Blind broad-band (0-10 Hz) numerical prediction of the 3-D near field seismic response of a Mw6.0 extended fault scenario: application to the nuclear site of Cadarache (France)

David Alejandro Castro Cruz, Filippo Gatti, Fernando Lopez-caballero, Fabrice Hollender, Elias El Haber, Mathieu Causse

Geophysical Journal International

Geophys. J. Int. (2023) **232,** 581–600
Advance Access publication 2022 September 5
GJI Seismology

https://doi.org/10.1093/gji/ggac346

Blind broad-band (0–10 Hz) numerical prediction of the 3-D near field seismic response of an M_W6.0 extended fault scenario: application to the nuclear site of Cadarache (France)

D. Castro-Cruz [1], F. Gatti [1], F. Lopez-Caballero,[1] F. Hollender,[2] E. El Haber[3] and M. Causse[3]

[1] *Université Paris-Saclay, CentraleSupélec, ENS Paris-Saclay, CNRS LMPS—Laboratoire de Mécanique Paris-Saclay,* 91190 *Gif-sur-Yvette, France. E-mail: filippo.gatti@centralesupelec.fr*
[2] *CEA, DES, DIMP, DCET, Cadarache, F-13108 Saint-Paul-Lez-Durance, France*
[3] *ISTerre, Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, Univ. Gustave Eiffel,* 38000 *Grenoble, France*

Accepted 2022 August 28. Received 2022 August 22; in original form 2022 January 13

SUMMARY

In this paper, physics-based numerical simulation (PBS) is employed to render a broad-band (0–10 Hz) realization of the near-field seismic response of the experimental nuclear site of Cadarache, located nearby the active Middle Durance Fault (southeastern France). The sensitivity of the earthquake numerical model to geological features is investigated by comparison with geophysical measurements and past aftershock and by highlighting the amplification induced by the soft sediments below Cadarache. The blind prediction of an M_W6 target earthquake is approached by synthesizing four different finite-fault scenarios. The outcome is compared to the standard ground motion prediction equations (GMPEs), unveiling a possible GMPE overestimation of the pseudospectral acceleration ordinates at short natural periods, supporting the actual need to integrate synthetic and empirical predictions when direct observations are not available.

Key words: Computational seismology; Earthquake ground motions; Earthquake interaction, forecasting, and prediction; Site effects; Wave propagation.

1 INTRODUCTION

In France, the seismic response of nuclear power plants (NPPs) is a topic of huge relevance, due to their strategic importance as prominent energy suppliers and due to the nefarious and long-lasting consequences associated with possible nuclear accidents. In the last decade, after the Fukushima accident (2011 Tohoku earthquake), the International Atomic Energy Agency (IAEA) steered the path towards the definition of new safety standards for the seismic risk associated with existing and modern nuclear plants (IAEA 2020), publishing and updating several safety guides and requirements that nuclear energy companies endeavoured to comply to, following stringent seismic re-assessment and checking if existing nuclear facilities can sustain the seismic loads provided by updated hazard estimations. The complementary post-2011 safety assessment induced major stakeholders into defining *hard-core* seismic levels beyond extreme situations for NPPs, more stringent than those originally considered for design or re-assessment of the facility seismic safety. Those levels were initially established on a deterministic basis and then justified through state-of-the-art probabilistic seismic hazard assessments in the last decade (Berge-Thierry *et al.* 2018). The *as-is* condition of the installation was considered the baseline condition, including the *as-built*, *as-operated* and *as-maintained* conditions and the ageing at the time of the assessment.

Seismic risk reassessment studies have been performed by either a deterministic seismic hazard analysis (DSHA) or a probabilistic seismic hazard analysis (PSHA) approach (IAEA 2003). Most of the existing NPPs have been designed according to DSHA, that is based on the worst-case scenario, in terms of maximum credible earthquake and operating basis earthquake. For each of them, reference intensity measures (IMs) on a single earthquake source are defined (known seismic sources sufficiently near the site) and available historical seismic and geological data are employed to estimate ground motion at the plant site (Andrews & Folger 2012). On the other hand, PSHA provides the exceedance probability of ground motion intensity within a given time frame (Bommer 2003; Andrews & Folger 2012; Khan *et al.* 2020). In the context of the nuclear industry, DSHA approach is preferred, based on the so-called basic safety rule (RFS 2001, 2001-01) and on the ECS ASN (2011) guide (ASN/2/01) that provides design rules for nuclear civil engineering structures.

Fault-to-structure physics-based simulation (PBS) thrives in this context nowadays, given the holistic modelling approach made available by the ever increasing availability of computational

resources (McCallen *et al.* 2020a,b). PBS is progressively catching on in both DSHA and PSHA studies (Milner *et al.* 2021; Stupazzini *et al.* 2021), providing the possibility of rendering hundreds of thousands of years of virtual earthquake scenarios in a non-ergodic framework. The relatively cheap access to huge computational power allows the construction of end-to-end digital twins of an NPP, of a neighbourhood, of an entire region, including large active faults, complex geological stratification and structural components (Fu *et al.* 2017; Milner *et al.* 2021). Realistic seismic scenarios can be constructed at different scales, coupling different methods and models, integrating the data into the prediction (Ichimura *et al.* 2018). The possibility of including large faults and structural components into the very same 3-D numerical model at the regional scale (10–100 km) is attractive (Touhami *et al.* 2020) yet hard to achieve at present with satisfactory and realistic results, in an engineering frequency range (0–30 Hz).

The apparently unlimited power of those digital laboratories is counterbalanced by the fact that they require detailed information over large regions. The overall epistemic uncertainty increases, while models are cumbersome and dragged into the curse of dimensionality. The major sources of uncertainty are hidden in the rupture mechanism and in the tectonic stress distribution. The crustal geological structures and their properties, are source of uncertainty too, particularly those pertaining the soft sediments and basin-like structures. Finally, the soil-structure interaction and the behaviour of structural components add some further uncertainty to the whole analysis. Scientists are faced to the lack of information required to constrain the model prediction within realistic trust regions, where safely vary the input parameters at stake. Therefore, PBS requires an extensive preliminary verification exercise at each site and for a multifold case-history (Bradley 2019).

In this study, an M_W6 earthquake hazard scenario is drawn via PBS, for the Cadarache nuclear research centre (with approximately 20 nuclear facilities) in southern France within the framework of the SINAPS@ project(ANR-11-RSNR-0022-04)[1]. The area of interest was struck by the 1909 Lambesc earthquake ($M_W6.0$ according to Baroux *et al.* (2003) and $M_W5.7$ according to the FCAT17 French seismic catalog Manchuel *et al.* (2018)), presumably the strongest known recorded earthquake in France. Previous studies (Dujardin *et al.* 2018, 2020) defined possible $M_W6.0$ rupture scenarios employing kinematic fault modelling and empirical Green's functions (EGFs), predicting realistic ground motion time histories for a hard rock site ($V_{S,30} = 1800$ m s^{-1}) and a sediment site ($V_{S,30} = 440$ m s^{-1}) nearby the Cadarache site. Synthetic ground motion parameters were found to be compatible with several ground motion prediction equations (GMPEs). The mentioned studies were extended by Castro-Cruz *et al.* (2021), who constructed a 3-D physics-based numerical model for a domain of approximately 50 km × 60 km and 60 km in depth and validated it against $M_L2.9$ earthquake occurred in 2010 in the surroundings (Guyonnet-Benaize *et al.* 2015). This weak seismic event allowed the authors to validate the preliminary geological and topographic model (Hollender *et al.* 2009; Guyonnet-Benaize *et al.* 2015) and to perform thousands of blind hybrid finite-fault ground motion predictions for the $M_W6.0$ target scenario defined by the hazard analysis. The hybrid seismic response analysis was obtained by blending low-frequency (0–5 Hz) PBS and broad-band (1–15 Hz) EGF method, providing reliable site response at the site (complying with the reference GMPE Berge-Thierry *et al.* (2003), widely used for DSHA of nuclear installations in France,

[1]https://www.institut-seism.fr/en/projects/sinaps

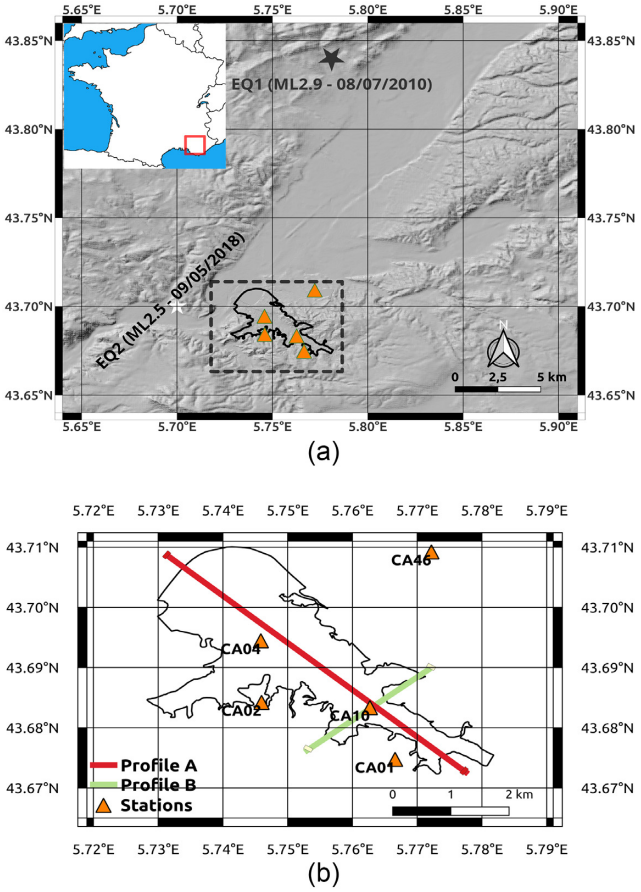

Figure 1. (a) Relief map of the surrounding areas around the Cadarache site. Orange triangles represent the available recording station. The elevation contour is expressed in meters above the mean sea level (MSL). The two small earthquakes recorded in the surroundings and indicated in Table 1 are reported [black (EQ1) and white (EQ2) stars, respectively]. The contour of the alluvial basin is indicated by the black line. (b) Detail of the sedimentary basin in the surrounding of the Cadarache area, including the recording stations available, employed by Perron *et al.* (2018) and the two cross-sections AA' and BB'.

and with the recent GMPE proposed by Kotha *et al.* (2020), based on an updated version of the European strong-motion data set).

The main objective of this paper is to unveil—via PBS—the complexity of plausible near-field strong ground motion events, triggered by one of the segments belonging to the nearby Middle Durance active fault system. Moreover, driven by previous geophysical campaigns (see Section 2), the present study aims at estimating the 3-D *basin-like* site-effects induced by the soft sediments underneath Cadarache. Four different extended-fault scenarios are compared, since they could not be described by the EGF method employed by Dujardin *et al.* (2020). The source modelling and its impact of the seismic site response are described in Section 3. In Section 3.2, the basin-like effects are analysed in detail, in order to check their compatibility with standard GMPEs.

2 DATA AND METHOD

A 3-D model, including the geometry of the main interfaces and the geophysical properties necessary for the numerical simulations (V_P, V_S, Q_P, Q_S, ρ) was built for the purpose of this study. The seismic regional response of the Cadarache site is simulated for

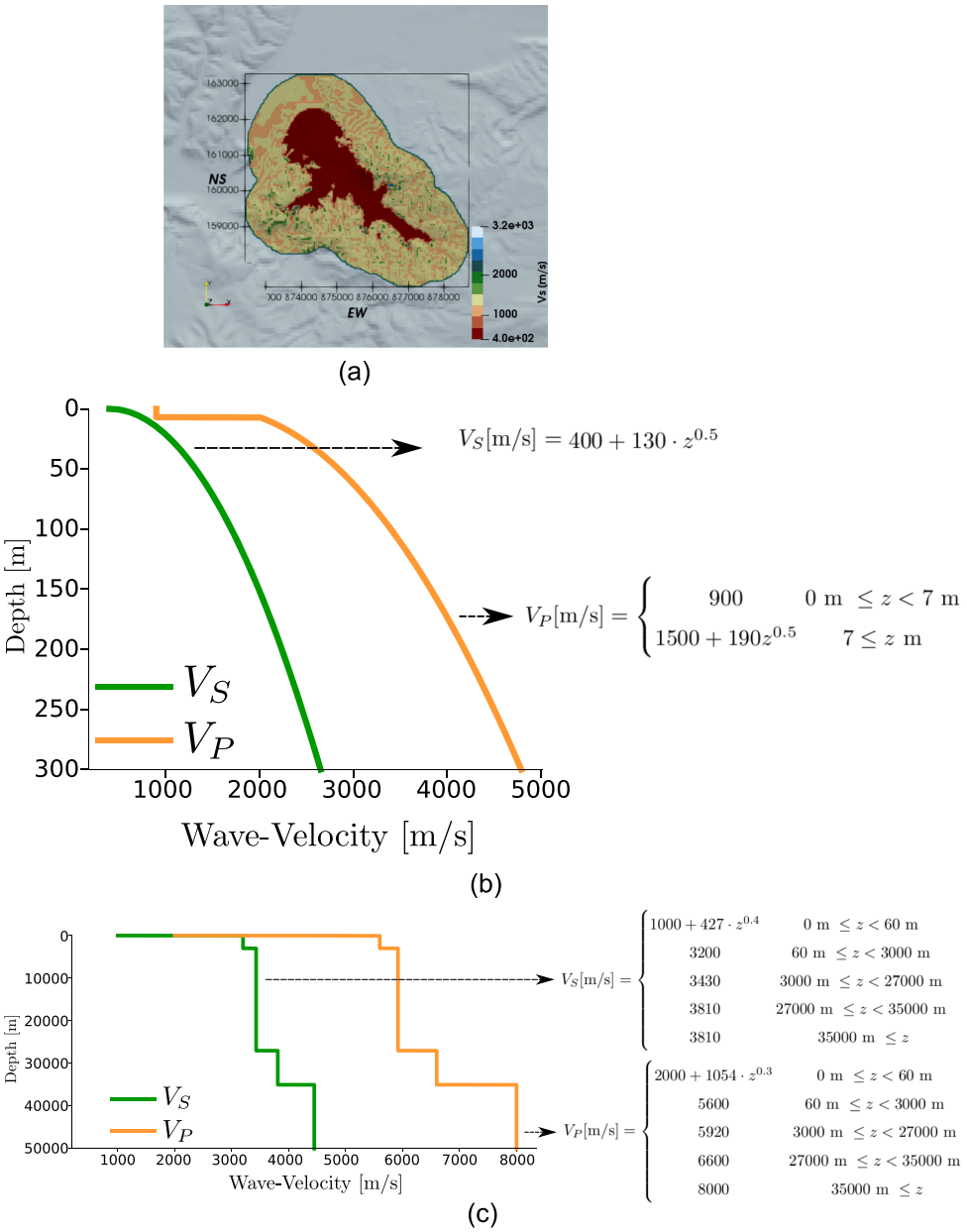

Figure 2. (a) Shear wave velocity V_S contour map at free surface, provided by CEA experimental campaign (see Fig. 1). (b) Wave velocity 1-D profile of the alluvial basin; (c) wave velocity 1-D profile of the outcropping bedrock.

the area depicted in Fig. 1, which shows the relief map of the area of interest, as well as a detailed representation of the boundaries of the sedimentary basin (Meso-Cenozoic sedimentary layers on Palaeozoic basement Guyonnet-Benaize *et al.* 2015).

