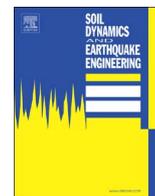




Contents lists available at ScienceDirect

## Soil Dynamics and Earthquake Engineering

journal homepage: [www.elsevier.com/locate/soildyn](http://www.elsevier.com/locate/soildyn)

## 3D numerical simulation and ground motion prediction: Verification, validation and beyond – Lessons from the E2VP project

E. Maufroy<sup>a</sup>, E. Chaljub<sup>a</sup>, F. Hollender<sup>b</sup>, P.-Y. Bard<sup>a,\*</sup>, J. Kristek<sup>c,d</sup>, P. Moczo<sup>c,d</sup>, F. De Martin<sup>e</sup>, N. Theodoulidis<sup>f</sup>, M. Manakou<sup>g</sup>, C. Guyonnet-Benaize<sup>h</sup>, N. Hollard<sup>a</sup>, K. Pitilakis<sup>g</sup>

<sup>a</sup> University Grenoble Alpes, ISTERre, CNRS, IRD, IIFSTTAR, Grenoble, France

<sup>b</sup> CEA, French Alternative Energies and Atomic Energy Commission, DEN, Saint Paul lez Durance, France

<sup>c</sup> Faculty of Math., Physics and Informatics, Comenius University, Bratislava, Slovakia

<sup>d</sup> Earth Science Institute, Slovak Academy of Sciences, Bratislava, Slovakia

<sup>e</sup> Bureau de Recherches Géologiques et Minières, Orléans, France

<sup>f</sup> ITSAK, Institute of Engineering Seismology and Earthquake Engineering, Thessaloniki, Greece

<sup>g</sup> Aristotle University of Thessaloniki, Thessaloniki, Greece

<sup>h</sup> Formerly CEA, Cadarache, France

### ARTICLE INFO

#### Keywords:

Ground motion  
Numerical simulation  
Validation  
Site effects  
Goodness-of-fit  
GMPE  
Aleatory variability

### ABSTRACT

The Euroseisest Verification and Validation Project (E2VP) is part of a series of complementary benchmarking exercises launched to better assess the ability of numerical simulation to accurately predict seismic ground motion. E2VP targeted more specifically the current, most-advanced numerical methods applied to realistic 3D, linear models of sedimentary basins through a quantitative comparison of the recorded and numerically-simulated ground motions. The target site, located within the Mygdonian basin near Thessaloniki, Greece, has been thoroughly investigated for two decades and a detailed, realistic 3D model has been derived from geological, geophysical and geotechnical investigations, while a dedicated instrumentation provided a significant number of surface and borehole recordings. Verification and validation tests up to a frequency of 4 Hz, much beyond the 0.4 Hz fundamental frequency of the deepest part of the graben, have been performed for a set of 19 local, small to moderate magnitude events. For careful and accurate enough computations, the model-to-model differences are smaller than the model-to-observations differences, the latter being controlled by uncertainties primarily in the crustal propagation model and source properties, and secondarily in the shallow structure. It is therefore recommended to prefer distant and/or deep events ( $R > 10\text{--}20$  km,  $Z > 8\text{--}10$  km) for validation exercises. Additional sensitivity tests illustrate the ability of carefully verified numerical simulation tools to provide an instructive insight at the structure of the so-called “aleatory” variability of ground motion, for both its within- and between-event components. The between-event variability is shown to be very sensitive to hypocenter location errors (even as low as  $\pm 2$  km), and to uncertainty in magnitude estimates. It explains the increase of aleatory variability for small magnitude events and emphasizes the usefulness of dense seismological networks. The within event, single-site variability is shown to be associated to an “epistemic” dependence of the 3D site response on the event back-azimuth, distance and depth, and calls for caution when interpreting single-station variabilities derived from a too small number of events.

### 1. Introduction

The rapid development of the simulation codes and computational facilities allowed considering the use of numerical-simulation tools as a valid option for predicting seismic ground motion, especially for poorly

instrumented or moderate-seismicity countries lacking representative earthquake recordings. However, such an approach requires a careful evaluation of the actual performance of numerical simulation codes. This issue has been the topic of a few international studies, including blind prediction tests or comparative exercises, focused on various

\* Correspondence to: ISTERre – Maison des Géosciences/Bâtiment OSUG C, CS 40700, 38058 GRENOBLE Cedex 9, France.

E-mail addresses: [emeline.maufroy@univ-grenoble-alpes.fr](mailto:emeline.maufroy@univ-grenoble-alpes.fr) (E. Maufroy), [emmanuel.chaljub@univ-grenoble-alpes.fr](mailto:emmanuel.chaljub@univ-grenoble-alpes.fr) (E. Chaljub), [fabrice.hollender@cea.fr](mailto:fabrice.hollender@cea.fr) (F. Hollender), [pierre-yves.bard@univ-grenoble-alpes.fr](mailto:pierre-yves.bard@univ-grenoble-alpes.fr) (P.-Y. Bard), [Jozef.Kristek@fmph.uniba.sk](mailto:Jozef.Kristek@fmph.uniba.sk) (J. Kristek), [moczo@fmph.uniba.sk](mailto:moczo@fmph.uniba.sk) (P. Moczo), [F.DeMartin@brgm.fr](mailto:F.DeMartin@brgm.fr) (F. De Martin), [ntheo@itsak.gr](mailto:ntheo@itsak.gr) (N. Theodoulidis), [manakou@civil.auth.gr](mailto:manakou@civil.auth.gr) (M. Manakou), [cedric.guyonnet-benaize@hotmail.fr](mailto:cedric.guyonnet-benaize@hotmail.fr) (C. Guyonnet-Benaize), [Niels.Hollard@e-ujf-grenoble.fr](mailto:Niels.Hollard@e-ujf-grenoble.fr) (N. Hollard), [kpitilakis@civil.auth.gr](mailto:kpitilakis@civil.auth.gr) (K. Pitilakis).

<http://dx.doi.org/10.1016/j.soildyn.2016.09.047>

Received 12 February 2016; Received in revised form 23 September 2016; Accepted 27 September 2016

Available online xxx

0267-7261/© 2016 Elsevier Ltd. All rights reserved.

sites. It started with the Turkey Flat, California (Cramer [1]), and Ashigara Valley, Japan (e.g., Bard [2]), blind tests focusing on effects of surface sediments, the results of which were presented during the first ESG conference in Odawara (Japan) in 1992. It was followed by more comprehensive comparison exercises on the Osaka/Kobe basin area in Japan (Kawase and Iwata, 1998 [3]), and on the Southern California area within the SCEC framework (Day et al. [4–6]; Bielak et al. [7]), which also included the effects of extended sources and regional propagation in the low frequency range ( $f < 1$  Hz). Each of these cases had its own specificities (for instance, very low frequencies for the Osaka and SCEC exercises). A request issued in late 2003 by the French Nuclear Authority (ASN) to perform a 3D, NL simulation of site response for specific sites, was the initial impetus for a dedicated R & D program funded by CEA Cadarache and ILL (Laue-Langevin Institute, an international research center on neutron science based in Grenoble, and operating the most intense neutron source on Earth). It started with an international benchmarking exercise on the Grenoble basin (Chaljub et al. [8]; Tsuno et al. [9]; Chaljub et al. [10]), and was further deepened through the Euroseistest Verification and Validation Project (E2VP). Considering the lessons of the ESG2006 Grenoble benchmark, the E2VP project was launched in 2007 with two main objectives: (a) a quantitative analysis of the accuracy of current, most-advanced numerical methods applied to realistic 3D models of sedimentary basins, in the linear, small strain domain (3DL verification); (b) a quantitative comparison of the recorded and numerically-simulated ground motions (3DL validation). The selected target site was an extensional graben located in the Mygdonian basin near Thessaloniki, Greece, located in a seismically active zone, belonging to both Serbomacedonian massif and Circum Rodope zone (Fig. 1). A detailed, realistic 3D model of the basin and surrounding area had already been derived from a comprehensive set of geological, geophysical and geotechnical investigations, and the site instrumentation installed for about two decades provided a significant number of surface and borehole recordings.

This paper is intended to present a concise overview of the work accomplished since the launching of the E2VP project. This project has been organized in two phases, E2VP1 (2007–2010) and E2VP2 (2012–

2014). As the main results of the first phase are reported in two recent papers (Chaljub et al. [11]; Maufroy et al. [12]), the present article puts more emphasis on the latest results, while reminding the overall process. The first section shortly reminds the main learnings of E2VP1, and its shortcomings as well. A few key issues were identified, which shaped the second phase E2VP2: its main components are presented in the following section, including an improvement of the source parameters for a larger set of validation events, an enlargement and refinement of the 3D model on the basis of newly compiled information and sometimes new measurements, and a comprehensive set of numerical simulations for close to 2000 point source locations and 15 receivers. These simulations aim first at the validation up to a frequency of 4 Hz, much beyond the 0.4 Hz fundamental frequency of the deepest part of the graben, for a set of 19 local, small to moderate magnitude events. The corresponding results are described in the following section, distinguishing the rock and sediment stations, and for the latter the absolute ground motion and the 3D site response. The next section is dedicated to the presentation of additional sensitivity tests, which illustrate the ability of carefully verified numerical simulation tools to provide an instructive insight at the structure of the so-called “aleatory” variability of ground motion: the between-event component is shown to be highly impacted by uncertainties in hypocentral location and magnitude, while the within event component is affected by the epistemic dependence of site response on source back-azimuth. The conclusion summarizes the main outcomes from the whole E2VP project, including recommendations regarding the organization of further validation exercises, the use of numerical simulation for ground motion prediction in engineering projects, and the analysis, interpretation and reduction of the aleatory variability in GMPEs.

## 2. From E2VP1 to E2VP2: the main steps

In short, the basic ideas of the project were, on the example of the Euroseistest site, to (1) quantify the “distance” between results of independent models and numerical schemes, and as much as possible to reduce them to the lowest possible level through a careful understanding of the differences; and (2) to compare this “cross-computation

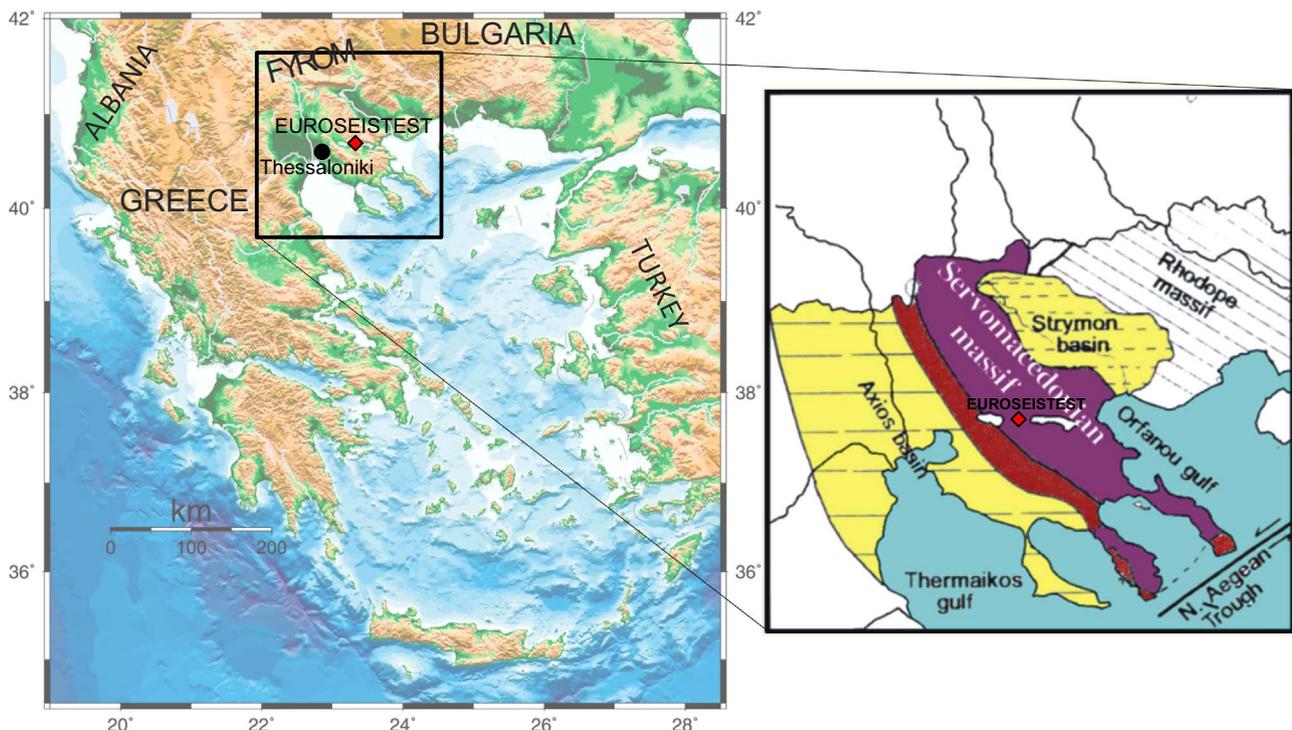


Fig. 1. Location of the Euroseistest site in North-Eastern Greece and first-order geological map of the surroundings of the Mygdonian basin. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 1**

Summary of main learnings from E2VP phase 1.

<p><b>Main lessons about verification and validation studies</b></p>	<ul style="list-style-type: none"> <li>● Careful verification requires time and often to “go back to basics”, while careful validation requires high quality data, i.e., including rich and high quality metadata.</li> <li>● No ground-motion simulation code accounting for wave propagation in complex media can be considered as press-button, neither in the linear, 3D domain, nor in the non-linear 2D - or even 1D - cases. The most common case is that, without iterations and cross-checking, different codes provide significantly different results when applied to the same case study.</li> <li>● Too fast applications of existing codes by non-expert users may yield wrong ground-motion estimates, potentially resulting in raising mistrust in end-users. This warning is applicable to all numerical simulation, including the simplest ones, but gets more and more important with increasing code sophistication.</li> <li>● Some codes currently used in engineering applications would deserve some significant improvements, or strong warnings on stringent validity limits, while even state-of-the-art codes (predominantly in the “academic” field) deserve constant upgrading.</li> </ul>
<p><b>Main recommendations for a wise use of numerical simulation codes</b></p>	<ul style="list-style-type: none"> <li>● One should never be satisfied with only one computation from one single team, but should request several teams (at least two) with different numerical schemes to perform parallel computations of the same case. Results should be considered as reliable only if they agree beyond some quantitative goodness-of-fit threshold.</li> <li>● These goodness-of-fit criteria should definitely be agreed upon by the engineering community in order to reach an objective of transparent quantitative comparison, which should replace sentences such as “one can see the very good agreement on the figure”</li> <li>● In the long run, it would be very valuable to assign a specific “quality label” to numerical codes and teams that did accept to run some of the now existing “canonical” cases with their own numerical code, which are freely available on web pages (<a href="http://www.sismowine.org/">http://www.sismowine.org/</a>). Maintaining this kind of internet facility in the long run will be beneficial for the whole community.</li> <li>● External peer reviews are always useful in assessing the quality of results derived from highly sophisticated numerical codes.</li> <li>● Comparison with actual data (in-situ earthquake recordings), whenever possible, are always useful. Having sensitive in-situ instrumentation (continuously recording broad-band velocimeters or sensitive accelerometers) proves to be invaluable for checking the reliability of numerical-simulation results.</li> </ul>

distance” to the “misfit” between simulation results and actual measured data for as many real events as possible. In order to keep track of overestimation or underestimation bias, a sign is included in this distance as detailed in Maufroy et al. [12]. The first phase E2VP1 (2007–2010) included a comprehensive series of cross-model verifications, with side computations on canonical models aimed at investigating the accuracy of numerical schemes under very stringent conditions – as detailed in Chaljub et al. [11] –, and a first round of comparison between observations and simulations for a small number (6) of local events, as reported in Maufroy et al. [12]. The computations were performed up to a frequency of 4 Hz: this remains limited compared to the frequency range of interest in earthquake engineering, but this is significantly higher than all previous similar exercises. It led to a number of lessons and recommendations on the use of the numerical-simulation approach, as listed in Table 1, but it also led to the identification of a few further issues that needed to be addressed in a second phase.

