Computing time-dependent activity rate using non-declustered and declustered catalogues. A first step towards time dependent seismic hazard calculations for operational earthquake forecasting

David Montiel-López¹, Sergio Molina¹,², Juan José Galiana-Merino³,⁴, Igor Gómez¹,², Alireza Kharazian¹, Juan Luis Soler-Llorens⁵, José Antonio Huesca-Tortosa⁶, Arianna Guardiola-Villora⁷, and Gonzalo Ortuño-Sáez⁸

¹Multidisciplinary Institute for Environmental Studies "Ramón Margalef" (IMEM), University of Alicante, Ctra. San Vicente del Raspeig, s/n, 03080 Alicante, Spain
²Department of Applied Physics, University of Alicante, Ctra. San Vicente del Raspeig, s/n, 03080 Alicante, Spain
³University Institute of Physics Applied to Sciences and Technologies, University of Alicante, Ctra. San Vicente del Raspeig, s/n, 03080 Alicante, Spain
⁴Department of Physics, Systems Engineering and Signal Theory, University of Alicante, Ctra. San Vicente del Raspeig, s/n, 03080 Alicante, Spain
⁵Department of Earth and Environmental Sciences, University of Alicante, Ctra. San Vicente del Raspeig, s/n, 03080 Alicante, Spain
⁶Department of Architectural Constructions, University of Alicante, Ctra. San Vicente del Raspeig, s/n, 03080 Alicante, Spain
⁷Department of Continuum Mechanics and Theory of Structures, Universitat Politècnica de València, Camino de Vera, s/n, 46022 Valencia, Spain
⁸Municipality of Orihuela, 03300 Orihuela, Spain

Correspondence: David Montiel-López (david.montlop@ua.es)

Abstract. Probabilistic Seismic Hazard Analysis (PSHA) typically requires tectonic b-values and seismic activity rates using declustered catalogues to compute the annual probability of exceedance of a given ground motion (for example, the peak ground acceleration or PGA). In this work, we propose a methodology that includes the spatially-gridded time-dependent b-value and activity rate computation using seismic clusters in PSHA calculations. To account for the spatial variability and the relationship of the earthquakes with the seismic sources, we incorporate the distance from the grid cell to the closest fault and the epicentre’s uncertainty into the smoothing kernel as the average distance and the variance, respectively. To illustrate this methodology, we selected two scenarios, one in central Italy where L'Aquila earthquake happened and one in south-eastern Spain, where several earthquakes with a moment magnitude (Mw) greater than 4.0 have taken place over the last 30 years, including two earthquakes with greater than or equal to 5.0 Mw. We compared three different seismic activity models based on the parameters considered in the calculations (distance from spatial cells to faults and epicentral distance uncertainty) and we defined and calculated the changes of the annual probability of exceedance for a given background PGA value. The results reveal an oscillation of the changes of the annual probability of exceedance in the proximity of the occurrence of significant events. The increase is more significant in high seismicity areas, such as Italy, but it is no so evident in moderate seismicity regions as Spain. However, we have observed how, for moderate to low seismicity regions, the use of a non-declustered catalogue can be appropriate when computing time-dependent PSHA, as in the case of Spain.
1 Introduction

Probabilistic Seismic Hazard Analysis (PSHA) has been the basis for seismic engineering design since Cornell (1968) proposed it in order to account for all the possible earthquake scenarios and ground motion levels that can occur in the different seismic sources affecting the site of interest. One of the key points of PSHA is how the uncertainties are incorporated into the ground-motion computation, so the results are much more appropriate for use in engineering decision-making for risk reduction. However, the procedure increases in complexity (Budnitz et al., 1997).

PSHA results will depend on combining the appropriate input models (those which, according to the scientific and engineering communities, represent the relevant phenomena in an appropriate way). Therefore, the choice of these models will evolve as our knowledge of the seismic activity and occurrence increases.

PSHA determines the probability of exceeding the ground motion level over a specified time period based on the occurrence rate of earthquakes and Ground Motion Prediction Equations (GMPEs). The occurrence rate of earthquakes is generally described by the truncated exponential model (Cosentino et al., 1977) and the characteristic earthquake model (Schwartz and Coppersmith, 1984). Additionally, this earthquake occurrence rate or activity rate is assumed constant during the computation process. Therefore, it provides results which can be used for the seismic design. Once the knowledge of the seismic activity and occurrence improves due to the recording of new rare events or new tectonic information and models, the PSHA can be calculated again, and the seismic building codes will be updated if needed.

On the other hand, many authors have begun to focus the PSHA computations from a temporal or ‘real-time’ perspective, so the term ‘time dependent probabilistic seismic hazard – TDPSHA’ is now widely used. They are based on how the probabilities of large events increase as stress builds up on a fault plane until it reaches the breaking strength of the rock (Kanamori and Brodsky, 2004) and also how the probabilities of large aftershocks are a decreasing function after the main large event (Ogata, 1988; Reasenberg and Jones, 1989). However, measuring changes in the stress caused by the main shock is possible only indirectly and with somewhat low precision.

