

The influence of source- and ground-motion model choices on probabilistic seismic hazard levels at 6 sites in France

Céline Beauval⁽¹⁾, Pierre-Yves Bard⁽¹⁾, Laurentiu Danciu⁽²⁾

Affiliations

⁽¹⁾ ISTERre, IRD, UGA, CNRS, CS 40700 - 38058 Grenoble Cedex 9, France

⁽²⁾ Department of Earth Science, ETH Zurich, Sonneggstrasse 5, 8092 Zurich, Switzerland

Corresponding author: C. Beauval

celine.beauval@univ-grenoble-alpes.fr

Tel: +33 (0)4 76 63 52 20

Abstract

In this study focused on France, we explore the uncertainties related to choices made while building a source model for hazard assessment and we quantify the impact on probabilistic hazard estimates. Earthquake recurrence models are initially built from the French Seismic CATalog (FCAT, Manchuel et al. 2018). We set up a logic tree that includes two alternative seismogenic source models (ESHM13 and Baize et al. 2013), two versions of FCAT catalog, two alternative declustering algorithms, and three alternative minimum magnitudes for earthquake recurrence modeling. We calculate the hazard for six cities (i.e. Nantes, Lourdes, Clermont-Ferrand, Briançon, Nice and Strasbourg) that are located in source zones with a minimum amount of data to work with. Results are displayed for the PGA and spectral period 0.2s, at return periods 475 and 5000 yrs. Exploration of the logic tree shows that the parameters with the most impact on hazard results are the minimum magnitude used in the recurrence modeling (up to 31%) and the selection of the seismogenic source model (up to 30%). We also use the SHARE European Earthquake Catalog (SHEEC, Woessner et al. 2015) to build earthquake recurrence models and compare hazard values obtained with the FCAT logic tree. Comparisons are limited because of the low number of events available in some sources in SHEEC ; however, results show that, depending on the site considered, the earthquake catalog selection can also strongly impact the hazard estimates (up to 50%). The FCAT logic tree is combined with four ground-motion models (Bindi et al. 2014, Boore et al. 2014, Cauzzi et al. 2015 and Drouet and Cotton 2015) to account for the epistemic uncertainty on the prediction of ground-motion. Exploration of the logic tree shows that the contribution of ground-motion model uncertainties can be larger than, equivalent to, or lower than the contribution of the source-model uncertainties to the overall hazard variability. Which component controls overall uncertainty depends on the site, spectral period and return period. Finally, exploring the logic tree provides a distribution for the ratios between hazard levels at 5000 and 475 years return periods, revealing that the ratios only slightly depend on source-model uncertainties, vary strongly from site to site, and can take values between 3 and 5, which is significantly higher than what is commonly assumed in the engineering community.

Keywords: Earthquake recurrence, probabilistic seismic hazard, uncertainties, France

1. Introduction

In a probabilistic seismic hazard study, a source model must be built for the region of interest that reflects the seismic potential of the region and must be combined with a ground-motion model (GMM) that is able to predict the ground motions produced by future earthquakes. The aim is to determine the probabilities of exceedance over future time windows for ground-motion levels of interest in earthquake engineering. Identification and quantification of uncertainties characterizing each step in the procedure is mandatory in order to understand the precision of the output given the current state-of-knowledge and to highlight which parameters control the hazard and would deserve more attention in the future.

One of the work packages of the SERA project (INFRAIA-01-2016-2017, <http://www.sera-eu.org>) aims to update the European seismic-hazard model. Both the source model and the ground-motion prediction models are currently being re-built, based on new data and research developments. The previous European seismic-hazard model published (ESHM13, Woessner et al. 2015) integrated models for France that are now superseded in terms of the delineation of seismogenic sources. Moreover, two earthquake catalogs have been published, one instrumental by Cara et al. (2015, SiHex) and one historical by Traversa et al. (2018), merged in the FCAT catalog (Manchuel et al. 2018). When building a source model for France, choices must be made regarding the data that are available to feed the earthquake recurrence model; these decisions may impact the final hazard estimates. To compare source-model uncertainties with ground-motion model uncertainties, the uncertainty related to the choice of the ground-motion model is also addressed. The present work is performed within the SERA project to gain experience on hazard estimates and associated uncertainties for France; however, this work is distinct from the exact procedure implemented for the update of the European seismic-hazard model that will be fully detailed in a future extended report.

2. Input data for building a source model for France: earthquake catalogs

The earthquake catalog is a key input when building a seismogenic source model within a probabilistic hazard assessment framework. The catalog must cover instrumental and historical periods to best possibly represent the seismic potential in the study region. In France, historical catalogs can extend up to the 5th century. The SHARE European Earthquake Catalog (SHEEC; Stucchi et al., 2012; Fig. 1) developed for ESHM13 and the French Seismic CAtalog (FCAT; Manchuel et al., 2018; Figs. 1 and 2) are those that are currently available. Although we provide a short description on the content of these catalogs, we encourage the reader to study the corresponding publications in detail. These two catalogs can be used to estimate seismic hazard; they represent two alternative ways of providing locations and magnitudes to instrumental and historical events in France. Manchuel et al. (2018, their figure 11) and Basili et al. (2018, pp 33-36) compare magnitude estimates in the FCAT and SHEEC in the historical period.

2.1 SHEEC catalog

The 1900 to 2006 time window is derived from the European-Mediterranean Earthquake Catalog (EMEC) (Grünthal and Wahlström, 2012, pp 542). The Laboratoire de Détection et de Géophysique (LDG) solutions are used from 1962 to 2004. A proxy for M_w is estimated by successively applying two conversion equations: the original M_L LDG is converted into a “Central European” M_L magnitude, which is then converted into an M_w with a dedicated conversion equation. From 1900 to 1961, moment magnitudes are estimated from epicentral intensity by applying an equation based on the “Central European” events.

The 1000 to 1899 time period was compiled at Istituto Nazionale di Geofisica e Vulcanologia (INGV) and described in Stucchi et al. (2012). Earthquake parameters were derived from the SisFrance macroseismic database by applying the Boxer method (Gasperini et al., 1999, 2010). One version of the Institut de Radioprotection et Sûreté Nucléaire (IRSN) catalog was also used (v. 1.1, 2011, D. Baumont and O. Scotti). For earthquakes with an IRSN solution, the equivalent moment magnitude of events corresponds to the weighted mean of the Boxer M_w and the IRSN M_w . Few events with magnitude (M_w) lower than 4.0 have been included in the catalog.

2.2 FCAT catalog

The FCAT catalog has a historical period that was built in the SIGMA project by Electricité de France (EDF) and Geoter companies (Traversa et al. 2018; Baumont et al. 2018) and of an instrumental period that was built in the SiHex project by a consortium of French seismological networks and observatories (Cara et al. 2015). These catalogs were merged into the FCAT catalogue (Manchuel et al. 2018); the instrumental period (1965-2009) has been appended to the historical period (463-1964). The SiHex catalog was built to provide the best possible solution for location and magnitude for each event (Cara et al. 2015). The catalog reports moment magnitude (M_w)

computed from coda-wave analysis for most events with M_L LDG > 4.0 ($\sim M_W > 3.4$). For events without an M_W (coda) estimation, a proxy M_W is inferred from M_L LDG using several conversions according to the magnitude range and the period of time. In some cases, magnitudes $M_W \leq 3.4$ are inferred from regional local magnitudes or from the duration magnitude of LDG. More details on the magnitude estimates in SiHex can be found in Denieul et al. (2015) and Cara et al. (2017). Laurendeau et al. (2019) analyzed the methods applied to obtain M_W proxies and discussed the uncertainties associated.

For the historical part, described in two companion papers (Traversa et al. 2018; Baumont et al. 2018), the method applied to determine the magnitude depends on the macroseismic data available (SisFrance). The epicentral locations were not revised in this work, they remain those of SisFrance. For well-described events, magnitudes and depths are jointly inverted through exploring a tree of intensity prediction equations (IPE). For events with a poor macroseismic dataset, magnitudes are estimated assuming an *a priori* depth. Overall, for 47% of the events in the final catalog, the macroseismic field is used to infer magnitude and depth (inversion strategy 3 in Traversa et al., 2018), for 37% of the events the magnitude is inferred from the epicentral intensity only, and for 16% of the events the magnitude is deduced from intensity levels inferred from felt testimonies only. These percentages do not account for FCAT historical events with $M_W < 3.5$. We believe that these events should not be used as their magnitudes are outside the IPE validity domain. In this study, we will test the inclusion of earthquakes that have extremely poorly constrained magnitudes deduced from felt testimonies and their impact on earthquake recurrence modeling.

2.3 FCAT behavior in time

Some unexpected tendencies are observed when the FCAT behavior is analyzed in time (Figs. 3-5). Figure 3a displays the cumulative number of events versus time for different magnitude intervals. Events with M_W 3.5-4.5 are distributed homogeneously over time from 1860 to 1965, but the number drops suddenly after 1965. Although the reasons for such discrepancies are not straightforward, the historical and instrumental periods in FCAT might not be homogeneous in terms of magnitude estimate. Figure 3b displays the cumulative number of events in FCAT and SHEEC over the interval M_W 4.1-4.4. Identical features can be observed for some individual source zones when analyzing the subcatalogs (Figs. 4 and 5 show examples for two seismogenic sources—in Brittany and in the Alps); there is a depleted number of earthquakes in the instrumental period with respect to the historical period. This tendency can also be observed in SHEEC but is less striking.

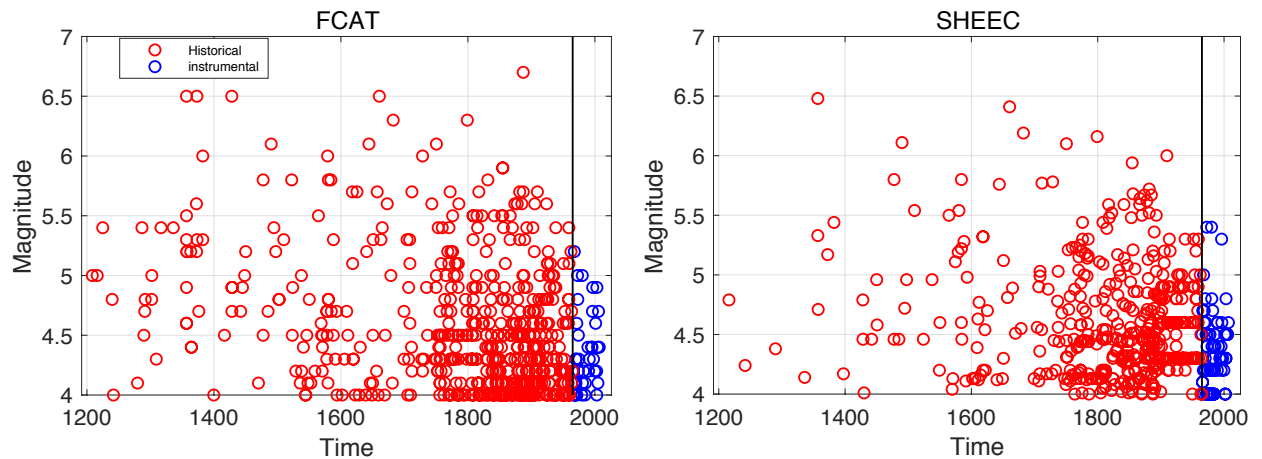


Figure 1 : Content of FCAT and SHEEC: magnitude versus time for events in the historical period (red) and in the instrumental period (blue). The black vertical line corresponds to 1965, which is when the FCAT historical part was appended to the SiHex instrumental catalog. The spatial window is the same for both catalogs, seen in the polygon enclosing Metropolitan French borders in Figure 2. FCAT extends to 2009, and SHEEC extends to 2006

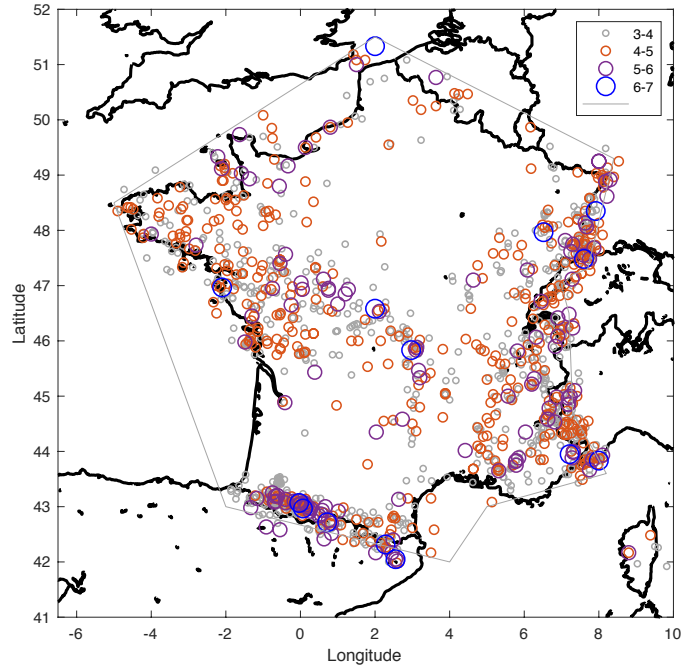


Figure 2 : FCAT: 1300-2009, declustered with the Reasenberg algorithm. The polygon enclosing Metropolitan French borders is used to evaluate earthquake recurrence at the country level (section 4.3).

