



HAL
open science

Influential parameters on 3-D synthetic ground motions in a sedimentary basin derived from global sensitivity analysis

Florent de Martin, E Chaljub, P Thierry, Pierre Sochala, Fabrice Dupros, Emeline Maufroy, B Hadri, A Benaichouche, F Hollender

► To cite this version:

Florent de Martin, E Chaljub, P Thierry, Pierre Sochala, Fabrice Dupros, et al.. Influential parameters on 3-D synthetic ground motions in a sedimentary basin derived from global sensitivity analysis. *Geophysical Journal International*, 2021, 227 (3), pp.1795-1817. 10.1093/gji/ggab304 . hal-03744613

HAL Id: hal-03744613

<https://brgm.hal.science/hal-03744613>

Submitted on 3 Aug 2022

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

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

Influential parameters on 3-D synthetic ground motions in a sedimentary basin derived from global sensitivity analysis

F. De Martin¹, E. Chaljub², P. Thierry³, P. Sochala¹, F. Dupros⁴, E. Maufroy², B. Hadri⁵, A. Benaichouche⁶ and F. Hollender^{2,7}

¹BRGM (French Geological Survey), Risks and Prevention Division, 45060 Orléans, France. E-mail: fdemartin@free.fr

²Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, UGE, ISTerre, Département Géophysique des Risques et de l'Environnement, 38041 Grenoble, France

³Intel Corporation, 92190 Meudon, France

⁴ARM Sophia-Antipolis, 0656 Valbonne, France

⁵King Abdullah University of Science and Technology (KAUST), Supercomputing Core Lab 23955, Thuwal, Saudi Arabia

⁶Group 42, Al Khaleej Al Arabi Street, Abu Dhabi, United Arab Emirates

⁷CEA Cadarache, 13115 Saint-Paul-lez-Durance, France

Accepted 2021 0. Received 2021 July 16; in original form 2020 October 7

SUMMARY

Which physical parameters are the most influential when predicting earthquake ground motions in a 3-D sedimentary basin? We answer quantitatively by doing a global sensitivity analysis of two quantities of interest: the peak ground motions (PGMs) and a time–frequency representation (the S transform) of ground motions resulting from the synthetic anelastic responses of the EUROSEISTEST. This domain of interest is modeled by two layers with uncertain depth-dependent mechanical properties and is illuminated by a plane S -wave propagating vertically upward in an uncertain homogeneous elastic bedrock. The global sensitivity analysis is conducted on 800+ physics-based simulations of the EUROSEISTEST requiring 8+ million core-hours (i.e. ≈ 900 yr of mono-core computation). The analysis of the PGMs at the free surface displays the spatial influence of the uncertain input parameters over the entire basin scale, while the analysis of the time–frequency representation shows their influence at a specific location inside the basin. The global sensitivity analysis done on the PGMs points out that their most influential parameter in the middle of the basin is the quality factor Q_S (it controls up to 80 per cent of the PGMs in certain locations where the sediments thickness is larger than 200 m). On the other hand, the geological layering configuration (here represented by the depth of a geological interface controlling the geological layering) strongly influences the PGMs close to the basin edges, up to 90 per cent. We also found that the shear wave velocity at the free surface of the basin and the one of the bedrock underlying the basin are to be considered on an equal footing, both influencing the PGMs in the middle of the basin and close to its edges. We highlight that the bedrock to basin amplification of the PGMs shows a clear increase with respect to the thickness of the sediments, but this amplification saturates from 200 m of sediments around the value of three and is frequency dependent. This PGMs amplification starts from about one tenth of the mean S -wavelength propagating in the basin. The global sensitivity analysis done on the S transform of the ground motions shows that (i) the own effect of the parameters fully controls the first S -wave train and mostly controls the direct arrival of the basin-induced surfaces waves, (ii) the quality factor Q_S controls 40–60 per cent of the decay of amplitude of coda waves, the remaining part being mainly controlled by interaction effects due to the coupling effect of several parameters and (iii) the interaction effects between the parameters increases with time, suggesting under the hypotheses of our study that the own effects control the ballistic wave propagation while the interaction effects control the diffusive wave propagation.

Key words: Probability distributions; Computational seismology; Earthquake ground motions; Site effects; Wave propagation.

1 INTRODUCTION

Sensitivity analyses applied to earthquake ground motion prediction is helpful to understand what mechanism controls the ground motion. This topic is particularly interesting because the mechanical and geometrical properties involved in the seismic-wave propagation are poorly known at the scales of our problems, ranging from several tens of kilometres to characterize sedimentary basins to tens of metres to detail local features. Despite colossal works have been done to build realistic 3-D earthquake hazard models, uncertainties naturally remain in the mechanical and geometrical properties of the geological strata due to the intrinsic complexity of the Earth and the difficulty to image it. Past studies have shown that physics-based ground motion predictions below 1 Hz can fairly reproduce the observations (e.g. Komatitsch *et al.* 2004; Iwata *et al.* 2008; Lee *et al.* 2008; Koketsu *et al.* 2009; Asano *et al.* 2016, 2017), however, predictions above 1 Hz still need a better understanding of the governing phenomena involved in the source, path and site effects. For instance, Graves & Pitarka (2016) showed that a key factor in matching the observed ground motion characteristics for frequencies larger than 1 Hz is obtaining the proper level of coherence in both the radiation and propagation of the wavefield; and Maufroy *et al.* (2015) pointed out that the differences between the observed and predicted ground motions have multiple origins, like the accuracy of source parameters or the uncertainties in the description of the geological medium. These uncertainties related to the geological medium have been investigated by Moczo *et al.* (2018) who confirmed for different sedimentary basins (including the EURO-SEISTEST) that the key structural parameters controlling the site amplification are the impedance contrast at the sediment/bedrock interface—as already shown by Bard & Bouchon (1980)—and the attenuation in sediments. The present work pursues the investigation of the uncertainties associated with the geological medium but does not limit to the identification of the influential parameters, it also quantifies their relative influence through the so-called ‘sensitivity indices’ resulting from a variance-based global sensitivity analysis proposed by Sobol (2001). Such a quantification helps to rank the uncertain parameters with respect to their degree of influence, revealing the important and negligible ones. This paper presents such analysis conducted up to 3 Hz in order to understand what controls the ground motion in a sedimentary basin below and above 1 Hz.

The impulsion that disseminated sensitivity analyses in various fields was given by the pioneer works of Sobol & Levitan (1999) and Sobol (2001). Among the existing sensitivity methods, such as linear regressions and local approaches, the global sensitivity analysis (GSA) has become popular since it measures the influence of individual input parameters (called ‘own effects’) or groups of input parameters (called ‘interactions effects’) when varying over the whole parametric domain without linearity or differentiability assumption of the model output. The own effects of a parameter i (among n) is measured by its first order sensitivity indices S_i and the interactions effects (i.e. the coupling of action of the parameters) are quantified by the second, third, ..., n order sensitivity indices that measure the influence of coupling together two, three, ..., n parameters. A first way of estimating the sensitivity indices is through the pick freeze Monte Carlo algorithm (e.g. Sobol 1993; Homma & Saltelli 1996) that requires an extensive sampling of the parametric domain (typically several hundreds of thousands) to estimate the partial variances. Another approach consists of building a polynomial chaos surrogate in order to analytically derive the partial variances (e.g. Sudret 2008; Crestaux *et al.* 2009).

We use here an alternative method proposed by Li & Mahadevan (2016) designed for estimating only the first-order sensitivity indices but from a limited number of samples. The main advantage of this method is that it directly uses the raw data of the simulations (e.g. the ground motions), avoiding the construction of a surrogate model.

Thanks to this sample-based method, we perform a GSA of the synthetic response of a 3-D sedimentary basin with respect to the uncertainties associated with the depth-dependent velocity structure of its sedimentary layers and of the underlying bedrock. Our work is motivated by the following question: what are the most influential parameters on the response of this sedimentary basin illuminated by a vertically incident plane shear wave? To answer, we focus on two quantities of interest commonly used in earthquake ground motion prediction: the peak ground motion (PGM) and the time–frequency representation of the ground motion, computed here with the S transform (Stockwell *et al.* 1996). The responses of the ensemble of the models of the basin rely on a high-performance 3-D physics-based solver (De Martin 2011) using a spectral-element method introduced in fluid dynamics by Maday & Patera (1989) and adapted for seismology by Komatitsch & Vilotte (1998); Komatitsch & Tromp (2002). Although our solver is efficiently parallelized using the message passing interface (e.g. Gropp *et al.* 1996), the cost of a single simulation solving 1.4 billion equations still amounts to 4 hr 45 min on 2048 Intel Haswell cores (corresponding to about 10 000 core-hours, i.e. ≈ 416 d on a single core). Because of this expensive computational cost, we carried out a limited ensemble of simulations, 400+ for the north–south (NS) polarized responses of the basin and 400+ for the east–west (EW) ones, on the supercomputers Shaheen II¹ and Occigen,² for a total computational budget of 8+ millions core-hours.

The PGMs extracted from the NS ensemble of simulations have already been studied by Sochala *et al.* (2020) with a polynomial chaos surrogate in order to compute various statistical information on the uncertain prediction of the PGMs, including marginal and joint probability distributions, interval probability maps and 2-D fields of global sensitivity indices. One goal of this paper is to extend this previous analysis with a time–frequency decomposition of the ground motions in order to have a full time–frequency visualization of the influence of the uncertain input parameters. The paper also completes the GSA presented by Sochala & De Martin (2017) in which the ground motion at the TST station (see Section 2) had been computed under 1-D wave propagation hypothesis and analysed by a GSA independently in the time and frequency domains. Some conclusions of these two articles will be discussed in the light of those presented here.

Concerning the uncertainties to be included in our study, four of the six classes of uncertainties defined by Kennedy & O’Hagan (2001) are present in our model: the *parameter uncertainty* corresponding to the stochastic description of the basin soil structure, the *model inadequacy* caused by the underlying hypotheses used in the construction of the geophysical model, the *parametric variability* due to the simplifying assumptions of the input uncertainty model, and the *code uncertainty* resulting from the incompleteness of sampling in the ensemble of simulations. The two other classes, not present in our study, are the *residual variability* that

¹Supercomputing Core Lab at the King Abdullah University of Science and Technology (KAUST, KSA).

²National Computing Center for Higher Education (CINES, France).

would be present in case of intrinsic stochastic process, like a random slip distribution over a fault, and the *observation error* that appears when solving inference (or calibration) problems with real data. While the *parameter uncertainty* is the most important class to answer our question, we mention along the text where the *model inadequacy*, the *parametric variability* and the *code uncertainty* are present.

Although we shall treat together the uncertainties present from the source to the site, we only focus on the uncertainties present in the sedimentary basin (the so-called ‘site effects’) in order to restrict the number of uncertain parameters and limit the ensuing number of model evaluations. Consequently, the path and source effects are excluded from our uncertainty model because they would add multiple parameters (see Section 6) and substantial computational cost to evaluate their relative influence. The former are avoided by considering a homogeneous half-space and the latter by representing the source as a unit impulse plane wave propagating vertically. Obviously, an exhaustive representation of the uncertainties would include these two effects because the earth contains heterogeneity on various scales (e.g. Phoon & Kulhawy 1999; Sato *et al.* 2012) and the source rupture process is known to be heterogeneous (e.g. Beroza 1991; Mai & Beroza 2002). The influence of the spatial heterogeneities have been quantified in various numerical studies, for instance, Frankel & Clayton (1986) showed that the propagation of high-frequency (>1 Hz) seismic waves in the earth’s crust gives significant information about the fluctuation spectrum of crustal heterogeneity on length scales from tens of kilometres to tens of metres; Iwaki *et al.* (2018) showed that the peak ground velocities are directly related to the structure of the random heterogeneity; and Takemura *et al.* (2020) demonstrated that the thick low-velocity accretionary prism has significant effects on high-frequency (>1 Hz) seismic wave propagation. Those spatial heterogeneities are often modeled by random fields generated either by a Karhunen–Loève expansion (e.g. Karhunen 1947; Loève 1968) or by a power spectral density function (e.g. Pardo-Iguzquiza & Chica-Olmo 1993). Such random fields are not considered in our study because solving the Karhunen–Loève expansion modal decomposition when the domain is much larger than the correlation length (which is our case) represents a computational issue that can quickly become unaffordable (e.g. Panunzio *et al.* 2018). In addition, considering a large number of Karhunen–Loève modes as uncertain parameters would compromise the feasibility of a GSA.

For our case study, we evaluate the impact of a probabilistic velocity structure defined by seven uncertain parameters and use a GSA to quantify their relative influence. Specifically, the backbone velocity structure of our domain of interest has two sedimentary layers with depth-dependent mechanical properties overlying a homogeneous elastic bedrock. The structure is described by seven independent random parameters, namely five shear wave velocities (two for each layer and one for the underlying bedrock), the depth of the interface between the two layers, and a scaling factor related to the shear-wave quality factor. As discussed in Section 6, other uncertainty models can be designed by using either new parameters (like the incidence of the input wave) or parameters derived from those selected (like the basin/bedrock impedance contrast).

The structure of the paper is organized as follow: we first introduce the site of study in Section 2 and the uncertainties of its geophysical model in Section 3. The numerical experimentation set up to compute the responses of the basin is described in Section 4 and the global sensitivity analysis is presented in Section 5. The results are discussed in Section 6 and conclusions are drawn in Section 7.

2 SITE OF STUDY

2.1 Overview

The site under investigation is the Mygdonian basin located in Greece, a few tens of kilometres away from the city of Thessaloniki (Fig. 1). This basin has been chosen since 1993 to be a multipurpose physical laboratory (called EUROSEISTEST) in the tectonically active graben of Mygdonia. Its web portal and ground-motion database³ are detailed in Pitolakis *et al.* (2013). The site is instrumented with an accelerometric network that consists of high-resolution, three-component accelerographs, five of them forming a vertical array at the TST site (Fig. 2a). As reported by Maufroy *et al.* (2015), the vertical array shows three *S*-wave resonant frequencies in the range [0–3] Hz (between the surface and down-hole sensors) that are attributable to the 1-D resonance, but the site is also prone to energetic surface waves around 1 Hz. The 1-D fundamental frequency is visible around 0.7 Hz and the first and second higher modes are visible around 1.75 and 2.75 Hz, respectively.

2.2 Geophysical models

Raptakis *et al.* (2000) built a 2-D geological and geophysical model from an extensive seismic campaign and a detailed geotechnical survey. A 3-D extension of this model is presented in Manakou *et al.* (2010) and Guyonnet-Benaize *et al.* (2013) in which the thickness of the sedimentary layers reaches about 450 m in the western and eastern sides of the basin (Fig. 3a). The velocity structure used in this study is the one updated by Maufroy *et al.* (2017) and contains two layers corresponding to the Mygdonian/Pre-Mygdonian (MP) geological formations, the former the layer L1 and the later the layer L2. At the TST station, the *S*-wave velocity starts from 130 m s⁻¹ at the free surface (top of L1) to reach 800 m s⁻¹ at the bottom of the sediments (bottom of L2), 200 m below (Fig. 4a, ‘*V_S* mean model’). The *S*-wave velocity in the underlying bedrock is 2400 m s⁻¹. The velocity model within the sedimentary layers is defined by two vertical gradients: one from the top to the bottom of the first layer L1, and a second one from the top to the bottom of the second layer L2. A schematic representation of this model is shown in Fig. 5 and the top and bottom values of the soil properties defining the double gradient are given in Table 1 (column ‘mean’).

The attenuation of the medium is defined through the empirical rule $Q = V/q$, with Q a quality factor, V a wave velocity and q a wave velocity scaling factor. The values of the *S*-wave quality factor Q_s , estimated by Jongmans *et al.* (1998) and Beauval *et al.* (2003) for the EUROSEISTEST, start around 20 at the free surface to reach approximately 100 at the bottom of the sediments. The mean Q_s structure of our study (Fig. 4b) is inspired by these studies and modified following (Maufroy *et al.* 2017). We observe on Fig. 4(b) that the empirical rule $Q_s = V_s/q_s$ used in this study is also broadly consistent with the relationship proposed by Brocher (2008) for the seismic model of northern California, but his Q_s structure mostly exhibits lower values than ours, that would lead to a larger attenuation of the amplitude of the ground motion.

The representation of the sedimentary basin exhibits significant differences between the 2-D and 3-D models, as shown by the cross-section in Fig. 2(b) done along the the AA’ line of Fig. 1. First, the

³<http://euroseis.civil.auth.gr> (last access 2021/08/02 17:41:57).

