23 Feb 2024

Editor decision: Publish subject to minor revisions (review by editor)

by Fabrice Cotton

Public justification (visible to the public if the article is accepted and published):

The paper is much improved and the results are better presented and discussed. One reviewer provides a solid list of recommendations/comments that should be considered in the final version.

We would like to thank you very much for handling our manuscript. We appreciate it very much.

I also have some suggestions/recommendations:

- The abstract should be strengthened to better reflect the lessons learned from this analysis. The current last sentence is weak and does not provide solid information to the reader. I suggest adding to the abstract the key "lessons learned" to improve future testing, which are listed in the conclusion.

Thank you for suggesting this improvement. The final lines of the revised abstract state the lessons learned from the comparisons in the paper.

- The authors compare many models, but do not provide enough information on which models are the most "solid/likely". Of course it is difficult to compare models, but the paper would really be more useful if the authors could give an expert opinion on them. For example, the paper states that the ESRM20 and brgm exp exposure models are different, but it is not really clear which of the two models is considered the most reliable (would the authors give similar weights to them in a logic tree approach?)

Thanks for your comment, which is obviously very important. This is quite a delicate issue and difficult to discuss, as the results are still open to debate despite the different models used. However, as we agree with you on the usefulness of the article in providing an expert opinion, we have tried to offer some guidance to the reader. (lines 670-687).

First of all, contrary to what might have been expected, the building-by-building scenario calculations did not provide enough reliable information to be able to calibrate the scenarios. This is due to the need to convert tags into degrees of damage, or to reinterpret inspection sheets. In France (as in Italy), emergency diagnostics (by the AFPS or by the firefighters) tag buildings in three colours (red, yellow and green), which is normal in an emergency context. One recommendation we can make, which will be useful for research and benchmarking work, is to add to the emergency inspection sheets the classification of the building according to the EMS-98 damage grade or to the damage scale in the ESRM20.

As far as the calculations based on the aggregate exposure models are concerned, we can say that, for site effects, the combination of BRGM V_{S30} model and BRGM's (infra-communal) exposure is the best choice at the commune scale. This choice is supported by the values of the V_{S30} in tables 3.3 and 3.4, where the values of this combination are closest to the site effects expected in the area. There are two reasons for this: the resolution of the exposure (nine points instead of one) and the resolution of the site effect zones in the BRGM V_{S30} model is better than that of the ESRM20, which is perfectly normal since the ESRM20 has been developed for application on the European scale.

Now, always in terms of aggregate models, from an exposure perspective, it appears that the simplification of the ESRM20 exposure and fragility model has a minor effect on the results (Figure 6a), at the municipal scale, if regional hazards and site effects have good description and resolution. Hence the complexity of drawing a general conclusion.

This is what emerges from this study and what is important to mention is that the resolution (extent) of the exposure is also important for the description of the effects of the site (as long as we use the centroid or even the mean within) not only for vulnerability calculation; if it is necessary to allocate resources, we would prioritize the detailed description of site effects and the assignment of the right typologies in the smallest exposure zones. This is quite consistent with other studies.

Finally, concerning our results, and considering that those of Armagedom software were the most calibrated with the observations through the AFPS studies, we think that, in our case, the best choice is the DS1 scenario which is the combination between the GM1 and BRGM exposure.

- It is not clear to me how low-cost sensors can contribute to a better explanation of damage levels (last sentence of the conclusion).

The last sentence of the conclusion has been revised and refers to Goulet et al. (2015), who proposed this.

- The conclusion discusses the use of the AS2000 model (Atkinson and Sonley, 2000 ?). This model is not mentioned in the analysis or in the bibliography. The use of many acronyms AS2000, FM2010, Ko2020 makes the paper difficult to read (the link with the reference list becomes difficult). I suggest to stick to the classical way of citing papers (Atkinson and Sonley, 2000 instead of AS2000).