For further details, the reader can refer to Castro-Cruz *et al.* (2021).

2.1 Geological model

The Cadarache geological model is based on several previous works. For the definition of the geometry of the basin, we adopted the local model built at the scale of the Cadarache Centre, the first version of which was carried out by Guyonnet-Benaize (2011) and which has since been periodically updated by the Commissariat á l'énergie atomique et aux énergies alternatives (CEA) by integrating the new data available. This local model was mainly built on the basis of boreholes (more than a thousand) and numerous horizontal-to-vertical ratio (H/V) measurements (Hollender *et al.* 2011) allowing the depth of the bedrock to be assessed in areas not covered by boreholes, following an approach also adopted by Cushing *et al.* (2020). Further details can also be found in Perron *et al.* (2018).

Concerning the properties of the geological formations, we proposed a simplified velocity model, both in the basin and in the bedrock, reported in Figs 2(b) and (c) (for V_P and V_S profiles).

For the velocity model in the basin, we used information from invasive [P and S suspension logger (PSSL), cross-hole and down-hole] and non-destructive (amplitude variation with angle, AVA) measurements, acquired at the margin of the InterPACIFIC

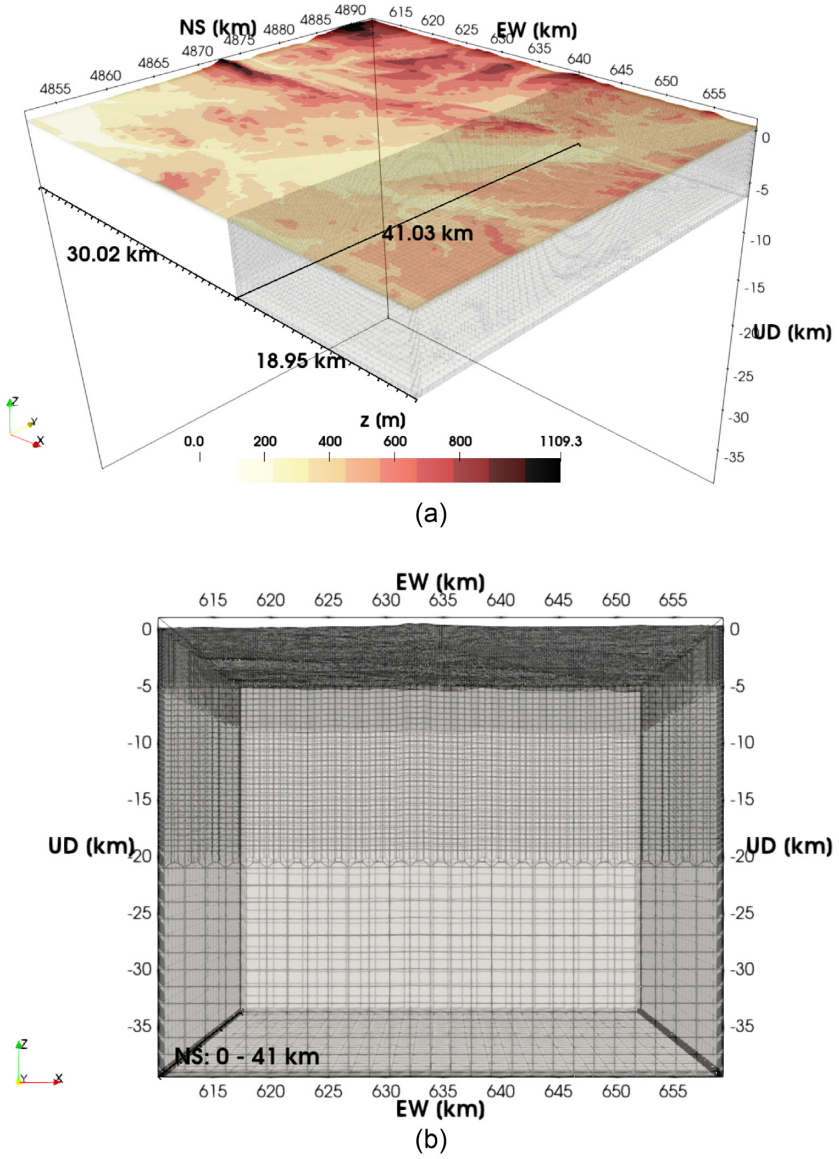

Figure 3. (a) Sketch of the numerical model of the Cadarache region with elevation. (b) Mesh lateral view.

Table 1. Details of the weak seismic events adopted in this paper, as defined by Dujardin *et al.* (2020). M_L: local earthquake magnitude; M_0: seismic moment magnitude; ϕ_S, δ, λ : strike, dip, rake angles.

Tag	Date	M_L (1)	M_0 (Nm)	Lat (°)	Lon (°)	Depth (km)	ϕ_S, δ, λ (°)
EQ1	08/07/2010	2.9	$1.8\ 10^{13}$	43.840	5.781	3.49	186, 19, -90
EQ2	09/05/2018	2.5	$8.0\ 10^{12}$	43.700	5.700	2.00	85, 75, 147.3

Table 2. Details of the four extended-fault seismic events adopted in this paper (target seismic moment 1.04×10^{18} Nm). EF1, EF2, EF3 and EF4 were generated according to Dujardin *et al.* (2020). L and W represent the along-strike fault length and along-dip width. Lon, lat and depth indicate the position of the nucleation point. ϕ_S, δ, λ: strike, dip, rake angles. $\Delta\sigma$ represent the average stress drop and V_R the average rupture velocity.

	L (km)	W (km)	Lon (°)	Lat (°)	Depth (km)	ϕ_S, δ, λ (°)	$\Delta\sigma$ (MPa)	V_R (m s^{-1})
EF1	13.860	7.490	5.78	43.91	4.91	34.6, 25.4, 0	1.15	2713
EF2	13.545	7.315	5.66	43.80	4.59	22.1, 18.0, 0	1.10	2611
EF3	13.965	7.525	5.65	43.81	4.88	35.8, 17.5, 0	0.91	2519
EF4	15.820	8.540	5.69	43.82	4.42	37.0, 23.3, 0	0.79	2727

Table 3. List of Anderson's criteria to assess the goodness-of-fit of two time histories.

Number	Symbol	Similarity of:	Definition
C1	SDa	Arias duration	$10 \cdot \left(1 - \max\left(\lvert \frac{I_{A,1}(t)}{I_{A,1}(T_{d,1})} - \frac{I_{A,2}(t)}{I_{A,2}(T_{d,2})} \rvert\right)\right)$
C2	SDe	Energy duration	$10 \cdot \left(1 - \max\left(\lvert \frac{I_{E,1}(t)}{I_{E,1}(T_{d,1})} - \frac{I_{E,2}(t)}{I_{E,2}(T_{d,2})} \rvert\right)\right)$
C3	SIa	Arias intensity	$S(I_{A,1}(T_{d,1}), I_{A,2}(T_{d,2}))$
C4	SIv	Energy integral	$S(I_{E,1}(T_{d,1}), I_{E,2}(T_{d,2}))$
C5	Spga	Peak acceleration	$S\left(\max_{t\in[0,T_d]}\lvert a_1(t)\rvert, \max_{t\in[0,T_d]}\lvert a_2(t)\rvert\right)$
C6	Spgv	Peak velocity	$S\left(\max_{t\in[0,T_d]}\lvert v_1(t)\rvert, \max_{t\in[0,T_d]}\lvert v_2(t)\rvert\right)$
C7	Spgd	Peak displacement	$S\left(\max_{t\in[0,T_d]}\lvert d_1(t)\rvert, \max_{t\in[0,T_d]}\lvert d_2(t)\rvert\right)$
C8	Ssa	Response spectra	$\mathbb{E}_{T_n}\left[S\left(Sa_1\left(T_n\right), Sa_2\left(T_n\right)\right)\right]$
C9	Sfs	Fourier spectra	$\mathbb{E}_f\left[S\left(FS_1\left(f\right), FS_2\left(f\right)\right)\right]$
C10	C∗	Cross correlation	$10 \cdot \max\left[C(a_1(t), a_2(t)), 0\right]$

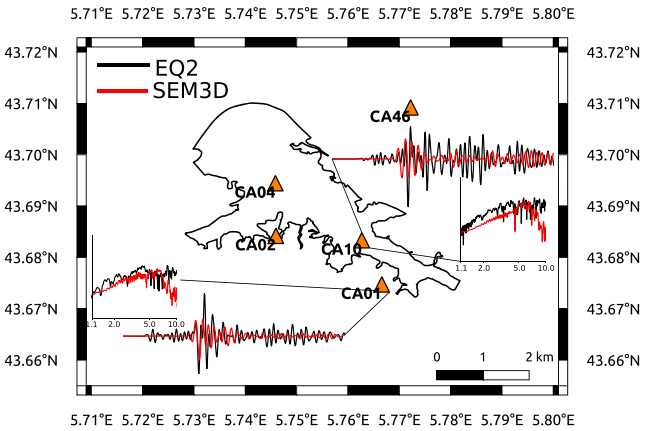

Figure 4. Map of the Cadarache sedimentary basin with comparison between recorded (black) and synthetic (red) time-histories and Fourier's amplitude at CA01 and CA10 for EQ2.

project (Garofalo *et al.* 2016a,b) at the location of the CA10 station (see Fig. 2). For the velocity model in the bedrock, we used the standard 1-D model used at the regional scale in seismicity location estimations. This model was completed at shallow bedrock depth (0–60 m) around the basin using data acquired in the Inter-PACIFIC project at station CA01. This addition allows the effects of the weathered part of the bedrock on the seismic motions to be taken into account for the stations located at the outcropping rock (CA01, CA02).

According to Castro-Cruz *et al.* (2021), the 3-D velocity model interpolating the outcrop bedrock and basin profiles includes a *transition* region in between the two geological models, to smooth out possible spurious reflections stemming from abrupt impedance contrast. In the *transition* region, the shear wave velocity at surface is lower than the outcropping bedrock model ($V_S \approx 1000$ m s^{-1} at surface, see Fig 2a). Concerning the viscous damping, Q_S was estimated as $V_S/10$ and Q_P as $V_P/5$.

Like any model, the present geological configuration represents a simplification of the reality. Nevertheless, we would like to draw the reader's attention to some important simplifications for a proper interpretation of the results:

(i) In reality, the Cadarache basin depicted in Figs 2(a) is open and it extends towards northwestern-ward. Since we have no information in this area, we have artificially *closed* the basin in the northwestern corner. The simulation results obtained in this area will not be therefore representative of the real ground motion.

(ii) The velocity profiles adopted for the basin are 1-D. They refer to station CA10 and they have been extrapolated to the whole basin. The available data is not sufficient to propose a site-wise model to take into account lateral variations. Similarly, the 1-D profiles themselves have been simplified by smoothing out vertical heterogeneities at the meter scale as they have no impact on the reference frequency band of interest for the present numerical simulations (see Section 2.3).

(iii) At the scale of the numerical model presented hereafter, only the Cadarache basin has been introduced. In this area, there are obviously other basins or low-velocity zones. Similarly, the outcropping bedrock is weathered over the whole area, and not only in a 1 km band around the Cadarache centre. These two simplifications were introduced to optimize the calculation mesh in order to reduce the overall computational burden.

2.2 Numerical model

Earthquake simulations were performed via the earthquake engine called SEM3D (CEA and CentraleSupélec and IPGP and CNRS 2017; Touhami *et al.* 2022), a high-performance code for elasto-acoustic wave propagation, based on the spectral element method (Faccioli *et al.* 1997; Komatitsch & Vilotte 1998). The numerical model is described in details by Castro-Cruz *et al.* (2021). In short, the region of interest (depicted in Fig. 1) is meshed with ≈ 13.5 million linear hexahedral elements for a total of 1.44×10^9 degrees of freedom (DOFs) with perfectly matched layers (PMLs) absorbing boundary conditions (Festa & Vilotte 2005). The mesh was designed to solve with high accuracy the wave-propagation problem up to a maximum frequency of 10.5 Hz, with a minimum wavelength of 38 m. In other words, hexahedra in the area around the basin have a maximal edge size of 105.8 m; within the basin, where soft deposits are located, the minimum element size is 49.1 m. A glimpse of the mesh detail around the sedimentary basin underneath the Cadarache site is provided in Fig 3(a). The basin/bedrock interface was not meshed directly, but the geological properties were assigned to a predefined structured mesh of hexahedra (as clearly depicted in Fig 3a).

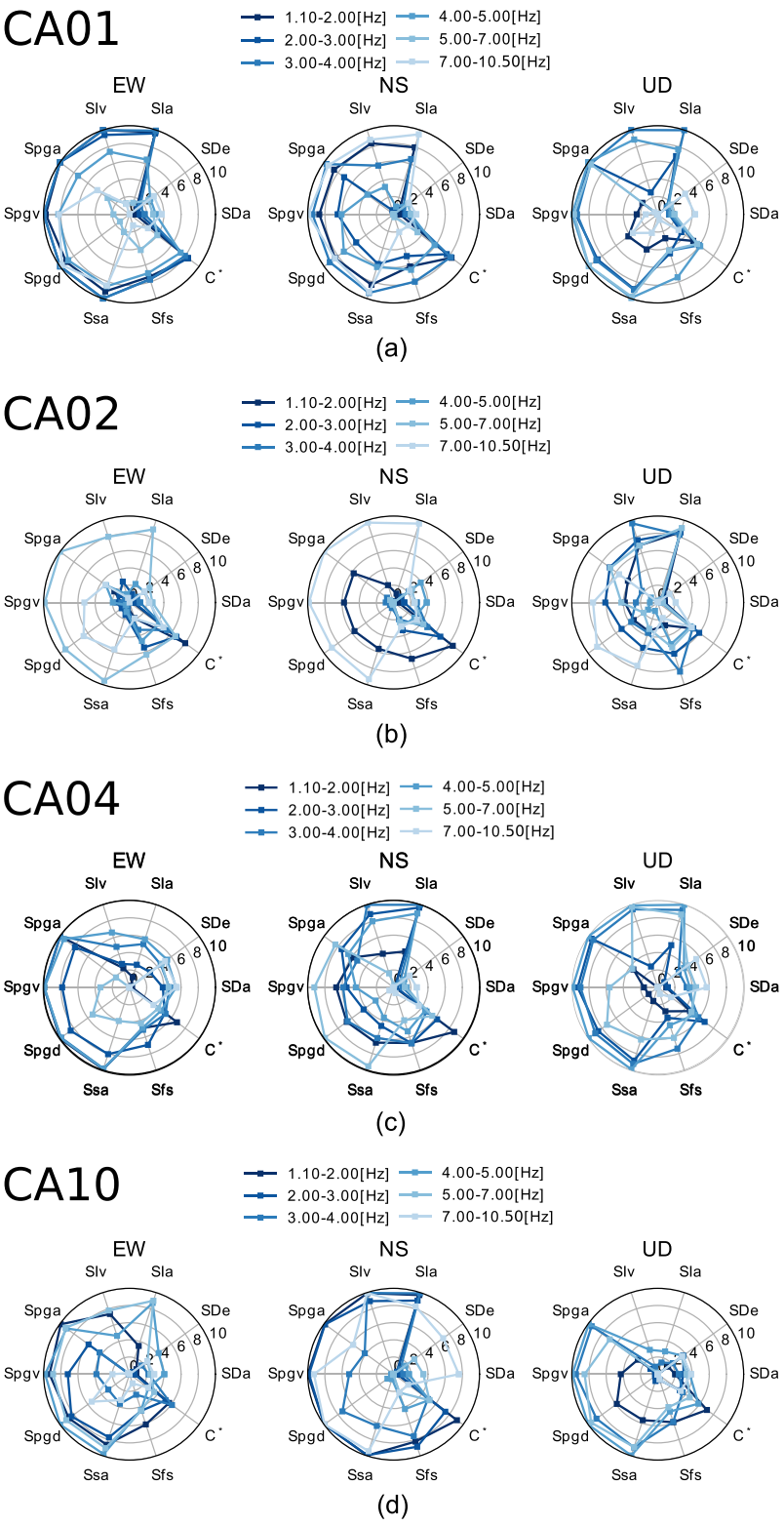

Figure 5. Modified Anderson's criteria (mAC) considered as goodness-of-fit 0–10 score (Anderson 2004). mAC were computed on the recorded and simulated time-histories at CA01, CA02, CA04 and CA10. Time-histories are band-pass filtered in the frequency range 1.1–10.5 Hz (fourth-order Butterworth filter). mAC were computed separately for different frequency ranges.