### 2.1. 3D linear modeling

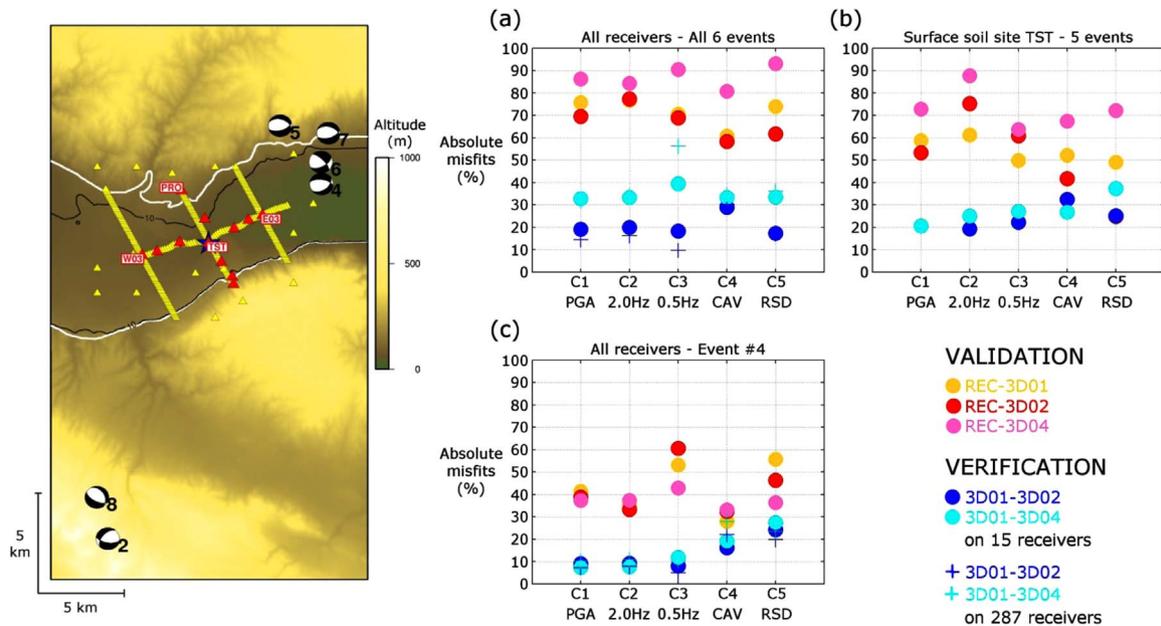
The main focus and success of E2VP1 was thus on the use of 3D, linear simulation. The main results are summarized in Fig. 2. The code-to-code differences could be drastically reduced by the consideration of dedicated canonical models and stringent goodness-of-fit criteria comparing the waveforms in the time-frequency domain (Kristekova et al. [13]), leading to significant improvements in the numerical schemes (Chaljub et al. [11]). The simulation-to-observation differences could be quantified for only a limited number of events (6) because of the moderate seismicity and the limited extension of the 3D model considered by that time. For those events, the simulated and observed waveforms remain so different that another metrics was adopted to quantify their differences, on the basis of “engineering” parameters. After a careful analysis of the original Anderson’s criteria (Anderson [14]), five parameters (C1 to C5) were selected (see Maufroy et al. [12]): pga (C1), the spectral acceleration at intermediate (C2) and low (C3) frequencies (averages in the [1.5–3 Hz] and [0.375–0.75 Hz] ranges, respectively), an “energy” indicator C4 (cumulative absolute velocity, CAV), and the Trifunac-Brady duration (RSD, Trifunac and

Brady [15]) for C5. Parameters C1–C3 evaluate the amplitude of the signal in different frequency bands. These frequency bands are chosen according to the observed characteristics of the real signals at the center of the Mygdonian basin: the frequency range evaluated by C3 includes the fundamental resonance frequency of the basin, while C2 covers the two higher modes. C1 evaluates the highest frequencies available in the synthetics. C4 and C5 evaluate the total energy of the signal and its duration, respectively. The misfit was computed for each parameter in terms of relative increase or decrease compared to the measured values. Fig. 2 indicates that such an “engineering” distance is around 10–25% between different simulations, to be compared with misfit values in the range 40–80% between observations and simulations. As detailed in Maufroy et al. [12], these numbers do vary depending on the considered receiver (rock or valley), on the considered event, and on the engineering parameter, but the overall trends are robust, and emphasize both the usefulness of the prior verification part and the difficulty to obtain satisfactory, unbiased numerical predictions of ground motion.

Only very few events could be used for the validation: this is a typical situation for moderate/weak seismicity areas. It was therefore considered useful to include more events [from 6 to 19] in the second phase of the validation exercise (those shown in Fig. 1), which led to increase the size of the 3D model, as illustrated later in Fig. 6. In addition, the significant misfit between observations and simulations was shown to be partly due to uncertainties or errors in source parameters: the misfits on the sole site response component were found lower than those on absolute motion (Maufroy et al. [12]). It was thus decided first to improve as much as possible the location of the 19 selected events, and second to investigate through numerical simulation how the uncertainties in source parameters map on the variability of site-specific ground motion from local earthquakes.

### 2.2. Non-linear (NL) modeling

The first phase also included a comparison of 2D, NL simulations on a NS cross-section of Euroseistest. This attempt for a verification of NL codes proved however to be a failure, as code-to-code differences were too large with too many, too poorly identified origins. Yet, it is



**Fig. 2.** Summary of horizontal absolute misfits obtained on the E2VP1 evaluation parameters C1–C5 (see text for their definition and Maufroy et al. [12] for more details) for the verification and validation exercises considering different configurations. Left: localization and focal mechanism of the 6 validation events (beachballs) and of the receivers used for the comparison (red and yellow triangles). Right: (a) average misfits for the 6 selected events at all receivers; (b) average misfits for the 5 events recorded at the central soil site TST; (c) average misfits for the biggest event #4 at all receivers. Synthetics-to-synthetics distances (verification, blue tones dots) are compared to recordings-to-synthetics misfits (validation, warm tones dots). The verification distances are computed for either the real array of 15 surface receivers (red triangles, solid circles) or the complete virtual array of 287 receivers (yellow triangles, crosses). A single value per array is obtained by calculating the weighted average of the absolute distances over the considered receivers (with weights proportional to the value of the corresponding ground-motion parameter). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

obvious that NL simulation codes deserve similar verification and validation efforts, especially as they are much more often used in engineering practice than 3D, linear simulation codes. Although non-linear site response was one of the major topics of the two pioneering blind tests initiated in the late 80's for the sites of Ashigara Valley (Japan) and Turkey Flat (California), they were inconclusive because both sites lacked strong motion records involving significant non-linear behavior. A new benchmarking of 1D NL codes was performed in the last decade, based on the same Turkey Flat site that experienced a 0.3g motion during the 2004 Parkfield earthquake, and a few other sites with vertical array data (La Cienega, California; the KGW02 KIK-net site in Japan, and Lotung in Taiwan). Its main findings, reported by Kwok et al. [16] and Stewart and Kwok [17], emphasized the key importance of the way these codes are used and of the required in-situ measurements. Significant differences between records and predictions have been identified as due to an incorrect velocity profile (although it was derived from redundant borehole measurements), a non-1D soil geometry (non-planar layers), and imperfections/deficiencies in the constitutive models, which were unable to represent the actual degradation curves for shear modulus and damping. The E2VP1 failure and these recent conclusions thus allowed to issue three main recommendations for future benchmarking exercises: a) NL verification should be performed on the simplest possible cases (1D soil columns and total stress, water-free, analysis); b) it should be performed on already instrumented sites having recorded large acceleration levels; c) it should be associated with careful in-situ surveys and lab tests designed in tight connection with the needs of the rheological models implemented in the various NL codes.

The second phase, E2VP2 (2012–2014), was thus designed to answer some of the identified issues related with 3D linear modeling, while two other projects were launched to address some other: the PRENOLIN project (Régner et al. [18–20]) was designed to start answering the issues about NL modeling according to the E2VP1 lessons, and another benchmarking exercise, named “INTERPACIFIC”, was launched for a comparative assessment of the performance of various in-situ geophysical and geotechnical survey techniques

(Garofalo et al. [21,22]). The present paper focuses exclusively on the new results related to 3D, linear modeling (E2VP2).

### 3. New validation phase E2VP2: model, data and simulations

This section presents the four main components of the additional work performed for this new phase, while the following sections will be dedicated to the presentation of the new results, in terms of validation and sensitivity analysis. The additional work started with the selection of a larger set of events and the re-assessment of their source parameters; the consideration of a larger set of events implied an enlargement of the 3D model, which was therefore updated and implemented in an improved numerical code. The latter was then used for computing the ground motion associated not only to the new set of real events, but also for sensitivity studies allowing the investigate the deterministic impact of source location (distance, depth and back-azimuth) on ground motion characteristics, together with the impact of source parameter uncertainties on ground motion variability.

#### 3.1. Improvement of source parameters for an increased number of local events (from 6 to 19)

This work included careful relocation and determination of focal mechanisms through waveform fitting of broad-band and accelerometric recordings with 1D synthetics computed by the discrete wavenumber method with the crustal velocity model proposed by Novotny et al. [23]. It is worth mentioning that for the events that were already considered in the first validation phase, the new source parameters can vary significantly with respect to the old ones: as an example, the largest event ( $M_w=4.4$ ) used in the first phase was moved by 5 km vertically and 4.5 km horizontally to define the S3 event of phase 2. The resulting event parameters are listed in Table 2 and their location and focal mechanism are displayed in Fig. 3. Details on the relocation work may be found in appendix 3 of Maufroy et al. [24].

**Table 2**

Characteristics of the 19 selected real events that occurred near the Mygdonian basin, whose recordings by the Euroseistest accelerometric array are compared to 3D numerical predictions in the validation phase 2. Only the preferred solutions of the inversion for source parameters are shown.

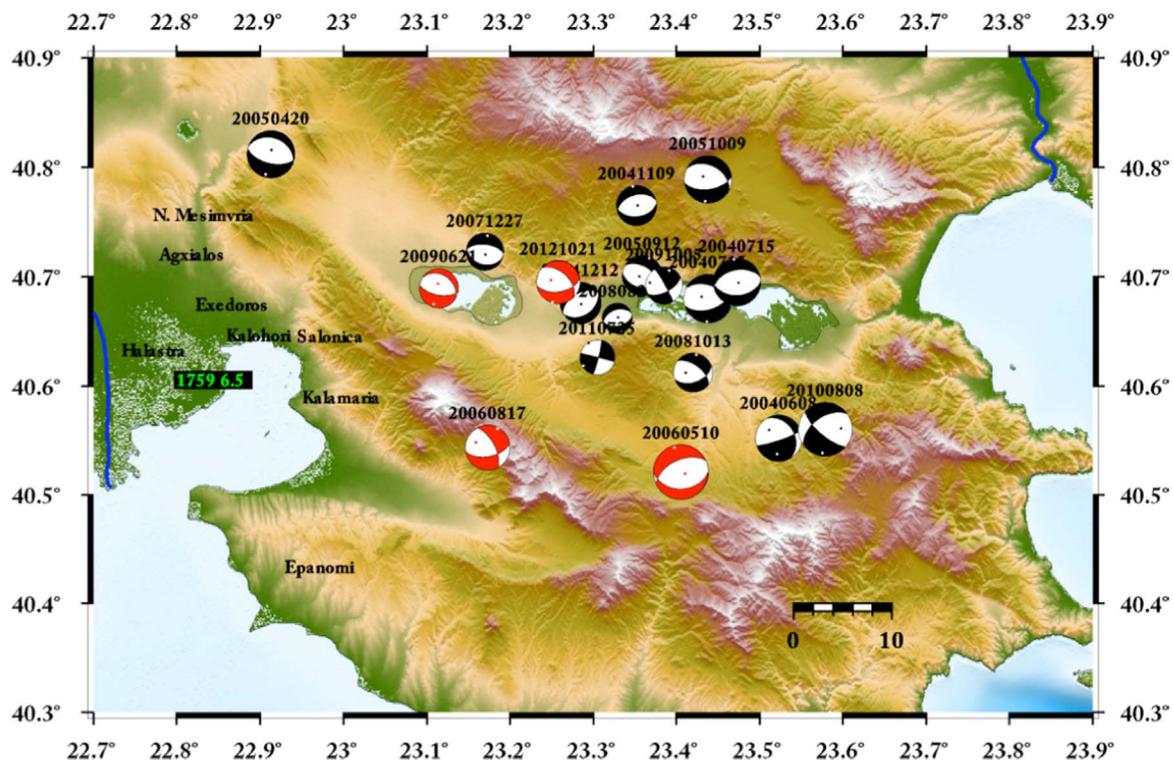
Event ID	Date	Lat. (°)	Long. (°)	Depth (km)	Mag. Mw	TST hyp. dist. (km)	Strike (°)	Dip (°)	Rake (°)
S1	2006/05/10	40.5208	23.4052	5	4.38	19.3	245	54	-105
S2	2006/08/17	40.5433	23.1732	11	3.59	20.0	80	57	-149
S3	2005/09/12	40.7255	23.3408	10	4.40	12.8	281	52	-98
S4	2009/06/21	40.6895	23.1148	11	3.14	18.7	100	61	-102
S5	2012/10/21	40.6950	23.2580	11	3.44	11.9	81	53	-127
S6	2004/06/08	40.5520	23.5233	9	3.30	25.0	71	82	-121
S7	2004/07/15	40.6800	23.4378	7	3.70	14.4	73	53	-118
S8	2004/07/15	40.6952	23.4733	9	3.70	18.2	258	47	-96
S9	2004/11/09	40.7648	23.3520	3	3.10	12.7	253	46	-98
S10	2004/12/12	40.6760	23.2853	4	2.70	4.2	240	51	-89
S11	2005/04/20	40.8121	22.9129	4	3.50	36.1	103	58	-94
S12	2005/09/12	40.7012	23.3586	4	3.00	8.1	301	52	-77
S13	2005/10/09	40.7889	23.4375	8	3.40	20.2	64	74	-116
S14	2007/12/27	40.7230	23.1700	11	3.50	16.4	276	59	-95
S15	2008/08/28	40.6617	23.3292	3	2.80	4.4	80	48	-83
S16	2008/10/13	40.6120	23.4200	9	2.90	15.3	306	58	-52
S17	2009/10/05	40.6920	23.3850	10	3.40	13.1	63	60	-174
S18	2010/08/08	40.5603	23.5785	8	4.60	28.1	235	52	-157
S19	2011/07/25	40.6265	23.3047	5	2.80	6.6	14	84	0

### 3.2. Update and extension of the 3D model

The E2VP-phase 1 was based on the preexisting 3D model as proposed by Manakou [25] and Manakou et al. [26]. For E2VP-phase 2, a new 3D model was built “from scratch” in order to avoid any bias due to pre-existing interpretation choices. This new model has been extended to the whole Mygdonian basin by gathering all the available information complemented by a few additional specific measurements to constrain the bedrock geometry, the sedimentary thickness and the seismic velocity. An important point to mention in order to understand the “philosophy” of this E2VP2 validation effort, is that we did our best to build the new 3D model only on the basis of the available geological,

geophysical and geotechnical data: there was no attempt to “retro-fit” the model in an *ad hoc* approach in order to optimize the fit between simulated ground motion and real records (an approach which was actually used in the first validation phase with the previous model). The objective is to be as close as possible to a realistic, blind prediction situation.

