May 2008 at 15:45 local time, the M_{w}6.3 Ölfus earthquake occurred in the western part of the SISZ. The small-aperture strong-motion array (ICEARRAY I) in the town of Hveragerði of the SISZ recorded the main-shock and more than 1700 of its aftershocks over the course of several months (see Fig. 1b). The first motion originated approximately 6.5 km east-southeast of the town and the spatial distribution of aftershocks outlined a pair of parallel, near-vertical north-south striking right-lateral strike slip faults that are approximately 4.5 km apart. While the epicenter was located on the eastern fault, ~ 6.5 km S-E of Hveragerði, the majority of the aftershocks occurred on the western fault that ruptured near-simultaneously, the surface trace of which lies only 1-2 km from Hveragerði [58,59]. The geometric mean of the horizontal PGA of the main-shock varies from about 38 to 88% of the acceleration of gravity (g) at distances ranging 0.9-2.3 km from the causative fault. Moreover, remarkable variations can be observed (i.e., 0.03 to 38% g) in recordings of aftershock having local magnitudes of 0.42-4.75 at epicentral distances of 1.6-15 km [58,60].

According to the geological information and borehole data in Hveragerði, the majority of ICEARRAY I stations (except IS609, IS610, and IS613) sit directly upon a lava-rock layer (~5000 years old, B/C in Fig. 2) which in turn lies on top of a softer sedimentary layer, introducing a significant shear wave velocity reversal. At deeper depths, this structure is essentially repeated with the underlying lava-rock layer (~10,000 years old) sitting on softer sediments on top of older bedrock [11,61]. After further reviewing the dataset to ensure the quality of the recordings, we selected motions from 610 aftershocks recorded by ICEARRAY I. Parametric information for these events is available from the National seismic network (SIL) of the Icelandic Meteorological Office [62–64] and hypocentral location estimates have been improved using data from a local and temporary seismograph network deployed six hours after the earthquake [65].

2.2. ICEARRAY II in North Iceland

In October 2012 and April 2013, two of the strongest earthquake sequences in North Iceland in over 30 years took place. The first sequence (October 2012) occurred on the western HFF at its junction to the Eyjafjarðaráll rift, which is a southward extension of the Kolbeinsey ridge (Fig. 1c). The earthquake activity consisted of several migrating bursts of seismicity that together illuminated several fault strands of the western HFF system and within the rift [66]. The results of the Global Centroid-Moment-Tensor (CMT) magnitudes show that the two largest events of the 2012 sequence occurred on 21 October and had moment magnitudes 5.3 and 5.6. The second earthquake sequence (April 2013) took place on the other transform structure in the TFZ, the Grímsey Oblique Rift, just east of Grímsey Island (Fig. 1c). The largest earthquake presented in Global CMT in this sequence occurred on 3 April with a moment magnitude of 5.3, with several smaller earthquakes occurring in the following two days [66]. The average distances to the 2012 and 2013 earthquake sequences from ICEARRAY II were about 71 km and 57 km, respectively. Due to seismic wave attenuation over these large event-station distances, the Icelandic strong-motion array in North Iceland (ICEARRAY II) only recorded the largest 26 earthquakes in the earthquake sequences. The recorded strong-motions are characterized with geometric mean of the horizontal PGA in the range of 0.03 to 3.0% g for earthquakes having local magnitudes in the range of 3.1-5.4 [25].

Although the geological features of Húsavík are characterized as glacial sediments, the topography and near-surface conditions are more complicated than for ICEARRAY I. The town is located in a small basin that is bounded to the north by the southwest-dipping HFF (Fig. 2). A recent compilation of geological information reveals that the near surface geology may be grouped into three main classes of different origins [29]. The northernmost part of the town sits on relatively hard Tillite, while the geology along the shoreline towards the south is characterized by horizontally layered fluvial sediments (Fig. 2). On top of the horizontal sediments lies a delta formation of glacial deposits formed at the end of the last glacial period, some 10,000 years ago. The oldest sediments are glacial deposits which have over time been altered to solid Tillite rock and underlies parts of Húsavík [67]. The sediments in the town are likely underlain by similar inter-glacial basalts as are seen to the southeast, while older Tertiary basalts form the basement north of the HFF.

2.3. The ICEARRAY I and II Datasets

In total, we analyzed 610 events recorded on 10 ICEARRAY I strong-motion stations, which resulted in a total of 4620 PGA values (geometric mean PGA from two horizontal components of motion); note that not all stations recorded motions from each event (for details see [68]). In this study we have excluded station IS610 due to its hillside location on a considerably older bedrock, along with station IS613 due to relatively few recordings and unreliable data. Analogously, for ICEARRAY II we only had 83 PGA values (14 events, red symbols in Fig. 1c) and 66 PGA values (11 events, green symbols in Fig. 1c) from the October 2012 and April 2013 sequences, respectively, recorded on 6 strong-motion stations. We considered using the ICEARRAY II dataset as a whole but due to the distinct and different locations of the October 2012 and April 2013 sequences we split the dataset into two to avoid issues with path effects [25]. Fig. 3 shows the characteristics of the ICEARRAY I and ICEARRAY II databases in terms of local magnitude vs. distance.

3. Bayesian Hierarchical Model

In order to model the spatial distribution of strong-motion over the ICEARRAY I and ICEARRAY II, respectively, we use a three-level BHM consisting of what is referred to as the data level, latent level, and hyperparameter level (see companion paper [52]). At the data level the probability distribution of the observed data is presented as a function of parameters. Most of these parameters are classified as latent parameters but a few are classified as hyperparameters. The prior probability distributions for the latent parameters are given at the latent level. The purpose of these prior distributions is to regularize or constrain the latent parameters in a sensible way to improve their inference. The prior distributions depend on hyperparameters. At the hyperparameter level, the prior distributions of the hyperparameters at the data level and the hyperparameters at the latent level are given. In the formulation convention of Bayesian statistics, the BHM model consists of the following terms:

$$\begin{aligned} Y_{es} = {} & \mu_{es}(M_e, R_{es}, D_e) + \delta B_e + \delta S_s + \delta WS_{es} + \delta R_{es}\,, \\ & e = 1, \dots, N\,, \;\; s = 1, \dots, Q \end{aligned} \tag{1}$$

where the indices e and s represent earthquake (event) e and station (recording site) s, respectively. The terms are assumed to be independent of each other, and for simplicity we assume that the database consists of strong-motions from N earthquakes that were recorded at Q stations. Y_{es} is the strong-motion parameter of interest for earthquake e recorded by station s. In this study it is the logarithm of peak ground acceleration i.e., $Y_{es} = \log_{10}(\mathrm{PGA}_{es})$ which is assumed to follow a normal distribution conditional on μ_{es}. The predictive model is

a function of independent variables (seismic parameters such as magnitude, distance, and depth) for event e at station s and is represented by μ_{es}. The overall effect of event e relative to the predictive model is denoted by δB_e (i.e., the event term or inter-event residuals), while the station term is given by δS_s. A spatially correlated event-station specific term is denoted by δWS_{es}, and δR_{es} accounts for effects that are not modeled by the other terms (i.e., unexplained terms) [52].

Due to the variability associated with the earthquake rupture process, the corresponding ground motions can, on average across all stations, be systematically higher or lower than the median value predicted by the model. For this reason, it is critical to separate the variabilities into inter-event and intra-event parts [4,69]. In the proposed BHM given by Eq. (1), δB_e represents the inter-event variability. The intra-event variability, on the other hand, describes how the ground motion observed at one station can be higher or lower than the median ground motions across all stations for a particular event [33]. In our BHM model, we further subdivide the intra-event variability into inter-station variability (i.e., station-to-station variability), event-station variability (i.e., variability between stations within an event), and other unaccounted variability (e.g., measurement and model error, etc.). One of the main benefits of the proposed BHM is estimating inter-station variabilities to obtain single-station sigma, which avoids double-counting some parts of the variability in a probabilistic seismic hazard analysis (PSHA) [8,42]. Effectively, therefore, the total aleatory variability of the model in Eq. (1) can be quantified with the variance of the sum of the last four terms of Eq. (1), which are assumed to be independent. Let σ_T^2 denote this variance:

$$\sigma_T^2 = \tau^2 + \phi_{S2S}^2 + \phi_{SS}^2 + \phi_R^2 \tag{2}$$

The inter-event variance (τ^2) quantifies the variation between events relative to the average ground motion level predicted by the GMM for each event. The inter-station variance (ϕ_{S2S}^2) quantifies the variability between stations, which then is primarily a manifestation of the localized variations such as the geological profiles beneath the stations. The event-station variance (ϕ_{SS}^2) can be defined as a measure of the spatial variability in the ground-motion amplitudes between stations within an event after taking into account the event and station terms. The purpose of this term is to quantify the remaining variations not already captured by the GMM or the event and station terms (i.e., a proxy for path effects). Finally, the unexplained term variance (ϕ_R^2) quantifies the variability in the measurement errors and other deviations that are not accounted for by other terms of the model.

We define the vector $\boldsymbol{\mu} = (\mu_{11}, \mu_{12}, \dots \mu_{es} \dots, \mu_{NQ})$ using a linear predictive function of independent variables as the GMM. The linear predictors are: the local magnitude of event e, M_e; source-to-station distance (hypocentral distance) from event e to station s, R_{es}; and the hypocentral depth of event e, D_e. The GMM has the simple and common functional form:

$$\mu_{es} = \beta_1 + \beta_2 M_e + \beta_3 \log_{10} R_{es} + \beta_4 D_e \tag{3}$$

The vector $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)$ contains the regression coefficients of the GMM, i.e., the region-specific vector containing the parameters that capture the characteristics of the seismic region and the geological structure. We assume a weakly informative prior distribution for $\boldsymbol{\beta}$ that is normal with mean $\boldsymbol{\mu}_\beta$ and a diagonal covariance matrix $\boldsymbol{\Sigma}_\beta$:

$$\boldsymbol{\beta} \sim \mathrm{N}(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta) \tag{4}$$

The vector $\boldsymbol{\delta B} = (\delta B_1, \dots \delta B_e, \dots, \delta B_N)$ quantifies the event-dependent deviation from the overall mean model for all events (read [52] for more details). The event terms are assumed to be independent of each other, normally distributed with zero mean and standard deviation of τ:

$$\boldsymbol{\delta B} \sim \mathrm{N}(\mathbf{0}, \tau^2 \boldsymbol{I}_N) \tag{5}$$

The station terms, δS_s, from stations $s = 1, \dots, Q$, which are gathered in the vector δS, are modeled a priori with a mean zero Gaussian distribution with covariance matrix $\boldsymbol{\Sigma}_{S2S}$. This covariance matrix is based on an exponential covariance function, a special case of the Matérn family with a smoothness parameter of $\nu_{S2S} = 0.5$, which is very flexible and suitable for a variety of applications [54,70]:

$$\boldsymbol{\delta S} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_{S2S}) \tag{6}$$

$$\{\boldsymbol{\Sigma}_{S2S}\}_{ij} = \left\{ \phi_{S2S}^2 \exp\left(-\frac{d_{ij}}{\Delta_{S2S}} \right) \right\}_{ij} \tag{7}$$

where in Eq. (7), d_{ij} is the inter-station distance between stations i and j $(i, j = 1, \dots, Q)$ and ϕ_{S2S}^2 is the marginal variance which describes the inter-station variability. The effect Δ_{S2S} is defined as the range parameter, and $1/\Delta_{S2S}$ is the decay rate [71], i.e., how fast the correlation between observations decreases with distance.

We define the event-station terms, δWS_{es}, as spatially correlated variables from a zero mean Gaussian field governed by a covariance function from the Matérn class. The covariance matrix for the δWS_{es} terms is a block diagonal matrix:

$$\boldsymbol{\Sigma}_{SS_{[(NQ)\times(NQ)]}} = \boldsymbol{I}_N \otimes \{\boldsymbol{\Sigma}'_{SS}\}_{ij} \tag{8}$$

$$\{\boldsymbol{\Sigma}'_{SS}\}_{ij} = \left\{\phi_{SS}^2\left(1+\frac{d_{ij}}{\Delta_{SS}}\right)\exp\left(-\frac{d_{ij}}{\Delta_{SS}}\right)\right\}_{ij} \tag{9}$$

Where, $i,j = 1,\dots,Q$, and the marginal variance ϕ_{SS}^2 describes the event-station variability, ν_{SS} is the smoothness parameter, taken as 1.5, and $1/\Delta_{SS}$ is the decay rate. The block matrix in Eq. (8) is defined with the Kronecker product, $\otimes$, in which the rows and columns are removed according to missing data-points (i.e., stations with no records for a specific event).

Finally, the unexplained terms, δR_{es} are independent mean zero Gaussian random variables with a standard deviation of ϕ_R, and when gathered in the vector $\boldsymbol{\delta R}$ then:

$$\boldsymbol{\delta R} \sim \mathrm{N}(\mathbf{0}, \phi_R^2 \boldsymbol{I}_N) \tag{10}$$

On the basis of the Bayesian hierarchical framework, the functions defined above, and our understanding of the physical process at the data level [54], our BHM can be set up as

$$\pi(\boldsymbol{y}|\boldsymbol{\beta},\boldsymbol{\delta B},\boldsymbol{\delta S}) = \mathrm{N}(\boldsymbol{y}|\boldsymbol{X\beta} + \boldsymbol{Z}_1\boldsymbol{\delta B} + \boldsymbol{Z}_2\boldsymbol{\delta S}, \phi_{\boldsymbol{R}}^2\boldsymbol{I}_{NQ} + \boldsymbol{\Sigma}_{SS}) \tag{11}$$

where $\boldsymbol{y}$ represents a vector containing all the observations, Y_{es}, corresponding to events, e, and stations, s. The matrices $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ are index matrices. If we consider that there are no missing data at the stations, then $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ are defined as

$$\boldsymbol{Z}_{1_{[(NQ)\times N]}} = \boldsymbol{I}_N \otimes \mathbf{1}_Q \tag{12}$$

$$\boldsymbol{Z}_{2_{[(NQ)\times Q]}} = \mathbf{1}_N \otimes \boldsymbol{I}_Q \tag{13}$$

where $\mathbf{1}_n$ is an *n*-element vector of ones and $\boldsymbol{I_n}$ is the $n \times n$ identity matrix. The matrix $\boldsymbol{X}$ is a $(NQ) \times p$ dimensional matrix ordered for all the stations and events, where from Eq. (3) $p = 4$ in this study. The matrix is such that the first column is the constant vector, $\mathbf{1}_N$, the next $(p-1)$ vectors contain the covariates corresponding to event e and station s. We note that in this study we used a Gibbs sampler to infer jointly all the hyperparameters at the third level of the proposed BHM model and latent parameters at the second level. The implementation procedure and convergence diagnostics are explained in detail in [52].

4. Results

In this section, we evaluate the reliability of the proposed BHM model using statistical inference of the model parameters. In the first sub-section, the model variabilities are assessed and physically interpreted for both arrays. Subsequently, the residual analysis is

presented to demonstrate the validity of the BHM model. A detailed inference of the station terms is presented next for both arrays where we highlight the importance of the site effects and localized geological impacts. Then, we present and compare the inter-event and intra-event variabilities between both arrays. Finally, the contribution of the source, path, and site effects are quantitatively determined and interpreted for both arrays.

4.1. Inference of Variabilities

The Markov Chain Monte Carlo (MCMC, see [52]) simulations provided estimates of the posterior distribution of the latent parameters and the hyperparameters of our BHM model for the ICEARRAY I dataset as a whole and for the ICEARRAY II dataset analyzed separately for the 2012 and 2013 earthquake sequences. The posterior distributions of the hyperparameters, $\boldsymbol{\theta} = (\tau, \phi_{S2S}, \phi_{SS}, \phi_R, \Delta_{SS})$, are shown in Fig. 4 and their statistics are listed in Table 1. It should be noted that we estimated the range parameter of the station term by trial and error and fixed it at $\Delta_{S2S} = 60$ meters for ICEARRAY I and $\Delta_{S2S} = 600$ meters for ICEARRAY II, corresponding to their respective minimum inter-station distances.

A striking feature of the posterior distributions of the hyperparameters of the ICEARRAY I dataset is how well they are determined compared to the ICEARRAY II datasets. For ICEARRAY I they represent nearly symmetric and unimodal posterior distribution with low uncertainty, while for ICEARRAY II they are mostly skewed, unimodal distributions with large uncertainty. For example, the posterior distribution of the range parameter Δ_{SS} in Fig. 4 is not only drastically wider for ICEARRAY II but it is non-symmetric while the posterior distribution for the same parameter in the case of ICEARRAY I is symmetric. The decay parameter, $1/\Delta_{SS}$, is significantly larger across ICEARRAY II which indicates that ICEARRAY I stations are correlated to each other to a much higher degree at shorter distances than those of ICEARRAY II. We expect that the geological differences of the two arrays are also mapped into the differences of the posterior estimate of Δ_{SS}.

The rest of the hyperparameters are the standard deviations of the event terms (τ), station terms (ϕ_{S2S}), event-station terms (ϕ_{SS}), and the remaining model error and unexplained variability (ϕ_R). As can be seen in Table 1, the posterior distribution of the standard deviation of the station term ϕ_{S2S} for ICEARRAY I has a posterior mean 0.092, and a 95% posterior interval [0.057, 0.151], while for ICEARRAY II the posterior means are 0.197 and 0.218 and the 95% posterior intervals are [0.099, 0.383] and [0.112, 0.427] for the

2012 and 2013 swarms, respectively. This effectively quantifies the difference between the two arrays in the overall variation of the posteriors for the individual δS's. The posterior mean and the 95% posterior interval of the standard deviation of the event term τ are 0.198 and [0.185, 0.211] for ICEARRAY I, while for ICEARRAY II they are 0.222 and [0.129, 0.359], and 0.148 and [0.043, 0.306] for the 2012 and 2013 sequences, respectively. The posterior mean and the 95% posterior interval of the standard deviation of the event-station term ϕ_{SS} are 0.116 and [0.110, 0.123] for ICEARRAY I, while for ICEARRAY II they are 0.052 and [0.009, 0.135], and 0.097 and [0.032, 0.179] for the 2012 and 2013 sequences, respectively. The difference between the ϕ_{SS} values for the 2012 and 2013 sequences effectively justifies our decision to separate the ICEARRAY II dataset by earthquake sequence. Then, the posterior mean of the unexplained term variability, ϕ_R, is approximately the same ($\sim 0.05 - 0.06$) for both arrays, but its standard deviation is about an order of magnitude ($\sim 4 - 11$ times) larger for ICEARRAY II than that for ICEARRAY I. Clearly the robust dataset recorded by ICEARRAY I allows a more reliable estimation of the parameters of the BHM, but the greater variability of the local geology and topography across ICEARRAY II also likely attributed to the reduced reliability of the estimates, relative to ICEARRAY I. Finally, the cross-correlation coefficients between the hyperparameters are shown in Fig. 5. Generally, the inter-station variation (ϕ_{S2S}) appears to be not correlated to any significant extent with any of the other hyperparameters (i.e., the standard deviation of: the inter-event term τ; the event-station term ϕ_{SS}; the unexplained term ϕ_R; and the range parameter Δ_{SS}). We are therefore confident in the results for the station terms effects to be unaffected by event characteristics.

4.2. Residual distribution

The inference on the latent parameters in $\boldsymbol{\beta}$, i.e., the regression coefficients (β_1: the constant coefficient; β_2: magnitude scaling coefficient; β_3: geometric attenuation coefficient; and β_4: depth scaling coefficient) of our GMM (represented by μ_{es}) resulted in well-simulated posterior samples that yield promising results. The regression coefficients for both arrays have nearly symmetric and unimodal distributions as can be seen in Fig. 6 (for the tabulated statistical results see Table 2). The analysis met the convergence criteria and resulted in reasonably low posterior standard deviations of the obtained β's for the ICEARRAY I dataset. For the ICEARRAY II datasets, in contrast, very large posterior standard deviations were obtained for the $\boldsymbol{\beta}$'s for the 2012 sequence, compared to the 2013

sequence for which the parameters are relatively well constrained. Moreover, the relative values of β_1 and β_2 are different between the sequences. While the cause cannot be determined at this point due to the overall small size of the dataset and/or the form of the GMM used, we suspect that the effects of fault-trapped waves along the HFF are largely responsible for the overall higher level of uncertainty for the 2012 sequence as recorded in Húsavík, relative to the 2013 sequence.

The event term (i.e., inter-event residuals) $\boldsymbol{\delta B}$ shows nearly symmetric and unimodal posterior distribution for both arrays (see Fig. 7 for ICEARRAY I, the posterior results for ICEARRAY II are not shown). Moreover, we did not observe a temporal correlation (clustering in time) between the event terms nor a strong posterior correlation between the event terms themselves (the posterior correlation is $\sim 20\%$). While there is clearly a variability in the individual estimates of the event terms δB_e, they do not exhibit any clear trends with magnitude or distance. As a result, the total model residuals, $Y_{es} - \mu_{es}$ (see Eq. (1)), are in general well behaved for all the datasets considered in this study. The intra-event terms are determined by additionally removing the individual earthquake specific effects from total residuals, or $Y_{es} - (\mu_{es} + \delta B_e)$ [40]. The intra-event residuals for the ICEARRAY I dataset are shown in Fig. 8. Their distribution vs. local magnitude, depth, hypocentral distance and back-azimuth show no sign of trend with respect to these four predictors. We note that the back-azimuth is not a dependent variable of the GMM, but is shown in Fig. 8 for the purpose of potentially identifying event differences based on what causative faults or parts of faults they occur (see Fig. 1). Both the inter-event and intra-event residuals (the latter is shown in Fig. 8) show unbiased distribution of residuals, which indicates that the relative ground motions over the array are not dependent on the causative fault. This effectively confirms that this predictor indeed does not need to be explicitly included in the GMM. Therefore, our BHM is unbiased with respect to the dataset with a total standard deviation of $\sigma_T = 0.25$ (base-10 logarithmic units) for ICEARRAY I and $\sigma_T = 0.30$ and 0.29 for the 2012 and 2013 earthquake sequences on ICEARRAY II, respectively.

4.3. Station term Inference

The posterior inference of the latent parameter that represents the station terms, $\boldsymbol{\delta S}$, for ICEARRAY I and ICEARRAY II is presented statistically in tabular format in Table 3, and the corresponding posterior distributions are shown in Fig. 9. The station terms serve as indicators to what extent the PGA can be expected to be either higher or lower than the mean

over the array. Their posterior probability distributions have shapes that are close to that of a Gaussian probability distribution. The station terms of ICEARRAY I are well constrained as is evident from their narrow posterior intervals which are associated with small posterior standard deviations that are similar, all in the range $0.0322 - 0.0325$. On the contrary, the posterior distributions of ICEARRAY II have about three times larger posterior standard deviations, all in the range $0.1126 - 0.1298$. Fig. 9 also highlights the smaller absolute range that the station terms of ICEARRAY I cover, relative to ICEARRAY II. For example, the largest relative station term for ICEARRAY I amounts to 0.15 ± 0.03 base-10 logarithmic units while for ICEARRAY II it ranges from 0.24 ± 0.11 to 0.28 ± 0.13. This is consistent with the expectation that for a relatively uniform site condition (in this case, rock) the corresponding station terms, δS_s, for ICEARRAY I, are relatively low and with relatively narrow marginal posterior distributions, the widths likely being a result of a large dataset. On the contrary, a larger range covered by station terms is expected where geologic profiles and topography vary considerably (ICEARRAY II δS_s values), along with relatively wider posterior distributions, determined from a relatively small dataset. This is consistent with the large differences in the standard deviations of the station terms, ϕ_{S2S}, between the two arrays.

As can be expected for seismic motions recorded over a relatively small area compared to the source-station distance, there is a very high overall positive correlation between station terms, as shown in Fig. 10a. For the ICEARRAY I the highest correlation (> 0.94) is found between IS602, IS603, IS604, IS605, and IS607, which are the subgroup with the smallest interstation distances, all located in the northwestern corner of the array. The very highest correlation is then found for stations IS603, IS604, and IS605 that form the densest part of the array, located 54-114 meters apart. Station IS611 (δS_9) sits in the middle of the array and has a similar correlation to all stations, while the cross-correlations between remaining stations are lower. For ICEARRAY II, the cross-correlations seen in Fig. 10b and c are in general much lower than for ICEARRAY I. This is consistent with the above-mentioned result of the decay parameter i.e., inverse of the hyperparameter Δ_{SS}. For ICEARRAY II, the two closest stations (IS703, δS_4 and IS705, δS_6) on the southern sediment terraces have relatively high-cross correlations in general, as can be expected both from topographical and sedimentary similarities, in addition to their proximity. Similarly, for the 2012 sequence, stations IS202 (δS_1) and IS704 (δS_5) also had relatively high cross-correlations, again as expected, being the two closest stations on the Tillite rock in the northern part. By visual comparison of Fig. 10b and c, higher correlations appear to exist overall between the station

terms for the 2012 sequence. In particular, the stations on the southern terraces (IS703, δS_4; IS705, δS_6) have relatively low correlations with the stations on the Tillite rock in the northern part (IS202, δS_1; IS704, δS_5) during the 2012 sequence. This indicates that the earthquake sources themselves or the direction of wave propagation may be an influencing factor.