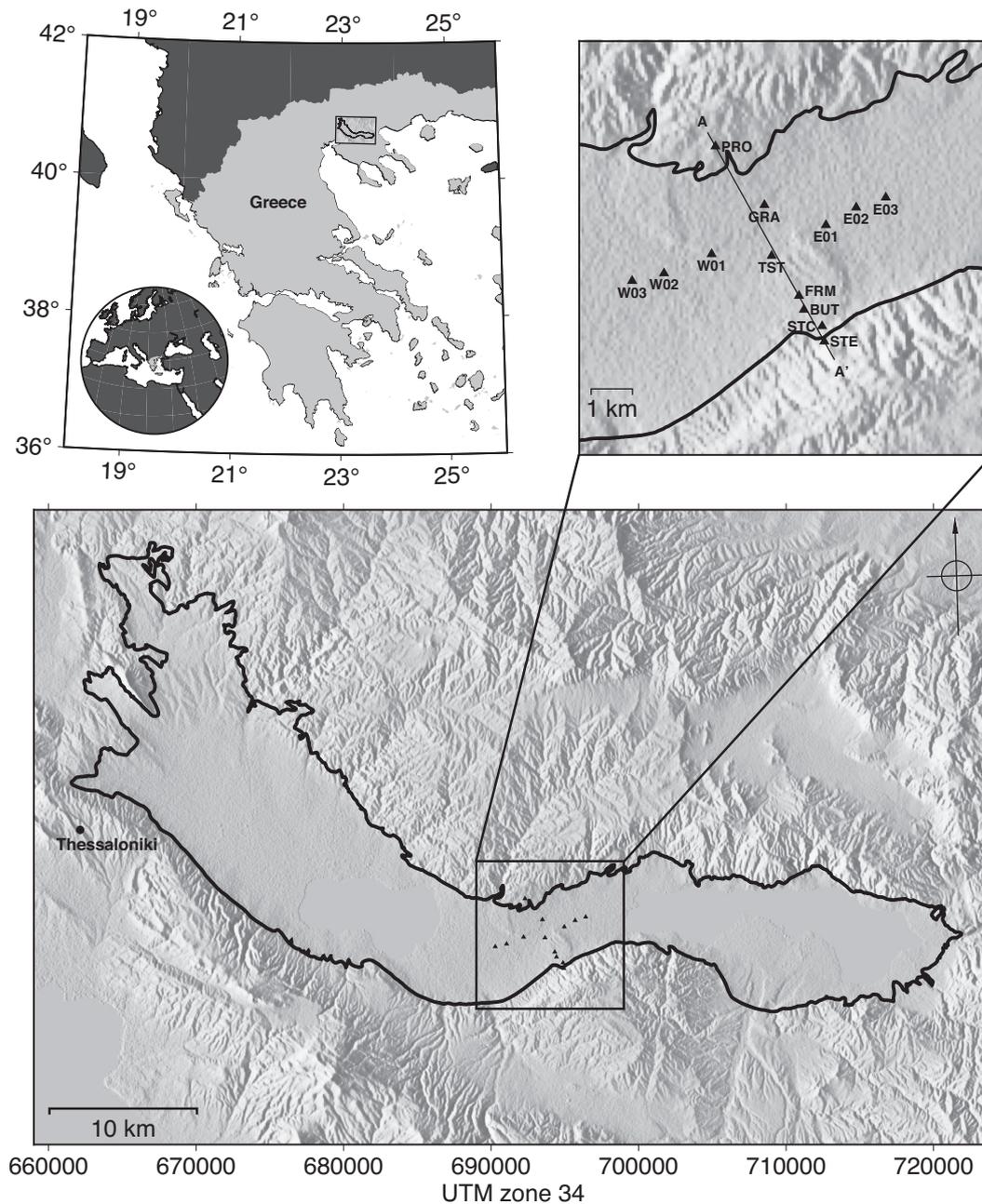


Figure 1. Top-left-hand panel: localization of the region of study in Greece. Bottom panel: domain of the numerical experimentation. The size of the box of simulations is $64 \times 42 \times 5 \text{ km}^3$. The limit sediment/bedrock at the surface is indicated by a solid line. Top-right-hand panel: zoom on the accelerometric network stations (shown as triangles) physically installed at the EUROSEISTEST.

number of layers is different since there are seven layers with piecewise constant properties for the 2-D model and two layers with piecewise linear properties for the 3-D model. The introduction of the gradient in each layer of the 3-D model allowed for a reduction of the number of layers while keeping a close velocity structure between the 2-D and 3-D models. Secondly, the shape of the geological strata is different between the two models: the SSE and NNW sides of the 3-D model are, respectively, deeper and gentler than those of the 2-D model. In the next section, we define the probabilistic framework associated with the 3-D geophysical model.

3 PROBABILISTIC FRAMEWORK

3.1 Choice of the uncertain parameters

3.1.1 Wave velocities

Because the S waves and surface waves are the most destructive waves for civil engineering infrastructures, our analysis naturally considers as uncertain the S -wave velocities. On the other hand, the P -wave velocities are here considered deterministic because their influence is assumed to be negligible on the response of a basin

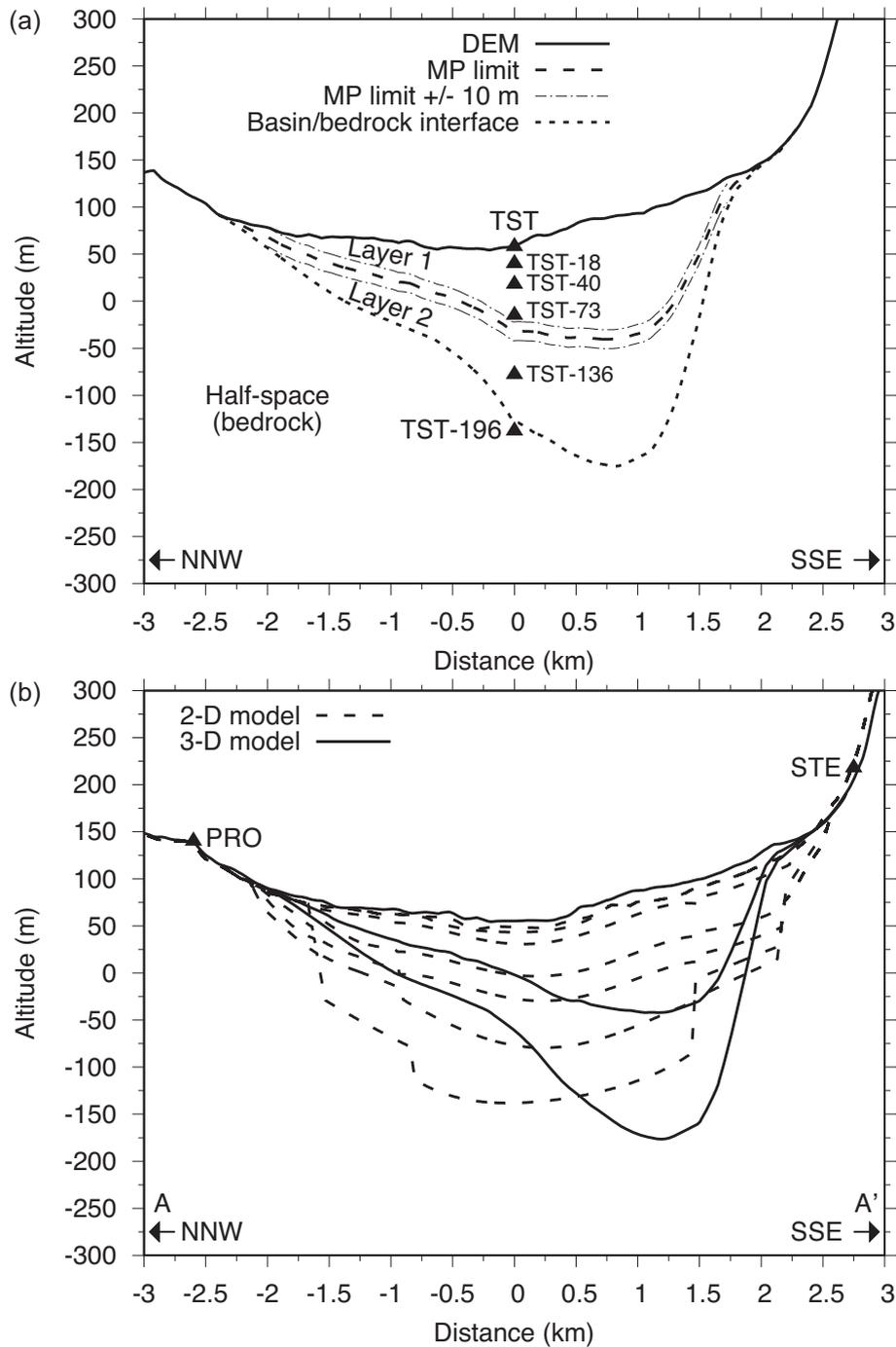


Figure 2. (a) Vertical cross-section of the 3-D model at the TST station. The velocity structure (not shown in this figure) is piecewise linear defined by values at the top and bottom of the layers 1 and 2. The half-space has a constant velocity over depth. (b) Vertical cross-section of the 2-D and 3-D models proposed by (Raptakis *et al.* 2000) and (Manakou *et al.* 2010; Guyonnet-Benaize *et al.* 2013), respectively, along the line A–A' passing through stations PRO and STE (see Fig. 1).

illuminated by a vertically incident plane shear wave. It is obviously possible to design a more general uncertainty model with a stochastic P -wave velocity but this option requires to define the uncertainty on the ratio V_S/V_P that depends on the Poisson coefficient as well as the rock lithology parameters. Indeed, for elastic linear media, the Poisson coefficient is defined by the relation $\nu = (V_P^2 - 2V_S^2)/(2(V_P^2 - V_S^2))$ and a more general model would then rely on two uncertain parameters among V_P , V_S and ν , namely (V_P, V_S) , (V_P, ν) or (V_S, ν) . In our case, a *parametric variability*

is present because only one parameter (V_S) instead of two is prescribed, nonetheless we have verified that the range of variation of the Poisson coefficient in the sedimentary layers and in the bedrock is physically admissible since $0 < \nu \leq 0.5$ for the 400+ models. The fact that the Poisson coefficient varies induces a variability on the fundamental and higher modes of Rayleigh wave phase velocities, as shown by Karray & Lefebvre (2008). This variability will be embedded into the sensitivity indices of the shear wave velocities. The Poisson coefficient could be considered as uncertain input

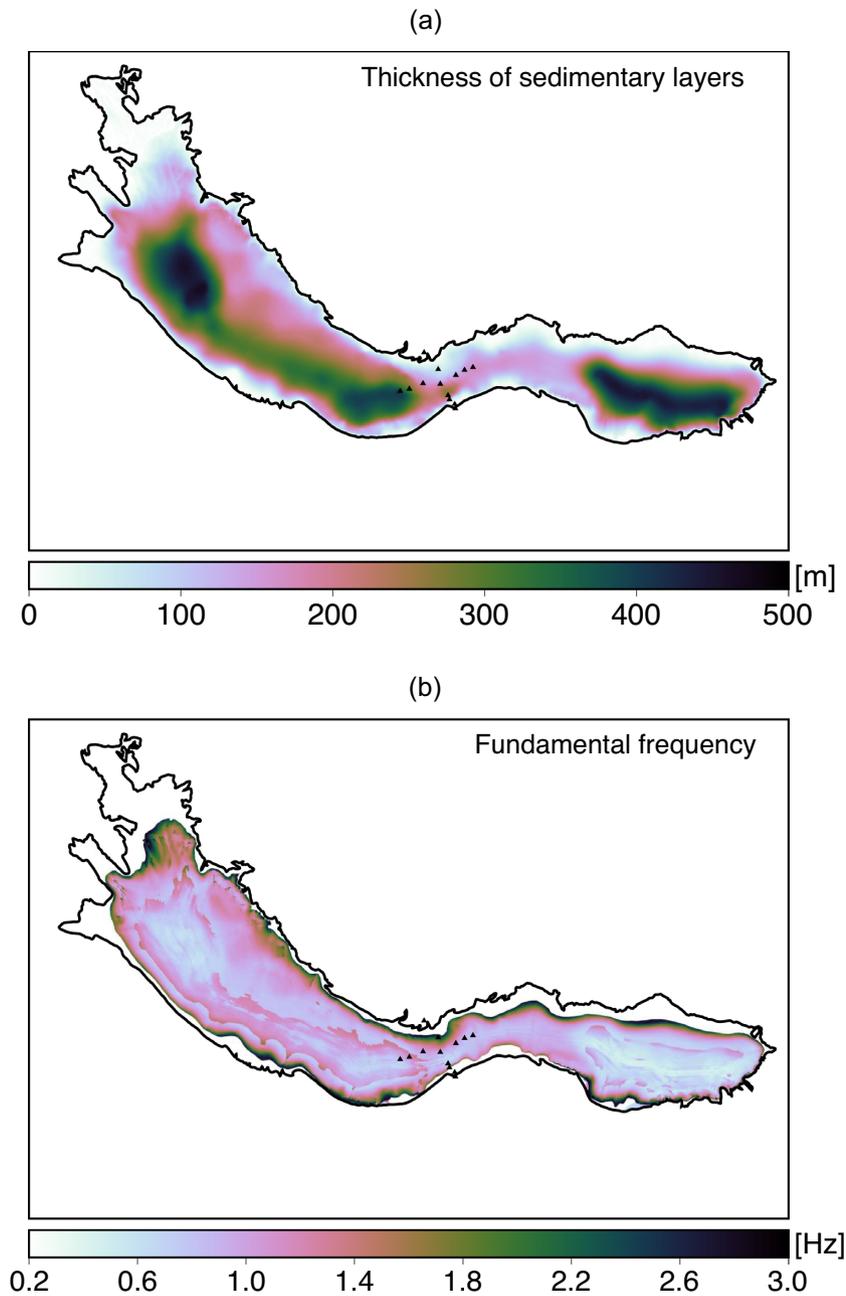


Figure 3. (a) Thickness of the sedimentary layers (L1 and L2) overlying the bedrock. (b) Fundamental frequency computed for the mean model presented in Section 2 using the u_{22} ground motion (see Section 4). Frequencies above 3 Hz are not shown because of the limit of validity of the mesh (See Fig. 10). The contour of the basin at the free surface is shown by a solid line. The stations of the permanent accelerometric network are represented by triangles. The cubehelix colour scheme is used because it is appropriate to the screen display and is monotonically increasing in terms of its perceived brightness (Green 2011; Moreland 2016).

parameter, but at the cost of adding five additional parameters into our model (four Poisson coefficients in the sediments and one in the bedrock). We preferred not to include them to ensure the convergence of the GSA, but a more complete uncertain model would require them.

The bulk densities is also deterministic because the elastic moduli present in the stress–strain relationship of the equations of motion vary linearly with them but quadratically with the wave velocity (i.e. $M_{S,P} = \rho V_{S,P}^2$). By doing this hypothesis, we prevent the uncertain model from five additional parameters (four densities in the sediments and one in the bedrock).

3.1.2 Quality factors

Another source of uncertainty is related to the medium attenuation, commonly represented by the quality factors. We consider an uncertain empirical scaling factor q_S that satisfies the rule $Q_S = V_S/q_S$ with Q_S the S -wave quality factor that we suppose constant in the frequency band of the simulations [0–3] Hz. Similarly to the Poisson coefficient, the distribution of Q_S is not prescribed and is source of *parametric variability*. However, the values of q_S are calibrated according to the values of Q_S found in the literature for the EURO-SEISTEST (see Section 3.2) in order to check the admissibility of

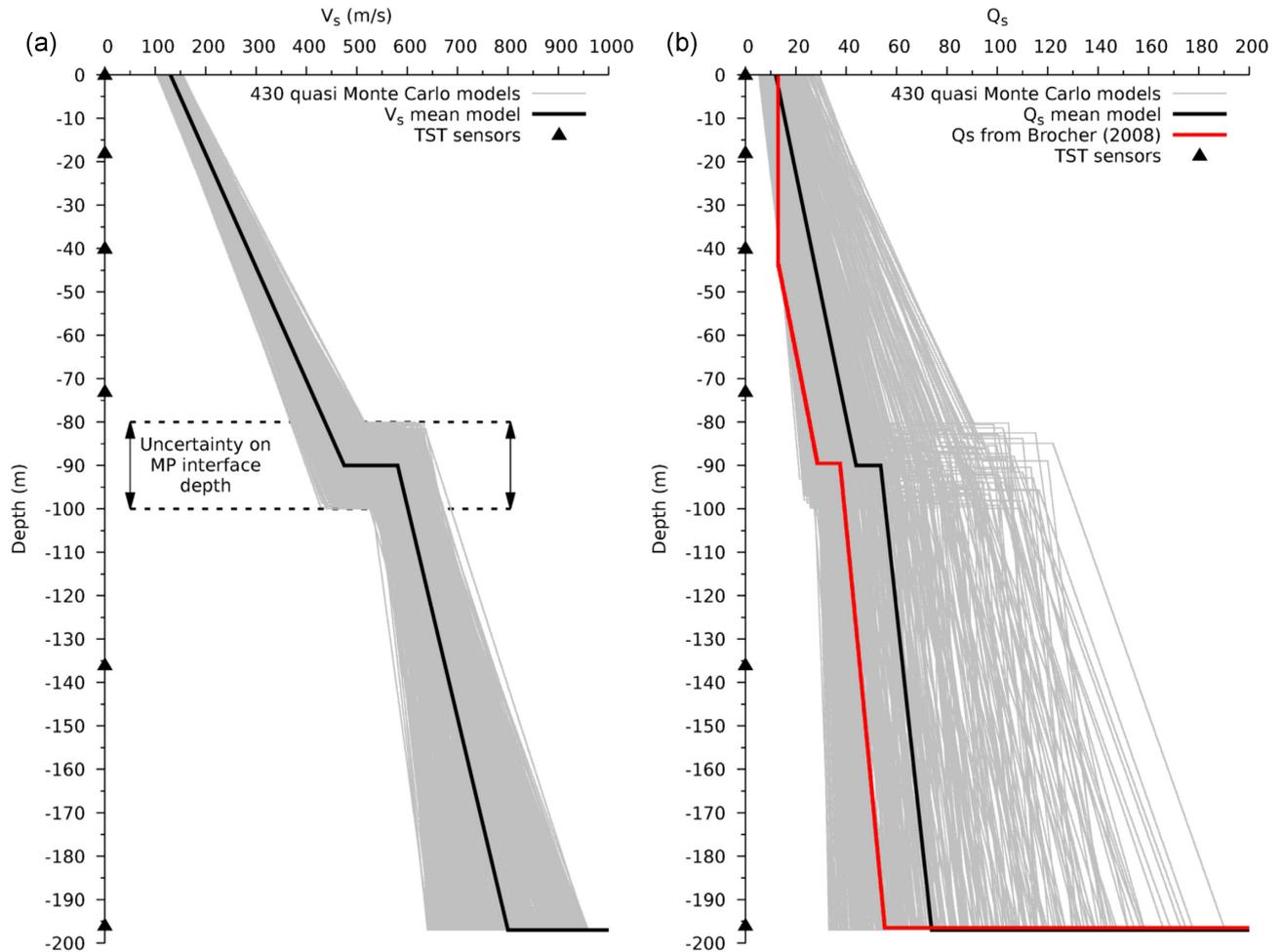


Figure 4. *S*-wave velocity structures (a) and *S*-wave quality factor structures (b) at the station TST used for the simulations. The random values of V_S are chosen by a quasi Monte Carlo method within the uncertain intervals shown in Table 1. Velocities and quality factors of the half-space are out of the maximal bound of the figure, their values are available in Table 1. The curve $Q_S = f(V_S)$ derived by Brocher (2008) (red solid line) has been computed using the V_S mean model.