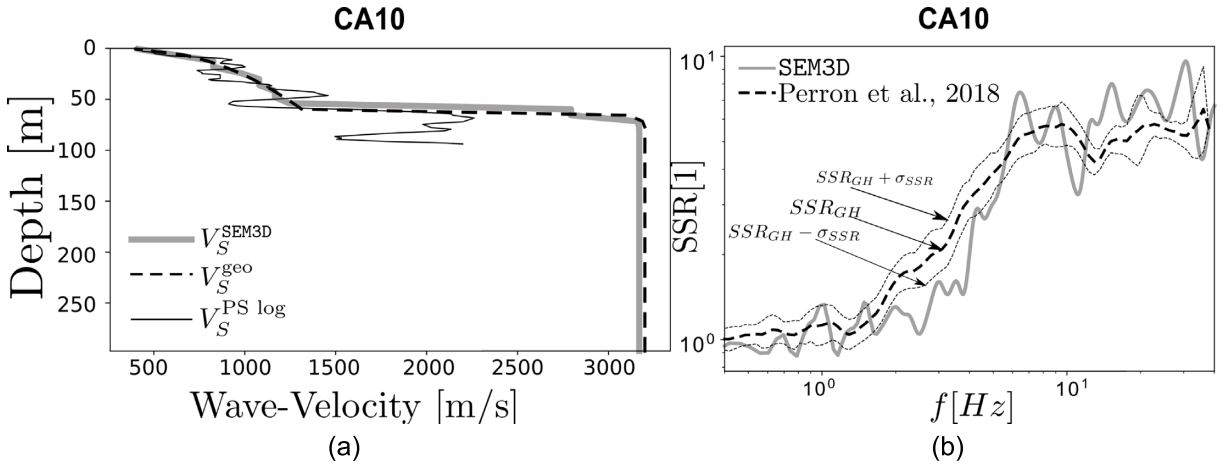

Figure 6. (a) Comparison between 1-D profiles at CA10 station provided by PS suspension logging (according to Perron *et al.* (2018), indicated by $V_S^{PS\ log}$), by the interpolated geological model (V_S^{geo}, provided by CEA) and the corresponding interpolation with SEM3D (V_S^{SEM3D}). (b) SSR computed at CA10, with respect to the referent site CA01. Solid black line represents the geometric average SSR (SSR_{GH}) empirically obtained by Perron *et al.* (2018) from *in situ* measurements. Dashed black lines represent the corresponding SSR_{GH} plus and minus standard deviation σ_{GH} computed from more than a hundred earthquakes. Red dashed line represent the synthetic SSR obtained via SEM3D simulation of EQ2.

This *not-honouring* approach simplifies the task of constructing the numerical model, but depending on the complexity of the interface, the interpolated geological properties may slightly differ from the original datum (see Section 2.3).

2.3 Validation of the geological model

In order to validate the synthetic geological model constructed with the purpose of accurately reproducing wave propagation in the Earth's crust and site effects, we compare point-source simulations with two small earthquakes recordings available for the region of interest (Fig. 1a). Details are reported in Table 1.

Castro-Cruz *et al.* (2021) validated the present numerical model for EQ1 in the 1.1–5 Hz frequency band. However, EQ1 epicentre is quite far away from the sedimentary basin. To extend the numerical simulations' validation in this work, the fidelity of the 3-D synthetic geological structure adopted to run each PBS is hereafter assessed in the 1.1–10.5 Hz frequency range, for EQ2 scenario (closer and more recent than EQ1). The available recordings have insufficient signal-to-noise ratio (SNR<3) for frequencies below 1.1 Hz (Dujardin *et al.* 2020). Therefore, the present validation focuses on the 1.1–10.5 Hz (upper accuracy limit of the numerical model). The small ground shaking was replicated with a point-wise source, characterized by a Gabor's function. Fig. 4 shows a comparison between recorded and synthetic acceleration time-histories (and Fourier's spectra) for EQ2, at CA01 and CA10, that is outside and inside the sedimentary basin. Fig. 5 shows the Anderson's criteria (Anderson 2004), modified according to Olsen & Mayhew (2010; mAC, see Appendix A.1), a comparative metrics to assess the goodness-of-fit (GoF) of synthetic and recorded time-histories on a scale from 0 to 10, based on ten different indicators listed in Table 3. See Appendix A.1 for further details on the mAC rating system. Fig. 5 depicts a polar representation of mAC, computed for different frequency bands between 1.1 and 10.5 Hz (each band being associated with a different colour) at CA01, CA02 (on outcropping bedrock) and CA04, CA10 (within the basin). Overall, the GoF is fair (i.e. GoF $>$ 4) at all stations and for a large band from 1.1 to 5 Hz. Arias duration $C_1 = SDa$, energy duration $C_2 = SDe$ and cross-correlation $C_{10} = C*$ overall score very poorly, due to unknown damping and to the difficulty in properly parametrizing the Gabor's source time function with unknown rise time (the aftershock being too small). The GoF at CA02 is poor, due to the lack of information on the transition zone between outcropping bedrock and soft sediments. This aspect should be further investigated.

Most of the mAC ranked above 5 in the 1.1–5 Hz frequency band, as already proven by Castro-Cruz *et al.* (2021). Interestingly enough, CA02 show lower scores compared to the validation performed by Castro-Cruz *et al.* (2021; based on EQ1), possibly proving the need to better constrain the transition zone around the soft basin, to effectively reproduce the near-fault ground motion close to the basin edges.

The validation exercise is undoubtedly a hard task to accomplish (Maufroy *et al.* 2015, 2016). As a matter of fact, Fig. 6(a) clearly shows that at station CA10 (within the basin, see Fig. 1b) the shear wave velocity profile obtained by PS Suspension Logging $V_S^{PS\ log}$, presented by Perron *et al.* (2018), largely differs from V_S^{geo}—the analytical velocity profile—and from the interpolated profile V_S^{SEM3D}, employed for the PBS run. This discrepancies have been highlighted by Castro-Cruz *et al.* (2021). They represent a well-known issue of numerical simulation, since they lead to spurious differences in the overall seismic response. Some authors (Wang *et al.* 2021) recently managed to introduce site-specific borehole information in 3-D average geological model. A further validation of the synthetic geological model is required in order to strengthen the validation process and check the fidelity of the numerical model to reproduce the site-effects. To this end, the *synthetic standard spectral ratio* (SSR), obtained from the SEM3D outcome, was compared to its empirical counterpart computed by Perron *et al.* (2018) via ambient noise geophysical campaign. The latter obtained the SSR between a generic site s_x and the reference site s_r following the expression:

$$\mathrm{SSR}_{x,r}(f) = \left\langle \frac{\{|U_x(f)|\}}{\{|U_r(f)|\}} \right\rangle_{\log} \tag{1}$$

where $|U_x(f)|$ and $|U_r(f)|$ are the amplitudes of the Fourier's acceleration spectra at s_x and s_r respectively. $\{\cdot\}$ represent the spectral smoothing operator and $\langle\cdot\rangle_{\log}$ represents the logarithmic (geometric) average operator. Fig. 6(b) shows the outcome of this comparison

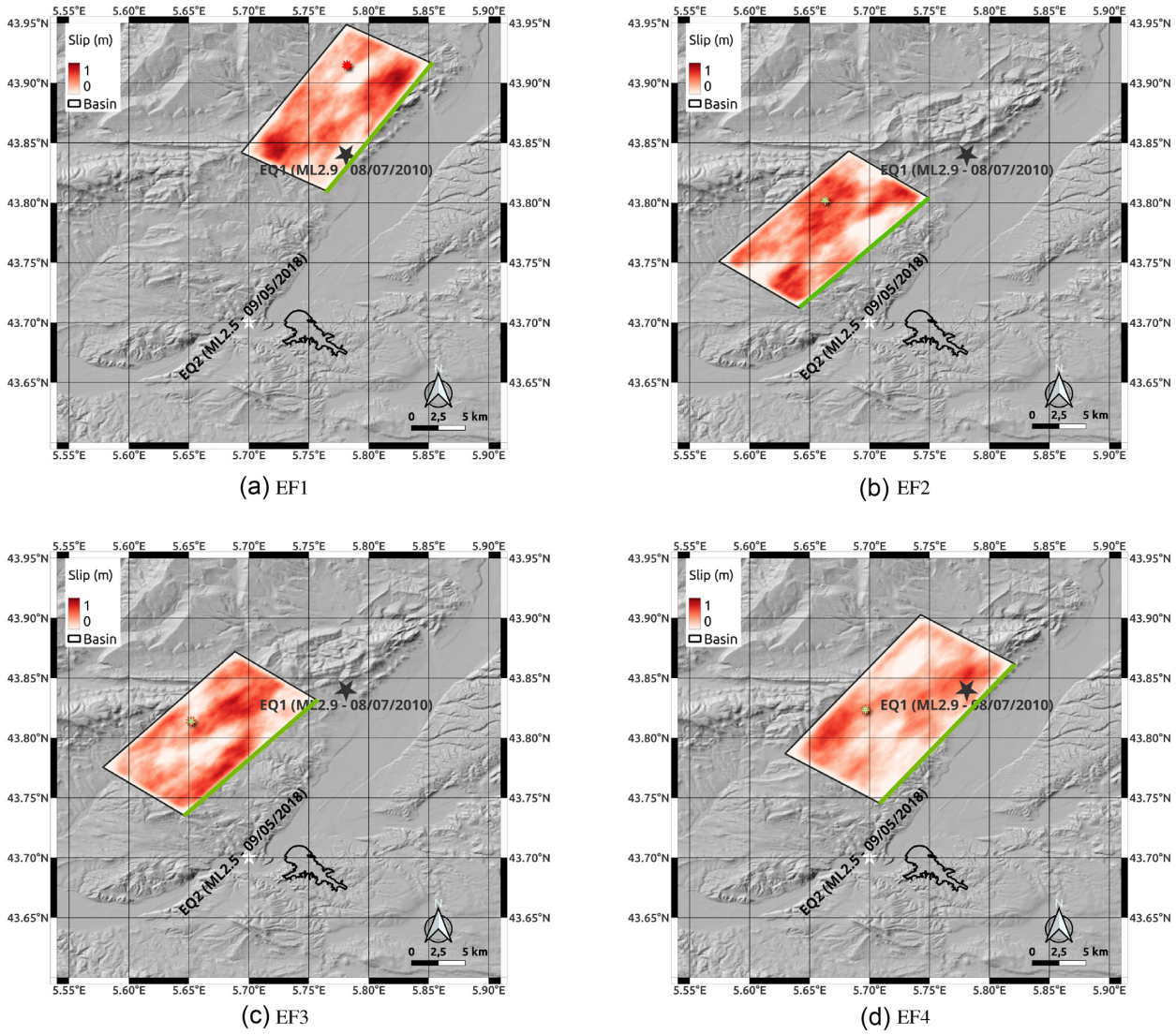

Figure 7. Map of the four kinematic source models adopted in this study, corresponding to an M_W6 target earthquake. The colour-scale represents the slip value. The stars indicate the epicentre's location. Contour lines indicate the depth. The green line indicates the top fault line.

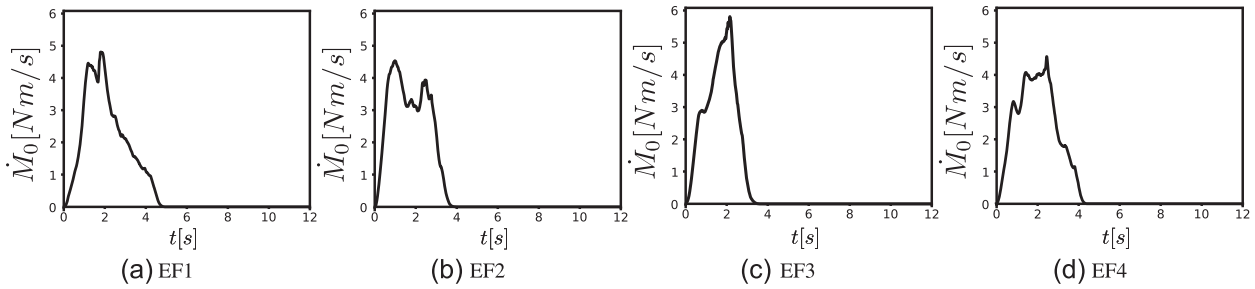

Figure 8. Cumulative moment rate for the four considered source models.

at station s_x CA10, adopting CA01 as reference site (given the good mAC scores at this station on outcropping bedrock). Despite some discrepancies, the synthetic SSR fits well the empirical one at CA10, suggesting that the numerical model has a good level of fidelity in a broad-band frequency range.