Despite the high correlation between station terms, the posterior density of the inter-station variance ϕ_{S2S}^2 is indeed different from zero, confirming that the amplitude of the spatial field $\delta S(\cdot)$ is different from zero. This strongly indicates that site effects influence the PGA parameter. On closer scrutiny, the station terms for stations IS601, IS604, and IS611 are significantly affecting (either positively i.e., amplifying, or negatively i.e., de-amplifying, relative to the mean) the earthquake strong-motion intensities (see Table 3 and Fig. 9), based on the 95% marginal posterior intervals for the individual δS's across ICEARRAY I. Here a station term is said to significantly affect the intensity if the 95% posterior interval does not contain zero. In particular, the impact of station term value for station IS601 (95% posterior interval is [0.0754; 0.2035]) is estimated to intensify the earthquake ground-motions, whereas δS values for stations IS604 (95% posterior interval is [-0.1471; -0.0190]), and IS611 (95% posterior interval is [-0.1483; -0.0207]) are expected to diminish the seismic ground motions. On the other hand, the rest of the station terms are not significantly affecting the intensity. For instance, a 95% posterior interval for station IS609 is [-0.0502; 0.0786] with an amplifying mean value of 0.0149. The same statistical analysis on ICEARRAY II's station terms with regards to their contribution to overall amplification or de-amplification, relative to the mean PGA across the array, did not significantly affect the intensity (i.e., negative 2.5% percentile and positive 97.5% percentile), primarily due to the large posterior standard deviations of the station term. As more data is being recorded however, it is expected that more significant site effects will be revealed for ICEARRAY II.

4.4. Inter-event vs. intra-event variability

The relative comparison of the standard deviations of the inter-event terms (τ), the inter-station terms (ϕ_{S2S}), the event-station terms (ϕ_{SS}), and the unexplained terms (ϕ_R) are shown graphically in Fig. 11. The results for ICEARRAY I show that the inter-event standard deviation ($\tau = 0.20$) is larger than the intra-event standard deviation ($\phi = \sqrt{\phi_{S2S}^2 + \phi_{SS}^2 + \phi_R^2} \approx 0.16$, see Table 1) which consists of the inter-station, the event-station, and the unexplained variabilities. For ICEARRAY II, on the other hand, the inter-event

standard deviation obtained for the 2012 sequence ($\tau = 0.22$) is noticeably higher than that for the 2013 sequence ($\tau = 0.15$), even though the total standard deviations are roughly the same ($\sigma_T = 0.30$ and 0.29 for 2012 and 2013 sequences, respectively). The intra-event standard deviation for the 2012 sequence ($\phi = 0.21$) is lower than that for the 2013 sequence ($\phi = 0.25$) while these are larger than the intra-event standard deviation of ICEARRAY I. However, it should be noted that while the posterior mean values of these three intra-event standard deviations are different, their statistical difference may not be significant, strictly speaking.

While the inter-event standard deviation mainly indicates to what extent the mean value of a given ground motion parameter corresponding to a particular earthquake is expected to vary relative to the overall mean of all earthquakes in the dataset, the station and path effects are mostly mapped into the inter-station and event-station standard deviations, respectively. In empirical GMMs, inter-event variability is generally found to be smaller than intra-event variability [e.g., 33,42]. Such GMMs are calibrated to strong-motion datasets that are characterized by regional distances (e.g., from around 150 m up to several hundred kilometers, depending on the tectonic region) and by large differences in station and path effects. On the contrary, the ICEARRAY I dataset consists of a large number of aftershock strong-motions recorded at multiple stations on uniform site conditions in a relatively small area (~ 1.23 km^2) compared to the local (hypocentral) distances of 2-20 km (see Fig. 3). The aftershocks for the most part occurred on two large and separate N-S faults (see Fig. 1) that ruptured during the main shock, and a smaller E-W segment, and the dataset is therefore associated with earthquakes having a relatively large variety of epicentral, depth and azimuthal values, and magnitudes [11,26,60,65]. In comparison, the dataset is much smaller for the ICEARRAY II (see Fig. 3). Additionally, the 2012 and 2013 earthquake sequences are characterized by a relatively small source region compared to the hypocentral distances of 50-80 km, and a narrow azimuthal range (see Fig. 1). This is particularly the case for the 2012 sequence where the hypocenters of the events are all located in the same small area on the HFF, while the 2013 earthquake hypocenters were scattered to a somewhat larger area on the Grímsey Oblique Rift and thus have a slightly wider azimuthal and distance coverage. It is of note though, that the 2012 sequence is associated with a larger inter-event standard deviation than the 2013 sequence, while the inter-station standard deviation is roughly the same for both sequences.

The standard deviation of the observed PGA values (base 10 logarithm of the geometric mean of two horizontal components) appears to be constant and intensity independent, as indicated in Fig. 12, where median observed PGA values over the ICEARRAY I are plotted for each event recorded by more than 5 stations and sorted by descending median amplitudes. The red points indicate the $\pm 1\sigma$ that is estimated with the standard deviation of the observations, and the estimates are such that $\sigma \simeq 0.17$. This is the BHM equivalence to the intra-event standard deviation (~ 0.16), a further confirmation of the model capturing the observed variability appropriately. It was to be expected that the intra-event variability of the dataset was smaller than the inter-event variability because the ground motions of each aftershock were recorded on relatively uniform site conditions, in addition to the motions at each station having travelled effectively the same propagation path. Given the uniformity of geologic profiles underlying the array stations used in this study, the corresponding intra-event variability may suggest a lower bound on the standard deviations attainable for GMMs [60]. In contrast, the ICEARRAY II results indicate that intra-event variability is higher than the inter-event variability (see Fig. 11).

4.5. Source, path and site contributions to the variability

The total standard deviation calculated using Eq. (2) is $\sigma_T \simeq 0.25$ for ICEARRAY I and $\sigma_T \simeq 0.30$ and 0.29 for the 2012 and 2013 sequences for ICEARRAY II, respectively. For comparison, the standard deviation reported for empirical and theoretical GMMs that have historically been applied in hazard analyses in Iceland, or have been calibrated either exclusively, or to a great extent, to strong-motion data from South Iceland is in the range $0.2 - 0.3$ [37,60,72,73]. Breaking up the total variability into subsets of standard deviations gives us the chance to quantitatively find out to what extent the respective earthquake source and site effects contribute to the overall variability of ground motion across the region. In particular, approximating the inter-station standard deviation makes it possible to estimate single-station sigma, which results in more accurate PSHA calculations by avoiding double counting of some parts of the variability [8]. As a result, in recent years many researchers have been trying to highlight the significant effect of single-station sigma by relaxing the ergodic assumption [8,56,74–76]. To obtain the single-station standard deviation we either need a station-specific dataset [8,74,75,77–79] or propose a non-ergodic model [42,56]. It should also be emphasized that allowing the standard deviation to vary between stations should not be the general assumptions unless there is sufficient data to estimate either empirically, theoretically, or numerically the variabilities. The proposed BHM modeling in

this study, can be used as a practical tool to estimate partially (single station) or fully (single path) non-ergodic GMMs and consequently standard deviations [80]. The inter-station variability represents the impacts of geological aspects and station conditions aside from the source and path characteristics. The inter-station standard deviation of PGA recorded on ICEARRAY I ($\phi_{S2S} \simeq 0.09 \pm 0.02$) is quite low and internally consistent, especially compared to the corresponding result for ICEARRAY II ($\phi_{S2S} \simeq 0.20 \pm 0.07$ and 0.22 ± 0.08 for the 2012 and 2013 swarms, respectively). The cause is undoubtedly the geological and topographical differences between the two arrays. Namely, the ICEARRAY I stations are located on flat lava-rock, and therefore, to at least the first approximation, are characterized as homogenous "rock" site condition [11]. On the contrary, available geological information of ICEARRAY II shows that different geological units of varying stiffness and thickness characterize the local station conditions in the town of Húsavík, in addition to significant topographical differences [29]. These variations of geology and relief may be the main reasons that led to higher variation in inter-station variabilities for ICEARRAY II relative to ICEARRAY I, in addition to the stark differences in the size of the datasets.

The contribution of source, path, and site variability can be numerically determined ($\sigma_i^2 / \sum_{i=1}^{4} \sigma_i^2$). In this regard, the low ratio of unexplained term variability (ϕ_R^2) over total variability ($3 - 5\%$) is almost constant for both arrays. We find this acceptable and an indicator that the BHM is well able to capture the salient characteristics of the PGA variation across the arrays. The ratio of inter-event variability (τ^2) to the total variability for ICEARRAY I is $\sim 61\%$ and for ICEARRAY II is $\sim 53\%$ for the 2012 sequence and $\sim 26\%$ for the 2013 sequence. The ratios of inter-station variability (ϕ_{S2S}^2) to total variability for ICEARRAY I is considerably lower, $\sim 13\%$, than that for ICEARRAY II, ($41 - 57\%$), using the 2012 and 2013 datasets. Finally, the ratio of event-station term variability (ϕ_{SS}^2) is approximately 21% for ICEARRAY I. The event-station term variability is largely different for the two datasets recorded by ICEARRAY II (3% and 11% for the 2012 and 2013 dataset, respectively).

5. Discussion

5.1. Spatial distribution of PGA

The proposed BHM model can be used to predict the spatial distribution of PGA at any location across the ICEARRAY I, within the accuracy-limits dictated by the results for each

array even though, strictly speaking, the results presented above pertain only to the stations where the data were recorded. The observed mean variation of normalized log PGA on the ICEARRAY I is shown in Fig. 13a as a spatial distribution where values between stations are obtained by interpolation. For comparison we plotted in a similar way the spatial distribution of the posterior mean of the station terms (not shown since it was indistinguishable from the observations). The data that Fig. 13a is based on are shown in Fig. 13b, wherein the observed motions ($\log_{10}$(PGA)), the mean inter-station residual, and the station terms (with error bars) are shown for each station. In all cases the estimates are captured by the $\pm$ one standard deviation of the corresponding station terms. Of note is that the apparent de-amplification of PGA relative to the mean takes place primarily at stations IS604 and IS611, while positive amplification takes place primarily at IS601. It is also interesting to note that IS609 sits on a small bedrock hill and has a slight positive relative station term, while the nearest stations IS611 and IS612 both sit on top of the lava structure and are associated with slightly negative terms. Further complicating the pattern is the fact that the centre of town (open area just northeast of IS611) is an active geothermal area (Fig. 2) [61].

We expect therefore that the relatively hotter and softer rock underlying this area will cause larger attenuation of seismic waves compared to other areas of the town. This may be the reason for the relatively lower PGA values that were consistently observed in the centre of town. Additionally, the lava structure has been shown to amplify the horizontal motions relative to the vertical motions primarily at the predominant frequencies of 3-4 and 8-9 Hz, respectively [11]. These frequencies are much lower than those that generally dominate PGA levels. We therefore expect that a future study analyzing pseudo-spectral accelerations as a function of frequency would provide station terms that are frequency dependent, consistent with the observed frequency dependent amplification of the lava structure [26]. It is of note, however, that in our results station IS605 shows a positive relative amplification of PGA compared to the relative de-amplification at IS603 and IS604, despite all these stations being very close to one another. Station IS605 had the clearest bimodal horizontal-to-vertical spectral ratio of all the ICEARRAY I stations, and the relative inconsistency of amplification between IS603 and IS605 stations may also be related to the general observation that the lava structure may vary greatly across short distances [11]. We note that interpolating the normalized amplitudes (or the station terms), as is done in Fig. 13a, extends the observed (or modeled) variation at discrete points into a continuous spatial distribution. While this

approach may be useful as a first approximation, it may not necessarily hold in places where the lateral variation in the geological structure is considerable. However, this can be expected to apply for parameters affected by high frequencies, such as the PGA.

5.2. Azimuth-dependent posterior of station term

The ICEARRAY II station term posteriors shown in Fig. 9 (see also Table 3) are quite comparable between the datasets for the two sequences, with the standard deviations of the posteriors being slightly larger for the 2013 sequence. A significant exception in terms of mean values, however, is seen for station IS703, which is located on top of the sedimentary terrace hills in the southern part of Húsavík. Along with IS705 (inter-station distance of 210 m), IS703 is affected by this type of sediments and topography. For the 2012 sequence the posteriors of the station terms for these two stations are effectively identical. However, for the 2013 sequence the mean value of the station term for station IS703 is nearly half of that for the 2012 sequence. The sudden change in the station term was quite unexpected because no changes had taken place at the site. The only alternative causative factor therefore lies with the 2013 earthquake sequence itself, compared to the 2012 sequence, given that the station term is sensitive to the earthquake parameters. Effectively, the only difference from a parametric point of view is the larger azimuthal angle of the earthquakes of the 2013 sequence. It is of significance that station IS703 sits on the steep edge of the terrace margin while station IS705 sits in the middle of the terrace. The seismic waves reaching the array during the 2013 sequence arrive at a different angle compared with the 2012 sequence. Moreover, the seismic waves from the 2012 sequence may be affected by trapped waves propagating along the HFF to the array. At present, given the limited dataset no emphasis has been placed on resolving this issue. However, once more data are recorded, especially from other azimuthal angles and source depths, a repeated analysis should shed light on the actual posterior distributions of the station terms for IS703 vs. IS705. Until then, we assume the posterior distributions of the station terms are similar and as modelled for the 2012 sequence, primarily due to the same geologic profiles and higher topography.

5.3. Aftershock ground motion scaling

The scaling of earthquake ground motion parameters with magnitude and distance is of great importance for seismic hazard assessment. It has been shown that the decay rate with distance, also known as geometrical spreading, is dependent on earthquake size, with smaller earthquakes having a more rapid spatial decay than large magnitude events [73,81,82]. In

addition, magnitude-dependent decay has also been observed in ground motion simulations from extended source models [83]. The ICEARRAY I dataset of aftershocks range over four-units of magnitude and can be used to provide insight into the validity of the magnitude-dependent decay rate. The mean values of the geometric mean PGA of two horizontal components at each recording station are plotted in Fig. 14 as a function of distance, and for completeness, magnitude as well. We separated the mean PGA values into five magnitude-bins and three distance-bins, as shown in the figure, thus capturing the scaling of each bin separately by simple functions, $y \sim r^{\bar{\beta}_3}$ and $\log y \sim \bar{\beta}_2 M_L$, where the constants ($\bar{\beta}$s) are expected to be the counterparts to those in Eq. (3) and summarized in Table 2. The scaling with distance in Fig. 14 appears to be slightly faster than to $1/r^2$, with a standard deviation of $\sim$0.3 for local magnitudes larger than 2. The 95% confidence bounds of $\bar{\beta}_3$ for different magnitude bins are also presented in Fig. 14. Moreover, as can be seen in Fig. 14, the dependency of PGA on local magnitude obtained by the simple standard regression in Fig. 14 appears to be relatively stable at $\sim(0.5 - 0.7)M_L$, with a standard deviation of roughly $\sim$0.4 which is in good agreement with the mean posterior of β_2 samples is equal to 0.7 ± 0.015.

5.4. Towards non-ergodic ground motion models

In a pioneering study, the variation in PGA recorded by SMART1, the first dense digital array in Taiwan, was empirically investigated [43] which highlighted the importance of site effects. Afterwards, many researchers focused their efforts to numerically partition the source, path, and site contribution to ground motion amplitudes with the particular aim of removing parts of the ergodic assumption from GMMs. Kotha et al. (2016) [84] and Ktenidou et al. (2017) [80] investigated ground motion uncertainty and variability via regression analysis and concluded that defining proxies for the site response and improved site information accounts for the reduction in systematic site-to-site variability, thus leading to a reduction of the aleatory variability. In this way, Kuehn and Scherbaum (2015, 2016) [42,56] proposed a multi-level Bayesian GMM which decomposes the observations into source, site, and path effects. However, for their regional dataset they proposed an empirical functional form to determine the station terms using a proxy for site response; we defined spatially correlated station terms because of our small-aperture array data. Moreover, from the outset we did not use a proxy for site response due to the generally unknown site effects on the two "rock" types under ICEARRAY I [11,26] and the multiple different conditions under

ICEARRAY II [26,29,85]. Thus, the site terms of our BHM have now effectively revealed the distinct contributions of the localized site condition under each station, which forms a basis for more detailed (preferably physical) modeling of the site response. Finally, in contrast to Kuehn and Scherbaum's studies, our BHM also includes an additional term to consider the unexplained terms in the BHM model.

6. Conclusions

The strong-motion data recorded on the dense urban strong-motion array in Hveragerði (ICEARRAY I) during the 29 May 2008, $M_{\mathrm{w}}6.3$ Ölfus earthquake and its aftershocks in South Iceland exhibited considerable variations in ground-motion amplitudes recorded on multiple stations over a small area. This variation is of particular interest because all of ICEARRAY I recording stations are located on "rock." Moreover, while the variation was observed for each event a systematic variation was not clear between events, even if they were associated with similar magnitudes and epicentral distances. Similarly, during the earthquake sequences in North Iceland in 2012-2013 significant variation of earthquake strong-motion amplitudes were observed on the dense urban strong-motion array in Húsavík (ICEARRAY II). In contrast to ICEARRAY I however, the geologic profile and topography varies significantly across ICEARRAY II.

For the purpose of systematically analyzing the strong-motions to explain their observed variations we modeled them using a multi-level Bayesian Hierarchical Model (for details see [52]) that partitions the motions into event, station, and event-station terms in order to quantify the relative contributions of source, site, and path effects. In this study, we analyze the peak-ground acceleration of the recorded motions using the BHM to facilitate our understanding of the spatial distribution of PGA and its variability. For that purpose, the total variability is partitioned into an inter-event and intra-event variabilities. Furthermore, the intra-event variability partitions into an inter-station, event-station, and unexplained term variabilities.

The resulting BHM is shown to capture the observed PGA variations with unbiased residuals. The total standard deviation is 0.25 base-10 logarithmic units for ICEARRAY I and ~ 0.30 for ICEARRAY II. The results clearly indicate significant site effects, as individual station terms are different from one another for both ICEARRAY I and II, while being of different levels in amplitude and associated with different levels of uncertainty. This

is manifested in the standard deviation of the overall station term being significantly different from zero according to the posterior density for both arrays. The standard deviation of the station term for ICEARRAY I is 0.09 ± 0.02 base-10 logarithmic units, while for ICEARRAY II it is 0.20 ± 0.07 and 0.22 ± 0.08 for the 2012 and 2013 swarms, respectively. For ICEARRAY I the largest relative station term amounts to 0.14 ± 0.03 logarithmic units, while for ICEARRAY II it ranges from 0.28 ± 0.13. In other words, the site effects on ICEARRAY I are less both in variation and amplitudes, primarily due to the relatively uniform site conditions across the array. Moreover, the ICEARRAY I results are much better constrained due to the large dataset, compared to ICEARRAY II. For the latter array, the variation in site effects is greater and has larger relative amplitudes, which was expected due to the varying geologic profiles and topography across the array. The BHM shows quantitatively to what extent the respective earthquake source, path, and site effects contribute to the overall variability of ground motions across the arrays. For ICEARRAY I, the results show that the station term is responsible for about 13% of the total variability, while the event term contributes to the vast majority of the total variability, or about 61%. The event-station term contributes about 21% to the total variability, with the remaining 5% being unexplained variability. For the ICEARRAY II, the results indicate that the site effects are as great or exceed the event effects (41% and 57% for the 2012 and 2013 datasets, respectively).

The comparison between the results for the two arrays explicitly quantifies the influence of the complexity in site effects on PGA. The computed station terms values for the stations are consistent with what was qualitatively expected for the site conditions, namely the computed station term values are smaller for the relatively uniform site conditions of ICEARRAY I, but larger for ICEARRAY II which is characterized by a variety of geologic profiles and topography. Moreover, detailed inspection of the spatial distribution of station terms shows that even over relatively small distances (tens of meters) on the same lava-structure of the ICEARRAY I, the site effects can vary significantly, which not only is of great practical significance for the microzonation of seismic hazard but also has important implications for locally varying seismic risk.

The results of this study have been produced using a novel application of the BHM which enabled the quantification of site effects, their relative amplitudes, and most importantly, their statistical uncertainty. While the PGA was the ground motion parameter studied herein, using the same methodology other ground motion parameters can be analyzed

such as peak ground velocity, pseudo-spectral acceleration/velocity/displacement, ground motion duration, etc. The analysis has improved our understanding of the key factors that affect the variation of seismic ground motions across a relatively small area. We believe our results find application in earthquake hazard assessment on a local scale, with practical implications for seismic risk and engineering decision making such as urban planning.

Acknowledgements

This work was supported by the Icelandic Centre for Research (Rannís Grant of Excellence no. 141261-051/52/53), the Doctoral grants from Eimskip Fund of the University of Iceland, and the University of Iceland Research Fund. The instruments were funded through a Rannís Equipment Grant and their deployment by the Icelandic Catastrophe Insurance (Grant no. S112-2013). The authors would like to express their gratitude to the inhabitants of Hveragerði and Húsavík, and their municipalities, for housing the recording equipment and for their dedication and support to the ICEARRAY project.

References

[1] Loh CH, Ang AH-S, Wen YK. Spatial Correlation Study of Strong Motion Array Data with Application to Lifeline Earthquake Engineering. University of Illinois Engineering Experiment Station, College of Engineering, University of Illinois at Urbana-Champaign; 1983.
[2] Seed HB, Romo MP, Sun JI, Jaime A, Lysmer J. The Mexico Earthquake of September 19, 1985—Relationships Between Soil Conditions and Earthquake Ground Motions. Earthq Spectra 1988;4:687–729.
[3] Aki K. Local site effects on weak and strong ground motion. Tectonophysics 1993;218:93–111.
[4] Al Atik L, Abrahamson N, Bommer JJ, Scherbaum F, Cotton F, Kuehn N. The variability of ground-motion prediction models and its components. Seismol Res Lett 2010;81:794–801.
[5] Baltay AS, Hanks TC, Abrahamson NA. Uncertainty, Variability, and Earthquake Physics in Ground- Motion Prediction Equations Uncertainty, Variability, and Earthquake Physics in Ground- Motion Prediction Equations. Bull Seismol Soc Am 2017;107:1754–72.
[6] Bessason B, Kaynia AM. Site amplification in lava rock on soft sediments. Soil Dyn Earthq Eng 2002;22:525–40.
[7] Di Giacomo D, Maria Rosaria G, Mucciarelli, M, Parolai S, Richwalski SM. Analysis and Modeling of HVSR in the Presence of a Velocity Inversion: The Case of Venosa, Italy. Bull Seismol Soc Am 2005;95:2364–72.
[8] Atkinson GM. Single-Station Sigma. Bull Seismol Soc Am 2006;96:446–55.
[9] Karabulut H, Bouchon M. Spatial variability and non-linearity of strong ground motion near a fault. Geophys J Int 2007;170:262–74.
[10] Abrahamson NA, Hollenback JC. Application of single-station sigma ground motion prediction equations in practice. Proc. 15th World Conf. Earthq. Eng. 15WCEE, Lisbon, Portugal: 2012, Paper No. 2536.
[11] Rahpeyma S, Halldorsson B, Olivera C, Green RA, Jónsson S. Detailed site effect estimation in the presence of strong velocity reversals within a small-aperture strong-motion array in Iceland. Soil Dyn Earthq Eng 2016;89:136–51.
[12] Bolt BA, Tsai YB, Yeh K, Hsu MK. Earthquake strong motions recorded by a large near-source array of digital seismographs. Earthq Eng Struct Dyn 1982;10:561–73.