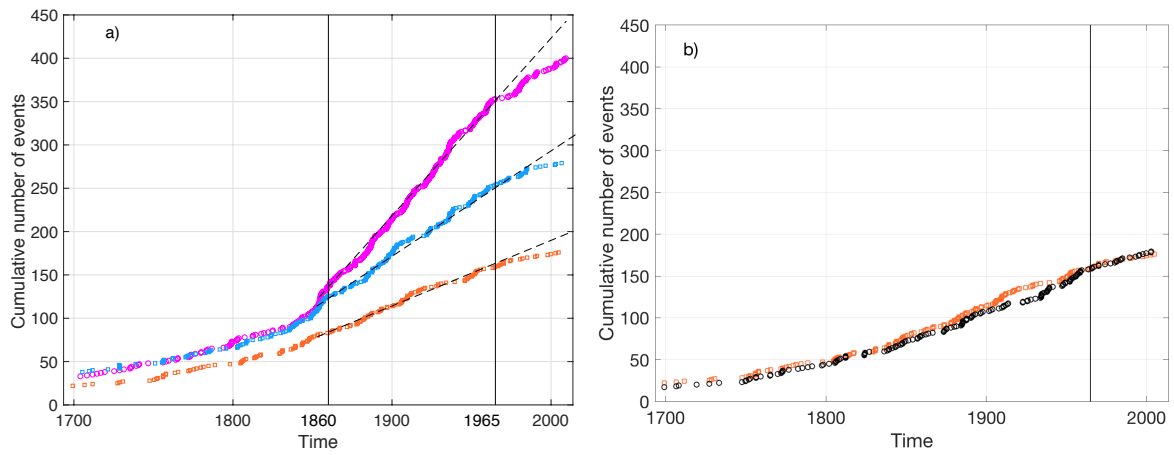


Figure 3 : Cumulative number of events versus time. (a) FCAT: $3.5 \leq M_W < 3.8$ (magenta), $3.8 \leq M_W < 4.1$ (blue), and $4.1 \leq M_W < 4.4$ (orange). (b) SHEEC: $4.1 \leq M_W < 4.4$ (black) superimposed to FCAT (orange). The spatial window is seen as a polygon enclosing Metropolitan French borders in Figure 2. The declustered catalogs are used.

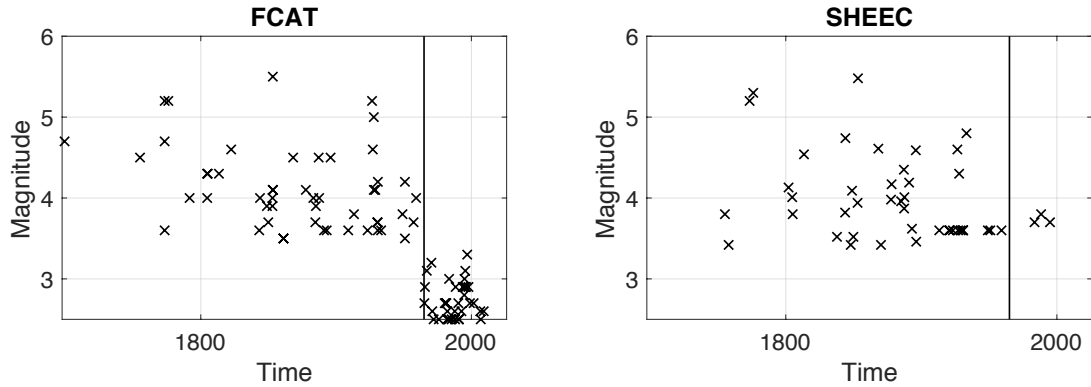


Figure 4 : FCAT and SHEEC subcatalogs for the Baize et al. (2013) source zone in Northern Brittany including the Cotentin peninsula (zones V-06 and V-07 in Baize et al., 2013; Figure 6). The black vertical line corresponds to 1965, which is when the FCAT historical part is appended to the SiHex instrumental catalog.

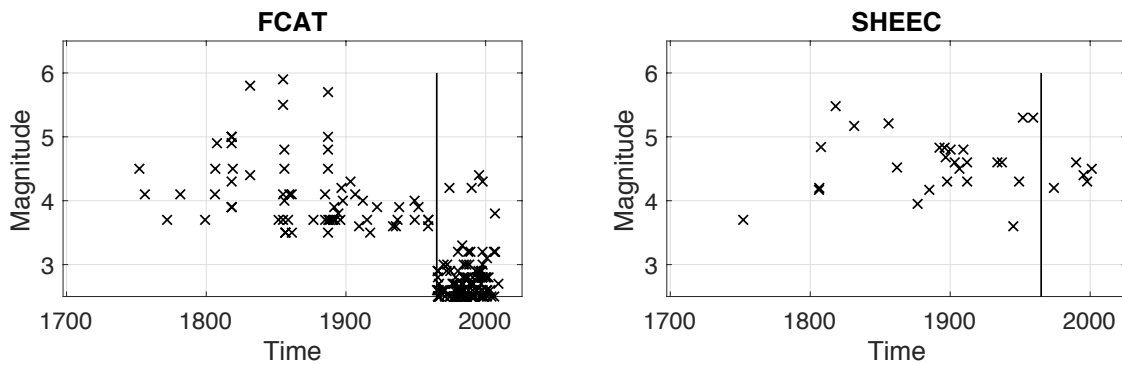


Figure 5 : FCAT and SHEEC subcatalogs for the Baize et al. (2013) source zone in the Southern Alps extending from Castellane to Nice (zones O-17, O-18, and O-21 in Baize et al. 2013; Figure 6). The black vertical line corresponds to 1965, which is when the FCAT historical part is appended to the SiHex instrumental catalog.

3. Input data for building an area source model for France: seismogenic source models

The seismogenic source model delineates area sources in space that will be considered homogeneous for earthquake recurrence modeling. This model results from a combined analysis of structural and rheological properties of the crust, geophysical data, and more dynamic data that includes geodynamics, seismicity, and neotectonics in the region of interest. The authors must hierarchize the criteria to be used in the delineation of sources; the final boundaries reflect their understanding of the seismotectonics of the region.

The seismogenic source model used for France in ESHM13 (Fig. 6a) relies on source zone limits derived by the EPAS group (Autran et al. 1998) for inner France and on source zone limits proposed by neighboring countries for sources on the border (i.e., United Kingdom, Belgium, Germany, Switzerland, Italy, Spain). This model is available and can be used for a probabilistic seismic hazard assessment (PSHA) study.

Baize et al. (2013) published a new seismogenic model (IRSN model) that proposes an update of Autran et al. (1999) by considering new data with corresponding new interpretations in the field of geodynamics, seismology, and active tectonics. The criteria used to delineate source zones are described in detail in Baize et al. (2013). Two versions are proposed to consider the uncertainty on source zone limits: a model with 66 source zones and an aggregated model with 38 source zones, which is considered here (Fig. 6b).

We make some modifications for two source zones. In the Western Pyrenees, source zone O-03 is characterized by a specific high level of seismicity in the instrumental and historical periods. We believe this seismicity should not be distributed inside source zone O-02; therefore, we maintain the boundary between source zones O-02 and

O-03 (boundary level 3, fig. 2 in Baize et al. 2013). A boundary is added in order to split the source zone that encloses Strasbourg (R-04, Upper Rhine graben, Baize et al. 2013) into a southern and northern zone. This horizontal boundary corresponds to the spatial coverage of FCAT. The FCAT delivered in the supplementary material of Manchuel et al. (2018) includes events up to 20 km and up to 40 km from the border for the instrumental and historical periods, respectively.

Because we are interested in understanding how the choices made while building the source model may impact hazard estimates, we consider both the ESHM13 and IRSN models as alternatives for the present hazard study.

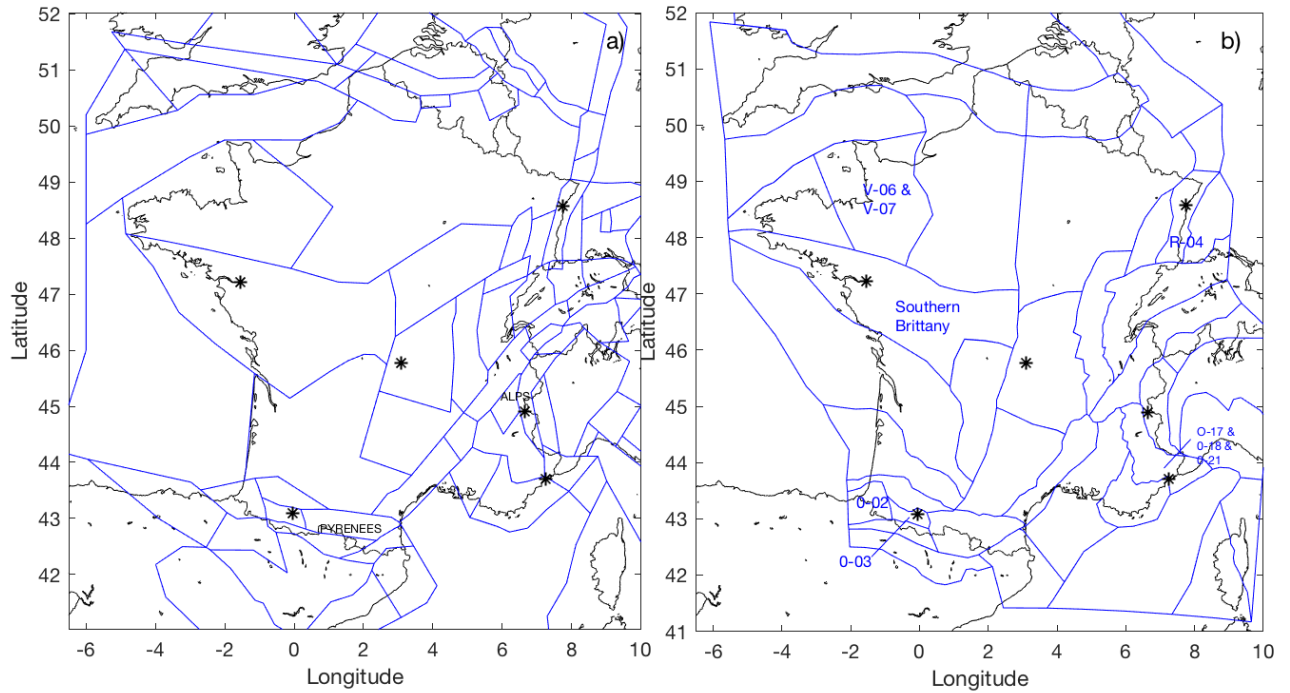


Figure 6 : Seismogenic source models used in the present study. (a) ESHM13 model (Woessner et al. 2015); (b) IRSN model (aggregated version, see the text for modifications in source zones R-04 and O-02/O-03) (Baize et al., 2013). Probabilistic seismic hazard calculations are performed for the six selected sites (stars) in Section 5.

4. Modeling earthquake recurrence

4.1 Declustering step

Earthquake catalogs are usually declustered before earthquake recurrence is modeled. Within the hazard calculation, an occurrence model in time must be assumed for earthquakes. Seismicity is rather diffuse in low-to-moderate seismicity regions, and few events can be associated to faults. The Poisson occurrence model is usually assumed for earthquakes, implying aleatory occurrences of events in time and space. Because the aim is to establish long-term recurrence models not affected by short-term fluctuations, clustered events should be discarded; however, we will show that this step has a negligible impact on recurrence models for regions of low-to-moderate seismicity.

