First of all, thank you, Prof. Chan, for your comments and suggestions in the review of the manuscript. In the following document, all the questions will be answered in order and quoting each item of the review using the yellow highlighted and bold font style. The numeration has changed and the figures' and tables' numeration refers to the last version of the manuscript.

The Italy case not discussed: The abstract and various sections of the manuscript mention the application of this approach to the L'Aquila, Italy, case. However, I am unable to find any corresponding results or discussion.

The corresponding paragraphs (section **3.1.2 Results** and **4.0 Conclusions**) have been modified in order to discuss the results from Central Italy case.

3.1.2 Results

*"**Figure 5** (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 variation of relative change reaches values higher than 35%. This sudden change is most probably due to the foreshock activity that preceded the mainshock, as a 4.1 M_L ground motion occurred on 30 March 2009. **Figure 6** (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 7**) provides the lowest values for the metrics (-3%, 4% and 3%, respectively). After this increase the RC slowly decreases over time for all three models."*

4.0 Conclusions

[…]

*"Regarding which of the proposed models can be more effectively used to describe these changes, we have to consider several factors. One could be how close are the computed PGA values to the national seismic hazard maps for each country. In the case of Central Italy models 1t, 2t and 3t provide the following background PGA values: 340.61, 359.72 and 334.28 cm/s^2. The ESHM20 model (Danciu et al., 2021) computes 334.38 cm/s^2. The closest match would be Model 3t followed by Model 1t. However, by looking at **Figure 5** and **Figure 7**, it can be seen that the Model 3t seems to be less responsive to the seismic activity than Model 1t, as the monthly and annual RC variations suggest (Model 1t monthly and annual variations are 4.5 times higher than Model 3t variations'). With this information Model 1t seems appropriate for the purpose of this work.*

*In general, this methodology benefits from complete catalogues in zones with increased seismicity - assuring less uncertainty in the b-value computation - and well-defined seismicity sources, where the seismicity smoothing is accurate. **Figure 16** shows this result, as the non-declustered catalogue (with weighted down cluster events) has less RC uncertainty and enables the use of the foreshocks in daily to weekly time scales.*

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 of Spain. As we saw, for Central Italy both the annual and monthly changes of the exceedance probability show important variations related to the foreshock activity preceding L'Aquila earthquake. This could be useful for OEF."

[…]

Definition of smoothing kernel: In this study, the smoothing kernel is determined by the average distance between all events surrounding an earthquake and the precision of the epicenter's location (Lines 68-70). When defining the second moment of the distribution (Sec. 2.1.2), this parameter is solely attributed to the precision of the earthquake epicenter measurement. However, I anticipate that the distance between all events is equally important for this parameter.

The smoothing kernel section and subsections have been rewritten in order to explain some of the features as well as to address this matter. Find here, the part of interest as well as the figure comparing the different cases (that arise from the different geometries and relations between faults). The rest of the modifications can be found in the modified version of the manuscript attached to the response.

- **Second moment of the distribution, σ** has been renamed to **Geophysical meaning of the parameter σ** and the contents of the subsection have been rewritten in order for them to be more clearly defined.

"This parameter accounts for the dispersion of the values of the distribution around the mean value. That is to say, how far one might expect to find earthquakes around the most probable value (of distance). Therefore, 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, on both the year and the location of the catalogue.

It should be noted that σ may depend on other geophysical parameters such as the characteristics of ground, the style of faulting and/or the tectonic stress regime, to cite a few. Nevertheless, in this work only the influence of the uncertainty in the epicentre's location will be considered in the smoothing process.

*As in the previous section, two different options regarding the epicentre uncertainty, **ε**, have been considered: either it depends on the year of occurrence (**ε_1**), or it is constant and computed as the mean value of the epicentral uncertainty for all the events (**ε_2**). "*

A new subsection has been added to show examples of the smoothing kernel implementation and a figure showcasing how the smoothing works is presented:

2.1.3 Examples of implementation

"In this section, some examples of how the smoothing kernels works are shown. There are three main different manners in which this smoothing is applied:

- *Usual 1D Gaussian filter,* $\mu = 0$

This is the case when using models 2 and 3 also when the distance from the centre of the spatial grid cell in which the seismic activity rate is computed to the nearest fault is greater than $\boldsymbol{d_c}$ *as defined in the section* ***2.1.1.*** *An example can be seen in* ***Figure 1a.***

- *Single fault,* $\mu \neq 0$

When the nearest fault is closer than $\boldsymbol{d_c}$ *from the centre of the spatial grid cell then the resulting function will provide a ring-shaped smoothed activity, the width of which will depend on* $\boldsymbol{\sigma}$*. Only the section of this ring in which the fault is located will be used in the smoothing. This can be achieved by considering the* $\boldsymbol{n}$ *closest points to the spatial grid cell centre and then computing the angles to define the ring arc (****Figure 1b****).*

- *Several faults,* $\mu \neq 0$

*This case is a generalization of the former with the exception that when spatial grid cell's centre is in between faults and at similar distances, then the full ring will be used as smoothing function (****Figure 1d****). On the other hand, if the distance to both faults is similar, but the spatial grid cell's centre is not in between the faults then the resulting smoothing is a ring arc (****Figure 1c****)."*