The conclusion no longer discusses the AS2000 model.
The acronyms have been replaced in the main body of the text. Now they only appear in Figure 4.

- Figure captions ("model") to the right of Figure 1 and 2 are not necessary.

The captions (“model”) have been removed from Figures 1 and 2.

Review of Manuscript egusphere-2023-1740 R1

Comparing components for seismic risk modelling using data from the 2019 Le Teil (France) earthquake

This manuscript is a revised version of the previously entitled "Testing the 2020 European Seismic Hazard and Risk Models using data from the 2019 Le Teil (France) earthquake" paper draft. The newly-proposed title ("Comparing components for seismic risk modelling using data from the 2019 Le Teil (France) earthquake") is much more appropriate and better reflects the contents of the manuscript. It is now clear that the work is not a "test" on the 2020 European Seismic Hazard and Risk Models.

The authors have done a commendable job reorganising the contents of the paper and labelling the different (combinations of) components being used for the different damage scenario calculations. It is now a lot clearer what comparisons are being made, which makes the work a lot more readable and understandable. The new figures and tables are very useful, as are the modifications introduced to the bar plots of the results.

There are, however, still some points that require clarification before publication. I recommend that the manuscript be considered for publication after the following minor revisions.

We would like to thank you very much for all the comments and the guidance, which were of major help in improving the manuscript.

Main Comments

1. Lines 22-23: The statement "an exposure and fragility model assembled herein leads to lower probabilities for damage grades 3-5 than the ESRM20 exposure and fragility model" suggests that the difference observed in Figure 6a is much larger than it looks (please see my comment on line 576 of the conclusions). Please re-phrase.

The revised abstract (line 24) states that the differences are small, which is in agreement with the revised conclusions.

2. Lines 53-55: The description of section 2 is focusing on just one of the aspects being presented there. I suggest either re-writing it as a more general statement (like "section 2 focuses on the interpretation of post-earthquake assessment damage data acquired for a small sample of buildings in terms of a 3-levelscale…"), without going into more details, or enumerating the several things being done (i.e., add that three different distributions of damage are defined, etc).

Lines 58-62 have been re-written based on your suggestion without going into more details.

3. Lines 69-71: I suggest to re-arrange the order of the three comparisons as they appear later in the paper (if so, the subsequent sentences need to be changed, e.g., "the last two types" □ "the first two…").

In the revised manuscript (lines 76-80), the comparisons are ordered as they appear in the paper, and the subsequent sentences have been revised.

4. Lines 127-129: Is the distribution of green/yellow/red tags across these 174 entries similar or different from the distribution of the other 327 entries (i.e., shown in Table 2-2)? This might give a hint on how the inspections were conducted (e.g., are undamaged buildings under-sampled?), which can inform the other two criteria (DD2 and DD3) defined in the paper, or indicate if any bias is introduced by removing these 174 entries (35%) of the damage dataset. Please comment in the paper.

We have added Table A5 and added a comment in the manuscript (lines 137-140) explaining that the removal of these entries does not introduce any significant bias.

5. Line 163 and Table 2-3: Please re-phrase. The word "extrapolation" immediately brings to mind that the same proportions of damage of the 327 buildings were applied to the 2,778 buildings, while a more complex combination of observed values and judgement-based decisions was applied for DD2 and DD3.

The word "extrapolation" has been replaced with the word "adjustment" (line 174).

6. Lines 176-178: Please explain the logic behind your judgement. This is relevant to give meaning to DD2, given that a weight of 0.82 is applied to those numbers, which means that DD2 ends up being almost a pure reflection of such judgement (i.e., the red/yellow/green tag-to-EMS98 conversion has little impact on DD2). It is clear that the process is inherently subjective, but explaining the rationale behind the subjectivity would make it more transparent and useful. Something that strikes the eye is the assumption that every single building of the 2,778 set was damaged at least non-structurally. This seems like a strong assumption, especially for buildings for which inspections were not requested. At the same time, nonstructural damage is difficult to assess, as it is quite common to encounter non-structural cracks in buildings, and very hard to determine if they were caused by that particular earthquake or not. Please discuss the rationale behind assuming no EMS-98 damage grade zero at all in the whole municipality.