Q_S values. The values of Q_P follow the simple rule $Q_P = \min(V_P/20, 2Q_S)$, as used by Maufroy *et al.* (2017). In the solver, the anelastic part of the stress–strain relationship follows the memory variables method proposed by Emmerich & Korn (1987) and the numerical implementation proposed by Ma & Liu (2006).

3.1.3 Geological interfaces

A third and last source of uncertainty is geometric and concerns the location of the MP interface z_I . A vertical translation of ± 10 m around the mean depth of the interface is introduced ($z_I = z_I^{mean} \pm 10$ m) to perturb the velocity structure (see Figs 2a and 4a). This uncertainty is relatively small compare to the sediments thickness because the MP interface is well constrained by the borehole data of the TST station. A more complex transformation (rather than just a vertical translation) could be investigated, by considering for instance a spline-based interface locally constrained by borehole data and free elsewhere, but such transformation is not considered in this paper. Although the uncertainty is relatively small compare to the sediments thickness, the group velocity of the fundamental mode of Rayleigh waves is impacted from 0.5 to 1.5 Hz (see Fig. 6). The Love waves group velocity is influenced locally

around 0.6 ± 0.1 Hz. It is also important to note that such a fluctuation of z_I modifies the geological layering configuration close to the edges of the sedimentary basin where three layering configurations are possible: [L1/L2/half-space] when $z_I = z_I^{mean}$, [L1/half-space] when $z_I < z_I^{mean}$ or [L2/half-space] when $z_I > z_I^{mean}$, as illustrated in Fig. 2(a). On the technical aspect, the geological formations L1, L2 and the half-space follow the *Onlap* and *Erode* relations (e.g. Calcagno *et al.* 2008) with L2 onlapping the half-space and L1 eroding L2. We note that the location fluctuations of the MP interface are smaller than the element size in the mesh (around 35 m, see Section 4). These relatively small fluctuations of the interface are accounted for by our solver owing to the exact integration of the velocity profile following a 1-D homogenization technique proposed by Capdeville *et al.* (2010a). The vertical translation of the MP interface, as well as the 1-D homogenization technique, can be categorized as a *model inadequacy*.

Besides, the basin/bedrock interface is not considered as uncertain because Guyonnet-Benaize *et al.* (2013) explicitly built their model to get an accurate 3-D geometry of it. The treatment of an uncertain basin/bedrock interface would deserve a particular attention of the inclination of the edges when this interface moves down. Indeed, in case of a downward vertical translation, the edges could be extended vertically (meaning that the basin's area is constant) or

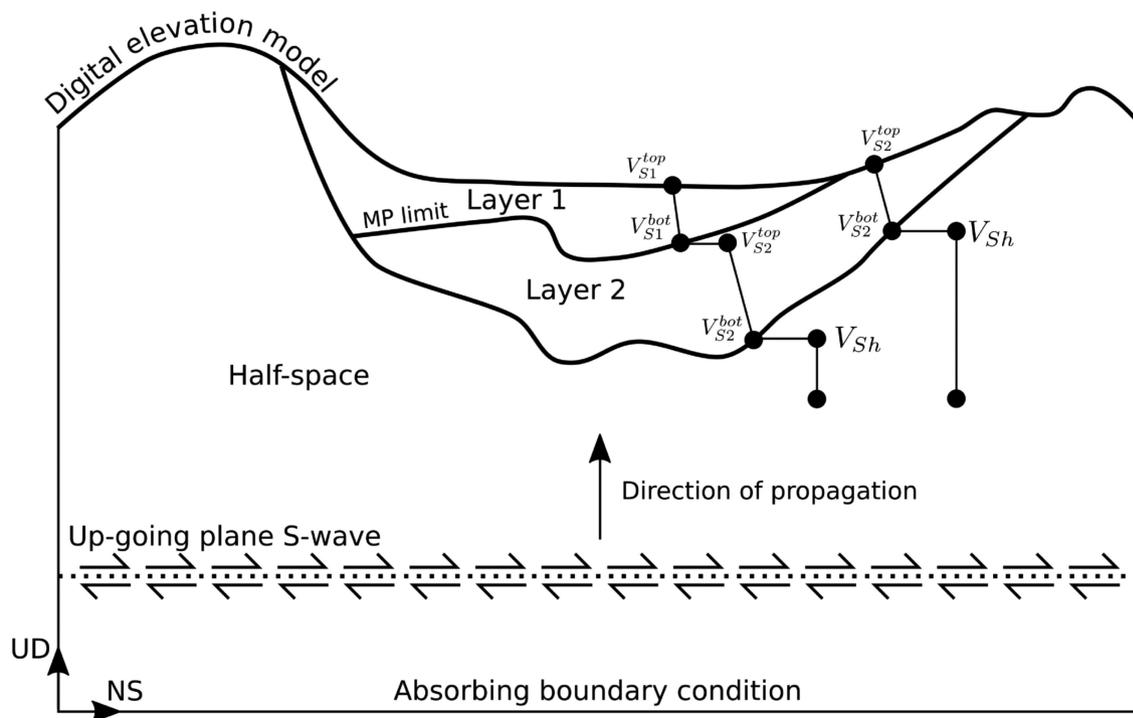


Figure 5. Schematic representation of the shear wave velocity structure. In the sedimentary basin, the profile is piecewise linear defined by values at top and bottom of layers. In the half-space, the value is constant (but uncertain).

Table 1. Mechanical and geometrical properties of layers. Values inside square brackets represent the interval of confidence of an uncertain parameter. The distribution law in the intervals is uniform. The profiles of V_S and Q_S are plotted in Fig. 4.

		V_S			V_P	ρ	Q_S	q_S			Q_P	z_I		
		Mean (m s^{-1})	Uncert. per cent	[min, max] (m s^{-1})				Mean	Uncert. per cent	[min, max]		Mean (m)	Uncert. per cent	[min,max] (m)
Layer 1	top	130	± 20 per cent	[104, 156]	1500	2075	V_S/q_S	12.5	± 60 per cent	[5, 20]	$\min(V_P/20, 2Q_S)$	-90	± 10 per cent	[-100, -80]
	bot	475	± 10 per cent	[428, 523]	2100	2130								
Layer 2	top	581	± 10 per cent	[523, 639]	2100	2130								
	bot	800	± 20 per cent	[640, 960]	2700	2250								
Half-space		2400	± 20 per cent	[1920, 2880]	4270	2500	5000				5000			

V_S , S-wave velocity; V_P , P-wave velocity; ρ , mass density; Q_S , S-wave quality factor; q_S , S-wave velocity scaling factor; Q_P , P-wave quality factor; z_I , depth of the Mygdonian/Pre-Mygdonian interface at TST station; bot, bottom; uncert., uncertainty. The Q values are assumed to be frequency independent.

with a given inclination (meaning that the basin's area is extended or reduced depending on the inclination of the edges), but both options are not realistic because they would modify the overall boundary of the basin that is generally well constrained by the outcropping bedrock. From Fig. 2(a), we can imagine that moving downward the basin/bedrock interface while keeping constant the inclination of the edges would lead to a basin extending to the mountainous areas (especially towards the SSE direction). In order to mitigate this problem, a spline-based transformation locally constrained by borehole data should be privileged to keep a fix extension (area) of the basin but allowing different angles of the edges. For instance, at the EUROSEISTEST, the gentle slope of the northern edge is more efficient to diffract local surface waves than the sharp southern edge (Chaljub *et al.* 2015, fig. 22).

3.2 Probabilistic distribution

The above-mentioned considerations led us to parametrized the uncertainty model by $N = 7$ independent random variables that can be collected in a random vector

$$\mathbf{m} := (V_{S1}^{\text{top}}, V_{S1}^{\text{bot}}, V_{S2}^{\text{top}}, V_{S2}^{\text{bot}}, V_{Sh}, q_S, z_I) \in \mathbb{M},$$

where \mathbb{M} is the parametric domain, V_{Si}^{top} (resp. V_{Si}^{bot}) the shear wave velocity at the top (resp. bottom) of the layer $i = 1$ or 2, V_{Sh} the half-space S-wave velocity, q_S the S-wave velocity scaling factor and z_I the vertical position of the MP interface (note that z_I depends on the two horizontal coordinates since the MP interface is not planar). The uncertain parameters follow a uniform distributions to produce equiprobable realizations over the parametric domain but other distributions with finite supports (e.g. beta or truncated Gaussian) could be tested in order to assess the robustness of the results with respect to the type of input distributions.

Fig. 4(a) shows the ensemble of V_S profiles at the TST station and Table 1 reports their ranges of variation. The percentage of uncertainty is set to ± 20 per cent on V_{S1}^{top} , V_{S2}^{bot} and V_{Sh} and to ± 10 per cent on V_{S1}^{bot} and V_{S2}^{top} . The Q_S profiles used in this study are shown in Fig. 4(b). As mentioned above, Jongmans *et al.* (1998) noted a large scattering on the Q_S values below 15 m. In order to take into account this scattering, the percentage of uncertainty on q_S has been set to ± 60 per cent.

Once the uncertainty model is settled, the next step is to sample the parametric domain \mathbb{M} according to the joint probability $p_{\mathbf{m}}$ of the inputs. For each type of polarization (NS and EW), the same ensemble $\mathbf{M} = \{\mathbf{m}^{(i)}\}_{1 \leq i \leq M}$ of $M = 400+$ realizations of the vector \mathbf{m} is generated by a quasi-Monte Carlo (QMC) procedure based on

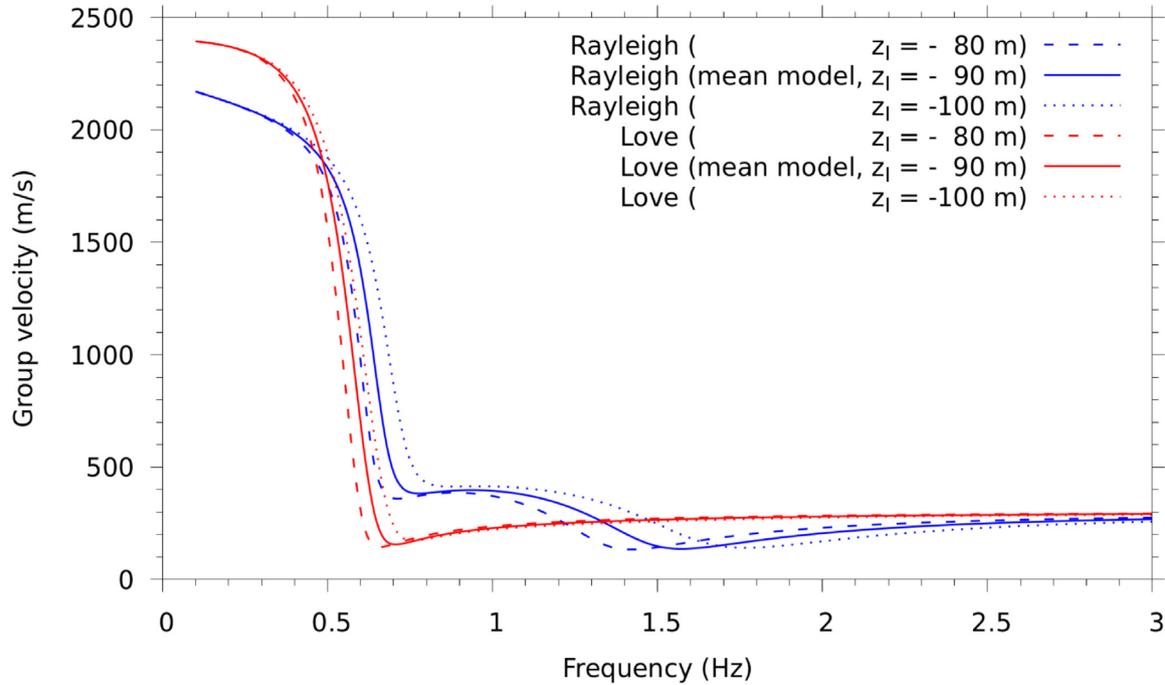


Figure 6. Influence of the depth of the MP interface (z_I) at the TST station on the group velocity of the fundamental mode of Rayleigh and Love waves computed in a 1-D layered medium with the velocity structure presented in Fig. 4.

a Sobol sequence with scrambling (Sobol 1976; Faure & Tezuka 2002). The next section introduces the numerical experimentation describing the simulations.

4 NUMERICAL EXPERIMENTATION

4.1 Overview

In the following, we use the notation $x = 1$ for the east direction, $y = 2$ for the north direction and $z = 3$ for the vertically upward direction. The ground motion $\mathbf{u}(\mathbf{x}, t)$ produced by a general three-component plane wave propagating vertically upward in the half-space is then defined by

$$u_i(\mathbf{x}, t) = u_{in}(\mathbf{x}, t) * s_n(t) \quad (1)$$

with $u_{in}(\mathbf{x}, t)$ the i th component of the ground motion due to a unit impulse plane wave polarized in the n -direction, $s_n(t)$ an arbitrary source time function in the n -direction and $*$ the convolution symbol. Through the superposition principle and the linearity of the elastodynamic wave equation, the summation convention is applied over n . For 3-D problems, i and n equal 1, 2 or 3. For instance, by omitting the space and time dependencies to simplify notations, the total ground motion along the east direction is $u_1 = (u_{11} * s_1 + u_{12} * s_2 + u_{13} * s_3)$, with u_{12} the ground motion along the east direction arising from a unit impulse plane wave polarized in the north direction, etc. We note that the u_{i3} components related to a P -wave illumination have been neglected hereafter because our study focuses on the response of a basin illuminated by vertically incident plane shear waves. We also purposely omitted the convolution with the arbitrary source time functions $s_n(t)$ to exclude source effects (e.g. frequency signatures and associated uncertainties, etc.). In the end, u_i represents the i th component of the ground motion resulting from the summation of the two unit impulse responses u_{i1} and u_{i2}

due to plane waves polarized in the east and north directions, respectively. The source time function of the plane waves is a low-pass filtered Dirac (see Section 4.3) that can be interpreted as a far-field displacement radiated by a Heaviside step function double-couple shear dislocation. This temporal relation between the source and its radiated far-field rests upon the double-couple shear dislocation theory stating that the temporal form of the far-field displacement is the time derivative of the source time function (e.g. Aki & Richards 2002).

For each QMC velocity model, we compute the unit impulse response of the EUROSEISTEST illuminated by a x - or y -polarized plane S -wave with vertical incidence. A full animation of a simulation showing $u_{22}(\mathbf{x}, t)$ is available at <http://efispec.free.fr/videos/shaheenII.mp4>. The Fig. 7 shows the ground motions u_{11} , u_{12} , u_{21} , u_{22} , u_1 and u_2 computed at the TST station. The uncertainty on the predicted motion can be appreciated from the QMC realisations depicted by grey lines: we observe a slight variability for the direct S -wave arrival (before 3 s) which contrasts with a substantial variability occurring after 3 s, when surface waves travel through the TST station. It is also interesting to note that the components u_1 and u_2 computed with the mean model (see Figs 7c and f) are similar until about 3 s (corresponding to the S -wave portion) but are very different after. This difference highlight the influence of the basin's 3-D geometry on the generation and propagation of the surface waves, despite the simple nature of the source illumination (see Section 4.3) and of the velocity structure (see Section 2.2). Finally, we also note from Fig. 7 that the total ground motion $u_i = u_{i1} + u_{i2}$ is mainly composed by its component parallel to the unit impulse input wave polarization.