[13] Loh C-H. Analysis of the spatial variation of seismic waves and ground movements from smart-1 array data. Earthq Eng Struct Dyn 1985;13:561–581.
[14] Abrahamson NA. Statistical properties of peak ground accelerations recorded by the SMART 1 array. Bull Seismol Soc Am 1988;78:26–41.
[15] Harichandran RS. Estimating the spatial variation of earthquake ground motion from dense array recordings. Struct Saf 1991;10:219–33.
[16] Chiu H-C, Amirbekian RV, Bolt BA. Transferability of strong ground-motion coherency between the SMART1 and SMART2 arrays. Bull Seismol Soc Am 1995;85:342–348.
[17] Theodulidis NP, Bard P-Y. Horizontal to vertical spectral ratio and geological conditions: an analysis of strong motion data from Greece and Taiwan (SMART-1). Soil Dyn Earthq Eng 1995;14:177–97.
[18] Dimitriu P, Theodulidis N, Bard PY. Evidence of nonlinear site response in HVSR from SMART1 (Taiwan) data. Soil Dyn Earthq Eng 2000;20:155–65.
[19] Özel O, Cranswick E, Meremonte M, Erdik M, Safak E. Site effects in Avcilar, west of Istanbul, Turkey, from strong-and weak-motion data. Bull Seismol Soc Am 2002;92:499–508.
[20] Fäh D, Kind F, Giardini D. Inversion of local S-wave velocity structures from average H/V ratios, and their use for the estimation of site-effects. J Seismol 2003;7:449–67.
[21] Giulio GD, Cornou C, Ohrnberger M, Wathelet M, Rovelli A. Deriving Wavefield Characteristics and Shear-Velocity Profiles from Two- Dimensional Small-Aperture Arrays Analysis of Ambient Vibrations in a Small-Size Alluvial Basin, Colfiorito, Italy. Bull Seismol Soc Am 2006;96:1915–33.
[22] Halldorsson B, Sigbjörnsson R, Schweitzer J. ICEARRAY: the first small-aperture, strong-motion array in Iceland. J Seismol 2009;13:173–8.
[23] Halldorsson B, Jónsson S, Papageorgiou AS, Green RA, Kalogeras I, Avery H, et al. ICEARRAY II: A new multidisciplinary strong-motion array in North Iceland. 15th World Conf. Earthq. Eng. 15WCEE, Lisbon, Portugal: 2012, Paper no. 2567.
[24] Metzger S, Jónsson S. Plate boundary deformation in North Iceland during 1992–2009 revealed by InSAR time-series analysis and GPS. Tectonophysics 2014;634:127–138.
[25] Olivera CI, Halldorsson B, Green RA, Sigbjörnsson R. Site effects estimation using ambient noise and earthquake data in Húsavík, north Iceland. Proc. 2nd Eur. Conf. Earthq. Eng. Seismol. 2ECEES, Istanbul, Turkey: 2014.
[26] Rahpeyma S, Halldorsson B, Green RA. On the Distribution of Earthquake Strong-motion Amplitudes and Site Effects Across the Icelandic Strong-motion Arrays. 16th World Conf. Earthq. Eng. 16WCEE, Santiago, Chile: 2017, Paper no. 2762.
[27] Sabetta F, Pugliese A. Attenuation of peak horizontal acceleration and velocity from italian strong-motion records. Bull Seismol Soc Am 1987;77:1491–513.
[28] Campbell KW, Bozorgnia Y. Empirical analysis of strong ground motion from the 1992 Landers, California, earthquake. Bull Seismol Soc Am 1994;84:573–88.
[29] Waltl P, Halldorsson B, Pétursson HG, Fiebig M, Sigbjörnsson R. The geological and urban setting of Húsavík, North Iceland, in the context of Earthquake Hazard and Risk Analysis. Proc. 2nd Eur. Conf. Earthq. Eng. Seismol. 2ECEES, Istanbul, Turkey: 2014.
[30] Abrahamson NA, Shedlock KM. Overview. Seismol Res Lett 1997;68:9–23.
[31] Abrahamson NA, Atkinson G, Boore DM, Bozorgnia Y, Campbell K, Chiou BS-J, et al. Comparisons of the NGA ground-motion relations. Earthq Spectra 2008;24:45.
[32] Douglas J, Aochi H. A Survey of Techniques for Predicting Earthquake Ground Motions for Engineering Purposes. Surv Geophys 2008;29:187–220.
[33] Strasser FO, Abrahamson NA, Bommer JJ. Sigma: Issues, Insights, and Challenges. Seismol Res Lett 2009;80:40–56.
[34] Douglas J. Consistency of ground-motion predictions from the past four decades. Bull Earthq Eng 2010;8:1515–26.
[35] Bommer JJ, Abrahamson NA, Strasser FO, Pecker A, Bard P-Y, Bungum H, et al. The Challenge of Defining Upper Bounds on Earthquake Ground Motions. Seismol Res Lett 2004;75:82–95.
[36] Bommer JJ, Abrahamson NA. Why do modern probabilistic seismic-hazard analyses often lead to increased hazard estimates? Bull Seismol Soc Am 2006;96:1967–1977.

[37] Kowsari M, Halldorsson B, Hrafnkelsson B, Snæbjörnsson JT, Ólafsson S, Rupakhety R. On the Selection of Ground-Motion Prediction Equations for Seismic Hazard Assessment in the South Iceland Seismic Zone. 16th World Conf. Earthq. Eng. 16WCEE, Santiago, Chile: 2017, Paper no. 2809.
[38] Joyner WB, Boore DM. Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake. Bull Seismol Soc Am 1981;71:2011–2038.
[39] Fukushima Y, Tanaka T. A new attenuation relation for peak horizontal acceleration of strong earthquake ground motion in Japan. Bull Seismol Soc Am 1990;80:757–783.
[40] Abrahamson NA, Youngs RR. A stable algorithm for regression analyses using the random effects model. Bull Seismol Soc Am 1992;82:505–10.
[41] Douglas J, Gehl P. Investigating strong ground-motion variability using analysis of variance and two-way-fit plots. Bull Earthq Eng 2008;6:389–405.
[42] Kuehn NM, Scherbaum F. Ground-motion prediction model building: a multilevel approach. Bull Earthq Eng 2015;13:2481–91.
[43] Chin B-H, Aki K. Simultaneous study of the source, path, and site effects on strong ground motion during the 1989 Loma Prieta earthquake: a preliminary result on pervasive nonlinear site effects. Bull Seismol Soc Am 1991;81:1859–1884.
[44] Boatwright J, Fletcher JB, Fumal TE. A general inversion scheme for source, site, and propagation characteristics using multiply recorded sets of moderate-sized earthquakes. Bull Seismol Soc Am 1991;81:1754–82.
[45] Moya A, Aguirre J, Irikura K. Inversion of Source Parameters and Site Effects from Strong Ground Motion Records using Genetic Algorithms. Bull Seismol Soc Am 2000;90:977–92.
[46] Shabestari KT, Yamazaki F, Saita J, Matsuoka M. Estimation of the spatial distribution of ground motion parameters for two recent earthquakes in Japan. Tectonophysics 2004;390:193–204.
[47] Ortiz-Alemán C, Reyes-Olvera M, Iglesias-Mendoza A, Orozco-Del-Castillo MG, Hernández-Gómez JJ. Estimation of source; path and site effects at MASE array stations: a comprehensive study. Geofísica Int 2017;56.
[48] Aki K. Space and Time Spectra of Stationary Stochastic Waves, with Special Reference to Microtremors 1957.
[49] Borcherdt RD. Effects of local geology on ground motion near San Francisco Bay. Bull Seismol Soc Am 1970;60:29–61.
[50] Nakamura Y. A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface. Q Rep Railw Tech Res Inst 1989;30:25–33.
[51] Bard PY. Microtremor measurements: A tool for site effect estimation? Proceeding Second Int. Symp. Eff. Surf. Geol. Seism. Motion Yokohama Japan; 1998, p. 1251–79.
[52] Rahpeyma S, Halldorsson B, Hrafnkelsson B, Jónsson S. Bayesian hierarchical model for variations in earthquake peak ground acceleration within small-aperture arrays. Environmetrics 2018;29.3, e2497.
[53] Vanem E. Bayesian Hierarchical Space-Time Models with Application to Significant Wave Height. Springer Science & Business Media; 2013.
[54] Banerjee S, Carlin BP, Gelfand AE. Hierarchical modeling and analysis for spatial data. Crc Press; 2014.
[55] Sigurdarson AN, Hrafnkelsson B. Bayesian prediction of monthly precipitation on a fine grid using covariates based on a regional meteorological model. Environmetrics 2016;27:27–41.
[56] Kuehn NM, Scherbaum F. A partially non-ergodic ground-motion prediction equation for Europe and the Middle East. Bull Earthq Eng 2016:1–14.
[57] Halldorsson B, Papageorgiou AS. Calibration of the Specific Barrier Model to Earthquakes of Different Tectonic Regions. Bull Seismol Soc Am 2005;95:1276–300.
[58] Halldorsson B, Sigbjörnsson R. The Mw6. 3 Ölfus earthquake at 15: 45 UTC on 29 May 2008 in South Iceland: ICEARRAY strong-motion recordings. Soil Dyn Earthq Eng 2009;29:1073–83.
[59] Sigbjörnsson R, Snæbjörnsson JT, Higgins SM, Halldorsson B, Ólafsson S. A note on the Mw 6.3 earthquake in Iceland on 29 May 2008 at 15:45 UTC. Bull Earthq Eng 2009;7:113–26.

[60] Douglas J, Halldorsson B. On the use of aftershocks when deriving ground-motion prediction equations. 9th US Natl. 10th Can. Conf. Earthq. Eng. 9USN10CCEE, Toronto, Canada: 2010, Paper no. 220.
[61] Sæmundsson K, Kristinsson S. Hveragerði. Soil temperature measurements and faults. Reykjavík, Iceland (in Icelandic): Iceland GeoSurvey (ÍSOR); 2005.
[62] Böðvarsson R, Rögnvaldsson ST, Slunga R, Kjartansson E. The SIL data acquisition system-at present and beyond year 2000. Phys Earth Planet Inter 1999;113:89–101.
[63] Panzera F, Zechar JD, Vogfjörd KS, Eberhard DAJ. A Revised Earthquake Catalogue for South Iceland. Pure Appl Geophys 2016;173:97–116.
[64] Panzera F, Mignan A, Vogfjörð KS. Spatiotemporal evolution of the completeness magnitude of the Icelandic earthquake catalogue from 1991 to 2013. J Seismol 2017;21:615–30.
[65] Brandsdóttir B, Parsons M, White RS, Guðmundsson Ó, Drew J, Thorbjarnardóttir BS. The May 29th 2008 earthquake aftershock sequence within the South Iceland Seismic Zone: Fault locations and source parameters of aftershocks. Jökull 2010;60:23–46.
[66] Passarelli L, Rivalta E, Jónsson S, Hensch M, Metzger S, Jakobsdóttir SS, et al. Scaling and spatial complementarity of tectonic earthquake swarms. Earth Planet Sci Lett 2018;482:62–70.
[67] Sæmundsson K, Karson JA. Stratigraphy and tectonics of the Húsavík-Western Tjörnes Area. Iceland Geosurvey; 2006.
[68] Halldorsson B, Avery H. Converting Strong-motion Networks to Arrays via Common Triggering. Seismol Res Lett 2009;80:572–8.
[69] Brillinger DR, Preisler HK. An exploratory analysis of the Joyner-Boore attenuation data. Bull Seismol Soc Am 1984;74:1441–50.
[70] Guttorp P, Gneiting T. Studies in the History of Probability and Statistics XLIX on the Matérn Correlation Family. Biometrika 2006;93:989–95.
[71] Stein ML. Interpolation of Spatial Data: Some Theory for Kriging. Springer Science & Business Media; 2012.
[72] Olafsson S, Sigbjörnsson R. Attenuation of Strong Ground Acceleration: A Study of the South Iceland Earthquakes 2000. Manuscr Submitt Publ Earthq Eng Struct Dyn 2000.
[73] Ambraseys NN, Douglas J, Sarma SK, Smit PM. Equations for the estimation of strong ground motions from shallow crustal earthquakes using data from Europe and the Middle East: horizontal peak ground acceleration and spectral acceleration. Bull Earthq Eng 2005;3:1–53.
[74] Morikawa N, Kanno T, Narita A, Fujiwara H, Okumura T, Fukushima Y, et al. Strong motion uncertainty determined from observed records by dense network in Japan. J Seismol 2008;12:529–46.
[75] Rodriguez-Marek A, Montalva GA, Cotton F, Bonilla F. Analysis of Single-Station Standard Deviation Using the KiK-net Data. Bull Seismol Soc Am 2011;101:1242–58.
[76] Lin P-S, Chiou B, Abrahamson N, Walling M, Lee C-T, Cheng C-T. Repeatable Source, Site, and Path Effects on the Standard Deviation for Empirical Ground-Motion Prediction Models. Bull Seismol Soc Am 2011;101:2281–95.
[77] Lin P-S, Chiou B, Abrahamson N, Walling M, Lee C-T, Cheng C-T. Repeatable Source, Site, and Path Effects on the Standard Deviation for Empirical Ground-Motion Prediction ModelsRepeatable Source, Site, and Path Effects on the Standard Deviation for Ground-Motion Prediction. Bull Seismol Soc Am 2011;101:2281–95.
[78] Chen L, Faccioli E. Single-station standard deviation analysis of 2010–2012 strong-motion data from the Canterbury region, New Zealand. Bull Earthq Eng 2013;11:1617–32.
[79] Zafarani H, Soghrat MR. Single-Station Sigma for the Iranian Strong Motion Stations. Pure Appl Geophys 2017:1–23.
[80] Ktenidou O-J, Roumelioti Z, Abrahamson N, Cotton F, Pitilakis K, Hollender F. Understanding single-station ground motion variability and uncertainty (sigma): lessons learnt from EUROSEISTEST. Bull Earthq Eng 2017:1–26.
[81] Cotton F, Pousse G, Bonilla F, Scherbaum F. On the Discrepancy of Recent European Ground-Motion Observations and Predictions from Empirical Models: Analysis of KiK-net Accelerometric Data and Point-Sources Stochastic SimulationsDiscrepancy of Recent European Ground-Motion Observations and Predictions from Empirical Models. Bull Seismol Soc Am 2008;98:2244–61.

[82] Mahani AB, Atkinson GM. Evaluation of Functional Forms for the Attenuation of Small- to-Moderate- Earthquake Response Spectral Amplitudes in North AmericaEvaluation of Functional Forms for Attenuation of Small- to- Moderate- Earthquake Response Spectral Amplitudes. Bull Seismol Soc Am 2012;102:2714–26.
[83] Anderson JG. Expected Shape of Regressions for Ground-Motion Parameters on Rock. Bull Seismol Soc Am 2000;90:S43–52.
[84] Kotha SR, Bindi D, Cotton F. Partially non-ergodic region specific GMPE for Europe and Middle-East. Bull Earthq Eng 2016;14:1245–1263.
[85] Rahpeyma S, Halldorsson B, Hrafnkelsson B. Bayesian Hierarchical Model of Peak Ground Acceleration for the Icelandic Strong-motion Arrays. Int. Conf. Earthq. Eng. Struct. Dyn. ICESD, Reykjavík, Iceland: 2017.
[86] Waltl P. Geomorphology and Building Stock of Húsavík, North Iceland: A Uniform GIS Database for Application in Hazard and Risk Modeling. Master's Thesis. BOKU University of Natural Resources and Life Sciences. Vienna, Austria, 2013.

FIGURES

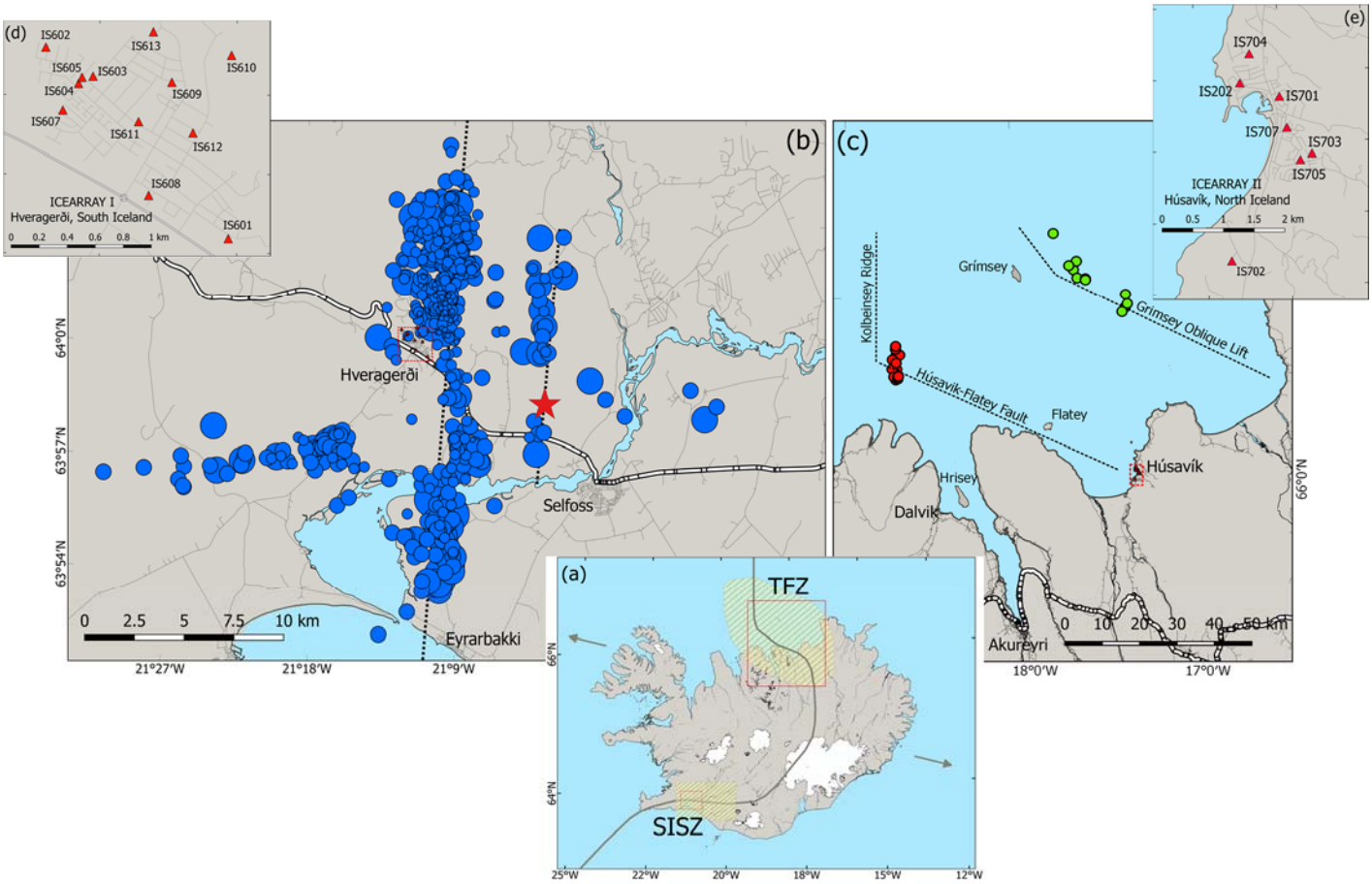

Fig. 1. (a) Map of Iceland with the approximate location of the Mid-Atlantic Ridge (grey line) and the South Iceland Seismic Zone (SISZ) and the Tjörnes Fracture Zone (TFZ) marked by hatched areas. Red rectangles indicate the areas shown in (b) and (c) with more details. (b) The aftershock distribution (blue circles) from 29 May 2008 Ölfus earthquake in southwest Iceland outlining the two causative earthquake faults (dotted lines). The twelve ICEARRAY I stations (red triangles shown in (d) along with station ID-codes) are located within the town of Hveragerði (red dashed rectangle shown in (b)). (c) Locations of the main events during the October 2012 (red circles) and April 2013 (green circles) earthquake swarms in the TFZ in addition to the main seismic lineaments of the TFZ as black dotted lines. The seven ICEARRAY II stations (red triangular shown in (e) along with station ID-codes) are located within the town of Húsavík (red dashed rectangle shown in (c)).

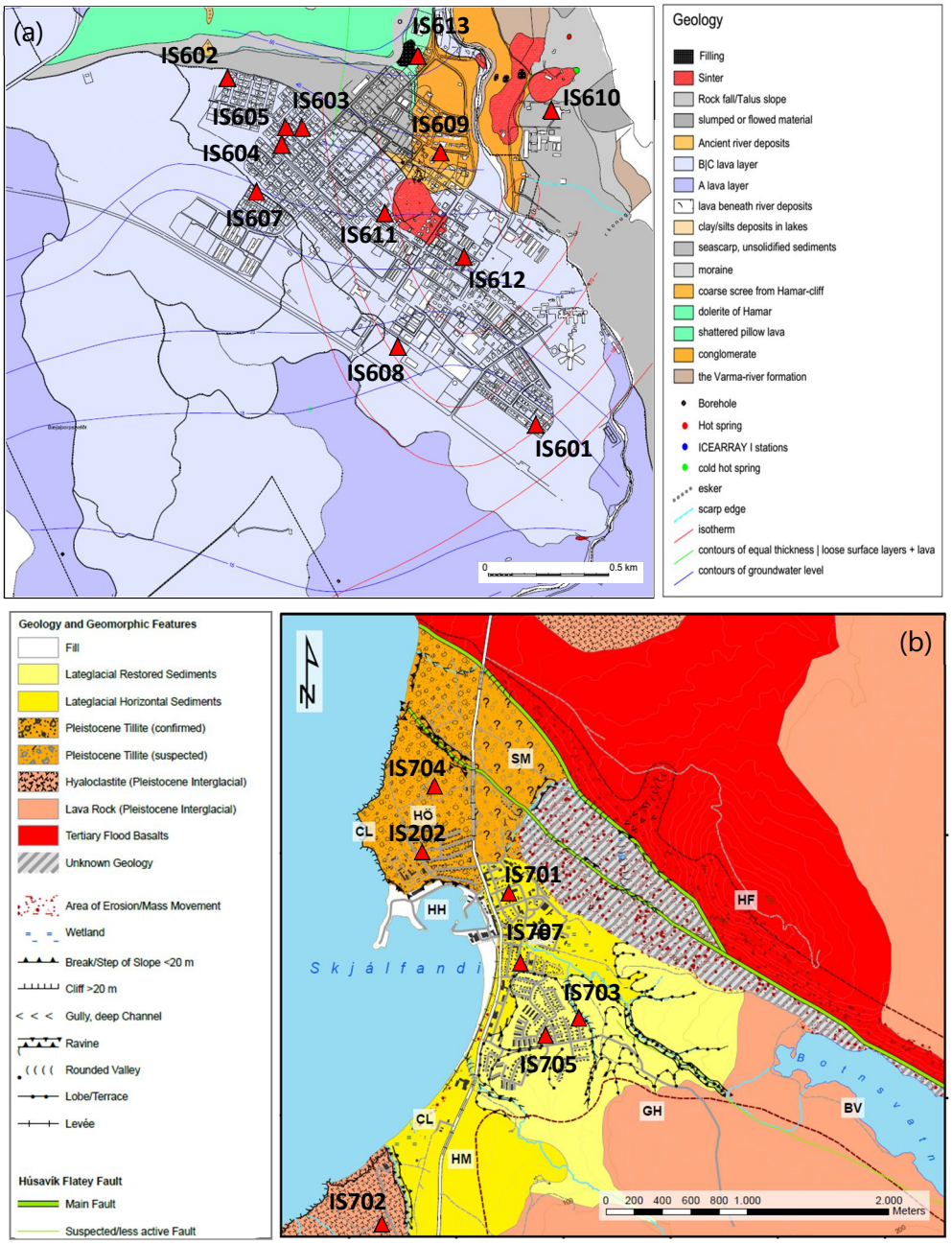

Fig. 2. Geological maps of (a) Hveragerði in southwest Iceland, where ICEARRAY I is located, adopted from [61] and of (b) Húsavík and surroundings, the location of ICEARRAY II, adopted from [29]. The array station locations are marked as red triangular.

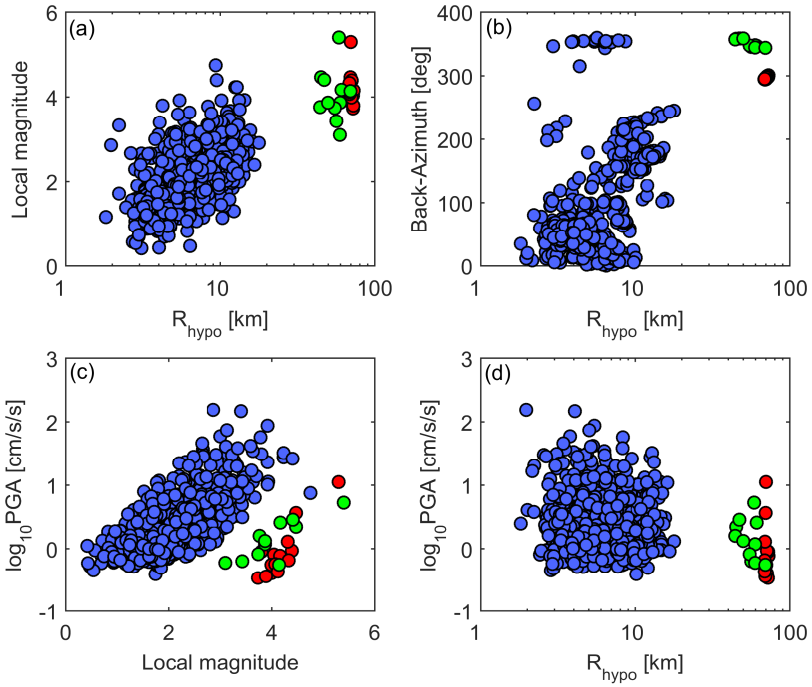

Fig. 3. Distribution of different earthquake parameters with (a) local magnitude and (b) back-azimuth plotted versus hypocentral distance (R_{hypo}), and then $\log_{10}$ PGA plotted versus (c) local magnitude and (d) hypocentral distance as recorded by ICEARRAY I (blue circles) and ICEARRAY II (red squares for the swarm in 2012 and green squares for the 2013 swarm).