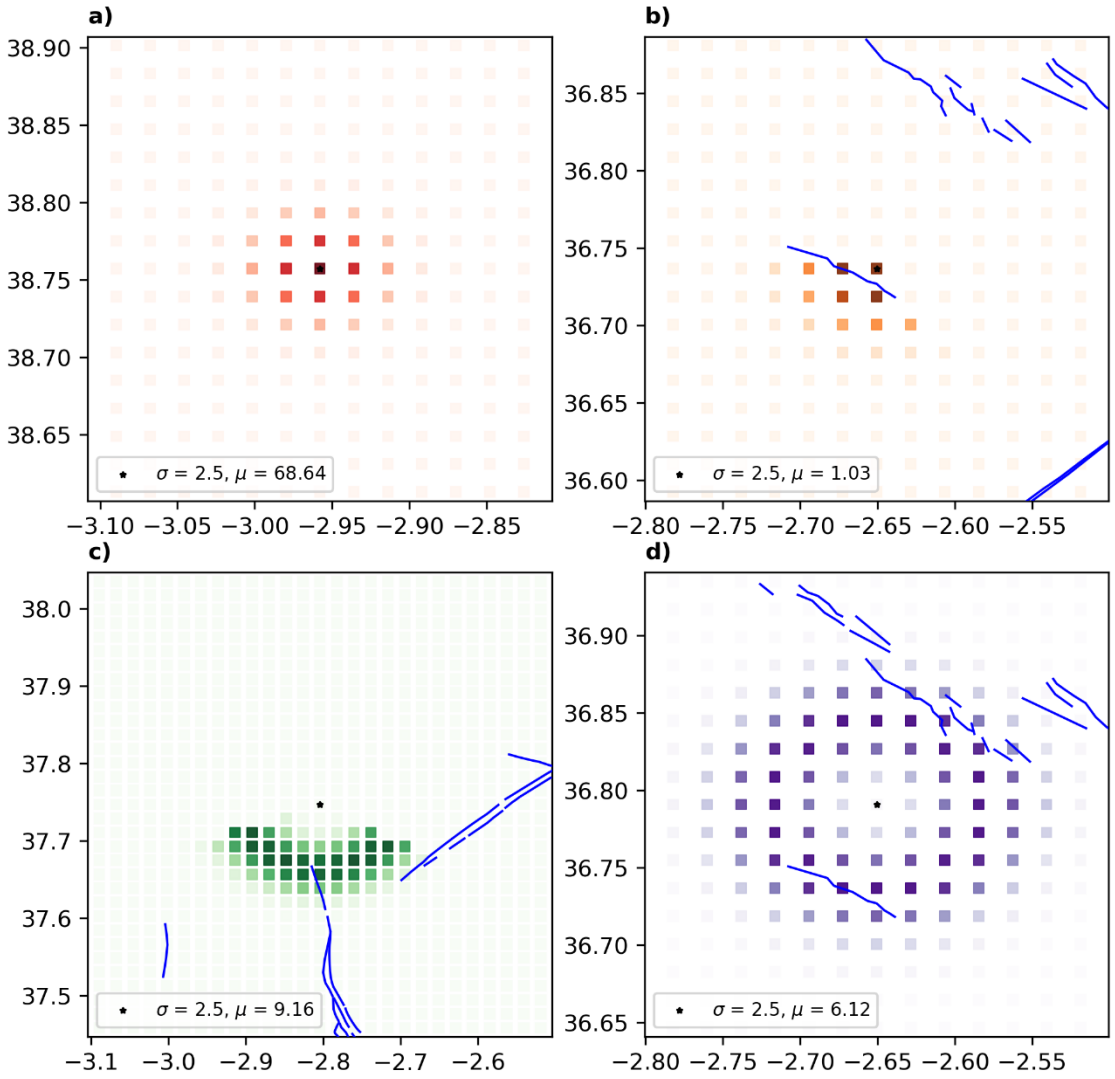

Figure 1. a) Smoothing function for $\mu = 0$. b) Smoothing function for $\mu \neq 0$ and a single fault. c) Smoothing function for $\mu \neq 0$ and several faults at similar distances. d) Smoothing function for $\mu \neq 0$ and the spatial grid cell in between faults at similar distances. The blue lines show the fault traces. In this example d_c equals 48 km.

Completeness magnitude: Based on my interpretation of Table 7, it appears that the magnitude range indicated in the top column has been complete from the year specified in the bottom column. Therefore, it should be that magnitudes of 3.0 and above are complete since 1978, rather than starting from magnitude 3.25 as stated in Lines 277-280. Additionally, the completeness magnitude typically decreases with upgrades to the seismic network, as such improvements generally enhance detection capabilities. It is customary for the completeness magnitude to remain stable for approximately a decade before decreasing sharply with a network upgrade. The gradual increase in the average completeness magnitude observed in Table 8 is unexpected. An explanation for the trend of decreasing completeness magnitude would be beneficial.