The compiled data include geology, hydrological and geotechnical boreholes (190 in total, out of which 59 reached the paleozoic basement at depths varying from 0 to 408 m), geophysical surveys (seismic refraction lines, controlled source/radio magneto-telluric surveys, array microtremor and H/V measurements). All available data were compiled in 3D using the geomodeller GOCAD (Caumon et al. [27];



**Fig. 3.** Map of the 19 seismic events that were considered for the validation part of E2VP2. The focal mechanisms are indicated with beach-balls, the size of which is proportional to the magnitude of the event. Most of the events exhibit normal faulting, consistently with the extension regime of the area. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

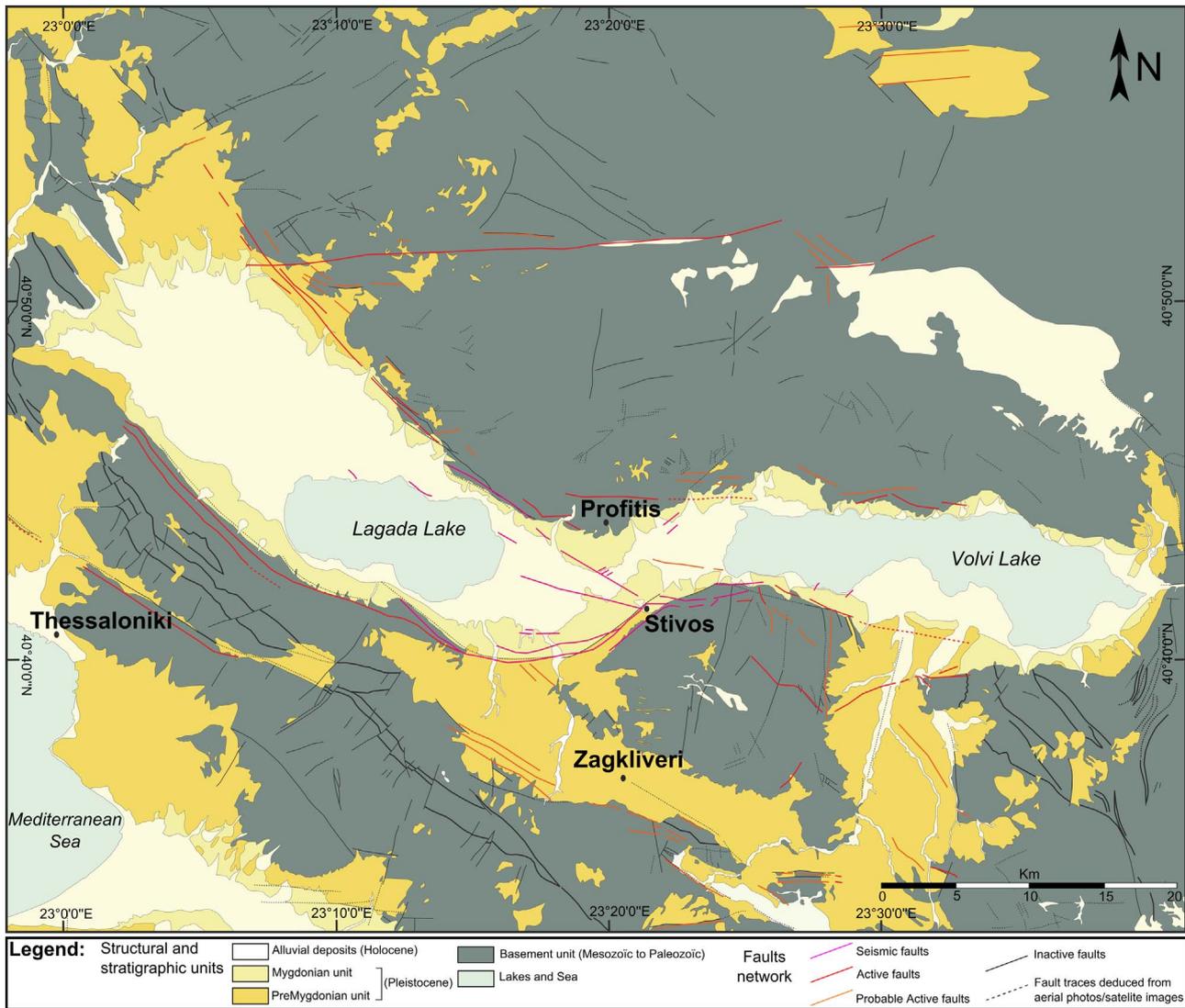


Fig. 4. Simplified structural sketch of the Mygdonian basin, modified from Mygdonian geological [29] and neotectonic maps [30], scales 1:50000–1:100000. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

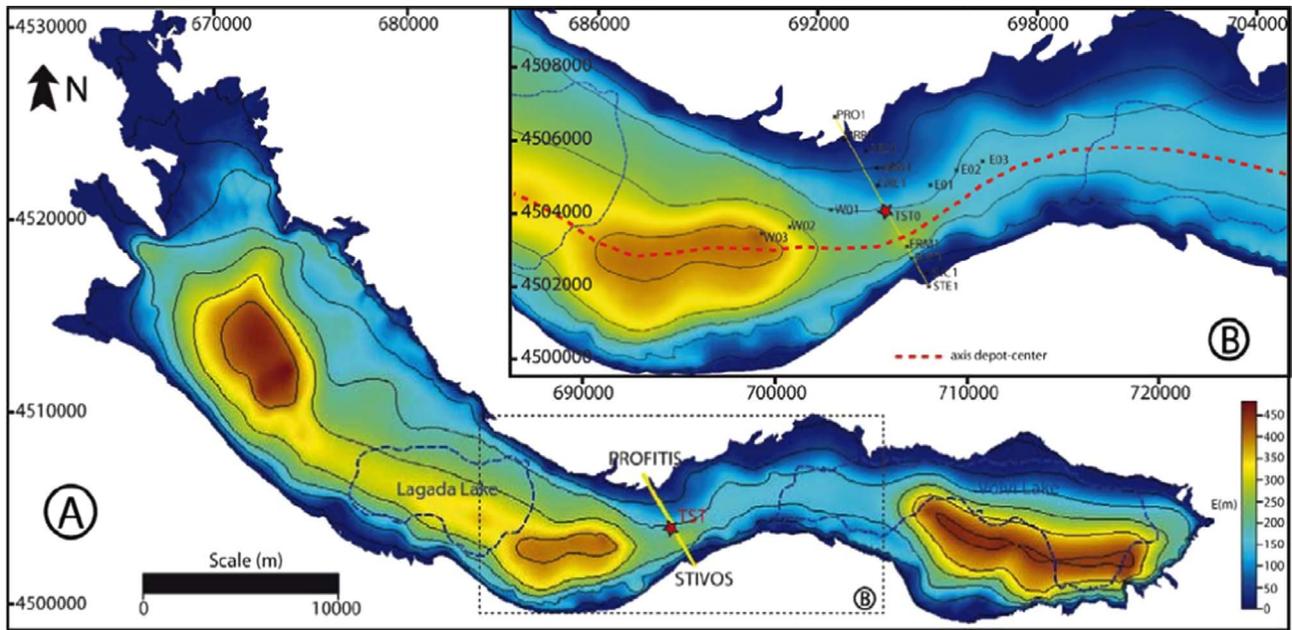
Mallet [28]) to build a 3D geological model for the whole Mygdonian basin.

The main features of the resulting model are 1/the 3D geometry of the “geophysical” bedrock (i.e. Paleozoic basement), 2/the main faults surfaces and 3/the thickness maps of the entire filling of the Mygdonian basin. The present-day structure of the basin (shown in Fig. 4) is composed of three structural units, from shallower to deeper unit: (1) the Mygdonian system (2) the ProMygdonian system and (3) the Paleozoic basement (Manakou [25]; Manakou et al. [26]). The Mygdonian and ProMygdonian systems are two sedimentary units presenting significant lateral thickness variations, from 140 m in the eastern part (close to the Volvi Lake) to 400 m in the western part (close to the Lagada Lake). The Mygdonian system is composed of quaternary fluvial-lacustrine, deltaic, lacustrine, lagoonal and estuarine deposits (Sotiriadis et al. [31]), corresponding to Pleistocene and Holocene age. The ProMygdonian system is composed of conglomerates, sandstones, silt-sand and red-beds sediments (Raptakis et al. [32]), with a Tertiary age. These two sedimentary units overlay the Paleozoic basement, composed of gneiss, amphibolites, two-mica schists and marble intrusions. The overall thickness of the Mygdonian and ProMygdonian units is mapped in Fig. 5, together with the surface topography outside of the Mygdonian sedimentary filling.

These structural units are affected by a complex fault system. In the entire basin, the faults are mostly striking NW-SE, excepted in the eastern part (Volvi Lake) where the faults strike E-W and N-S. The main features are the 12 km long Vasiloudi - Gerakarou - Nikomidino - Stivos fault system, running through the southern and western part of the basin (F-GNSP for the main fault system and F-VL & F-Sx for its two segments, see Fig. 4). This fault system presents a constant dip to the North (70–80°), reduced to about 35° with increasing depth.

### 3.3. Update of the 3D simulation model (Spectral Element method)

3D simulations were performed with a spectral element code including an improved meshing and velocity homogenization strategy, surface topography and intrinsic attenuation. The size of the computational domain is 69 km×69 km in the horizontal directions, and extends from the surface (with an elevation with respect to sea level varying from –6 m to 1181 m) down to a constant elevation plane at 30 km depth. The spectral element mesh was obtained using a robust, semi-automated procedure which produces a geometrically conforming, unstructured mesh of hexahedral elements, the sizes of which were tuned for a maximum frequency of 4 Hz and the associated wavelengths. The new model is characterized by a “double-gradient” velocity model, characterized by a first linear gradient from 130 m/s at surface



**Fig. 5.** Thickness of the whole sedimentary (quaternary+tertiary) series as derived from the 3D compilation of available data in 3D geomodelling. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 3**

$V_s$ ,  $V_p$ ,  $\rho$ ,  $Q_s$  and  $Q_p$  “anchor” values used to build E2VP2 properties model within the basin.

	$V_s$	$V_p$	$\rho$	$Q_s$	$Q_p$
Surface	130	1500	2075	$=V_s/10$	$=\max [V_p/20, V_s/5]$
M/P limit	475	2100	2130		
Bedrock top	800	2700	2250		

to 475 m/s at an intermediate surface within the basin, and a second linear gradient from this intermediate surface (475 m/s) to the bedrock interface (with a sediment base velocity of 800 m/s). This intermediate surface, corresponding to a gradient change without velocity jump, could be associated to the Mygdonian/Premygdonian (M/P) limit. The corresponding velocity, unit mass and quality factors are listed in Table 3. The only abrupt velocity contrast thus corresponds to the sediment/bedrock interface, where the S-velocity jumps from 800 m/s to 2400 m/s. As specified in Maufroy et al. [33], the mesh does not follow neither the discontinuity of material parameters at the sediment-bedrock interface, nor the discontinuity of their first-order derivatives at the limit between the Pre-Mygdonian and Mygdonian sediments. Instead, those interfaces were homogenized by vertically averaging the S- and P- slownesses and mass density, as suggested in [10]. The size of the homogenization window is  $L=25$  m, about half the minimum element size. The mesh contains about 6.5 million elements and, since the polynomial order is set to  $N=4$ , about 435 million grid points. The element size is kept smaller than the local minimum wavelength, that is, smaller than 50 m in the shallower basin part and 450 m in the bedrock, insuring at least 5 grid-points per wavelength for frequencies up to about 4 Hz.

For consistency, the bedrock velocity model has been taken identical to the velocity model used for the event relocation [23].

### 3.4. Ground motion simulations

This updated and improved model was then used for the simulation of the expected ground motion for various sets of events and receivers:

- A first set (“S1”) was dedicated to the validation, i.e., the comparison between predictions and observations. It consists of the 19 selected

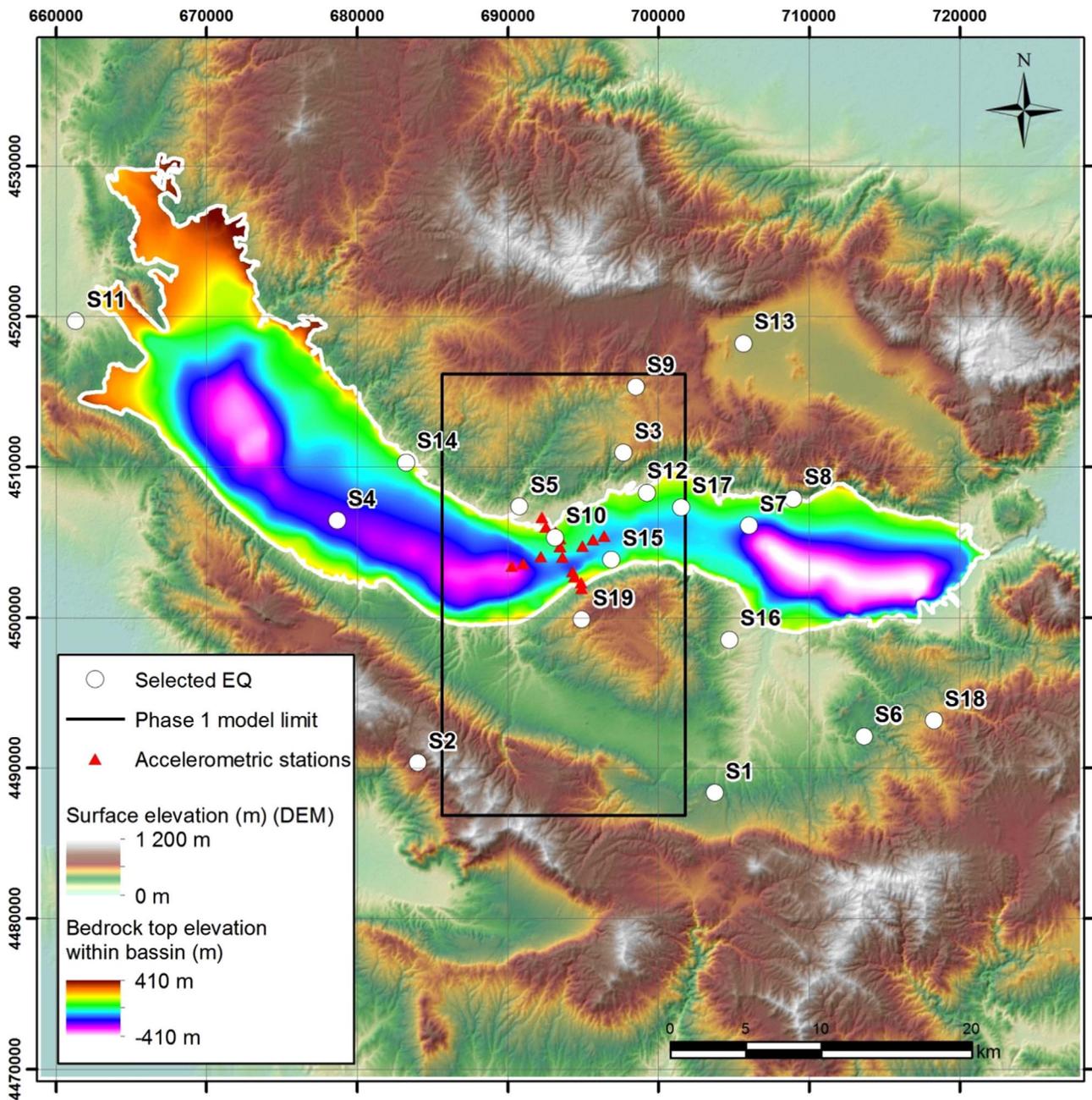
events, with their actual, improved source parameters (magnitude range: 2.7–4.6; distance range: 0–30 km, as listed in Table 2), computed at the 15 receivers corresponding to actual strong motion instruments, as shown in Fig. 6.

- A second set (“S2”) was dedicated to the investigation of the sensitivity of ground motion to the exact source location: each of the 19 events was recomputed for a total of 27 hypocentral positions, considering all combinations with a shift of the actual hypocenter by  $\pm 2$  km in each X, Y and Z direction. The resulting shifted hypocenters are thus located within a cubic box centered on the actual hypocenter location (as indicated in Table 2), and with a 4 km long edge.
- A large set (“S3”) of  $7 \times 36 \times 5 = 1260$  virtual events arranged in 7 concentric circles from 2.5 to 30 km, 36 back-azimuths ( $10^\circ$  step) and at 5 different depths from 2.5 to 15 km was considered to perform a comprehensive investigation of the sensitivity of ground motion and site response to source location (i.e., distance, depth, and backazimuth) in a fully 3D environment. The corresponding focal mechanisms were randomly generated following a Gaussian distribution around the “average” normal faulting parameters in the Mygdonian basin area: strike= $86 \pm 18^\circ$ , dip= $52 \pm 15^\circ$ , rake= $-101 \pm 51^\circ$ .
- Out of this “S3” set, a subset “S4” was extracted corresponding to a set of 52 actually occurred events (“real catalog”, see below), which could not however be all used for the validation as a) the corresponding number of recordings was often too small, and b) the focal mechanism could not be determined with enough accuracy.