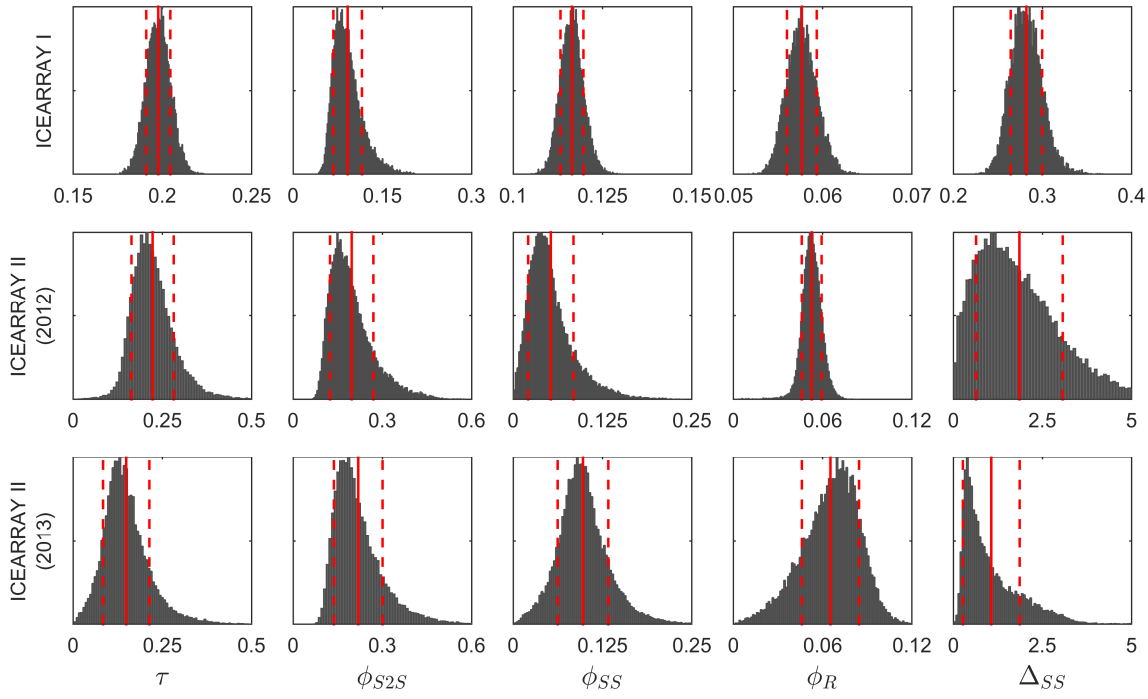

Fig. 4. Posterior histograms of hyperparameters, $\boldsymbol{\theta}$, for ICEARRAY I (1st row) and ICEARRAY II (2nd and 3rd rows for swarm 2012 and swarm 2013, respectively). The red solid and dashed lines show posterior mean and posterior mean ± 1 posterior standard deviation, respectively.

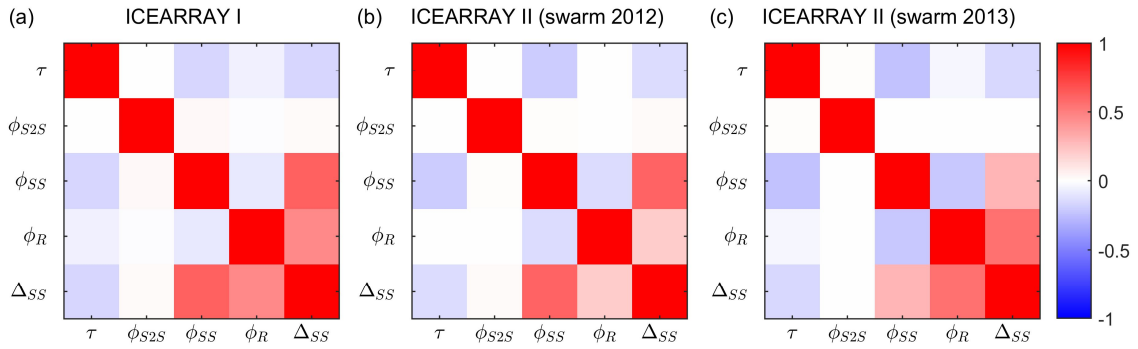

Fig. 5. Correlation coefficient matrices of hyperparameters, $\boldsymbol{\theta}$, across (a) ICEARRAY I, (b) ICEARRAY II for swarm 2012, and (c) ICEARRAY II for swarm 2013.

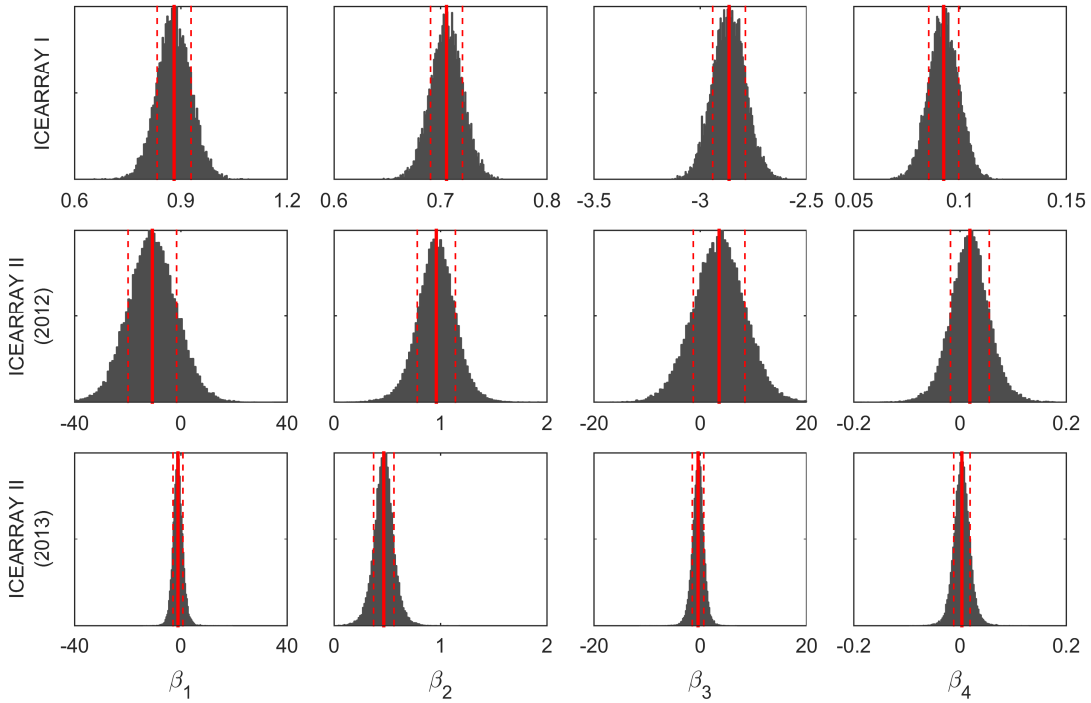

Fig. 6. Posterior histograms of model coefficients, $\boldsymbol{\beta}$, for ICEARRAY I and ICEARRAY II using four chains within the MCMC simulations. The red solid and dashed lines show posterior mean and posterior mean ± 1 posterior standard deviation, respectively.

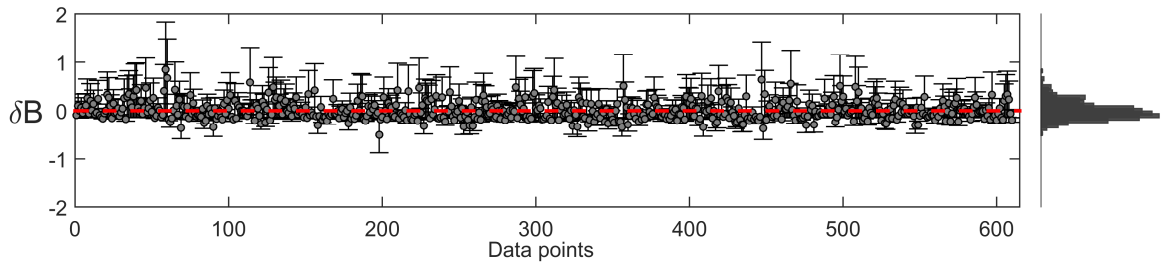

Fig.7. The individual earthquake event terms (of $\boldsymbol{\delta B}$) for the ICEARRAY I. The plot shows the posterior mean (grey dots) and 95% posterior intervals (error bars) by using four chains in the MCMC simulations. The histogram on the right represents the overall distribution of posterior means of the elements in $\boldsymbol{\delta B}$. The red dashed line represents the zero-reference line.

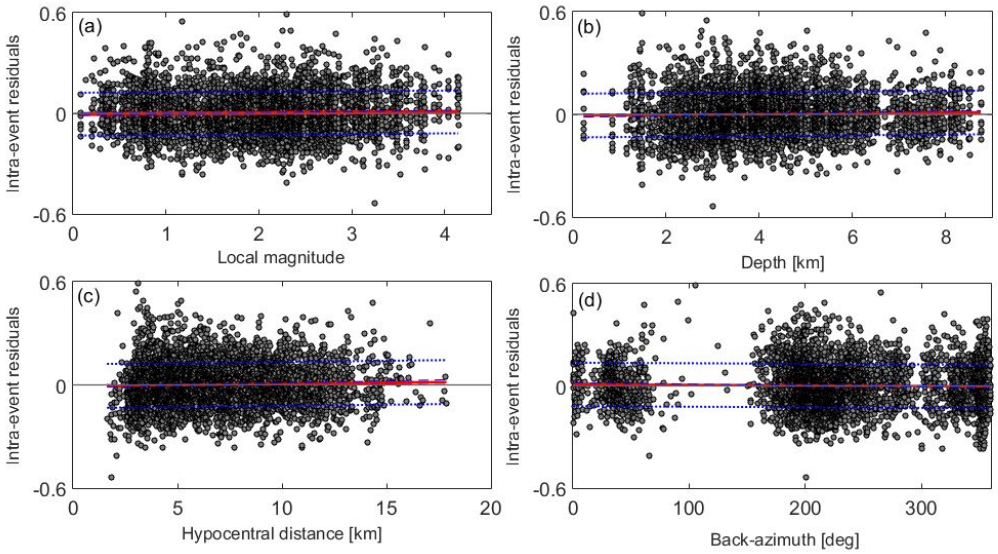

Fig. 8. Distribution of intra-event residuals versus (a) local magnitude (b) depth (c) hypocentral distance and (d) back-azimuth for ICEARRAY-I. The red thick, blue dashed, and blue dotted lines show the sample regression lines, 95% confidence limits for the ordinate to the true regression lines, and regression line ± 1 standard deviations, respectively.

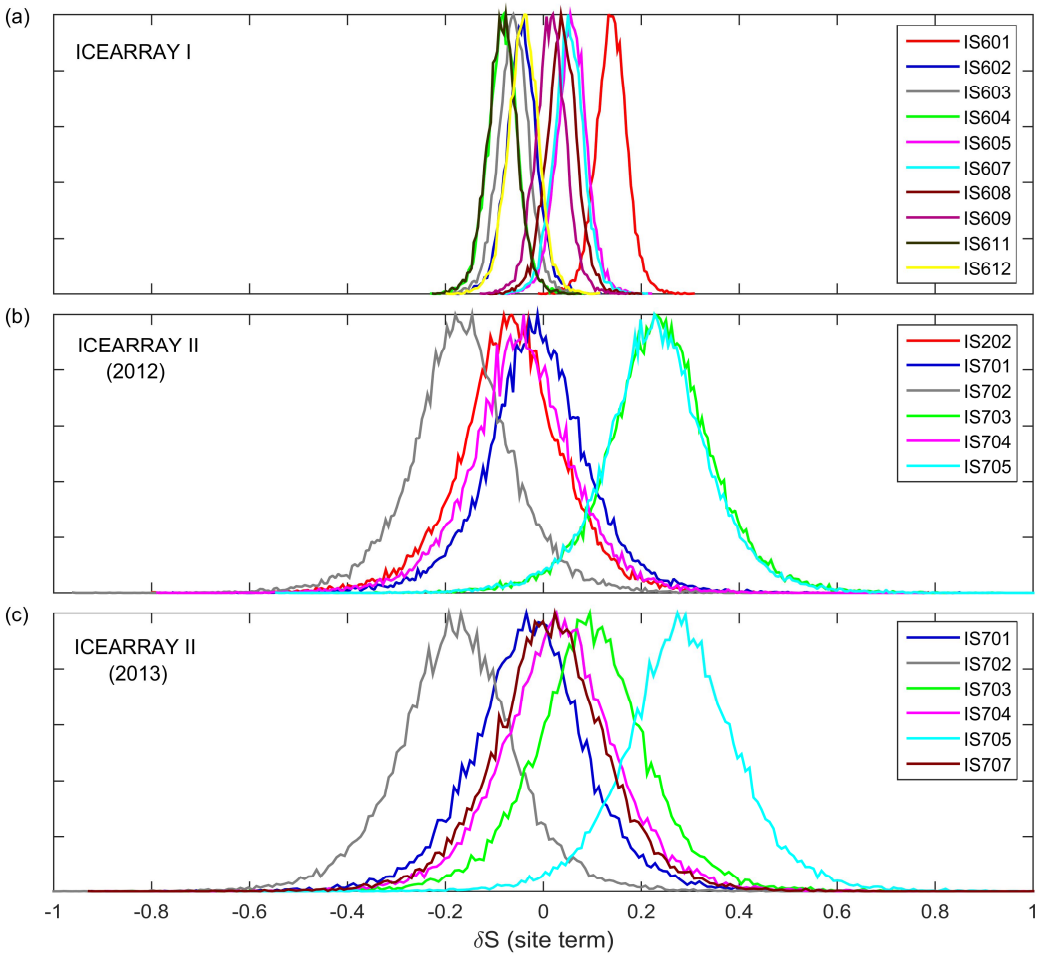

Fig. 9. Posterior distributions of station terms, $\boldsymbol{\delta S}$, for ICEARRAY I (a) and ICEARRAY II 2012 and 2013 sequences (b and c, respectively) using four chains in the MCMC simulations.

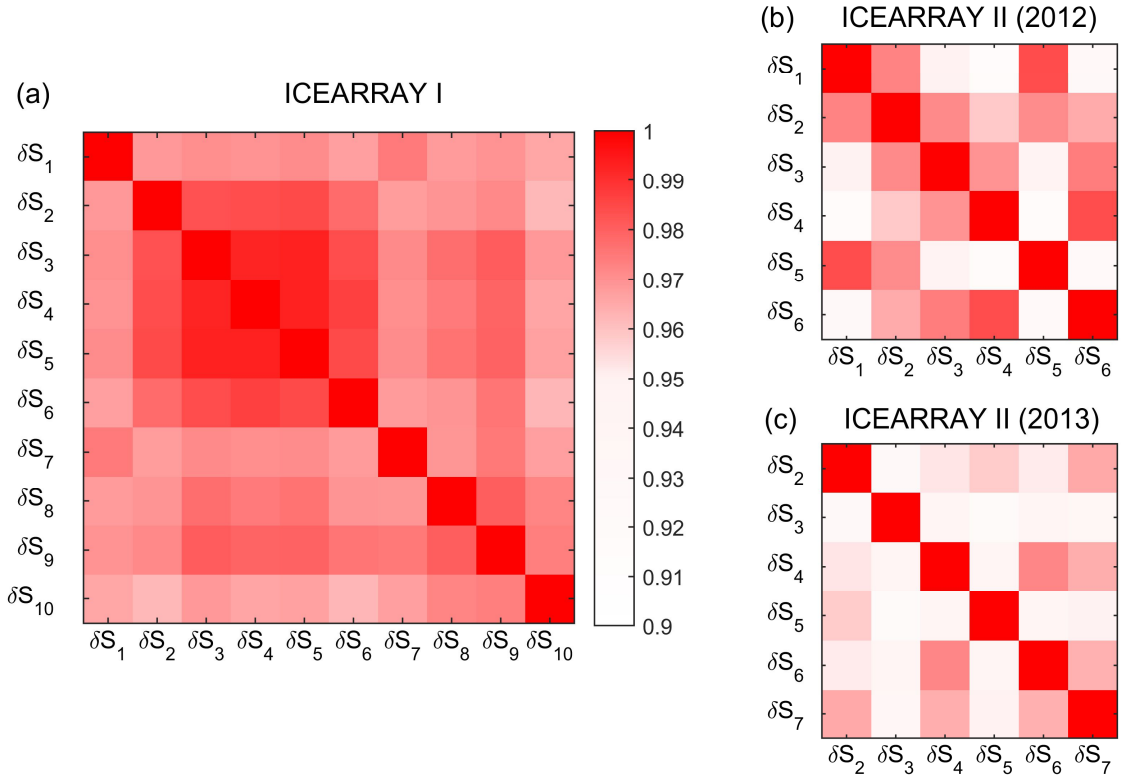

Fig. 10. Station effect correlation matricies for (a) ICEARRAY I and (b-c) ICEARRAY II strong-motion stations. See Table 3 for the denotation of δS's with station numbers, and **Error! Reference source not found.** for station locations.

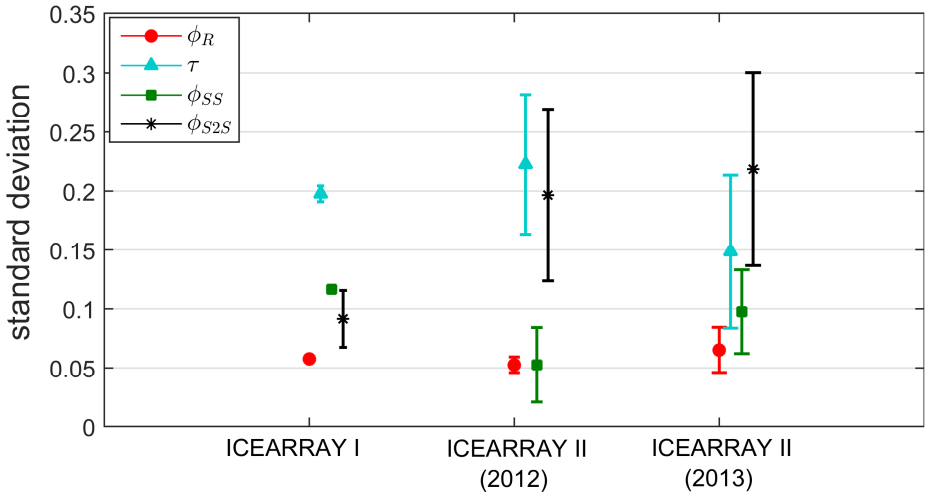

Fig. 11. Comparing means and standard deviations of the inter-event (τ, cyan triangle), inter-station (ϕ_{S2S}, black asterisk), event-station (ϕ_{SS}, green square), and unexplained term (ϕ_R, red circle) variability for ICEARRAY I and ICEARRAY II.

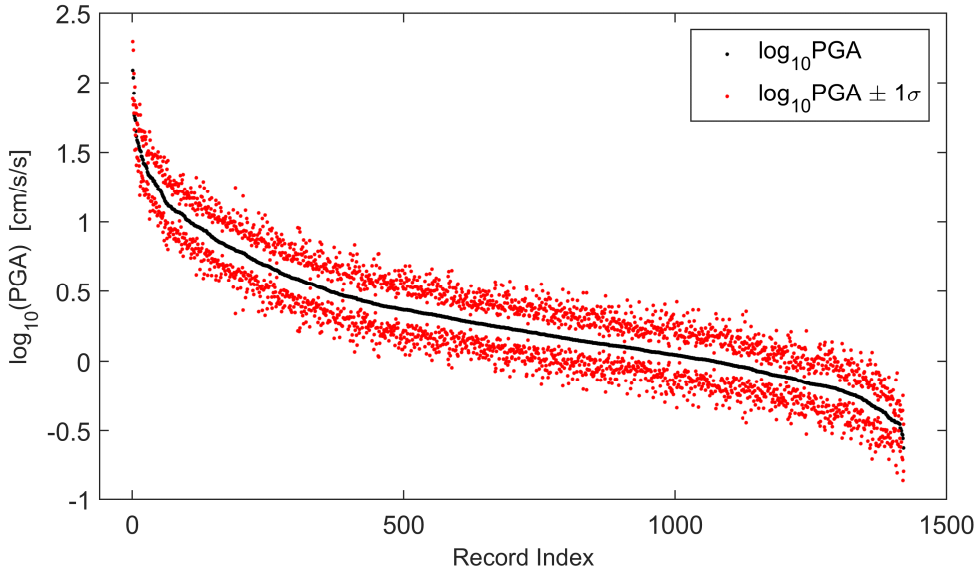

Fig. 12. The median of the geometric mean PGA (black symbols) for each aftershock recorded by more than 5 stations on the ICEARRAY I $\pm 1\sigma$ (red symbols).

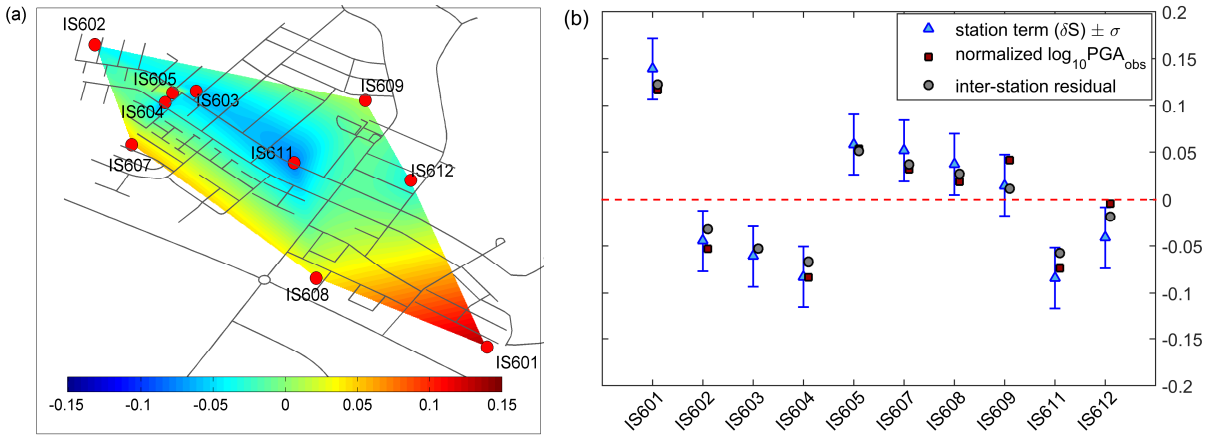

Fig. 13. Normalized observations of $\log_{10}$ PGA (a) interpolated across ICEARRAY I (b) and compared (as squares) with the posterior mean of station effects, $\delta S \pm 1\sigma$, at each station (triangles with error bars). Also shown are the means of inter-station residuals for each station. The plots show that the BHM captures the observed spatial distribution of $\log_{10}$ PGA completely.

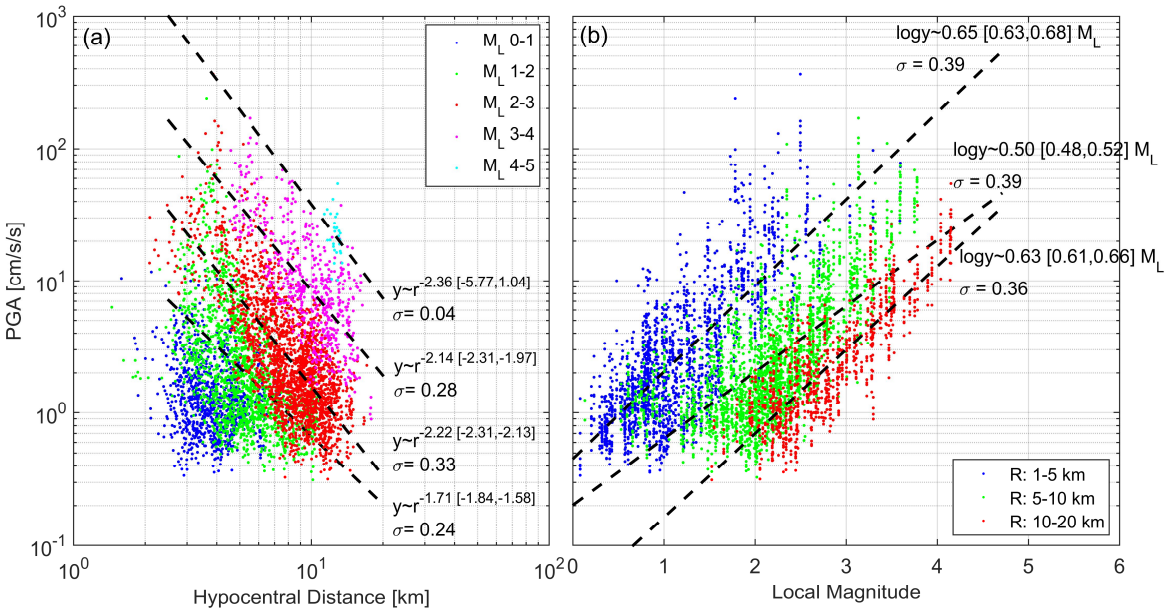

Fig. 14. The geometric mean PGA of horizontal components across the ICEARRAY I plotted vs. (a) hypocentral distance and clustered by local magnitude, M_L, along with straight lines fitted through the PGA for four magnitude clusters and (b) vs. M_L clustered by distance with fitted lines. The scaling with distance and magnitude is given along with 95% confidence bounds for the coefficients and associated standard deviations.