We agree that this should be the case. The data used for the tables has two different sources with an overlapping period (1962-1979). The difference in the values is due to the extent of each period. For instance, in **Table 7** the values 3.25 and 3.75 Mw

are the class marks for the years 1978 and 1975, respectively. Meanwhile, in **Table 8** the corresponding completeness magnitude value for the period from 1962 to 1979 is 3.4 Mw and has been computed as the spatial average for the South-eastern Spain area using the data from ***(González, 2017).*** This procedure combined with an inequal development of the seismic network could explain the smooth changes in the completeness magnitude.

The completeness magnitude values used in this work are the following and now appear as a single table in the manuscript, note that the values from ***Gaspar-Escribano et al. (2015)*** corresponding to the years 1975 and 1978 have not been used as the data from ***(González, 2017)*** allows to further discretize the timeline:

Years	1048-1520	1521-1800	1801-1883	1884-1908	1909-1962	1963-1979	1979-1984
Completeness magnitude [M_w]	6.25	5.75	5.25	4.75	4.25	3.4	3.3

Years	1984-1992	1993-1998	1998-2002	2002-2010	2010-2013	2013-2023
Completeness magnitude [M_w]	3.0	2.9	2.3	2.1	1.9	1.8

The abrupt change between the period 1909-1962 and 1963-1979 is to be expected, as more development in the seismic network had been made in the mid-20th century than in the early 20th century given the political and historical context of Spain.

The paragraph before the table has been modified to explain the values used in this work.

*"**Gaspar-Escribano et al. (2015)** defined different threshold magnitudes for different regions around Spain. The class marks of these magnitude intervals for the zone of interest (South-eastern Spain) have been selected as the completeness magnitudes up until 1962. From 1962 on, the completeness magnitude has been computed by spatially averaging the gridded completeness magnitude results available from **González (2017)** over the area of study. The completeness magnitude values used in this work are presented in **Table 7**."*

*Table 7. Completeness magnitude for each period according to **Gaspar-Escribano et al. (2015)** in the top tabular and the spatially averaged completeness magnitude for each period using the results from **González (2017)** in the bottom tabular.*

Completeness magnitude [M_w]	6.25	5.75	5.25	4.75	4.25
From year	1048	1521	1801	1884	1909
to year	1520	1800	1883	1908	1962

Completeness magnitude [M_w]	3.4	3.3	3.0	2.9	2.3	2.1	1.9	1.8
From year	1963	1980	1985	1993	1999	2003	2011	2013
To year	1979	1984	1992	1998	2002	2010	2013	2023

Model validation: Based on the results presented for the three models concerning both annual and monthly variations in the change of exceedance probability (as seen in Figures 12 and 13 and discussed in Lines 335-337), the authors assert that Model 1t outperforms the others. However, discerning significant differences is challenging, whether in the monthly variations of the relative change in annual probability of exceedance (Figure 12) or in the annual variations of the same (Figure 13). Moreover, I question the approach of basing model validation solely on 'greater changes before and after selected earthquakes' without incorporating statistical analyses. I believe that a more rigorous statistical evaluation is necessary to substantiate the claimed superiority of Model 1t.

We agree that the discussion of the models' performance should be further explained. To do so, the figures of the results have been modified in order to be clearer and the results have commented in their respective sections and also in the conclusions.

