Referee #2

Thank you for your comments and remarks, which imply a considerable improvement to the manuscript. We give response to all the points raised by referee 2 below.

 # General comments

This paper is interesting and covers an important topic of interest to NHESS readers.

However, there are several significant weaknesses in the paper, which I estimate will require *major revisions* to address.

Although the major revisions may involve considerable work from the authors, I do think it should be possible to revise the paper to meet the standards of NHESS publication.

Before going into detail below, I note some 'overall' problematic aspects of the current draft:

- Poor referencing. The paper contains no references in the Introduction, despite discussing other work. Other areas also need more references (e.g. often non-obvious 'facts' are stated, without information on where they come from). I note a number of examples below, but not all – in general, the authors need to be more rigorous about justifying statements with references.

We include 33 new references in the document, specially in the introduction.

- There is insufficient explanation of the earthquake catalogue analysis methods (i.e. the discussion on page 3 in the paragraph around line 20 needs to be expanded).

This paragraph is changed, including the appropriate references and explaining the method. The graph of figure 2 illustrates this point.

- A number of statements in the paper are unclear or lacking in rigor (details below).

- I think there is a (fixable) logical flaw in the method. The authors are assuming (equations 8 and 9)

1: region_seismicity = zone_seismicity + fault_seismicity; and furthermore, at various points in the analysis they assume (equations 10 and 4):

2: all of these seismicity sources can individually be represented with GR curves;

The combination of (1) and (2) would be fine *if* all the seismicity sources have the same 'b-value', but the authors suggest different 'b-values' on fault vs region/zone.

Mathematically, I don't think (1) holds under this assumption (i.e. you can't sum 2 GR curves with different b values and expect to get another GR curve). I have suggested a possible fix below, which looks like it may be reasonable based on the data in the paper – but I can't be sure. Anyway, the authors should address this somehow.

Equations 8 and 9 (new 12 and 13) add cumulative seismic moment rates and cumulative seismicity rates. Both are point values and not functions (as GR curves).

The b-value give the ratio between events of different magnitudes and thus it changes the rate of earthquakes produced of each magnitude, but not the total earthquake rate nor the total released moment rate. We adjust the b-value that keeps those values (cumulative seismic moment rates and cumulative seismicity rates) constant for each source.

We agree with you when you state that the GR curves cannot be just added, as they have different b-value.

On the other hand, I expect these issues can be addressed. Furthermore I note there are good points to the paper: it treats an important topic, provides a good application of the methods (which I found much easier to read than the methods description itself), and includes a reasonable attempt to quantify the uncertainties (Section 2.4).

 # Specific Comments

- The Introduction does not contain any citations of other work. However, there is repeated mention of 'other studies' {e.g. p1 Line 25; p1 Line 27; p2 Line 3; p2, Line 22; p2, Line28 ... and many other instances, I will not list them all}. Throughout this section, the authors need to provide references to justify their descriptions of other work (even when they refer to general issues that are not specific to a particular study).

References have been included for all cases

- P 3, Line 15: ** Magnitudes values above MMaxC present recurrence periods higher than the catalog OP or constitute a sample an insufficient number of records to apply statistics. ** – This statement needs to be rephrased to correct the grammar. More importantly I think it is factually incorrect.

For illustration, suppose we have a catalogue of 100 years duration, and this catalogue contains zero events with Mw>=X. Then I claim that we **can apply statistics** to place bounds on the rate of events with Mw>=X, and furthermore, that **the return period of such events MIGHT be less than the observational period**. Both of these points contradict the author's statement. For example, with zero events in 100 years, an exact Poisson test of the true rate of events with Mw>=x (e.g. poisson.test(x=0,T=100) in the R software) gives a 95% confidence interval on the true rate of [0, 0.037] events/year.

So the return period is probably greater than around 1/0.037=27 years. This makes sense – if the return period were very small then it is likely an event would occur in the data. Furthermore, it is obviously possible that the return period is < 100 years (we don't get a '100 year event' in every hundred year observational period).

I think the authors probably want to say that it is increasingly difficult to constrain rates of rarer events – that's correct - but the authors need to weaken the statement accordingly.

We wanted to refer to the fact that when recurrence periods are long, there are a few records of those events in the catalog, and consequently, the catalog presents a small sample, with limited statistical significance to establish recurrence periods. We changed the phrase for:

"Magnitude values above MMaxC present recurrence periods higher than the catalogue OP. These values usually constitute a sample that does not include a high enough number of records to clearly establish the recurrence period, as this makes it increasingly difficult to constrain rates for rarer events."

However, we want to indicate that we refer to recurrence periods (as the inverse of the earthquake occurrence rate) and we do not consider return periods at this point of the work (it is considered in hazard calculations).

- P 3, lines 17-18: The authors should explain the 'Stepp (1972)' approach to estimating the reference years for different return periods. Although Figure 2 is related to this, I cannot understand the method from the figure.

This paragraph is changed, references are included and the method is explained to understand figure 2 better. The reference to Stepp was a mistake that we have changed.

- P3, Lines 19-20: **Then, it is possible to estimate ...** – the authors should explain how they do this estimation. There could be a range of methods (depending on the extent to which one makes parametric assumptions, like that the data-generating process has a GR distribution).

We include two equations to explain how both parameters are estimated. However, we want to indicate that we extract these data from the catalog, and we do not assign a GR distribution in this part of the calculation.