The revised manuscript discusses the reasoning behind the selection of the probabilities based on our judgement, and the assumption that every building has at least non-structural damage. We claim that this assumption is reasonable in the case of the inspected buildings. Moreover, we acknowledge that this assumption may lead to an overestimation of non-structural damage in the uninspected buildings, but this overestimation should not be excessive due to pre-existing non-seismic damage (lines 190-198; 474-478; 529-533).

7. Lines 184-188: These lines explain how to use Table 2-8 in combination with the P(tag) column of Table 2-4 to obtain a final damage distribution, but not where the numbers in Table 2-8 are coming from. The caption of Table 2-8 says they were defined by expert judgement. For the same reasons stated in the previous comment (regarding DD2), please explain the rationale behind your expert judgement. Please highlight that this criterion applies to this particular earthquake, to avoid an erroneous interpretation that in any earthquake a red tag would imply only a 5% probability of DG5.

The revised manuscript (lines 208-210) explains that the Table 2-8 reflects the judgement of experts who participated in the post-seismic emergency assessments. It also states that they apply only in the case of this earthquake.

8. Lines 264-265: Given that the date of construction was available, the lateral force coefficients could have been estimated (following Crowley et al., 2021, https://doi.org/10.1007/s10518-021-01083-3, for example). Please rephrase to simply say that the lateral force coefficient was not considered, but avoid saying it cannot be considered.

The manuscript has been rephrased as suggested (lines 292-294).

9. Line 276: Please clarify these are the OpenQuake names of these ground motion models and add the corresponding citations:

- KothaEtAl2020Site: Kotha et al. (2020) (it's already in the references)
- KothaEtAl2020SlopeGeology: Weatherill et al. (2023) (already in the references too)
- KothaEtAl2020ESHM20 would be Weatherill et al. (2020) but, as it is not being used in the paper, there is no need to mention it (see comment below).

10. Lines 276-278: The statement "which were developed in the context of the development of the GMM KothaEtAl2020ESHM20" is not accurate (please see the explanation in my previous review). However, it is not necessary to explain all the alternative versions of this GMM, given that not all of them are being used in the paper and it is now clear that the paper is not "testing ESHM20/ESRM20", as it was presented in the previous version of the manuscript. As per my previous comment, please focus on the two GMMs used in the comparisons and provide the corresponding citations (e.g., like done in lines 318-319, "a version of…", which could be moved to section 3.3). Lines 279-281 are fine.

The manuscript has been revised according to comments 9 and 10. It no longer refers to KothaEtAl2020ESHM20 (lines 306-308, 351-353).

11. Lines 289, 291, 295-296: If the values of Vs30 used result from "averaging over the polygon of the municipality" (line 291), then they are not "the values […] at the coordinates of the exposure centroids" (line 289). Of course, once the Vs30 values are retrieved, they are used in OpenQuake as if located in the centroids, but they are not the values of Vs30 at the centroids. Please correct. Lines 295-296 present the same issue.

The revised manuscript (lines 320-322) has been corrected. It states that the values are those returned for the coordinates of the centroids by the *point* workflow of the *exposure to site tool*.

12. Lines 289-294: I assume the same procedure was used to extract slope and geology. Please explain in the paper.

The revised manuscript explains this (lines 328-330).

13. Line 290: Please add the following citation for the ESRM20 exposure-to-site tool:

Dabbeek, J., Crowley, H., Silva, V., Weatherill, G., Paul, N., and Nievas, C.I.: Impact of exposure spatial resolution on seismic loss estimates in regional portfolios, Bulletin of Earthquake Engineering, 19(14), 5819-5841, 2021.

This citation has been added (line 321).