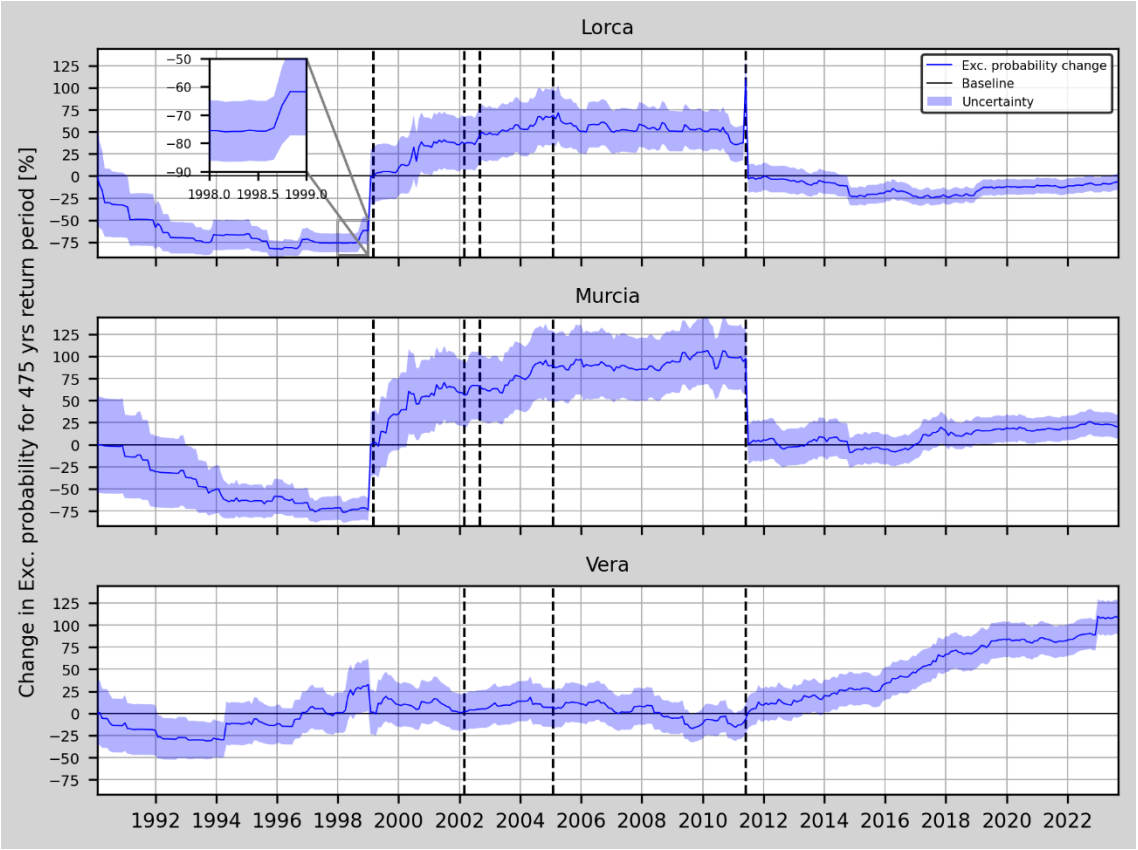

Figure 11. Relative change (RC) of the annual exceedance probability and corresponding uncertainty for Model 1t in Lorca, Murcia and Vera (from top to bottom). A zoom in on the mentioned increase in the RC during 1998 in Lorca's site appears in the upper left side of the graph.

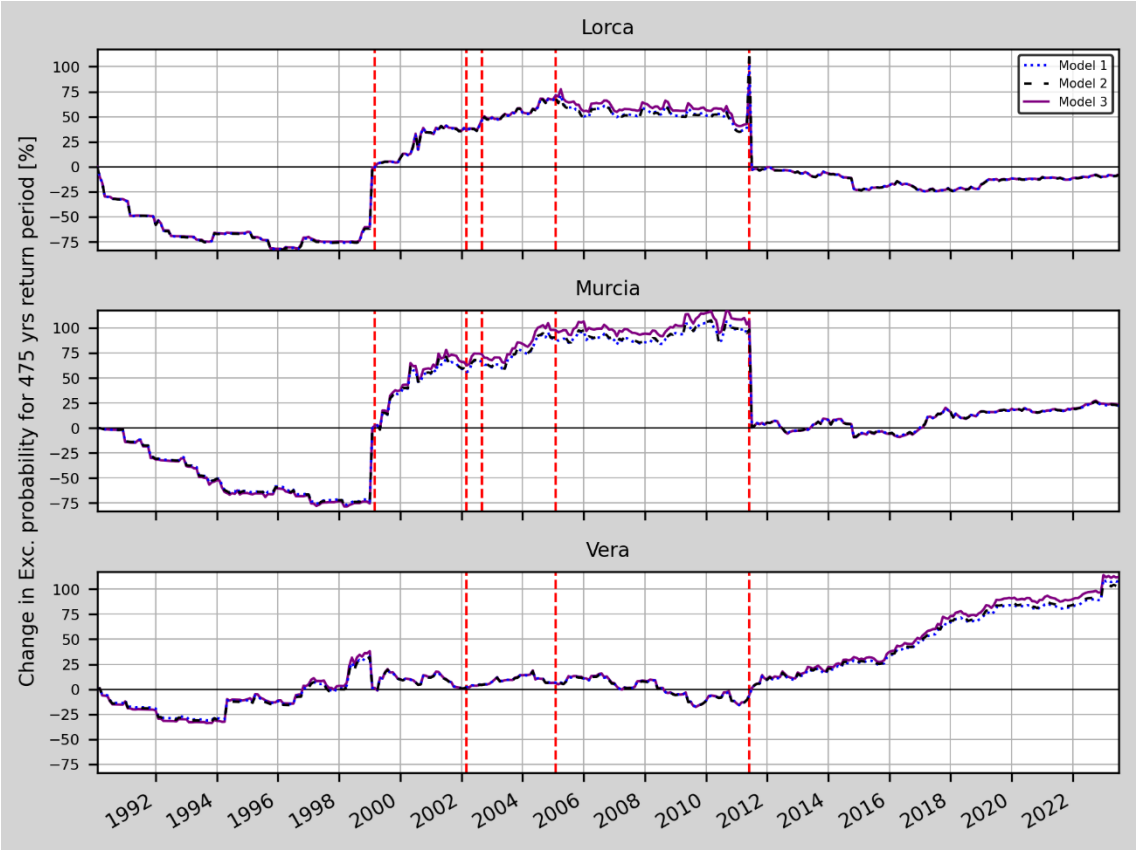

Figure 12. Mean value of the relative change (RC) of the annual exceedance probability for models 1t, 2t and 3t in Lorca, Murcia and Vera (from top to bottom). At the right side, a zoom in on the peak due to Lorca earthquake.