TABLES

Table 1. Posterior statistics for hyperparameters $\boldsymbol{\theta}$ for ICEARRAY I and ICEARRAY II. The statistics are given as posterior means and standard deviations along with posterior distribution percentiles. The ICEARRAY II statistics are provided separately for the 2012 and 2013 swarms.

θ	Mean	SD	Percentiles				
			2.5%	25%	50%	75%	97.5%
ICEARRAY I							
τ	0.1977	0.0067	0.1847	0.1930	0.1977	0.2022	0.2112
ϕ_{S2S}	0.0915	0.0240	0.0571	0.0741	0.0873	0.1039	0.1510
ϕ_{SS}	0.1164	0.0032	0.1104	0.1142	0.1164	0.1185	0.1231
ϕ_R	0.0577	0.0017	0.0546	0.0565	0.0576	0.0588	0.0611
Δ_{SS}	0.2819	0.0177	0.2502	0.2697	0.2810	0.2932	0.3193
ICEARRAY II (swarm 2012)							
τ	0.2225	0.0592	0.1293	0.1826	0.2150	0.2538	0.3586
ϕ_{S2S}	0.1966	0.0729	0.0999	0.1442	0.1807	0.2319	0.3834
ϕ_{SS}	0.0525	0.0318	0.0092	0.0305	0.0460	0.0668	0.1349
ϕ_R	0.0527	0.0067	0.0400	0.0487	0.0526	0.0568	0.0654
Δ_{SS}	1.8512	1.2159	0.1800	0.9278	1.6235	2.5388	4.7670
ICEARRAY II (swarm 2013)							
τ	0.1486	0.0649	0.0426	0.1067	0.1401	0.1808	0.3058
ϕ_{S2S}	0.2185	0.0818	0.1121	0.1611	0.2003	0.2566	0.4277
ϕ_{SS}	0.0976	0.0354	0.0316	0.0754	0.0950	0.1170	0.1789
ϕ_R	0.0653	0.0192	0.0227	0.0532	0.0674	0.0791	0.0980
Δ_{SS}	1.0617	0.7966	0.2038	0.4546	0.7945	1.4762	3.0261

Table 2. Posterior statistics for the latent parameters $\boldsymbol{\beta}$, i.e., the coefficients of the GMM, for ICEARRAY I and ICEARRAY II. The statistics are given as posterior means and standard deviations along with posterior distribution percentiles. The ICEARRAY II statistics are provided separately for the 2012 and 2013 swarms.

β	Mean	SD	Percentiles				
			2.5%	25%	50%	75%	97.5%
ICEARRAY I							
β_1	0.8826	0.0479	0.7875	0.8505	0.8827	0.9130	0.9800
β_2	0.7058	0.0150	0.6768	0.6957	0.7059	0.7160	0.7355
β_3	-2.8658	0.0764	-3.0142	-2.9161	-2.8647	-2.8156	-2.7175
β_4	0.0923	0.0070	0.0787	0.0876	0.0923	0.0969	0.1062
ICEARRAY II (swarm 2012)							
β_1	-10.8730	8.5864	-27.4308	-16.6367	-10.9166	-5.2297	6.0997
β_2	0.9631	0.1636	0.6338	0.8589	0.9643	1.0692	1.2824
β_3	3.6068	4.5663	-5.4280	0.6100	3.6319	6.6671	12.4481
β_4	0.0183	0.0331	-0.0467	-0.0027	0.0181	0.0394	0.0849
ICEARRAY II (swarm 2013)							
β_1	-1.2448	1.5694	-4.3252	-2.2195	-1.2707	-0.2923	1.9319
β_2	0.4674	0.0807	0.3061	0.4182	0.4671	0.5169	0.6292
β_3	-0.3346	0.9106	-2.1771	-0.8873	-0.3236	0.2365	1.4587
β_4	0.0032	0.0132	-0.0229	-0.0049	0.0031	0.0111	0.0300

Table 3. Posterior statistics for the latent parameters $\boldsymbol{\delta S}$, i.e., the station terms for ICEARRAY I and ICEARRAY II. The statistics are given as posterior means and standard deviations along with posterior distribution percentiles. The ICEARRAY II statistics are provided separately for the 2012 and 2013 swarms.

θ	Station	Mean	SD	Percentiles				
				2.5%	25%	50%	75%	97.5%
ICEARRAY I								
δS_1	IS601	0.1393	0.0323	0.0754	0.1187	0.1395	0.1596	0.2035
δS_2	IS602	-0.0443	0.0323	-0.1090	-0.0645	-0.0443	-0.0239	0.0194
δS_3	IS603	-0.0610	0.0322	-0.1248	-0.0813	-0.0608	-0.0406	0.0023
δS_4	IS604	-0.0828	0.0322	-0.1471	-0.1029	-0.0825	-0.0627	-0.0190
δS_5	IS605	0.0580	0.0322	-0.0057	0.0377	0.0582	0.0783	0.1210
δS_6	IS607	0.0518	0.0322	-0.0126	0.0317	0.0520	0.0720	0.1154
δS_7	IS608	0.0373	0.0324	-0.0277	0.0171	0.0375	0.0580	0.1015
δS_8	IS609	0.0149	0.0324	-0.0502	-0.0053	0.0153	0.0354	0.0786
δS_9	IS611	-0.0841	0.0323	-0.1483	-0.1042	-0.0837	-0.0636	-0.0207
δS_{10}	IS612	-0.0409	0.0325	-0.1063	-0.0614	-0.0405	-0.0200	0.0244
ICEARRAY II (swarm 2012)								
δS_1	IS202	-0.0666	0.1147	-0.3039	-0.1320	-0.0657	0.0004	0.1591
δS_2	IS701	-0.0139	0.1126	-0.2453	-0.0781	-0.0138	0.0515	0.2119
δS_3	IS702	-0.1667	0.1128	-0.3978	-0.2309	-0.1670	-0.1013	0.0603
δS_4	IS703	0.2391	0.1146	0.0081	0.1718	0.2381	0.3060	0.4708
δS_5	IS704	-0.0461	0.1146	-0.2832	-0.1119	-0.0455	0.0215	0.1807
δS_6	IS705	0.2323	0.1139	0.0006	0.1657	0.2312	0.2984	0.4626
ICEARRAY II (swarm 2013)								
δS_2	IS701	-0.0280	0.1289	-0.2881	-0.1031	-0.0269	0.0471	0.2327
δS_3	IS702	-0.1794	0.1298	-0.4449	-0.2555	-0.1775	-0.1008	0.0782
δS_4	IS703	0.0948	0.1290	-0.1695	0.0195	0.0950	0.1710	0.3535
δS_5	IS704	0.0316	0.1294	-0.2305	-0.0445	0.0316	0.1077	0.2945
δS_6	IS705	0.2827	0.1291	0.0239	0.2069	0.2818	0.3589	0.5455
δS_7	IS707	0.0129	0.1288	-0.2480	-0.0617	0.0133	0.0887	0.2715

Paper IV

Estimate of Shear-wave Velocity profile using microseismic Horizontal-to-Vertical Spectral Ratios Inversion

Sahar Rahpeyma, Benedikt Halldorsson, Birgir Hrafnkelsson, Sigurjon Jónsson, and Orhan Polat

Rahpeyma S, Halldorsson B, Hrafnkelsson B, Jónsson S, and Polat O (2018b) Estimation of Shear-wave velocity profile using microseismic Horizontal-to-Vertical-Spectral Ratios Inversion. *Soil Dynamics and Earthquake Engineering, (Manuscript completed)*

Estimate of shear-wave velocity profile using microseismic Horizontal-to-Vertical Spectral Ratios inversion

Sahar Rahpeyma[1], Benedikt Halldorsson[2*,3], Birgir Hrafnkelsson[4], Russel A. Green [5], Sigurión Jóhnsson[6], and Orhan Polat[7]

Abstract: The shear-wave velocity (V_s) model plays an important role in many engineering applications. The V_s profile can be retrieved either with invasive or non-invasive techniques. The non-invasive methods are known as practical and cost-efficient alternatives; however, the inversion problems are highly nonlinear and can be influenced by non-uniqueness solution which leads to high level of uncertainties. It is well established that the Horizontal-to-Vertical Spectral Ratio (HVSR) results represent the subsoil characteristics, particularly when there is a significant impedance contrast between layers. In this paper, we model the seismic body waves in the layered medium and simulate the theoretical HVSR on the basis of an approximate initial soil profile model. We then apply a robust and computationally efficient Bayesian inversion technique which provides an estimate of the corresponding soil structure. We test the proposed technique using microseismic recordings at two nominated reference sites from IzmirNet stations in Turkey and Mirandola station in Italy. The Markov Chain Monte Carlo (MCMC) technique with embedded Metropolis steps is employed in order to obtain the best fitting family of V_s profiles along with their uncertainties. The theoretical HVSRs are calculated through body-wave approximation as a reliable estimate for subsoil structure of the sedimentary layers overlaying the half-space. A blind test is conducted over the number of layers to consistently investigate the best resolution of model parametrization. The reliable agreement between the results and available soil properties information confirms the applicability of the proposed approach in this study.

Keywords: Bayesian approach, HVSR inversion, Markov Chain Monte Carlo, Body-waves approximation, Uncertainties.

[1] Earthquake Engineering Research Centre & Faculty of Civil and Environmental Engineering, School of Engineering and Natural Sciences, University of Iceland, Selfoss, Iceland. sahar@hi.is
[2*] Earthquake Engineering Research Centre & Faculty of Civil and Environmental Engineering, School of Engineering and Natural Sciences, University of Iceland, Selfoss, Iceland. skykkur@hi.is (corresponding author)
[3] Geoscience Research Group, Division of Processing and Research, Icelandic Meteorological Office, Reykjavík, Iceland
[4] Faculty of Physical Science, School of Engineering and Natural Sciences, University of Iceland, Iceland. Birgirhr@hi.is
[5] Department of Civil & Environmental Engineering, Virginia Tech, Blacksburg, VA, USA. rugreen@vt.edu
[6] King Abdullah University of Science and Technology, Physical Science and Engineering Division, Thuwal, 23955-6900, Saudi Arabia. sigurjon.jonsson@kaust.edu.sa
[7] Dokuz Eylul University, Department of Geophysics, Izmir, Turkey, orhan.polat@deu.edu.tr

1. Introduction

Comprehensive site-specific studies from seismic wave amplification of the recent destructive earthquakes such as Michoacan earthquake (Mexico City, 1985); Loma Prieta earthquake (Northern California, 1989); Kobe earthquake (Japan, 1995); İzmit earthquake (Turkey, 1996) and Chi-Chi earthquake (Taiwan, 1999) have been drawing attention to the significant impact of the near-surface soil characteristics on ground motion and resulting damage distribution. In geotechnical as well as engineering applications, shear-wave velocity (V_s) is known as the most common physical indicator of the soil dynamic properties which characterizes the localized site conditions. In particular, seismic building codes such as Eurocode 8 (European Committee for Standardization 2003), American Society of Civil Engineers (ASCE 2007, 2010), and also National Earthquake Hazard Reduction Program, NEHRPs (Building Seismic Safety Council, BSSC 2003; Boore 2004) use V_{s30} (i.e., the average shear-wave velocity down to 30 m depth) to classify soil categories for evaluating the dynamic behavior of the soil in shallow surface. In addition, V_{s30} is considered in just about all Ground Motion Prediction Equations (GMPEs) as a key component in seismic hazard analysis which reflects the site amplification (Power et al. 2008; Boore et al. 2013). However, what is often missing in V_s modeling is reliable estimate of uncertainties due to non-uniqueness results and non-linear inversion techniques (Foti et al. 2009). Hence, in this study, we aim to propose a practical and reliable framework to find the V_s model with reasonable resolution of layering and associated uncertainties.

Estimation of the V_s profile (i.e., shear-wave velocity vs. depth) has been accomplished using a variety of either invasive (e.g., down-hole or cross-holes seismic surveys) or non-invasive (e.g., surface-wave or body-wave approaches, seismic travel-time inversions, and refraction-reflection analyzes) processing tools. Although the classical in-situ material testing can provide the key subsoil properties and consequently approximate the V_s profile with reasonable resolution using closely spaced boreholes (Kramer 1996), it would be rather expensive and time-consuming in order to amass the measurements covering the whole area under study. Therefore, the in-situ measurements are primarily recommended in projects of relevant importance.

In contrast, non-invasive techniques have long been recognized as functional and cost-efficient alternatives to obtain V_s profile. To a large extent, ground motion recordings on the ground surface comprise important subsoil characteristics. Thus, implementing different theoretical and numerical algorithms have been largely developed on the basis of wave

propagation techniques to determine the main subsoil characteristics. In this regard, surface-wave analysis, founded on inversion the dispersion curves, has been becoming more common and resulting in significant applications (Aki 1957; Capon 1969; Nolet and Panza 1976; Arai and Tokimatsu 2004, 2005; Castellaro and Mulargia 2009; Bard et al. 2010; Socco et al. 2010; Molnar et al. 2010). The approach consists of estimating the dispersive characteristics at a receiver by acquisition and processing of seismic data and then inverting the obtained data to estimate the subsoil properties. However, data processing and inversion of the experimental data using surface-wave methods are mainly computationally intensive in comparison to the invasive techniques besides results can be largely influenced by prior assumptions (Scherbaum et al. 2003; Molnar et al. 2010; Garofalo et al. 2016a, b). Moreover, surface-wave inversion is characterized as non-linear and ill-posed problem which can be strongly affected by solution non-uniqueness (Luke et al. 2003; Scherbaum et al. 2003; Foti et al. 2009; Teague and Cox 2016). Hence, several different V_s profiles can be found having an equally good agreement with the experimental model cause a high certain level of ambiguity in the final V_s structure and associated model parameters' uncertainties using dispersion curve inversion methods. Apart from this issue of non-uniqueness, mis-interpretation in the identification of dispersion curve modes (fundamental mode or higher modes) can bias the results (e.g., Forbriger 2003; O'Neill and Matsuoka 2005).

Arai and Tokimatsu (2004, 2005) and Herak (2008) successfully retrieve V_s model without engaging the dispersion curves by inverting the microseismic Horizontal-to-Vertical Spectral Ratio (HVSR). The HVSR technique also known as Nakamura's method is a well-known and practical alternative which reliably characterizes resonance phenomena (Nogoshi and Igarashi 1970; Lachet and Bard 1994; Mucciarelli 1998; Nakamura 2008). It has been well investigated that the shape of the HVSR amplification curves are reasonably speak for the subsoil characteristics and the fundamental frequency associated to the maximum amplitude of the HVSR curve can be a representative of the strong enough velocity contrast between soil layers (Nakamura 2000b; Bard and SESAME-Team 2005; Nakamura 2008). Hence, due to the adequate accuracy, easy and fast implementation of the HVSR technique, it has been attracting the attention of numerous researchers who applied the method to estimate the soil amplification properties (Konno and Ohmachi 1998; Mucciarelli and Gallipoli 2001; Di Giacomo et al. 2005; Sylvette et al. 2006; Herak 2008; D'Amico et al. 2008b; Bonnefoy-Claudet et al. 2009; Gallipoli and Mucciarelli 2009; Rahpeyma et al. 2016, 2017). Many researchers have investigated the reliability of using ambient noise, both numerically and experimentally, for

the quantifications of site effects (Bard 1998; Di Giacomo et al. 2005; Sylvette et al. 2006; D'Amico et al. 2008a; Pilz et al. 2009; Rahpeyma et al. 2016). It is noteworthy that in the absence of earthquake strong-motion recordings, microseismic data can be easily obtained and subsequently provides additional constraints and spatial resolution of site effects via HVSR method. The basic statement of the HVSR technique is that the vertical component of the ground motion in cases where the soil stratigraphy is flat and horizontal is supposed to be free of any kind of influence related to the site conditions at the recording site. Nevertheless, there are different physical interpretations for the fundamental concepts of the Nakamura's technique (Arai and Tokimatsu 2004; Parolai et al. 2005; Arai and Tokimatsu 2005; Picozzi and Albarello 2007; Herak 2008; D'Amico et al. 2008b; Sánchez-Sesma et al. 2011). The critical debate over the underlying theory of the Nakamura's method focused on this hypothesis that the obtained spectral ratio is chiefly determined by whether body-waves, approaching vertically the surface (Herak 2008; Nakamura 2008), or surface-waves, Rayleigh and Love waves with relevant upper modes (Arai and Tokimatsu 2004, 2005; Lunedei and Albarello 2010). The numerical comparison of different interpretations revealed that the surface-wave approach presents more reliable results for frequencies larger than fundamental resonance frequency of the sedimentary layer over the bedrock. Whereas, the body-wave approach provides more consistent results around the resonance (predominant) frequency (Albarello and Lunedei 2010). Despite a wide debate over different concepts, by and large the HVSR method is considered as a reliable and practical tool to obtain the V_s profile.

Another critical concern involved in estimation of the most-reliable V_s profile is implementing an appropriate inversion technique which can highly impact the consistency of the unknown model parameters in addition to the associated uncertainties. There is a large number of inversion procedures to invert the experimental data mainly based on minimizing misfit function considering an arbitrary level of best-fit model acceptance ratio (Rothman 1985; Yamanaka and Ishida 1996; Lai and Rix 1998; Parolai et al. 2007; Herak 2008; Socco and Boiero 2008). However, a basic concern with misfit-minimization approaches can be enlightened as ignoring quantitative uncertainty in the model parameter estimates which may cause optimistic assessments of predictive accuracy. In other words, the prior information for the model parameters are mainly neglected in the minimization. Furthermore, due to the non-linearity and non-uniqueness results of the inversion problems, true quantitative evaluation of confidence intervals (i.e., uncertainties) on V_s profile is problematic and of great importance. In order to overcome the deficiency of the classical misfit-minimization methods, Bayesian

inversion technique was implemented recently by Schevenels et al. (2008) to active-source dispersion data; however, due to over-emphasis on low misfit regions the final results are expected to be biased and not reliable (Schevenels et al. 2008). Afterwards, Molnar et al. (2009) applied Bayesian inversion of dispersion curves of microtremor array data to estimate the V_s profile where Bayesian technique found to be a reliable inversion technique to reliably infer soil properties (Molnar et al. 2010). In general, the Bayes theorem attempts to statistically update data and make inferences in the light of the observations (Gelman and Rubin 1992; Diggle et al. 1998; Congdon 2014; Gelman et al. 2014). The Bayesian methodology principally differs from the classical frequentist methods in that all of the unknown parameters in the underlying probability model are treated as random variables, in contrast to unknown constants.

In this study, therefore, we take the advantage of the Bayesian inversion approach to invert the experimental microseismic HVSR in order to approximate the most probable subsoil structures (a set of optimal subsoil V_s profiles) at nominated test stations across the areas with high level of seismicity. The test stations are located in the Izmir city, Turkey, and Mirandola, Italy, where the localized geological features are characterized as "soft soil". The theoretical HVSR technique considering body-wave approximations is applied to a benchmark/test site to verify the efficiency of the proposed approach. A Markov-Chain Monte Carlo (MCMC) method with Metropolis steps is considered to simulate the posterior distribution of the unknown model parameters. The convergence diagnostics are also applied to find the most reliable set of shear-wave velocity models for the nominated stations. Although there are still many sources of uncertainties, such as lack of good prior information and extremely simplifying assumptions about the ground structure, the results are in good agreement with available geological information. The findings of this project lead not only to reliable estimate of the subsoil physical parameters but also of their associated uncertainties, that are key factors in site classification guidelines, seismic design criteria, microzonation studies, seismic hazard analysis and provide significant information for many earthquake engineering applications.

2. Physical Theory

2.1. The Analytical Transfer Function for Body-waves Approximation

It is essential in non-linear Bayesian inversion to consider a theoretical model based on the available geological properties. In this study, the theoretical transfer functions of a set of horizontally stratified, linearly elastic layers overlaying a uniform half-space, excited by

vertically incident, transient plane waves is modeled based on the fast recursive algorithm proposed by (Tsai 1970) and modified later by Herak (2008) to consider frequency-dependent attenuation and body-wave dispersion (Herak 2008). The soil structure consists of any number of viscoelastic layers overlying a half-space and each layer is characterized by its thickness, velocity of the wave propagation, density, and elastic properties or Q-factor for seismic body wave which controls the anelastic properties. The amplification spectrum, $F_\kappa(f)$, can be calculated as Eq. (1)

$$F_\kappa(f) = \left\{\cos\left[2\pi f \frac{H}{\beta_{\kappa,a}(2\pi f)}\right] + i\frac{\rho_a\beta_{\kappa,a}(2\pi f)}{\rho_b\beta_{\kappa,b}(2\pi f)}\sin\left[2\pi f \frac{H}{\beta_{\kappa,a}(2\pi f)}\right]\right\}^{-1} \quad (1)$$

where f is the frequency, H is the thickness of the sedimentary layers over the half-space, ρ is the density, β_κ is the complex velocity of phase κ (P- or S-wave), and indexes a and b represent the sedimentary layer and bedrock, respectively (Albarello and Lunedei 2010). The complex velocity, β_κ, accounts for anelastic properties and can be defined as Eq. (2)

$$\beta_\kappa(2\pi f) = \frac{\beta_\kappa^e}{1 - \frac{1}{\pi Q_\kappa}\log\left(\frac{f}{f_{ref}}\right)}\left(1 + \frac{1}{2Q_\kappa}i\right) \quad (2)$$

where β_κ^e is the elastic velocity of the body-waves and f_{ref} is a reference frequency (in this study is considered as 1.0 Hz). It should be highlighted that the transfer function in Eq.(1) characterizes linear estimation of the amplification as it does not consider non-linear behavior of soil for large seismic vibrations (Herak 2008). Assuming that the microseismic vibrations are founded by vertically incident body waves. Consequently, P- and S-phase amplitudes control vertical and horizontal ground motion components, respectively. Moreover, if P- and S-phases have the same amplitude in the bedrock, HVSRs based on microseismic recorded on the surface are obtained by the respective amplifications of these phases induced by seismic properties of the soft sedimentary layers overlaying the rigid, stiff bedrock. In the light of the aforementioned assumptions, the HVSR at the surface can be determined as the ratio of the S-waves transfer functions (horizontal components), $F_S(f)$, to the P-waves transfer function (vertical component), $F_P(f)$, presented as Eq. (3) (Nakamura 2000b; Herak 2008; Albarello and Lunedei 2010).

$$HVSR(f) = \frac{|F_S(f)|}{|F_P(f)|} \tag{3}$$

Hence, on the basis of the theoretical amplification function presented in Eq. (3), the HVSR curves can uniquely defined by sets of soil properties characterizing each layer.

3. Bayesian Modeling Setup

3.1. Bayesian statistical inference

The Bayesian statistical methodology provides a robust statistical structure for making inference on different independent quantities of interest (i.e., model parameters, $\boldsymbol{\theta}$, or unobserved data, $\tilde{\boldsymbol{y}}$) in the light of observations using an underlying probability statements (Berger 2013; Congdon 2014; Gelman et al. 2014). In the context of the Bayesian framework, the unknown model parameters are assumed to be random variables and assigned prior probability distribution logically defined based on available information or a priori subjective beliefs. As such, the prior probability distributions of the unknown model parameters aid as probabilistic descriptions of what is known about the unknown parameters before data are collected and processed. However, it is noteworthy that prior information for model parameters is not always available and many times we only can consider a uniform distribution as the prior assumptions. Therefore, a Bayesian inversion approach formulates the non-linear inverse problem in terms of posterior probability distribution of the model parameters, $\pi(\boldsymbol{\theta}|\boldsymbol{y})$, which are considered random variables constrained by observed data, $\pi(\boldsymbol{y}|\boldsymbol{\theta})$, and prior knowledge about model parameters, $\pi(\boldsymbol{\theta})$. As can be seen in Eq. (4), inference on the posterior of model parameter vector given data contains the information about the sampling distribution, $\pi(\boldsymbol{y}|\boldsymbol{\theta})$, also known as likelihood function and assumptions about the prior information $\pi(\boldsymbol{\theta})$:

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{\pi(\boldsymbol{y}|\boldsymbol{\theta})\,\pi(\boldsymbol{\theta})}{\pi(\boldsymbol{y})} \tag{4}$$

where $\pi(\boldsymbol{y})$ denotes the marginal density function of the observation values and is independent of model parameters and can be obtained as $\pi(\boldsymbol{y}) = \int \pi(\boldsymbol{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\theta$. In this vein, the knowledge about model parameters will be updated by conditioning the underlying probability model on the observed data and prior information as it is shown in Eq. (5). Since the marginal data density function $\pi(\boldsymbol{y})$ is independent of parameters $\boldsymbol{\theta}$, the posterior distribution of the

model parameters can be acquired through employing the simplified Bayesian theorem as presented in Eq. (5):

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) \propto \pi(\boldsymbol{\theta})\pi(\boldsymbol{y}|\boldsymbol{\theta}) \tag{5}$$

where, $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ is the joint posterior distribution of the model parameters, $\boldsymbol{\theta}$, given the derived theoretical transfer function of the subsoil as data $\boldsymbol{y}$ requires information about the sampling distribution $\pi(\boldsymbol{y}|\boldsymbol{\theta})$ and also a sensible assumption about the prior distribution $\pi(\boldsymbol{\theta})$ if exists. In this regard, the obtained posterior distribution integrates updated knowledge about the unknown parameters considering knowledge found from the observed data.