In general small earthquakes are more frequent than large earthquakes. This is quantitatively stated in the Gutenberg–Richter law (Gutenberg and Richter, 1956) (G-R from now on) that can be seen in Eq. 1:

\[ \log_{10} N(M \geq m) = a - b \cdot m \] (1)

where \( N \) is the cumulative number of earthquakes with magnitude \( M \) above \( m \), the \( b \) value is the average size distribution of earthquakes (which expresses the ratio between high magnitude and low magnitude earthquakes) and \( a \) is the productivity, or more precisely, \( 10^a \) is known as the seismic activity rate. As the PSHA results are given as an annual probability of exceedance for a given intensity of the ground motion, the most common way to work with the activity rate is using the annual activity rate, which is obtained by dividing \( 10^a \) by the duration in years of the seismic catalogue. So, if we can identify seismic sources a priori, then the seismic data inside each seismic source is used to compute a source-specific (\( a \) and \( b \)) magnitude frequency distribution. However, it is often challenging to identify the corresponding boundaries and to have enough data allowing a significant statistical fitting.
Therefore, Frankel (1995) instead of specifying spatial borders for each seismic source adopted a boundary-less source model when computing the PSHA for central and eastern United States. Under this approach, the historical seismicity is spatially smoothed, and activity rates are computed at a grid of locations through the analysis domain. First, he divided the region into a grid, and then he counted for each cell of the spatial grid the number of earthquakes greater than a reference magnitude ($M_{ref}$) depending on the occurrence year of the event (1700 for magnitudes greater than 5.0 Mw and 1924 for magnitudes greater than 3.0 Mw). Next, the author obtained a maximum likelihood estimate for a ($Weichert, 1980$) that he would then smooth using a Gaussian kernel with a correlation distance, $c$, of 50 km. This normalised smoothed value, $\tilde{n}_i$, was calculated as follows (Eq. 2):

$$\tilde{n}_i = \frac{\sum_j n_j \cdot \exp(-\Delta_{ij}^2/c^2)}{\sum_j \exp(-\Delta_{ij}^2/c^2)}$$

(2)

where $\Delta_{ij}$ is the distance between the $i^{th}$ and the $j^{th}$ cells of the grid and then the summation of the counts, $n_j$, over $j$ is done considering cells within distance equal to 3 times $c$ (being $c$ the aforementioned correlation distance) from the $i^{th}$ cell.

Later, Woo (1996) proposed an alternative finite-range form for the kernel, based on the fractal dimension of epicentres and shown in Eq. 3:

$$K(M, x) = \frac{D}{2\pi \times \eta(M)} \left( \frac{h(M)}{r} \right)^{2-D} \text{ if } r \leq h(M)$$

$$K = 0 \text{ if } r > h(M)$$

(3)

where $M$ is magnitude for a location $x$, $r$ is the radial separation distance and $h(M)$ is a magnitude-dependent bandwidth parameter which can be parametrised as $h(M) = H \cdot \exp(k \cdot M)$ where $H$ and $k$ are regionally estimated constants using seismological and geological considerations and $D$ is the fractal dimension of the epicentres.

Subsequently, Helmstetter et al. (2006) proposed a model for the seismicity density calculation by means of an isotropic adaptive kernel (Izenman, 1991) that smoothed the seismicity depending on the number of events (in order to increase or decrease the detail in the seismicity calculations). Here, the parameter used for the smoothing kernel depends on the average distance between all the events around an earthquake but also, on the accuracy of the epicentre location in the first instrumental era of the earthquake catalogue.

Hiemer et al. (2014) created a model based on the seismicity and the fault moment release in order to consider the active mechanisms that generate seismicity in a more direct manner in order to smooth the seismicity (i.e., the locations of the earthquakes). They use the kernel defined by Helmstetter et al. (2006) and, similarly, the fault moment rate was smoothed with an isotropic kernel (Eq. 4):

$$k_i(r) = \frac{C(d)}{(r^2 + d^2)^{1.5}} \cdot M_i$$

(4)

where $k$ is the value of the smoothing kernel for a fault point $i$, $M_i$ is the fault moment at that point, $d$ is the constant smoothing distance, $C(d)$ is the normalisation constant and $r$ is the epicentral distance.
More recently, the 2020 European Seismic Hazard model (ESHM20) has been released (Danciu et al., 2021). The authors combine the smoothing seismicity algorithms with active fault models. In this case, they point out the challenge of avoiding double counting events around faults when they consider the background seismicity and the one linked to the fault’s activity. Another example of this approach is shown in the work by Pandolfi et al. (2023), where the authors combine 3D information of the seismic sources with the data in the seismic catalogue to calculate the seismic rate.

The works cited in the previous paragraph showcase the importance of considering the active seismogenic sources when computing the activity rate. A common assumption within PSHA is that seismicity can be well-described by a Poisson process (Cornell and Winterstein, 1987). A fundamental property of Poisson processes is that the instantaneous rate of events is constant and does not depend upon the occurrence of other events located close in either space or time. However, earthquake sequences feature a significant number of aftershocks, and these events are dependent upon the main shock. The purpose of declustering seismicity data is to remove these dependent events so that the underlying long-term average rate of occurrence can be estimated.

Taroni and Akinci (2021) proposed the use of aftershocks and foreshocks in the seismic activity calculation since removing such events from seismic catalogues may lead to underestimating seismicity rates and, consequently, the final seismic hazard in terms of ground shaking. To do this, they used as kernel a simple weight function of the form (Eq. 5):

\[ k_i(N) = \frac{1}{N} \]  

(5)

where \( N \) is the number of events in the seismic series in which the event \( i \) belongs.

This weight function ensures that the contribution of each event will be the same for the activity rate computation, regardless of its association with a seismic series.

With all the exposed factors, we are going to investigate the sensitivity of the activity rate computation model to both the proximity of the spatial cells to the seismic sources and the epicentral uncertainty related to each event of the catalogue and its influence on a time-dependent seismic hazard, so we will evaluate if the obtained values may be used as a decision factor on Operational Earthquake Forecasting (OEF). We will also consider the foreshocks and aftershocks in order to calculate this activity rate by means of a previous clustering process so that main shock and corresponding foreshocks and aftershocks are grouped in a given cluster, but we will also compare the results with the ones obtained by using a declustered catalogue. To do this, we will consider two case studies: Central Italy, a high seismicity area, which will help calibrate the models proposed, and Southeastern Spain, a moderate seismicity area, in which different treatments in the catalogue will be tested (declustering and using tectonic b-value vs time-dependent b-value).