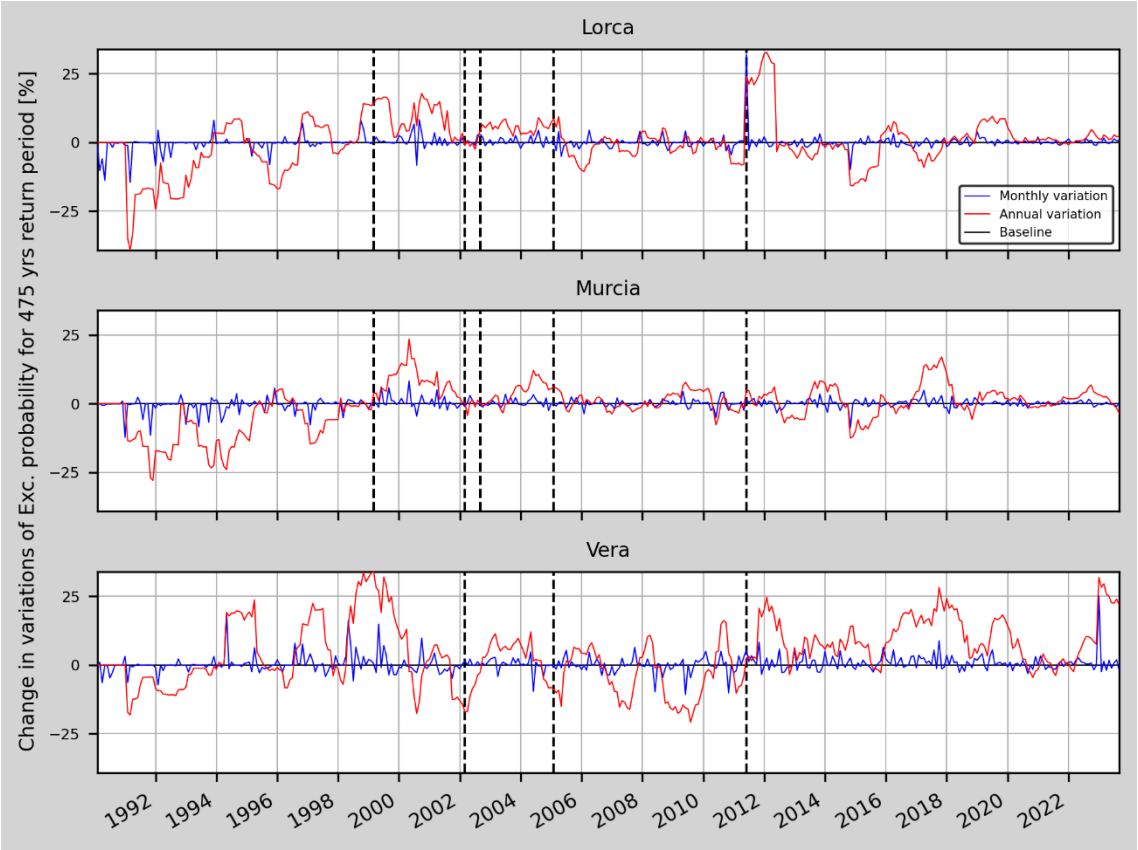

Figure 13. Annual and monthly variations of the relative change of the annual probability of exceedance for Model 1t in Lorca, Murcia and Vera (from top to bottom).