The governing equations of motions are solved using EFISPEC3D⁴, an open-source parallel computer program. This

⁴Available at <http://efispec.free.fr>, co-developement at <https://gitlab.brgm.fr/brgm/efispec3d>

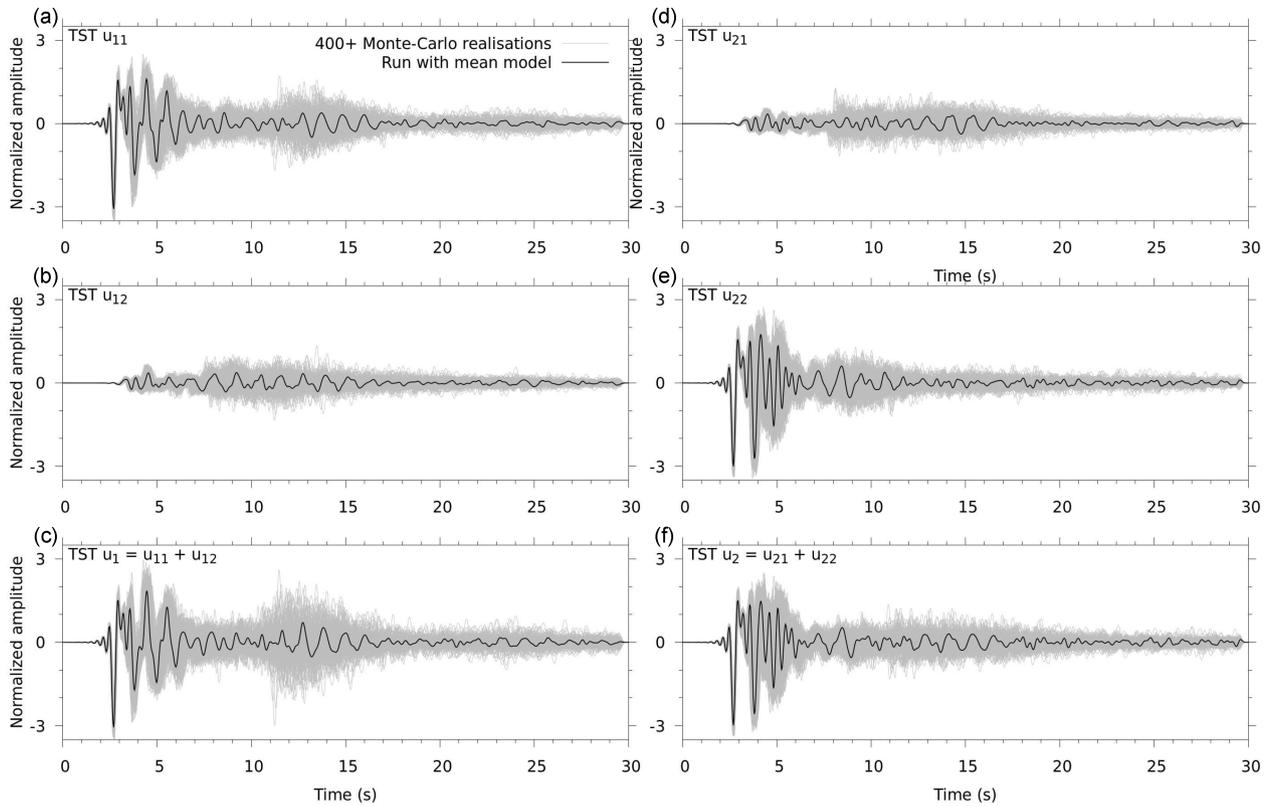


Figure 7. Ground motions u_{11} (a), u_{12} (b), $u_1 = u_{11} + u_{12}$ (c), u_{21} (d), u_{22} (e), $u_2 = u_{21} + u_{22}$ (f) computed at the TST station. Ground motions obtained from mean model (black line). Ground motions obtained from QMC models (grey lines).

solver is by default compiled in single precision (i.e. 32 bits) to solve the equations faster, without loss of accuracy. The accuracy of EFISPEC3D has been verified with international benchmarks (Maufroy *et al.* 2015; Chaljub *et al.* 2015) and semi-analytical solutions (De Martin 2011; Matsushima *et al.* 2014). Its scalability on Shaheen II supercomputer is presented in Sochala *et al.* (2020). Its most computationally intensive kernel—the computation of the internal forces—has been optimized for the single instruction multiple data (SIMD) vectorization (Jubertie *et al.* 2018; Sornet *et al.* 2018).

4.2 Mesh

Although 400+ models are generated, a unique mesh of the computational domain has been built based on the most constraining model in terms of numerical dispersion. This model is the one with the minimal S -wave velocity in the domain, that is 104 m s^{-1} (see Table 1). To ensure no numerical dispersion (e.g. De Basabe & Sen 2007) below 3 Hz, the mesh is designed with (at least) five Gauss–Lobatto–Legendre (GLL) computational points at the minimal wavelength. The size of the elements is adapted over the domain of simulation to propagate a minimal wavelength of 35 m within the sedimentary basin, which corresponds to a S -wave velocity of 104 m s^{-1} at 3 Hz. The mesh is shown in Fig. 8 and contains 7 053 889 unstructured hexahedral elements. The SEM used to solve the weak form of the equations of motion is based on fourth-order \mathbb{Q}_4 Lagrange basis functions yielding 476 426 167 GLL computational points and approximately 1.4 billion degrees of freedom (three degrees of freedom per GLL points). Regarding the time integration scheme, the

Courant–Friedrichs–Lewy (CFL) stability condition imposed by the explicit Newmark-beta scheme yields a time step equal to 0.2 ms.

The procedure for the mesh generation contains three steps: (i) A hexahedral mesh \mathcal{M}_1 of the domain of simulation with a flat free surface is generated by CUBIT (Blacker *et al.* 1994). The length of the edges of the hexahedral elements varies from 315 m at the bottom of the domain to 105 m at the free surface by using the tripling technique implemented in CUBIT. (ii) A mesh \mathcal{M}_2 is obtained by a homemade code deforming the flat free surface of \mathcal{M}_1 to fit the topography provided by a discrete elevation model. (iii) The final mesh \mathcal{M} is produced by CUBIT by refining with the tripling technique the hexahedral elements of the mesh \mathcal{M}_2 located within the sedimentary basin to obtain elements with edges length equal to 35 m.

As the mesh generation is burdensome, especially the meshing of the geological interfaces with hexahedral elements, we avoid the option to build a coincident mesh with the MP and basin/bedrock interfaces for each of the 400+ profile. By doing so, the boundaries of the hexahedral elements do not follow the interfaces and therefore discontinuities of mechanical properties are present inside the elements cut by the interfaces. This configuration violates the fundamental hypothesis of the standard finite-element method that implies continuous mechanical properties per element. To mitigate this issue, we use a 1-D vertical homogenization method that smooths the discontinuities by applying a boxcar filter over the mechanical properties. The homogenization (or upscaling) methods rely on a simple physical consideration stating that the wavefield is not (or weakly) perturbed by heterogeneities much smaller than the minimal wavelength of interest (Backus 1962). Generalization of the homogenization method applied to finite-element meshes is presented by Capdeville *et al.* (2010a, b). In our study, the boxcar

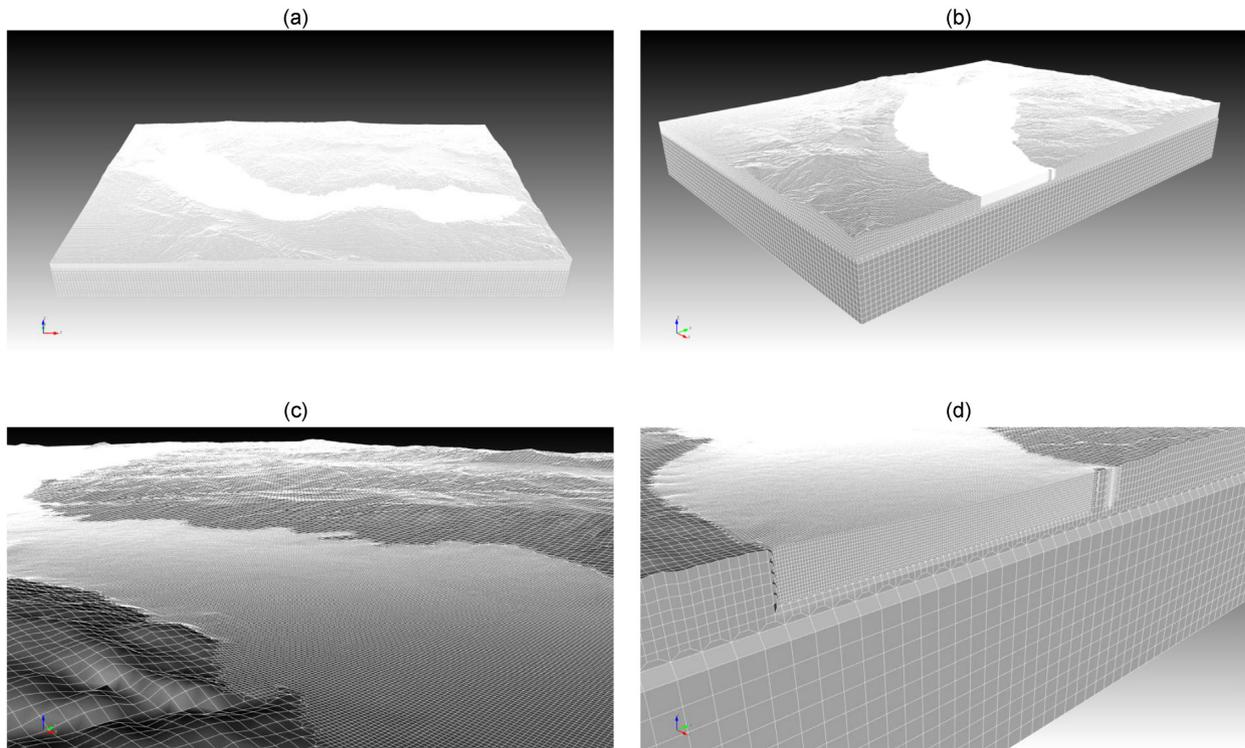


Figure 8. Mesh of the model [(a) general view, (b) zoom on the basin, (c) view of the truncated mesh and (d) zoom on the truncated mesh] designed with two refinements of the size of the elements edge: one from ≈ 315 to 105 m and one from ≈ 105 to 35 m. Within the contour of the basin, the mesh is refined over a constant depth that includes the bottommost point of the basin.

function length is equal to the minimal wavelength of our simulations (i.e. 35 m). This regularization using a 1-D homogenization technique rather than a 3-D one, can be viewed as a *model inadequacy*.

4.3 Source illumination

The seismic source is a vertically upgoing plane S -wave polarized along the x or y axis traveling in a homogeneous linear elastic half-space. To adapt the frequency band of the source to the discrete mesh, we use a unit pseudo-impulse function (i.e. a pseudo-Dirac) defined as the unit impulse response of a Butterworth low-pass filter whose time and frequency functions are shown in Figs 9 and 10, respectively. The unit source spectrum (equal to [0–5] Hz) is designed to be a bit broader than the mesh spectral validity [0.33–3] Hz to ensure a flat unit response over it.

The plane wave source is imposed over a xy -plane located at -4320 m depth via the displacement field in the first second of the time integration scheme. An absorbing condition, based on the zero-order paraxial approximation (i.e. Engquist & Majda 1977), is active at the bottommost boundary of the domain (-5040 m) to mimic a semi-infinite half-space and absorb downgoing waves. On the vertical boundaries, no paraxial approximation is implemented because their presence would lead to artificial waves propagating from the wave injection plane towards the centre of the domain (e.g. Maeda *et al.* 2017, Fig. 15b). To avoid such artificial waves generation during the plane wave injection phase, the degrees of freedom on the vertical boundaries were locked except in the direction of the shear motion by setting to zero the displacements perpendicular to the polarization of the input wave. For instance, when injecting a plane wave polarized in the x -direction, the y - and

z -displacements were set to zero on the four vertical boundaries. This technique does not absorb the waves propagating towards the vertical boundaries, but it guarantees no artificial waves generation during the plane wave injection phase that is more important for our case. Indeed, since the energy of the upgoing input wave is mostly reflected downward and scarcely reflected horizontally (see animation <http://efispec.free.fr/videos/shaheenII.mp4>), we have neglected the reflections of small amplitude surface waves at the vertical boundaries. Although this technique of plane wave injection has proven to be satisfactory for our specific 3-D case, a more careful treatment of the effects of the boundary conditions would be necessary. For instance, Maeda *et al.* (2017) showed on a 2-D example that the so-called ‘Perfectly Match Layer’ boundary condition minimizes the occurrence of artificial reflections on the vertical boundaries.

At last, for the time–frequency analysis, the ground motions at each station are delayed or advanced in time to synchronize them at the basin/bedrock interface with respect to the mean S -wave velocity of the bedrock. This procedure is implemented in order to eliminate path effect due to the stochastic nature of the S -wave velocity of the bedrock.

4.4 Output saving strategy

The ground motions are saved at the stations of the permanent accelerometric network shown in Fig. 1 and at the 12 297 150 GLL points located at the free surface of the mesh. Saving all the GLL points of the free surface allows for the reconstruction of the surface ground motion at any location during the post-processing phase. In the current paper, we only study the station of permanent accelerometric network but the use of the entire bank of ground motions is discussed in Section 6.

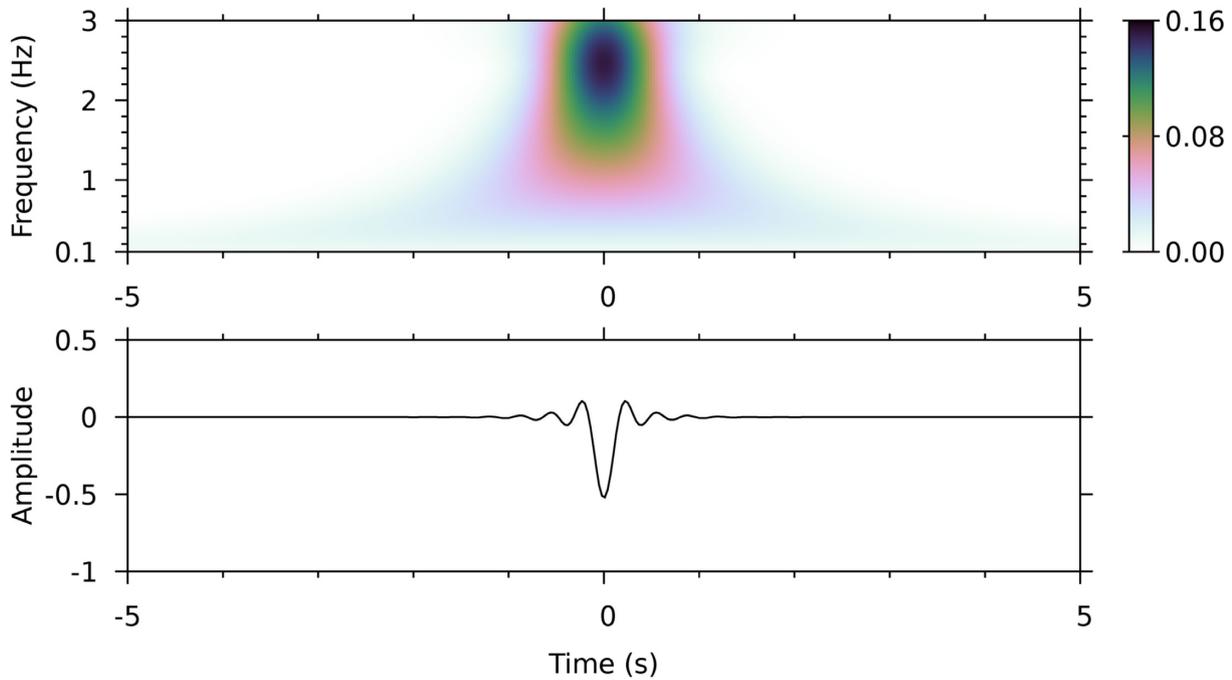


Figure 9. Source time function and its S transform envelope plotted in the time–frequency domain.

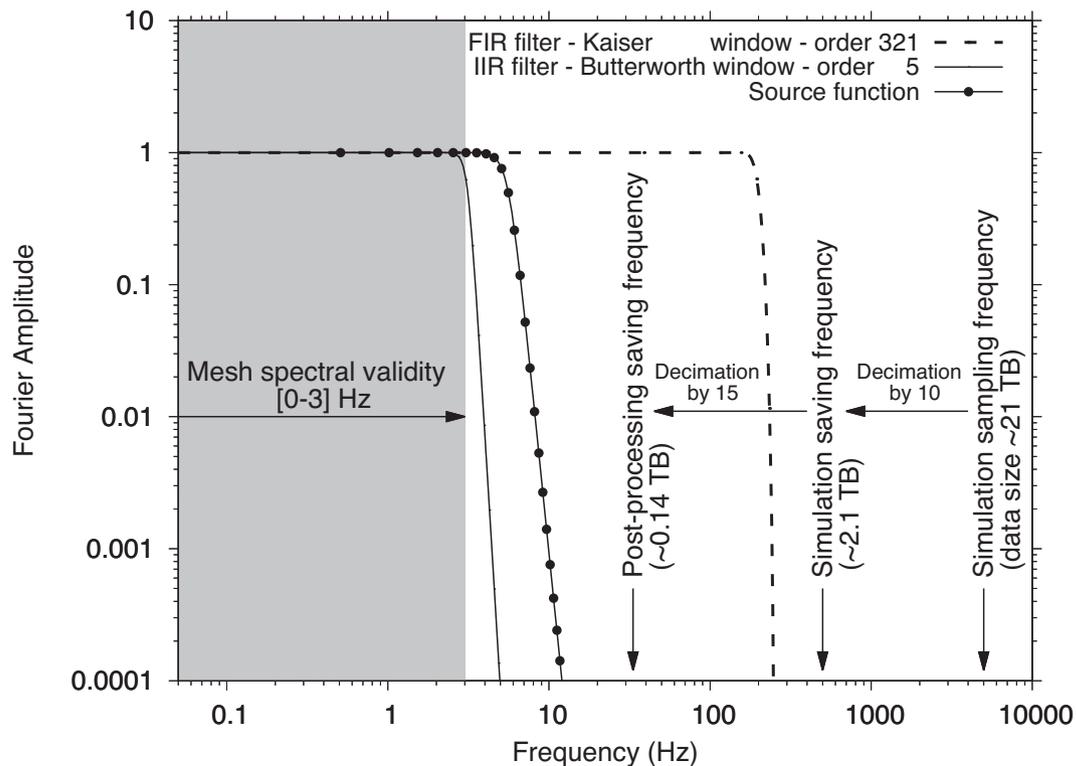


Figure 10. Fourier amplitude of the source and filters functions. Vertical arrows indicate the different sampling frequencies used in this numerical experimentation.

