Review of Manuscript NHESS-2019-331

"Induced seismicity risk analysis of the hydraulic stimulation of a geothermal well on Geldinganes, Iceland"

M. Broccardo, A. Mignan, F. Grigoli, D. Karvounis, A.P. Rinaldi, L. Danciu, H. Hoffmann, C. Milkereit, T. Dahm, G. Zimmermann, V. Hjörleifsdóttir & S. Wiemer

This is an interesting paper that narrates the development of an induced seismic risk model for planned geothermal stimulations through fluid injections in Iceland. The topic would clearly seem to be within the scope of *Natural Hazards and Earth System Sciences* and, as the authors point out, it may be the first published induced seismic risk study performed prior to the commencement of injections. I would therefore conclude that the paper is a potentially valuable contribution and should be published, but I believe that it can also be improved. In particular, some information that is extraneous given the title and focus of the paper could be removed, and some more detail provided on essential features. I also think the authors could do a slightly better job of citing relevant literature and acknowledging work by others. In the following paragraphs I offer my observations and suggestions in the order in which they refer to the manuscript rather than any hierarchy of importance. I am not correcting inconsistencies in the citation of references within the text—years sometimes in brackets, sometimes not—because I do not consider that part of my work as a peer reviewer.

Line 28 (and many other locations): The Latin phrase "*a priori*" should be written in italics and never hyphenated.

Line 43: Please explain what straddle packers are since this may not be common knowledge for all readers (myself included).

Line 58: "Adding an additional" is very inelegant, please re-word.

Line 59: Heat cannot be effectively transmitted over distance so such systems will always be close to consumers out of sheer necessity.

Line 61: From the references cited here, the reader is left with the impression that other than Dr Mignan and his co-workers, nobody has ever written about the assessment and management of induced seismicity.

Line 68: "even more importantl<u>y</u>"

Line 69: "are the sizes and stress state"

Lines 76-77: Successful? Majer et al. (2007) documents several geothermal projects in which the largest induced events occurred after pumping had stopped, highlighting the limitation of TLS. The recent paper in SRL by Baisch et al. (2019) on the efficacy of TLS is probably worth citing in this regard

Lines 81-82: It is, at least, the first being made public—there may well be commercial and confidential studies that have not been released.

Line 84: The Latin phrase “*in situ*” should be written in italics and never hyphenated.

Lines 104-105: Why or how is this being suggested? Is M 2 the detectability threshold?

Line 122: “a strike-slip/thrust regime”: I am not clear if this means one or the other, or an oblique combination of the two? If the latter, do you mean reverse rather than thrust? Thrust specifically refers to very shallow dipping reverse faults, and strike-slip on such faults is unusual.

Lines 147-148: “The concept…….(Pittore et al., 2018)” – unless you tell the reader what the outcome of the DESTRESS evaluation was, this sentence does not really serve any purpose.

Line 165: Suggestion: “is planned to be approximately equal” (current wording is strange)

Line 196-236: This text is not needed since it is perfectly summarised in Figure 3 and is not central to the title and focus of the paper.

Line 238: Bommer et al. (2017) is not listed in the references—and I don’t know what paper it refers to!

Line 292: “(and generally unknown *a priori*)”

Line 293: Langenbruch et al., 2018.: Nature Communications 9:3946, doi:10.1038/s41467-018-06167-4 have made the connection to pressure changes explicit in their modification of the seismogenic index.

Lines 298-299: No, it is not also known as the seismogenic index, it is the seismogenic index and you are choosing to give it another name, for reasons that are not clear to me at all. This was pioneering work by Serge Shapiro and there is no need whatsoever to assign it a new name—unless the plan henceforth is to cite this paper as the origin of the concept, which would be totally unacceptable.

Lines 314-315: What weights are assigned to each of the models? This is not indicated in Figure 4. The rationale for the weights also needs to be presented.

Lines 352-354: Again, the weights need to be reported, together with their rationale. Why are these parameter combinations not depicted as logic-tree nodes and branches in Figure 4?

Line 361: Is this really the appropriate reference for the concept of maximum magnitude?! Did nobody work on these topics before you?