2 Methodology

In this section, the procedure used to obtain the parameters that will be used inside the smoothing kernel is described. The purpose of this kernel is to smooth the gridded seismic activity (for which a spatial grid is previously defined), helping to improve the description of the seismicity in the area. These parameters will define the different models to be tested in the different areas.
2.1 Smoothing kernel parameters

For this work, a well known smoothing function has been selected to smooth the gridded seismicity, the 2D Gaussian function (Eq. 6):

\[ f(r) = A \cdot \exp \left( -\frac{(r - \mu)^2}{2 \cdot \sigma^2} \right) \]  

where \( r \) is the distance between spatial grid cells, \( \mu \) is the first moment of the distribution, \( \sigma \) is the second moment of the distribution and \( A \) is the normalization constant.

As expressed in the last paragraph of the introduction, we will avoid any arbitrary choice in the definition of these parameters \((\mu \text{ and } \sigma)\) by assuming they have a geophysical meaning.

2.1.1 First moment of the distribution, \( \mu \)

The mean value of the Gaussian distribution, \( \mu \), is the value of the independent variable (in this case the distance between spatial grid cells) that corresponds to the maximum value of the distribution. In a more geophysical description, it should correspond with the distance from a given cell in which the probability of having an earthquake is higher.

Two options have been considered in this work: either the location at which the earthquakes have occurred is the one with higher probability, or the most prone for earthquakes to happen is defined by the distance from a given cell to the nearest fault/seismic source \((d_f \text{ in this work})\).

For areas in which the tectonic structures are only present in part of the region, an hybrid approach may be used by calculating the mean distance between the faults and using a cut-off distance. That is, if the distance from the cell to the nearest fault is higher than the cut-off distance then the location at which earthquakes would happen in that cell is the one with the highest probability \((\mu = 0)\)

2.1.2 Second moment of the distribution, \( \sigma \)

The standard deviation of the Gaussian distribution, \( \sigma \), is the width or dispersion of the values of the distribution around the mean value. That is, how far one might expect to find earthquakes around the most probable value. We have considered that this second parameter is related to the accuracy of earthquake’s epicentre measurement. This means that it would depend on the methodologies and instrumentation used for the calculation of the epicentre, and thus, depend on both the year and the location of the catalogue.

As in the previous section, two different options regarding the epicentre uncertainty, \( \varepsilon \) in this work, have been taken into account: either it depends on the year of occurrence, or it is constant and computed as the mean value of the epicentral uncertainty for all the events.

Three different models have been proposed to account for the variations in these parameters (Table 1), where \( \varepsilon_1 \), refers to the different epsilon values depending on the period of the catalogue, and \( \varepsilon_2 \) refers to the fixed value for all the catalogue:
Table 1. Models for the seismic activity smoothing.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>$d_{fi}$</td>
<td>0 km</td>
<td>0 km</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>$\varepsilon_1$</td>
<td>$\varepsilon_1$</td>
<td>$\varepsilon_2$</td>
</tr>
</tbody>
</table>

2.2 Foreshock and aftershock smoothing

Firstly, the spatial grid is defined by creating a rectangle spanning the maximum and minimum longitudes and latitudes of the catalogue with the desired resolution. Then, all the events of the catalogue must be assigned to each cell. This is done by calculating the minimum distance of each event to all the cells in the spatial grid. The most important step, regarding the activity rate calculation, is the identification of the seismic clusters present in the area in the selected period of time. As indicated in the introduction, we do not pretend to remove the foreshocks and aftershocks but to identify the main event and all related events in the corresponding cluster. There are several options for this task: a) using the Reasenberg and Jones (1989) algorithm -RJ by applying the ZMAP software (Wiemer, 2001), or b) or using the Afteran -A- algorithm (Musson, 1999) or the Gardner and Knopoff (1974) -GK74- declustering algorithms by applying the Python libraries included in OpenQuake’s (Pagani et al., 2014).

These algorithms will flag each event from each cluster with an identifier which will be added as a column to the catalogue, and the events that do not belong to any series will have a value of zero for this field. In order to decide which algorithm carries out the process better on the catalogue, a comparison between them has been made using default parameters.

Then, once the events belonging to each cluster have been identified (a comparison between different algorithms and the selection criteria will be explained in Spain’s case study) so each event has a label identifying to which cluster they belong to (an example can be seen in Figure 1). The procedure explained by Taroni and Akinci (2021) has been adapted to work with as many clusters can be found in each spatial grid cell.
To do this, all the events belonging to each cluster are counted (Eq. 7):

\[ c_j = \sum_{1}^{q} \cdot 1 \]

where the sum goes over the \( q \) events belonging to a cluster \( c_j \). It can be seen that if an event does not belong to any cluster (i.e., the cluster label for that event is set to zero) then \( c_j \) equals 1. The weighted counts for each spatial grid cell are calculated as the summation of all the events over the different clusters (Eq. 8):

\[ k_i = \sum_{1}^{j} \left( \frac{1}{c_j} \sum_{1}^{m} \cdot 1 \right) \]

where \( k_i \) is the weighted count of events inside the cell \( i \), the first sum goes over the number of clusters, \( j \), inside said cell, and the second summation goes over each event, \( m \), of the cluster \( j \) that is inside the cell \( i \).