Saving the ground motions at the 12+ million GLL points during 30 s at a sampling frequency $F_{\text{simu}} = 5000$ Hz (the inverse of the time step) would lead to 21 TB single precision uncompressed binary data per simulation. To avoid such large amount of storage, the ground motion time-series are low-pass filtered and decimated twice: at runtime (i.e. ‘on the fly’) during a simulation and after the end of

a simulation. The runtime filter is a Kaiser window finite impulse response filter whose ideal cut-off frequency is 200 Hz. Once the filter has been applied, the time-series are decimated at runtime by a factor 10 by saving 1 every 10 samples (Fig. 10). The time-series saved at all the GLL points are saved at a sampling frequency $F_{\text{save}} = 500$ Hz, leading to a 2.1 TB file per simulation. This file is written

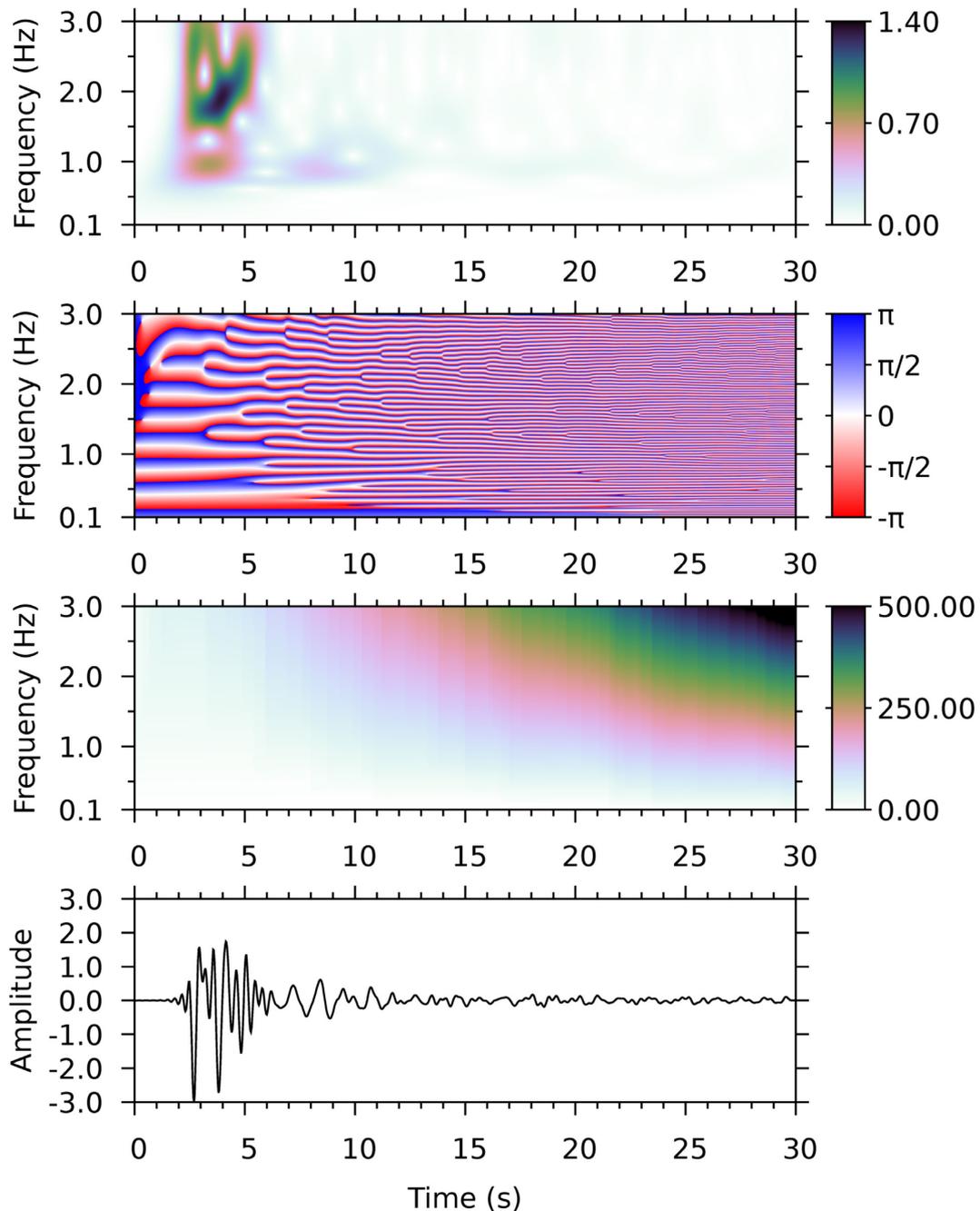


Figure 11. From top to bottom: envelope, wrapped phase, unwrapped phase of the S transform computed from the ground motion u_{22} at the TST station, for the mean soil structure of the QMC models.

in parallel at runtime on a Lustre file system for which EFISPEC3D code has been tested and optimized (Paciucci *et al.* 2016) in order to avoid performance degradation due to large amount of I/O. After the end of a simulation, the 2.1 TB of time-series are read and filtered for a second time by a Butterworth window infinite impulse response filter whose ideal cut-off frequency is 3.47 Hz. Once this filter has been applied, the time-series are decimated by a factor 15 by saving 1 every 15 samples (Fig. 10). At the end, the simulation sampling frequency imposed by the explicit time step has been divided by a factor 150 and the post-processing sampling frequency is equal to $F_{\text{post}} = 33.3$ Hz, leading to 140 GB file per simulation. In theory, the optimal frequency F_{post} is two times the Nyquist frequency ($2 \times 3 = 6$ Hz), but we noted on preliminary runs that using this frequency

underestimates the PGMs by about 10 per cent justifying the use of a more restrictive frequency that correctly estimates the PGMs.

5 GLOBAL SENSITIVITY ANALYSIS

5.1 Quantities of interest

The global sensitivity analysis is conducted on two quantities of interest: the peak ground motion computed at the 12+ million free surface GLL points and the S transform of the ground motions computed at the permanent accelerometric stations. The analysis of the PGM shows the spatial influence of the uncertain parameters,

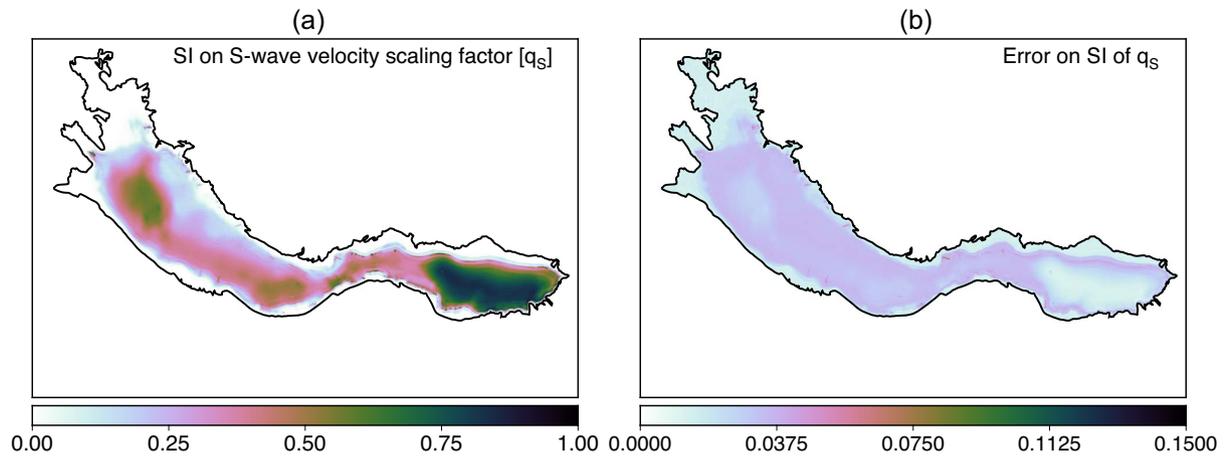


Figure 12. First-order SI on the S -wave scaling factor q_s computed on p_{22} (a) and its error (b) computed from bootstrapping. The colour scale of the error is adjusted to its maximum value (0.15). Note that outside the basin, the error equals about 0.02 but has been set to zero because the variance of p_{22} is negligible outside the basin (see Fig. 13d).

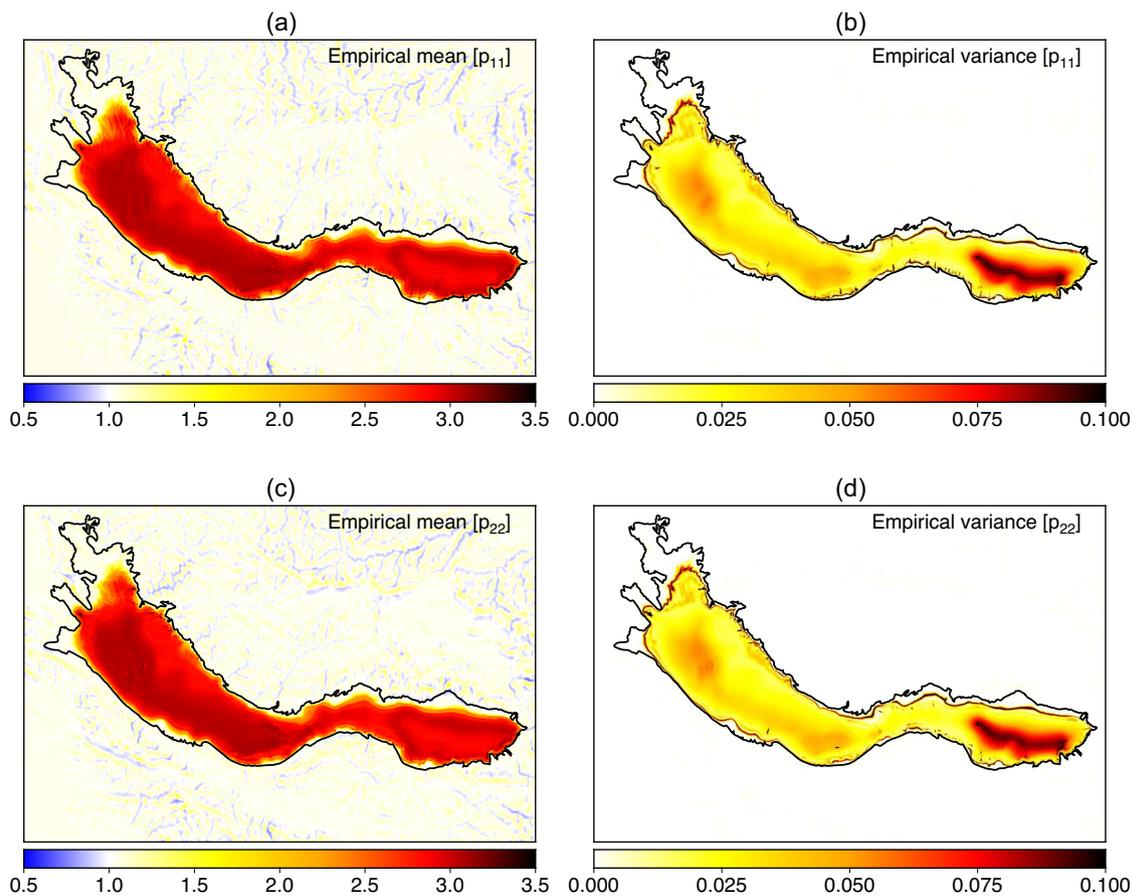


Figure 13. Empirical mean and variance of p_{11} and p_{22} . The colour scale of the mean (resp. variance) has been clipped at 3.5 (resp. 0.1) to mask high values (black areas). The maximum of the mean (resp. variance) is around 4.0 (resp. 1.0) for p_{11} and around 4.6 (resp. 1.4) for p_{22} .

while the analysis on the S transform depicts their influence on a time–frequency representation of the ground motion and will help to understand their influence on the different type of waves composing the ground motion.

The PGM field, $p_{in}(\mathbf{x})$, is defined as the absolute maximum displacement $u_{in}(\mathbf{x}, t)$ over the simulation time $T = 30$ s, that is

$$p_{in}(\mathbf{x}) = \max_{0 \leq t \leq T} |u_{in}(\mathbf{x}, t)|.$$

The S transform is defined as a continuous wavelet transform with a specific mother wavelet (the Gaussian function) multiplied by the phase factor $\exp(-i2\pi ft)$. The S transform of a function $u(t)$ is defined by Stockwell *et al.* (1996) as

$$S_u(\tau, f) = \int_{-\infty}^{\infty} u(t) \frac{|f|}{\sqrt{2\pi}} \exp\left(-\frac{(\tau - t)^2 f^2}{2}\right) \exp(-i2\pi ft) dt.$$

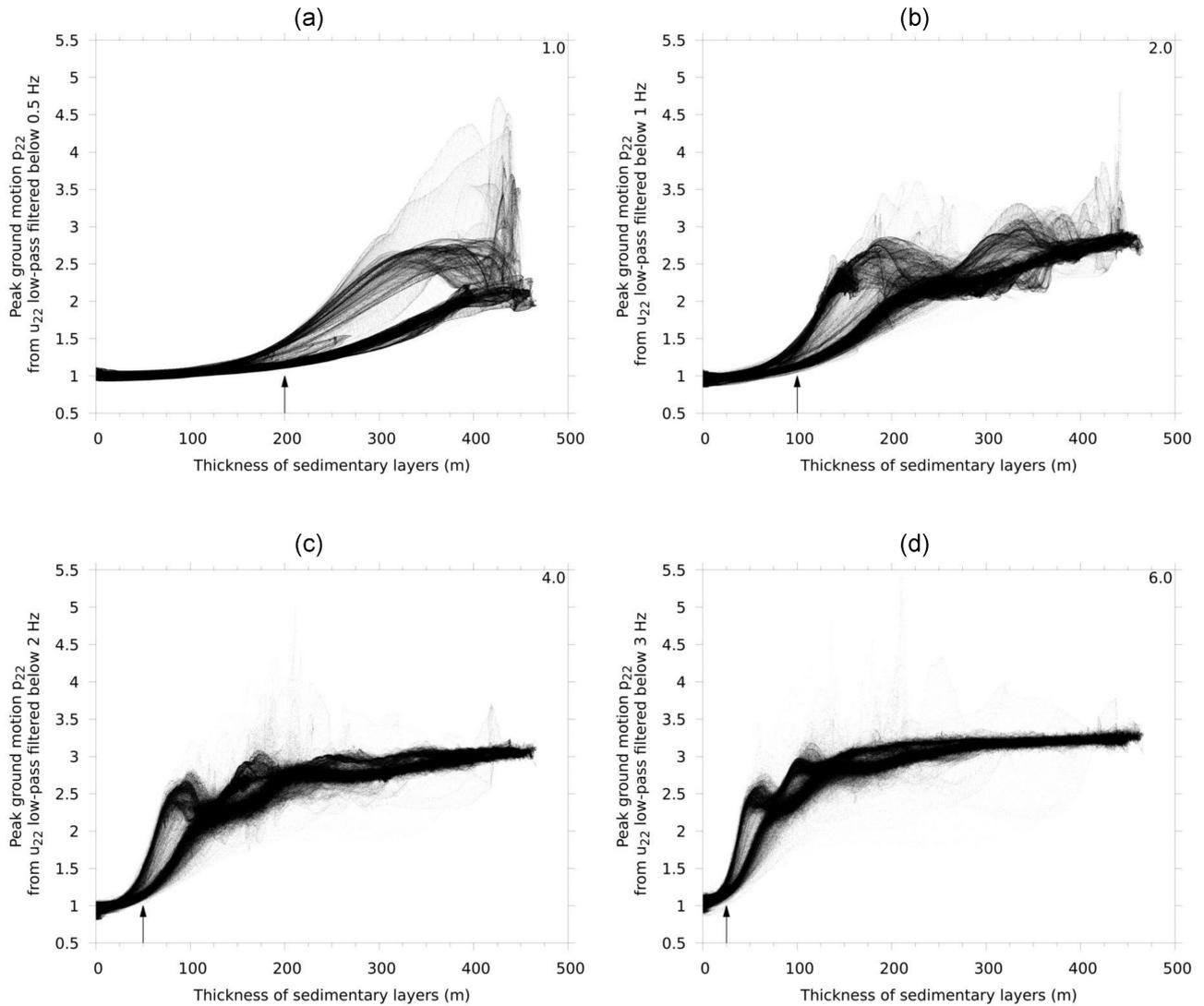


Figure 14. Normalized peak ground motions p_{22} (≈ 12 million points per panel) computed from low-pass filtered ground motion u_{22} of realization $n^\circ 1$ plotted versus the thickness of the sedimentary layers. Normalization has been done using the theoretical (filtered) peak ground motion on flat outcropping bedrock. The number at the top-right of each panel indicates the ratio between a theoretical filtered peak ground motion with the one filtered at 0.5 Hz. An upward arrow shows the PGMs amplification onset.