There are several declustering algorithms available to identify clustered events. The Gardner and Knopoff (1974, GK) algorithm relies on simple magnitude-dependent windows in time and space that are used in most hazard studies by the United States Geological Survey (e.g., Petersen et al. 2014). The Reasenberg (1985, RE) algorithm relies on different assumptions with clusters linked by smaller events that are allowed to grow in time and space (see, e.g., Christophersen and Smith 2008). The time interaction window is based on the Omori decay, whereas the spatial interaction zone is based on the magnitude of prior events. We apply both algorithms to the $M_w \geq 2.0$ events of the FCAT, the GK algorithm with windows proposed by Burkhard and Grünthal (2009), and the RE algorithm with the original parameters and errors on location accounted for. The Reasenberg algorithm identifies around 20% of clustered events. The Gardner and Knopoff simple window method identifies around 33% clustered

events ($M_w \geq 2.0$). Considering the instrumental period of FCAT (SiHex catalog, 1965-2009), the impact on the cumulative number of events with magnitude larger or equal to M_w is displayed in Figure 7. As expected, the proportion of identified clustered events decreases with increasing magnitude. For the recurrence modeling magnitude range of interest (e.g., $M_w \geq 3.0$) the difference in obtained annual rates is small. Although the choice of declustering algorithm is thus likely to have a limited impact on hazard, it will be tested for the FCAT catalog. We also note a break in the exponential decrease around magnitude 3.4: magnitudes of events $M_w \leq 3.4$ result from a conversion equation and are thus not strictly homogeneous with magnitudes of events $M_w > 3.4$.

For SHEEC, we only use the mainshocks identified during the SHARE project using the GK algorithm and the Burkhard and Grünthal (2009) window parameters (Woessner et al. 2015).

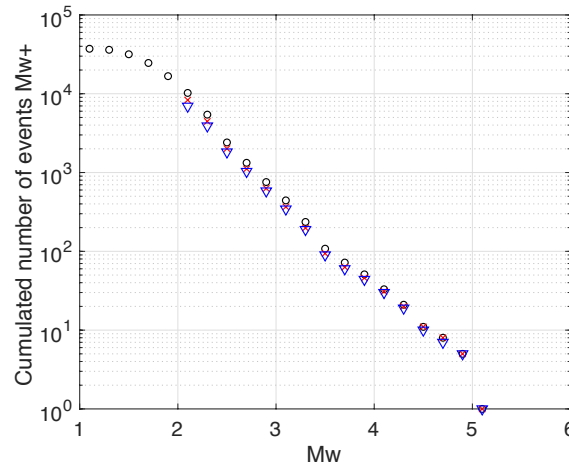


Figure 7: Magnitude-frequency distribution considering the SiHex catalog (1965-2009, instrumental part of FCAT). Black symbols: full catalog; red crosses: catalog declustered with Reasenberg algorithm; blue triangles: catalog declustered with Gardner & Knopoff windows in time and space.

4.2 Time windows of completeness (Table 1)

Time windows of completeness are determined from the cumulative number of events versus time plots (visual inspection). The time windows are estimated considering the entire country. We also perform the analysis on a regional scale (e.g., Alps, Pyrenees) to ensure the time windows are approximately valid. In the low-to-moderate magnitude range (3.2-5.0) identifying the time windows is fairly straightforward; however, for larger magnitudes, identifying the time windows is more difficult because of the restricted number of events. Other methods, independent from the data, should be tested to evaluate completeness, including estimating the detection capacity of a seismological network (Schorlemmer and Woessner 2008) or performing an in-depth historical analysis (Stucchi et al. 2004). However, such studies do not yet exist for France. The uncertainty on the determination of time windows of completeness is not explored here.

Table 1 : Time periods of completeness – SHEEC and FCAT. FCAT declustered with Reasenberg algorithm.

Magnitude interval	3.2-3.5	3.5-3.8	3.8-4.1	4.1-4.4	4.4-4.7	4.7-5.0	5.0-5.3	5.3-5.6	5.6-5.9	5.9-6.2	6.2-6.5	6.5-6.8
Time period - SHEEC	-	-	-	1800	1800	1800	1750	1500	1300	1300	1300	1300
Time period - FCAT	1965	1860	1850	1830	1830	1750	1750	1750	1600	1310	1310	1310

4.3 Modeling earthquake recurrence at the country level

At first, earthquake recurrence is modeled considering the entirety of France. Accounting for events within the complete time windows, observed annual exceedance rates are estimated from the FCAT catalog for magnitudes $M_w \geq 3.2$. We do not include magnitudes below 3.2 for various reasons:

- The cumulative number of events versus time plots for magnitude intervals $2.6 \leq M_w < 2.9$ and $2.9 \leq M_w < 3.2$ display unexpected features, indicating that M_w might not be homogeneous over time within these intervals.
- With low magnitudes, it is more difficult to discriminate events of tectonic origin from events of anthropogenic origin.

- Below M_w 3.4, most instrumental SiHex magnitudes are obtained by converting M_L LDG magnitudes (as well as, for some events, regional local magnitudes or duration LDG magnitude). Two distinct slopes can be observed in Fig. 7, indicating that the magnitude range below 3.4 is not homogeneous with $M_w > 3.4$. Above M_w 3.4, the instrumental dataset is rather homogeneous, although only a fraction of the events have a magnitude determined from coda waves.

We use a magnitude bin equal to 0.3. The magnitude bin width should be approximately larger than or equal to the uncertainty on the magnitude of events; 0.3 appears to be a reasonable compromise considering the uncertainty level on the instrumental magnitudes (around 0.1-0.2) and the uncertainty level on the historical magnitudes, which is strongly event dependent and can reach values as high as 0.4-0.5.

The earthquake recurrence model considered here is form 2 in Anderson and Luco (1983):

$$N(m) = 10^{a-bm} - 10^{a-bM_{max}} \quad \text{for magnitude } m \leq M_{max} \quad (\text{eq. 1})$$

Considering the FCAT catalog declustered with the Reasenberg algorithm (Fig. 8), the logarithm of the annual number of events decreases linearly with increasing magnitude over the interval 3.2-5.6 (Gutenberg-Richter 1944 model). For larger magnitudes, the annual rates are more unstable. Weichert's (1980) method yields a b-value of 0.82 ($\Delta b = 0.02$) using 879 events with $M_w \geq 3.2$. At such a large spatial scale, a b-value within the interval 0.9-1.0 would be expected. One explanation might be the heterogeneity between the historical and the instrumental period, with fewer events in the magnitude range 3.5 - 4.5 in the instrumental time window (1965-2009) than in the corresponding complete historical time window (1850-1965). Numbers of events per magnitude interval are reported in Table 2.

Considering the SHEEC catalog (Fig. 8), the magnitude threshold is higher, and a narrower magnitude range can be used. The logarithm of the annual number of events decreases linearly with increasing magnitude over the interval 4.1-5.3. For larger magnitudes the annual rates are lower than expected from the extrapolation of the lower magnitude range. Weichert's (1980) method yields a b-value of 0.93 ($\Delta b = 0.04$) using 325 events with $M_w \geq 4.1$. Numbers of events per magnitude interval are reported in Table 3. At the country level, over the magnitude interval 4-5, the rates based on SHEEC are similar to the rates based on FCAT. This is not always the case at the source zone levels (see Section 4.4).

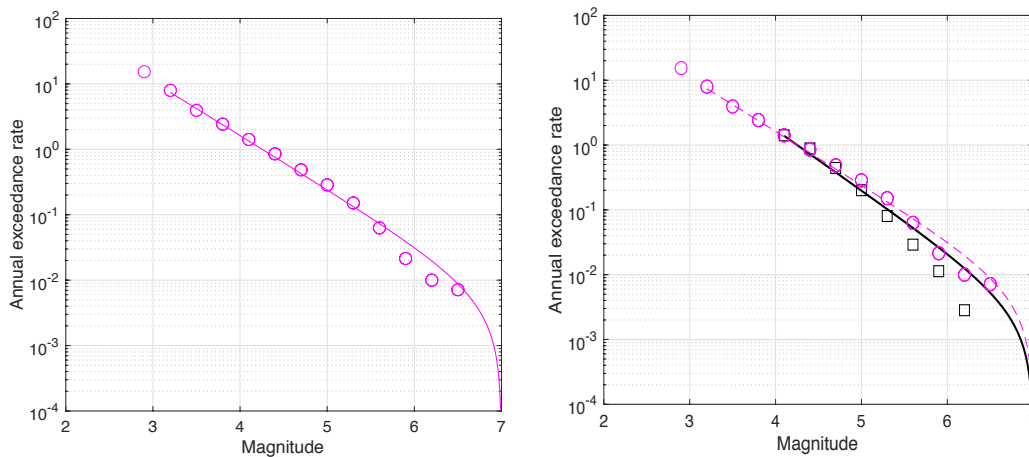


Figure 8: Magnitude-frequency distributions for FCAT (magenta) and SHEEC (black), considering the completeness time windows in Table 1.

Table 2: Number of events used to determine observed rates, per magnitude interval and within time periods of completeness – SHEEC

Minimum bound of magnitude interval	4.1	4.4	4.7	5.0	5.3	5.6	5.9	6.2
# of events	109	92	50	31	26	9	6	2
Time period (years)	207	207	207	257	507	507	707	707

Table 3: Number of events used to determine observed rates, per magnitude interval and within time periods of completeness – FCAT (Reasenbergl declustering)

Minimum bound of magnitude interval	3.2	3.5	3.8	4.1	4.4	4.7	5.0	5.3	5.6	5.9	6.2	6.5
# of events	181	227	162	101	66	52	35	23	17	8	2	5
Time period (years)	45	150	160	180	180	260	260	260	410	700	700	700

4.4 Modeling earthquake recurrence at the level of the source zones

Earthquake recurrence must be modeled from past seismicity within each seismic source zone. The number of events to use varies among sources. France is a low-to-moderate seismicity country and many sources contain few earthquakes. Because the minimum magnitude of completeness of SHEEC is high, the earthquake recurrence models inferred from SHEEC are more poorly constrained than the models inferred from FCAT. The criteria followed to derive earthquake recurrence parameters a and b from the source sub-catalogs are below:

- Because the minimum magnitude chosen may impact the recurrence models, three alternative minimum magnitudes are used for the FCAT catalog (3.2, 3.5, 3.8).
- For SHEEC, there are too few events to evaluate the impact of the minimum magnitude; the minimum magnitude used is 4.1 or 4.4 depending on the source zone (4.4 is used if the number of events in the interval $4.1 \leq M_w < 4.4$ is lower than the number of events in the interval $4.4 \leq M_w < 4.7$).
- If there are at least 15 events within the completeness time windows (and above the minimum magnitude considered), a and b-values are provided by the Weichert’s method (1980).
- If there are less than 15 events within the completeness time windows, the b-value estimated at the country level (called “regional” b-value) is applied, and the a-value is calculated from the mean of the rates of the two lowest magnitude bins.

Considering two alternative seismogenic source models and two alternative earthquake catalogs, these criteria are applied systematically for recurrence modeling in all seismogenic source zones. Figure 9 displays the recurrence models obtained for six example zones of the IRSN seismogenic model (host zones for the six sites considered in the hazard calculations [Section 5]). The number of events used and recurrence parameters obtained are reported in Tables 4 (SHEEC) and 5(FCAT). The earthquake recurrence models established from FCAT are rather well-constrained in these zones, with the logarithm of the annual exceedance rates linearly decreasing over an interval that extends over at least two magnitude degrees. Recurrence models based on SHEEC are established from a more restricted magnitude interval and are thus less constrained, except for the source zone hosting the city of Nantes (Southern Brittany). The recurrence models obtained for the source zones that are hosting the six sites considering the ESHM13 seismogenic source model are displayed in the Appendix (Fig. S1).

The recurrence models are well-constrained in the southern Brittany source zone (Nantes host zone). The model relies on 28 events in the case of SHEEC and 71 to 117 events in the case of FCAT depending on the minimum magnitude used. Varying the minimum magnitude for FCAT has only a minor impact on the model. FCAT leads to much higher rates than SHEEC (factor 2). Because the models are well-constrained, the differences obtained between both catalogs can be related to the different ways both catalogs have been built, particularly the moment magnitudes estimated for each earthquake. The b-value inferred from FCAT is rather low (0.75 to 0.86), which is representative of most b-values estimated from FCAT throughout France. Similar observations can be made considering the Nantes host zone in the ESHM13 seismogenic source model (Fig. S1), with rates based on FCAT much higher as compared with SHEEC.