For instance, if inside a cell there are 20 events that belong to a cluster composed of 100 events in total and 13 events that do not belong to a cluster, the weighted number of events for that cell will be:

\[ k_i = 13 + \frac{20}{100} = 13.2 \]
2.3 Seismic activity rate computation

With the two smoothing stages exposed before, we will consider both the nature and source of the earthquakes and the uncertainty related to the earthquake location. Thus, the seismic activity rate is calculated as the product of the weighted counts and the smoothing kernel (Eq. 9):

\[ \lambda_i = w_i \cdot k \]  \hspace{1cm} (9)

where \( w \) is an \( n \times n \) matrix, being \( n \) the number of cells of the spatial grid, that for each cell, \( i \), contains the \( n \) values of the smoothing kernel associated to the cell. On the other hand, \( k \) is a vector containing the weighted count for each cell \( i \) as defined in Eq. 8. So, for each cell the vector product between the smoothing kernel \( (w_i \) can be seen as vector) and the weighted count is done. This means all cell counts are added in each cell activity rate computation and the smoothing function works similarly to the cut-off distance from Frankel (1995).

2.4 Exceedance probability calculation

The annual exceedance probability of a given PGA has been obtained by developing a Python script based on Openquake (Pagani et al., 2014). Two models have been tested for the computation of the needed b-value: a fixed (time-independent) b-value assigned from the tectonic zones of each country and a gridded (time-dependent) b-value calculated using the methodology proposed by Montiel-López et al. (2023). In both cases, another Python script was developed to obtain a gridded time-dependent seismic activity rate for moment magnitudes greater than or equal to 4.0 Mw for each cell.

The Akkar and Bommer (2010) empirical equation has been used as a ground motion prediction model since it is appropriate for Mediterranean regions as Spain, whereas the Akkar et al. (2014) ground motion empirical equation has been used for Central Italy since it was used in the EHSM20 (Danciu et al., 2021). An example of the workflow for this work is represented in Figure 2.

Figure 2. Workflow diagram for the exceedance probability computation. Examples of the main inputs are given as the three spatial mappings of a) b-value, b) seismic activity rate and c) tectonic zones’ parameters.
Since our goal is to investigate if the temporal changes in the seismic activity and b-value can be observed as a trend in the PSHA results that can be used as an indicator for OEF, the temporal evolution of the annual probability of exceedance (PoE) of a background PGA corresponding to 475 years return period (i.e., 0.002 PoE) has been computed as a time-dependent value. The results are expressed as a relative change (RC, Eq. 10) when compared with the background annual exceedance probability (long-term value).

\[
RC = 100 \cdot \left( \frac{PoE}{0.002} - 1 \right)
\]

This background value is also related to the year of the seismic hazard updates in the seismic normative. In order to save computation time, the annual exceedance probability is only calculated for the main cities located inside the spatial grid.

Additionally, we have computed the annual variation of the RC in the exceedance probability \((RC_i - RC_{i-12})\), with \(i\) the computed month) and the monthly variation \((RC_i - RC_{i-1})\) to investigate if any of these metrics is more effective as an indicator for OEF.

3 Case studies

3.1 Central Italy

3.1.1 Catalogue preparation and parameters for computation

As mentioned before, since Central Italy is a very active region, this case study will help us to decide which of the models (Table 1) performs better. Central Italy (Abruzzo, Campania, Lazio, Marche, Moise, Toscana and Umbria) is a region where several high magnitude earthquakes and significant seismic series have occurred in the past. The main focus is on L’Aquila, where a 6.1 Mw earthquake (Table 2) struck the area in 2009 and caused 309 deaths and 1500 injured. Therefore, the city of L’Aquila has been selected as the site for the hazard computation.

Table 2. L’Aquila earthquake data and distance to hazard computation site.

<table>
<thead>
<tr>
<th>Location</th>
<th>Lat. (°N)</th>
<th>Long. (°E)</th>
<th>Depth (km)</th>
<th>Mw</th>
<th>Int. (EMS-98)</th>
<th>Date</th>
<th>Epicentral distance (km) to L’Aquila</th>
</tr>
</thead>
<tbody>
<tr>
<td>L’Aquila (AB)</td>
<td>42.334</td>
<td>13.334</td>
<td>10</td>
<td>6.1</td>
<td>VIII</td>
<td>6 Apr 2009</td>
<td>5.2</td>
</tr>
</tbody>
</table>

Figure 3 shows location of the area of study (a rectangular area that with longitudes from 11.392 to 15.372° E and latitudes from 40.374 to 44.354° N) and the tectonic zones and main faults as defined in the EHSM20 (Danciu et al., 2021). For this area the Italian HORUS catalogue has been used (Lolli et al., 2020) as it has been homogenised and comprises events from
1960 to 2012 (in order to study this particular seismic series). It has a total of 49112 events with maximum depth of 30 km and maximum magnitude of 6.1 Mw.

![Image](https://example.com/image.png)

**Figure 3.** Map showing the tectonic zones as defined by Danciu et al. (2021) in Central Italy. The star marks the epicentre of L’Aquila earthquake (Table 2) and the numbers under the acronyms in each tectonic zone show the tectonic b-value.

The events that are not earthquakes (such as quarry blasts, eruptions, explosions, etc.) have been filtered out from 2012 on (as the catalogue has such information). In order to consider the influence of such events prior 2012 the area that has been selected for this study does not show important changes in the b-value according to the results of Taroni et al. (2022). In this case the catalogue has not been declustered, but the clusters have been identified by using Gardner and Knopoff (1974) algorithm and this information has been used to weight down the influence of the non-independent events towards the seismic parameters’ computation.