In statistical modeling in order to numerically approximate posterior density function of model parameters, Markov Chain Monte Carlo (MCMC) simulation is known as one of the most practical and common simulation tools (Gelman and Rubin 1992; Smith and Roberts 1993; Gilks 2005). The MCMC is basically applicable to almost any Bayesian modeling and includes general algorithm for simulating independent Markov chains which has a desired target density. This procedure is mainly carried out using the Gibbs sampling framework (Geman and Geman 1984; Casella and George 1992) and the Metropolis algorithm (Metropolis et al. 1953). The Gibbs sampler, also called alternating conditional sampling, is a randomized algorithm that mainly applied to obtain a sequence of samples which are approximated from a specified probability distribution. Theoretically, a Markov chain is a sequence of random variables $\theta^1, \theta^2, \ldots, \theta^t$ (for any t) such that the current iteration depends only on the most recent (latest) iteration i.e., θ^{t-1}. This technique results in a set of Markov chains including samples gathered in every iteration, that can be shown to converge to the target posterior distribution $\pi(\boldsymbol{\theta}|\boldsymbol{y})$.

Within the MCMC, the Metropolis algorithm can be applied as an updating strategy which tracks adaptation of a random symmetric walk in parameters space to define the acceptance or rejection of the samples to converge to the specified target distribution. One of the main benefits of Metropolis algorithm is that the algorithm converges to the target density for different proposal density functions. The following steps summarize the applied MCMC in this study to sample the posterior density of model parameters:

1. Initialize the MCMC process with preliminary estimates of parameters $\boldsymbol{\theta}^0$.

2. At step *k* sample a proposal value $\boldsymbol{\theta}^*$ from a given proposal density (e.g., in this study normal distribution is used) with mean value $\boldsymbol{\theta}^{k-1}$ and predefined fixed variance.

3. Calculate the ratio

$$\alpha = \min\left\{1, \frac{\pi(\boldsymbol{\theta}^* \quad |\boldsymbol{y})}{\pi(\boldsymbol{\theta}^{k-1}|\boldsymbol{y})}\right\} \tag{6}$$

4. Sample u_k from uniform density on [0,1]. Accept or reject the proposed values of model parameters according to:

$$\boldsymbol{\theta}^k = \begin{cases} \boldsymbol{\theta}^{k-1} & if \quad \alpha \leq u_k \\ \boldsymbol{\theta}^* & if \quad \alpha > u_k \end{cases} \tag{7}$$

In other words, if the state transition leads to higher probability value than the previous state, the proposed value is accepted, but if the transition produces a lower probability, then the proposed value is only accepted with a probability of α. Upon rejection, the simulated value in the current.

5. Run the MCMC algorithm with the updated estimates and repeat steps 2-4.

6. Obtain posterior summaries for model parameters $\boldsymbol{\theta}$ using their posterior samples.

3.2. Bayesian convergence diagnostics

Although the MCMC approach using Metropolis steps guarantees convergence to the objective posterior density to a large extent, the rate of convergence and degree of dependency between successive samples can notably affect the convergence diagnostics. In general, fast convergence and low dependency between successive samples yield higher quality of the MCMC parallel chains. First of all, a rational practice influencing convergence to an unbiased estimate include deleting early samples of the Markov chain, commonly referred to as "burn-in" (a burn-in length of at least 25% of total samples is applied here). We run many sets of combination with different number of layers, various prior assumptions and start points, number of iterations, number of chains, and burn-in sample size to find the most consistence results. All chains would be analyzed together after simulating the desired number of iterations by removing burn-in part.

Several convergence diagnostics need to be derived from the simulated samples using Markov chains to evaluate the chain's convergence trends. Operationally, actual convergence of MCMC simulation has been obtained when inference of variables do not depend on the starting point of simulations (Brooks and Gelman 1998). For this purpose, in this study, we used three different convergence assessment methods: First visually, since the convergence of the Gibbs sampler is monitored by inspecting the trace plots of the MCMC simulations for each model parameter, and the behavior of the posterior sample chains can expose bad mixing of chains or chaotic behavior of separate chains. Secondly, we used Gelman-Rubin statistics to assess the convergence of iterative MCMC simulations. The confidence limits are based on the assumption that the stationary distribution of the variable under examination is normal and approximate convergence is diagnosed when the upper limit is close to 1.00. Hence, we set Gelman-Rubin statistic in the range of $\sim 1.00 - 1.10$ in order to find the most applicable length of sampling and fast convergence to a steady state. Large value of the Gelman-Rubin statistic, which expresses the potential scale reduction factor and typically is greater than 1.10, indicates that simulations have not converged to the target density. This is important because it is critical to know how many iterations are essential to achieve the desired level of accuracy, due to the Gibbs sampler being extremely computationally demanding, even for a relatively small-scale statistical problem. Finally, the accuracy of the statistics based on the simulated posterior samples and the dependence between successive samples of the Markov chain was estimated from the autocorrelation function. The obtained autocorrelations at each lag can be displayed graphically. If the level of autocorrelation is very high, then the parameter will be a poor diagnostic for convergence. In this way we generate stable parallel Markov chains and ensure efficient sampling in the final stage for posterior inference to yield well-defined marginal distributions of the model parameters.

3.3. Implementation

The MCMC method was chosen owing to the problem's nonlinearity, which is the best accounted for by multiple parallel Markov chains. The parametrization of the theoretical HVSR is based on the assumptions of 2-D layered models consisting of a stack of homogenous linear elastic layers over a half space. Subsoil physical properties such as thickness (H), density (ρ), shear-wave velocity (V_s), compressional velocity (V_p), and elastic properties for S- and P-waves (Q_s and Q_p) are considered as model parameters for the theoretical HVSR. Trial inversions for all the unknown model parameters together reveled that due to non-uniqueness results and

considerable trade-off between parameters the obtained results are not informative. Then some parameters are gradually fixed to assess their influences. We observed that thickness and shear-wave velocity are the most influential variables among the all model parameters which are highly correlated. It has been also proven that the theoretical transfer function chiefly depends on shear-wave velocity and depth of the subsoil and negligibly on the other soil properties (Foti et al. 2009; Molnar et al. 2010). Hence in this study, $\boldsymbol{\theta} = (H, V_s)$, are assumed to be unknown and the rest of parameters are defined as known and fixed parameters which can be approximated on the basis of available geological information at the nominated stations.

As can be seen in Eq. (8), at each iteration k of the MCMC simulation process, the unknown variables, θs, are drawn, as input to calculate the theoretical HVSR, from a normal distribution centered at an adaptive mean value, $\mu_{\theta_{p,l}}$, and pre-defined fixed standard deviation, $\sigma_{\theta_{p,l}}$, for each layer and all chains.

$$\theta_{p,l}^{k} \sim \mathrm{N}\left(\mu_{\theta_{p,l}}, \sigma_{\theta_{p,l}}^{2}\right) \tag{8}$$

where subscripts $p = 1,2$ and $l = 1, \dots, L$ indicate the model parameters indicator $\boldsymbol{\theta} = (H, V_s)$ and the layer, respectively. The mean value of the model parameters distribution is conditioned on the mean value of the latest iteration i.e., ($\mu_{\theta_{p,l}} \sim \theta_{p,l}^{k-1}$) and the random walk (standard deviation) is defined based on 5% of the mean value of the model parameters. At each layer, the lower $\left(\theta_{p,L}\right)$ and upper $\left(\theta_{p,U}\right)$ bounds for each model parameter and each layer are chosen reasonably in order to limit parameters' space ($\theta_{p,L} < \theta_{p,l} < \theta_{p,U}$) to avoid the inversion stick into a wrong convergence track due to the trade-off between model parameters; however, the boundaries should be wide enough to allow the data to determine the model parameters. The initial values for the model parameters, which are used to produce the initial theoretical HVSR, are instinctive approximations. It is noteworthy that the proposed methodology will be more accurate if we have prior knowledge (e.g. borehole data or in-situ material testing results) about the shallow subsoil stratigraphy.

As can be seen in Eq. (5), in order to setup the Bayesian framework we need to define the prior probabilities for all model parameters and a likelihood function. Due to the lack of precise information particularly on the deeper layers, the prior probability density function of each parameter is chosen as a uniform probability density function for each model parameter on bounded intervals $\theta_{p,L} \leq \theta_{p,l} \leq \theta_{p,U}$, such that

$$\pi(\theta_{p,l}) = \begin{cases} 1 & \theta_{p,L} \le \theta_{p,l} \le \theta_{p,U} \\ 0 & otherwise \end{cases} \quad (9)$$

The join prior probability density function for all the model parameters is the product of the individual prior densities as Eq. (10).

$$\pi(\boldsymbol{\theta}) = \prod_{p=1}^{P} \pi(\theta_{p,l}) \quad (10)$$

It is assumed that the probability density function of the spectral amplitudes $HVSR_i$ (i.e., y_i) in each frequency bin f_i with $i = 1, \dots, N_f$ (N_f is the total number of frequency bins) is lognormal with the expected value $\mu_{HV}(f_i)$ and the variance $\sigma_{HV}^2(f_i)$ of $\log(HVSR_i^{obs})$. Consequently, the probability density function for each y_i at frequency bin i is given by:

$$\pi(y_i|\theta_i) \sim \mathrm{LN}\left(y_i \mid \mu_{HV}(f_i), \sigma_{HV}^2(f_i)\right) \quad (11)$$

In order to assess the efficiency of our proposed method to estimate the V_s profile and associated uncertainties, benchmark/test stations preferably with available geological information are required. In this study, Bayesian inversion of microseismic HVSR is applied at two stations in Turkey and Italy. Both stations are characterized as soft sedimentary layers overlaying the bedrock. The subsoil geological structure is investigated for both nominated stations.

4. Application to Izmir, Turkey

4.1. Location and geological setting

The Aegean region in western extremities of Turkey is known as one of the seven geographical regions of Turkey and one of the most seismically active region of the Eastern Mediterranean region. Izmir, the capital of the Aegean region of Turkey, is known as one of the most populated (more than 3.5 million inhabitants) with dense industrial infrastructures in the country. Essentially, seismic hazard analysis across Izmir region is of paramount importance due to large and growing population and infrastructures which are surrounded by active faults. Furthermore, as can be seen in Figure 1, the majority of the settlements (industrial and populated areas) are collocated on top of Quaternary alluvial deposits around the Gulf of Izmir. Other prominent units are Miocene-aged sandstones, mudstones, andesitic volcanic, and Paleocene limestones (Polat et al. 2009, 2012; Gok and Polat 2014). Hence, due to the

unconsolidated deposits in the Izmir basin, a significant site effect as a result of strong changes in seismic wave propagation and consequently ground motion amplitude is largely expected.

In 2008, a small aperture local seismic network, IzmirNet, consists of 16 stations was established in order to enhance the understanding of the seismic wave propagation across the region and also provide imperative information for various earthquake engineering applications (Polat et al. 2009). Borehole drilling, microseismic analysis, electrical resistivity, and seismic refraction techniques, as well as multichannel analysis of surface waves (MASW) measurements in the IzmirNet stations were performed in order to determine the geological, geophysical, and geotechnical site characterization (Gok 2011; Polat et al. 2012; Eskişar et al. 2013; Gok et al. 2014; Uzel and Özkaymak 2014). The acceleration recordings were mainly installed at sites with thick Quaternary Neogene formations (see Figure 1). Borehole drilling were performed up to 30 meters in IzmirNet stations in order to determine the geological information. They are in Bornova basin at the east, the Karsiyaka-Mavisehir area at the north, and between Balcova and Urla at the south. The Quaternary sediments and cretaceous flysch are the main units near the Balcova area, with a sedimentary fill up to 180 m. Figure 1 shows the distribution of IzmirNet stations across the Izmir Gulf (Polat et al. 2009; Gok and Polat 2014).

4.2. Microseismic data collection and processing

The continuous microseismic recordings of a minimum 30-minute duration and sampled at 100 sps (samples per second) were recorded at IzmirNet stations. The microseismic measurements across IzmirNet stations were carried out during late March and early April 2008. All stations are free-field and equipped with three-component CMG-5TD accelerographs (Guralp Systems, Reading, UK) with CMG-5T force balance accelerometer and built-in 24-bit AD converter for data acquisition. An asymmetric digital subscriber line (ADSL) system controls the stations and downloads real-time continuous data. The IzmirNet stations use the ADSL technology to download and store data in the central processing laboratories in Izmir and Ankara. For each station, a minimum 15 minutes of microseismic measurements were recorded at the sampling rate of 100 Hz. The data processing to obtain the HVSR at each site was performed using GEOPSY software in the following routine: the data was filtered between 0.20 and 25 Hz by a band-pass 4 poles Butterworth filter after the mean and a linear trend were removed; then each component of the recorded signal was windowed in a time series of 30 second length without overlapping; use cosine taper 5% and for each time window Fast Fourier

Transform (FFT) was calculated and smoothed using Konno and Ohmachi logarithmic window function, $B = 20$ (Konno and Ohmachi 1998). For each time window the spectral ratio between the root-mean square average spectrums of two horizontal components over the spectrum of the vertical component was calculated and, finally, the station representative average HVSR and the standard deviation were computed and plotted as a function of frequency. This approach is also the recommended method by SESAME (Bard and SESAME-Team 2005) and is the most commonly used method for HVSR analyses. It is worth mentioning that we examined the sensitivity and stability of the HVSR results for several effective input parameters. For instance, the smoothing coefficient, different length of time windows, and different combinations of horizontal components (e.g. squared average, total horizontal energy, and directional energy).

The obtained microseismic HVSR results over IzmirNet local accelerometers network in the current study are found to be in good agreement with local site effect investigations using strong motion earthquake (Polat et al. 2012). The analysis revealed that the majority of the stations (e.g., BYN, CMD, KSK, MNV, MVS, and URL) explicitly produce a fundamental peak at low frequencies in the range of $0.4 - 0.8$ Hz while the other stations (e.g., BRN, BUC, BYR, GZL, YMN, and YSL) have a broad peak at frequencies larger than 1.0 Hz (e.g., YMN). On the other hand, some stations exhibit bimodal amplification curves with one mode being more dominant and of relatively larger amplitude than the other (e.g., BLC, BOS, KON). These results indicate that significant site response variation exists on a spatial scale across the IzmirNet network. In this study, we selected station BYN located on soft soil in the eastern part of Izmir Bay with a clear fundamental frequency peak at $0.7 - 0.8$. As can be seen in the insert figure at bottom left in Figure 1, the HVSR curve at station BYN has a predominant frequency at ~ 0.76 Hz. The HVSR characteristics for BYN station suggests that the soil column acts as a single layer on top of a high impedance contrast between layers, where strong amplification and frequency dependent resonance are known to occur. To apply the MCMC approach we should remember that Bayesian inversion approach essentially represents data uncertainty distribution into parameter uncertainty distributions. In this study, the convergence of thickness and shear-wave velocity infers concurrently to achieve better understanding of exist inter-correlation as well as uncertainties of the model parameters.

4.3. Bayesian inversion results

In order to generate stable parallel Markov chains and ensure the efficient sampling in the final stage for posterior inference of the model parameters we analyze different configurations for the MCMC setup (e.g., number of parallel chains, length of sampling, burn in period, and etc.). The sensitivity analysis and convergence diagnostics revealed that due to ill-posed and non-linear inversion problem as well as non-uniqueness results the number of the parallel chains and sampling iterations need to be increased. We investigated that increasing the number of chains provide better inference of the model parameters due to considering different sets of values and diminishing the results of non-uniqueness result. A grid search of the MCMC initiated with a starting model whose parameters are randomly perturbed within the wide enough bounds defined $[\theta_L, \theta_U]$ results in posterior probability distribution of the model parameters. To avoid sticking into localized minima besides the non-uniqueness effects which can lead to unreliable V_s profile estimation the boundary should be wide enough and avoids unrealistic limitations.

On the other hand, in the context of microseismic inversion problems, adopting enough parameters (e.g., number of layers, fixed or dynamic model parameters) to obtain the subsoil structure is known to be an influential parameter. Essentially, implementing not enough parameters can conspicuously result in under-fitting the data, biasing parameter estimation and under-estimating the associated uncertainties. In contrast, considering too many parameters can over-fit the data result in under-determined parameters and excessive variations of parameters (Molnar et al. 2010). Hence, in order to find the most probable parametrization with the most reliable resolution of the subsoil structure we propose a practical strategy of initializing the inversion process with the simplest (e.g., a single-layer) subsoil structure overlaying the bedrock and then gradually increase the number of stack layers over the half space and ensure the informative posterior distribution as well the associated uncertainties.

Figure 2 represents the visual inspection and convergence diagnostics of the $N_c = 20$ sampled chains of the length of $N_L = 20{,}000$ with $N_B = 5{,}000$ burn-in samples for the subsoil properties of a single sedimentary layer overlying the half-space (bedrock) obtained at BYN station. In general, fast convergence and low dependency between successive samples yield higher quality of the MCMC parallel chains. The visual inspection of trace plots (Figure 2) shows good mixing in successive samples to estimate the model parameters. The posterior histograms of posterior samples after burn-in period represent unimodal posterior distribution

for thickness and shear-wave velocity. The Gelman-Rubin diagnostic ($\hat{R}$) or the potential scale reduction is in less than 1.10; however, the degree of autocorrelation for non-thinned simulations is around 50% at lag 50 for the model parameters. Thinning (i.e. discard all but every 5^{th} samples with the goal of reducing autocorrelation) the posterior samples is one the techniques which is recommended to decrease the autocorrelation and also speed up the posterior analysis. As can be seen in Figure 2c, thinning by factor 5 reduces the autocorrelation (red bars), it is not very recommended due to reducing precision of MCMC approximations (Link and Eaton 2012).

Figure 3a shows the observed (cf. solid black spectrum) with associated standard deviation (cf. dashed red spectrum), the initial theoretical (cf. solid blue spectrum), and the posterior (cf. solid pink spectrum) HVSR amplification curves. The clear peak in the observed HVSR at the nominated station (BYN) is as indicator of a strong velocity contrast within depth which is linked to the fundamental mode of the subsoil structure. Theoretically, we can assume the simplest subsoil structure with a single predominant frequency as a sedimentary column sits on top of a hard layer (bedrock). Figure 3b shows the initial and the posterior V_s profile (blue and pink, respectively). We note that the inversion of the observed HVSR is defined over the nominated range of frequencies around the fundamental frequency (gray area shown in Figure 3a). It may be argued that at relatively high frequency, no HVSR peak associated to a shallow stratigraphic horizon can be observed; thus, the basic requirement for the proposed procedure to concentrate around the fundamental mode could be satisfied. Furthermore, in the seismic microzonation practice, attention has generally only been paid to the main resonance frequency, which is the largest HVSR peak, while other stable humps and troughs in the curve were not considered. The marginal probability distribution for individual parameters (H and V_S) is shown in Figure 3c to highlight the trade-off between model parameters. As can be seen in Figure 3b, the initial model is defined as a $\sim 260\ m$ sedimentary layer with shear-wave velocity of $\sim 600\ m/s^2$ overlaying the half-space that produces the theoretical fundamental frequency approximately equal to 0.5 Hz. However, according to the localized geological/geostructural data, BYN station sits on the alluvial deposit layer of around $180 - 200\ m$ depth (Polat et al. 2009; Gok et al. 2014). Explicitly, the posterior mean of the V_s profile for a single-layer model of sedimentary layer over the half-space is estimated around 200 meter with standard deviation of $\sim 50\ m$. The results indicate that the uncertainty associated to the thickness is higher than the shear-wave velocity.

Although with a single-layer subsoil structure we could quantitatively estimate model parameters' characteristics, the key question is that to what extent a detailed V_s profile can be extracted from the recorded data? As can be seen in Figure 4, a group of different subsoil structures at station BYN can result in the same HVSR (both fundamental frequency and amplitude) due to the possible trade-offs between model parameters and non-uniqueness results. Hence, in this study, with the aim to determine the most probable subsoil parametrization with reasonable resolution, initially, a single-layer model is considered and through gradually increasing the number of stack layers over the half-space the posterior distributions of the model parameters acknowledge to what extent the current subsoil structure is informative (i.e., a uniform posterior distribution of model parameter is non-informative). We conduct a blind test over the number of layers to consistently investigate the best resolution of model parametrization. Therefore, we continue adding the number of layers as far as the posterior distributions of model parameters do not provide any consistent information.

The effect of adding the number of layers is shown in Figure 5 in which the marginal posterior probability distributions of a five-layer subsoil model over a uniform underlying half-space for model parameters $\boldsymbol{\theta} = (H, V_S)$ are illustrated. We used a Gibbs sampler with Metropolis steps for total iteration of $N_L = 20{,}000$ through $N_c = 20$ parallel chains, and $N_B = 5{,}000$ burn-in samples which in total creates $300{,}000$ random samples (i.e., theoretical HVSR) across the nominated range of frequencies. The convergence of the obtained model parameters is evaluated by controlling the traceplots of the Markov chains, calculating the Gelman–Rubin ($\hat{R}$) statistic for all parameters. It should be noted that due to the predefined boundary for model parameters the autocorrelation is naturally higher than what we expected. On a close scrutiny, the results of posterior histograms of thickness and shear-wave velocity at each layer explicitly highlight that increasing the number of layers leads to less informative posterior distribution with large uncertainties for deeper layers comparing to the shallower layers. The initial HVSR model (blue HVSR curve) and the final HVSR model (pink HVSR curve) obtained based on the posterior mean values of the subsoil properties in addition to the associated V_s profiles are shown in Figure 5a and Figure 5b, respectively. However, considering thinning factor of 5 can help to decrease the autocorrelation. Nonetheless, although the initial V_s profile and HVSR model behaves differently, the posterior model fits the observed HVSR very well and the estimated depth of the sedimentary layer over the half-space is around $180 - 190\ m$. Figure 5c illustrates the matric correlation and evaluate the correlation between posterior model parameters. Contrary to the relatively high correlation for a single-layer subsoil structure (cf.

$\sim 80\%$ in Figure 5c), the correlation between model parameters of a multilevel subsoil structure is not large ($\sim \pm\ 30\ -\ 40\%$). The negative and positive correlation can be observed between thicknesses and shear-wave velocities of layers.

Figure 6 highlights the correlation between model parameters in each layer using the normalized joint marginal probability distribution for a five-layer sub-soil structure overlaying the half-space. A more detailed look at Figure 6 shows that correlation between the posterior values of the thickness parameters of different layers are negative while there is mainly positive correlation between the posterior values of the shear-wave velocity of the layers. Comparing the trade-off between model parameters of a single-layer model and a multi-level model indicate that although increasing the number of layers can improve the resolution, it can lead to increasing the uncertainties and confidence region of the model parameters.

5. Application to Mirandola, Italy

5.1.Location and geological setting

Mirandola city sits on the Po river plain in North Italy. As can be seen Figure 7, the subsoil structure of the region is mainly characterized by layers of alluvial deposits with sandy horizons and silty-clayey layers overlaying a stiff rock layer of marine and transitional deposits of lower-middle Pleistocene age (Garofalo et al. 2016a, b; Cox and Teague 2016; Tarabusi and Caputo 2017). Furthermore, historically the region is affected by moderate to strong earthquakes. As can be seen in Figure 7, the city is located near the epicenter of the recent Emilia seismic sequence in 2012 (Anzidei et al. 2012). Due to industrial districts in addition to the growing fault-propagation anticline close to the Mirandola city and its surroundings, different site effect investigations and microzonation studies have been done in the region. (see Figure 7). Mirandola station was a part of the InterPACIFIC (Intercomparison of methods for site parameter and velocity profile characterization) project recently with the aim of comparing non-invasive technique with invasive ones at three European stations with different site characterization. In addition to the surface waves analysis several invasive tests (e.g. seismic boreholes and cross-holes) were performed across the Mirandola to provide a reliable benchmark of direct V_s and V_p measurements. The invasive tests indicate that V_s profile is rather plain and smooth with no abrupt velocity contrasts until the top soft sedimentary layer is reached the stiff soil at a depth between 110 and 120 m (Griffiths et al. 2016; Cox and Teague 2016).

5.2. Microseismic data collection and processing

The continuous microseismic recordings of a minimum 60-minute duration were collected at Mirandola station (MIR). The noise measurements were made using different geometric arrays (triangle, L-shape, and circle shape). All passive measurements were made with Güralp broadband CM6TD seismometers with integrated digitizer. The signals at three components were sampled at 200 Hz sampling frequency with continues GPS synchronization. The microseismic measurements across MIR station were carried out during September 2013. The same data processing routine as previously described for BYN station in Turkey was implemented for MIR station. The Mirandola station is located on flat ground and it is rather quiet and at the boarder of a residential area, without noisy facilities. As can be seen in the insert figure on right top of Figure 7, there is a clear and sharp peak at fundamental frequency at $\sim 0.7 - 0.8$ Hz with the amplification of 4.46 which indicates there is a velocity contrast within depth. It should be noted that the maximum H/V ration from the three-component ambient data reported by other site effect studies at MIR station (e.g., (Griffiths et al. 2016; Griffiths Shawn C. et al. 2016)) found to occur at 0.73 Hz with a standard deviation of 0.03 Hz.