3 BLIND PREDICTION OF TARGET SCENARIO

In order to assess the complex interaction between near-fault ground motion and site effects at Cadarache, four blind predictions of plausible M_W6 target earthquake are presented hereafter. For the

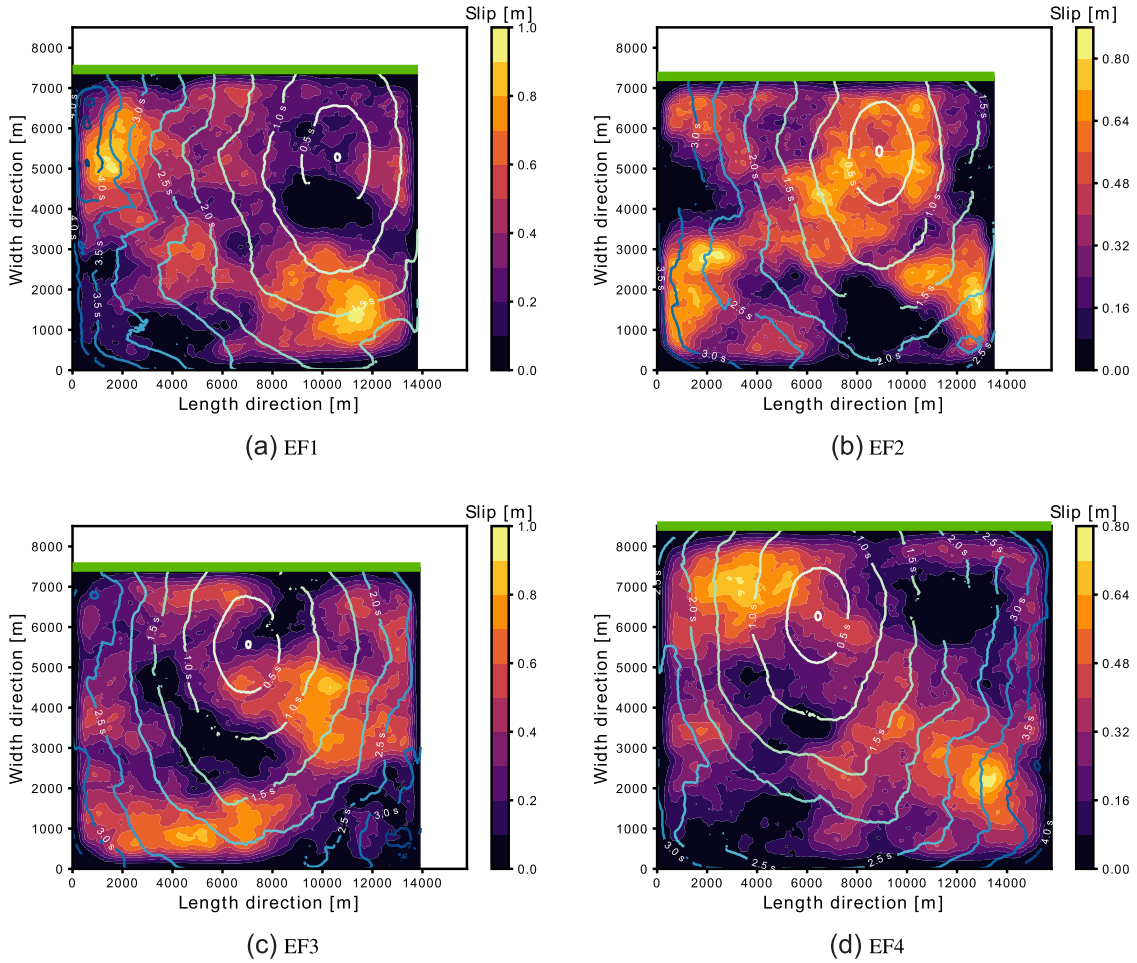

Figure 9. Final slip distributions associated with the four kinematic source models considered in this study. The contour lines represent the rupture front (every 0.5 s). The green line indicates the top fault line.

geometry of the different segments of the Middle Durance Fault, we used the model proposed by Guyonnet-Benaize *et al.* (2015).

3.1 Description of the source models

Four kinematic source models (EF1, EF2, EF3 and EF4, see Fig. 7) were simulated, corresponding to northwest dipping segments of the active fault zone reported by Guyonnet-Benaize *et al.* (2015) The target seismic moment is fixed at 1.04×10^{18} Nm, corresponding to the maximum value reported the French seismic catalogue (Baroux *et al.* 2003) and adopted in other studies in the area (Dujardin *et al.* 2020). The kinematic models are generated using the methodology proposed by Dujardin *et al.* (2020), who performed ground motion simulations for a suite of 100 rupture scenarios on the Middle Durance Fault. The model considers a slip and a rupture time distributions that follow a k^{-2} model, at constant rise-time (Hisada 2000; Dujardin *et al.* 2018; Causse *et al.* 2021). In this paper, four source models were generated, each representing a segment of the Middle Durance Fault (see Table 2). A wide variety of rupture parameters is spanned, namely geometry, position, stress drop, rupture velocity and position of the rupture initiation. The formulation of Dujardin *et al.* (2020) ensures that the generated moment rate functions follow a standard Brune's model (Brune 1970) for a target corner frequency value (or equivalently stress drop value). Once fixed the value of the frequency corner f_C and the characteristic rupture length L_C, the rupture velocity V_R is determined as $V_R = f_C \cdot L_C$. As such, the values of stress drop and of the rupture velocity are implicitly anticorrelated, as reported in recent analyses of databases of source time functions (Chounet *et al.* 2018) or in the analyses of high-frequency ground motion variability (Causse & Song 2015). No correlation between other local source parameters is introduced either, such as between local slip velocity an rupture velocity. In this sense, the adopted approach differs from pseudo-dynamic models, in which local correlation between source parameters are calibrated from spontaneous dynamic rupture simulations (e.g. Schmedes *et al.* 2010; Song 2016). Moreover, provided that the PGA is strongly controlled by the stress drop value $\Delta\sigma$, the latter has been chosen of the order of 1 MPa, as per Dujardin *et al.* (2020). This value is based on the study by Drouet *et al.* (2010), who estimated mean stress drop of $\approx$ 1 MPa for small to moderate earthquakes in the French Alps area using generalized inversion technique. Fig. 8 depicts the seismic moment rate time-history across the fault plane, as the sum of the contributions of each sub-asperity. On the other hand, Fig. 9 shows the final slip distribution and the rupture propagation on the fault plane.

As depicted in Fig. 9, all the rupture paths nucleate at depth, as commonly observed in past earthquakes occurred in the region (Mai 2005), but the position of the rupture initiation varies along the strike, with bilateral propagation.

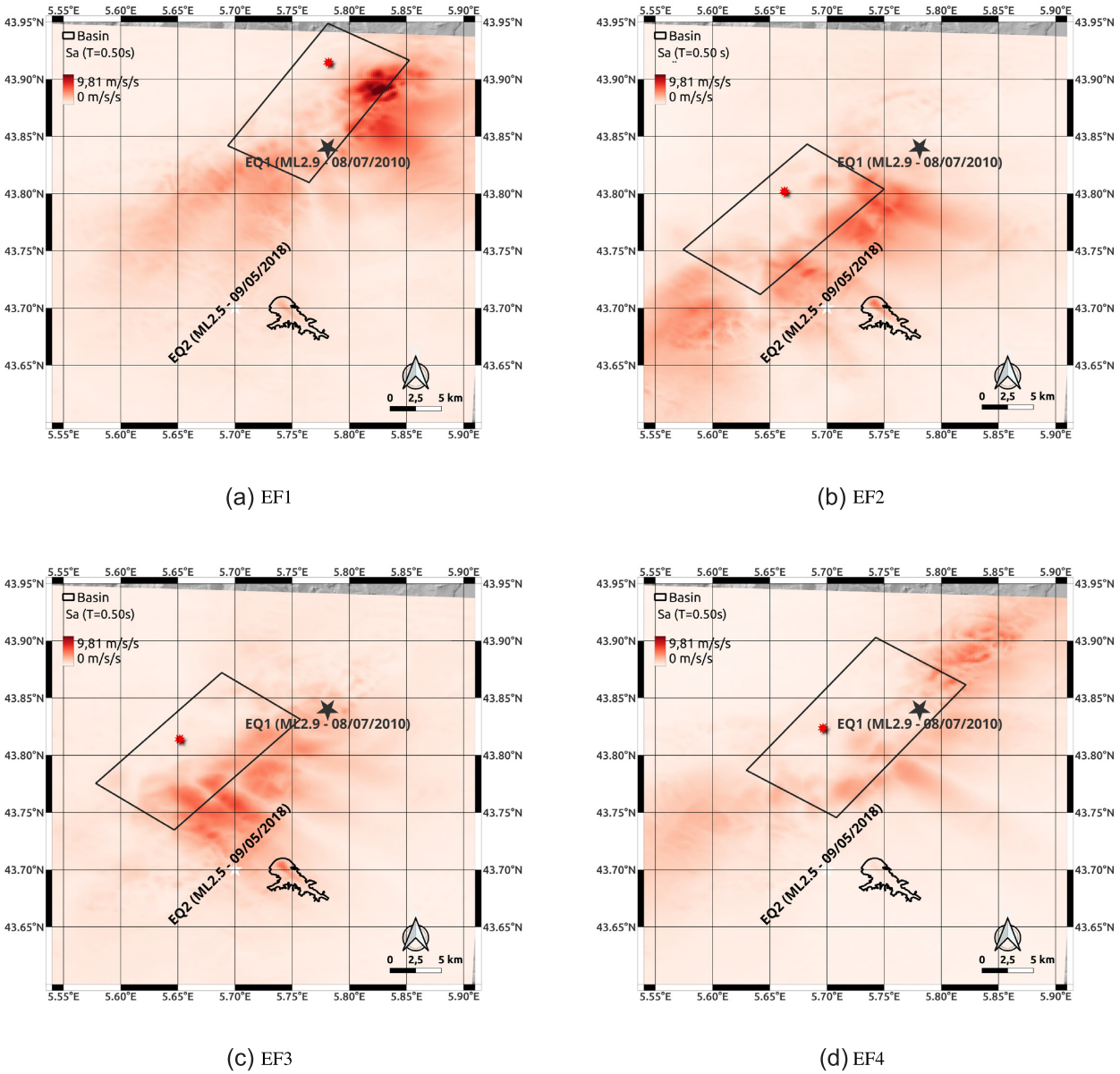

Figure 10. Spatial distribution of the *Sa* values at natural period T=0.5 s for the four different PBS run with different extended faults: (a) EF1, (b) EF2, (c) EF3 and (d) EF4.

Fig. 10 shows the pseudospectral acceleration values *Sa* for a natural period $T = 0.5$ s, close to the natural period computed by Perron *et al.* (2018) from 1-D wave-propagation analysis at the borehole scale within the soft sedimentary basin. Fig. 10 reveals an interesting aspect: besides the expected amplification within the basin, EF2 and EF3 (the closest to the basin) induce large ground shaking in the surrounding (northwestern-ward) of the basin, due to strong updip directivity effects. However, EF1 generate a wider area with large amplification around its hypocentre, mainly due to larger average stress drop and directivity effects towards southwest.

To clear this aspect, Fig. 11 synthesizes the *Sa* variation at each location, with the respect to the site distance to sources' hypocentres and for different natural periods $T = 0.1$, 0.5, 1 and 3 s, respectively. Fig. 11 unveils the fact that EF1 induces the largest *Sa* values at different natural periods, in the distance range of 0–30 km. EF4 induces instead the lower amplification, despite being closer to the CEA site. This is because the CEA site is located in the backward rupture direction. In addition, the large slip area in the southern part of the fault is mainly located beneath the rupture initiation, resulting in very weak updip directivity effects toward the site.

3.2 Assessment of the site effects

In order to assess the possible site effects induced at Cadarache by the four sources, the synthetic strong ground motion time-histories at CA02 (depicted in Fig. 12) and CA04 (depicted in Fig. 14) are analysed. It is evident that EF2 produces the most intense ground shaking, being the closest source to the sedimentary basin, as witnessed by the clear amplification, when comparing CA02 and CA04 response. This amplification can be further appreciated by looking at the *Sa* values in Figs 13 and 15, referring to CA02 and CA04 stations respectively. EF4 showcases the weakest ground motion

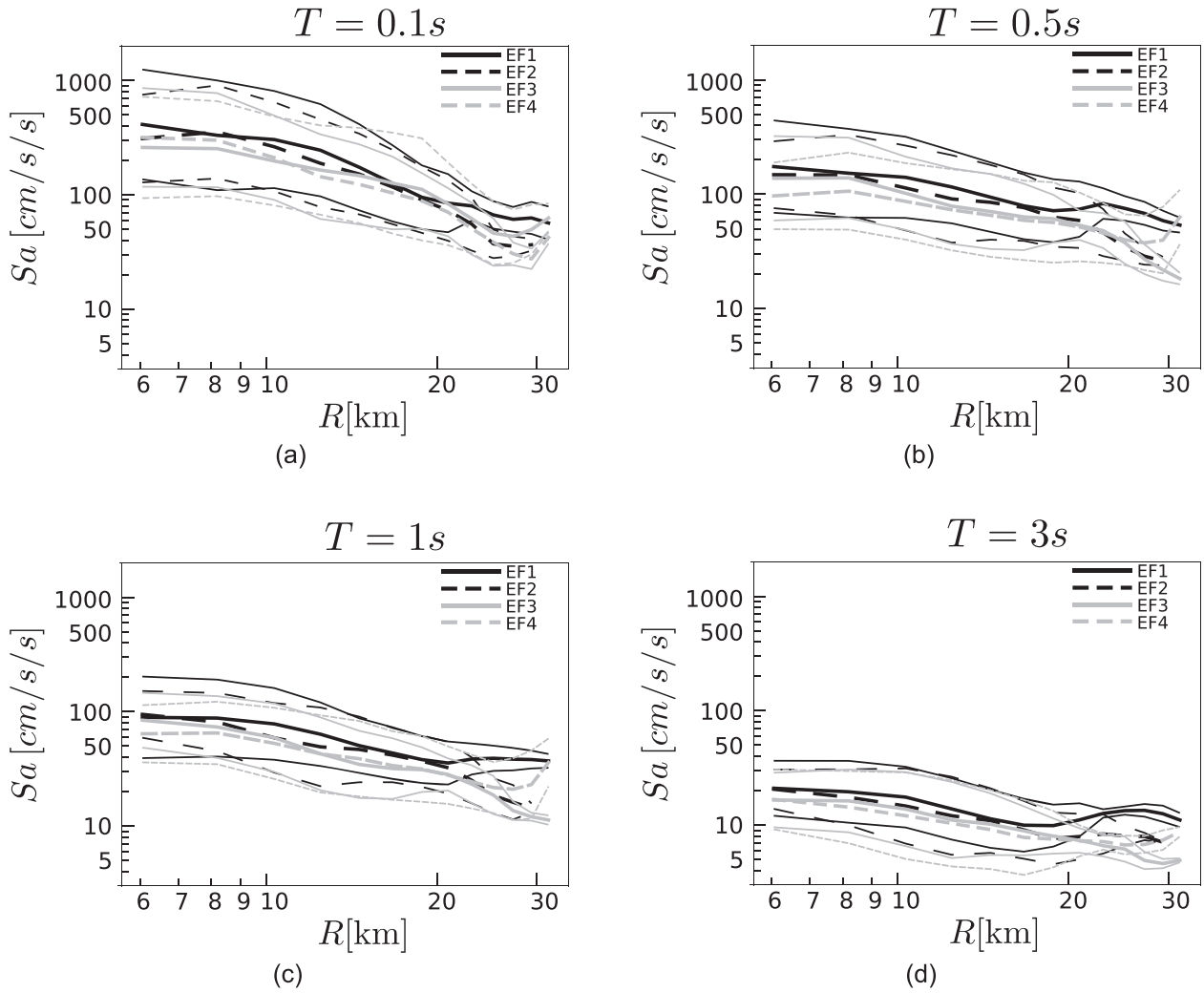

Figure 11. Comparison of the pseudospectral acceleration *Sa* with respect to the station's distance *R* from the hypocentre for a period of $T = 0.1$ s (a), 0.5 s (b), 1 s (c) and 3 s (d). Thicker lines correspond to the geometric mean values, whereas thinner lines correspond to ± a standard deviation.

response, given the position of the nucleation, resulting in weak directivity effects.