A spatial cell grid of 40000 points (200 x 200) has been created spanning the above longitude and latitude ranges. The completeness magnitude (Table 3), it has been retrieved from Taroni et al. (2021).

**Table 3.** Completeness magnitude values proposed by Taroni et al. (2021).

<table>
<thead>
<tr>
<th>Year</th>
<th>1960</th>
<th>1980</th>
<th>1990</th>
<th>2003</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completeness magnitude (Mw)</td>
<td>4.0</td>
<td>3.0</td>
<td>2.5</td>
<td>2.1</td>
<td>1.8</td>
</tr>
</tbody>
</table>

The epicentral uncertainty has been computed by calculating the mean value of the maximum horizontal error (the maximum between the latitude and longitude errors (Klein, 2002, p. 88)) on the Italian Parametric Earthquake Catalogue 4.0 (CPTI15) (Rovida et al., 2022, 2020). A mean value of 6 km is obtained for all the periods considered in this study. Additionally, a
second value of 30 km has been selected, following the work of Taroni et al. (2021), in order to define the three models. Table 4 contains the three models considered for this test.

Table 4. Models for the activity smoothing in Central Italy.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Model 1t</th>
<th>Model 2t</th>
<th>Model 3t</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mu)</td>
<td>(d_{f_i})</td>
<td>0 km</td>
<td>0 km</td>
</tr>
<tr>
<td>(\sigma)</td>
<td>6 km</td>
<td>6 km</td>
<td>30 km*</td>
</tr>
</tbody>
</table>

*From Taroni et al. (2021).

3.1.2 Results

Figure 4 (Model 1t) presents a moderate increase in the annual exceedance probability (25%) one month before L’Aquila earthquake occurred, and not only the annual, but also the monthly variations of relative change attained values higher than 35%. Figure 5 (Model 2t) shows a similar trend in all the metrics as the previous model with a slightly lower value for the exceedance probability change before the earthquake (22%) and the annual and monthly variations (32%). The Model 3t (Figure 6) provides the lowest values for the metrics (-3%, 4% and 3%, respectively).

![Figure 4. Model 1t. Relative change (RC) of the annual exceedance probability (top) and its annual and monthly variation (bottom).](https://doi.org/10.5194/egusphere-2023-2818)
Figure 5. Model 2t. Relative change (RC) of the annual exceedance probability (top) and its annual and monthly variation (bottom).

Figure 6. Model 3t. Relative change (RC) of the annual exceedance probability (top) and its annual and monthly variation (bottom).
Given that the main objective is to be able to perform OEF in the area of study, the Model 1t (Figure 4) performs the best in terms of exceedance probability change and its annual and monthly variations since the obtained values are the highest one month before the main shock (when compared with the models 2t and 3t).

### 3.2 Southeastern Spain

#### 3.2.1 Catalogue preparation and parameters for computation

The south and south-east of Spain are the regions with a higher seismic hazard in Spain (IGN-UPM Working Group, 2013; Kharazian et al., 2021) with values reaching 0.23 g for a 10% probability of exceedance in 50 years (i.e., return period of 475 years). Although Spain is a moderate to low seismic region compared to other European countries such as Italy or Greece, it has been exposed to several damaging earthquakes in the past being the most representative the 1829 Torrevieja earthquake and the 1884 Arenas del Rey earthquake, both with a maximum intensity IX-X. Additionally, in the last 25 years, south-east Spain has suffered seven earthquakes with Mw above 4.5 (Table 5 and Figure 7), being the 2011 Lorca earthquake the most relevant since it was the most recent earthquake causing damage to buildings and injuries to the population. The seismicity is usually very shallow (mainly lower than 10 km). Three main cities (Murcia, Lorca and Vera from North to South) have been chosen as representative of the region in terms of decreasing seismic hazard values for a 475 years return period.

Table 5. Damaging earthquakes in the last 25 years and epicentral distance to some chosen cities in the area of study.

<table>
<thead>
<tr>
<th>Location</th>
<th>Lat. (°N)</th>
<th>Long. (°E)</th>
<th>Depth (km)</th>
<th>Mw</th>
<th>Int. (EMS-98)</th>
<th>Date</th>
<th>Epicentral distance (km) to</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Murcia</td>
</tr>
<tr>
<td>N. Mula (MU)</td>
<td>38.0963</td>
<td>-1.5014</td>
<td>1.1</td>
<td>4.9</td>
<td>VI</td>
<td>2 Feb 1999</td>
<td>34.5</td>
</tr>
<tr>
<td>S. Gergal (AL)</td>
<td>37.0931</td>
<td>-2.5379</td>
<td>0.8</td>
<td>4.6</td>
<td>V</td>
<td>4 Feb 2002</td>
<td>159.4</td>
</tr>
<tr>
<td>SW. Bullas (MU)</td>
<td>37.8925</td>
<td>-1.8353</td>
<td>1.2</td>
<td>5.0</td>
<td>V</td>
<td>6 Aug 2002</td>
<td>62.9</td>
</tr>
<tr>
<td>NW. Aledo (MU)</td>
<td>37.8535</td>
<td>-1.7555</td>
<td>10.9</td>
<td>4.8</td>
<td>VII</td>
<td>29 Jan 2005</td>
<td>20.6</td>
</tr>
<tr>
<td>Lorca (MU)</td>
<td>37.7175</td>
<td>-1.7114</td>
<td>4.0</td>
<td>5.1</td>
<td>VII</td>
<td>11 May 2011</td>
<td>59.5</td>
</tr>
</tbody>
</table>