5.3. Bayesian inversion results

The sharp peak of the HVSR curve at MIR station can be modeled by a column of soft soil sits over the hard rock. Sensitivity analysis revealed that due to more complicated subsoil structure of the MIR station comparing to the BYN station more synthetic samples are required. Hence, we used minimum $N_C = 30$ chains of a total length of $N_L = 30{,}000$ iterations by considering $N_B = 5{,}000$ samples as burn-in period for different model sets. A grid search of adapting Monte Carlo is initiated with a starting model whose parameters are randomly perturbed within the bounds defined $[\theta_L\ , \theta_U]$ for unknown model parameters and the number of layers are gradually increasing to capture more details within depth. The initial results from a single-layer soil structure in addition to the convergence diagnostics results for MIR station is shown in Figure 8. As can be seen in Figure 8 the posterior distributions of the model parameters are well determined with unimodal distribution and the convergence diagnostics emphasize on the convergence of the model parameters.

Figure 9a and Figure 9b represent the initial and the obtained HVSR and subsoil V_s structure using mean posterior of model parameters (represented in Figure 8) and Figure 9c highlights the strong trade-off between model parameters using the marginal probability distribution for individual parameters (H and V_s). Although the available geological

information and previous studies at the MIR station indicated that the subsoil structure consists of sedimentary layer with the depth of $50 - 150\ m$ we set the initial model at $200\ m$ to control the reliability of the results. As can be seen in Figure 9, although the trade of between model parameters is very large, the obtained V_s profile approximate the expected depth of sedimentary layer overlaying the uniform bed-rock.

We gradually increase the number of layers as far as the posterior distributions of model parameters do not provide any consistent information for the layers. Figure 10 compares the initial and final posterior HVSR and the associated V_s profiles for a three-layer subsoil structure. Figure 10a and Figure 10b explicitly indicate that contrary to the very ambiguous and different initial subsoil structure of a complex V_s profile with a velocity reversals, the obtained V_s profile is in a very good agreement with the findings of the other studies (Garofalo et al. 2016a, b). The obtained V_s profile shows that there is a sedimentary layer with approximately $\sim 120\ m$ depth and monotonically increasing shear-wave velocity within depth. The correlation between model parameters is shown in Figure 10c. Furthermore, the less informative posterior distributions of the deeper layers are highlighted at thickness and shear-wave velocity posterior probability distribution.

Figure 11 shows the normalized joint marginal probability distribution between and within the model parameters for a three-layer subsoil structure overlaying the half-space which clarifies the trade-off between model parameters at different layers. A more detailed look at Figure 11 shows that inter-parameters correlations of the thickness mainly have negative correlation while the inter-parameters correlations of the shear-wave velocity have positive correlation. The trade-offs between model parameters are inherent in the physics of inversion problem and cannot be solved either by data processing or applying different inversion strategies. We found that by adding more reliable information and limiting the model parameters the convergence and consequently retrieve posterior samples are more accurate.

6. Discussion and Conclusion

This study applies a non-linear Bayesian inversion framework in order to retrieve the most-probable subsoil V_s structure and the associated uncertainties to the unknown model parameters using microseismic HVSR. In this framework, we invert the experimental microseismic HVSR for most-probable shear-wave velocity profile estimation that can be obtained from conventional measurements using only a three-component receiver. Considering Bayesian

inversion technique, we determined posterior probability distribution of the subsoil geophysical parameters as well as the associated uncertainties intervals. We applied MCMC technique with Metropolis sampling steps to systematically search the model parameters space. Furthermore, we explored a reliable resolution of the subsoil layering structure by progressively increasing the number of subsoil layers can reliably infer the posterior probability distribution of model parameters. The validation and convergence of the posterior samples were fully examined using trace plots visualizing, Gelamn-Rubin statistics, and the autocorrelation plots. The proposed Bayesian inversion of microseismic HVSR was applied at two reference stations in IzmirNet local accelerometric array, Turkey, and Mirandola station, Italy, to investigate the ability of the proposed scheme to recover an accurate V_s profile in relatively deep geological settings.

According to the available geological data (Sözbilir et al. 2008; Uzel and Özkaymak 2014), the base of the basin in the investigation region is formed by Bornova Melange (i.e., quartzous fly-ash). The site effect investigations across IzmirNet clearly show there is a strong velocity contrast within the depth of approximately $200\,m$. However, the variation of the model parameters has not been statistically investigated. In this study, a well-resolved velocity profile to approximately $180 - 200\,m$ is determined at the BYN station in Turkey considering different layering. On the other hand, the site effect characteristics, main geological features, and subsoil structure of the Mirandola station (MIR) has been analyzed in great details since the station was nominated for the InterPACIFIC project (see Garofalo et al. 2016a, b) in which, 14 independent teams of experts analyzed the experimental surface-wave datasets in addition to a set of quality control tests using the borehole data. Furthermore, Griffiths et al. (2016) and Griffiths Shawn C. et al. (2016) used experimentally-measured dispersion data was used to investigate common methods of 466 accounting for uncertainty in V_s profiles at sites in Mirandola. Comparing the results of the proposed Bayesian inversion in this study with the other studies confirms the reliability of the proposed method and show a very good agreement and the obtained velocity profile resolved a sedimentary layer of approximately $110\,m$ overlying the bed-rock in addition to the model parameters' uncertainties at each layer which is not accurately reported by previous studies (e.g., Garofalo et al. 2016a, b).

Shapiro (1996) showed that the V_s profiles determined from the classical surface-wave inversion schemes (e.g., damped least-square algorithms) are too restrictive and uncertainties are not correctly estimated. Moreover, the results are very sensitive to the selection of the initial

model and the inversion process can easily be biased by wrong choices in terms of model parameterization that lead the solution into local minima (Sambridge 2001; Luke et al. 2003; Wathelet et al. 2004; Herak 2008). Although we believe that a reliable prior information on the model parameters can conspicuously enhance the V_s profile estimate, as can be seen in Figure 5a-b and Figure 10a-b, the results of the Bayesian inversion proposed in this study for BYN station and MIR station, respectively, indicate that although we chose a not very realistic initial model, the posterior V_s profiles could reasonably project the true and expected V_s profile for the nominated stations.

One of the main obstacles embedded in any V_s profile inversion strategies is non-uniqueness solutions that is mainly related to the ill-posed mathematical formulation. In other words, the HVSR inversion process suffers from trade-off between the shear-wave velocity and layer thickness and therefore may not be very reliable in terms of absolute velocity-depth values (Scherbaum et al. 2003). In general, the higher uncertainty for deeper layers reflects the non-unique nature of the inverse problem. Non-uniqueness and errors in the identification of propagation modes are probably the main cause for doubts in the ability of surface-wave methods to recover a realistic model of the site. Many authors have suggested different strategies for the inversion of surface-wave data. Some authors applied stochastic methods (e.g., Beaty et al. 2002; Dal Moro et al. 2007; Socco and Boiero 2008 among the others), some adopted deterministic methods like the least- squares algorithm (e.g., Lai and Rix 1998; Xia et al. 1999; Socco et al. 2009), while others focused on multi-mode, effective mode or joint inversion of Rayleigh and Love wave dispersion curves for better constraining the solution (e.g., Xia et al. 2003; Parolai et al. 2005; Arai and Tokimatsu 2005; Picozzi et al. 2009; Dal Moro and Ferigo 2011; Poggi et al. 2012). The findings of this project, however, shows that increasing the number of parallel chains in the proposed Bayesian inversion can reasonably results in better convergence due to expansion of the input parameters combinations and considering as much as possible potential synthetic HVSR models. In this case, within the MCMC grid search all perturbations start around different initial parameters values. Hence, the determined V_s profiles with this approach and the associated uncertainties would be reliable.

In order to verify the efficiency of the inversion results is finding the predominant frequency from posterior simulations of model parameters we calculate the natural frequency of the soil, f_n, using the harmonic average defined in Eq. (12):

$$f_n = {(2n-1)} \Big/ {4\sum_{i=1}^{N_L}\left(\frac{H_i}{V_{s_i}}\right)} \tag{12}$$

where n is the mode number, H_i is the thickness and V_{s_i} is the shear-wave velocity for the i^{th} soil layer, and N_L refers to the total number of layers overlaying the half-space. The estimated mean posterior predominant frequency (f_0) for a five-layer subsoil structure over half-space estimated in Figure 5 for station BYN is 0.63 Hz with the posterior 95% confidence interval of $[0.58 - 0.72]$ Hz which is in approximately good agreement with the experimental fundamental frequency of 0.70 Hz. The posterior mean and 95% confidence intervals for a three-layer subsoil structure over half-space presented in Figure 10 for station MIR can be approximated as 0.60 Hz and $[0.52 - 0.65]$ Hz, respectively.

As this is a non-linear inversion problem, we suggest that using the Bayesian inference by MCMC algorithms for posterior simulation can point out important inter-parameter relations (trade-off) and also irregular non-Gaussian distributions which would otherwise lead to faulty conclusions when treated through simple linear regressions. The findings of this project lead not only to reliable estimates of the subsoil physical parameters but also of their associated uncertainties, that are key factors in site classification guidelines, seismic design criteria, microzonation studies, and seismic hazard analysis. Furthermore, this approach can be considered as a reliable estimate to compare with the other procedure such as InterPACIFIC project.

Acknowledgments

This work was supported by the Icelandic Centre for Research (Grant of Excellence no. 141261-051/52/53), the University of Iceland Research Fund, and a Doctoral grant from Eimskip Fund of the University of Iceland.

References

Aki K (1957) Space and Time Spectra of Stationary Stochastic Waves, with Special Reference to Microtremors. Bulletin of Earthquake Research Institute 35:415-456.

Albarello D, Lunedei E (2010) Alternative interpretations of horizontal to vertical spectral ratios of ambient vibrations: new insights from theoretical modeling. Bulletin of Earthquake Engineering 8:519–534

Anzidei M, Maramai A, Montone P (2012) Preface. Annals of Geophysics 55.

Arai H, Tokimatsu K (2004) S-Wave Velocity Profiling by Inversion of Microtremor H/V Spectrum. Bulletin of the Seismological Society of America 94:53–63.

Arai H, Tokimatsu K (2005) S-Wave Velocity Profiling by Joint Inversion of Microtremor Dispersion Curve and Horizontal-to-Vertical (H/V) Spectrum. Bulletin of the Seismological Society of America 95:1766–1778.

ASCE (2007) Seismic rehabilitation of existing buildings. Reston, VA

ASCE (2010) Minimum Design Loads for Buildings and Other Structures. Reston, VA

Bard PY (1998) Microtremor measurements: A tool for site effect estimation? In: Proceeding of the Second International Symposium on the Effects of Surface Geology on Seismic Motion. Yokohama, Japan. pp 1251–1279

Bard P-Y, Cadet H, Endrun B, et al (2010) From Non-invasive Site Characterization to Site Amplification: Recent Advances in the Use of Ambient Vibration Measurements. In: Garevski M, Ansal A (eds) Earthquake Engineering in Europe. Springer Netherlands, pp 105–123

Bard PY, SESAME-Team (2005) Guidelines for the implementation of the H/V spectral ratio technique on ambient vibrations: measurements, processing, and interpretations. SESAME European research project. SESAME European research project

Beaty KS, Schmitt DR, Sacchi M (2002) Simulated annealing inversion of multimode Rayleigh wave dispersion curves for geological structure. Geophys J Int 151:622–631.

Berger JO (2013) Statistical Decision Theory and Bayesian Analysis. Springer Science & Business Media

Bonnefoy-Claudet S, Baize S, Bonilla LF, et al (2009) Site effect evaluation in the basin of Santiago de Chile using ambient noise measurements. Geophys J Int 176:925–937.

Bonnefoy-Claudet S, Cécile C, Pierre-Yves B, et al (2006) H/V ratio: a tool for site effects evaluation. Results from 1-D noise simulations. Geophysical Journal International 167:827–837.

Boore DM (2004) Estimating Vs(30) (or NEHRP site classes) from shallow velocity models (depths < 30 m). Bulletin of the Seismological Society of America 94:591–597

Boore DM, Stewart JP, Seyhan E, Atkinson GM (2013) NGA-West 2 Equations for Predicting Response Spectral Accelerations for Shallow Crustal Earthquakes. PEER 2013/05, Pacific Earthquake Engineering Research Center, Berkeley, California 135

Brooks SP, Gelman A (1998) General methods for monitoring convergence of iterative simulations. Journal of computational and graphical statistics 7:434–455

Building Seismic Safety Council, BSSC (2003) NEHRP recommended provisions for seismic regulations for new buildings and other structures. Washington

Capon J (1969) High-resolution frequency-wavenumber spectrum analysis. Proceedings of the IEEE 57:1408–1418.

Casella G, George EI (1992) Explaining the Gibbs Sampler. The American Statistician 46:167–174.

Castellaro S, Mulargia F (2009) VS30 Estimates Using Constrained H/V Measurements. Bulletin of the Seismological Society of America 99:761–773.

Congdon P (2014) Applied bayesian modelling. John Wiley & Sons

Cox BR, Teague DP (2016) Layering ratios: a systematic approach to the inversion of surface wave data in the absence of a priori information. Geophys J Int 207:422–438.

Dal Moro G, Ferigo F (2011) Joint analysis of Rayleigh- and Love-wave dispersion: Issues, criteria and improvements. Journal of Applied Geophysics 75:573–589.

Dal Moro G, Pipan M, Gabrielli P (2007) Rayleigh wave dispersion curve inversion via genetic algorithms and Marginal Posterior Probability Density estimation. Journal of Applied Geophysics 61:39–55.

D'Amico V, Picozzi M, Baliva F, Albarello D (2008a) Ambient noise measurements for preliminary site-effects characterization in the urban area of Florence, Italy. Bulletin of the Seismological Society of America 98:1373–1388

D'Amico V, Picozzi M, Baliva F, Albarello D (2008b) Ambient Noise Measurements for Preliminary Site-Effects Characterization in the Urban Area of Florence, Italy. Bulletin of the Seismological Society of America 98:1373–1388.

Di Giacomo D, Maria Rosaria G, Mucciarelli, M, et al (2005) Analysis and Modeling of HVSR in the Presence of a Velocity Inversion: The Case of Venosa, Italy. Bulletin of the Seismological Society of America 95:2364–2372

Diggle PJ, Tawn JA, Moyeed RA (1998) Model-based geostatistics. Journal of the Royal Statistical Society: Series C (Applied Statistics) 47:299–350.

Douglas J (2010) Consistency of ground-motion predictions from the past four decades. Bulletin of Earthquake Engineering 8:1515–1526.

Eskişar T, Özyalin Ş, Kuruoğlu M, Yilmaz HR (2013) Microtremor measurements in the northern coast of İzmir Bay, Turkey to evaluate site-specific characteristics and fundamental periods by H/V spectral ratio method. J Earth Syst Sci 122:123–136.

European Committee for Standardization (2003) Eurocode 8: Design of Structures for earthquake resistance - Part1: General rules, seismic actions and rules for buildings. European Standard

Forbriger T (2003) Inversion of shallow-seismic wavefields: II. Inferring subsurface properties from wavefield transforms. Geophys J Int 153:735–752.

Foti S, Comina C, Boiero D, Socco LV (2009) Non-uniqueness in surface-wave inversion and consequences on seismic site response analyses. Soil Dynamics and Earthquake Engineering 29:982–993.

Gallipoli MR, Mucciarelli M (2009) Comparison of Site Classification from VS30, VS10, and HVSR in Italy. Bulletin of the Seismological Society of America 99:340–351

Garofalo F, Foti S, Hollender F, et al (2016a) InterPACIFIC project: Comparison of invasive and non-invasive methods for seismic site characterization. Part II: Inter-comparison between surface-wave and borehole methods. Soil Dynamics and Earthquake Engineering 82:241–254.

Garofalo F, Foti S, Hollender F, et al (2016b) InterPACIFIC project: Comparison of invasive and non-invasive methods for seismic site characterization. Part I: Intra-comparison of surface wave methods. Soil Dynamics and Earthquake Engineering 82:222–240.

Gelman A, Carlin J., Stern HS, Rubin DB (2014) Bayesian data analysis. Chapman & Hall/CRC, London

Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. Statistical Science 457–472

Geman S, Geman D (1984) Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-6:721–741.

Gilks WR (2005) Markov Chain Monte Carlo. In: Encyclopedia of Biostatistics. John Wiley & Sons, Ltd

Gok E (2011) Investigation of Earthquake Hazard and Seismic Site Characteristic in the Examples of Bursa and Izmir. PhD thesis, Graduate School of Natural and Applied Science, Dokuz Eylul University, Izmir.

Gok E, Chávez-García FJ, Polat O (2014) Effect of soil conditions on predicted ground motion: Case study from Western Anatolia, Turkey. Physics of the Earth and Planetary Interiors 229:88–97.

Gok E, Polat O (2014) An assessment of the microseismic activity and focal mechanisms of the Izmir (Smyrna) area from a new local network (IzmirNET). Tectonophysics 635:154–164.

Griffiths SC, Cox BR, Rathje EM, Teague DP (2016) Surface-Wave Dispersion Approach for Evaluating Statistical Models That Account for Shear-Wave Velocity Uncertainty. Journal of Geotechnical and Geoenvironmental Engineering 142(11):04016061.

Griffiths Shawn C., Cox Brady R., Rathje Ellen M., Teague David P. (2016) Mapping Dispersion Misfit and Uncertainty in Vs Profiles to Variability in Site Response Estimates. Journal of Geotechnical and Geoenvironmental Engineering 142:04016062.

Herak M (2008) ModelHVSR-A Matlabs tool to model horizontal-to-vertical spectral ratio of ambient noise. Computers & Geosciences 34:1514–1526

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

Kramer SL (1996) Geotechnical earthquake engineering. Prentice Hall, USA: New Jersey

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

Lai CG, Rix GJ (1998) Simultaneous inversion of Rayleigh phase velocity and attenuation for near-surface site characterization. School of Civil and Environmental Engineering, Georgia Institute of Technology

Link WA, Eaton MJ (2012) On thinning of chains in MCMC. Methods in Ecology and Evolution 3:112–115.

Luke B, Calderón-Macías C, Stone R, Huynh M (2003) Non-Uniqueness in Inversion of Seismic Surface-Wave Data. In: Symposium on the Application of Geophysics to Engineering and Environmental Problems 2003. Environment and Engineering Geophysical Society, pp 1342–1347

Lunedei E, Albarello D (2010) Theoretical HVSR curves from full wavefield modelling of ambient vibrations in a weakly dissipative layered Earth. Geophys J Int 181:1093–1108.

Metropolis N, Rosenbluth AW, Rosenbluth MN, et al (1953) Equation of state calculations by fast computing machines. The journal of chemical physics 21:1087–1092

Molnar S, Dosso SE, Cassidy JF (2010) Bayesian inversion of microtremor array dispersion data in southwestern British Columbia. Geophys J Int 183:923–940.

Mucciarelli and Gallipoli (2001) A critical review of 10 years of microtremor HVSR technique.pdf. Bollettino di Geofisica Teorica ed Applicata 42:255–266

Mucciarelli M (1998) Reliability and applicability of Nakamura's technique using microtremors: an experimental approach. Journal of earthquake engineering 2:625–638

Nakamura Y (2008) On the H/V spectrum. In: Proceedings of the 14th World Conference on Earthquake Engineering. Beijing, China

Nakamura Y (2000a) Clear identification of fundamental idea of Nakamura's technique and its applications. In: Proceedings of the 12th World Conference on Earthquake Engineering. p 2656

Nakamura Y (2000b) Clear Identification of Fundamental Idea of Nakamura´s Technique and Applications

Nogoshi M, Igarashi T (1970) On the propagation characteristics estimations of subsurface using microtremors on the ground surface. Journal of Seismological Society of Japan 23:264–280

Nolet G, Panza GF (1976) Array analysis of seismic surface waves: Limits and possibilities. PAGEOPH 114:775–790.

O'Neill A, Matsuoka T (2005) Dominant Higher Surface-wave Modes and Possible Inversion Pitfalls. Journal of Environmental & Engineering Geophysics 10:185–201.

Parolai S, Mucciarelli M, Gallipoli MR, et al (2007) Comparison of Empirical and Numerical Site Responses at the Tito Test Site, Southern Italy. Bulletin of the Seismological Society of America 97:1413–1431.

Parolai S, Picozzi M, Richwalski SM, Milkereit C (2005) Joint inversion of phase velocity dispersion and H/V ratio curves from seismic noise recordings using a genetic algorithm, considering higher modes. Geophys Res Lett 32:L01303.

Picozzi M, Albarello D (2007) Combining genetic and linearized algorithms for a two-step joint inversion of Rayleigh wave dispersion and H/V spectral ratio curves. Geophys J Int 169:189–200.

Picozzi M, Strollo A, Parolai S, et al (2009) Site characterization by seismic noise in Istanbul, Turkey. Soil Dynamics and Earthquake Engineering 29:469–482.

Pilz M, Parolai S, Leyton F, et al (2009) A comparison of site response techniques using earthquake data and ambient seismic noise analysis in the large urban areas of Santiago de Chile. Geophysical Research Letters 178:713–728.

Poggi V, Fäh D, Burjanek J, Giardini D (2012) The use of Rayleigh-wave ellipticity for site-specific hazard assessment and microzonation: application to the city of Lucerne, Switzerland. Geophys J Int 188:1154–1172.

Polat O, Ceken U, Uran T, et al (2009) IzmirNet: A strong-motion Network in Metropolitan Izmir, Western Anatolia, Turkey. Seismological Research Letters 80:831–838.

Polat O, Gok E, Ceken U, Kuruoglu M (2012) Seismic studies from small aperture accelerometric network in Izmir metropolitan city, Turkey. In: 2012 IV International Conference "Problems of Cybernetics and Informatics" (PCI). pp 1–5

Power M, Chiou BS-J, Abrahamson NA, et al (2008) An Overview of the NGA Project. Earthquake Spectra 24:3–21

Rahpeyma S, Halldorsson B, Green RA (2017) On the Distribution of Earthquake Strong-motion Amplitudes and Site Effects Across the Icelandic Strong-motion Arrays. In: 16th World Conference on Earthquake Engineering (16WCEE). Santiago, Chile, p Paper no. 2762.

Rahpeyma S, Halldorsson B, Olivera C, et al (2016) Detailed site effect estimation in the presence of strong velocity reversals within a small-aperture strong-motion array in Iceland. Soil Dynamics and Earthquake Engineering 89:136–151.

Rothman D (1985) Nonlinear inversion, statistical mechanics, and residual statics estimation. GEOPHYSICS 50:2784–2796.

Sambridge M (2001) Finding acceptable models in nonlinear inverse problems using a neighbourhood algorithm. Inverse Problems 17:387.

Sánchez-Sesma FJ, Rodríguez M, Iturrar´an-Viveros U, et al (2011) A theory for microtremor H/V spectral ratio: application for a layered medium. Geophysical Journal International 186:221–225

Scherbaum F, Hinzen K-G, Ohrnberger M (2003) Determination of shallow shear wave velocity profiles in the Cologne, Germany area using ambient vibrations. Geophys J Int 152:597–612.

Schevenels M, Lombaert G, Degrande G, François S (2008) A probabilistic assessment of resolution in the SASW test and its impact on the prediction of ground vibrations. Geophys J Int 172:262–275.

Shapiro N (1996) Etude de l'intéraction des ondes sismiques guidées régionales avecles hétérogénéités delacroute. PhDthesisUniversite´ Joseph Fourier

Smith AFM, Roberts GO (1993) Bayesian Computation Via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods. Journal of the Royal Statistical Society Series B (Methodological) 55:3–23

Socco L, Boiero D, Foti S, Wisén R (2009) Laterally constrained inversion of ground roll from seismic reflection records. GEOPHYSICS 74:G35–G45.

Socco L, Foti S, Boiero D (2010) Surface-wave analysis for building near-surface velocity models — Established approaches and new perspectives. GEOPHYSICS 75:75A83-75A102.

Socco LV, Boiero D (2008) Improved Monte Carlo inversion of surface wave data. Geophysical Prospecting 56:357–371.

Sözbilir H, Uzel B, Sümer Ö, et al (2008) Evidence for a kinematically linked EW trending İzmir Fault and NE-trending Seferihisar Fault: Kinematic and paleoseismogical studies carried out on active faults forming the İzmir Bay, Western Anatolia. Geological Bulletin of Turkey (in Turkish) 51:91–114

Tarabusi G, Caputo R (2017) The use of HVSR measurements for investigating buried tectonic structures: the Mirandola anticline, Northern Italy, as a case study. Int J Earth Sci (Geol Rundsch) 106:341–353.

Teague DP, Cox BR (2016) Site response implications associated with using non-unique Vs profiles from surface wave inversion in comparison with other commonly used methods of accounting for Vs uncertainty. Soil Dynamics and Earthquake Engineering.