The three sets (S1–S3, including thus S4) were computed for the 15 receivers using the reciprocity theorem, which allows to limit the number of simulations to 3 times the number of receivers and thus saves computational time when the number of sources exceeds the number of receivers. As detailed in Causse et al. [34], the spatial derivatives of the 3D Green’s functions are then convolved with the moment tensors to obtain the time series at the 15 receivers. The epicenter locations of sets S3 and S4 are displayed in Fig. 7.

## 4. New validation results

The set S1, corresponding to the nineteen seismic events listed in



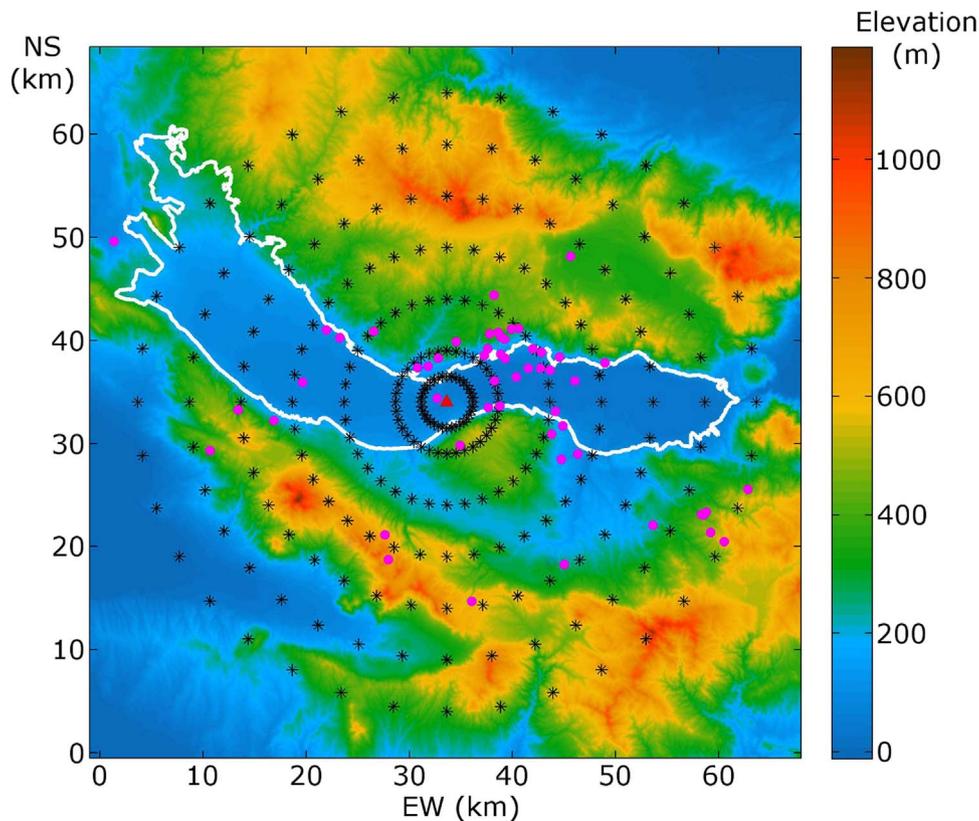
**Fig. 6.** Map of the whole model used for E2VP phase 2 modeling (box of 69×69 km), compared with the location of the area of the “phase 1” modeling box. The surface topography DEM and elevation of the top of the bedrock within the basin are also displayed. Also shown are the location of the 15 accelerometric stations used for the validation (red triangles), the 19 real events selected for the comparison between numerical predictions and actual recordings (white circles, with numbers referring to event IDs in Table 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2 and Fig. 3, well recorded by the Euroseistest accelerometric array, is considered in this section. The 3D numerical simulations of the 19 events are performed with the code EFISPEC3D (De Martin [35]) implementing the Spectral Element Method. They include the effects of surface topography and of frequency independent intrinsic attenuation, which was modeled using a Zener body with 3 relaxation mechanism distributed between 0.1 Hz and 4 Hz, as detailed in Moczo et al. [36].

Pre-processing of the data to perform the validation exercise includes: (1) filtering the real data with a Butterworth filter between 0.3 Hz (order 6) and 3.0 Hz (order 10), in order to get a similar frequency content between the recordings and the synthetics without contamination by high energy transients at frequencies higher than the maximum simulation frequency (4 Hz); (2) synchronization of record-

ings with the corresponding synthetics on the first *P*-wave arrival; (3) all couples of signals to be compared are cut to the same length in duration; and (4) a study of the signal-to-noise (SNR) ratio is performed on all recordings to determine the frequency band where that ratio is greater than 3: only recordings fulfilling such a SNR criterion over a frequency band [0.7–4 Hz] have been considered for the validation.

Differences between numerically-simulated seismograms and earthquake recordings were quantified as for E2VP Phase 1 on the basis of the 5 ground-motion parameters mentioned above. Arguments for the selected characteristics, details on their computations and on the way to handle the horizontal components are provided in Maufroy et al. [12]. The main results of this “blindly- oriented” validation exercise are summarized below, starting with the rock sites.



**Fig. 7.** Epicentral location of the “virtual” seismic sources considered in the numerical study. The response of the Mygdonian basin (bold white line) is computed for 1260 virtual sources (black circular crosses) located on seven concentric circles around the central soil site TST indicated by the red triangle, with back-azimuths equally distributed every 10°, and five different depths. 52 of these virtual locations are very close to real earthquakes recorded at some of the accelerometric stations: the epicentral locations of those 52 real events are depicted by the magenta dots. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

#### 4.1. Validation results at rock sites

Results of the comparison between actual recordings and their numerical predictions at 3 rock sites in Euroseistest are given in Table 4. PRO is the northernmost station on a rock outcrop in the Profitis village, STE is the southernmost station on a rock outcrop south of the Stivos village, and TST5 is the deepest borehole sensor (depth 196 m) in the central vertical array. The level of misfit obtained on the 5 ground motion parameters  $C_i$  of the E2VP evaluation procedure is expressed by the average misfit computed for each of them from the selected events that were recorded at the corresponding rock site. The misfit values obtained here are similar or below in absolute value to the misfits obtained in the first validation phase at PRO and STE rock sites, giving a first confirmation that the surface ground motion outside the basin is in general well predicted by the numerical simulations. Only the borehole site TST5 exhibits anomalously high misfit values at the highest frequencies considered in the validation.

To get another viewpoint on the level of misfit outside the basin, Fig. 8 displays the Fourier spectral ratios computed between recordings and their numerical predictions for the events recorded at these 3 rock

**Table 4**

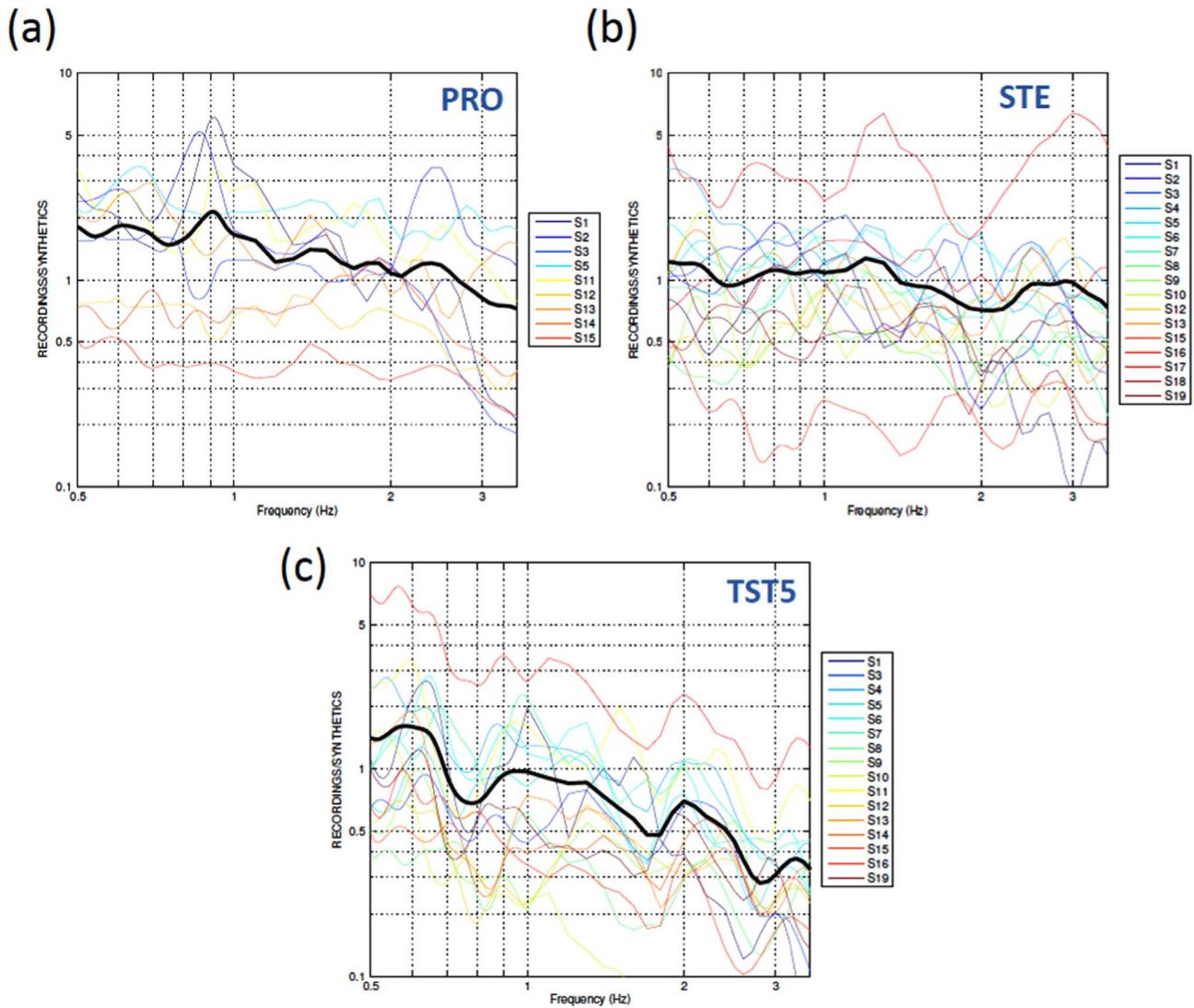
Average in % of horizontal misfits on the E2VP evaluation criteria between the actual recordings and their numerical predictions at 3 rock sites: northern rock site PRO, southern rock site STE and borehole TST5. The number of events recorded by each station and considered in the average is indicated in the last column.

	C1 PGA	C2 2.0 Hz	C3 0.5 Hz	C4 CAV	C5 RSD	Number of events
PRO	-21	-24	-54	-9	-73	9
STE	34	39	1	8	-124	17
TST5	128	129	53	88	-161	16

sites. Concerning northern rock site PRO (see Fig. 8a), the median ratio of observed ground motion over predicted is globally satisfactory (i.e., around and close to value 1), except for the lowest frequencies that are under-estimated. Reason for that low-frequency under-estimation of the ground motion at PRO is not yet understood. At southern rock site STE ( Fig. 8b) the median ratio is satisfactorily close to 1 in all frequencies. It is noteworthy however that the validation results at STE can be bad for a few events (the colored lines giving the result for each individual event are far from the value 1 in a few cases).

At borehole site TST5 (Fig. 8c) that is located close to the sediment-basement interface at 196 m depth, the actual ground motion appears to be significantly over-estimated by the synthetics at high frequencies. This can obviously significantly impact the ground motion prediction at the surface and center of the basin. Several tentative explanations can be considered. One possibility could be the new crustal-propagation model, as the model from Papazachos [37], used in E2VP phase 1, was later replaced by the model from Novotný et al. [23] also used in the improved characterization of the seismic sources. A comparative study for the two models, detailed in Maufroy et al. [24], indicates the E2VP2 crustal model induces only a slight increase (around 20% in general, possibly reaching up to 60% over some narrow frequency bands) of rock motion compared to the previous E2VP model. This effect, although it could participate, cannot fully explain the high-frequency overestimation found at borehole site TST5.

Another possibility is that borehole site TST5, located right below the sediments, might be significantly and inappropriately affected by the basin propagation in the new basin model of E2VP phase 2. The strongest argument in favor of that hypothesis is found in Fig. 8c. The average ratio (solid black line) exhibits, in addition to the overestimation trend at high frequencies, three troughs at frequencies around 0.75 Hz, 1.7 Hz and 2.7 Hz: these frequencies do coincide with the first 1D resonance frequencies in the basin at TST0. Such an excess



**Fig. 8.** Fourier spectral ratios between recordings and their numerical predictions at 3 rock sites: (a) northern rock site PRO, (b) southern rock site STE and (c) borehole TST5. Each colored line corresponds to one seismic event of the validation. The solid black line indicates the average ratio in each panel. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

footprint of the frequency-dependent site effect at TST, together with the overestimation trend at high frequency at TST5, might be indicative of a too weak attenuation in the new basin model, resulting in a too energetic feedback of energy radiating from the basin into the bedrock, especially at high frequency. Such an hypothesis should be also associated with an overamplification at TST0, which is partly the case, as seen in the following.

Finally a third possibility comes from the absence of scattering in the numerical model, the bedrock consisting of horizontally layered, homogeneous media. The signal duration (parameter C5) is found indeed to be under-predicted for each rock site, which is consistent both with the over prediction of amplitude parameters especially at high frequencies, and the absence of scattering.

#### 4.2. Validation results in the Mygdonian basin

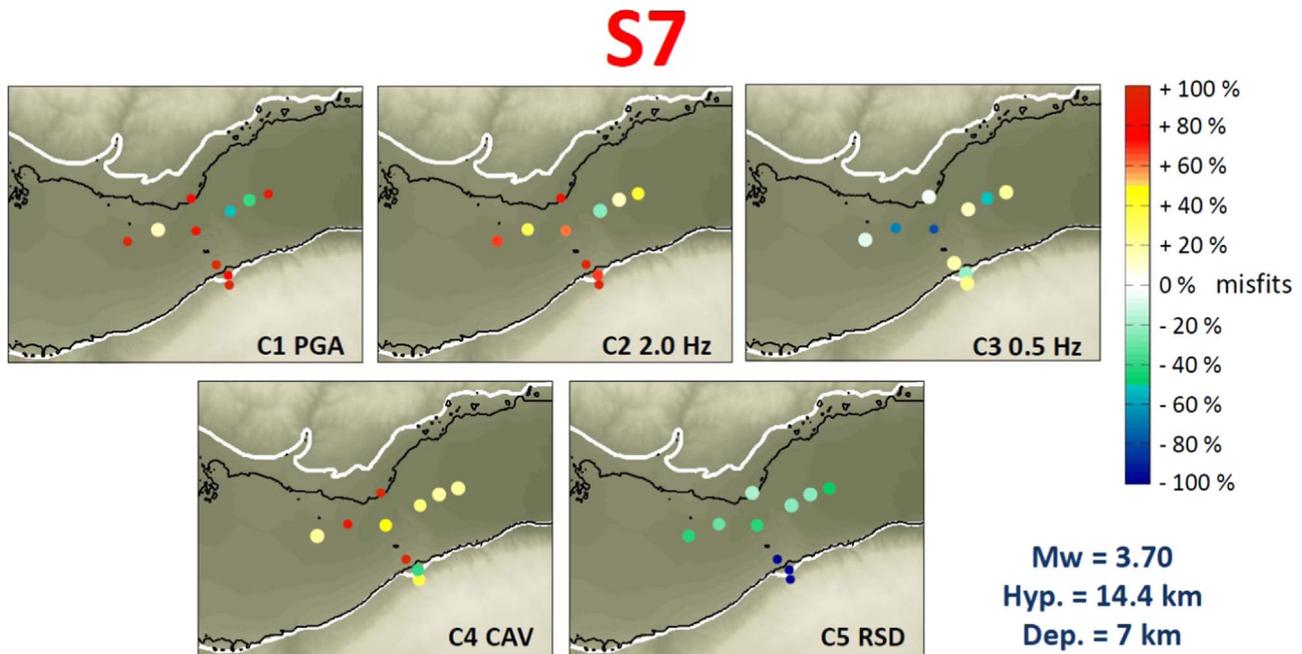
The same evaluation procedure has been applied to all stations for all selected events. The validation results have been detailed for each event, each site and each considered ground motion parameter as illustrated in Fig. 9 on the example of event S7 (one of the most satisfactory as seen in Table 5). They can also be summarized for each site and ground motion parameter as listed in Table 5 for the central site TST0. The detailed results for each event are available in Maufroy

et al. [24].