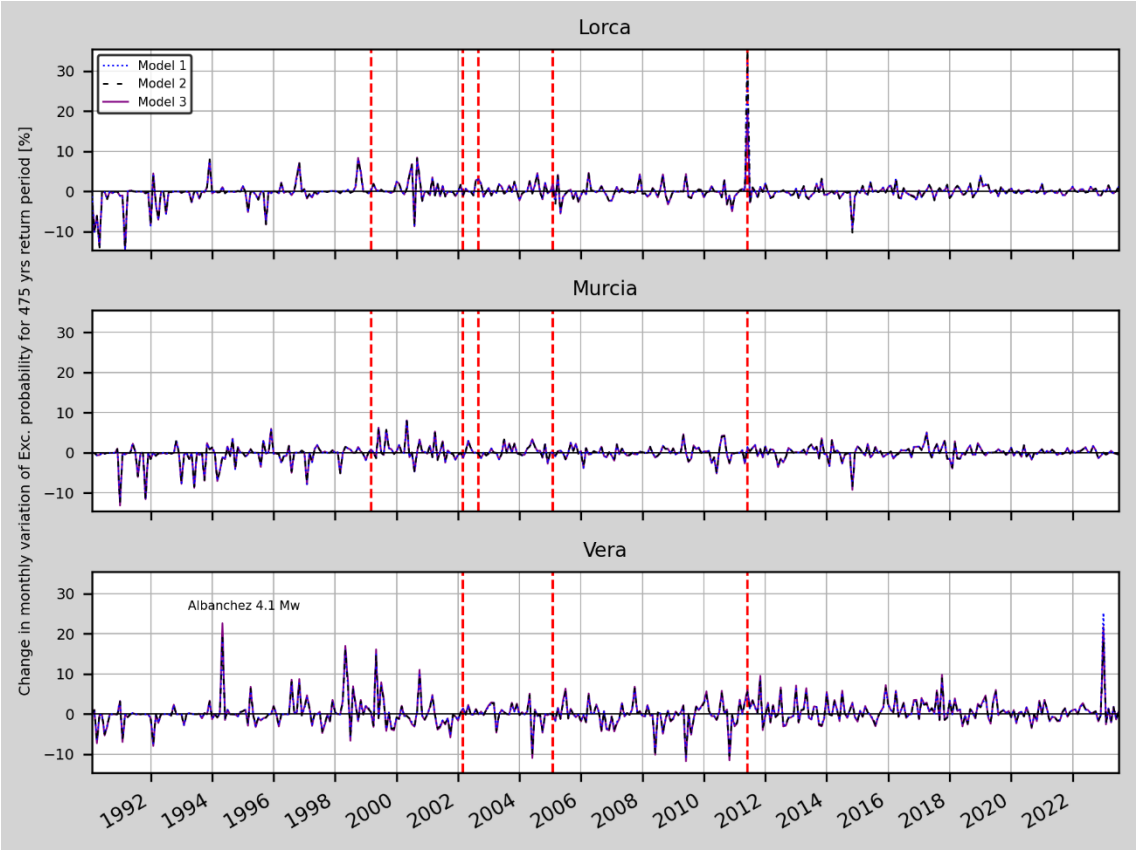

Figure 14. Monthly variations of the relative change of the annual probability of exceedance for models 1t, 2t and 3t in Lorca, Murcia and Vera (from top to bottom). Two peaks have been selected for a zoom in detail (after Lorca earthquake in Lorca site and in December 2022 in Vera site). The earthquake that causes the peak in 1994 at Vera site has also been indicated.

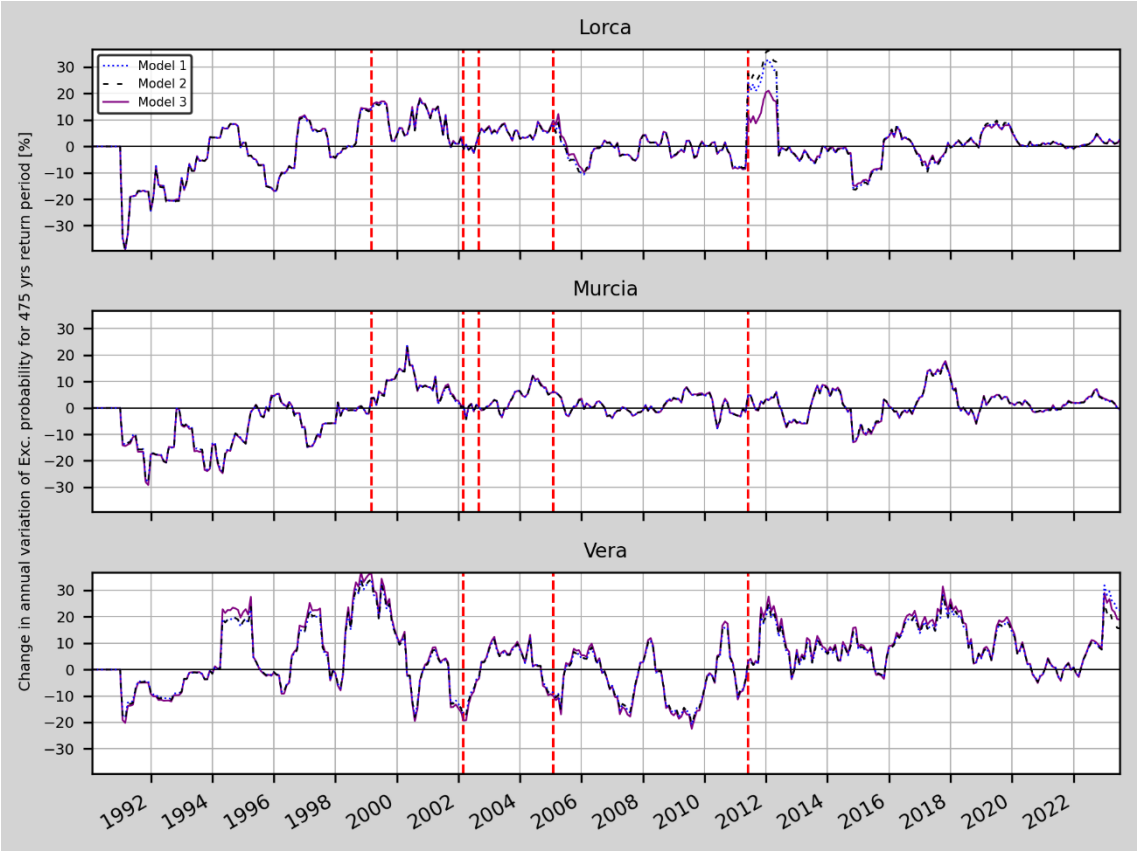

Figure 15. Annual variations of the relative change of the annual probability of exceedance for models 1t, 2t and 3t in Lorca, Murcia and Vera (from top to bottom).