The source time function and its S transform envelope are plotted in Fig. 9. At the TST station, this input motion turns into the ground motion shown in Fig. 11 due to the presence of the sedimentary basin. The S transform envelope shows a first energetic wave train in the time–frequency box [2–3] s [1.4–3.0] Hz corresponding to the upgoing S wave. The second wave train at [3–6] s [1.4–3] Hz is more complex since it contains both the surface waves generated by the edges of the basin and the upgoing body wave already reflected one time at the bottom of the basin. After 6 s, the envelope amplitude is much smaller and corresponds to the coda waves.

The phase of the S transform (Fig. 11) presents strong discontinuities due to (purely mathematical) jumps around $\pm\pi$. Because of these jumps, we choose to unwrap the phase following the algorithm proposed by Itoh (1982) in order to prevent the GSA from artificial discontinuities. After applying the unwrapping algorithm, the phase is represented by a smoother function as illustrated in Fig. 11. Nonetheless, the phase unwrapping problem is a ill posed, inverse problem (e.g. Ghiglia & Pritt 1998), thus we have tested several algorithms (including the minimum cost flow algorithm, the Flynn’s

min discontinuity algorithm, the unweighted multigrid algorithm, the residues and branch cuts algorithm, as well as the minimum L_2 norm algorithm) but without substantial improvement over the Itoh algorithm that seems well suited to our case, probably because the synthetic signals are free of noise.

5.2 Preamble about sensitivity indices

Before physically interpreting the sensitivity indices (SI) in the next Sections, we recall that these indices are commonly used to rank the effect of an input parameter or group of input parameters on an output. By definition, each index belongs to [0,1] and the sum of all the indices is equal to one. When the sum of the first-order SI for an output equals one, then all the sensitivity is governed by the own effect of each parameter. When the first-order sensitivity index S_i of an input parameter i equals one, this parameter controls all the sensitivity of the output and all other SIs equal zero. The interaction effects can be quantified with higher order indices and are globally

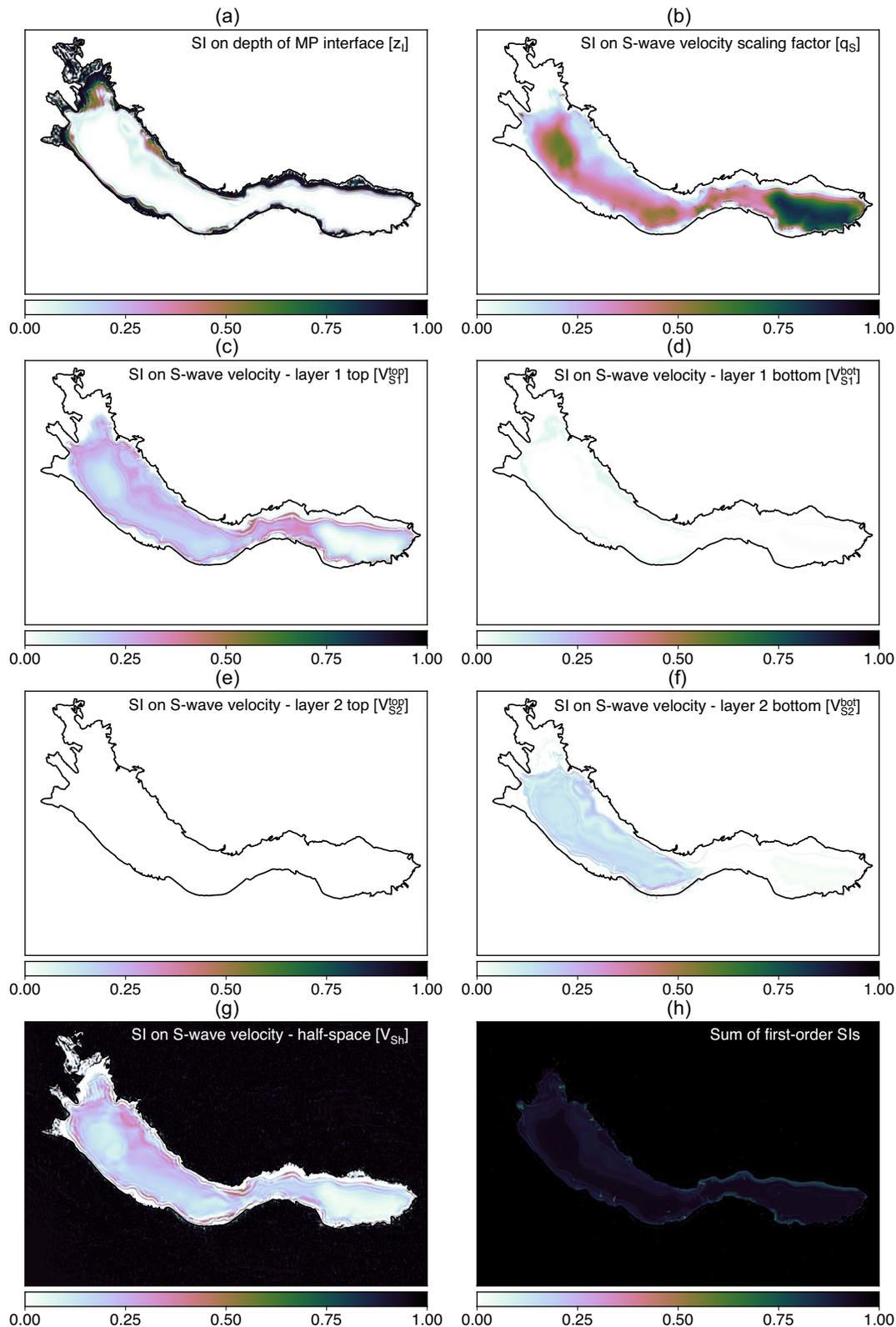


Figure 15. First-order sensitivity indices and their summation for the PGM. A value close to zero (white) indicates that a parameter has no influence on the PGM.

measured by $1 - \sum_{i=1}^N S_i$. In this study, we compute the first-order SIs from the sample-based method proposed by Li & Mahadevan (2016) (see the appendices for more details). We found that this sample-based method leads to almost similar first-order SIs for the

PGM as those derived by a polynomial chaos surrogate presented by Sochala *et al.* (2020).

To check if the ensemble of 400+ simulations is large enough to correctly estimate the first-order SIs with a low number of samples,

we have computed a confidence interval from bootstrapping (e.g. Efron & Tibshirani 1994). The principle of the bootstrap method is to generate R replicas $\{\mathbf{M}_r\}$ by resampling the ensemble \mathbf{M} with replacement, and then to compute the mean squared error between the moment estimator (mean or variance) computed from \mathbf{M} and its replicas obtained from \mathbf{M}_r . This technique thus provides an estimations of the variability due to the finite nature of the sampling. If the bootstrap error exhibits low values, the sample size to estimate first-order partial variances is large enough. In this study, small bootstrap errors computed from $R = 50$ replicas have been observed for all the quantities of interest (maps of PGM as well as the envelope and phase of the S transform) as illustrated by Fig. 12 where the first-order SI map of the S -wave scaling factor q_S is plotted with its error. We observe that the bootstrap error is much smaller than the SI itself meaning that the *code uncertainty* is here innocuous since the size of the QMC ensemble is large enough to capture the main variations of the quantities of interest.

5.3 Peak ground motion analysis

5.3.1 Mean and variance

We start by the analysis of the p_{ii} extracted from u_{ii} . The analysis of the total PGM $p_i := \max_i(u_{i2} + u_{i1})$ is not presented because the summation of the 400+ NS and EW-polarized simulations has not been performed yet since it requires to manipulate the two sets of 400 files \times 140 GB. We can speculate from Fig. 7 that p_i and p_{ii} shall be quite similar over the free surface.

The empirical mean and variance of p_{11} and p_{22} computed from the 400+ ensemble are represented in Fig. 13. The values are normalized by the theoretical bedrock outcropping motion on a flat free surface, therefore, no amplification of the motion are represented by values equal to one, de-amplification by values smaller than one and amplification by values greater than 1. The map of the mean shows that the PGMs are mostly amplified by a factor of three inside the basin, but higher amplification values (up to four) are reached at some specific locations. The mean and variance of p_{11} and p_{22} present similar general patterns. In details, we see some differences close the basin's edges because the generation of basin-induced Rayleigh or Love waves depends on the polarization of the input motion with respect to the orientation of the edges. The variance maps show that the PGMs inside the basin can have very different spatial variabilities. Inside the basin, a variance of 0.025 (standard deviation $\sigma \approx 0.16$) is predominant except in the eastern side of the basin where the variance is close to 0.1 ($\sigma \approx 0.32$). Close to the edges, where the generation of surface waves occurs, the variance of the PGM can reach 1.4, meaning that the standard deviation ($\sigma \approx 1.2$) has the same order of magnitude as the mean values at these locations. The values of $\sigma \approx 0.16$ found in our numerical experimentation are similar to those reported by Iwaki *et al.* (2018) who have investigated the variability of the the peak ground velocity in heterogeneous media. Outside the basin, amplification and de-amplification by the topography is clearly seen.

The scatter plots of Fig. 14 highlight the variability of the PGMs with respect to the thickness of the sedimentary layers. The p_{22} values have been extracted from u_{22} low-pass filtered at 0.5, 1, 2 and 3 Hz. When filtered below 3 Hz, the PGMs increase with the thickness of sediments but saturate around the value of 3.2 for thicknesses larger than 200 m. Such saturation of waves' amplitude with respect to the thickness of sediments has also been reported for Love waves in the Kanto basin where Yoshimoto & Takemura (2014) observed

that the fundamental Love wave's amplitude is similar among the different structure models for deep bedrock depths (>4.0 km). From these scatter plots, we can also observe a relation between the onset of the amplification of the PGMs and the thickness of sediments. When filtered below 3, 2, 1 and 0.5 Hz, the amplification starts from about 25, 50, 100 and 200 m, respectively. Those thicknesses correspond roughly to one tenth of the mean S -wavelength propagation in the basin ($\lambda_{\text{mean}} = V_S^{\text{mean}}/f_{\text{mean}}$). At 3 Hz: $\lambda_{\text{mean}}/10 \approx 500/1.5/10 \approx 33$ m, at 2 Hz: $500/1.0/10 \approx 50$ m, at 1 Hz: $500/0.5/10 \approx 100$ m and at 0.5 Hz: $500/0.25/10 \approx 200$ m.

5.3.2 Sensitivity indices

The map of first-order SIs are presented in Fig. 15. We see that the scaling factor (q_S) induces a significant fraction of the PGMs variance within the deepest part of the basin. This fraction reaches 0.8 (over a maximum of 1, see Section 5.2) on the eastern side of the basin where the variance is particularly large. The importance of the scaling factor over the PGMs is not surprising as it modifies the energy dissipation of the medium, with a direct impact on the amplitude of the ground motion. These values around 0.8, obtained here from 3-D simulations, are similar to the ones found by Sochala & De Martin (2017) in the 1-D case at the TST station using the uncertain soil profiles presented in Table 1.

The first-order SI associated with the depth of the MP interface (z_I) mainly controls the uncertainty at the edges of the basin, consistently with the fact that this parameter controls the geological strata at the border of the basin (either L1/L2/bedrock, or L1/bedrock, or L2/bedrock, see Fig. 5). In contrast, the impact of this parameter quickly drops to zero inside the basin. Recalling that the variance of the PGMs is low at the boundary of the basin, the effect of the uncertainty in the MP interface location is overall quite limited. This point is discussed in Section 6.

The next two influential parameters are the S -wave velocity at the top of the first layer, V_{S1}^{top} , and bedrock velocity V_{Sh} . The first-order SI of these two inputs are similar in their magnitude and spatial pattern. These parameters are both responsible for 20–25 per cent of the variance, except on the eastern side of the basin where the influence is mainly governed by the scaling factor q_S . The velocity at the bottom of the second layer, V_{S2}^{bot} , exhibits a similar influence but with slightly lower values. The velocities at the bottom of the first layer, V_{S1}^{bot} , and the one at the top of the second layer, V_{S2}^{top} , are much less influential.

It is important to note that the sum of the first-order SIs is mostly higher than 0.8 inside the basin (see Fig. 15h), meaning that the own effects control 80 per cent of the PGM. The sensitivity indices being additive (see Section 5.2), this result means that the PGMs are 80 per cent controlled by a summation of own effects.

5.4 S transform analysis

5.4.1 PRO station (on top of bedrock)

Fig. 16 shows that the envelope and phase of the ground motion $u_2 = u_{21} + u_{22}$ at the PRO station are mainly influenced by the bedrock velocity V_{Sh} (note that the SIs of the other parameters are not plotted because their values are lower than 0.1). This influence is particularly visible on the S -wave train around 2.5 s, in all the frequency band. Moreover, the first-order SI of V_{Sh} is almost everywhere (in the time–frequency space) equal to the sum of the first-order SIs, meaning that the own effects of the other parameters are negligible;

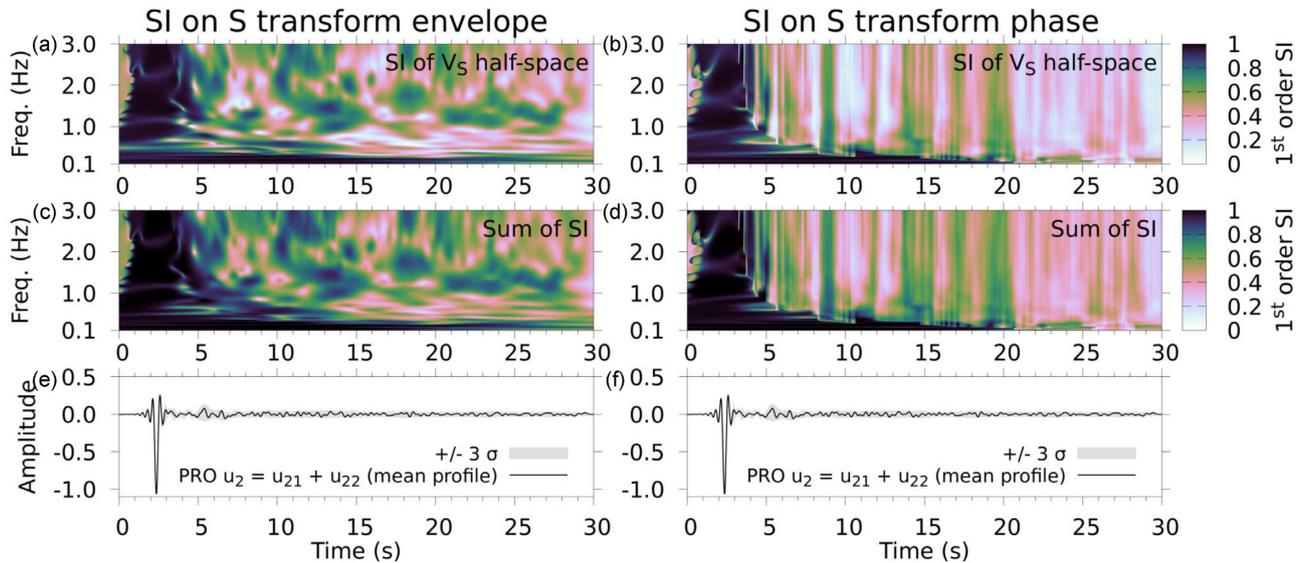


Figure 16. First order sensitivity indices and sum of the indices computed on the envelope (left-hand panels) and phase (right-hand panels) of the S transform at the PRO station installed on rock. The bottommost panels shows $u_2 = u_{21} + u_{22}$ computed at this station.

in other words, when the first-order SI of V_{Sh} is lower than one, interaction effects are involved in the control of the ground motion. Although the main influence of V_{Sh} was expected because the PRO station is located on top of the bedrock, it shows the relevance of the GSA to determine important and negligible parameters, as well as the proportion of the own and interaction effects controlling the ground motion.

After about 6 s, the bedrock velocity V_{Sh} is not the sole predominant parameter in the full frequency band (see time–frequency locations after 6 s in Fig. 16 where the first-order SI of V_{Sh} is lower than 1.0), meaning that the ground motion at the PRO station is also controlled by the interaction effects. Is it also interesting to note that below 0.5 Hz and up to 15 s, V_{Sh} is the unique parameter controlling the ground motion because its first-order SI is close to one. Besides, strong influences of V_{Sh} are also seen at specific time–frequency spots where the SI of V_{Sh} is larger than 0.8, for instance around 14 s and 1 Hz.

5.4.2 TST station (on top of sediment)

Regarding now the TST station, the first-order SIs of $u_2 = u_{21} + u_{22}$ are shown in Fig. 17 and those of $u_1 = u_{11} + u_{12}$ and $u_3 = u_{31} + u_{32}$ are presented in the Supporting Information (see Figs S1 and S2). For the three components (u_1 , u_2 and u_3) and similarly to the PRO station, the velocity of the half-space V_{Sh} still has a major influence, especially below 0.5 Hz. Nonetheless, the depth of the MP interface z_I , the S -wave scaling factor q_S or the S -wave velocity at the top of the first layer V_{S1}^{top} influence the ground motions as well because the TST station is located on top of the sedimentary layers.