Tsai NC (1970) A Note on the Steady-State response of an Elastic Half-Space. Bulletin of the Seismological Society of America 60:795–808

Uzel B, Özkaymak HS& Ç (2014) Neotectonic Evolution of an Actively Growing Superimposed Basin in Western Anatolia: The Inner Bay of İzmir, Turkey. Turkish Journal of Earth Sciences 21:439–471.

Wathelet M, Jongmans D, Ohrnberger M (2004) Surface-wave inversion using a direct search algorithm and its application to ambient vibration measurements. Near Surface Geophysics 2:211–221.

Xia J, Miller R, Park C (1999) Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves. GEOPHYSICS 64:691–700.

Xia J, Miller RD, Park CB, Tian G (2003) Inversion of high frequency surface waves with fundamental and higher modes. Journal of Applied Geophysics 52:45–57.

Yamanaka H, Ishida H (1996) Application of genetic algorithms to an inversion of surface-wave dispersion data. Bulletin of the Seismological Society of America 86:436–444.

Figures

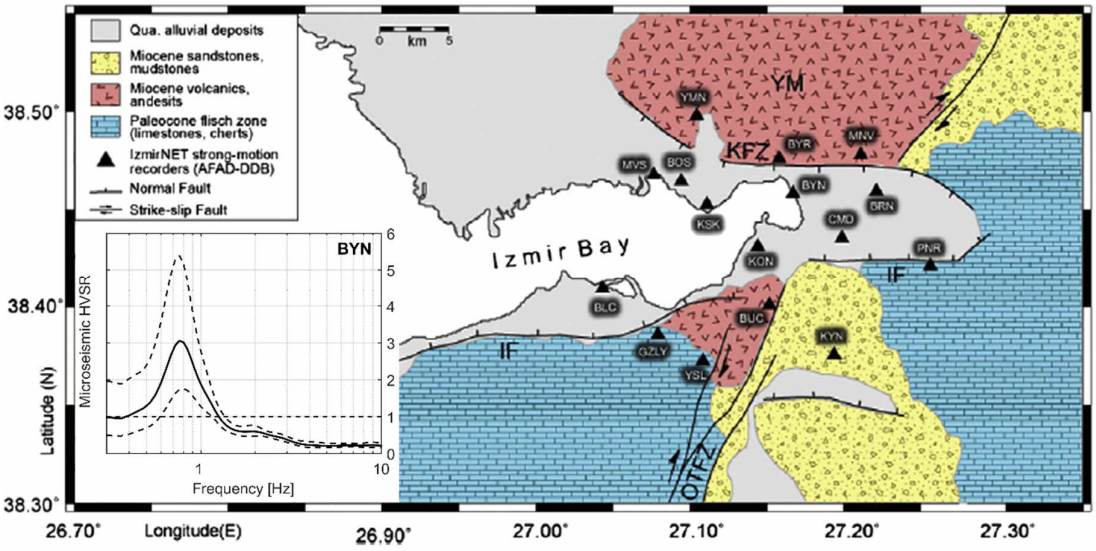

Figure 1. Location of IzmirNet array (filled triangles) on geology of Izmir and simplified geological features. YM: Yamanlar mountain, IF: Izmir Fault, KFZ: Karsiyaka Fault Zone, OTFZ: Orhanli-Tuzla Fault Zone, SFZ: Seferihisar Fault Zone (Gok and Polat 2014). The insert figure on bottom left shows the observed HVSR form microseismic measurements with a Konno and Ohmachi smoothing coefficient B=20 for BYN station.

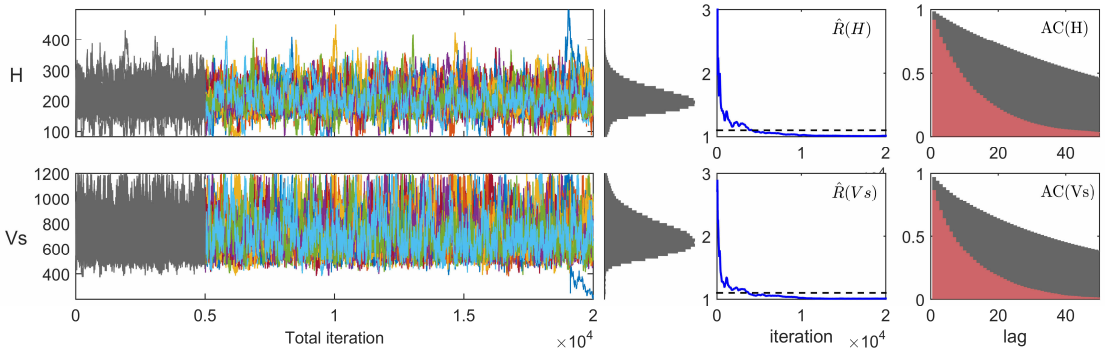

Figure 2. Convergence diagnostics: trace plots of all sampling chains ($N_c = 20$) with total iteration ($N_T = 20{,}000$) and considering burn-in period ($N_B = 5000$) in gray based on the MCMC simulations; model parameters posterior histogram; Gelman-Rubin plots, the blue line denotes the median of Gelman-Rubin statistics as a function of iterations; autocorrelation plots based on the MCMC simulations until lag 50 for a single-layer soil structure (gray bars), autocorrelation function after thinning of scale 5 (red bars).

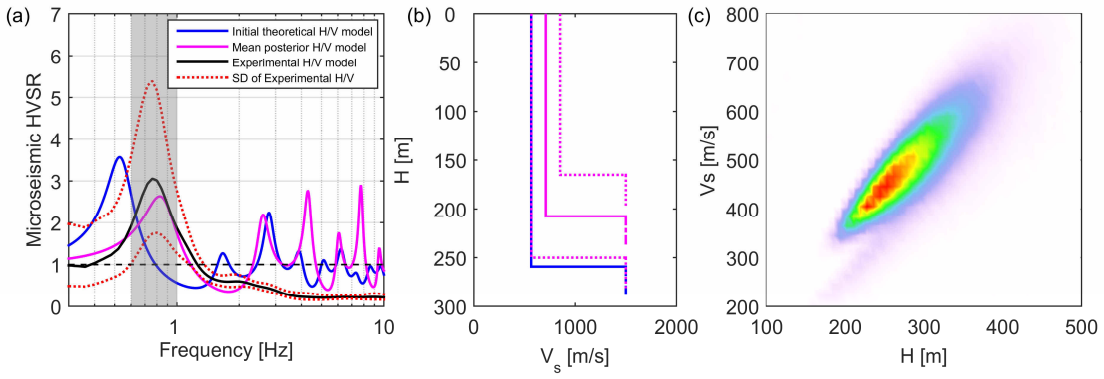

Figure 3. (a) S-wave velocity profile for initial (blue) and final (pink) model obtained from MCMC simulation presented in Figure 2 mean HVSR (solid black line) and standard deviation (dotted red line) of microseismic recordings, initial HVSR model (solid blue line), and final HVSR obtained from posterior simulations (solid pink line). (c) joint marginal probability distribution from inversion of the microseismic data.

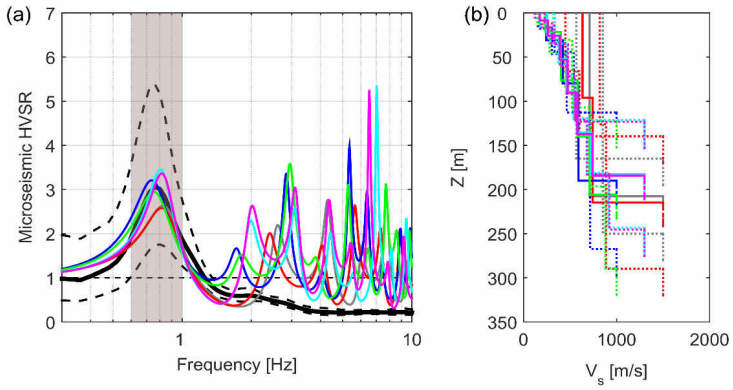

Figure 4. (a) Different theoretical HVSR curves and (b) different Vs profiles obtained by posterior mean values of model parameters (H and Vs) using different layering.

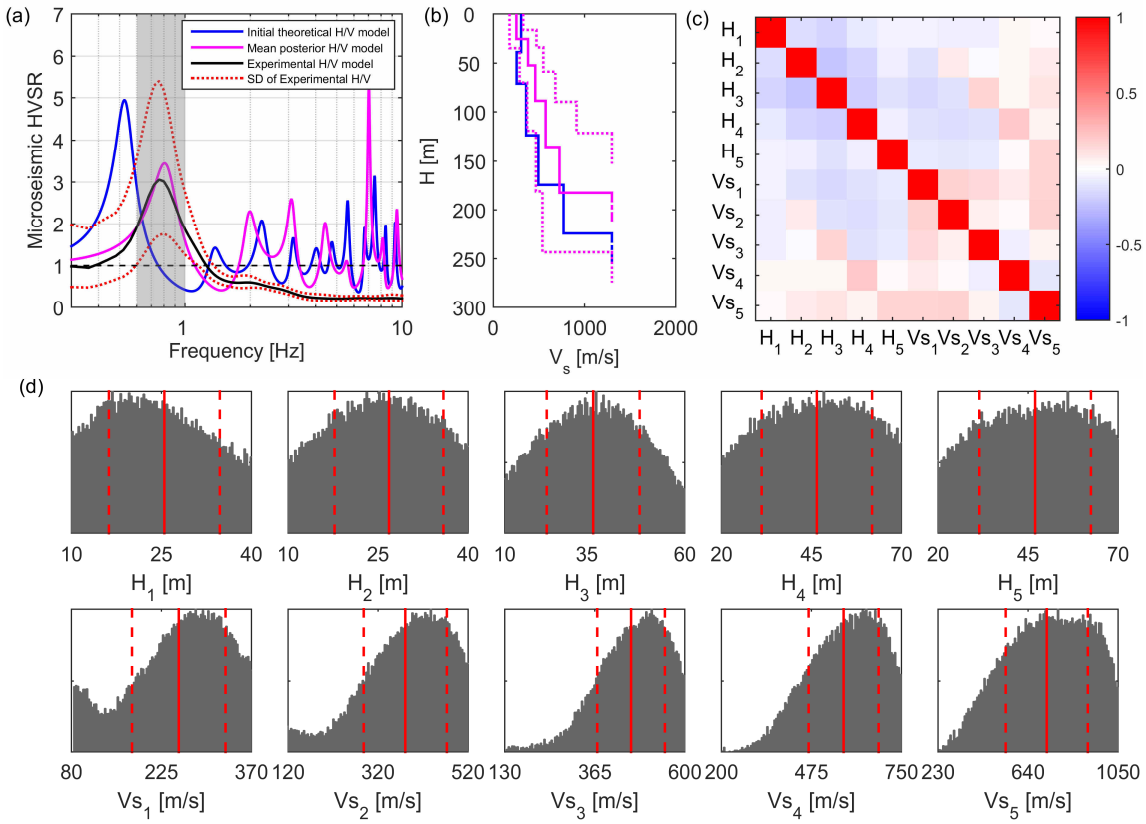

Figure 5. (a) The observed, initial, and posterior HVSR obtained from Bayesian MCMC inversion for a 5-layer subsoil structure at BYN station, Turkey; (b) S-wave velocity prrofile for the initial (blue) and posterior (magenta) model; (c) Correlation matrix of posterior samples of model parameters; (d) posterior histograms for thickness (H) and S-wave velocity (Vs).

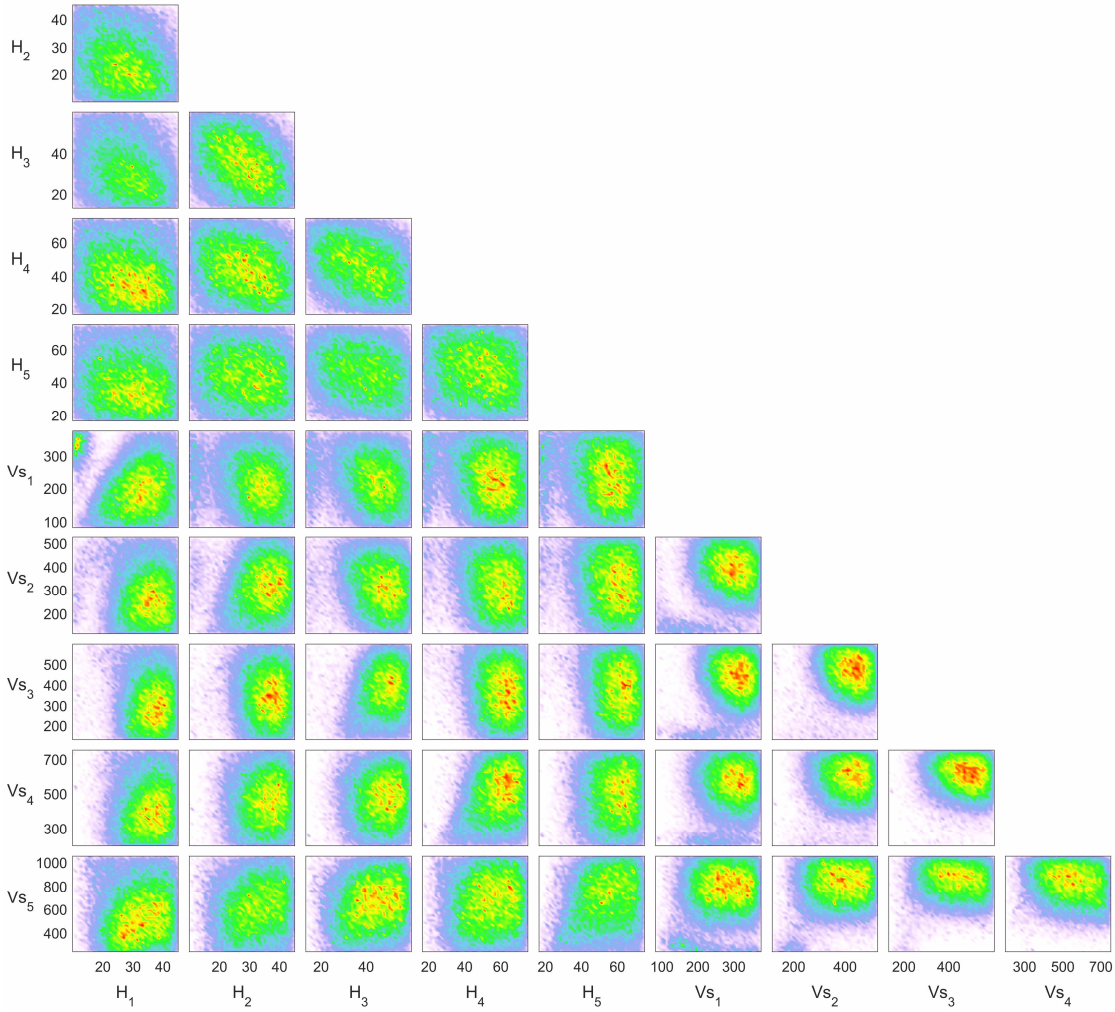

Figure 6. Normalized joint marginal probability distribution from inversion of microseismic HVSR at BYN station for a five-layer subsoil model overlaying half-space presented in Figure 5.

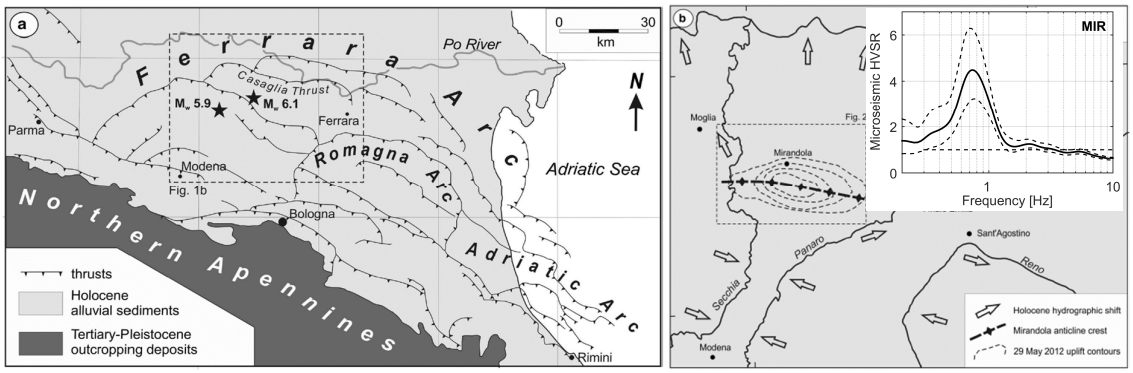

Figure 7. Geological and tectonic sketch map of the buried Northern Apennines fold-and-thrust belt. Stars represent the epicenters of 20 May (*M*w 6.1) and 9 May (*M*w 5.9) 2012 earthquakes (Tarabusi and Caputo 2017). The insert figure on right top shows the observed HVSR form microseismic measurements with a Konno and Ohmachi smoothing coefficient B=20 for MIR station.

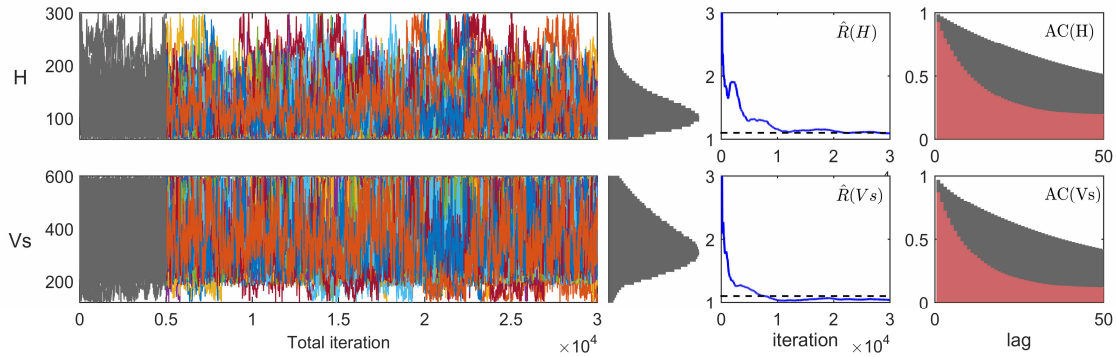

Figure 8. Convergence diagnostics: trace plots of all sampling chains ($N_C = 30$) with total iteration ($N_T = 30{,}000$) and considering burn-in period ($N_B = 5000$) in gray based on the MCMC simulations; model parameters posterior histogram; Gelman-Rubin plots, the blue line denotes the median of Gelman-Rubin statistics as a function of iterations; autocorrelation plots based on the MCMC simulations until lag 50 for a single-layer soil structure (gray bars), autocorrelation function after thinning of scale 5 (red bars).

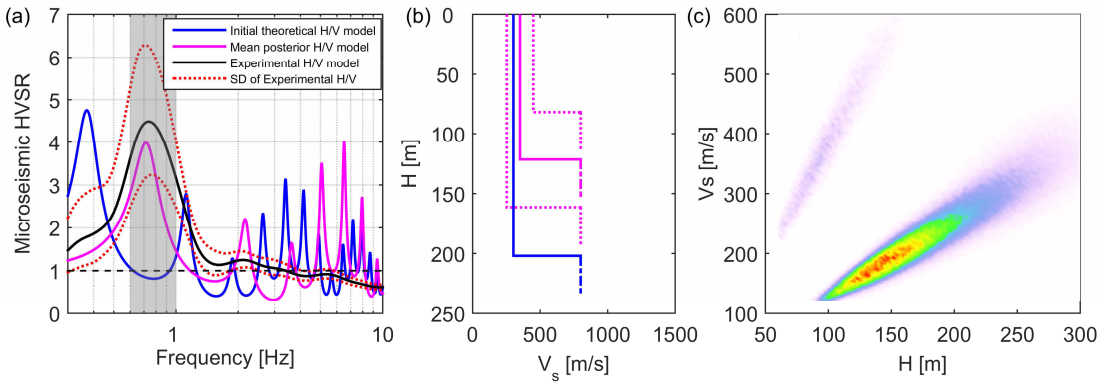

Figure 9. (a) S-wave velocity profile for initial (blue) and final (pink) model obtained from MCMC simulation presented in Figure 8. mean HVSR (solid black line) and standard deviation (dotted red line) of microseismic recordings, initial HVSR model (solid blue line), and final HVSR obtained from posterior simulations (solid pink line). (c) joint marginal probability distribution from inversion of the microseismic data.

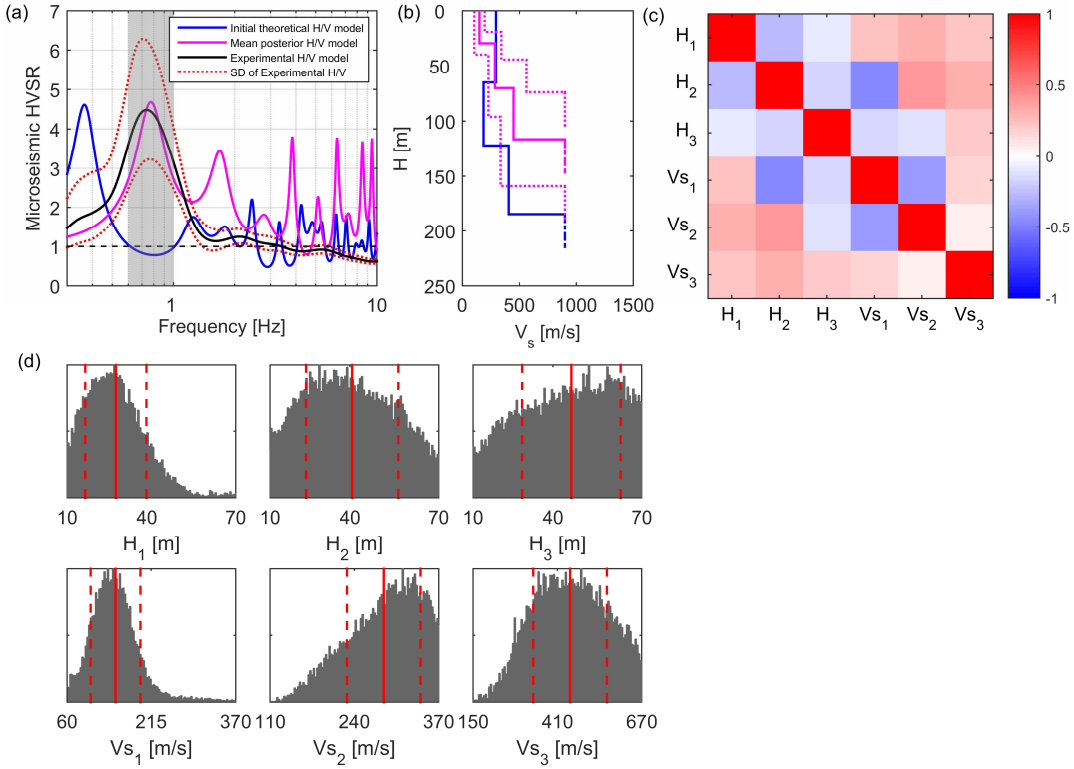

Figure 10. (a) The observed, initial, and posterior HVSR obtained from Bayesian MCMC inversion for a 3-layer subsoil structure at MIR station, Italy; (b) V_s velocity profile for the initial (blue) and posterior (magenta) model; (c) Correlation matrix of posterior samples of model parameters; (d) posterior histograms for thickness and shera-wave velocity.

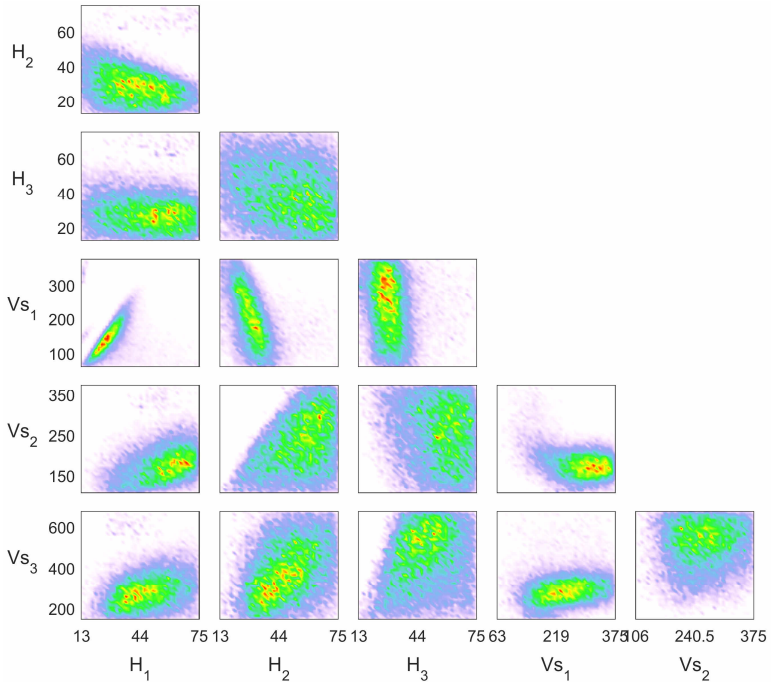

Figure 11. Normalized joint marginal probability distribution from inversion of microseismic HVSR at BYN station for a five-layer subsoil model overlaying half-space presented in Figure 10.