In general, the response spectra *Sa* is peaking at around 0.3 Hz for the four scenarios (see Fig. 15). This is in accordance with Figs 10(a)–(d), which already showed that all four EF scenarios entailed an important amplification. The 3-D site-effects taking place in the surrounding of the Cadarache site can be assessed at two cross-sections of the sedimentary basin (depicted in Fig. 1), one (named AA', see Fig 16a) striking along the normal direction and one (named BB', see Fig. 16b) striking along transversal one. From a qualitative standpoint, Figs 16(c)–(f) show the interferograms generated by EF3 on the two cross-sections AA'and BB', along the two horizontal dimensions. The wave field interaction with the basin traps it in and it causes the well-known multiple reflection phenomenon. Site-effects are quantified according to several classical indicators: the horizontal average amplitude of Fourier's transform of the acceleration time-history ($\mathrm{FSA_H}$), the horizontal geometric average response spectra (Sa_H), the average synthetic horizontal spectral ratio ($\mathrm{SSR_H}$, see eq. 4) and the H/V ratio. $\mathrm{FSA_H}$, Sa_H, $\mathrm{SSR_H}$ and H/V reads respectively (Field & Jacob 1995; Castro-Cruz *et al.* 2020):

$$\mathrm{FSA_H}(f) = \frac{1}{2}\left\{\sqrt{|U_{\mathrm{EW}}(f)|^2 + |U_{\mathrm{NS}}(f)|^2}\right\} \tag{2}$$

$$\mathrm{Sa_H}(T) = 10^{\frac{1}{2}\left(\log_{10} Sa_{\mathrm{EW}}(T) + \log_{10} Sa_{\mathrm{NS}}(T)\right)} \tag{3}$$

$$\mathrm{SSR_H}(f) = \frac{\left\{\sqrt{|U_{\mathrm{EW}}(f)|^2 + |U_{\mathrm{NS}}(f)|^2}\right\}}{\left\{\sqrt{|U_{\mathrm{EW},r}(f)|^2 + |U_{\mathrm{NS},r}(f)|^2}\right\}} \tag{4}$$

$$\mathrm{H/V}(f) = \frac{1}{2}\frac{\left\{\sqrt{|U_{\mathrm{EW}}(f)|^2 + |U_{\mathrm{NS}}(f)|^2}\right\}}{\{|U_{\mathrm{UD}}(f)|\}} \tag{5}$$

Figs 17 and 18 show—for profile AA' and profile BB' respectively—the spatial variation of the four indicators above mentioned. Quite interestingly, in the case of profile AA', there is a twofold amplification corresponding to the deepest point in the basin, as witnessed by $\mathrm{FSA_H}$, Sa_H and $\mathrm{SSR_H}$, very similar outside and inside the basin, showing a peak around 0.05 Hz. The dominant frequency within the basin is ≈5 Hz, whereas where the basin gets deeper, a second main peak appears at ≈7 Hz.

The $\mathrm{SSR_H}$ depicted in Figs 17 and 18 shows that the basin amplifies the incident wave motion at frequencies larger than 2 Hz. As expected, the fundamental frequency highlighted by the H/V ratio changes as a function of the thickness of the basin. For example, at the deepest part (profile AA' around +1 km), the fundamental frequency ranges around 3 Hz and at the shallow part of the basin

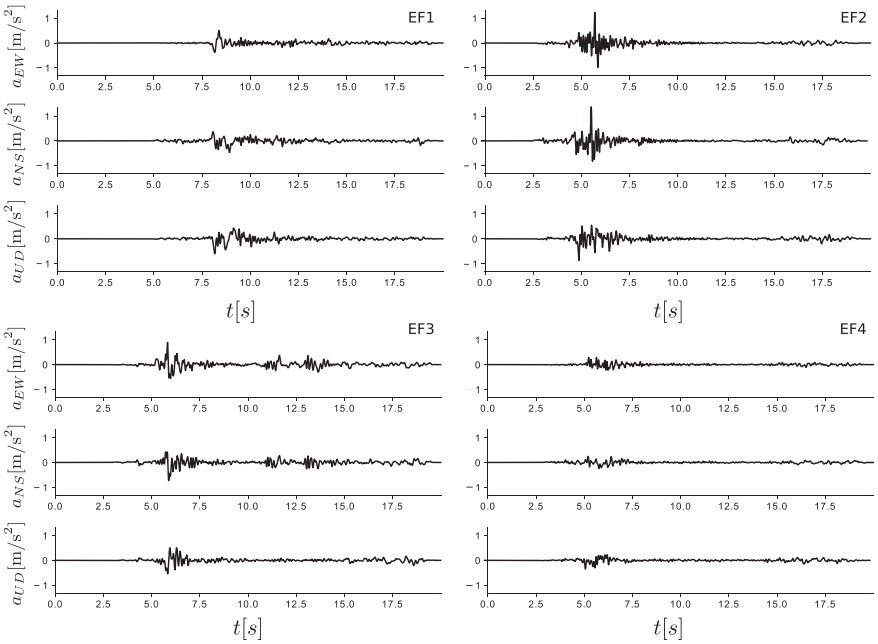

Figure 12. Synthetic acceleration time-histories $a(t)$ along EW, NS and UD direction, respectively, obtained for the four EF scenarios at CA02 station (located on outcrop bedrock).

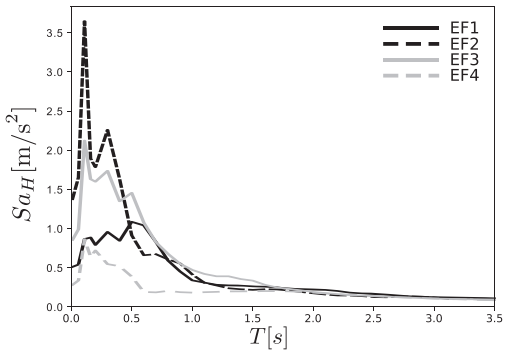

Figure 13. Synthetic Sa_{H} spectra (horizontal geometric average) obtained at CA02 station (located on outcrop bedrock) for the four EF scenarios.

around 8 Hz. The presence of the basin affects the response spectrum around the basin itself, with dominant peaks at ≈0.2 s (AA') and ≈0.1 s (BB'), respectively. For long periods, the basin does not induce remarkable amplifications.

3.3 Comparison of synthetic and empirical Ground Motion Models

To check the predictive capability of the synthetic earthquake generator, the synthetic *Sa* values (Sa_{model}) at different natural periods and for each scenario is hereafter compared to the GMPE developed by Berge-Thierry *et al.* (2003), widely used for the area of interest (Dujardin *et al.* 2020) and to the more recent GMPE proposed by Kotha *et al.* (2020) within the framework of a recent European effort to update the seismic hazard and risk assessment tools. In both cases, the reference intensity measure adopted in this paper is the *Sa* at different natural periods.

Comparison with Berge-Thierry et al. (2003). The functional form of median intensity measure of the Berge-Thierry *et al.* (2003) GMPE reads:

$$\ln Sa_{\mathrm{GMPE}}(T) = a(T)M + b(T)d_{\mathrm{hyp}} - \ln d_{\mathrm{hyp}} + c_1(T) + c_2(T) \tag{6}$$

with $a(T)$, $b(T)$, c_1 and c_2 are deterministic regression coefficients. Berge-Thierry *et al.* (2003) calibrated the GMPE in eq. (6) upon recorded time-histories on European territory and in the United States (an early version of the ESM database Ambraseys *et al.* 1996), within the range of surface wave magnitude 4 $\leq M_{\mathrm{S}} \leq$ 7.9 and in a range of hypocentral distance 4 km $\leq d_{\mathrm{hyp}} \leq$ 330 km and depths below 30 km (shallow crustal earthquakes, according to the average French seismotectonic context). This GMPE discriminates between two site conditions based on the value of $V_{\mathrm{S,\,30}}$: a site is classified as *bedrock* whether $V_{\mathrm{S,\,30}} \geq 800$ m s^{-1} ($c_1(T) \neq 0$, $c_2(T) = 0$); a site is classified as *alluvium* whether 300 m s^{-1} $\leq V_{\mathrm{S,\,30}} <$ 800 m s^{-1} ($c_1 = 0$, $c_2(T) \neq 0$). Fig. 19 shows the log-difference between synthetic and empirical ground motion prediction, averaged over d_{hyp}. Stations within the basin's borders (depicted in Fig. 19b) were considered as alluvial sites, whereas stations outside the basin were considered as rock sites (see Fig. 19a), in agreement with the $V_{\mathrm{S,\,30}}$ classification made by Berge-Thierry *et al.* (2003).

Comparison with Kotha et al. (2020). The functional form of median intensity measure of the Kotha *et al.* (2020) GMPE reads:

$$\ln Sa_{\mathrm{GMPE}}(T) = e_1 + f_{R,g}(M_{\mathrm{W}}, R_{\mathrm{JB}}, T) + f_{R,a}(R_{\mathrm{JB}}, T) + f_M(M_{\mathrm{W}}, T) \tag{7}$$

with e_1 the generic offset, $f_{R,\,g}$ the functional form representing the geometric spreading, $f_{R,\,a}$ the functional form representing the apparent anelastic attenuation (region-wise) and f_M the magnitude scaling. The R_{JB} represents the Joyner–Boore distance. Without entering into much detail (we remand to the original publication Kotha *et al.* 2020), eq. (7) represents the *fixed* effects of the GMPE, with deterministic coefficients. Those coefficients are then adjusted in order to take into account the random effects related to the event, to the tectonic locality, to the attenuation region, and to the site specific ground-motions.

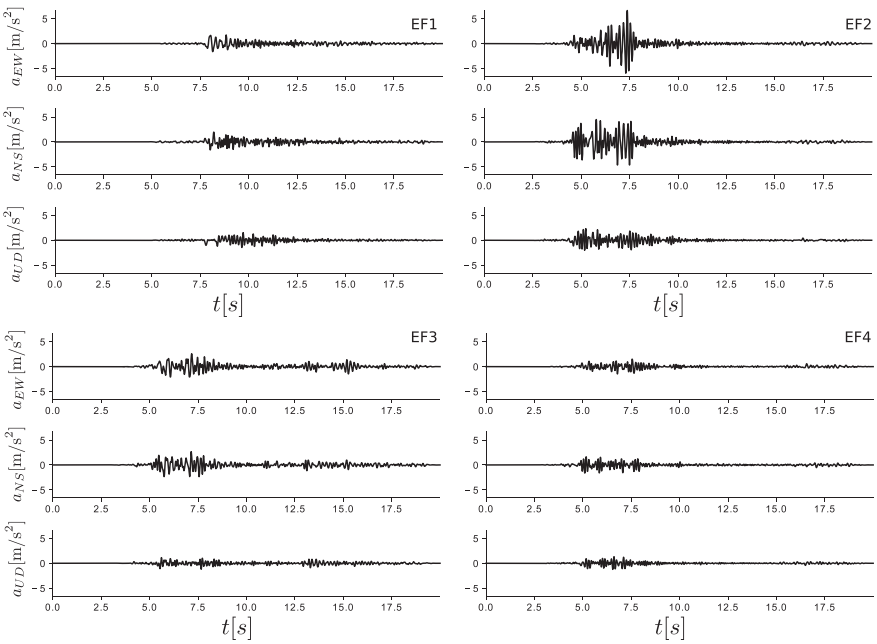

Figure 14. Synthetic acceleration time-histories $a(t)$ along EW, NS and UD direction respectively, obtained for the four EF scenarios at CA04 station (located within sedimentary basin).

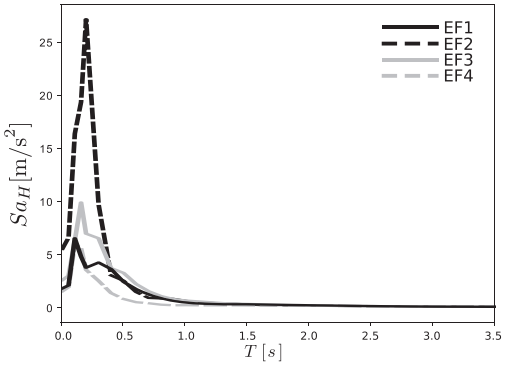

Figure 15. Synthetic Sa_{H} spectra (horizontal geometric average) obtained at CA04 station (located within sedimentary basin) for the four EF scenarios.

Compared to Berge-Thierry *et al.* (2003), Kotha *et al.* (2020) calibrated their GMPE on more than 18 000 records from 927 shallow crustal events ($3.1 \leq M_{\mathrm{W}} \leq 7.4$) recorded at 1829 stations (with $0 \leq R_{JB} < 545$ km). The records were selected within the new European strong-motion data set (Bindi *et al.* 2019), based on a harmonized M_{W} estimate and selecting stations that recorded more than 3 events. Fig. 20 shows the comparison between the GMPE et the numerical simulations performed in this study, inside and outside the basin. The ratios are averaged over R_{JB}, at each natural period. Fig. 20 depicts the comparison between the synthetic results obtained via PBS and the GMPE by Kotha *et al.* (2020). The GMPE proposed by Kotha *et al.* (2020) does not discriminate the records based on the available (or not) $V_{\mathrm{S,\,30}}$ value. On the contrary, site-to-site response variability is captured by the site-specific random-effects. However, in Fig. 20 stations were distinguished as *inside* and *outside* the basin according to the same classification adopted in Fig. 19 for the Berge-Thierry *et al.* (2003) GMPE model.

4 DISCUSSION

The comparisons in Figs 19(a) and 20(a) for a dense grid of stations located on the outcrop-bedrock outside the basin (but in its surroundings) stresses that for all four PBS, the numerical prediction lies within one σ variability of both the GMPEs, for periods $T > 1.5$ s. For lower natural periods, the EF4 response underestimates the GMPE value $Sa_{\mathrm{GMPE}} - \sigma$. However, Berge-Thierry *et al.* (2003) defined as *rock site* the stations at which the $V_{\mathrm{S,\,30}} > 800$ m s^{-1} and *alluvium site* those with 300 m s^{-1} $< V_{\mathrm{S,\,30}} <$ 800 m s^{-1}. Given the available geological profiles, depicted in Fig. 2, only the first 10 m down-hole can be considered as *alluvium site* in the present case. Berge-Thierry *et al.* (2003) commented on the definition of the bedrock condition in their original work, mentioning the fact that the site coefficients as a function of frequency show very small difference between rock and alluvium sites, with an average amplification of 21 per cent. They further admitted that their *rock records* have a specific intrinsic site effect and underlined the difficulty to qualify a site as *a rock site*. Alternatively, the classification proposed by Boore (2003) was applied, using the available geological and geophysical information. This uncertainty can explain the discrepancy with our model.