More specifically, for the u_1 and u_2 components (see Fig. 17 for u_2 and Fig. S2 for u_1), the depth of the MP interface z_I influences the first 7 s of the envelope and the first 5 s of the phase of the motion. These time frames contain the first arrival of the S wave (2–3 s) and the first arrival of the surface waves (3–7 s). Its influence is particularly visible on the S -wave arrival around 2 s at 1 Hz (SI \approx 40–60 per cent) and on the fundamental mode of the surface waves around 3–7 s at 1.25 Hz (SI \approx 60 per cent). This frequency of 1.25 Hz highlighted by large values of the first-order SI of z_I is

coherent with the variability of the group velocity of the Rayleigh wave with respect to z_I , as shown in Fig. 6. After 7 s from which the waves could be considered as coda waves, the first-order SI of z_I shows a lesser influence of the MP interface. A complementary understanding of the contribution of edge-diffracted surface waves could be done using advanced seismic array processing technique (e.g. Imtiaz et al. 2021).

Another important influential parameter is the S -wave velocity scaling factor q_S which clearly governs the envelope of the S transform of the three components (see Figs 17c, S1c and S2c). A particularly strong effect of q_S is observed between 2 and 5 s for frequencies higher than 1 Hz. A clear influence is also seen on the coda-waves, between 9 and 13 s.

The S -wave velocity at the top of the first layer V_{S1}^{top} has also an influence on the envelope and phase of the motion before 10 s (see Figs 17e, S1e and S2e). For the u_2 component (Fig. 17e), this velocity seems to control the second and third resonant frequencies of the S -wave portion visible around 2.5 s at 1.5 and 2.5 Hz (Maufray et al. 2015).

Finally, as for the PRO station, we observe at the TST station that the sum of the first-order SIs on the envelope (Figs 17i, S1i and S2i) and phase (Figs 17j, S1j and S2j) are close to one before 5 s and decrease over time. It means that the first arrival of the S wave and the first arrival of surface waves (0–5 s) are mainly controlled by the own effects of the parameters while the coda waves (5 s and later) are mainly governed by interaction effects.

6 DISCUSSION

6.1 About the GSA

The implementation of a global sensitivity analysis raises a number of questions, chief among them concerns the modelling of the uncertainty sources. We reiterate that the construction of the uncertainty model is crucial when performing a GSA since the sensitivity of the outputs depend on the choice and number of the uncertain independent parameters as well as their intervals and distributions. Indeed, a change in the definition of an uncertain parameter may

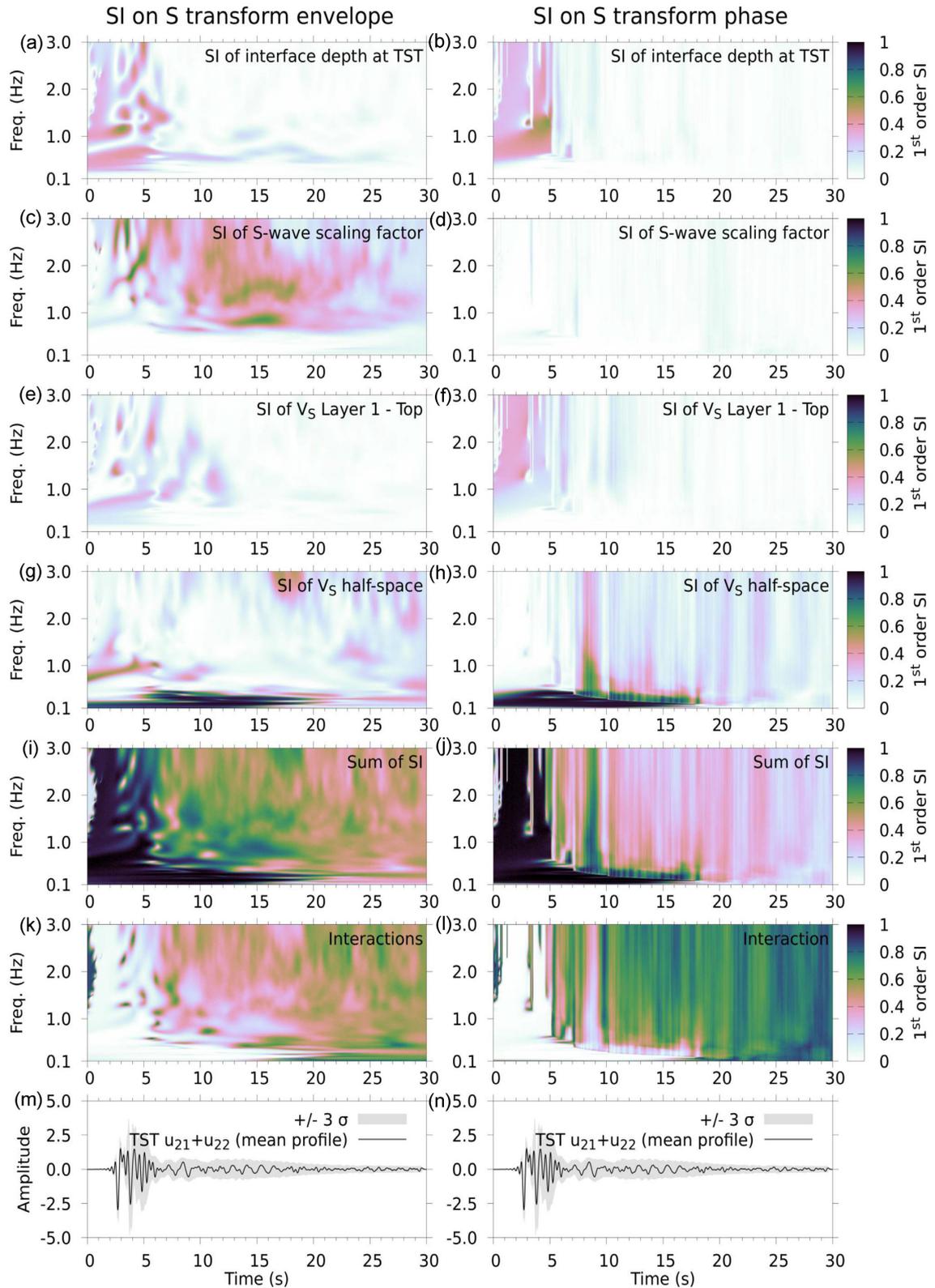


Figure 17. First order sensitivity indices and their sum and interaction computed on the envelope and phase of the S transform of $u_2 = u_{21} + u_{22}$ at the TST station installed on sediments. The ground motion u_2 of the mean profile is plotted with its standard deviation (panels m and n).

alter the sensitivities, and therefore it would be valuable for our study to repeat the GSA with different uncertainty models.

A second point to be aware of is the interpretation of the sensitivity indices that is only valid for the chosen parametric domain

and cannot be extrapolated for different supports and probabilistic laws. For instance, we observe that the depth of the MP interface z_I has a minor influence on the coda waves in our study because its fluctuation is small (± 10 m) compare to the mean S -wavelength in

the basin (let say approximately $500/1.5 \approx 333$ m). Unfortunately, the current GSA does not say what it would be if the interval of fluctuation was larger.

A third important feature to keep in mind when dealing with a quantity of interest which depends on a physical variable (space, time or frequency) is the local aspect of the normalization in the computation of the SI. In other words, the partial variances are normalized by the total variance that can strongly vary with respect to the physical variable and the interpretation of the sensitivity indices must therefore be done jointly with those of the partial variances. For example, the analysis of the sensitivity indices from 0 s to 2 s is of little interest in our case since the total variance of the ground motion is close to zero (see for instance the first 2 s in Fig. 17m).

Last but not least, the choice of the uncertain parameters is not fully obvious and one may want to explore the effects of new parameters derived from the already chosen ones. For instance, is it common in earthquake ground motion prediction to consider the impedance contrast between the bedrock and the sediments. In our case, this contrast is defined as $Z := (\rho_h V_{sh}) / (\rho_2^{bot} V_{S2}^{bot})$ and is uncertain via V_{S2}^{bot} and V_{sh} . It is possible to choose Z instead of V_{S2}^{bot} and V_{sh} in the current GSA but with full knowledge of the fact that the probability distribution of Z , constructed as a ratio of uniform distributions, is not uniform. The derivation of new uncertain parameters was not investigated in this paper, but Z , V_{S30} , Q_S30 or the inclination of the incident wave would be pertinent candidates.

6.2 Completion of the uncertainty model

As already mentioned, the GSA presented here does not consider the source effects or the path effects as uncertain, thus limiting the number of uncertain parameters and guarantying accurate results in regards of the affordable computational cost. Nevertheless, Maufroy *et al.* (2017) showed at the EUROSEISTEST a pronounced north-south asymmetry of both amplification and duration lengthening caused by non-isotropic excitation of surface waves at the basin edges. The current assumption of plane wave with vertical incidence likely underestimates these amplifications and elongations of basin-induced surface waves and would deserve to be relaxed. In that case, the depth, the azimuth and the focal mechanism of a point source would be relevant additional uncertain parameters. Kinematic extended fault with non-uniform random slip distribution could also be considered. This would increase the number of uncertain parameters because the subfaults as well as the rupture velocity shall be considered uncertain, nonetheless, such a GSA would judge the relative importance between the source and site effects. Concerning the source time function, any temporal structure can be changed retroactively (by a convolution with the desired spectrum).

6.3 Perspectives

A future seismic hazard assessment (SHA) in the tectonically active graben of Mygdonia shall rely on the GSA presented here. Although common SHA parameters like V_{S30} are not included here, other parameters, such as the S -wave quality factor, the S -wave velocity of the bedrock or the thickness of sediments (at least up to 200 m) shall be considered for predicting ground motion amplifications.

Finally, the results discussed here exploits 200 GB of data, which represents 0.14 per cent of the whole data generated during the numerical experiments. This fraction is small but the goal of present

work is to highlight the potential of a GSA in order to identify the most influential parameters governing the site effects. The extension of the time–frequency analysis to the 12+ million ground motions over the entire spatial domain is challenging because it necessitates to post-process and analyse 120 TB of data.

7 CONCLUSION

The global sensitivity analysis presented in this study helps to decipher what controls the ground motion stemming from an uncertain geological model. It allows not only the identification of the controlling parameters, but also the quantification of their relative importance owing to the derivation of the variance-based sensitivity indices. Some of the influential parameters may be intuitively known (as seen at the PRO station) but other are less obvious and can only be revealed by such analysis. Although our study is specific for the EUROSEISTEST illuminated by a plane S -wave travelling vertically upward, general tendencies emerged from the ground motions analysed on top of sediments:

Concerning the PGMs analysis, they are mainly controlled in the centre of the basin by the S -wave quality factor (Q_S). The first-order SI associated to Q_S quantifies its influence around 40–60 per cent in the eastern side of the basin and up to 70–80 per cent in the western side of the basin. This conclusion is coherent with the one stated by Moczo *et al.* (2018) concerning the influence of the attenuation on the amplification. The GSA also shows that the S -wave velocity of the bedrock (V_{sh}) and the S -wave velocity at the free surface (V_{S1}^{top}) should not be neglected to predict the PGMs. We also found that the sum of the first-order sensitivity indices on the PGMs is mostly higher than 0.8 inside the basin, meaning that the own effect of the parameters control 80 per cent of the PGM. In other words, coupling two, three, up to the seven parameters all together only represent 20 per cent of the total influence on the PGMs. Our simulations also confirm that the PGMs are correlated with the thickness of the sediments with a saturation around three occurring above 200 m of sediments for the PGMs extracted from ground motions low-pass filtered below 3 Hz. We also observe that the amplification of the PGMs starts from about one tenth of the mean S -wavelength.

Concerning the S transform analysis, the direct S -wave train of the ground motion is almost fully controlled by the own effect of the parameters, the most influential being the S -wave quality factor (Q_S), the depth of the MP interface (z_l), the S -wave velocity at the free surface (V_{S1}^{top}) and the one of the bedrock (V_{sh}). This conclusion had been already stated in 1-D at the TST station by Sochala & De Martin (2017). The similarity between the 1-D and 3-D results probably comes from the source illumination and we can speculate that the interaction effects would increase for inclined plane waves or point sources arbitrarily distributed due to the 3-D nature of the site effects. We also found that the interaction effects between the parameters grow over time, suggesting that the own effects control the ballistic wave propagation while the interaction effects control the diffusive wave propagation. The interaction effects control 40 to 60 per cent of the envelope after 10 s, but the own effect of the quality factor is still influent (about 40–60 per cent as well). In other words, the decay of coda waves amplitude cannot be solely attributed to the quality factor Q_S , whose influence is about 40–60 per cent of the whole influence, the remaining part being due to the interaction effects between the parameters. This conclusion was not achievable by Sochala & De Martin (2017) because the SIs had been computed independently in the time and frequency domains. The authors observed that the quality factor had a tiny influence in

the time domain but a significant impact on the high frequencies. The time–frequency representation presented here helps to understand how the influence of the parameters are shared between these two domains. The study of the interaction effects would require the computation of high-order indices through additional 3-D simulations and/or the design of surrogate models for the S transform.

ACKNOWLEDGEMENTS

This work was performed using HPC resources from KAUST Supercomputing Laboratory, Saudi Arabia (Shaheen II) and from GENCI-CINES, France (Occigen) for the Grand Challenges 2016. FDM, FD and PT thank D. Keyes and H. Ltaief for their support to access to the supercomputer Shaheen II. FD thanks S. Requena and E. Boyer for their invitation to participate to the Occigen Grand Challenges 2016 and V. Cameo Ponz for its technical support on the supercomputer Occigen. FDM thanks K. Meza-Fajardo for sharing an algorithm to compute the S transform. FDM and FD thank F. Boulahya, L. Santeramo, F. Guiet and A. C. Winfer for their help to transfer the large amount of data. Figures have been plotted with Gnuplot (<http://www.gnuplot.info/>) and GMT (<https://www.generic-mapping-tools.org/>). This work was performed under the auspices of BRGM fundings obtained from the French Ministry for Higher Education, Research and Innovation (MESRI, program 172). Very constructive reviews provided by two anonymous reviewers were helpful in revising and improving the paper.

DATA AVAILABILITY

The synthetic ground motions u_{21} , u_{22} and u_{23} computed at the stations TST and PRO, the GSA code (in Fortran/MPI) to compute the first-order SI on the S transform and the script to plot them are available at <http://efispec.free.fr/downloads/gji-demart-in-et-al-2020.tar.gz>. A README inside the tarball explains its content. Other data or codes are available on demand to F. De Martin. EFISPEC3D is an open-source code available at <http://efispec.free.fr>. Co-developement of EFISPEC3D can be done at <https://gitlab.brgm.fr/brgm/efispec3d>.