In order to compute the seismic activity rate to be used in a PSHA, first we need to compile an homogeneous and complete seismic catalogue in the influence area, needed for the chosen locations. This catalogue comprises all the events from 1396 to August 2023 in south-eastern Spain inside the tectonic zones of Eastern Betic Shear Zone (ZCBOR), Eastern Inner Betics (BIOR), Valencian Plateau and Alicante’s Prebetic (PVPA), Murcian Prebetic (PM), Sierra Nevada-Filábrides and Guadix-Baza (SNFCGB), Central Inner Betics (BIC), Southern Plateau (MS), Cazorla-Segura and Albacete’s Prebetic (CSPA), Central Guadalquivir and Algerian-Balearic Basin (CAB), as defined by García-Mayordomo (2015) to create the Spanish Seismic Hazard Map (IGN-UPM Working Group, 2013).
Figure 7. Map showing the tectonic zones as defined by García-Mayordomo (2015) in the South-eastern Spain region. The numbers indicate the order in which the earthquakes appear in Table 4 and the numbers under the acronyms in each tectonic zone show the tectonic b-value.

The catalogue contains a total of 20279 events that span from 1396 to August 2023. Their moment magnitudes range from 0.1 to 6.8 Mw after being homogenised using the magnitude correlation equations for this region (IGN-UPM Working Group, 2013). Their depth goes up to 90 km, although in the calculations only the earthquakes shallower than 30 km are considered (which amount for a total of 20168 events). A rectangle defined by the corners $\left[(-3.3750^\circ, 36.3938^\circ), (-0.0817^\circ, 39.1000^\circ)\right]$ in (longitude, latitude) is created with a total of 40401 cells ($201 \times 201$) to contain all the events. This grid involves a step of 0.01647° for the longitude and 0.01353° for the latitude.

Table 6 presents the number of clusters and the events in clusters for the whole seismic catalogue. As can be seen, there are significant differences when applying the three algorithms, not only in the number of clusters but also in the number of events inside said clusters. The RJ algorithm classifies a total of 652 clusters in the catalogue while GK74 detects 1012 clusters. The A algorithm is in between, identifying 863 clusters. Considering that Cabañas et al. (2011) carried out a detailed study on the 2011 Lorca’s earthquake seismic series, we have used their results to validate the best algorithm. According to Cabañas et al. (2011) the cluster corresponding to Lorca’s series is composed of 143 events (including the foreshock, the main shock and the aftershocks). As can be seen from Table 6, the GK74 algorithm seems to work better on the data, since it identifies 136 events (less than a 5% difference compared with Lorca’s series).

Zaliapin and Ben-Zion (2020) also pointed out these problems with the identification of aftershocks and main shocks and proposed an algorithm to discriminate between background and clustered events by randomly thinning a complete catalogue by removing nearest-neighbour earthquakes. Moreover, Anderson and Zaliapin (2023) examine the effect on the hazard estimation...
They conclude that hazard estimates are most sensitive to the catalogue thinning near the aftershock zone, and less sensitive elsewhere.

**Table 6.** Comparison of cluster identification and total events inside clusters among three declustering algorithms: an analysis using Lorca’s seismic series.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>RJ</th>
<th>A</th>
<th>GK74</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of clusters</td>
<td>652</td>
<td>863</td>
<td>1012</td>
</tr>
<tr>
<td>Events in clusters</td>
<td>7143</td>
<td>14822</td>
<td>7552</td>
</tr>
<tr>
<td>Events inside Lorca’s series</td>
<td>123</td>
<td>3394</td>
<td>136</td>
</tr>
</tbody>
</table>

Our catalogue starts in the historical period (when there was a lack of instrumentation and procedures to accurately locate the epicentres and evaluate the magnitude of the earthquakes) and ends in the present days. This implies that not all the magnitude values will be complete in the catalogue (low magnitudes are missing in the historical period) and the location uncertainty will also differ depending on the year of detection. First, we will characterise the completeness magnitude - the minimum magnitude from which the catalogue is not missing any record- and periods for the Spanish seismic catalogue.

Gaspar-Escribano et al. (2015) defined different threshold magnitudes for different regions around Spain. Table 7 compiles the values obtained for our study area. As we can see the lowest magnitude interval, represented by its class mark 3.25 Mw, is complete after 1978, so after this year all the earthquakes with magnitude higher or equal than 3.25 Mw will appear in the catalogue.

**Table 7.** Completeness magnitude for each period according to Gaspar-Escribano et al. (2015).

<table>
<thead>
<tr>
<th>Completeness magnitude (Mw)</th>
<th>[3.0 - 3.5)</th>
<th>[3.5 - 4)</th>
<th>[4 - 4.5)</th>
<th>[4.5 - 5)</th>
<th>[5, 5.5)</th>
<th>[5.5 - 6)</th>
<th>[6 - 6.5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>1978</td>
<td>1975</td>
<td>1908</td>
<td>1883</td>
<td>1800</td>
<td>1520</td>
<td>1048</td>
</tr>
</tbody>
</table>

Magnitudes below 3.25 Mw will have different completeness years so from 1978 to 2023 the completeness magnitude has been calculated by averaging the cut-off magnitude results available from González (2017) over the Eastern Betic Shear Zone (Table 8).

---

https://doi.org/10.5194/egusphere-2023-2818
Preprint. Discussion started: 2 January 2024
© Author(s) 2024. CC BY 4.0 License.
Table 8. Averaged completeness magnitude for each period using the results from González (2017).