The comparison between the four PBS and the GMPE proposed by Kotha *et al.* (2020) seems slightly improved, compared to the GMPE proposed by Berge-Thierry *et al.* (2003), considering the tighter uncertainty margins. This is probably due to the flexibility of the Kotha *et al.* (2020) model, which introduces site-to-site random variability and that is well calibrated on small-moderate magnitude ranges. Moreover, despite being associated with a rather low maximum slip (compared to other sources) EF3 scenario displays the second largest intensity for all periods, possibly due to a larger average rupture velocity (see Fig. 10c). Within the basin (see Figs 19b and 20b), the simulated results approach the GMPE for scenario EF1, mainly at high periods ($T > 2$ s).

EF4 showcases the lowest intensity both inside and outside the basin and for both GMPEs (see Fig. 9d). One can argue that this

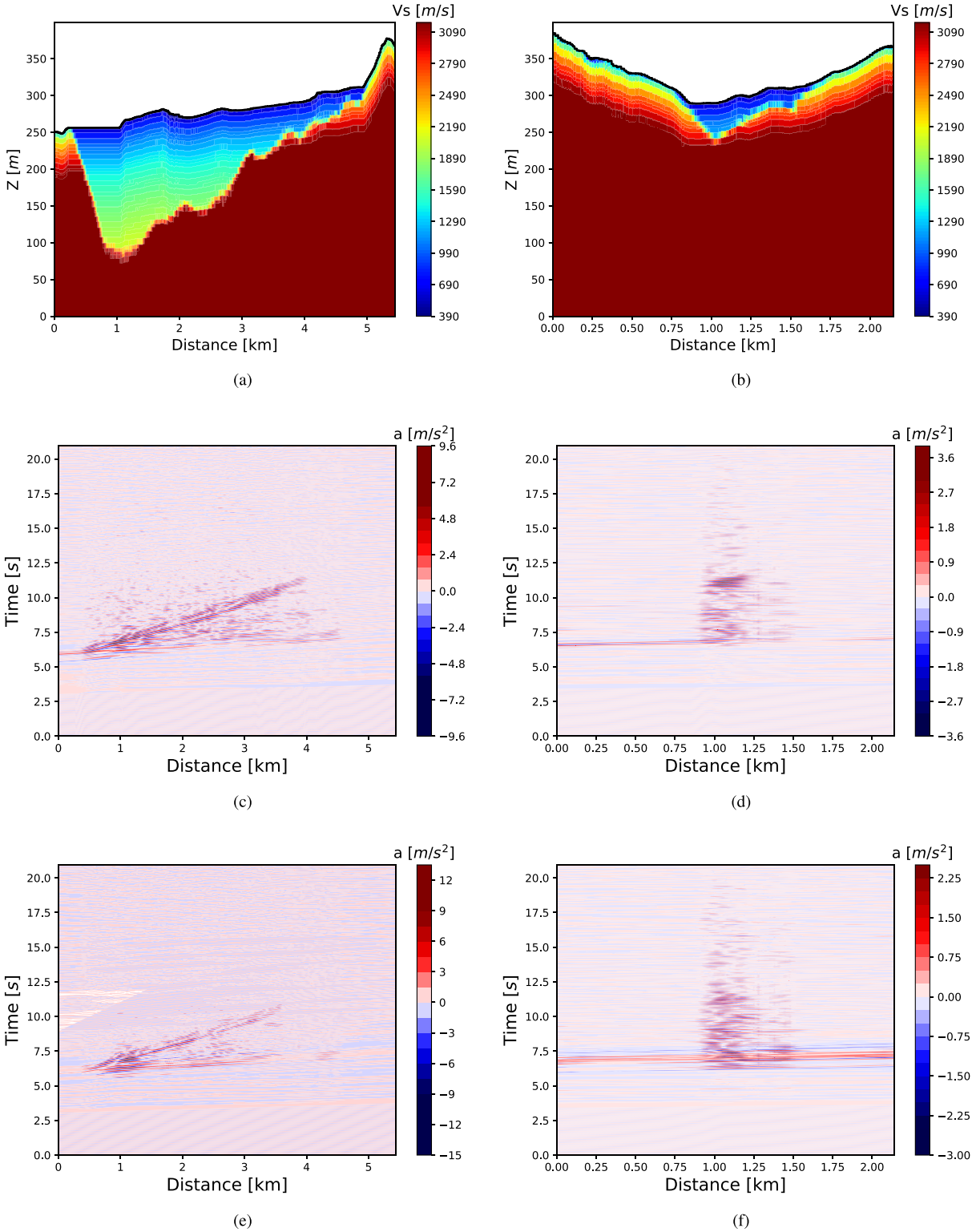

Figure 16. Profile AA' (a) and Profile BB' (b) shear wave velocity (V_S) across the two 2-D cuts. Profile AA' (c–e) and profile BB' (d–f) interferogram across the normal SH direction (c and d) and transverse *P*–*SV* direction (e and f).

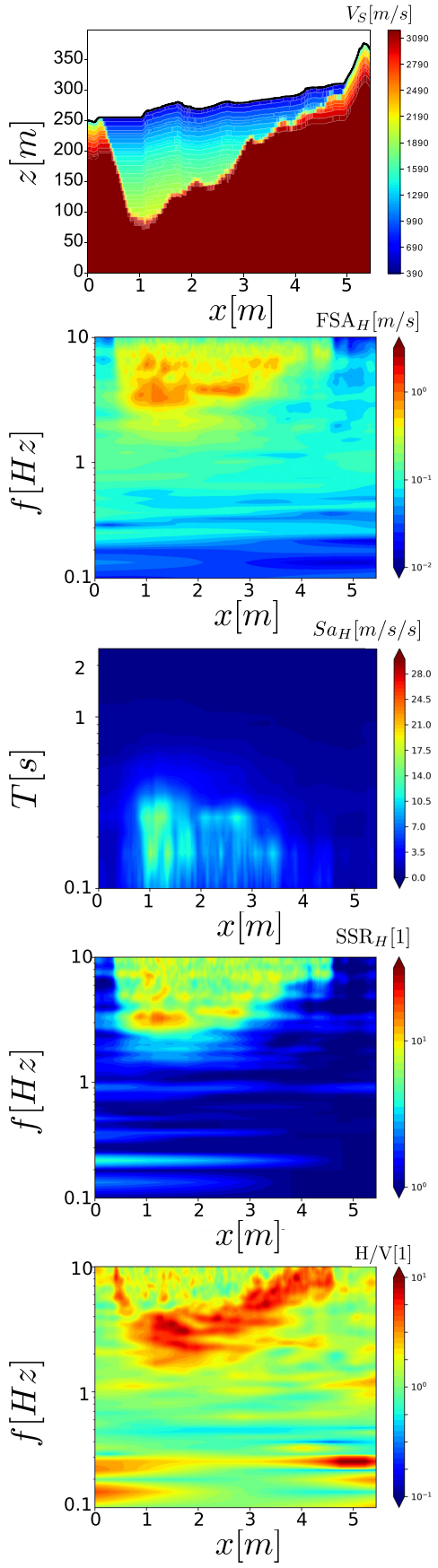

Figure 17. Profile AA': 2-D V_S profile; $\mathrm{FSA_H}$; Sa_H; $\mathrm{SSR_H}$; H/V ratio.

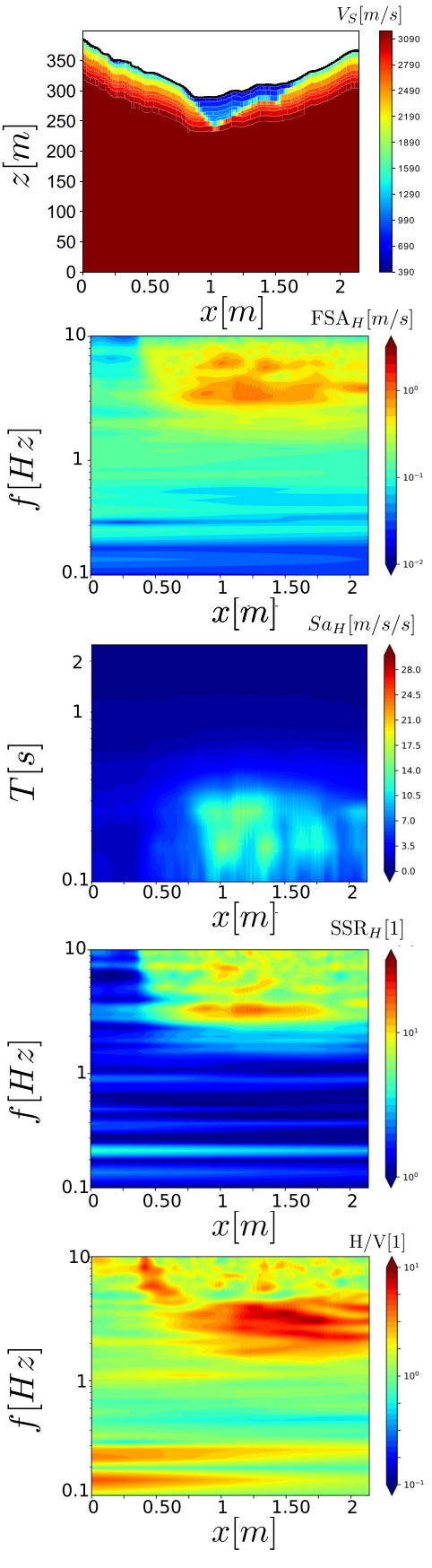

Figure 18. Profile BB': 2-D V_S profile; $\mathrm{FSA_H}$; Sa_H; $\mathrm{SSR_H}$; H/V ratio.

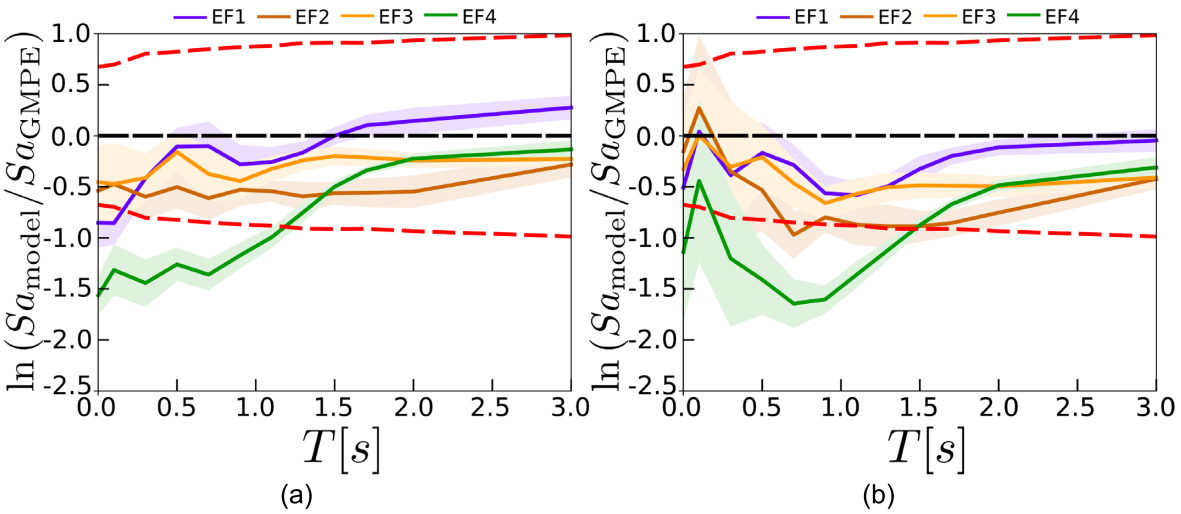

Figure 19. Residual between the simulated Sa_{model} and the Berge-Thierry *et al.* (2003) GMPE prediction Sa_{GMPE}. Red dashed lines indicate the residual for one standard deviation $Sa_{\text{GMPE}} \pm \sigma$. (a) Residuals computed from stations outside the sedimentary basin; (b) residuals computed for all the stations inside the sedimentary basin (based on the $V_{\text{S, 30}}$ values).

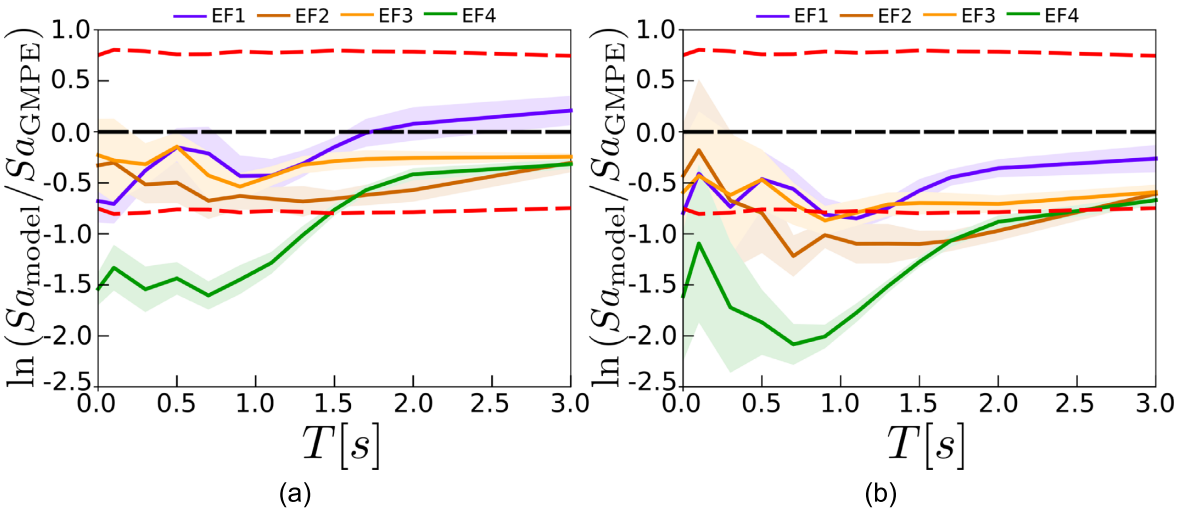

Figure 20. Residual between the simulated Sa_{model} and the Kotha *et al.* (2020) GMPE prediction Sa_{GMPE}. Red dashed lines indicate the residual for one standard deviation $Sa_{\text{GMPE}} \pm \sigma$. (a) Residuals computed from stations outside the sedimentary basin; (b) residuals computed for all the stations inside the sedimentary basin (as in Fig. 19).

outcome is related to the choice of the stress drop. However, the adopted stress drop $\approx$ 1 MPa corresponds to a corner frequency value of 0.2 Hz, which is consistent with worldwide observations for an $M_{\text{W}}6$ earthquake (e.g., Allmann & Shearer 2009). Thus, there is no evidence that the chosen stress drop value explains the difference between simulations and GMPEs at low periods. A fine analysis of the stress drop values of earthquakes used in the databases adopted to calibrate the GMPEs would be necessary to address this issue.