In most cases the ground motion in the basin is significantly overestimated, while the signal duration (parameter C5) is almost systematically under-estimated. The last parameter is tightly related both to regional scattering – not accounted for in the modeling –, and to damping within the sediments, which is not constrained by any measurement. One way to investigate the preferential origin of this duration under-prediction is to have a special focus on the basin response.

Following the procedure of the first validation phase (Maufroy et al. [12]), “hybrid” time histories are computed to further investigate the ability of the numerical predictions to predict the site-response component of the ground motion. Hybrid time histories at TST0 are obtained by convolving the recorded signal at TST5 (which thus includes the actual source and path terms, in particular all the delayed arrivals from crustal scattering) with the site-effect part coming from the simulation (synthetic borehole-surface transfer function computed for the same event).

Hybrid time histories maximize the impact of numerical estimate of site-effect component and minimize the effect of uncertainties in source description or in crustal propagation. The validations misfits corresponding to such hybrid time histories are also listed in Table 5: such an approach can be found to significantly improve the fit on the



**Fig. 9.** Example map of horizontal misfits on the E2VP evaluation criteria between the recordings of the real event S7 (see Table 2 and Fig. 6) and its 3D numerical prediction. C1 is based upon peak ground acceleration, C2 upon elastic spectral acceleration ranging 1.5–3.0 Hz, C3 upon elastic spectral acceleration ranging 0.375–0.750 Hz, C4 upon cumulative absolute velocity and C5 upon 5–95% relative significant duration (see [12] for details). Each colored dot corresponds to the misfit obtained at the corresponding real surface receiver. Red/yellow tones are for overestimation of the recordings by the prediction; blue/green tones are for underestimation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

amplitude-frequency parameters C1–C3, and on the energy parameter C4 as well – with however a remaining trend for over-prediction –, while the average misfit on duration C5 is moved from under-prediction to over-prediction. The misfit values obtained on criteria C1 to C3 for the hybrid time histories are significantly lower than for the full synthetics, being in absolute value closed to the values obtained in validation phase 1 (see Fig. 14 in Maufroy et al. [12]). The typical average misfit values encountered in phase 1 for hybrids were ranging from –20% to –50%. In the second phase, the corresponding average values equal 27%, 29% and 1% on C1–C3 respectively. However, when not taking into account events S10 and S15 (identified as being bad candidates for validation due to their proximity and low magnitude), those values decrease to 19%, 16% and –22% on C1–C3 respectively (–

48% on C3 when considering only events with signal-to-noise ratio greater than 3). This represents a general improvement of the validation results on the site-effect component.

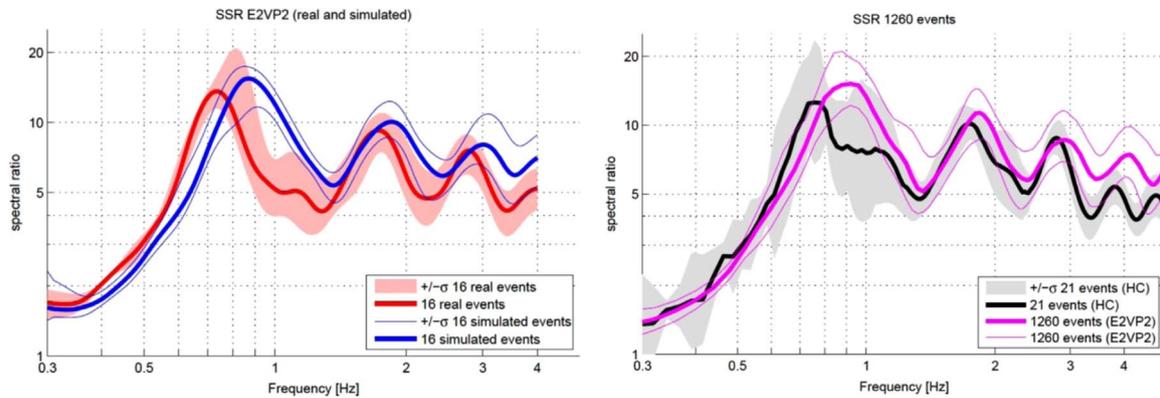
One important change from validation phase 1 to phase 2, is that the site-effect component was globally under-estimated in phase 1 (negative average misfit values for hybrid time histories at TST0), while it is now mostly over-estimated in phase 2 (positive average misfit values on C1 and C2). This observation, combined with the average misfit on duration, also supports the probable under-estimation of the damping value within the sediments.

The slight over-prediction of site response estimates also appear clearly on Fig. 10, which displays a comparison between the observed and computed surface/downhole transfer functions for various event

**Table 5**

Values of average horizontal misfits on the five engineering parameters  $C_i$  between the actual recordings at central soil site TST0 and their numerical predictions. Values in % evaluate the predictions by full synthetics vs. hybrid time histories (i.e. site response only). For each parameter  $C_i$ , the average is computed over the 16 events that were recorded both at TST0 and TST5, and over the 14 events excluding the worst ones S10 and S15.

	FULL SYN. C1	HYBRIDS C1	FULL SYN. C2	HYBRIDS C2	FULL SYN. C3	HYBRIDS C3	FULL SYN. C4	HYBRIDS C4	FULL SYN. C5	HYBRIDS C5
S1	154	10	129	11	27	–58	196	72	–39	92
S3	177	19	186	39	89	–48	156	13	–93	–63
S4	142	62	128	84	78	52	42	95	–169	–6
S5	147	32	166	35	40	–12	76	–4	–93	1
S6	128	81	117	104	–22	106	68	198	–44	98
S7	87	–24	63	–19	–77	–90	47	25	–41	64
S8	130	12	173	12	156	–73	192	62	–61	47
S9	186	–3	192	20	164	15	152	34	–70	27
S10	471	75	506	98	431	203	365	94	–205	101
S11	68	35	63	32	–65	4	61	66	4	47
S12	201	5	201	29	136	–72	146	–16	–48	–26
S13	243	46	153	–47	5	–134	123	–47	–109	–19
S14	234	2	214	–5	188	–1	184	–15	–57	15
S15	310	92	338	134	262	129	253	114	–102	29
S16	2	9	–119	–120	–143	–17	–72	31	–71	47
S19	99	–15	139	50	58	19	126	146	–69	53
AVERAGE	173	27	166	29	83	1	132	56	–79	32
AVERAGE without S10 S15	143	19	129	16	45	–22	107	50	–69	27



**Fig. 10.** Left: Median of SSR (Standard Spectral Ratios) at TST0 with TST5 as reference station, derived from the actual recordings of the 16 events of set S1 (solid red line, associated variability shown in pink), and for the corresponding E2VP2 simulations (bold blue line, associated variability shown by thin blue lines). Right: the same for another set of recordings (black line, 21 events with the best sigma to noise ratio, including more distant events outside the model box) compared with the predicted response for set S3 (solid magenta line, 1260 events on concentric circles and with various depth). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

subsets. Fig. 10a, corresponding to the 16 events of set S1 for which the TST0/TST5 ratio is available, exhibits a slight over-estimation of the site response by the E2VP2 numerical simulation, especially at frequencies between 1 and 1.5 Hz, together may be with a slight over estimation of the fundamental frequency. Fig. 10b compares the “best” instrumental estimate of TST5/TST0 transfer function, derived from a set of 21 pairs of recordings with high signal-to-noise ratio (including events from outside the model box), with the average transfer function obtained for the S3 simulation set (1260 events). An over-prediction in the same frequency range is still present, but it is also associated with a significantly larger variability of the site response (especially the observed one), and significant differences in the average instrumental ratios from the two sets of recordings. Although different, the average numerical and instrumental transfer functions are almost entirely located within the 16–84% variability range (i.e. within  $\pm 1$  standard-deviation) of each other, except around 1 Hz and beyond 3 Hz. The latter is probably indicative of an under-estimation of the damping in the sediments, while the former is probably a consequence of the high sensitivity of ground motion and site response to the source location (distance, depth and back-azimuth), which is investigated further in the next section.

## 5. Sensitivity studies and insight into the structure of the aleatory variability

One of the important outcome of the verification and validation exercise is the significantly smaller code-to-code distance compared to the code-to-data misfit. The latter is interpreted as resulting from errors or uncertainties in the source parameters and on the propagation model. It is therefore fully legitimate to use the numerical simulation approach to investigate, in a relative way, the sensitivity of ground motion and site response both to the variability in source parameters (i.e., the variability for a wide range of different hypocentral locations: epicentral distance, depth and backazimuth) and to the uncertainty in source parameters (i.e., to small changes in the magnitude and location parameters comparable to hypocentral location error). The synthetics obtained with the simulation approach for a large number of sources and receivers can also be used to generate some synthetic GMPEs, and to analyse the impact of the source uncertainties on the value of the aleatory variability in the light of the results of the sensitivity studies, investigating in particular how epistemic variabilities affect the within and between-event aleatory variability components.

### 5.1. Epistemic variability of site response

It is most often considered - at least implicitly - that site response

can be decoupled from source and path effects. This section takes advantage of the S3 simulation set to investigate the sensitivity of the site response to some simple source-receiver attributes (back-azimuth, depth and distance) linked to the source-site crustal path and incoming wavefield, which may be a priori thought to impact the site response in a highly 3D environment. As a comprehensive analysis of this simulation set is provided in Maufroy et al. [33], only a short insight is provided here in relation with the impact of epistemic variability of site response on the “apparent” within-event aleatory variability. Fig. 11 displays the variability of surface/borehole transfer function with source back-azimuth and epicentral distance:

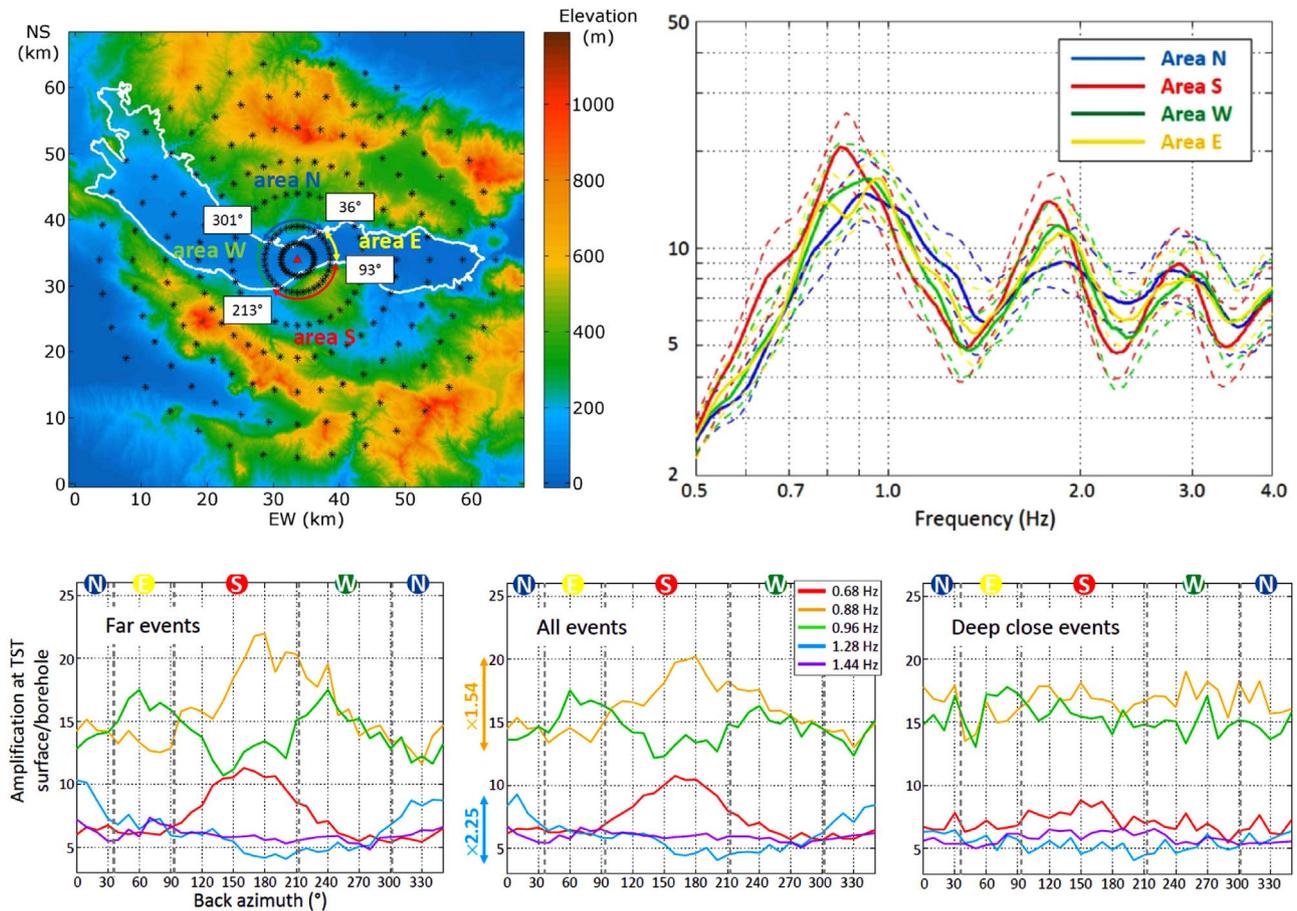
The amplification is found to exhibit a noticeable dependence on source back-azimuth. It is slightly larger around the two main peaks (0.7 Hz and 2 Hz) for southern events, and exhibits a smoother frequency-dependence for northern sources. The largest back-azimuth dependence is found for shallow, far sources while the smallest one consistently corresponds to deep, close sources (Fig. 11 bottom). This variability is related to the actual 3D geometry of the sediment-basement interface, with gentler slopes on the northern edge, and steeper slopes on the southern edges.

This dependency is frequency dependent: the variations with back-azimuth are the largest for the intermediate frequency range (around and just above the basin fundamental frequency), in between the fundamental and first higher 1D resonances, corresponding to the band mostly affected by edge-generated surface waves: their energy strongly depends on the incidence (source-depth) and back-azimuth of the incoming wavefield, in relation with the complex 3D geometry of the sediment-basement interface.

Though not shown here, the variability in site response is also found to be significantly larger when the reference is an outcropping rock at 2–3 km distance from the considered site, while it is minimum when the reference is at depth in the bedrock beneath the considered site. Vertical arrays are thus to be recommended, even though the “reference” motion at deep bedrock may significantly differ from an outcropping rock motion.