Both recurrence models are well-established and consistent in the source zone that is hosting Lourdes (Fig. 9). There are more events M_w 4.5-5.5 in the case of FCAT than expected from the extrapolation of the lower magnitude range. We observe this bump around M_w 5.0 throughout the Pyrenees; the bump remains even after excluding historical events with magnitude M_w inferred from felt testimonies or events with magnitude M_w inferred from a unique (epicentral) intensity. An in-depth analysis of the FCAT historical period is needed. There is no such bump in the SHEEC catalog. We believe that this anomalous number of events around M_w 5.0 results from a bias in the estimation of FCAT historical magnitudes in the area. As a consequence, we trust only the earthquake recurrence model based on a minimum magnitude of 3.2.

In the source zone hosting Nice (in the Southern Alps), there are 23 and 52 events available to model recurrence for SHEEC ($M_w \geq 4.4$) and FCAT ($M_w \geq 3.2$), respectively. The recurrence model relying on SHEEC is less constrained; however, the annual exceedance rates over the magnitude range 4-5 are clearly higher than the FCAT rates. This same observation can be made for the recurrence models in the source zone hosting Strasbourg. Because the FCAT b-value is lower than the SHEEC b-value, rates predicted for magnitude $M_w \geq 5.0$ are close.

In the source zones hosting Clermont-Ferrand and Briançon, the recurrence models relying on FCAT are well-constrained. Considering SHEEC, only 9 ($M_w \geq 4.4$) and 14 ($M_w \geq 4.4$) events fall within the periods of completeness. Recurrence models are built assuming a regional b-value (0.93) and relying on the rates of the two lowest magnitude bins. Observed annual rates estimated from FCAT and SHEEC roughly superimpose below M_w 5. The difference in the earthquake recurrence modeling is mostly due to the low number of events available for SHEEC.

Because we have observed clear differences in the magnitude estimates from one catalog to the other in some regions, we did not merge SHEEC and FCAT catalogues. To model the recurrence for sources beyond the French border we use only one recurrence model based on SHEEC. This decision has a negligible impact for the sites considered here. FCAT provided in the supplementary material of Manchuel et al. (2018) stops at around 20 km beyond the border for the instrumental part. For sources on the border (e.g., Strasbourg or Nice host zones), we have also modeled recurrence using an extended version of SiHex (obtained on request from authors) to check that seismic rates were not underestimated due to missing earthquakes.

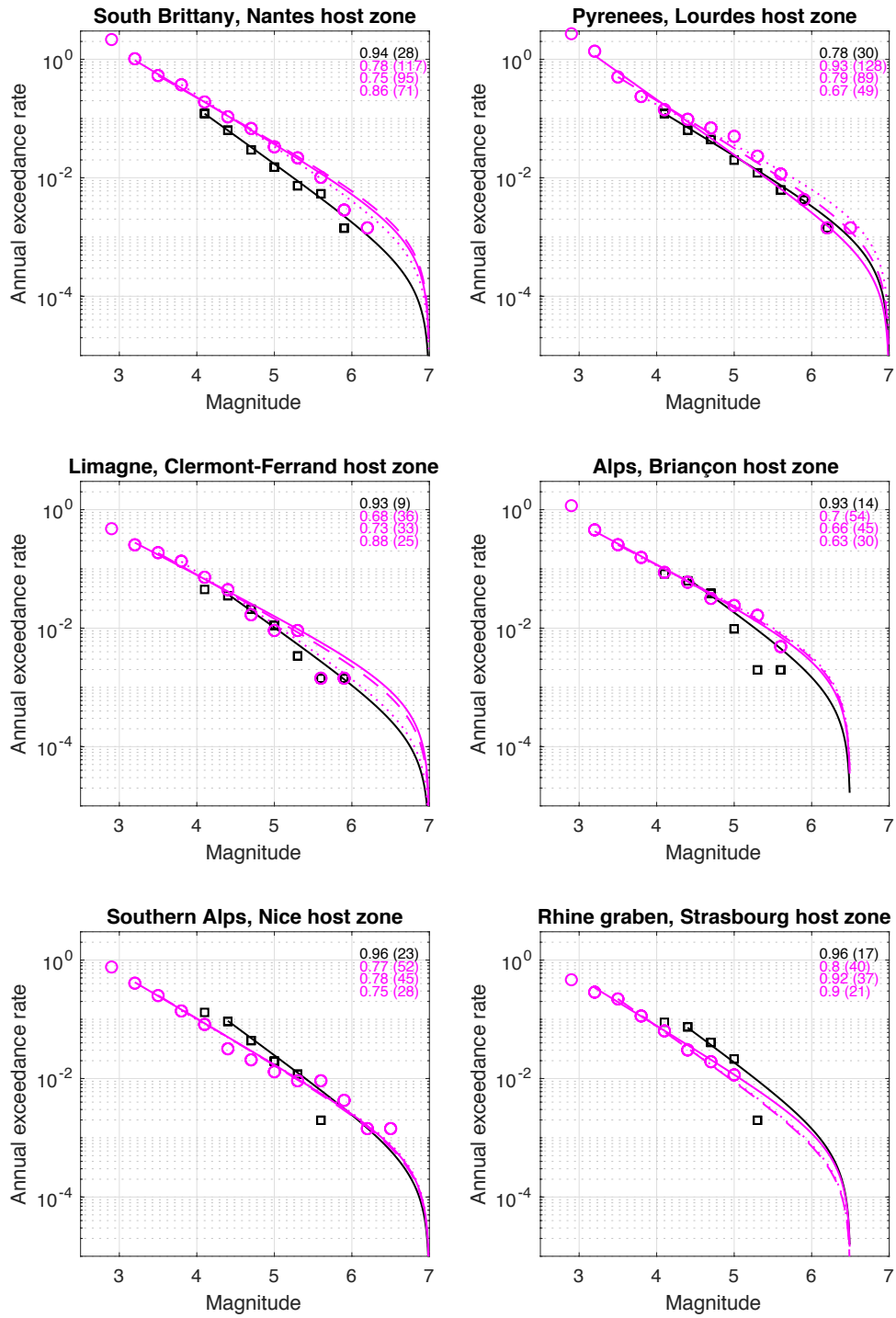


Figure 9 : Magnitude-frequency distributions for the seismogenic source zones of the IRSN model hosting the 6 sites selected for hazard calculations, established from the FCAT (in magenta, Reasenberg declustering) and SHEEC (in black). Three minimum magnitudes are considered for modeling using FCAT (3.2: solid line, 3.5: dashed line, 3.8: dotted line). The b-values are indicated; the total number of events used for modeling recurrence are in parenthesis (black, SHEEC; magenta, FCAT; considering a minimum magnitude from 3.2 to 3.8).

Table 4 : Modeling of earthquake recurrence for the source zones of the IRSN model that is hosting the six sites considered in the hazard calculation: number of events within periods of completeness– SHEEC catalogue. HZ stands for Host Zone.

Magnitude interval	4.1-4.4	4.4-4.7	4.7-5.0	5.0-5.3	5.3-5.6	5.6-5.9	5.9-6.2	6.2-6.5	M_{\max_obs}	a-value	b-value

HZ Nantes	12	7	3	2	1	2	1	0	6.2	2.93	0.94
HZ Lourdes	12	4	5	2	3	1	2	1	6.4	2.3	0.78
HZ Clermont	2	3	2	2	1	0	1	0	6.1	2.66	0.93
HZ Briançon	4	5	6	2	0	1	0	0	5.7	2.93	0.93
HZ Nice	8	10	5	2	5	1	0	0	5.8	3.2	0.96
HZ Strasbourg	3	7	4	5	1	0	0	0	5.3	3.08	0.96

Table 5 : Modeling of earthquake recurrence for the source zones of the IRSN model that is hosting the six sites considered in the hazard calculation, considering a minimum magnitude of 3.2: number of events within periods of completeness – FCAT catalogue (Reasenbergl declustering).

Magnitude interval	3.2-3.5	3.5-3.8	3.8-4.1	4.1-4.4	4.4-4.7	4.7-5.0	5.0-5.3	5.3-5.6	5.6-5.9	5.9-6.2	6.2-6.5	6.5-6.7	M _{max_obs}	a-value	b-value
HZ Nantes	22	24	29	15	7	9	3	3	3	1	1	0	6.3	2.48	0.78
HZ Lourdes	39	40	15	8	5	5	7	3	3	2	0	1	6.5	3.03	0.93
HZ Clermont	3	8	10	5	5	2	0	2	0	1	0	0	6.1	1.62	0.7
HZ Briançon	9	15	11	5	5	2	2	3	2	0	0	0	5.6	1.86	0.7
HZ Nice	7	17	9	9	2	2	1	0	2	2	0	1	6.7	2.08	0.77
HZ Strasbourg	3	16	8	6	2	2	3	0	0	0	0	0	5.6	2.09	0.8

4.5 Exploring uncertainties on recurrence modeling from FCAT: different FCAT versions and declustering algorithms

Two versions of FCAT are considered in hazard calculations to understand the impact of including very uncertain events in recurrence modeling: the full catalog and the catalog with historical magnitudes relying on felt testimonies removed. To understand the impact of the choice of the declustering algorithm, these two catalog versions are alternatively declustered with Reasenbergl and Gardner and Knopoff algorithms, producing four alternative earthquake catalogs for recurrence modeling. The completeness time windows per magnitude interval are re-evaluated for each version of the catalog from visual inspection of cumulative number of events versus time plots. Combining these four catalogs with three possible minimum magnitude for the modeling of the recurrence, twelve recurrence models are derived per source zone of a given seismogenic source model. These models populate a source-model logic tree to quantify the impact on hazard, which is described in section 5.

5. Estimating PSHA and quantifying impacts of source and ground-motion model uncertainties

5.1 PSHA calculation

Probabilistic ground shaking can be determined since recurrence parameters have been estimated for all source zones. Depth distributions for earthquakes within each source zone must be assumed; probability density functions for depth are established from the source zone sub-catalogs. Our aim is not to achieve a complete PSHA study but to perform sound tests in order to understand the impact some key decisions made while building the source model have on hazard. We make simple assumptions for M_{max}, which is the maximum magnitude bounding the upper magnitude range of the recurrence models. M_{max} is fixed to 6.5 in all sources except those in which a magnitude larger than 6.0 has been observed in the historical period in at least one of the catalogs (FCAT version with all events or SHEEC), in which case M_{max} is fixed to 7.0. We do not explore the uncertainty related to the M_{max} value. Table 6 summarizes the parameters considered in the probabilistic hazard calculation. Calculations are performed for a generic rock site (V_{S30} = 760 m/s).

Four ground-motion models for shallow crustal sources are used to explore the uncertainty for ground-motion prediction:

- Boore et al. (2014) built from the Next Generation Acceleration database (western United States and worldwide datasets);
- Bindi et al. (2014) built from a pan-European dataset;
- Cauzzi et al. (2015) built from a global dataset (~50% of data from Japan and ~30% from pan-European regions);
- Drouet and Cotton (2015) built from a stochastic model developed for the French Alps.

All models use the Joyner and Boore distance measure except Cauzzi et al. (2015), which uses rupture distance.

Table 6 : Parameters used in the PSHA calculation (minimum magnitude used for integrating the frequency-magnitude distributions, maximum source-site distance taken into account, truncation level of the Gaussian predicted by the GMM, and V_{S30} of the generic rock sites)

Parameter	Value used
M_{min}	$M_w=4.5$
Maximum distance	250km
Truncation of σ	+4
V_{S30}	760 m/S

Six cities are selected in which hazard calculation will be led (Fig. 10). Briançon, Nice, and Lourdes are localized in a region of “medium hazard” in the actual French regulation (NF EN 1998-1 [2005]; EN 1998-1 [2004]); level 4 out of 5 on a scale (in the Antilles). Clermont-Ferrand, Nantes, and Strasbourg are localized in a region of “moderate hazard” (level 3). With the exception of Nice and Lourdes, the cities are located roughly in the middle of a source zone (host zone for the site). Hazard at the site is controlled by the seismicity within the host zone, as confirmed by disaggregation studies in space (example in Fig. 10).