*"Similar results are obtained for all three locations, although higher changes in the monthly variations (**Figure 14**) can be seen for Model 1t and Model 2t after Lorca's earthquake in Lorca site, and for Model 1t in December 2022 at Vera site. Overall, the monthly variations do not show changes preceding relevant earthquakes. One of the possible explanations is the lack of foreshocks in most of the main shocks. In Lorca earthquake, even though there was a 4.5 Mw earthquake almost two hours before the main-shock, the one-month increments on the computation process are not able to show any change in RC.*

*The annual variations on the other side (**Figure 15**), show periods of increased RC before some of the selected earthquakes. An example is seen in Lorca site where a 15% increase is seen before Mula earthquake from June 1998 (the earthquake occurred in February 1999). Another example can be seen in both Murcia and Lorca sites, where a 10% increase can be seen before Aledo earthquake from May 2004 until the earthquake occurrence in January 2005.*

The most prominent increase on the annual variation occurs after Lorca earthquake (32.8%, 36.2% and 21% for models 1t, 2t and 3t)."

In the conclusions, also:

4.0 Conclusions

[…]

"Regarding which of the proposed models can be more effectively used to describe these changes, we have to consider several factors. One could be how close are the computed PGA values to the national seismic hazard maps for each country. In the case of Central Italy models 1t, 2t and 3t provide the following background PGA values: 340.61, 359.72 and 334.28 cm/s^2. The ESHM20 model (Danciu et al., 2021) computes 334.38 cm/s^2. The closest match would be Model 3t followed by Model 1t. However, by looking at ***Figure 5*** *and* ***Figure 7****, it can be seen that the Model 3t seems to be less responsive to the seismic activity than Model 1t, as the monthly and annual RC variations suggest (Model 1t monthly and annual variations are 4.5 times higher than Model 3t variations'). With this information Model 1t seems appropriate for the purpose of this work."*

Figure 1 caption: Please provide clear definitions for each symbol and color used.

The figure's caption has been updated to include clear definitions for each element of the map. Some changes in the figure have also been made, as some items where not able to be correctly read.

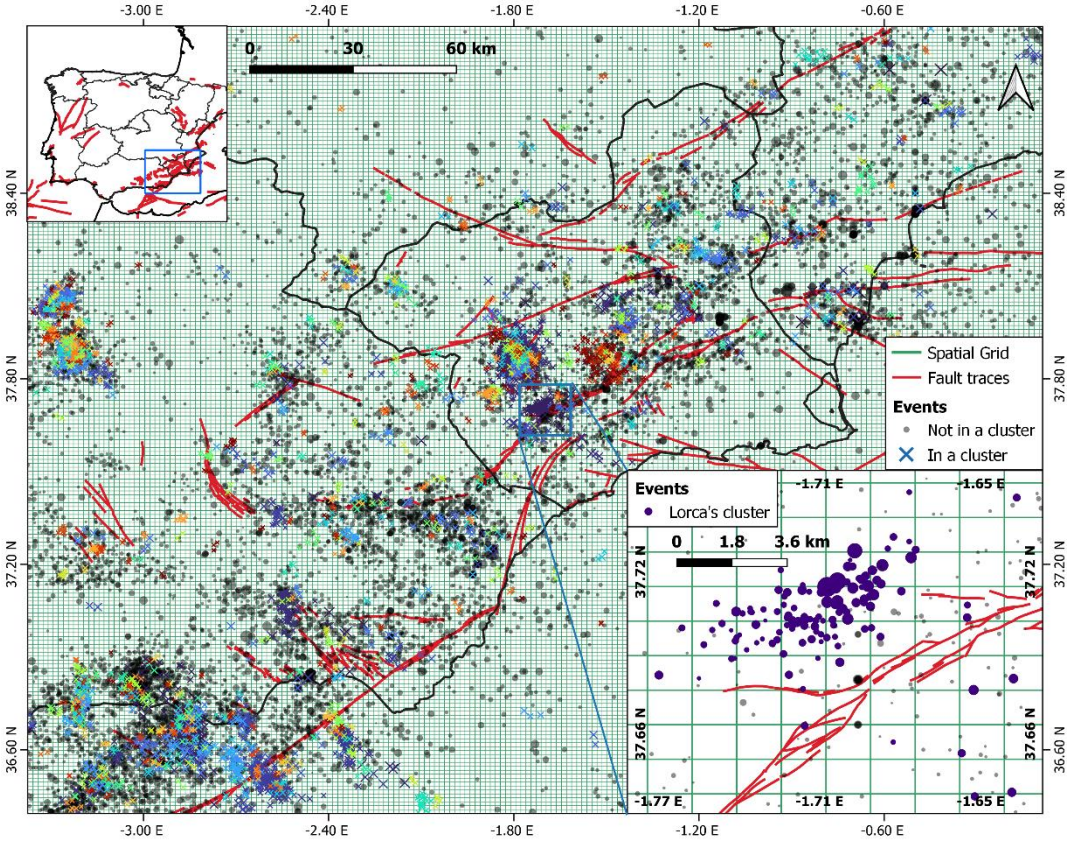

Figure 2. Example of cluster identification in South-eastern Spain. ***The cross markers represent the earthquakes and are coloured depending on the cluster they belong to. The events of the catalogue that do not belong to any cluster are***

represented with grey circles. The fault traces are presented in red-coloured lines and the spatial grid cell's limits with green lines. A zoom in on Lorca's cluster is shown in the bottom right corner, where the events of the Lorca's series are plotted using purple circle symbols.