Another possible explanation of the underestimation of the empirical GMPE models could be the poor geological information introduced in the numerical model in order to successfully constrain the ground motion prediction (with numerical accuracy up to $\approx$ 0.1 s). As a matter of fact, both the Berge-Thierry *et al.* (2003) and Kotha *et al.* (2020) GMPE estimation of the site effect are rather simplified compared to numerical simulations.

To appreciate the spatial variability of the site response, Figs 21 and 22 show the spatial distribution of the GMPE residuals, of *Sa* values at T =0.5 s and the empirical prediction obtained with the GMPE by Berge-Thierry *et al.* (2003) and by Kotha *et al.* (2020) respectively.

Results are rather similar, despite the fact that the former GMPE takes into account the site effects through $V_{\text{S, 30}}$. The numerical simulation overestimates the GMPEs at the deepest points within the basin (along AA' profile) for all scenarios, whereas at the edges of the basin simulations and GMPEs seem to be in agreement. A general underestimation is observed for EF4 simulations, all over the area surrounding the basin. This seems to confirm that EF4 is strongly affected by the particular rupture scenario, but a refined numerical model is however required to render a smoother transition from outcropping bedrock and sedimentary basin.

It must be noted that comparing GMPE with the average outcome of PBS is always a rather tricky exercise. This is due to an ontological difference between the two methods, developed to solve the same problem: the lack of massive site-specific seismic observations over a dense spatial grid. GMPE are the result of complex nonlinear regressions on recorded data, but which in general do not include

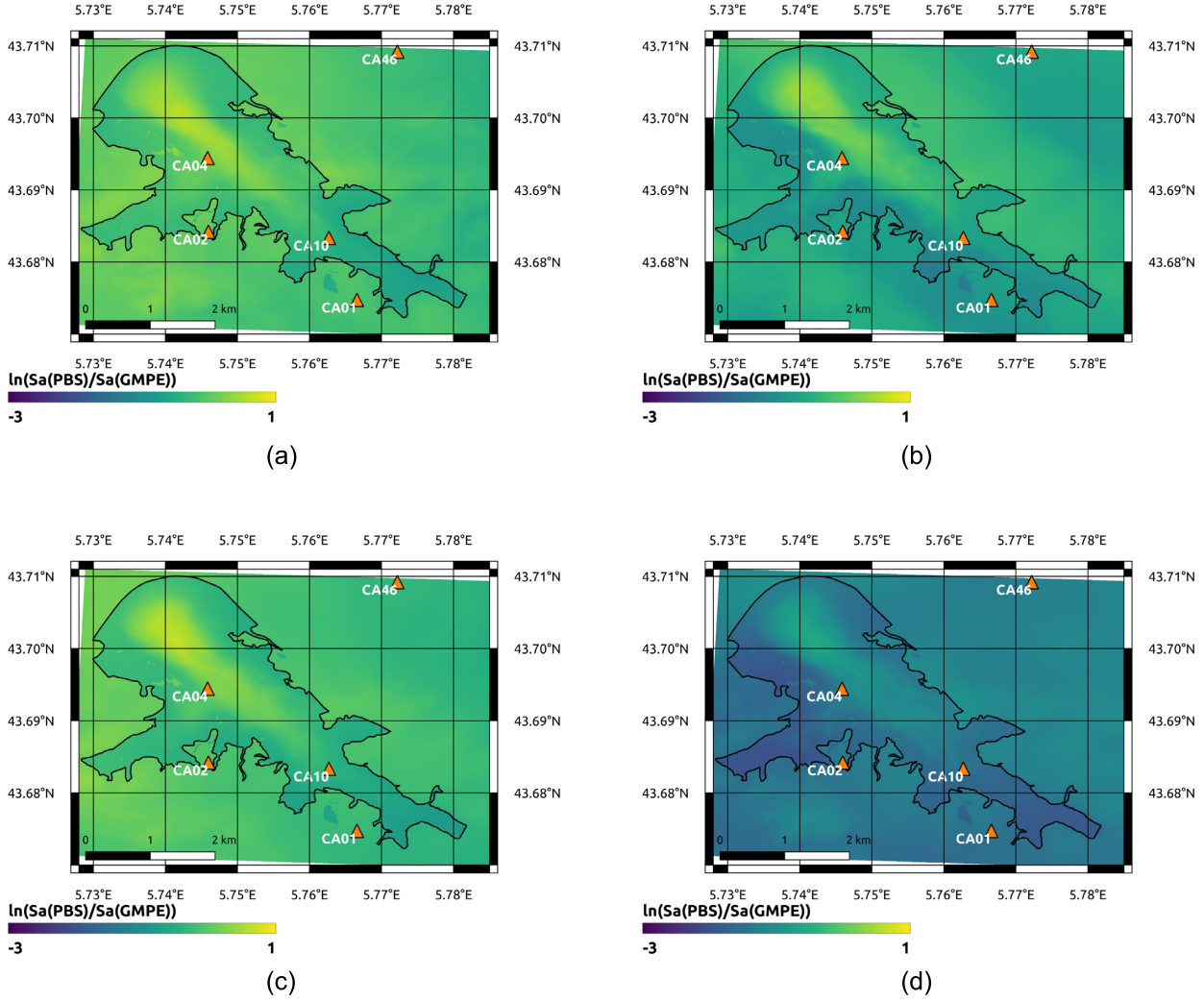

Figure 21. Spatial distribution of the residual $\ln \frac{Sa_{\mathrm{PBS}}(T=0.5\ \mathrm{s})}{Sa_{\mathrm{GMPE}}(T=0.5\ \mathrm{s})}$ considering the GMPE model by Berge-Thierry *et al.* (2003) for EF1 (a), EF2 (B), EF3 (c) and EF4 (d).

the ones recorded at the site of interest. Hopefully, along with general purpose GMPEs, region-specific GMPEs have being developed, targeting specific tectonic context. However, GMPEs cannot estimate complex site and basin effects, neither the spatial variability and incoherence of ground motion, essentially because they aim at condensing the large variability and high complexity of the earthquake phenomenon into an equation. Despite being user-friendly, especially when compared to PBS, the GMPE formulation overlooks some complex site-wise characteristics of the ground motion, which on the contrary might be modelled (*a priori*) and estimated with PBS (along with basin effects, surface waves and several other intricate issues of seismic wave propagation). PBS require quite an expertise in numerical modelling, as well as consequent computational resources. Moreover, PBS must be duly validated beforehand, provided that the numerical model can be updated as long as new geophysical and seismological data are acquired. The representative power of PBS is potentially infinite, which makes it extremely hard to constrain to realistic predictions.

When estimating the seismic response of critical sites, such as Cadarache, the two methods must be adopted as complementary predictions, rather than aiming at making the two of them converging to similar results.

5 CONCLUSIONS

This paper compares synthetic and empirical blind predictions of a target $M_{\mathrm{W}}6.0$ earthquake at the nuclear site of Cadarache (France). The numerical model is an improvement of the one constructed and validated by Castro-Cruz *et al.* (2021). The mesh is designed so as to reach a maximum accuracy of 10.5 Hz. The model implements the latest update of the regional geological model constructed and maintained by the CEA, based on *in situ* experimental campaigns. The model is at first validated against a small $M_{\mathrm{L}}2.5$ earthquake recordings, leading to a fair comparison (assessed via modified Anderson's criteria) despite the intrinsic difficulty and uncertainty of validating a geological model against such a small earthquake and in such a large frequency band. Four alternative $M_{\mathrm{W}}6.0$ extended fault blind scenarios are numerically generated (using a k^{-2} kinematic model) and compared to the empirical ground motion prediction equation (GMPE) proposed by Berge-Thierry *et al.* (2003) and by Kotha *et al.* (2020), widely employed for the French and European seismic contexts. The synthetic predictions comply with the GMPE median $\pm\sigma$ estimation, although for high periods, the synthetic prediction is highly affected by the directivity effect proper to each extend fault scenario. Moreover, the physics-based

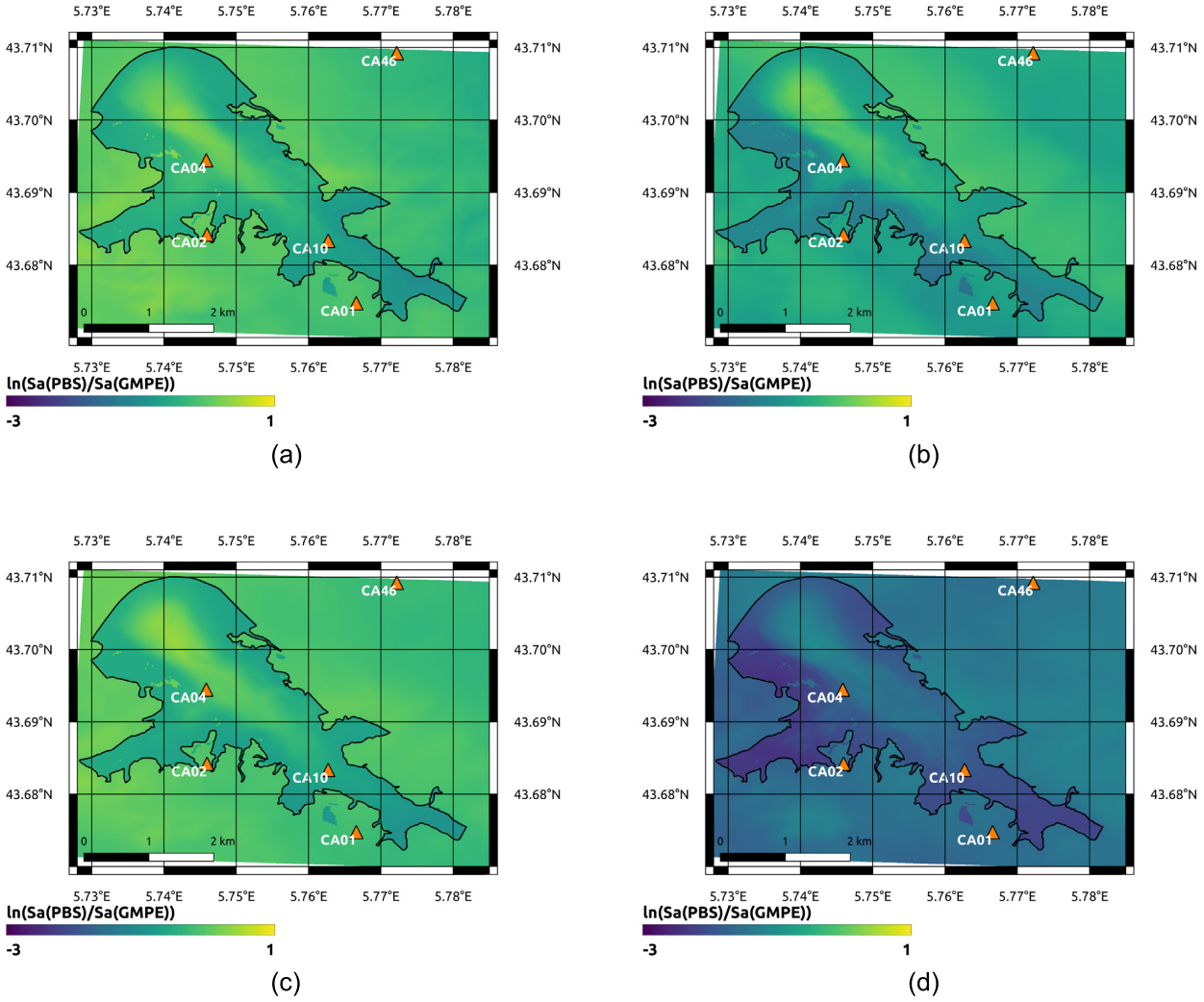

Figure 22. Spatial distribution of the residual $\ln \frac{Sa_{\mathrm{PBS}}(T=0.5\ \mathrm{s})}{Sa_{\mathrm{GMPE}}(T=0.5\ \mathrm{s})}$ considering the GMPE model by Kotha *et al.* (2020) for EF1 (a), EF2 (B), EF3 (c) and EF4 (d).

simulations capture the spatial distribution of 3-D site effects, differing from the GMPE estimation which simply discerns between rock and alluvium sites based on the $V_{S,30}$ estimation. Because of the shape of the basin, a stronger amplification is obtained at high frequencies across the transverse basin axis, whereas lower amplification is observed along the longitudinal one. The numerical model represents the first milestone predictive tool for the seismic site response of the critical Cadarache nuclear site.

The outcome of this paper backs the complementary role played by synthetic and empirical GMPEs with the respect to an improved site response analysis in low-moderate seismic contexts.

DATA AVAILABILITY

All simulated data generated for the sake of editing this paper are available upon contacting the corresponding author and upon reasonable request.

ACKNOWLEDGMENT

Numerical simulations were performed by David Castro-Cruz, who post-processed the results and generated the figures. The source models were provided by Elias El Haber and Mathieu Causse. David Castro-Cruz wrote the first draft of the paper. Filippo Gatti (corresponding author) edited the manuscript, from the original draft to its actual version. Mathieu Causse collaborated with the paper drafting, focusing on the paragraphs dealing with the seismic source models and comparison with empirical ground motion models. Fabrice Hollender provided the geological data. He and Fernando Lopez-Caballero supervised the analysis of the results and the paper drafting and editing.

REFERENCES

Allmann, B.P. & Shearer, P.M., 2009. Global variations of stress drop for moderate to large earthquakes, *J. geophys. Res.,* **114**(B01310), 1–22.

Ambraseys, N.N., Simpson, K.A. & Bommer, J.J., 1996. Prediction of horizontal response spectra in Europe, *Earthq. Eng. Struct. Dyn.,* **25**(4), 371–400.

Anderson, J.G., 2004. Quantitative measure of the goodness-of-fit of synthetic seismograms, in *13th World Conference on Earthquake Engineering,* Paper No. 243, Vancouver, BC, Canada, pp. 775–784.

Andrews, A. & Folger, P., 2012. Nuclear power plant design and seismic safety considerations, Tech. rep., Congressional Research Service.

Baroux, E., Pino, N.A., Valensise, G., Scotti, O. & Cushing, M.E., 2003. Source parameters of the 11 June 1909, Lambesc (Provence, southeastern France) earthquake: a reappraisal based on macroseismic, seismological, and geodetic observations, *J. geophys. Res.,* **108**(B9, 2454), 1–23, doi:10.1029/2002JB002348.

Berge-Thierry, C., Cotton, F., Scotti, O., Griot-Pommera, D.-A. & Fukushima, Y., 2003. New empirical response spectral attenuation laws for moderate European earthquakes, *J. Earthq. Eng.,* **7**(2), 193–222.

Berge-Thierry, C., Hollender, F., Guyonnet-Benaize, C., Baumont, D., Ameri, G. & Bollinger, L., 2017. Challenges ahead for nuclear facility site-specific seismic hazard assessment in France: the alternative energies and the Atomic Energy Commission (CEA) vision, in *Pure and Applied Geophysics,* pp. 3609–3633 , doi:10.1007/s00024-017-1582-2.