<table>
<thead>
<tr>
<th>Completeness magnitude (Mw)</th>
<th>3.4</th>
<th>3.3</th>
<th>3.0</th>
<th>2.9</th>
<th>2.3</th>
<th>2.1</th>
<th>1.9</th>
<th>1.8</th>
</tr>
</thead>
</table>

The uncertainty of the epicentral location (ε) varies with time, showing a decreasing behaviour since the techniques and instrumentation have been continuously improved. The appropriate estimation of this uncertainty is very important in order to correctly assign the location of each earthquake to a given seismic source.

Following the research of Peláez and López (2002), the ε values for each period are presented (Table 9). The period 1990-2023 has been obtained as the average epicentral uncertainty using the data provided by the national seismic network. A second fixed ε value of 7.5 km has been computed as the mean value for all the uncertainties in the catalogue.

Table 9. ε values proposed by Peláez and López (2002).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ε (km)</td>
<td>20</td>
<td>15</td>
<td>10</td>
<td>5</td>
<td>2.5*</td>
</tr>
</tbody>
</table>

* Calculated as the average epicentral uncertainty for the 1990-2023 period events in our catalogue.

The three models to be evaluated are presented in Table 10. Fixed model implies a fixed tectonic b-value while time-dependent model indicates a time-dependent b-value.

Table 10. Models for the exceedance probability calculation in south-eastern Spain.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Fixed model</th>
<th>Time-dependent model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1f</td>
<td>Model 2f</td>
</tr>
<tr>
<td>µ</td>
<td>d_f1</td>
<td>0 km</td>
</tr>
<tr>
<td>σ</td>
<td>Table 9</td>
<td>Table 9</td>
</tr>
</tbody>
</table>

3.2.2 Results

After computing the time-dependent PSHA for the different models shown in Table 10, we have observed that Model 1t provides better results than the others, similarly to what we obtained for Italy. Therefore, with the exception of the fixed models (models 1f, 2f and 3f, where we present a general comparison between all the models), we will present the results of Model 1t in detail along with the general comparison. The results for the rest of the models can be found in the Appendix section.
The time-dependent PSHA (PGA for a return period of 475 years) has been computed using the proposed methodology, for the compiled non-declustered catalogue, in one-month increments starting from 1990. The b-value is constant and given by the zonation proposed by García-Mayordomo (2015) (models 1f, 2f and 3f). This background PGA value will be a long term PSHA which varies each time that the seismic normative has changed in Spain. Our first background PGA corresponds to the PSHA computed using a catalogue with the same length as the one used for the NCSE-94 (1994). This background PSHA value will be used from 1990 to December 1998 (since the next code updated the seismic hazard map using a seismic catalogue up to 1999). The second background value will correspond to the PSHA computed with a catalogue of the same length as the one used for the NCSE-02 (2002) and it will be used from 1999 to May 2011 (since that is the year when the seismic hazard map was updated again). Finally, the last background value, corresponds to PSHA computed using a catalogue of the same length as the one used for the current seismic hazard map for Spain (IGN-UPM Working Group, 2013). This last value is used from June 2011 to August 2023.

Figure 8 represents the temporal evolution of the results for each tested model. The vertical lines correspond to the main earthquakes from Table 5 and with epicentral distance lower than 75 km from the chosen city, as the contribution of the events further than this distance is negligible towards hazard computation. As can be seen, the behaviour is similar for the three models. The exceedance probability decreases continuously since 1990 except for Lorca in 1997, and 2006; Murcia in 1996; and Vera in 1994 and 1999. These variations are due to seismic activity changes, but they do not appear to be related with the occurrence of any of the main earthquakes from Table 1. On the other hand, all the models provide similar changes in the exceedance probability although in the city of Lorca, Model 3f is the one with the lowest percentage change and Model 2f is the one with the highest percentage change.
Therefore, the results indicate that using a non-declustered catalogue with a fixed tectonic b-value and a temporal activity rate for PSHA is not appropriate to correctly identify an increase to be used for OEF.

**Time-dependent PSHA using time-dependent b-value (time-dependent model)**

In this section, models 1t, 2t and 3t (Table 10) are tested using the same PGA background values explained previously. As we can see from Figure 9 the annual probabilities decrease prior to the Mula earthquakes for Lorca site. However, close to the earthquake, it shows a slight increase even in Vera site, although it is 101.4 km away from the earthquake’s epicentre. In Murcia site, it decreases continuously until one month before the earthquake, when it shows a sharp increase (from -75% to almost 10% in the change of exceedance probability). After the Mula earthquake the change in probability exceedance remains higher than 20% (even increasing up until 50% in the case of Lorca and 100% in the case of Murcia) for both Lorca and Murcia sites until Lorca earthquake happens. In Vera site, this parameter oscillates about the baseline. After 2011 it increases steadily in Vera, whereas in Lorca and Murcia stays constant after 2019. On the other hand, Model 2t and Model 3t (Figure 10) demonstrate a similar behaviour as Model 1t, although Model 3t showcases higher exceedance probability before Lorca 2011 earthquake.
Figure 9. Relative change (RC) of the annual exceedance probability and corresponding uncertainty for Model 1t in Lorca, Murcia and Vera (from top to bottom).
Figure 10. Mean value of the relative change (RC) of the annual exceedance probability for models 1t, 2t and 3t in Lorca, Murcia and Vera (from top to bottom).

Figure 11 shows that one month before Mula earthquake a slight change in the annual variation (over 15%) can be seen for Lorca site. After the earthquake, the PSHA remains high for a few months and it oscillates around the baseline after Gergal and Bullas earthquakes. A slight change of the values can be seen prior to Lorca earthquake in Lorca site (around 15% in absolute value), but the tendency is not as clear as in the Italy case study. This means, no clear conclusions can be drawn.
Figure 11. Annual and monthly variations of the relative change of the annual probability of exceedance for Model 1t in Lorca, Murcia and Vera (from top to bottom).