Lines 159: I am not quite sure if equation (7) is essential.

We agree that it is indeed not essential, the equation 7 has been discarded, although the explanatory paragraph has been kept in order to support **Equation 8** (now **Equation 7**) explanation.

*"To do this, all the events belonging to each cluster are counted and define the clustering weight, c_j. 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. 7**):"*

Line 182: Revise 'time-dependent' as 'stability'.

In order to be clearer about the b-value options the previous line has been rewritten as:

*"[...] a fixed and **constant** (time-independent) b-value assigned from the tectonic zones of each country [...]"*

We think it is better to keep the time-dependent adjective to the time-dependent model to emphasize the aim of the work, TDPSHA.

Figure 3 caption: Is the tectonic b-value obtained by EHSM20 or by this study?

Yes, the b-values are defined within the tectonic zonation from Danciu et al. (2021) -the EHSM20. In order to clarify this, the caption has been modified.

"The map shows the tectonic zones (green lines) in Central Italy with their acronyms and tectonic b-values as computed in the EHSM20 (Danciu et al., 2021). The star marks the epicentre of L'Aquila earthquake (Table 2), and the red lines represent the fault traces."

Figure 4 caption: Please provide clear definitions for the dashed line.

Figures from Figure 4 until the ones from the appendix have gotten their captions updated in order to have all the information needed for their analysis.

Table 6: The effectiveness of the declustering approaches can be assessed with a confusion matrix, which provides not only the number of events but also the counts of true positives, true negatives, false positives, and false negatives.

A figure presenting the confusion matrix for each of the declustering methods and the discussion, is now presented:

*"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 them, the cluster corresponding to Lorca's series, from 11 May 2011 until 19 July 2011, is composed of 146 events (including the foreshock, the main shock and the aftershocks). In order to test the performance of the methods, the confusion matrices for each one have been computed. In the area of study, a total of 249 events have been recorded, which means a total of 103 background events should be identified. For this analysis, all the events classified in a cluster different from the one of Lorca series have been considered as background for simplicity. **Figure 9** shows that GK74 method is the most adequate (with a 94.43% mean for the metrics compared with the 92.88% for RJ and a 74.54% for A) and also the one that is able to identify more events belonging to Lorca's series."*

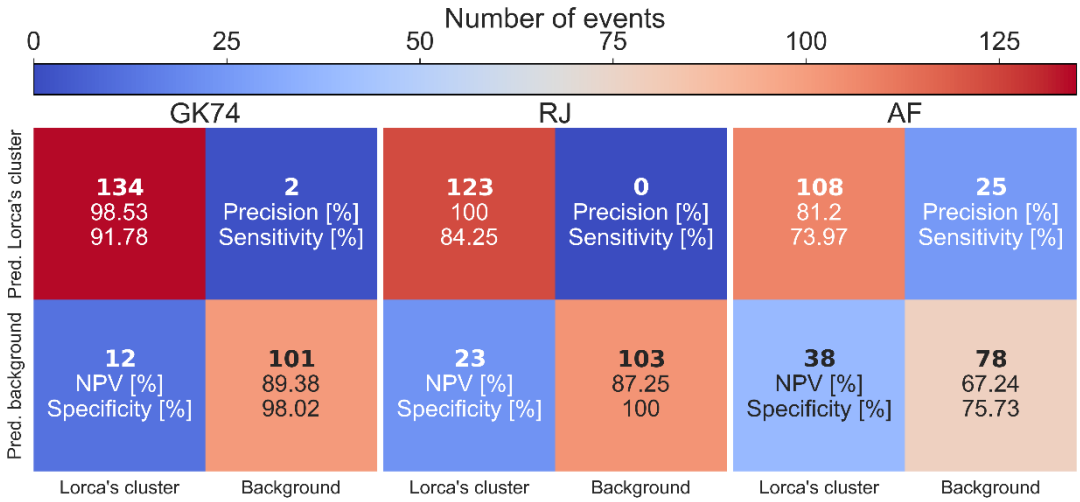

Figure 9. Confusion matrices for the tested declustering methods. Inside each square, the number of events (bold) and some metrics computed using the data are presented (NPV stands for Negative Predictive Value).

Figures 14 and A1, Lines 370-372: Why is there an increase in the expected hazard and rate in Vera? Does this suggest that a large earthquake is anticipated in the future? An explanation based on the data and/or methodology used would be helpful.

The sudden increase in December 2022, that adds up to the increasing RC tendency from 2012 onwards, can be related to the recent earthquake in a municipality 14 km from Vera site. This has been now addressed in the results' comments:

"[...] Lastly, the peak in the annual and monthly variation at Vera in 2022 appears due to the seismic activity in Turre (a town 14 km south from Vera) where a 4.02 Mw earthquake struck on 31 December 2022."

As for the increasing tendency itself, there is not enough data in order to formulate a hypothesis. It could be due to changes in the background seismicity (meaning a new update is due for the PGA background in Vera). In the last paragraph of the conclusions:

"Finally, in the case of south-eastern Spain, the PSHA kept high in the region after the Mula earthquake and 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, more time and data are needed to confirm this."