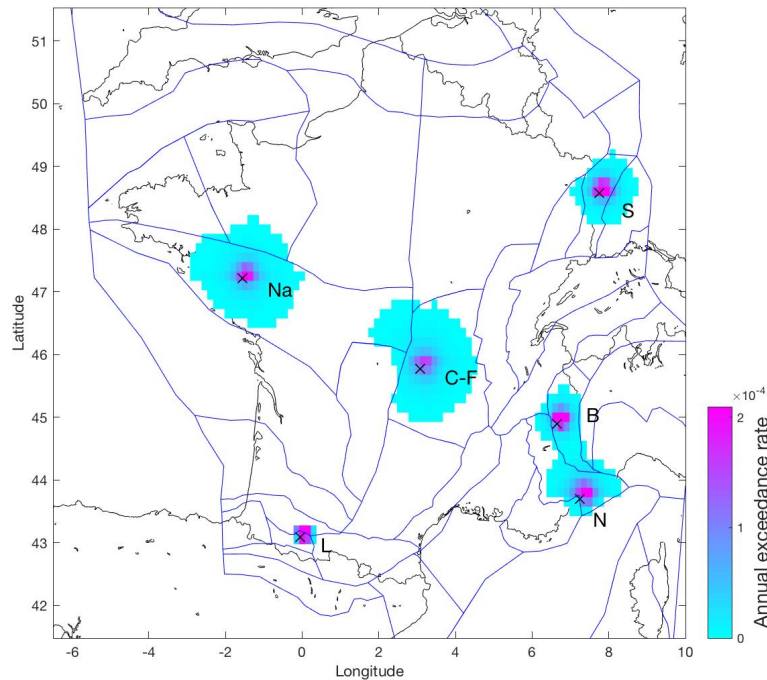


Figure 10 : Spatial disaggregation for the six cities studied for the PGA at 475 year return period. The spatial cells that contribute up to 98% of the total exceedance rate are displayed. The Baize et al. (2013) seismicogenic model is combined to the FCAT; Bindi et al. (2014) is used to predict ground-motion exceedances. Na, Nantes; L, Lourdes; C-F, Clermont-Ferrand; S, Strasbourg; B, Briançon; N, Nice.

5.2 Variability of hazard estimates considering FCAT

We set up a logic tree to explore the uncertainties on the earthquake recurrence model based on FCAT, and we propagate these uncertainties up to the hazard estimates. The logic tree includes two alternative seismicogenic source models (IRSN and ESHM13), two versions of FCAT (with and without uncertain felt-based M_w events), two alternative declustering algorithms, and three alternative minimum magnitudes for earthquake recurrence modeling (Fig. 11). Exploration of the logic tree leads to 24 alternative earthquake recurrence models per source zone that are combined with the Boore et al. (2014) ground-motion model to calculate 24 hazard curves at each site. For a given return period, the hazard curves are then interpolated and provide 24 accelerations, which are obtained for a spectral period 0.2 s at a 475 year return period and are displayed in Figure 12.

For each site, acceleration variability is displayed, first considering the full logic tree (black, 24 branches) and then grouping the accelerations estimates to understand the impact of choices made while building the source model.

We choose this representation rather than an impact analysis (as in Beauval and Scotti, 2004) or a Tornado diagram (e.g., Anderson et al. 2018), to determine an absolute impact that is not dependent on a specific set of reference parameters. Mean acceleration values are reported with the full range (minimum and maximum) of acceleration values.

- In order to quantify the impact of the seismogenic source model choices, the 24 accelerations are split into two groups: accelerations obtained with the IRSN seismogenic model (dark blue) and accelerations obtained with the ESHM13 seismogenic model (light blue).
- In order to quantify the impact of the minimum magnitude choices used to model recurrence, the 24 accelerations are split into three groups (dark to light grey, corresponding to a minimum magnitude equal to successively 3.2, 3.5, and 3.8).
- Similarly, the impact of the declustering algorithm can be evaluated by redistributing the 24 accelerations in two groups according to the employed declustering algorithm (purple, Reasenber; pink, Gardner and Knopoff).
- The impact of the FCAT version is also assessed: grouping the accelerations that rely on a recurrence model based either on all FCAT events (dark green) or rely on the same catalog with uncertain felt-based Mw events removed (light green).

Our aim is to identify which parameter choice(s) control(s) the hazard at a given site. For a given parameter choice, the larger the distance between two alternative mean values, the larger the impact on the hazard. The impact is estimated with respect to mean values of distributions. Let's A_{min} be the lower mean value and A_{max} be the larger mean value, the impact calculated is the difference between A_{max} and A_{min} normalized by A_{min} and expressed in percentage: $\frac{(A_{max}-A_{min})\times 100}{A_{min}}$.

For 4 cities (i.e., Nantes, Lourdes, Clermont-Ferrand, Strasbourg), the minimum magnitude chosen for recurrence modeling is the parameter that most influences the hazard estimate (e.g., ~31% impact in Clermont-Ferrand, ~11% in Strasbourg). For Briançon and Nice, the seismogenic source model choice is the parameter that most influences the hazard (e.g., ~26% impact in Briançon). The seismogenic source model choice also has a significant impact in Clermont-Ferrand (i.e., 30%). The declustering algorithm choice has a small influence on the hazard and is the parameter that least influences the hazard for 4 of the 6 sites (i.e., Nantes, Clermont-Ferrand, Briançon, Strasbourg). The FCAT version chosen also has a small influence on the hazard, except in Briançon (i.e., ~13% impact) and Strasbourg (i.e., ~9% impact).

In Lourdes, the minimum magnitude chosen fully controls the variability on the hazard results. The mean acceleration value is found to vary from 0.45g to 0.68g (0.2s) based on the minimum magnitude selected; however, this is an effect of the anomalous “bump” around magnitude Mw 5.0 that leads to low b-values (see Figs. 9 and S1). Exploring the 24 branches, the logic tree leads to a mean value of 0.56g, which we consider to be unrealistic. As underlined in Section 4.4, we trust only the earthquake recurrence model obtained with a minimum magnitude of 3.2, controlled by seismic rates in the magnitude range 3.2-4.4, corresponding to a mean acceleration of 0.45g.

Results for the PGA at 475 and 5000 year return periods are displayed in the Appendix (Figs. S2 and S3). Although the acceleration values are shifted toward lower or higher values, the relative trends are similar to the previous results at 0.2 s and a 475 year return period. By considering the PGA rather than 0.2 s, or increasing the return period from 475 years to 5000 years, the same observations can be made on the relative impacts on hazard of the source-model parameters.

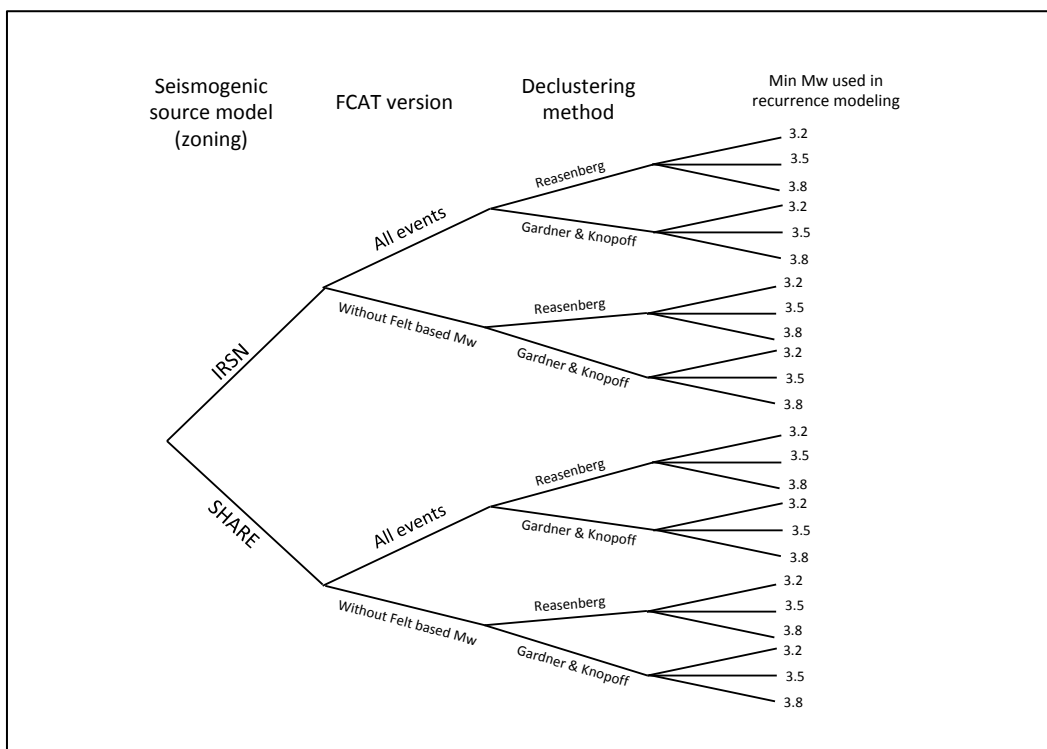


Figure 11 : Logic tree exploring uncertainties on earthquake recurrence models considering FCAT, including two alternative seismogenic source models, two alternative versions for FCAT (all events and without felt-based Mw events), two alternative declustering algorithms, and three alternative minimum magnitudes.

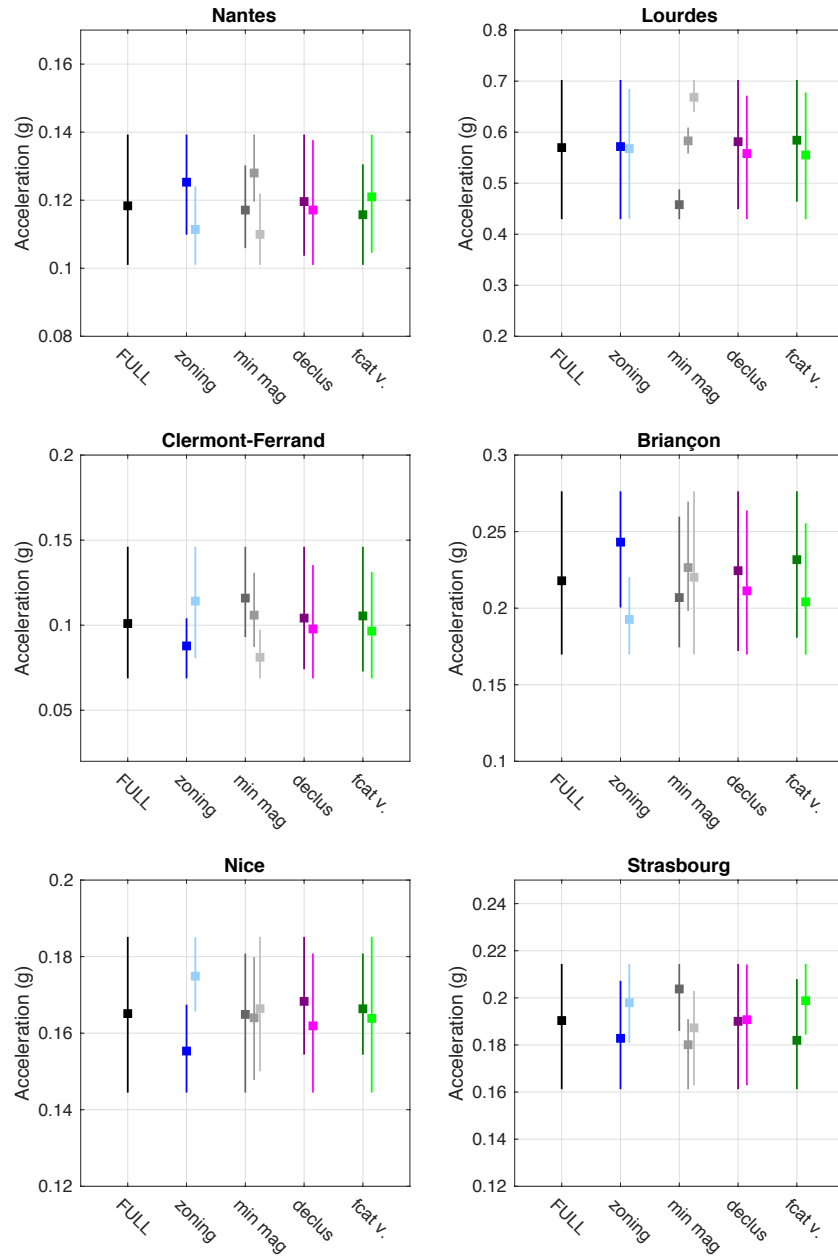


Figure 12: Hazard results exploring the FCAT source-model logic tree, for spectral period at 0.2s at 475 year return period. Mean value (square), minimum and maximum values (vertical bar). ‘FULL’: full logic tree: black (24 branches). ‘zoning’: choice of the seismogenic source model, dark blue: IRSN seismogenic source model (12 branches); light blue: ESHM13 model (12 branches). ‘min mag’: choice of the minimum magnitude for modeling earthquake recurrence: dark to light grey (either 3.2, 3.5, or 3.8, eight branches each). ‘declus’: choice of the declustering algorithm: Reasenberg (purple, 12 branches) or Gardner and Knopoff (pink, 12 branches). ‘fcatt v.’: choice of the FCAT version: with (dark green) or without (light green) very uncertain felt-based M_w events. Calculations performed with the ground-motion model Boore et al. 2014. For Lourdes, we trust only the results relying on a recurrence model derived with $M_w \geq 3.2$ (dark grey, see the text).