### 5.2. Sensitivity of ground motion to source location uncertainties

Besides the variability of site response related to large changes in back-azimuth and distance, we also investigated the impact of small changes in source location, mimicking the actual uncertainty in source location, on the variability of ground motion. The background objective is to provide a better guidance for the selection of appropriate events for future validation exercises, but it also improves our understanding of the aleatory variability of GMPEs. In that aim, we have considered the “S2” set, where the hypocentral locations of the 19 events of set S1 are moved at 27 different positions within a 4 km edge cubic box



**Fig. 11.** Example impact of the source distance and backazimuth on the average amplification at the central TST site. Top left=Map of the basin and source location considered in this study. The basin contours are indicated by the bold white line and the surface elevation is given by the color scale. The location of the central receiver TST is shown by the red triangle, and the sources epicenters are shown by the circular setting of black crosses. Ground motions at TST are analyzed by considering 4 back-azimuth areas (N, E, S, and W) as described by the colored circular arrows; the areas are separated by the 4 back-azimuth values at TST (degrees labels) that correspond to the basin edges. Top right=average surface/borehole transfer functions for the 4 different back-azimuth ranges (solid lines=average, dotted lines=average  $\pm$  one standard deviation). Bottom frames: details on the sensitivity to backazimuth for five different frequencies (color code), and three different event subsets, as indicated in the different frames (far events, all events and deep close events. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

centered on the best location estimate. The  $\pm 2$  km variability in each  $x$ ,  $y$  and  $z$  direction is considered a reasonable and probably minimum estimate of the actual location uncertainty, especially for small magnitude events. For each event and each site, the variability of ground motion was estimated from the standard deviation of the acceleration response spectra for the 27 different hypocentral locations, at each frequency from 0.5 to 3.5 Hz, and then averaged over the whole frequency band. For each event, the site variabilities were averaged to provide an index of the “event” variability linked to the source location; this averaging was done separately for the four rock stations and the 11 sediment stations within the graben. The resulting values are summarized in Table 6, Fig. 12 displays the variability of acceleration spectra at site TST0 on the example of 4 different events corresponding to different depths and epicentral distances: event S10 is shallow (4 km) and close (4.2 km hypocentral distance), event S05 is deep and close (11 km and 11.9 km, respectively), event S11 is shallow and more distant (4 km, and 36 km, respectively), and event S02 is deep and distant (11 km and 20 km, respectively, see Table 2 and Fig. 3). Significant differences appear between the 4 events: the smallest variability is found for distant, deep, event S02, while the largest corresponds to the closer, shallow event S10.

A detailed look at Table 6 indicates that the variabilities at sediment and rock sites are very similar, and seem to be much more related to source location than to site response. This is further illustrated in Fig. 13, which compares the overall variabilities at sediment sites and at rock sites, together with the variability of the site transfer function. It

shows that sediment and rock variabilities are comparable: their differences remain small compared to the large event-to-event variability of the sensitivity of ground motion to exact source location. Moreover, the ground motion variability is systematically larger, to much larger, than the corresponding variability of the rock-to-sediment transfer function: the former range from 0.07 to 0.30, while the latter lie between 0.04 and 0.10. Fig. 14 confirms that the ground motion variability is tightly related to the source distance and hypocentral depth: it is the largest for shallow, near sources, (variabilities larger than 0.25 for depths smaller than 5 km and epicentral distances smaller than 10 km) and the smallest for distant, deeper sources (variabilities below 0.10 for distances larger than 10 km and depth beyond 10 km, or distances beyond 20 km and depths larger than 8 km). One may notice that the site response variability also exhibits a slight decrease with increasing epicentral distance and hypocentral depth; it is however much less pronounced than for the response spectra. It may therefore be concluded that validation exercises focusing on absolute ground motion are much more difficult for very close events or very shallow, local events, unless there is a very dense local seismological array that allows to locate the events with a precision much smaller than the  $\pm 2$  km uncertainty considered here. Nevertheless, if the validation target is the site transfer function, even close and shallow events can be used.

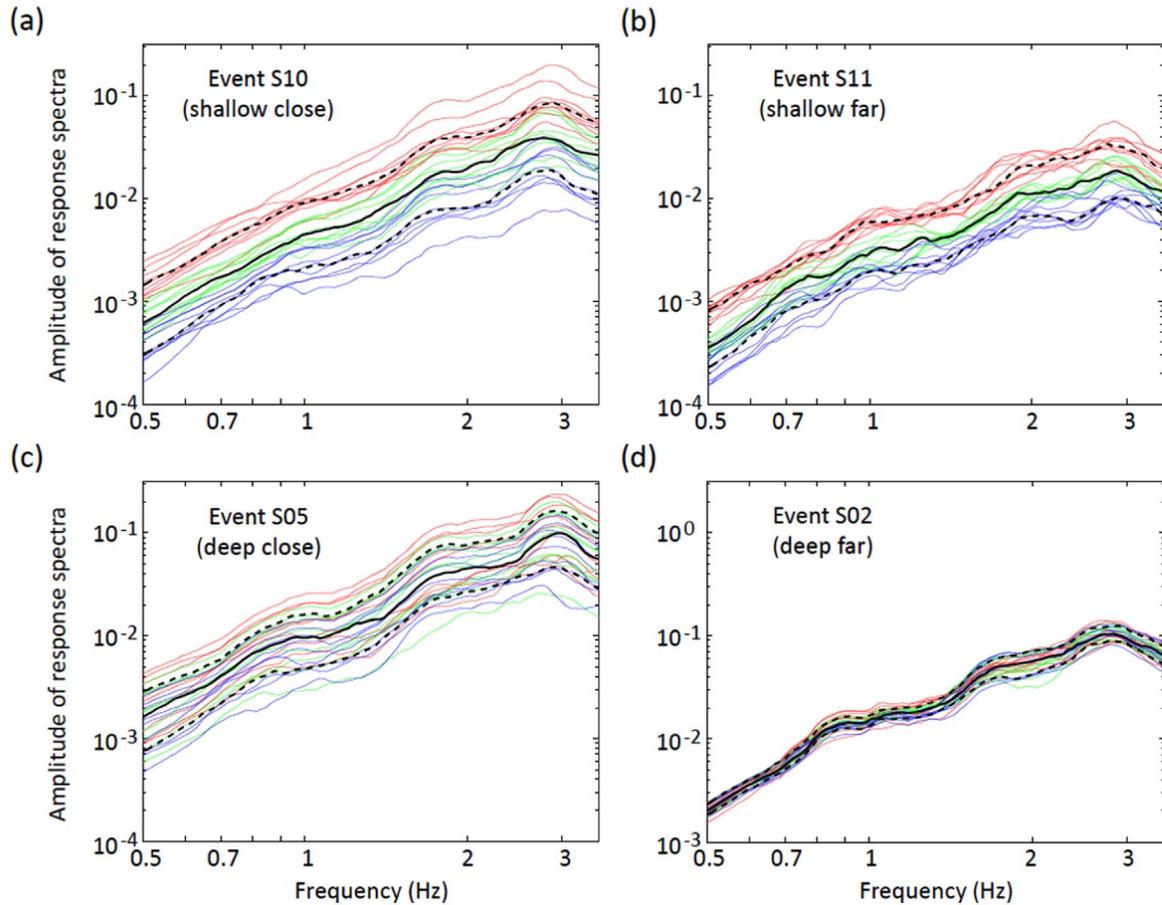
**Table 6**

Sensitivity of ground motion to hypocentral location uncertainty. For the 19 events of set S1 (rows), the table lists the value of the average standard deviation (log10 scale) of the computed acceleration response spectra over the frequency range [0.5–3.2 Hz] for the 4 rock stations (Column #2), the 11 sediment stations (column #3) the deep borehole site at graben center (TST5, column #4) and the surface site at graben center (TST0, column #5); finally column #6 lists the corresponding values of the average variability of TST0/TST5 Fourier transfer functions over the same frequency range.

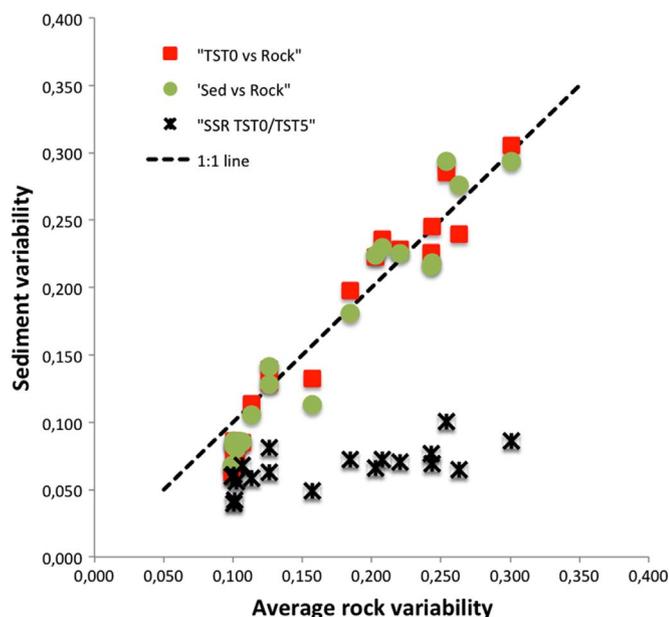
Event ID (S1 set)	Average rock	Average sediment	TST5	TST0	TST0/TST5
S01	0,126	0,141	0,131	0,129	0,081
S02	0,100	0,068	0,074	0,062	0,060
S03	0,101	0,082	0,089	0,086	0,040
S04	0,101	0,085	0,109	0,079	0,042
S05	0,244	0,219	0,294	0,246	0,069
S06	0,102	0,081	0,074	0,072	0,056
S07	0,208	0,230	0,241	0,236	0,072
S08	0,126	0,128	0,162	0,140	0,063
S09	0,263	0,276	0,234	0,240	0,065
S10	0,301	0,294	0,308	0,305	0,086
S11	0,203	0,224	0,217	0,223	0,066
S12	0,221	0,225	0,225	0,228	0,070
S13	0,114	0,106	0,124	0,114	0,058
S14	0,107	0,086	0,093	0,085	0,068
S15	0,254	0,294	0,271	0,286	0,101
S16	0,184	0,181	0,217	0,198	0,072
S17	0,157	0,113	0,171	0,132	0,049
S18	0,101	0,086	0,077	0,076	0,061
S19	0,243	0,216	0,242	0,226	0,077

### 5.3. Contributions to the understanding of aleatory variability

The previous sections show that the predicted ground motion exhibits an epistemic dependence on the source location, that the 3D site response varies with the source back-azimuth, and that the sensitivity to tiny changes in the source position very significantly impact the ground motion at short distances or for shallow sources. Such a complex and coupled dependence is not taken into account even in the latest, most sophisticated GMPEs. Considering that results of carefully verified numerical simulation codes are reliable, we thus used the results of the comprehensive set of 1260 virtual sources to derive synthetic GMPEs and to investigate how the above mentioned, well-identified, epistemic variability maps into the aleatory variability  $\sigma$  (and its within-event,  $\phi$ , and between-event,  $\tau$ , components). Basically, for each considered source, the numerically derived Green's functions were convolved with *ad hoc* source functions in order to simulate earthquakes with magnitude in the range 2–5. For the subsets corresponding to real events (S4 set corresponding to 52 real events), the magnitude was tuned to the real one; for all the other “virtual” events of the S3 set, the magnitude was assigned arbitrarily, in the limited range 2–5 however to be consistent with the point source assumption. These various sets of scaled synthetics were then used to derive GMPEs using the artificial neural network (ANN) approach described in Derras et al. [38] and Derras et al. [39]. The “standard” explanatory variables were the moment magnitude  $M_w$ , the epicentral distance  $D_{epi}$ , the hypocentral depth  $Z$ , and the  $V_{S30}$  site proxy. Alternative site proxies were also considered (fundamental frequency  $f_0$ , local sediment thickness  $h$ , average sediment velocity  $V_{sh}$ ), and additional source parameters as well (mainly the source back-azimuth



**Fig. 12.** Variability of the acceleration response spectra at TST0 site resulting from a  $\pm 2$  km variability in source location, for 4 different events of the S1 set. The color code corresponds to the hypocentral depth  $Z$  (red for  $Z=Z_0-2$  km, green for  $Z=Z_0$  and blue for  $Z=Z_0+2$  km). The median response spectra is given in each panel by the solid black line, surrounded by the upper (84%) and lower (16%) percentiles as dashed black lines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 13.** Variability of the ground motion predictions resulting from a  $\pm 2$  km variability in source location, for the 19 different events of the S1 set. This variability is expressed in terms of standard deviations over the 0.5–3.3 Hz frequency range, of the predicted response spectra and TST0/TST5 transfer functions as listed in Table 6. The different symbols correspond to the average variability of response spectra at the central TST0 site (red squares), and for the 11 sediment sites (green circles), together with the variability of the TST0/TST5 transfer functions (black stars) plotted as a function of the corresponding average variability of response spectra for the 4 rock sites (abscissa). Each symbol corresponds to one of the 21 events of the S1 set. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

BAZ). The objectives were multifold:

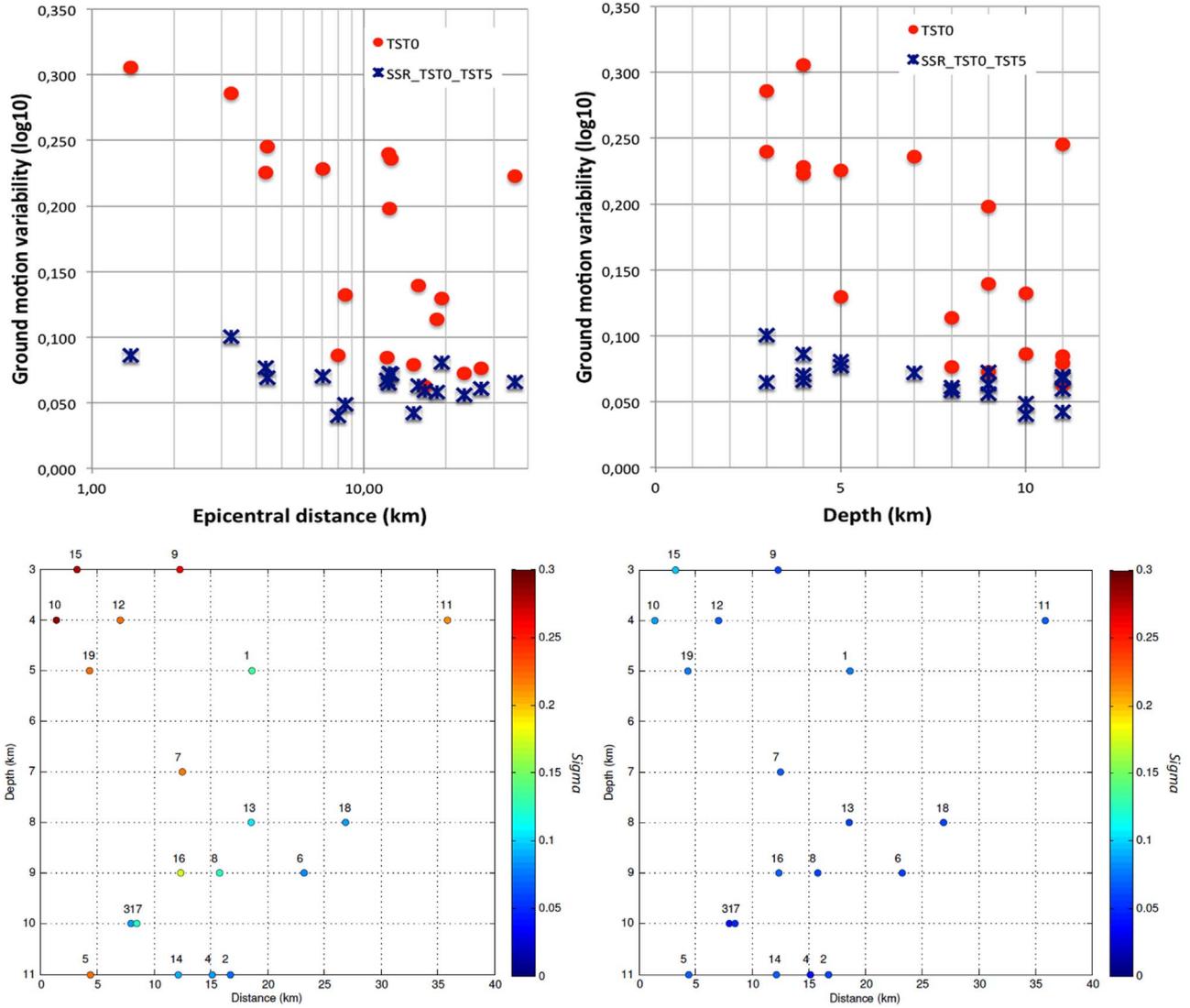
- Compare the within and between- event variability levels between synthetics and real data (set S4).
- Investigate the effect of the size of the data set (small, S4, versus large, S3) on the values of within- and between-event variabilities.
- Investigate the effect of “controlled uncertainty” on magnitude or source localization on the between-event variability.
- Investigate the impact of various site proxies ( $V_{S30}$ , fundamental frequency  $f_0$ , local sediment thickness  $H$ , average sediment velocity  $V_{SH}$ ) and of the corresponding uncertainties on the within-event variability.
- In such 3D basins, investigate the possible use of other source/site information, such as back-azimuth, in view of somewhat reducing the aleatory variability.