14. Line 306, Table 3-3: Are the values in the "ESRM20" column calculated as weighted averages of the district polygons or are those the values for the points themselves (without averaging)? If so, which polygons, given that centroids 0-8 are sub-municipal divisions of Le Teil?

The values for the ESRM20 column are the values for the points themselves without averaging. This is specified in the revised manuscript (line 321).

15. Section 3.3: In their reply to point 16a of my previous review, the authors indicated they have used the ESHM20/ESRM20 ground motion model logic tree, keeping only the branches associated with their tectonic setting (active shallow crust), which is correct. The active shallow crust branch of the logic tree still contains 15 sub-branches. I agree with the authors that all these details are likely not so relevant for most readers, but then the paper talks about 1,000 realisations of ground motion when, in reality, it looks like 15,000 realisations (1,000 x 15) were used. I suggest to add a small comment that clarifies that the ESHM20/ESRM20 ground motion model logic tree for active shallow crust area sources was used, that it consists of 15 branches (and their associated weights), and that 1,000 realisations of ground motion were sampled for each of the 15 branches. These 15 branches stem from a 5-branch discrete approximation to the Gaussian distribution describing the regional variability of the earthquake source, which effectively represent five different levels of stress drop and thus "account for the possibility of having extreme stress parameter values", as written in line 105 of the present manuscript.

In the original submission, we used the ground motion logic tree in the comparison with Armagedom. In the revised version submitted in the previous round of review in January, Armagedom is compared to the DS1 calculation that does not use a ground motion logic tree. We consider that the paper cannot afford to expand on this subject. We hope that you will approve this change.

Clarification regarding the authors' reply to main comment 4c of my previous review: In my original review I did not explicitly say that I was referring to the ground motion logic tree, not the source model logic tree (e.g., when saying "The ESRM20 logic tree input file and its "cut" version used for shallow crustal areas when comparing against past earthquakes indicate that this is the GMPE used in ESRM20 to calculate losses"), though the links to the files implicitly referred to the ground motion logic tree. The authors' reply referring to the collapsed version of the logic tree for ESRM20 refers to the source model logic tree, not the ground motion logic tree. The ground motion logic tree was not collapsed for use in ESRM20.

Thank you for clarifying this. Our reply in the previous round of review was incorrect. Please accept our apologies.

16. Lines 318-319: Please specify which site model was used for comparing ground motions and macroseismic intensities.

Line 353 of the revised manuscript specifies that the BRGM V_{S30} site model was used.

17. Lines 330-337: Was correlation between spectral periods not considered in OpenQuake? If so, this can be another source of difference in the results obtained using the same rupture as the shake-map. Please comment if that is the case. Moreover, two alternative correlation models are mentioned as being used with the shake-map (BJ2008+JB2009 vs the Nearest Positive Definite Matrix). Please clarify how the two alternatives co-exist (i.e., are they all grouped together and averaged out in the results?).

The correlation between spectral periods was considered using the default correlation model in OpenQuake *BakerJayaram2008* by Baker and Jayaram (2008). For the spatial correlation, the by Jayaram and Baker is used. The correlation models are used to create the correlation matrix. The Nearest Positive Definite Matrix is used during the sampling process. The revised manuscript gives these details (lines 371-378).

18. Lines 376-377: The equation says σMCS but the text says σMMI. Please correct.

This has been corrected (line 416-418).

19. Lines 373-394:

- According to the text, FM2010 is equation 1 and CA2015 is equation 2.
- In equation 1, σ_{MCS} or σ_{MMI} are used.
- In equation 2, $\sigma_{singleline}$ is used.
- The caption of Table 4-2 says it refers to the FM2010 model, which according to equation 1 uses σ_{MCS} or σ_{MMI}, but the table says σsingleline.
- The caption of Table 4-1 says it refers to the CA2015 model, which according to equation 2 uses $\sigma_{singleline}$, but the table says σ_x.