5.3 Differences in the hazard levels considering FCAT or SHEEC

There is no unique way of building an earthquake catalog that covers several centuries. Considering two catalogs built within two separate projects can enable an estimate of the uncertainty on hazard related to catalog choice. For FCAT, a complete logic tree has been set up to explore uncertainties in the different steps that lead to a recurrence model per source zone (Section 5.2). Using SHEEC, less can be done because the minimum magnitude of completeness is high ($M_w > 4.0$) and, for some sources, few events are available. Only two alternative hazard

values (one per seismogenic source model [Fig. 13]) are calculated here. It is instructive to analyze the differences or similarities considering FCAT and SHEEC in light of the recurrence models displayed in Figures 9 and S1.

For Clermont-Ferrand and Briançon, not much can be inferred from the comparison. The accelerations obtained from SHEEC are slightly below the 16th percentile that rely on FCAT in Clermont-Ferrand and are within the 16th-84th percentiles in Briançon. Based on inspection of host zone recurrence models and the underlying observed rates (Figs. 9 and S1), the difference between the recurrence models that rely on FCAT and SHEEC is mostly due to the low number of events and the restricted magnitude interval available for SHEEC ($M_w \geq 4.4$).

For Nice and Strasbourg, the recurrence model based on SHEEC is better constrained, with observed rates based on SHEEC clearly larger than rates based on FCAT in the magnitude interval 4-5 (Figs. 9 and S1). In Nice, the accelerations obtained from SHEEC are within the mean and 84th percentile calculated from the FCAT logic tree (Fig. 13). In Strasbourg, the accelerations inferred from the SHEEC catalog are larger than the 84th percentile from FCAT.

Nantes and Lourdes are both located within source zones that have a wealth of data in which the recurrence models inferred from FCAT and SHEEC are both well-constrained (Figs. 9 and S1). In Lourdes, the accelerations calculated from SHEEC are within the 16th percentile and the mean inferred from the FCAT logic tree (Fig. 13). In the source zone enclosing Lourdes, ignoring the bump around 5.0 for FCAT (i.e. considering a minimum magnitude of 3.2), the recurrence models from FCAT and SHEEC roughly overlap. Contrarily, in Nantes, using SHEEC rather than FCAT leads to twice as low hazard levels (around 0.06g for SHEEC with respect to a mean value of 0.12g for FCAT, at 0.2 s for the return period 475 years). In the source zone that encloses Nantes, over the magnitude range 4-5.5, seismic rates based on SHEEC are divided by 2 to 2.5 with respect to the seismic rates based on FCAT. These results show that the impact of the earthquake catalog chosen varies considerably with the site being studied; it can either be negligible or can significantly impact the hazard estimates.

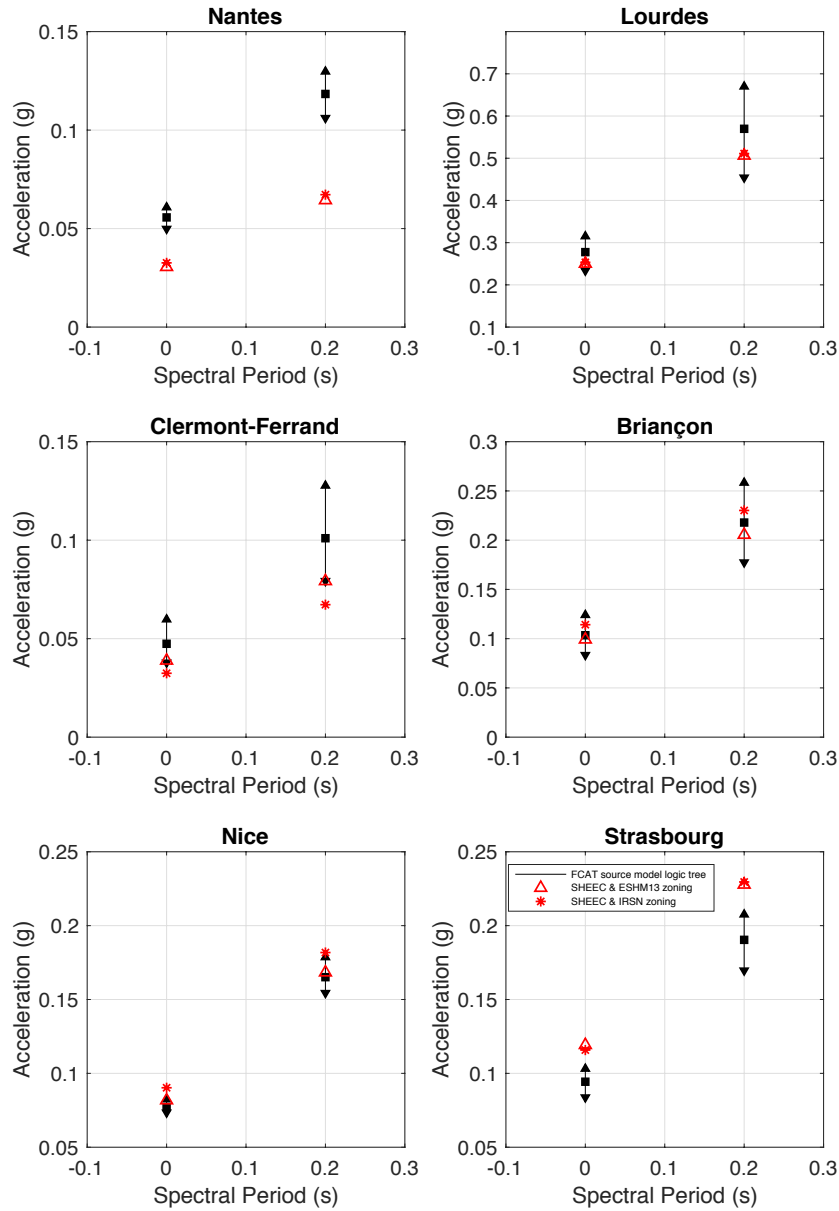


Figure 13: Comparison between hazard estimates based on recurrence models inferred from FCAT (black) or from SHEEC (red) for the spectral periods PGA (0s) and 0.2 s, at 475 year return period. The FCAT logic tree is explored (see Fig. 11), mean values (black squares) and 16th and 84th percentiles (black triangles) are displayed. Two hazard values are based on SHEEC: one for each seismogenic source model (triangle, ESHM13 model; star, IRSN model). Calculations are performed with the Boore et al. (2014) ground-motion model.

5.4 Source-model uncertainties versus ground-motion model uncertainties: impact on the hazard estimates

The source-model logic tree based on FCAT (Fig. 11) is combined with a ground-motion model logic tree. Four ground-motion models potentially adapted to the French context are included, based on four different strong-motion databases. The variability of hazard estimates related to the uncertainties on the source model is now compared to the variability related to the uncertainties on the ground-motion model (following Beauval et al. 2018). As the source model logic tree includes only one earthquake catalog (i.e., FCAT), the earthquake catalog choice is not included in the uncertainty exploration, and the variability on the hazard estimate must be considered as a lower bound.

Figure 14 displays the complete acceleration distribution, exploring the source-model logic tree based on FCAT combined with the ground-motion logic tree ($24 \times 4 = 96$ branches). The distribution of accelerations obtained for a given branch of the source model is also displayed to identify the variability related to uncertainties on the prediction of ground motions. Results are shown for only two branches of the FCAT source model (see Fig. 14

legend), but we have checked that the variability is stable regardless of the branch selected. The distribution of accelerations obtained for a given branch of the ground-motion model is also displayed, reflecting the variability related to the uncertainties on the FCAT source model. Results are shown only for two ground-motion models, and we have checked that the variability obtained only slightly depends on the ground-motion model selected.

For the PGA at 475 year return period, variability on the hazard estimates related to uncertainties on the ground-motion model are found to be equivalent to (Strasbourg), higher than (Nice, Nantes), or lower than (Clermont-Ferrand, Briançon) the variability related to the uncertainties on the source model. For the spectral period 0.2 s at 475 year return period (Fig. 15), exploration of the FCAT source-model logic tree leads to a higher hazard variability than exploration of the ground-motion logic tree in Clermont-Ferrand, Briançon, and Strasbourg. For Nantes and Nice, source model and ground-motion model uncertainties equally contribute to the overall uncertainty. Results for the PGA at 5000 years are displayed in the Appendix (Fig. S4). Our results show that the relative impact on hazard estimates of source-model uncertainties with respect to ground-motion model uncertainties varies with the site, spectral period, and return period.

For Lourdes, the variability related to the source-model uncertainties is much higher than the variability associated to the selection of the ground-motion model. These results are due to the low b-values obtained using minimum magnitudes of 3.5 and 3.8 for recurrence modeling (anomalous high seismic rates around M_w 5.0 in the observed seismic rates), we only trust the results for minimum magnitude 3.2. The historical part of the next version of FCAT in the Pyrenees should be analyzed in detail.