Lines 363-365: This statement may not hold for induced seismicity. In some hazard and risk studies for induced earthquakes, the results are very sensitive to the maximum magnitude.

Lines 371-372: The phrase in parentheses does not currently make sense, re-word.

Lines 386-391: A clearer explanation of the two concepts could be given. Is this the difference between maximum possible and maximum expected, as discussed, for example by Zöller & Holschneider, 2016, BBSA, 106(6), 2917–2921? As currently written it will create confusion because in common practice, M_{max} is used as the symbol for the parameter you refer to as m_{sup}.

Lines 422-423: This is a poor reason for using only PGA unless it is also found to be an efficient and sufficient parameter for the risk calculations.

Lines 423-424: Given that R_{jb} ignores focal depth, which is critical for induced earthquakes (as note in the next paragraph) what was the rationale for this choice?

Lines 425-426: In Figure S1, the range of epistemic uncertainty appears to decrease with increasing magnitude, which is counter intuitive. Do the authors have an explanation for this feature?

Lines 428-429: I don't think it's probably biased, it is biased—see Bommer et al., 2007, BSSA, 97(6), 2152-2170 for discussion of this issue. Baltay & Hanks, 2014, BSSA, 104(6), 2851–2865 provide a clear physical explanation for this behaviour.

Lines 432-434: Depth influences the length of the travel path and the stress drop, so this dismissal of the issue seems rather casual.

Line 435: "over instrumental intensity measures" (physics-based is not appropriate).

Line 439: You claim to be using EMS98 but choose a GMICE calibrated in terms of MCS intensities. Musson et al., 2010, JOSE, 14, 413-428, have pointed out that while these two scales appear equivalent in their definitions, they are generally not used in an equivalent manner.

Line 440: Please explain how the increase sigma value is applied in the hazard and risk calculations—is the PGA to intensity conversion performed within the hazard integral?

Line 468: I would suggest mentioning physical damage to buildings, from which injuries and costs are both then derived.

Lines 473 & 475: What is meant by "statistically average" in these phrases? Is individual risk average over the exposed population or reported for individual locations?

Line 481: Risk targets for individuals in the Netherlands are define at 10^{-4} and 10^{-5}.

Line 481: the reference is van Elk et al. (2017).

Line 495-496: On what basis is this assumption made? Is there any basis for its justification? Or you have to assume it to make the results meaningful?

Section 4: There is nothing about the exposed building stock in terms of numbers of buildings and their location, just as there is nothing about the exposed population. More information is needed and must be added.

Section 4.1: You also do not explain how hazard and risk are convolved—do you convolve hazard and fragility curves or use stochastic event sets? If the former, any spatially aggregated risk metric will be over-estimated.

Lines 565-585: Should the objectives not be stated at the beginning rather than the end?

Line 568: The concept of "informed technical community" is now considered confusing and unhelpful, and in both NUREG-2117 and NUREG-2213—which effectively supersede NUREG/CR-6372)—the expression is "centre, body and range of the technically-defensible interpretations".

Lines 603-604: Why? If there is a ground-motion recording network and large uncertainty in the choice of GMPEs, what is the basis for dismissing *a priori* any update of this part of the model?

Appendix A: I propose that you remove this section to make room for some information about the exposure model. I especially think this should come out since the recommendations only focus on modifying the model as the project proceeds, but no mention is made here on in Section 5 about how the results of the *a priori* study should be used <u>before</u> the study begins. Do the results provide confidence to proceed with the project? Do the results indicate that some changes could be made, such as to location or even building strength? Surely the value of a risk model is to inform decision-making? I would refer the authors to van Elk et al., 2019, Earthquake Spectra, 35(2), 537-564 for an example of an induced seismic risk model designed for such a purpose. If the value of the authors' model is to have a starting point for updating and refining during the project procedures, then this should be stated. And how are the thresholds on the TLS in Figure 3 linked to the risk estimates? Where are the disaggregation results that support these magnitude thresholds? The authors could tie the various elements of the paper together much more than is currently done.

Julian J Bommer
Imperial College London
13th January 2020