Please revise and adjust where needed. Moreover, I suggest keeping only the coefficients for PGA in Tables 4-1 and 4-2, as the ones for PGV are not used in the paper.

The manuscript lines 416-425 have been corrected. The coefficients for PGV have been deleted from Tables 4-1 and 4-2.

20. Line 417: It would increase clarity if section 4.2 concluded by stating which rupture model the rest of the paper is going to be based on, as selecting it was the purpose of section 4.2 (instead of stating it only later in section 5).

The revised manuscript (lines 449-450) states that the Ritz et al. rupture model is used in the rest of the paper.

21. Lines 425-429, Table 5-1: For GM2, please re-phrase the site model as “ESRM20 site model (slope & geology)”. For GM3, please re-phrase it as “ESRM20 Vs30 model”. None of the two are data.

This has been corrected (line 450, Table 5-1).

22. Lines 456-458: In the sentence “The fact that in these cases a green tag...”, I suggest to refer back to Table 2-2, otherwise the reader might not follow where this statement is coming from.

A reference to Table 2-2 has been added (line 507).

23. Lines 471-473: I understand Table A4 was also used to convert results in section 5.1, but it was never mentioned there. I suggest stating this in section 5.1 as well. Moreover, please comment somewhere in the main text on the decision to assume that ESRM20's D0 (no damage) translates to EMS-98 damage grade 1, i.e. slight non-structural damage (see comment 6 about lines 176-178 above).

We hope that you will agree that the revisions with respect to comment 6 have covered this comment too.

24. Line 486-487: Please re-phrase "which utilize the damage observations and expert judgement, respectively", given that DD2 is also heavily influenced by expert judgement (as explained in lines 177-178).

The revised manuscript (lines 543-545) explains that DD2 depends mostly on expert judgement and on the damage observation on a lesser extent, while DD3 is entirely based on expert judgement.

25. Lines 479-500: The discussion on the effect of the exposure models focuses on the results obtained with GM3 (i.e., DS5 vs DS6). Please comment if the same trends are observed when using the other GMs (i.e., DS1 vs DS2, DS3 vs DS4, DS7 vs DS8).

The same trends are observed with the other GMs too (see figure below). This is mentioned in the revised manuscript (lines 551-552).