Bindi, D., Kotha, S.-R., Weatherill, G., Lanzano, G., Luzi, L. & Cotton, F., 2019. The pan-European engineering strong motion (ESM) flatfile: consistency check via residual analysis, *Bull. Earthq. Eng.,* **17**(2), 583–602.

Bommer, J.J., 2003. Uncertainty about the uncertainty in seismic hazard analysis, *Eng. Geol.,* **70**(1), 165–168.

Boore, D.M., 2003. Simulation of ground motion using the stochastic method, *Pure appl. Geophys.,* **160**(3), 635–676.

Bradley, B.A., 2019. On-going challenges in physics-based ground motion prediction and insights from the 2010-2011 Canterbury and 2016 Kaikoura, New Zealand earthquakes, *Soil Dyn. Earthq. Eng.,* **124,** 354–364.

Brune, J.N., 1970. Tectonic stress and the spectra of seismic shear waves from earthquakes, *J. geophys. Res.,* **75**(26), 4997–5009.

Castro-Cruz, D., Régnier, J., Bertrand, E. & Courboulex, F., 2020. A new parameter to empirically describe and predict the non-linear seismic response of sites derived from the analysis of Kik-Net database, *Soil Dyn. Earthq. Eng.,* **128,** 105833.

Castro-Cruz, D., Gatti, F. & Lopez-Caballero, F., 2021. High-fidelity broadband prediction of regional seismic response: a hybrid coupling of physics-based synthetic simulation and empirical Green's functions, *Nat. Hazards,* **108,** 1997–2031.

Causse, M. & Song, S.G., 2015. Are stress drop and rupture velocity of earthquakes independent? Insight from observed ground motion variability, *Geophys. Res. Lett.,* **42**(18), 7383–7389.

Causse, M., Cornou, C., Maufroy, E., Grasso, J.-R., Baillet, L. & El Haber, E., 2021. Exceptional ground motion during the shallow M_w 4.9 2019 Le Teil earthquake, France, *Commun. Earth Environ.,* **2**(1), 1–9.

Commissariat à l'énergie atomique et aux énergies alternatives and CentraleSupélec and Institut de Physique du Globe de Paris and CNRS, 2017. *SEM3D Ver 2017.04 Registered at French Agency for Protection of Programs (Dépôt APP)*.

Chounet, A., Vallée, M., Causse, M. & Courboulex, F., 2018. Global catalog of earthquake rupture velocities shows anticorrelation between stress drop and rupture velocity, *Tectonophysics,* **733,** 148–158.

Cushing, E.M. *et al.*, 2020. Building a three dimensional model of the active Plio-Quaternary basin of Argostoli (Cephalonia Island, Greece): an integrated geophysical and geological approach, *Eng. Geol.,* **265,** 105441.

Drouet, S., Cotton, F. & Guéguen, P., 2010. v_{S30}, k, regional attenuation and M_w from accelerograms: application to magnitude 3-5 French earthquakes, *Geophys. J. Int.,* **182**(2), 880–898.

Dujardin, A., Causse, M., Berge-Thierry, C. & Hollender, F., 2018. Radiation patterns control the near-source ground-motion saturation effect, *Bull. seism. Soc. Am.,* **108**(6), 3398–3412.

Dujardin, A., Hollender, F., Causse, M., Berge-Thierry, C., Delouis, B., Foundotos, L., Ameri, G. & Shible, H., 2020. Optimization of a simulation code coupling extended source (k-2) and empirical Green's functions: application to the case of the Middle Durance Fault, *Pure appl. Geophys.,* **177,** 2255–2279.

ECS ASN, 2011. Evaluations Complémentaires de Sûreté, rapport de l'Autorité de Sûreté, Tech. rep., Nuclear Authority Safety website.

Faccioli, E., Maggio, F., Paolucci, R. & Quarteroni, A., 1997. 2D and 3D elastic wave propagation by a pseudo-spectral domain decomposition method, *J. Seismol.,* **1**(3), 237–251.

Festa, G. & Vilotte, J.-P., 2005. The newmark scheme as velocity–stress time-staggering: an efficient pml implementation for spectral element simulations of elastodynamics, *Geophys. J. Int.,* **161**(3), 789–812.

Field, E.H. & Jacob, K.H., 1995. A comparison and test of various site-response estimation techniques, including three that are not reference-site dependent, *Bull. seism. Soc. Am.,* **85**(4), 1127–1143.

Fu, H. *et al.*, 2017. 18.9Pflopss Nonlinear Earthquake Simulation on Sunway TaihuLight: Enabling Depiction of 18-Hz and 8-meter Scenarios, in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, SC '17,* pp. 2:1–2:12, ACM, New York, NY, USA.

Garofalo, F. *et al.*, 2016a. InterPACIFIC project: Comparison of invasive and non-invasive methods for seismic site characterization. Part II: Inter-comparison between surface-wave and borehole methods, *Soil Dyn. Earthq. Eng.,* **82,** 241–254.

Garofalo, F. *et al.*, 2016b. InterPACIFIC project: Comparison of invasive and non-invasive methods for seismic site characterization. Part I: Intra-comparison of surface wave methods, *Soil Dyn. Earthq. Eng.,* **82,** 222–240.

Guyonnet-Benaize, C., 2011. Modélisation 3D multi-échelle des structures géologiques de la région de la faille de la moyenne Durance (SE France), *PhD thesis*, Ecole Doctorale Sciences de l'Environnement (Aix-en-Provence).

Guyonnet-Benaize, C., Lamarche, J., Hollender, F., Viseur, S., Münch, P. & Borgomano, J., 2015. Three-dimensional structural modeling of an active fault zone based on complex outcrop and subsurface data: the Middle Durance Fault Zone inherited from polyphase Meso-Cenozoic tectonics (southeastern France), *Tectonics,* **34**(2), 265–289.

Hisada, Y., 2000. A theoretical omega-square model considering the spatial variation in slip and rupture velocity, *Bull. seism. Soc. Am.,* **90**(2), 387–400.

Hollender, F. *et al.*, 2009. Deep geometry of the Middle Durance Fault system (SE of France): reprocessing and new interpretation of reflection seismic sections, in *International Conference Provence,* pp. 6–8, Aix-en-Provence (France).

Hollender, F., Andre, M., Guyonnet-Benaize, C., Cornou, C., Caillot, V. & Bard, P.-Y., 2011. Can high daily-variation of noise level alter results of ambient vibration H/V technique?, in *Conference: ESG4-4th IASPEI / IAEE International Symposium on Effects of Surface Geology on Seismic Motion,* pp. 1–6, University of California Santa Barbara.

IAEA, 2003. *Seismic Design and Qualification for Nuclear Power Plants, no. NS-G-1.6 in Specific Safety Guides,* International Atomic Energy Agency.

IAEA, 2020. *Methodologies for Seismic Safety Evaluation of Existing Nuclear Installations, no. 103 in Safety Reports Series,* International Atomic Energy Agency.

Ichimura, T. *et al.*, 2018. A fast scalable implicit solver with concentrated computation for nonlinear time-evolution problems on low-order unstructured finite elements, in *2018 IEEE International Parallel and Distributed Processing Symposium (IPDPS),* pp. 620–629, IEEE.

Khan, S., Muhammad, S., *et al.*, 2020. Scenario-based seismic hazard analysis using spectral element method in northeastern Pakistan, *Nat. Hazards,* **103**(2), 2131–2144.

Komatitsch, D. & Vilotte, J.-P., 1998. The spectral element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures, *Bull. seism. Soc. Am.,* **88**(2), 368–392.

Kotha, S.R., Weatherill, G., Bindi, D. & Cotton, F., 2020. A regionally-adaptable ground-motion model for shallow crustal earthquakes in Europe, *Bull. Earthq. Eng.,* **18**(9), 4091–4125.

Mai, P.M., 2005. Hypocenter locations in finite-source rupture models, *Bull. seism. Soc. Am.,* **95**(3), 965–980.

Manchuel, K., Traversa, P., Baumont, D., Cara, M., Nayman, E. & Durouchoux, C., 2018. The French seismic CATalogue (FCAT-17), *Bull. Earthq. Eng.,* **16**(6), 2227–2251.

Maufroy, E. *et al.*, 2015. Earthquake Ground Motion in the Mygdonian Basin, Greece: The E2VP Verification and Validation of 3D Numerical Simulation up to 4 Hz, *Bull. seism. Soc. Am.,* **105**(3), 1398–1418.

Maufroy, E. *et al.*, 2016. 3D numerical simulation and ground motion prediction: Verification, validation and beyond-Lessons from the E2VP project, *Soil Dyn. Earthq. Eng.,* **91,** 53–71.

McCallen, D., Petrone, F., Miah, M., Pitarka, A., Rodgers, A. & Abrahamson, N., 2021. EQSIM-A multidisciplinary framework for fault-to-structure earthquake simulations on exascale computers, part II: Regional simulations of building response, *Earthq. Spectra,* **37**(2), 736–761.

McCallen, D. *et al.*, 2021. EQSIM-A multidisciplinary framework for fault-to-structure earthquake simulations on exascale computers part I: Computational models and workflow, *Earthq. Spectra,* **37**(2), 707–735.

Milner, K.R., Shaw, B.E., Goulet, C.A., Richards-Dinger, K.B., Callaghan, S., Jordan, T.H., Dieterich, J.H. & Field, E.H., 2021. Toward Physics-Based Nonergodic PSHA: A Prototype Fully Deterministic Seismic Hazard Model for Southern California, *Bull. seism. Soc. Am.***111**(2), 898–915.

Olsen, K.B. & Mayhew, J.E., 2010. Goodness-of-fit criteria for broadband synthetic seismograms, with application to the 2008 Mw 5.4 Chino Hills, California, earthquake, *Seismol. Res. Lett.,* **81**(5), 715–723.

Perron, V., Gélis, C., Froment, B., Hollender, F., Bard, P.-Y., Cultrera, G. & Cushing, E.M., 2018. Can broad-band earthquake site responses be predicted by the ambient noise spectral ratio? Insight from observations at two sedimentary basins, *Geophys. J. Int.,* **215**(2), 1442–1454.

RFS-2001-01, 2001. *French Nuclear Safety Authority,* Nuclear Authority Safety website https://www.asn.fr/content/download/53897/367951/version/1/.../RFS-2001-01.pdf.

Schmedes, J., Archuleta, R.J. & Lavallée, D., 2010. Correlation of earthquake source parameters inferred from dynamic rupture simulations, *J. geophys. Res.,* **115**(B03304), 1–12.

Song, S.G., 2016. Developing a generalized pseudo-dynamic source model of Mw 6.5-7.0 to simulate strong ground motions, *Geophys. J. Int.,* **204**(2), 1254–1265.

Stupazzini, M., Infantino, M., Allmann, A. & Paolucci, R., 2021. Physics-based probabilistic seismic hazard and loss assessment in large urban areas: A simplified application to Istanbul, *Earthq. Eng. Struct. Dyn.,* **50**(1), 99–115.

Touhami, S., Lopez-Caballero, F. & Clouteau, D., 2021. A holistic approach of numerical analysis of the geology effects on ground motion prediction: Argostoli site test, *J. Seismol.,* **25,** 115–140, doi:10.1007/s10950-020-09961-0.

Touhami, S., Gatti, F., Lopez-Caballero, F., Cottereau, R., de Abreu Corrêa, L., Aubry, L. & Clouteau, D., 2022. SEM3D: a 3D high-fidelity numerical earthquake simulator for broadband (0-10 Hz) seismic response prediction at a regional scale, *Geosciences,* **12**(3), https://www.mdpi.com/2076-3263/12/3/112.

Wang, X.-C., Wang, J.-T., Zhang, L. & He, C.-H., 2021. Broadband ground-motion simulations by coupling regional velocity structures with the geophysical information of specific sites, *Soil Dyn. Earthq. Eng.,* **145,** 106695.

APPENDIX

A.1 Modified Anderson's criteria

The GOFs system defined by Anderson (2004) is listed in the Table 3.

$d_i(t)$, $v_i(t)$ and $a_i(t)$ represent the displacement, velocity and acceleration time histories of the ith signal and with:

$$S(p_1, p_2) = 10 \cdot e^{-\left(\frac{p_1 - p_2}{\min(p_1, p_2)}\right)^2} \tag{A1}$$

f represents the frequency and T_n represents the natural period. $T_{d,i}$ is the duration, FS_i is the acceleration Fourier's spectrum, Sa_i is the pseudospectral acceleration, $I_{A,i}(t)$ is the Arias intensity of the ith signal, respectively. $I_{E,i}$ is the integral square velocity of the ith signal respectively. Finally, C represents the cross-correlation and $\mathbb{E}$ represents the empirical average. C9 defined by Anderson (2004) reads:

$$\begin{aligned} \mathrm{C9}_{\mathrm{original}} &= \mathbb{E}_f \left[S\left(F_{S,1}(f), F_{S,2}(f)\right)\right] \\ &= \frac{1}{n_f} \sum_{i=1}^{n_f} \left[S\left(F_{S,1}(f_i), F_{S,2}(f_i)\right)\right] \end{aligned} \tag{A2}$$

In other words, the C9 GOF is measured individually and the final score is set to the average of these values. However, adopting the version in eq. (A2), the difference among different cases cannot be captured due to the general large variance in the Fourier spectrum (Olsen & Mayhew 2010). To deal with this issue, the Fourier spectrum is smoothed by averaging the values across a specific bandwidth in the GOF method proposed by Olsen & Mayhew (2010):

$$C9_{\mathrm{modified}} = S\left(\frac{1}{n_{fu} - n_{fl}} \sum_{i=n_{fl}}^{n_{fu}} F_{S,1}(f_i), \frac{1}{n_{fu} - n_{fl}} \sum_{i=n_{fl}}^{n_{fu}} F_{S,2}(f_i) \right) \tag{A3}$$

where $f_{n_{fl}}$ corresponds to the lower limit of the considered frequency band and $f_{n_{fu}}$ is the upper limit of the considered frequency band.

A.2 Performance of the numerical analysis

Given the scalability properties of `SEM3D`[2], each run (30 s earthquake simulation, on the template model) took approximately 80 min (≈960 hr CPU-time) on 720 MPI cores Intel Xeon Gold 6230 20C @2.1 GHz Cascade Lake. To synthesize a 20 long earthquake event, an `SEM3D` run of the described numerical model took ≈10 hr (wall-time) on 4000 MPI processes deployed on 200 nodes Bull B720 Haswell 12 cores 2.6 Ghz.

This work was granted access to the HPC resources of CINES under the allocation 2018-A0040410444, 2019-A0060410444, 2020-A0070411083 and 2020-A0080410444, made by GENCI (Grand équipement national de calcul intensif). Computations were also performed using HPC resources allocated by the Mésocentre Moulon, the supercomputer of CentraleSupélec and École Normale Supérieure de Paris-Saclay, Paris-Saclay University.

[2] https://www.researchgate.net/publication/349254101_SEM3D-High-resolution_seismic_wave_propagation_modelling_from_the_fault_to_the_structure_for_realistic_earthquake_scenarios_GENCI_Allocation_A0080410444/stats