The results obtained so far are partly illustrated in Figs. 15 and 16 for two simple ground motion parameters, pga and pgv. Similar analyses were performed also for spectral ordinates at various oscillator frequencies (from 0.5 to 3.5 Hz), but are not significantly different from what is obtained for pga and pgv (note however that, as numerical predictions are valid only up to 4 Hz, “pga” corresponds to a limited frequency range). The main findings are summarized below.

- i) The within and between event variabilities do exhibit a significant dependence on the data set. Fig. 15 displays the total, between- and within-event variabilities obtained for pga and pgv with a neural network considering moment magnitude, epicentral distance, depth and  $V_{S30}$  as explanatory variables, and for the two sets S3 and S4. All variabilities are shown to increase with the size of the data set. The between-event variability is found to be much larger (i.e., by about 50%) for the full S3 set than for the reduced S4 set; the size of the data set also slightly impacts the within-event variability, but to a much smaller extent (around 10–20%). The large increase of the between-event variability can be explained by

the strong sensitivity of ground motion to hypocentral depth and distance. The increase of the within-event variability is consistent with the fact the S4 subset exhibits a skewed azimuthal distribution (see Fig. 7) and the observed sensitivity of site response to back-azimuth (Fig. 11). Such findings raise the attention on the need to consider a large set of recordings at a given site to have a reliable estimate of the within-event variability: in S3 set there are 1260 recordings per site, and 52 in S2 set. It thus invites to be cautious when working on too small data sets: both the between- and within-event variabilities may be underestimated, which may in turn also impact the estimates of single-site sigma. This should definitely be kept in mind in the derivation of GMPEs.

- ii) It is also worth mentioning (see Maufroy et al. [40] for more details) that, for the S4 subset, the values of within and between-event variabilities are found somewhat lower on synthetics compared to their values derived from actual recordings. The corresponding values are listed in Table 7 (second and third columns). One of the reason of the lower within-event variabilities on the synthetics may come from the absence of uncertainty on the source parameters (magnitude and location). Therefore, different levels of uncertainties were artificially introduced in the magnitude and location values, without changing the corresponding synthetics computed with real, unperturbed magnitude and location. Magnitude values were randomly modified using a uniform distribution within  $[m-\Delta m, m+\Delta m]$ , with  $\Delta m$  taken equal successively to 0.1, 0.2, 0.3, 0.4 and 0.5. For each  $\Delta m$  value, ten random sets of magnitude values were generated, and GMPEs derived on the corresponding sets of unchanged synthetics, with unchanged locations and distances, and modified magnitude values. The results are illustrated, on the example of the peak ground velocity PGV, in Fig. 16 left: as expected, the within-event estimate is left basically unchanged, while the between-event variability exhibits a significant, quasi-linear increase with  $\Delta m$ : it is doubled (from 0.12 to 0.24, log 10 values) for  $\Delta m=0.25$ , and tripled (up to 0.36) for  $\Delta m=0.5$ . Similar trends were obtained for other ground motion parameters (pga, spectral accelerations at various periods) and are not indicated here for conciseness.
- iii) A similar analysis was performed to investigate the impact of uncertainties in source location. The actual locations were randomly modified using a uniform distribution centered on the actual one, with maximum deviations  $\Delta l$  in the x, y and z directions varying from 1 to 10 km. For each  $\Delta l$  value, ten random sets of modified locations were generated to avoid any set-specific bias. The right frame of Fig. 16 indicates that the within-event variability remains basically unchanged, while the between-event variability increases noticeably with source location uncertainty: the  $\tau$  value increases by 25% (from 0.12 to 0.15) for a location uncertainty of  $\pm 3$  km, and by 50% (up to 0.18) for a location uncertainty of  $\pm 5$  km. Considering the average magnitude uncertainty (especially for moderate magnitude events) is at least 0.2, and the average location uncertainty is probably ranging from 2 to 5 km, our results indicate that a non negligible amount (i.e., increase by about 50–100%) of the between-event variability comes from source parameter uncertainties. The increase of variability at small magnitude reported in many recent GMPEs may therefore be a consequence of a larger location uncertainty, and investing in dense seismological networks for more precise localization may therefore constitute one of the most efficient ways to reduce sigma, especially for GMPEs including a large number of recordings from small magnitude events.
- iv) The other types of source parameters that were considered for these synthetic GMPEs are the source depth  $Z$  and back-azimuth BAZ. It was found that, for the S4 data set, the  $\tau$  value is significantly reduced when considering  $Z$  (from 0.2 to 0.10, i.e. by about 50%), and further reduced (from 0.10 to 0.075, 25%) when considering BAZ, while the  $\phi$  value remains unchanged. Such



**Fig. 14.** Control of the ground motion variability due to source location uncertainty by epicentral distance and depth. The top row displays the dependence of variability on response spectra at TST0 (red circles) and Fourier transfer functions TST0/TST5 (blue stars) as a function of epicentral distance (left) and hypocentral depth (right). The bottom row displays the values of the variabilities of TST0 response spectra (left) and TST0/TST5 transfer functions (right) in the (epicentral distance/depth) plane: the color code indicates the corresponding variability (same scale for both plots); the labels on each symbol indicate the S1 event ID. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

results cannot be directly extrapolated to classical data sets used for the derivation of GMPEs; they however indicate that the twin parameter set (epicentral distance, depth) performs better in ground motion prediction than the sole hypocentral distance, and that, when considering a single site with a pronounced 3D underground structure, the source back-azimuth could be considered to further reduce the prediction uncertainty.

- v) Finally, it was found that two site proxies,  $V_{S30}$  and  $f_0$ , perform almost equally well to account for site conditions, while the local sediment thickness  $H$  and average sediment velocity  $V_{SH}$ , which are sometimes proposed for alternative site classifications, perform much more poorly. A similar analysis was performed on the impact of uncertainties in the estimates of these site proxies on the within-event variability  $\varphi$ : it was found – on this particular data set – that the impact is very small, much smaller than the uncertainties on source parameters. As the number of sites is limited, and the associated geological conditions as well, further investigations are needed to generalize such results.

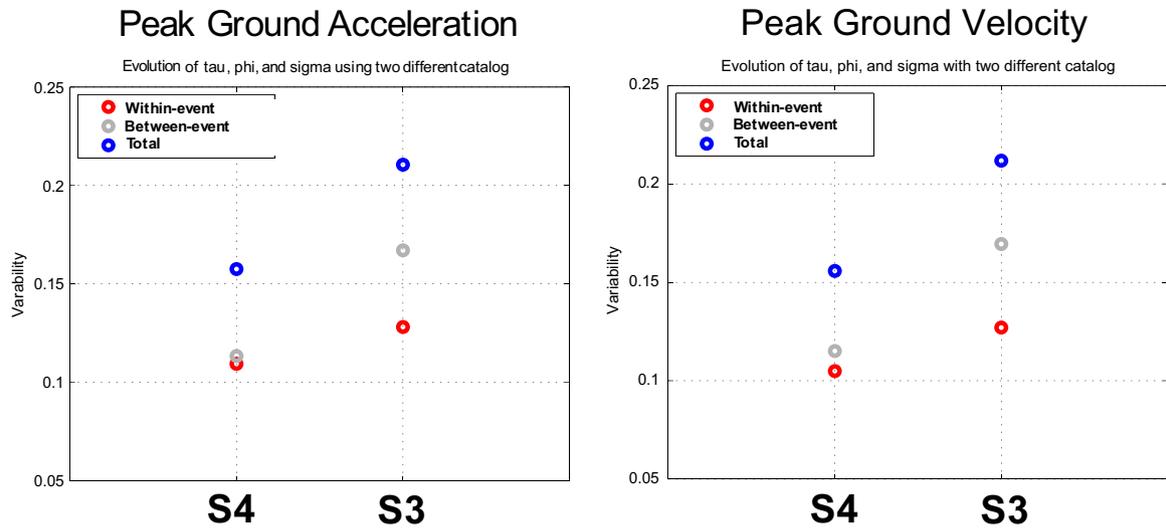
## 6. Conclusions

The use of numerical simulation has proved extremely powerful and useful for improving the understanding of the physics of ground motion from source to site. Using the simulation approach for design purposes requires much care and is much more demanding, especially when going to frequencies beyond 1 Hz. Verification and validation exercises such as E2VP will certainly be repeated in the future on other sites and datasets. In the same way as E2VP benefitted from lessons of previous similar benchmarking exercises, it is important to summarize the main lessons from the present project, and more specifically from its second phase, as the main results from E2VP1 are already summarized in Table 1.

Two important recommendations must first be mentioned, because they appeared as recurrent issues all along the 8-year life of the E2VP:

Numerical simulation codes require careful use and regular cross-checking, which proves to be a very efficient tool in securing the quality of the results, especially after code updates or improvements.

The most important aspects of accuracy of any numerical method and code that is applied for numerical prediction of earthquake ground motion in engineering projects, may be valuably verified through some



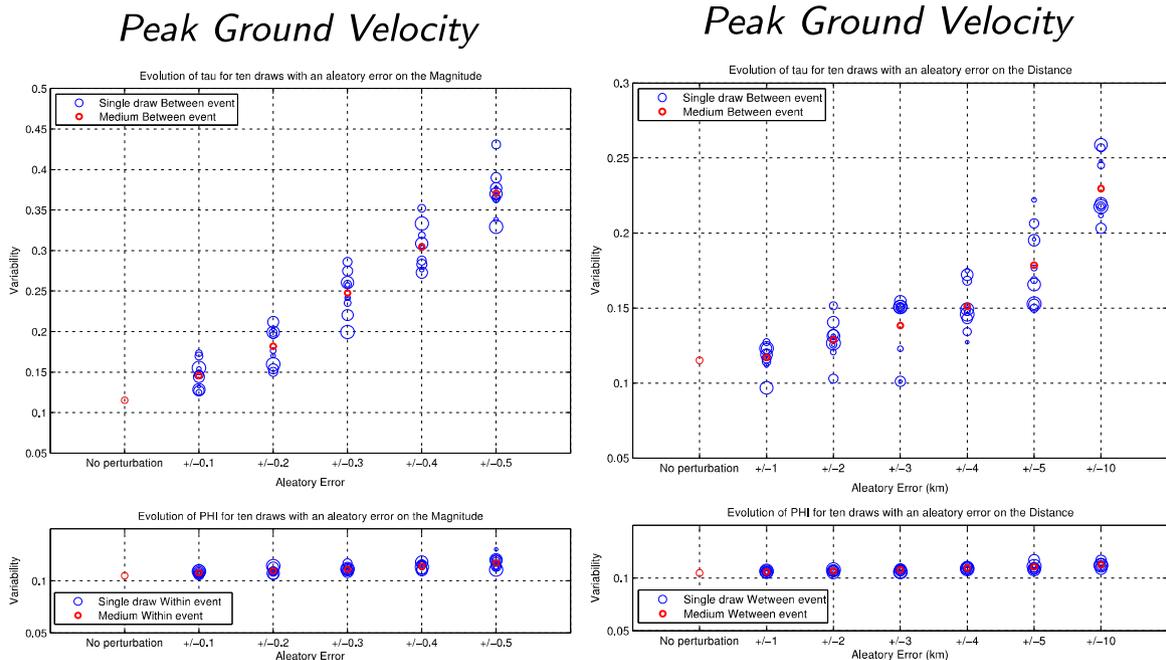
**Fig. 15.** Impact of data set size on the total (blue circle), within-event (red circle) and between-event (grey circle) variabilities. The S3 set and S4 subsets consist of 1260 events and 52 events, respectively, both recorded at 15 sites. The aleatory variabilities are displayed for the peak ground acceleration (left) and peak ground velocity (right).

stringent cases with already existing solutions. The canonical cases developed within E2VP, which are freely available to the seismological community (<http://www.sismowine.org>), can serve this purpose.

Most of the new work achieved during E2VP2 was related to validation. The comparison for the 19, relocated events has thus been found in average slightly improved for rock sites, and slightly deteriorated within the Mygdonian basin, with an overall trend for an overestimation. It was found once again that the sole site response (“hybrids”) is better estimated. The significant overestimation in terms of signal amplitude (parameters C1–C4) thus comes mainly for the overestimation of the rock motion, also associated with an underestimation of signal duration (C5): both may come from the absence of scattering in the considered crustal model. When considering all the real receivers, the overall E2VP2 misfit values range between +50 and +150%, to be compared with the +40–80% of E2VP1 (on only 6

events), while the “site-response only” misfits now range around +20%, while they were around –40% for E2VP1. The modifications in the basin model have slightly improved the site response estimate, but the sensitivity to the source parameters and the associated uncertainties leads to conclude that the feasibility of validation up to frequencies around a few Hz (4 Hz in the present case) is still a real challenge, for several reasons.

- The predicted ground motion proves to be very sensitive to the exact position of the source – especially its depth and distance – for very close events and for local, shallow events: as it is unrealistic to expect a precision on localization smaller than 2 km (especially for the depth), it is therefore recommended, for validation purposes, to select events with epicentral distance  $R$  larger than 20 km and hypocentral depths  $Z$  larger than 8 km. Closer events ( $R > 10$  km)



**Fig. 16.** Influence of the uncertainty on source parameters (magnitude, left; and location, right) on the aleatory variability components (between-event on top and within event on bottom). This influence is illustrated here for the PGV. The amount of uncertainties considered for the magnitude values and the source location are indicated on the abscissa. Each open blue symbol corresponds to one random generation of perturbed magnitude or location sets, while red circles correspond to the median values for the 10 random sets for one level of magnitude or location uncertainty. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 7**

Comparison of observed and “synthetic” variabilities for the S4 set. The observed values (second column) correspond to a specific GMPE derived on the basis of the S4 set by O.J. Ktenidou et al. (personal communication, see also Ktenidou et al. [41]). The synthetic values (columns 3–5) are derived with the neural network approach using the exact magnitude and location (column 3), randomly modified magnitudes ( $\Delta m = \pm 0.2$ ) and exact locations (column 4), and randomly modified locations ( $\Delta l = \pm 3$  km) and exact magnitudes (column 5).

	Observed, after relocation and with site information (VS30)	Synthetics, exact M and location	Synthetics, with exact location and magnitude error $\Delta m = \pm 0.2$	Synthetics, with exact magnitude and location error $\Delta l = \pm 3$ km
Between-event $\tau$	0.18	0.12	0.18	0.15
Within-event $\phi$	0.15	0.12	0.12	0.12
Total $\sigma$	0.24	0.16	0.22	0.19

can also be used, provided they are deeper ( $Z > 10$  km)

- The misfit between observations and numerical predictions remain significantly larger than the distance between carefully selected, up-to-date, and carefully implemented numerical simulation codes. For the prediction of ground motion for expected events with a priori defined source characteristics, the numerical-simulation approach is fully legitimate in the toolbox for site-specific ground-motion estimation.
- The predictions-to-observations differences are significantly lower when considering only the site amplification, especially when the reference is at depth within a vertical array. This emphasizes the added value of “hybrid” approaches made possible by the availability of down-hole recordings and the invaluable usefulness of in-situ recordings: it seems today very difficult to predict site effects in a complex geometry context with only geological, geophysical and geotechnical information. Site instrumentation is strongly recommended, including also due attention to reference sites (downhole and outcropping rock as much as possible) for a proper “calibration” of the reference motion.
- The 3D site response however exhibits a significant dependency on the source back-azimuth, which partly explains the event-to-event variability of instrumental site-to-reference spectral ratios.