REFERENCES

- Aki, K. & Richards, P. G., 2002. *Quantitative Seismology*, 2nd edn, University Science Books.
- Asano, K., Sekiguchi, H., Iwata, T., Yoshimi, M., Hayashida, T., Saomoto, H. & Horikawa, H., 2016. Modelling of wave propagation and attenuation in the Osaka sedimentary basin, western Japan, during the 2013 Awaji Island earthquake, *Geophys. J. Int.*, **204**(3), 1678–1694.
- Asano, K., Iwata, T., Sekiguchi, H., Somei, K., Miyakoshi, K., Aoi, S. & Kunugi, T., 2017. Surface wave group velocity in the Osaka sedimentary basin, Japan, estimated using ambient noise cross-correlation functions, *Earth, Planets Space*, **69**(1), 1–20.
- Backus, G. E., 1962. Long-wave elastic anisotropy produced by horizontal layering, *J. geophys. Res.*, **67**(11), 4427–4440.
- Bard, P.-Y. & Bouchon, M., 1980. The seismic response of sediment-filled valleys. part 1. the case of incident SH waves, *Bull. seism. Soc. Am.*, **70**(4), 1263–1286.
- Beauval, C., Bard, P.-Y., Moczo, P. & Kristek, J., 2003. Quantification of frequency-dependent lengthening of seismic ground-motion duration due to local geology: applications to the Volvi area (Greece), *Bull. seism. Soc. Am.*, **93**(1), 371–385.
- Beroza, G. C., 1991. Near-source modeling of the Loma Prieta earthquake: evidence for heterogeneous slip and implications for earthquake hazard, *Bull. seism. Soc. Am.*, **81**(5), 1603–1621.
- Blacker, T. D., Bohnhoff, W. & Edwards, T., 1994. CUBIT Mesh Generation Environment, Vol. 1: Users manual, Tech. rep., Sandia National Labs., Albuquerque, NM, USA.
- Brocher, T. M., 2008. Key elements of regional seismic velocity models for long period ground motion simulations, *J. Seismol.*, **12**(2), 217–221.
- Calcagno, P., Chilès, J.-P., Courrioux, G. & Guillen, A., 2008. Geological modelling from field data and geological knowledge: Part I. Modelling method coupling 3D potential-field interpolation and geological rules, *Phys. Earth planet. Inter.*, **171**(1), 147–157.
- Capdeville, Y., Guillot, L. & Marigo, J.-J., 2010a. 1-D non-periodic homogenization for the seismic wave equation, *Geophys. J. Int.*, **181**(2), 897–910.
- Capdeville, Y., Guillot, L. & Marigo, J.-J., 2010b. 2-D non-periodic homogenization to upscale elastic media for P–SV waves, *Geophys. J. Int.*, **182**(2), 903–922.
- Chaljub, E. *et al.*, 2015. 3-D numerical simulations of earthquake ground motion in sedimentary basins: testing accuracy through stringent models, *Geophys. J. Int.*, **201**(1), 90–111.
- Crestaux, T., Le Maître, O. & Martinez, J.-M., 2009. Polynomial chaos expansion for sensitivity analysis, *Reliab. Eng. Syst. Saf.*, **94**(7), 1161–1172.
- De Basabe, J. & Sen, M., 2007. Grid dispersion and stability criteria of some common finite-element methods for acoustic and elastic wave equations, *Geophysics*, **72**(6), T81–T95.
- De Martin, F., 2011. Verification of a spectral-element method code for the southern California earthquake center LOH.3 viscoelastic case, *Bull. seism. Soc. Am.*, **101**(6), 2855–2865.
- Efron, B. & Tibshirani, R. J., 1994. *An Introduction to the Bootstrap*, CRC Press.
- Emmerich, H. & Korn, M., 1987. Incorporation of attenuation into time-domain computations of seismic wave fields, *Geophysics*, **52**(9), 1252–1264.
- Engquist, B. & Majda, A., 1977. Absorbing boundary conditions for numerical simulation of waves, *Proc. Natl. Acad. Sci.*, **74**(5), 1765–1766.
- Faure, H. & Tezuka, S., 2002. Another random scrambling of digital (t, s)-sequences, in *Monte Carlo and Quasi-Monte Carlo Methods 2000*, pp. 242–256, Springer.
- Frankel, A. & Clayton, R. W., 1986. Finite difference simulations of seismic scattering: Implications for the propagation of short-period seismic waves in the crust and models of crustal heterogeneity, *J. geophys. Res.*, **91**(B6), 6465–6489.
- Ghiglia, D. C. & Pritt, M. D., 1998. *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*, 1st edn, Wiley-Interscience.
- Graves, R. & Pitarka, A., 2016. Kinematic ground-motion simulations on rough faults including effects of 3D stochastic velocity perturbations, *Bull. seism. Soc. Am.*, **106**(5), 2136–2153.
- Green, D. A., 2011. A colour scheme for the display of astronomical intensity images, preprint ([arXiv:1108.5083](https://arxiv.org/abs/1108.5083)).
- Gropp, W., Lusk, E., Doss, N. & Skjellum, A., 1996. A high-performance, portable implementation of the MPI message passing interface standard, *Parall. Comput.*, **22**(6), 789–828.
- Guyonnet-Benaize, C. *et al.*, 2013. Imaging 3D geological structure of the Mygdonian basin (northern Greece) with geological numerical modeling and geophysical methods, in *Proceedings of the EGU General Assembly Conference*, held 7–12 April, 2013, Vienna, Austria, Vol. 15, id. EGU2013-9839.
- Homma, T. & Saltelli, A., 1996. Importance measures in global sensitivity analysis of nonlinear models, *Reliab. Eng. Syst. Saf.*, **52**(1), 1–17.
- Imtiaz, A., Cornou, C., Bard, P. & Hobiger, M., 2021. Diffracted wavefield decomposition and multidimensional site effects in the Argostoli valley, Greece, *Geophys. J. Int.*, **224**(3), 1849–1869.
- Itoh, K., 1982. Analysis of the phase unwrapping algorithm, *Appl. Opt.*, **21**(14), 2470–2470.
- Iwaki, A., Maeda, T., Morikawa, N., Takemura, S. & Fujiwara, H., 2018. Effects of random 3D upper crustal heterogeneity on long-period (≤ 1 s) ground-motion simulations, *Earth, Planets Space*, **70**(1), 156.

- Iwata, T., Kagawa, T., Petukhin, A. & Ohnishi, Y., 2008. Basin and crustal velocity structure models for the simulation of strong ground motions in the Kinki area, Japan, *J. Seismol.*, **12**(2), 223–234.
- Jongmans, D. et al., 1998. EURO-SEISTEST: determination of the geological structure of the Volvi basin and validation of the basin response, *Bull. seism. Soc. Am.*, **88**(2), 473–487.
- Jubertie, S., Dupros, F. & De Martin, F., 2018. Vectorization of a spectral finite-element numerical kernel, in *Proceedings of the 2018 4th Workshop on Programming Models for SIMD/Vector Processing*, pp. 8, ACM.
- Karhunen, K., 1947. *Über lineare Methoden in der Wahrscheinlichkeitsrechnung*, Vol. 37, Sana.
- Karray, M. & Lefebvre, G., 2008. Significance and evaluation of Poisson's ratio in Rayleigh wave testing, *Can. Geotech. J.*, **45**(5), 624–635.
- Kennedy, J. R. C. & O'Hagan, A., 2001. Bayesian calibration of computer models, *J. R. Stat. Soc., B*, **63**(3), 425–464.
- Koketsu, K., Miyake, H., Afnimar & Tanaka, Y., 2009. A proposal for a standard procedure of modeling 3-D velocity structures and its application to the Tokyo metropolitan area, Japan, *Tectonophysics*, **472**(1–4), 290–300.
- Komatitsch, D. & Tromp, J., 2002. Spectral-element simulations of global seismic wave propagation-I. Validation, *Geophys. J. Int.*, **149**(2), 390–412.
- Komatitsch, D. & Vilotte, J. P., 1998. The spectral-element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures, *Bull. seism. Soc. Am.*, **88**(2), 368–392.
- Komatitsch, D., Liu, Q., Tromp, J., Suss, P., Stidham, C. & Shaw, J. H., 2004. Simulations of ground motion in the los angeles basin based upon the spectral-element method, *Bull. seism. Soc. Am.*, **94**(1), 187–206.
- Lee, S.-J., Chen, H.-W., Liu, Q., Komatitsch, D., Huang, B.-S. & Tromp, J., 2008. Three-dimensional simulations of seismic-wave propagation in the taipei basin with realistic topography based upon the spectral-element method, *Bull. seism. Soc. Am.*, **98**(1), 253–264.
- Li, C. & Mahadevan, S., 2016. An efficient modularized sample-based method to estimate the first-order Sobol' index, *Reliab. Eng. Syst. Saf.*, **153**, 110–121.
- Loève, M., 1968. *Probability Theory: The University Series in Higher Mathematics*, East-West-Press Pvt. Limited.
- Ma, S. & Liu, P., 2006. Modeling of the perfectly matched layer absorbing boundaries and intrinsic attenuation in explicit finite-element methods, *Bull. seism. Soc. Am.*, **96**(1), 1779–1794.
- Maday, Y. & Patera, A. T., 1989. Spectral element methods for the incompressible Navier-Stokes equations, IN: State-of-the-art surveys on computational mechanics (A90-47176 21-64). New York, American Society of Mechanical Engineers, 1989, pp. 71–143. Research supported by DARPA.
- Maeda, T., Takemura, S. & Furumura, T., 2017. OpenSWPC: an open-source integrated parallel simulation code for modeling seismic wave propagation in 3D heterogeneous viscoelastic media, *Earth, Planets Space*, **69**(1), 1–20.
- Mai, P. M. & Beroza, G. C., 2002. A spatial random field model to characterize complexity in earthquake slip, *J. geophys. Res.*, **107**(B11), ESE 10–1-ESE 10-21.
- Manakou, M., Raptakis, D., Chávez-García, F., Apostolidis, P. & Ptilakis, K., 2010. 3D soil structure of the Mygdonian basin for site response analysis, *Soil Dyna. Earthq. Eng.*, **30**(11), 1198–1211.
- Matsushima, S., Hirokawa, T., De Martin, F., Kawase, H. & Sánchez-Sesma, F. J., 2014. The effect of lateral heterogeneity on horizontal-to-vertical spectral ratio of microtremors inferred from observation and synthetics, *Bull. seism. Soc. Am.*, **104**(1), 381–393.
- Maufroy, E. et al., 2015. Earthquake ground motion in the Mygdonian basin, Greece: the E2VP verification and validation of 3D numerical simulation up to 4 Hz, *Bull. seism. Soc. Am.*, **105**(3), 1342–1364.
- Maufroy, E. et al., 2017. Source-related variability of site response in the Mygdonian basin (Greece) from accelerometric recordings and 3D numerical simulations, *Bull. seism. Soc. Am.*, **107**(2), 787–808.
- Mocco, P., Kristek, J., Bard, P.-Y., Stripajová, S., Hollender, F., Chovanová, Z., Kristeková, M. & Sicilia, D., 2018. Key structural parameters affecting earthquake ground motion in 2D and 3D sedimentary structures, *Bull. Earthq. Eng.*, **16**(6), 2421–2450.
- Moreland, K., 2016. Why we use bad color maps and what you can do about it, *Electr. Imag.*, **2016**(16), 1–6.
- Paciucci, G., De Martin, F. & Thierry, P., 2016. Enhancing scalability and performance of parallel file systems, White paper, Tech. rep., Intel.
- Panunzio, A. M., Cottreau, R. & Puel, G., 2018. Large scale random fields generation using localized Karhunen–Loève expansion, *Adv. Model. Simulat. Eng. Sci.*, **5**(1), 1–29.
- Pardo-Iguzquiza, E. & Chica-Olmo, M., 1993. The fourier integral method: an efficient spectral method for simulation of random fields, *Math. Geol.*, **25**(2), 177–217.
- Phoon, K.-K. & Kulhawy, F. H., 1999. Characterization of geotechnical variability, *Can. Geotech. J.*, **36**(4), 612–624.
- Ptilakis, K., Roumelioti, Z., Raptakis, D., Manakou, M., Liakakis, K., Anastasiadis, A. & Ptilakis, D., 2013. The EUROSEISTEST strong-motion database and web portal, *Seismol. Res. Lett.*, **84**(5), 796–804.
- Raptakis, D., Chávez-García, F., Makra, K. & Ptilakis, K., 2000. Site effects at EUROSEISTEST–I. Determination of the valley structure and confrontation of observations with 1D analysis, *Soil Dyn. Earthq. Eng.*, **19**(1), 1–22.
- Sato, H., Fehler, M. C. & Maeda, T., 2012. *Seismic Wave Propagation and Scattering in the Heterogeneous Earth*, Springer Science & Business Media.
- Sobol, I., 2001. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, *Math. Comput. Simulat.*, **55**(1), 271–280.
- Sobol, I. & Levitan, Y. L., 1999. On the use of variance reducing multipliers in monte carlo computations of a global sensitivity index, *Comput. Phys. Commun.*, **117**(1), 52–61.
- Sobol, I. M., 1976. Uniformly distributed sequences with an additional uniform property, *USSR Comput. Math. Math. Phys.*, **16**(5), 236–242.
- Sobol, I. M., 1993. Sensitivity estimates for nonlinear mathematical models, *Math. Model. Comput. Exp.*, **1**, 407–414.
- Sochala, P. & De Martin, F., 2017. Surrogate combining harmonic decomposition and polynomial chaos for seismic shear waves in uncertain media, *Comput. Geosci.*, **22**, 125–144.
- Sochala, P., De Martin, F. & Le Maitre, O., 2020. Model reduction for large-scale earthquake simulation in an uncertain 3D medium, *Int. J. Uncertain. Quant.*, **10**(2), 101–127.
- Sornet, G., Jubertie, S., Dupros, F., De Martin, F. & Limet, S., 2018. Performance analysis of SIMD vectorization of high-order finite-element kernels, in *Proceedings of the 2018 International Conference on High Performance Computing & Simulation (HPCS)*, pp. 423–430, IEEE.
- Stockwell, R. G., Mansinha, L. & Lowe, R., 1996. Localization of the complex spectrum: the S transform, *IEEE Trans. Signal Process.*, **44**(4), 998–1001.
- Sudret, B., 2008. Global sensitivity analysis using polynomial chaos expansions, *Reliab. Eng. Syst. Saf.*, **93**(7), 964–979.
- Takemura, S., Yabe, S. & Emoto, K., 2020. Modelling high-frequency seismograms at ocean bottom seismometers: effects of heterogeneous structures on source parameter estimation for small offshore earthquakes and shallow low-frequency tremors, *Geophys. J. Int.*, **223**(3), 1708–1723.
- Yoshimoto, K. & Takemura, S., 2014. A study on the predominant period of long-period ground motions in the Kanto basin, Japan, *Earth, Planets Space*, **66**(1), 1–7.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Figure S1. First order sensitivity indices and their sum and interaction computed on the envelope and phase of the S transform of $u_3 = u_{31} + u_{32}$ at the TST station. The ground motion u_3 of the mean profile is plotted with its standard deviation (panels m and n).

Figure S2. First order sensitivity indices and their sum and interaction computed on the envelope and phase of the S transform of u_1

= $u_{11} + u_{12}$ at the TST station. The ground motion u_1 of the mean profile is plotted with its standard deviation (panels m and n).

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.

APPENDIX A: VARIANCE DECOMPOSITION

Let $\mathbf{m} := (m_1, \dots, m_N)$ be a vector of N uncertain *independent* parameters. The joint distribution $p_{\mathbf{m}}$ and the parametric domain \mathbb{M} are defined as

$$p_{\mathbf{m}} := \prod_{k=1}^N p_k \quad \text{and} \quad \mathbb{M} := \prod_{k=1}^N I_k, \quad (\text{A1})$$

where p_k and I_k are the distribution and range of the parameter m_k .

The total variance \mathbb{V} of a quantity of interest $u(\mathbf{m})$, $\mathbf{m} \in \mathbb{M}$, can be decomposed into 2^{N-1} terms (Sobol 1993) ($N = 7$ in our study),

$$\mathbb{V}(u(\mathbf{m})) = \sum_{i=1}^N \mathbb{V}_i + \sum_{i < j} \mathbb{V}_{i,j} + \dots + \mathbb{V}_{1,\dots,N}, \quad (\text{A2})$$

where $\{\mathbb{V}_i\}$ denote the first-order interaction terms, $\{\mathbb{V}_{i,j}\}$ the second order terms, and so on. In particular, the partial variance \mathbb{V}_i is equal to the variance of the conditional expectation of u given the i th parameter m_i ,

$$\mathbb{V}_i := \mathbb{V}[\mathbb{E}(u|m_i)]. \quad (\text{A3})$$

The expectation operator is defined over the domain $\mathbb{M}_{\sim i} := \prod_{\substack{k=1 \\ k \neq i}}^N I_k$ while the variance operator is defined over the range I_i of m_i ,

$$\mathbb{E}(u|m_i) = \int_{\mathbb{M}_{\sim i}} u(\mathbf{m}) \prod_{\substack{k=1 \\ k \neq i}}^N dm_k, \quad \text{and} \quad \mathbb{V}(v) = \int_{I_i} \left(v - \int_{I_i} v dm_i \right)^2 dm_i. \quad (\text{A4})$$

We can note that $\mathbb{M}_{\sim i}$ is a $N - 1$ dimensional domain excluding the i th dimension whereas I_i is the one dimensional range of m_i , so that we have $\mathbb{M}_{\sim i} \times I_i = \mathbb{M}$. As the parameters have uniform distributions, their density functions p_{m_i} have been omitted in the integrals for sake of simplicity.

APPENDIX B: METHOD TO COMPUTE FIRST-ORDER SENSITIVITY INDICES

The first order sensitivity indices can be estimated from an extensive sampling of the parametric domain or derived from a surrogate model. By focusing only on the first-order indices, we rely on a sample-based method (Li & Mahadevan 2016) for its ease of implementation and fast computational time.

Dividing both sides of eq. (A2) by the variance lead to the so-called sensitivity indices (SI). The first-order indices $\{S_i\}$ measure the relative contributions of each parameter alone onto the variance,

$$S_i := \frac{\mathbb{V}_i}{\mathbb{V}(u(\mathbf{m}))}. \quad (\text{B1})$$

These indices are commonly used since they allow to rank the effect of each uncertain input on the output. By definition, each index belongs to $[0,1]$ and their sum is equal to one.

The method proposed by Li & Mahadevan (2016), which is based on a discretization of I_i into L equally probable intervals $\{\Phi_i^l\}$, computes a set of L local empirical conditional expectation $\{\widehat{\mathbb{E}}_i^l(u)\}_{1 \leq l \leq L}$ and then calculates the empirical variance of this set. The empirical estimators are formally written as

$$\mathbb{E}(u|m_i \in \Phi_i^l) \approx \widehat{\mathbb{E}}_i^l(u) := \frac{1}{|\mathbf{M}_i^l|} \sum_{\mathbf{m} \in \mathbf{M}_i^l} u(\mathbf{m}) \quad \text{where} \quad \mathbf{M}_i^l := \{\mathbf{m} \in \mathbf{M} \text{ s.t. } m_i \in \Phi_i^l\}, \quad (\text{B2})$$

$$\mathbb{V}[\mathbb{E}(u|m_i)] \approx \widehat{\mathbb{V}}_L(\widehat{\mathbb{E}}_i^l(u)) := \frac{1}{L-1} \sum_{l=1}^L \left(\widehat{\mathbb{E}}_i^l(u) - \frac{1}{L} \sum_{l=1}^L \widehat{\mathbb{E}}_i^l(u) \right)^2, \quad (\text{B3})$$

where $|\mathbf{M}_i^l|$ denotes the number of realizations associated to \mathbf{M}_i^l . The calculation procedure has been applied separately to each coefficient of the PGM vector as well as the S transform envelope and phase matrices.