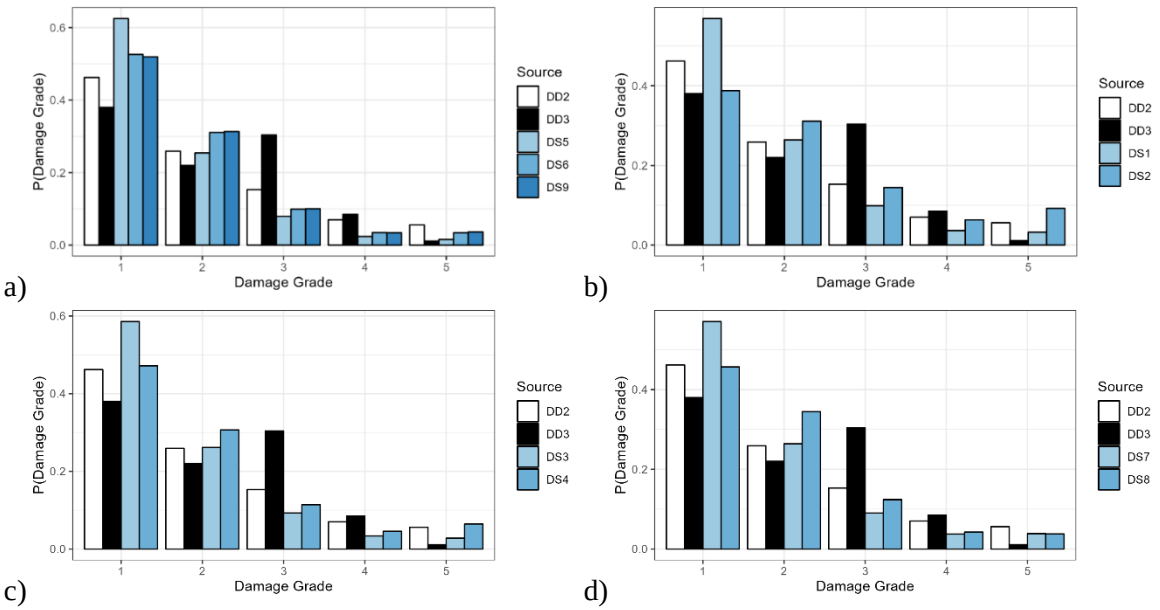

Figure R1 Effect of the exposure model on the probabilities per EMS-98 damage grade for the analyses calculations with an aggregated exposure including the total number of buildings in Le Teil based on the ground motion map a) GM3 b) GM1 c) GM2 d) GM4

26. Line 510: It is stated that scenario DS5 uses the KothaEtAl2020ESHM20SlopeGeology GMM, but Table 5-2 indicates that DS5 uses GM3, which is based on KothaEtAlSite (Table 5-1) instead. Please revise and adjust.

The manuscript has been corrected (lines 551-552). It states: "The damage grade probabilities in the scenario DS3, which uses the KothaEtAl2020ESHM20SlopeGeology GMM, are between the results for DS1 and DS5 for all damage grades."

27. Figures 6 and 7: The large peak of damage grade 3 for DD3 and the associated extreme valley of damage grade 3 (also for DD3) suggest that perhaps it was too optimistic to assume 55% of red tags meaning damage grade 3 and only 5% of red tags meaning damage grade 5 (Table 2-8). It is of course impossible to know for sure, as DD2 is also heavily influenced by expert judgement and all other values are estimates, not observations (i.e., it is not possible to pinpoint one "correct" value). It might be worth including a comment on this.

Indeed, in DD3, the probabilities for damage grades 3-5 depend heavily on the probabilities of these damage grades conditioned on a red tag, which were assigned based on expert judgement. Indeed, it was too optimistic to assign a 55 % probability of damage grade 3 in case of a red tag. Alternative assignments of the probabilities for a red tag may smooth out in DD3 the peak for damage grade 3. The Table below is an example (may be added to the manuscript if you deem it necessary). These comments have added to the manuscript (lines 545-548).

tag	P(DG1|tag)	P(DG2|tag)	P(DG3|tag)	P(DG4|tag)	P(DG5|tag)
Green	0.80	0.20	0	0	0
Yellow	0	0.40	0.60	0	0
Red	0	0	0.05	0.70	0.25

tag	P(DG1|tag)·P(tag)	P(DG2|tag)·P(tag)	P(DG3|tag)·P(tag)	P(DG4|tag)·P(tag)	P(DG5|tag)·P(tag)
Green	0.380	0.095	0.000	0.000	0.000
Yellow	0.000	0.125	0.188	0.000	0.000
Red	0.000	0.000	0.011	0.148	0.053
Sum:	0.380	0.220	0.199	0.148	0.053

28. Line 559: The comparison of ground motions and macroseismic intensities was not "based on components of the ESRM20". The GMM used was KothaEtAl2020Site, which was not used in ESHM20 or ESRM20 (please see main comment 4a in my previous review). The rupture models were obtained from the literature, not from the ESHM20 source model. The site model used is not specified (see my comment above regarding lines 318-319). Please re-phrase.

This sentence has been rephrased (lines 625-626) and it mentions neither the ESHM20 nor the ESRM20.

29. Line 576: Assuming this statement comes from analysing Figure 6a, the difference in the proportion of buildings with damage grades 3-5 obtained with DS5 and DS6 looks very small. Please re-phrase the conclusion so that it better reflects the difference observed, and/or provide numbers that justify the statement.

Indeed the differences are very small. Thank you for this comment. The revised manuscript has been corrected (line 644).