It is also worth to discuss these results in the light of a few recent studies reporting comparisons between strong motion recordings and low to intermediate frequency ground motion simulation for various moderate to significant size earthquakes: Chino Hills, California (2008, Mw 5.4, 336 receivers within 100 km distance, 0.1–4 Hz; Tabora and Bielak [42]), Po Plain, Italy (2012, Mw 6.0, 34 receivers within 30 km distance, 0.1–1.5 Hz; Paolucci et al. [43]), and South Napa, California (2014, Mw 6.0, 10 receivers within 20 km distance, 0.1–5 Hz; Galovic [44]). The first two report Anderson-like [14] “goodness-of-fit” scores predominantly in the range 4–8 (i.e., basically “fair” and “good”), corresponding to simulation to recordings differences in the range 50–100%, while the last indicates PGV and response spectra ratios in the range [0.25–4]- thus with some gof scores below 4 (“poor” fit)-, most of them being in the range [0.5–2] (i.e., at least “fair”). Such average misfit values turn out to be slightly smaller than ours when considering full synthetics (see Table 5), and significantly larger when considering hybrid synthetics at the single receiver TST0, i.e., when focusing only on the 3D site response. One must keep in mind however that such a comparison must be done with caution. The values reported in these recent post-earthquake simulation studies correspond to one particular, rather large, event, while ours are an average for many smaller events. Another difference is that most often, in those studies, the target records have already been used to constrain the rupture history, while in our case the determination of source parameters is independent of the target recordings. The comparisons reported in these papers are performed after the event, and thus do not correspond to the case of a fully blind situation where the ground motion prediction is requested in advance of the event (which was actually the initial request from the French nuclear authority that launched all these benchmarking exercises).

In addition, the comprehensive sensitivity study also showed that, beyond the deterministic prediction of ground motion for a given earthquake scenario, carefully verified numerical simulation can provide a very instructive insight at the structure of the so-called “aleatory” variability of ground motion, for both its within- and between-event components. For the E2VP case, the between-event variability has been found very sensitive to hypocenter location errors (25% increase of  $\tau$  for a location uncertainty of  $\pm 3$  km), and to uncertainty in magnitude estimates (doubling of  $\tau$  for a  $\Delta m$  uncertainty of  $\pm 0.25$ ). Such a finding explains the increase of aleatory variability for small magnitude events in most recent GMPEs, and emphasizes the usefulness of dense seismological networks in order to reduce  $\tau$  and  $\sigma$ . The within event, single-site variability is shown to be associated to an “epistemic” dependence of the 3D site response on the event back-azimuth (predominantly), and on distance and depth (more slightly). This result calls for caution in the interpretation of single-station variabilities derived from a too small number of events: in the E2VP case, even 52 recordings at one site are not enough to fully capture the whole variability of site response when the azimuthal distribution is skewed.

## Acknowledgements

We thank A. Kiratzi, Z. Roumelioti from AUTH for their efficient help in relocating the events, and O.J. Ktenidou for her precious contribution about aleatory variability. We also thank an anonymous reviewer for his/her relevant comments and suggestions which helped improving the quality and clarity of the paper. The funding of this project was provided by CEA within the framework of the CASHIMA and SIGMA projects. The SEM calculations were performed using HPC resources from GENCI-TGCC under grants (2013–046060, 2014–046060) and using the Froggy platform from the CIMENT HPC center of Grenoble Alpes University (<https://ciment.ujf-grenoble.fr>, last accessed February 2016) which is supported by the Rhône-Alpes region (GRANT CPER07\_13 CIRA), the OSUG@2020 labex (reference ANR10 LABX56) and the Equip@Meso project (reference ANR-10-EQPX-29-01) of the Programme Investissements d’Avenir supervised by the Agence Nationale pour la Recherche. The SPECSEM3D code used to produce the 3-D synthetic Green’s functions is available at <https://geodynamics.org/cig/software/specsem3d> (last accessed September 2016).

## References

- [1] Cramer CH. Weak-motion observations and modeling for the Turkey Flat, U.S., site-effects test area near Parkfield, California. *Bull Seismol Soc Am* 1995;85(2):440–51.
- [2] Bard P-Y. Discussion session: lessons, issues, needs and prospects, special theme session 5: Turkey Flat and Ashigara Valley experiments. In: Proceedings of the 10th world conference on earthquake engineering conference proceedings. 1992. p. 6985–88.
- [3] Kawase H, Iwata T. A report on submitted results of the simultaneous simulation for Kobe. In: The effects of surface geology on seismic motion, recent progress and new horizon on ESG study. vol. 3. Balkema/CRC Press.
- [4] Day SM, Bielak J, Dreger D, Graves R, Larsen S, Olsen KB, Pitarka A. Tests of 3D elastodynamic codes: final report for lifelines project 1a01, tech. rep., Pacific

- Earthquake Engineering Research Center and Southern California Earthquake Center. 2001.
- [5] Day SM, Bielak J, Dreger D, Graves R, Larsen S, Olsen KB, Pitarka A. Tests of 3D elastodynamic codes: final report for lifelines project 1a02, tech. rep., Pacific Earthquake Engineering Research Center. 2003.
- [6] Day SM, Bielak J, Dreger D, Graves R, Larsen S, Olsen KB, Pitarka A. 3D ground motion simulations in basins: final report for lifelines project 1a03, tech. rep., Pacific Earthquake Engineering Research Center. 2005.
- [7] Bielak J, Graves RW, Olsen KB, Taborda R, Ramírez-Guzmán L, Day SM, et al. The shakeout earthquake scenario: verification of three simulation sets. *Geophys J Int* 2010;180:375–404. <http://dx.doi.org/10.1111/j.1365-246X.2009.04417.x>.
- [8] Chaljub E, Tsuno S, Bard P-Y. Numerical Benchmark of 3D ground motion simulation in the valley of Grenoble, French Alps. In: Bard P-Y, Chaljub E, Cornou C, Cotton F, Guéguen P (eds.), ESG2006 Proceedings. LCPCE Editions, ISSN 1628-4704. Vol. 2. 2009. p. 1365–75.
- [9] Tsuno S, Chaljub E, Bard P-Y. Grenoble simulation benchmark: comparison of results and main learnings. In: Bard P-Y, Chaljub E, Cornou C, Cotton F, Guéguen P (eds.), ESG2006 Proceedings. ISSN 1628-4704. Vol. 2. 2009. p. 1377–433.1436.
- [10] Chaljub E, Moczo P, Tsuno S, Bard P-Y, Kristek J, Kaser M, et al. Quantitative comparison of four numerical predictions of 3D ground motion in the Grenoble valley, France. *Bull Seism Soc Am* 2010;100(4):1427–55.
- [11] Chaljub E, Maufroy E, Moczo P, Kristek J, Hollender F, Bard P-Y, et al. 3-D numerical simulations of earthquake ground motion in sedimentary basins: testing accuracy through stringent models. *Geophys J Int* 2015;201(1):90–111. <http://dx.doi.org/10.1093/gji/ggu472>.
- [12] Maufroy E, Chaljub E, Hollender F, Kristek J, Moczo P, Klin P, et al. 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* 2015;105:1398–418. <http://dx.doi.org/10.1785/0120140228>.
- [13] Kristeková M, Kristek J, Moczo P. Time-frequency misfit and goodness-of-fit criteria for quantitative comparison of time signals. *Geophys J Int* 2009;178(2):813–25.
- [14] Anderson JG. Quantitative measure of the goodness-of-fit of synthetic seismograms. Proceedings of the 13th world conference on earthquake engineering, 2004 August 16. Canada: Vancouver, British Columbia; 2004, [Paper No. 243].
- [15] Trifunac MD, Brady AG. On the correlation of seismic intensity scales with the peaks of recorded strong ground motion. *Bull Seism Soc Am* 1975;65(1):139–62.
- [16] Kwok AO, Stewart JP, Hashash YM. Nonlinear ground-response analysis of Turkey Flat shallow stiff-soil site to strong ground motion. *Bull Seismol Soc Am* 2008;98:331–43.
- [17] Stewart J, Kwok A. Nonlinear seismic ground response analysis: protocols and verification against array. Data PEER Annu Meet San Franc Present 2009:84.
- [18] Régnier J, Bonilla LF, Bard PY, Kawase H, Bertrand E, Hollender F, Marot M, Sicilia D, Nozu A. PRENOLIN Project: a benchmark on numerical simulation of 1D non-linear site effect. 1 – verification phase based on canonical cases. 6ICEGE, Christchurch, New-Zealand; November 2015.
- [19] Régnier J, Bonilla LF, Bard PY, Kawase H, Bertrand E, Hollender F, Marot M, Sicilia D, Nozu A. PRENOLIN Project: a benchmark on numerical simulation of 1D non-linear site effect. 2 – Results of the validation phase. 6ICEGE, Christchurch, New-Zealand; November 2015.
- [20] Régnier J, Bonilla L-F, Bard P-Y, Bertrand E, Hollender F, Kawase H, Sicilia D, Arduino P, Amorosi A, Asimaki D, Boldini D, Chen L, Chiaradonna A, DeMartin F, Ebrille M, Elgamal A, Falcone G, Foerster E, Foti S, Garini E, Gazetas G, Gélis C, Ghofrani A, Giannakou A, Gingery JR, Glinsky N, Harmon J, Hashash Y, Iai S, Jeremić B, Kramer S, Kontoe S, Kristek J, Lanzo G, di Lernia A, Lopez-Caballero F, Marot M, McAllister G, Mercierat ED, Moczo P, Montoya-Noguera S, Musgrove M, Nieto-Ferro A, Pagliaroli A, Pisanò F, Richterova A, Sajana S, Santisi d'Avila MP, Shi J, Silvestri F, Taiebat M, Tropeano G, Verrucci L, Watanabe K. International benchmark on numerical simulations for 1D, non-linear site response (PRENOLIN): verification phase based on canonical cases. *Bull Seism Soc Am* 2016;106(5):2112–35. <http://dx.doi.org/10.1785/0120150284>.
- [21] Garofalo F, Foti S, Hollender F, Bard PY, Cornou C, Cox BR, et al. InterPACIFIC project: comparison of invasive and non-invasive methods for seismic site characterization. Part I: intra-comparison of surface wave methods. *Soil Dyn Earthq Eng* 2016. <http://dx.doi.org/10.1016/j.soildyn.2015.12.010>.
- [22] Garofalo F, Foti S, Hollender F, Bard P-Y, Cornou C, Cox BR, et al. Interpacific project: comparison of invasive and non-invasive methods for seismic site characterization. Part II: intercomparison between surface-wave and borehole methods. *Soil Dyn Earthq Eng* 2016. <http://dx.doi.org/10.1015/j.soildyn.2015.12.009>.
- [23] Novotný O, Zahradník J, Tselentis G-A. Northwestern Turkey earthquakes and the crustal structure inferred from surface waves observed in western Greece. *Bull Seism Soc Am* 2001;91(4):875–9.
- [24] Maufroy E, Chaljub E, Hollender F, Bard P-Y, De Martin F, Roumelioti Z, Theodoulidis N, Guyonnet-Benaize C, Ktenidou O, Perron V. 3D ground motion simulations for site effects assessment: Learnings from Euroseistest Verification and Validation Project. Cashima/Sigma internal report number SIGMA-2014-D3-137. 2014. 79 pages (main texte) +131 pages (appendixes).
- [25] Manakou M. Contribution to the determination of a 3D soil model for site response analysis. The case of the Mygdonian basin [Ph.D. thesis, in Greek with English abstract]. Thessaloniki (Greece): Department of Civil Engineering, Aristotle University of Thessaloniki; 2007.
- [26] Manakou M, Raptakis D, Chávez-García FJ, Apostolidis PI, Pitilakis K. 3D soil structure of the Mygdonian basin for site response analysis. *Soil Dyn Earthq Eng* 2010;30:1198–211.
- [27] Caumon G, Collon-Drouaillet P, Le Carlier de Veslud C, Viseur S, Sausse J. Surface-based 3D modelling of geological structures. *Math Geosci* 2009;41:927–45.
- [28] Mallet JL. {G}eomodeling. Applied Geostatistics. New York, NY: Oxford University Press; 2002.
- [29] I.G.M.E. Geological Maps of Greece, Scale 1:50.000. Sheets: Thessaloniki, Epanomi, Vasilika, Thermi, Polygiros, Zagliveri.
- [30] Mountrakis D, et al. Special edition of the neotectonic map of Greece, Earthquake Planning and Protection Organization (OASP). Scale 1:100.000. Sheets: Langadas, Thessaloniki; 1995–1996.
- [31] Sotiriadis L, Psilovikova A, Vavliakis E, Syrides G. Some tertiary and quaternary basins of Macedonia/Greece. Formation and evolution. *Clausthal Geol Abh* 1983:21.
- [32] Raptakis D, Manakou M, Chávez-García FJ, Makra K, Pitilakis K. 3D configuration of Mygdonian basin and preliminary estimate of its site response. *Soil Dyn Earthq Eng* 2005;25:871–87.
- [33] Maufroy E, Chaljub E, Theodoulidis N, Roumelioti Z, Hollender F, Bard P-Y. Source-related variability of site response in the Mygdonian basin, Greece, from accelerometric recordings and 3D numerical simulations. *Submit Bull Seism Soc Am*, Minor Revis 2016.
- [34] Causse M, Chaljub E, Cotton F, Cornou C, Bard P-Y. New approach for coupling k-2 and empirical Green's functions: application to the blind prediction of broadband ground motion in the Grenoble basin. *Geophys J Int* 2009. <http://dx.doi.org/10.1111/j.1365-246X.2009.04354.x>.
- [35] De Martin F. Verification of a spectral-element method code for the Southern California Earthquake Center LOH.3 viscoelastic case. *Bull Seismol Soc Am* 2011;101(6):2855–65.
- [36] Moczo P, Kristek J, Galis M. The finite-difference modelling of earthquake motions: waves and ruptures. Cambridge University Press; 2014.
- [37] Papazachos CB. Crustal P- and S-velocity structure of the Serbomacedonian Massif (Northern Greece) obtained by non-linear inversion of traveltimes. *Geophys J Int* 1998;134:25–39. <http://dx.doi.org/10.1046/j.1365-246x.1998.00558.x>.
- [38] Derras B, Bard PY, Cotton F, Bekkouche A. Adapting the neural network approach to PGA prediction: an example based on the KiK-net data. *Bull Seism Soc Am* 2012;102(4):1446–61. <http://dx.doi.org/10.1785/0120110088>.
- [39] Derras B, Bard P-Y, Cotton F. Towards fully data driven ground-motion prediction models for Europe, in press and on line. *Bull Earthq Eng* 2014;12-1:495–516. <http://dx.doi.org/10.1007/s10518-013-9481-0>.
- [40] Maufroy E, Chaljub E, Hollender F, Bard P-Y, Kristek J, Moczo P, De Martin F, Theodoulidis N, Manakou M, Guyonnet-Benaize C, Pitilakis K, Hollard N. Validating the numerical simulation approach for ground motion prediction: General framework and latest lessons from the E2VP project. In: vited theme lecture, 6ICEGE (Proceedings of the 6th international conference on earthquake geotechnical engineering). Christchurch, New-Zealand; November 1–4, 2015. paper #452, 19 pages.
- [41] Ktenidou OJ, Roumelioti Z, Abrahamson NA, Cotton F, Pitilakis K. Ground motion uncertainty and variability (single-station sigma): insights from Euroseistest, Greece, 2014, Abstract S42B-03 presented at 2014 Fall Meeting, AGU (American Geophysical Union), San Francisco, Calif., 15–19 Dec. 2014
- [42] Taborda R, Bielak J. Ground-motion simulation and validation of the 2008 Chino Hills, California, Earthquake. *Bull Seismol Soc Am* 2013;103(1):131–56.
- [43] Paolucci R, Mazzieri I, Smerzini C. Anatomy of strong ground motion: near-source records and three-dimensional physics-based numerical simulations of the Mw 6.0 2012 May 29 Po Plain earthquake, Italy. *Geophys J Int* 2015;203:2001–20.
- [44] Galovic F. Modeling velocity recordings of the Mw 6.0 South Napa, California, Earthquake: unilateral event with weak high-frequency directivity. *Seismol Res Lett* 2016;87(1). <http://dx.doi.org/10.1785/0220150042>.