Then, Figure 12 and Figure 13 sum up the results for the three models regarding the annual and monthly change in variations in the change of the exceedance probability. It can be seen that Model 1t seems to perform better as it presents greater changes before and after the selected earthquakes.
Figure 12. Monthly variations of the relative change of the annual probability of exceedance for models 1t, 2t and 3t in Lorca, Murcia and Vera (from top to bottom).
Figure 13. Annual variations of the relative change of the annual probability of exceedance for models 1t, 2t and 3t in Lorca, Murcia and Vera (from top to bottom).

□ Effect of the declustering on the results

In order to compare the effect of the catalogue declustering on the results, the Model 1t has been plotted using both the declustered catalogue (with a total of 13841 events) and the full catalogue (with the clusters identified and weighted down accordingly).

Figure 14 represents the changes in the annual exceedance probability when using Model 1t. As can be seen, the results using a non-declustered catalogue provide, in general, lower changes in the exceedance probability. However, it seems to be more stable regarding the magnitude of the oscillations. This can be seen in Lorca site, where the exceedance probability changes are greater during the 2002-2011 period.
4 Conclusions

This methodology considers the influence of all the events in the seismic clusters and also the location of the seismic sources (corresponding active faults) for seismic activity rate smoothing and b-value computation, showing that when computing a time-dependent PSHA the use of a non-declustered catalogue will provide similar results to using a declustered catalogue. Therefore, if we compute the changes of the annual probability of exceedance for a given PGA value (fixed as a background value which may change according to the updates in the seismic normative), we will be able to provide how this probability is changing with time. Although our results are not significant to relate these changes to the occurrence of a main earthquake for low to moderate seismicity areas, the methodology can be useful for other countries with a higher seismicity, or in the future if new significant earthquakes occur in the studied region. In this case, for Central Italy both the change of the exceedance probability and the annual and monthly changes of this parameter show important changes that could enable OEF.

The oscillation of the change in the annual probability of exceedance (increases and decreases) can be more accurately described using a spatially gridded time-dependent b-value instead of a fixed one for each tectonic zone, this can be seen when comparing Figure 8 with Figure 10. Therefore, we suggest using always spatially gridded b-values for the corresponding period (time-dependent) when computing the background PGA value and the corresponding changes in the annual probability of exceedance in the time-dependent PSHA.
Regarding which of the proposed models can be more effectively used to describe these changes, we have to consider several factors. For instance, if we compare how close the results are to the values calculated for the current national seismic hazard maps, then Model 1t is the one that performs best. This means that considering proximity to faults as well as accuracy in epicentral measurements of the catalogue has a positive impact towards seismic hazard assessment.

This methodology benefits from complete catalogues in zones with increased seismicity - assuring more stability in the b-value calculation and even weekly or daily updates in the exceedance probability - and well-defined seismicity sources, where the seismicity smoothing is accurate. Figure 14 in a sense, shows this result, as a declustered catalogue has fewer events than a non-declustered catalogue (but with clusters weighted down), and the latter, more stable results for Model 1t.

The result may indicate that after the Mula earthquake the PSHA kept high in the region and it did not decrease until the occurrence of the Lorca earthquake. However, the continuous increase of the PSHA in Vera after the Lorca earthquake cannot be directly related to a potential upcoming earthquake similar to the one from Lorca. Therefore, we need more time and data to confirm this.

*Code and data availability.* The data and code used in this study is available upon request.
Appendix A: Other results for Southeastern Spain

A1 Time-dependent PSHA using time-dependent b-value (time-dependent model)

Figure A1. Relative change (RC) of the annual exceedance probability and corresponding uncertainty for Model 2t in Lorca, Murcia and Vera (from top to bottom).
**Figure A2.** Relative change (RC) of the annual exceedance probability and corresponding uncertainty for Model 3t in Lorca, Murcia and Vera (from top to bottom).
Figure A3. Annual and monthly variations of the relative change of the annual probability of exceedance for Model 2t in Lorca, Murcia and Vera (from top to bottom).
Figure A4. Annual and monthly variations of the relative change of the annual probability of exceedance for Model 3t in Lorca, Murcia and Vera (from top to bottom).
A2 Effect of the declustering on the results

Figure A5. Model 2t. Relative change (RC) of the annual exceedance probability for a non-declustered and a declustered catalogue in Lorca, Murcia and Vera (from top to bottom).
Figure A6. Model 3t. Relative change (RC) of the annual exceedance probability for a non-declustered and a declustered catalogue in Lorca, Murcia and Vera (from top to bottom).
Author contributions. Conceptualisation and original idea: DML and SM; methodology: DML, SM, JJGM, IG, AK, JLSL, JAHT, AGV and GOS; DML wrote the code and tested the different components; DML and SM performed the data curation; writing the original draft: DML, SM and JJGM; writing review and editing: DML, SM, JJGM, IG, AK, JLSL, JAHT, AGV and GOS. All authors have read and agreed to the published version of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This study was supported by the Spanish Government through research project PID2021-123135OB-C21, and Research Group VIGROB-116 (University of Alicante) and by the Conselleria de Innovación, Universidades, Ciencia y Sociedad Digital de la Generalitat Valenciana through research project CIAICO/2022/038.
References


García-Mayordomo, J.: Creación de un modelo de zonas sismogénicas para el cálculo del mapa de peligrosidad sísmica de España, https://www.igme.es/publicaciones/publiFree/Creaci%C3%B3n%20de%20un%20modelo%20de%20zonas%20sismog%C3%A9nicas%20de%20Espa%C3%B1a.pdf, 2015.