30. Line 577: Assuming this statement comes from analysing Figure 6b, the number of buildings is larger in DS5 than in DS6 for damage grade 3, slightly larger (almost the same) in DS5 than in DS6 for damage grade 4, and slightly smaller (almost the same) in DS5 than in DS6 for damage grade 5. Moreover, lines 494-496 state that "the results of the damage scenarios for damage grades 3-5 [in terms

of number of buildings] present minor differences". This contradicts the statement on line 577 of the conclusions. Please revise.

The sentence in line 577 has been deleted.

31. Lines 579-580: Please clarify what "the dataset based on the emergency post-seismic diagnosis" refers to. Damage dataset? Exposure dataset?

This sentence has been replaced with the following "…with the calculations DD2 and DD3 (see Section 2.2 for the details of the calculations)." (lines 647-648).

32. Lines 587-594: This is a very relevant point to make. However, these lines seem to refer only to DD1 and not to DD2 or DD3 (e.g., "the proposed rule"). Line 591 states that "the effect of possible alternative conversion rules" was not studied, which I assume refers to potential alternatives of Table 2-1, but it can be confusing for the reader because DD2 and DD3 are alternative conversion rules. Please re-phrase so that it is clear when you are referring only to DD1 and consider adding some comments regarding DD2/DD3.

The revised manuscript has been rephrased to make clear that it refers to Table 2-1 and DD1 (lines 659-662). Comments with respect to DD2/DD3 have been added (lines 662-665).

33. Lines 598-599: This is a very relevant point to make as well. Given the large weight that expert opinion had in the definition of DD2 and DD3, it would be important to remind the reader here of these subjective assumptions and the fact that all comparisons carried out against DD2 and DD3 have this inherent limitation.

We hope that you will agree that the revision in response to comment 32 has covered this comment too.

34. Lines 599-603: Given that inspections were carried out upon request from the owners, the assumption that undamaged buildings are underrepresented in the sample of 327 buildings seems quite reasonable. This statement appears as misaligned with the assumption in the conversion between ESRM20 and EMS- 98 damage grades that ESRM20 damage grade 0 equates to EMS-98 damage grade 1 (Table A4). Regarding completely destroyed buildings not being inspected, this looks like a more complicated assumption. Post-earthquake damage assessments tend to be carried out not only "to inform about the risk associated with the use of impacted buildings" (as stated in line 603), but also to understand future housing needs and for governments to make an estimate of the need for relief funds. I would say it makes sense that completely destroyed buildings may not have been inspected in detail (to classify damage to different structural components, for example), but they may have been tagged in terms of the global state of the building. This is why it would be important to comment in the paper about the 174 entries of the damage assessment for which only colour tags for the buildings were available (as per my comment earlier about lines 127-129). Are completely destroyed buildings better represented in these 174 entries? Please consider all this to potentially re-phrase this last paragraph of the conclusions.

We hope that you will agree that the revisions in reply to comment 6 have discussed the conversion between the ESRM20 and the EMS-98 damage scales.

First, as a side note, we would like to note that, to the best of our knowledge, the post-seismic emergency inspections by the AFPS have as exclusive purpose to warn with respect to the risk presented to the users of the buildings. Therefore, please allow us to claim that it is not unreasonable to assume that completely destroyed buildings may have not been inspected by the AFPS. However, we suppose that completely destroyed buildings may have been included in surveys by others such as the firefighters and researchers working on macroseismic intensity.

Moreover, the entries that were filtered out of the dataset have a similar distribution of colour tags (Table A5) to the remaining entries. We assume that they are incomplete for reasons unrelated to the level of damage.

Minor Comments/Edits

Unless otherwise indicated, all minor comments/edits have been made.

1. Line 13: To my knowledge, "emergency post-seismic diagnosis" is not a standard term. Consider replacing it with "emergency post-earthquake assessment" or "emergency post-earthquake inspection" (here and all throughout the paper).