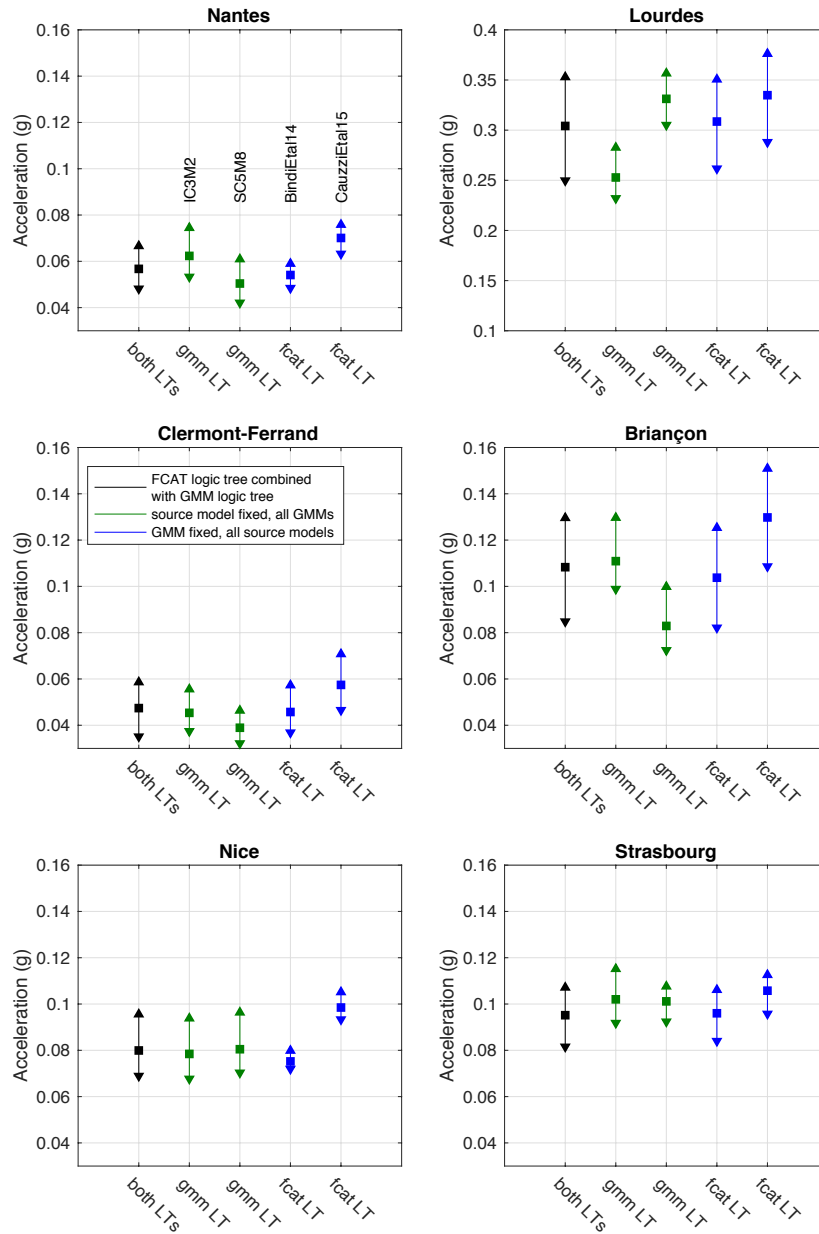


Figure 14 : Sensitivity of PGA estimates to the source model and ground-motion model uncertainties at 475 year return period (mean values and 16th and 84th percentiles). ‘both LTS’ : Exploring both the FCAT source-model logic tree and the ground-motion model logic tree (black, 96 branches). ‘gmm LT’ : For a fixed source-model branch (IC3M2, SC5M8), exploration of the full ground-motion logic tree (green, 4 branches). ‘fcap LT’ : For a fixed ground-motion model (Bindi et al. 2014, then Cauzzi et al. 2015), exploration of the full FCAT source-model logic tree (24 branches, blue). IC3M2 : IRSN seismicogenic source model + FCAT version without very uncertain felt-based M_w events + Reasenberg declustering + minimum magnitude 3.2 for recurrence modelling ; SC5M8 : ESHM13 seismicogenic source model + FCAT version without very uncertain felt-based M_w events + Gardner & Knopoff declustering + minimum magnitude 3.8.

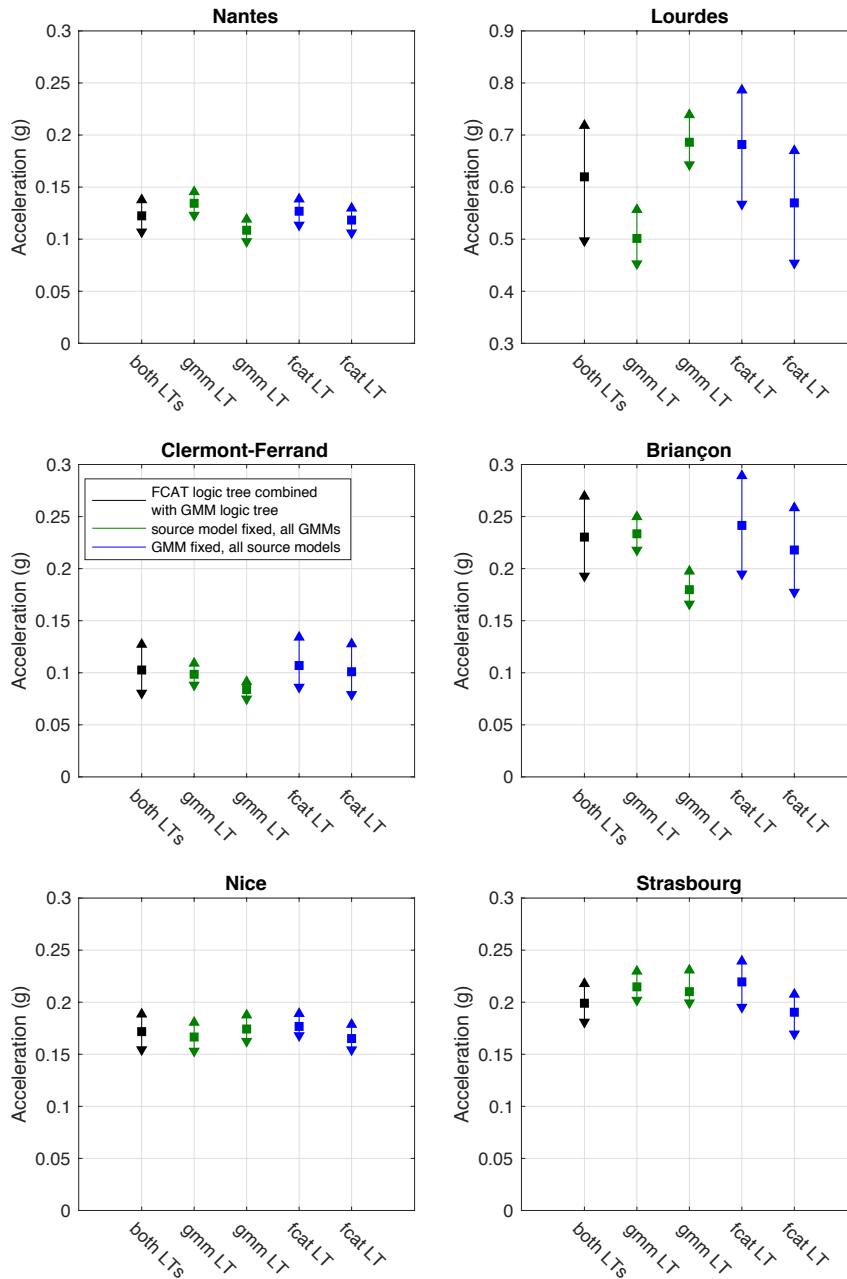


Figure 15 : Sensitivity of the acceleration for spectral period 0.2 s to the source model and ground-motion model uncertainties at a 475 year return period. See Figure 14 legend.

5.5 Impact on the power-law decay of the hazard curves

An important characteristic of hazard curves is their evolution with return period, which is often approximated as a power law in the form:

$$Sa(T, RP) = Sa(T, RP_0) * (RP/RP_0)^\alpha,$$

with $Sa(T, RP)$ the spectral acceleration at spectral period T and return period RP ,

and $Sa(T, RP_0)$ the spectral acceleration at spectral period T and return period of reference RP_0 (usually 475 year).

Such an approximation is used (e.g., in Bazzurro and Cornell, 2004) to convolve the distribution of rock hazard with the distribution of site amplification in order to derive an analytic formula for site-specific PSHA (see, e.g., Aristizabal et al., 2018). It is also used in earthquake regulations to derive the coefficients associated with various

building importance classes in relation to the corresponding exceedance probability levels (i.e., to the corresponding return periods). The value commonly accepted, especially in the French and EC8 building codes, is $\alpha = 1/3$, following (among others) the simple theoretical derivations by Betbeder-Matibet (2008). This exponent value is hidden behind most of the importance coefficients listed in EC8. The α value of $1/3$ leads to a ratio of ~ 2.2 between the acceleration at 5000 and 475 years.

The value of this exponent potentially depends on the magnitude scaling of spectral ordinates (ground motion prediction equations [GMPEs]), on the aleatory variability of GMPEs, and on the magnitude-frequency distributions (a and b-values, M_{\max}); thus, it could depend on the assumptions regarding the source models. We have therefore also analyzed the impact of the various assumptions considered in this work on the ratio $S_a(T, RP=5000)/S_a(T, RP=475)$ for PGA (0s) and 0.2 s periods. Figure 16 displays the ratios obtained exploring the FCAT logic tree, as well as considering the SHEEC catalog combined with both seismogenic source models (IRSN and ESHM13), and predicting ground motions with the Boore et al. (2014) model. Results show that exploration of the FCAT logic tree has a limited impact on the ratios, which vary between 2.8 and 4.2 depending on the source model and the site. Ratios that rely on SHEEC can be close to the values that result from FCAT (e.g., in Nantes, Lourdes, Clermont-Ferrand) or slightly lower than the 16th percentile obtained from FCAT (e.g., in Briançon, Nice, and Strasbourg). Ratios are shifted toward slightly larger values for all sites when the Bindi et al. (2014) ground-motion model is considered (varying from 2.8 to 5.1, Fig. S5).

Overall, it is worth noticing that the ratios between hazard levels at 5000 and 475 year return periods exhibit two clear features (Fig. 16):

- a) Large site-to-site variability of the ratio;
- b) Values obtained are systematically much higher than what is commonly accepted (~ 2.2). Observed values of 3 to 5 correspond to an exponent α between 0.5 and 0.7.

The commonly accepted exponent value $1/3$ is based on theoretical developments ignoring the ground-motion model aleatory variability (Betbeder-Matibet, 2008). More work is required to understand which parts of the source and GMM models have an influence on this exponent. Future efforts should also be dedicated to clarify the basis on which the importance coefficients and corresponding return periods recommended in the building codes have been derived.

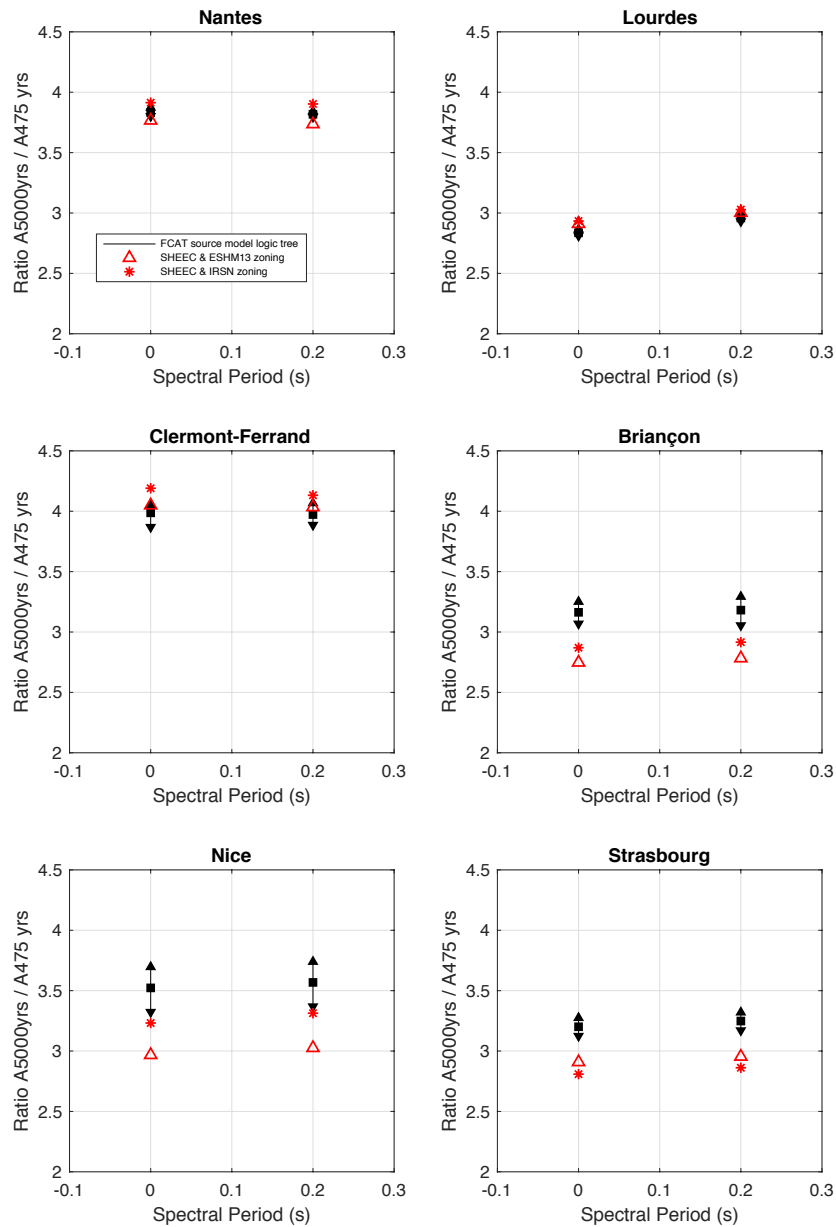


Figure 16: Ratios between accelerations at 5000 and 475 year return periods for the PGA (0s) and the spectral period 0.2 s. Square and triangles: exploration of the FCAT source-model logic tree, mean and 16th and 84th percentiles. Red symbols: ratios estimated from two source models based on SHEEC (triangle: combined with ESHM13 seismicogenic source model; star: with IRSN model). Calculations performed with the Boore et al. (2014) ground-motion model.

6. Conclusions

Our aim is to explore uncertainties related to the source model in hazard calculations for France. We study the impact on hazard when using alternative seismicogenic source models, alternative earthquake catalogs, alternative declustering algorithms, and different minimum magnitudes for earthquake recurrence modeling. We use the FCAT and SHEEC earthquake catalogs. Alternative source models are combined with four ground-motion models (i.e. Bindi et al. 2014, Boore et al. 2014, Cauzzi et al. 2015, Drouet and Cotton 2015) to estimate probabilistic seismic hazard. We focus on sites located in source zones with a minimum amount of data to work with, within areas of moderate seismicity in France (Alsace, Alps, Pyrenees, Massif Central and Southern Brittany).

Considering FCAT, a logic tree can be set up to quantify the overall impact of source-model uncertainties on hazard. The results show that the parameters that have the highest impact on hazard vary among sites. Nonetheless, the parameters that most impact hazard are the minimum magnitude used to model earthquake recurrence (up to

31% impact on the mean hazard estimate) and the seismogenic source model (up to 30%). Considering different spectral periods (PGA, 0.2s) and different return periods (475 and 5000 yrs), the controlling parameters remain identical, and the estimated impacts remain stable.

Because SHEEC has a high minimum magnitude of completeness ($M_w > 4.0$), most zones are left with few events. This lack of data prevents proper exploration of uncertainties. Nonetheless, there are a few source zones where the recurrence model is well-constrained. Earthquake recurrence models based on FCAT and SHEEC are found comparable in some sources (e.g., the Pyrenees). In other sources, seismic rates inferred from FCAT can be either lower (e.g., in the Lower Rhine graben) or much higher (e.g., in Southern Brittany) than the rates inferred from SHEEC. For Nantes, choosing FCAT rather than SHEEC leads to twice as large mean hazard values (PGA and 0.2 s, 475 year return period).

Four ground-motion models that rely on four different databases are selected as an attempt to sample the epistemic uncertainty on the prediction of ground motions for France. The overall variability on hazard estimates can be obtained for the sites under study by combining these four ground-motion models with the logic tree established from FCAT. We also assess the respective contribution of the source-model uncertainty and ground-motion model uncertainty to the overall uncertainty. The contribution of the source-model uncertainty can be lower than, equivalent to, or higher than the contribution of the ground-motion model uncertainty and varies with the site, spectral period, and return period.

Some anomalous features in the FCAT earthquake catalog have been highlighted (e.g., problem of homogeneity in magnitude between the historical and instrumental periods, apparently too many events around $M_w 5.0$ in the Pyrenees) that might impact the hazard estimates obtained in the present study. There is no unique method to infer earthquake parameters from macroseismic data or to homogenize an instrumental catalog in terms of moment magnitude. Earthquake catalogs built by different teams may lead to different seismic rates for seismogenic sources. One way to handle this uncertainty in seismic hazard modeling would be to integrate alternative earthquake catalogs in the source-model logic tree.

Acknowledgements

This work has been partially funded by the European Research program H2020, entitled Seismology and Earthquake Engineering Research Infrastructure Alliance for Europe (SERA). The OpenQuake engine has been used to run hazard calculations (Pagani et al. 2014). We are thankful to the reviewers. They performed an in-depth analysis of the manuscript; their questions and remarks led to a more complete exploration of uncertainties and to an improved manuscript.

References

- Akkar S., Sandikkaya M.A., Bommer J.J. (2014): Empirical Ground-Motion Models for Point- and Extended- Source Crustal Earthquake Scenarios in Europe and the Middle East. *Bulletin of Earthquake Engineering*, 12(1), 359-387.
- Anderson G. (2018). Quantifying the Epistemic Uncertainty in the Probabilistic Seismic Hazard from Two Major Faults in Western Nevada, *Earthquake Spectra*, Vol. 34, No. 2, pp. 549-568, <https://doi.org/10.1193/080717EQS156M>.
- Anderson, J. G., and J. E. Luco (1983). Consequences of slip rate constraints on earthquake occurrence relations, *Bull. Seismol. Soc. Am.* 73, no. 2, 471–496.
- Aristizabal, C., PY. Bard, C. Beauval and JC. Gomez (2018). Integration of site effects into PSHA : a comparison between two fully probabilistic methods on the Euroseistest site, *Geosciences*, 8 (8), ISSN 2076-3263.
- Autran A., Blès J.-L., Combes P., Cushing M., Dominique P., Durouchoux C., Gariel J.-C., Goula X., Mohammadioun B. & Terrier M. (1998). – Probabilistic seismic hazard assessment in France. Part One: seismotectonic zonation. – 11th European Conference on Earthquake Engineering, Balkema, Rotterdam.
- Baize S., E.M. Cushing, F. Lemeille H. Jomard (2013). Updated seismotectonic zoning scheme of Metropolitan France with reference to geologic and seismotectonic data, *Bull. Soc. géol. France*, 2013, t. 184, no 3, pp. 225-259.
- Basili R., Carafa M., Kastelic V., Maesano F., Tiberti M., Rovida A., Antonucci A., Weatherill G., Lammers S., Danciu L. (2018). SERA Deliverable D25.2, Updated databases of seismicity, faults, and strain rates for ESHM20, 70 pages.
- Baumont D, Manchuel K, Traversa P, Baumont D, Durouchoux C, Nayman E, Ameri G Intensity attenuation models calibrated in M_w for metropolitan France (2018). *Bull Earthq Eng*, 16:2285–2310, <https://doi.org/10.1007/s10518-018-0344-6>.
- Beauval, C., and O. Scotti (2004). Quantifying sensitivities of Probabilistic Seismic Hazard Assessment for France to earthquake catalog uncertainties, truncation of ground-motion variability and magnitude limits, *Bull. Seismol. Soc. Am.* 94, 1579–1594.
- Beauval C., J. Marinière, H. Yepes, L. Audin, JM. Nocquet, A. Alvarado, S. Baize, J. Aguilar, JC Singaicho, and H. Jomard (2018). A New Seismic Hazard Model for Ecuador, *BSSA*, 108 (3A) : 1443-1464.

- Betbeder-Matibet, J., 2008, Seismic Engineering, ISTE Ltd and John Wiley & Sons Inc, 957 pages, DOI:10.1002/9780470611135.
- Bindi D., Massa M, Luzi L, Ameri G, Pacor F, Puglia R, Augliera P (2014). Pan-European ground motion prediction equations for the average horizontal component of PGA, PGV and 5 %-damped PSA at spectral periods of up to 3.0 s using the RESORCE dataset, *Bulletin of Earthquake Engineering*, 12(1), 391-340.
- Boore, D. M., J. P. Stewart, E. Seyhan, and G. Atkinson (2014). NGA-West2 equations for predicting PGA, PGV, and 5% damped PGA for shallow crustal earthquakes, *Earthq. Spectra* 30, no. 3, 1057–1085.
- Burkhard M and Grünthal G (2009) Seismic source zone characterization for the seismic hazard assessment project PEGASOS by the Expert Group 2 (EG 1b). *Swiss J Geosci* 102(1):149–188.
- Cauzzi, C., E. Faccioli, M. Vanini, and A. Bianchini. 2015. “Updated Predictive Equations for Broadband (0.01–10 s) Horizontal Response Spectra and Peak Ground Motions, Based on a Global Dataset of Digital Acceleration Records.” *Bulletin of Earthquake Engineering* 13(6):1587–1612. doi: 10.1007/s10518-014-9685-y.
- Cara M, Cansi Y, Schlupp A et al (2015). Si-Hex: a new catalogue of instrumental seismicity for metropolitan France. *Bull Soc Géol Fr*, 186, 1, 3–19.
- Cara M., Denieul M., O. Sèbe, B. Delouis, Y. Cansi, A. Schlupp (2017). Magnitude M_w in metropolitan France, *Journal of Seismology*, 21, 551-565, DOI 10.1007/s10950-016-9617-1.
- Christophersen, A., and E. G. C. Smith (2008). Foreshock rates from aftershock abundance, *Bull. Seismol. Soc. Am.* 98, no. 5, 2133–2148.
- Denieul M, Sèbe O, Cara M, Cansi Y (2015) Mw from crustal coda waves recorded on analog seismograms. *B Seismol Soc Am* 105:831–849. doi:10.1785/0120140226.
- Drouet S. and F. Cotton (2015). Regional Stochastic GMPEs in Low-Seismicity Areas: Scaling and Aleatory Variability Analysis—Application to the French Alps, *Bulletin of the Seismological Society of America*, Vol. 105, No. 4, pp. 1883–1902.
- EN 1998-1 (2004). Eurocode 8: Design of structures for earthquake resistance – Part 1: General rules, seismic actions and rules for buildings [Authority: The European Union Per Regulation 305/2011, Directive 98/34/EC, Directive 2004/18/EC, 231 pages.
- Gardner J.K. and L. Knopoff (1974). Is the sequence of earthquakes in southern California, with aftershocks removed, Poissonian? *BSSA*, 64, 5, 1363-1367.
- Gasperini P, Bernardini F, Valensise G, Boschi E (1999). Defining seismogenic sources from historical earthquake felt reports. *Bull Seismol Soc Am* 89:94–110.
- Gasperini P, Vannucci G, Tripone D, Boschi E (2010) The location and sizing of historical earthquakes using the attenuation of macroseismic intensity with distance. *Bull Seismol Soc Am* 100:2035–2066.
- Gutenberg, B., and F. Richter (1944). Frequency of earthquakes in California, *Bull. Seismol. Soc. Am.* 34, 185–188.
- Grünthal G, Wahlström R (2012) The European-Mediterranean Earthquake Catalogue (EMEC) for the last millennium. *J Seismol*. doi:10.1007/s10950-0129302-y.
- Laurendeau, A., C. Clément, O. Scotti (2019). Catalogue de sismicité instrumentale unifié en M_w pour la France métropolitaine : la problématique des conversions de magnitude. 10^{ème} Colloque National de l’AFPS, 24 au 27 Septembre 2019, Strasbourg. 10 pages.
- Manchuel K, Traversa P, Baumont D, Cara M, Nayman E, Durouchoux C (2018). The French seismic CATalogue (FCAT-17). *Bull Earthq Eng*. <https://doi.org/10.1007/s10518-017-0236-1>.
- NF EN 1998-1 (2005). Eurocode 8 - Calcul des structures pour leur résistance aux séismes - Partie 1 : règles générales, actions sismiques et règles pour les bâtiments (Indice de classement : P06-030-1), ISSN 0335-3931, 186 pages.
- Pagani, M., D. Monelli, G. Weatherill, L. Danciu, H. Crowley, V. Silva, P. Henshaw, L. Butler, M. Nastasi, L. Panzeri, et al. (2014). OpenQuake-engine: An open hazard (and risk) software for the global earthquake model, *Seismol. Res. Lett.* 85, 692–702.
- Petersen, M. D., M. P. Moschetti, P. M. Powers, C. S. Mueller, K. M. Haller, A. D. Frankel, Y. Zeng, S. Rezaeian, S. C. Harmsen, O. S. Boyd, et al. (2014). Documentation for the 2014 update of the United States national seismic hazard maps, *U.S. Geol. Surv. Open-File Rept.* 2014-1091, 243 pp., doi: 10.3133/ofr20141091.
- Reasenber, P. A. (1985). Second-order moment of central California seismicity, *J. Geophys. Res.* 90, 5479–5495.
- Schorlemmer D. and J. Woessner (2008). Probability of detecting an earthquake, *Bull Seismol Soc Am* 98:2103–2117.
- Stucchi M, Rovida A, Gomez Capera AA et al (2012) The SHARE European Earthquake Catalogue (SHEEC) 1000–1899. *J Seismol* 17:523–544. doi:10.1007/s10950-012-9335-2
- Traversa P, Baumont D, Manchuel K, Nayman E (2018). Exploration tree approach to estimate historical earthquakes M_w and depth, test cases from the French past seismicity. *Bull Earthq Eng*. doi:10.1007/s10518-017-0178-7
- Weichert, D. H. (1980). Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes, *Bull. Seismol. Soc. Am.* 70, no. 4, 1337–1346.
- Woessner J., L. Danciu, D. Giardini et al. (2015). The 2013 European Seismic Hazard Model : key components and results, *Bull. Earthquake Eng.*, DOI 10.1007/s10518-015-9795-1.