2. Line 14: "shake-map analyses" can be replaced by simply saying "shake-maps", while "scenario damage analyses" can be replaced by "scenario damage calculations", here and all throughout the paper. The word "analysis" does not imply a computation/calculation.

3. Line 18: No need for inverted commas around "building-by-building" (here and all throughout the paper). Many inverted commas throughout the paper could be removed to improve legibility.

4. Line 56: "to investigate as of components".

5. Lines 62: "Shake-maps are employed".

6. Line 71: "comparisons based on a building-by-building exposure model".

This has been replaced with "comparisons based on probabilities of EMS-98 damage grades".

7. Line 75: Mentioning the GMM is too much detail at this stage (the other GMM model is not being mentioned, there is no reference or explanation, all this comes later), I suggest to just focus on the site models.

8. Line 93: "and a moment magnitude Mw 4.9".

9. Lines 97-98: "the municipality of Le Teil, that they cannot".

10. Lines 115 & 117: erase "the" in front of "Le Teil".

11. Line 131: "to EMS-98 damage grades".

12. Line 158: “compare results of analyses against three different sets”.

13. Line 173: A reference to Table 2-2 at the end of the sentence would improve clarity.

14. Line 180: I believe the intended reference is Table 2-6, not 2-7.

15. Captions of Tables 2-6 and 2-7: “of the EMS-98 damage grades”.

16. Lines 190-205: Please consider labelling the final damage probabilities for DD1, DD2 and DD3, so that the reader can easily come back to them once they reach the sections with the comparisons. Please add these final damage probabilities to Table 2-8.

The damage probabilities for DD1, DD2 and DD3 have been labelled in Tables 2-5, 2-7, and 2-8. The final damage probabilities have been added to Table 2-8.

17. Line 244: “similarity of their lateral load-bearing systems”.

18. Line 249 and several instances along the paper: There is an error with the references. The text “Erreur! Source du renvoi introuvable” appears.

19. Line 260: Eliminate “building class”.

20. Line 284: Please define “EC8” and add the appropriate citation.

21. Line 286: “extracted”.

22. Line 322: “was re-calculated with the”.

23. Line 331: At the end of the line it says “Moreover, the”, but then the next line contains a different sentence.

24. Line 332: What do you mean by “updated” parameters? “Updated” with respect to what? Please revise this sentence.

The word “updated” has been deleted (line 333).

25. Line 421: I suggest adding “which includes 327 buildings with classes defined in Table 3-2”, to ease readability.

26. Line 423: It is three different site models being used (BRGM Vs30, ESRM20 Vs30, and ESRM20 slope and geology), not two (as the sentence says). Please be consistent with the use of upper/lower case for BRGM (sometimes it appears in lower case and sometimes in upper case).

27. Line 450: Should the reference be Table 2-2 instead of Table 3-2?

We intended to add a reference to the table that gives more information on DD1. This table is Table 2-3 and it has been placed immediately after DD1 (line 457) in the revised manuscript.

28. Line 455: I suggest adding “the horizontal structural elements have a yellow tag (see Table 2-1)”.

29. Line 481: I suggest adding “the same rupture model, GMM, and site model (GM3)”, to ease readability.

30. Lines 500-501: I suggest adding to the end of the caption “using ground motion map ID GM3”, to ease readability.

31. Lines 531-532: I infer that Armagedom takes the map of macroseismic intensity as a user input. As it is currently phrased, it can be interpreted that this particular map (of Schlupp et al., 2022) is the only map of macroseismic intensity that is hard-coded in Armagedom. I suggest re-phrasing.

Lines 540-542 have been revised.

32. Lines 538-540: I suggest erasing this last sentence. Now that the goals of the paper have been re-phrased and it is clear that the paper is not a test of ESHM20/ESRM20 components, the purpose of this comparison has become more self-explanatory.

33. Line 576: “The scenario damage analyses leads to”.

34. Lines 581-582: “The estimation based on the Armagedom tool results in probabilities”.