- P 3, Line 24: ** This way we avoid miscalculation problems ...**. I think this statement is too simplistic, and should be rewritten. In reality, errors in the estimated rates may be quite significant even for events with true return period significantly less than the catalogue duration. There are statistical methods for quantifying such uncertainties in different contexts.

For example, for individual rates one might assume a Poisson process and estimate confidence intervals as I suggested in an earlier comment. Or for a full magnitude frequency distribution, one might estimate uncertainty in the rates for any magnitude by assuming a GR distribution, and using maximum likelihood with profile likelihood confidence intervals - or using Bayesian techniques, etc. If the authors google-search 'Gutenberg Richter maximum likelihood estimation' they will find many related references.

The text has been modified: "In this way, we avoid using magnitudes with long recurrence periods that have not been recorded in the catalogue within the completeness periods". Again, we make reference to recurrence periods and not to return periods.

- P 4, Section 2.2: It's not clear to me whether the authors assume there is only 1 fault per region (this is how the math appears, and consistent with the use of the word 'fault' rather than 'faults' in most places), or if there are multiple faults and GR curves which are summed over (that would seem sensible, but no summation is mentioned), or if there are multiple faults which are all treated with a single GR model (but there is no summation mentioned in equation 3, and the word 'fault' is mostly used, rather than 'faults'). This needs to be made much clearer to the reader, with 'summations over faults' included in the equations as appropriate. Table 4 makes me think it is 'multiple faults with single GR model', but it is far from clear.

We consider as many faults per region as available in the fault database. All faults are assigned the same b-value and different occurrence rate (as each fault presents different moment rate). We corrected the ambiguity in the manuscript.

- P 4, line 12: It is unclear how 'v' is estimated (I suppose from other geophysical studies? Or paleo work?). Please provide references. Also, how do you choose the shear modulus? References.

The slip rate (v), can be obtained from paleoseismicty studies and GNSS measurements. We include a reference to database of active faults providing this information.

The shear modulus value is $\mu = 3.2 \times 10^{10}$ Pa (Walters et al., 2009; Martínez-Díaz et al., 2012)

- P 4, line 16: Note that 'm=0' is not the lowest value of m. Maybe just say 'from very low up to the maximum ...'.

The text is modified

- P 4, line 26: Related to the above comment, you might make the lower summation index in the first right-hand-side term be "negative infinity" rather than zero. Also, I think you don't model earthquakes with m < M_min later in the paper? If so, I suggest mentioning that here. Indeed you might re-write this part to avoid mention of events below M_min, perhaps it is just confusing?

We do not include events with magnitude below Mmin to estimate seismic hazard, but we take into account the moment rate related to events with magnitudes below Mmin to estimate the seismic moment rate in the interval Mmin-MmaxC.

$\sum_{\mathrm{M_{Min}}}^{\mathrm{M_{MaxC}}} \dot{\mathrm{n}}(\mathrm{m}) \cdot \mathrm{Mo}(\mathrm{m}) = \dot{\mathrm{M}}\mathrm{o}_{\mathrm{fault}} - \sum_{\sim 0}^{\mathrm{M_{Min}}} \dot{\mathrm{n}}(\mathrm{m}) \cdot \mathrm{Mo}(\mathrm{m}) + \sum_{\mathrm{M_{MaxC}}}^{\mathrm{M_{Max}}} \dot{\mathrm{n}}(\mathrm{m}) \cdot \mathrm{Mo}(\mathrm{m})$ of Eq. (6)

The moment rate assigned to each interval (0- M_{Min}), (M_{Min}, M_{MaxC}), (M_{MaxC}, M_{Max}) depends on the b-value estimated for the fault. Hence, we must take into account this interval until a b-value is fixed.

This part is changed for a better understanding.

- P 4, line 22: I'm not sure if you have defined the $noverline{d}$ variable. Also I don't understand the notation "Mo|" in the numerator (what is the bar? what is Mo? Should it be for faults only?)

The notation is changed and all parameters are defined

- P 5, line 9: Here, I think you should remind the reader that the 'region' parameters are assumed known, based on Equations 1 and 2.

A comment is included

- P 5, line 14: **Note that the b-value of the zone appears in this equation can be equaled to the b-value of the region as both sources present similar seismic nature.** Mathematically I think this is problematic. Question: Is a distribution defined by summing 2 GR distributions with different 'b-values' still a GR distribution? I don't think so. However, that seems to be an implicit assumption of your method (i.e. because the regional seismicity is the sum of 2 different GR distributions, fault and zone, each having different b values). This suggests an internal contradiction in the methods, which should be fixed. Suggestion for a fix: Consider modifying your method so that b-fault is equal to b-region and b-zone. Then I think the issue would be avoided? Furthermore, from Table 4, that would look to be an OK approximation, given the statistical uncertainties in Table 1.

This topic is answered in "General comments". We remark that we do not equal the GR distributions, but only the total seismicity rate and moment rate, both values are concrete and constant for each source independently the ratio of different magnitudes that are calculated later on.

- P5, Section 2.4. I like this analysis overall. However, from Table 1 it seems like you are using a single b value in the simulations (?), in contrast to the above-mentioned 'distinct fault/region b values'. The testing here needs to be consistent with the methodology.

So if you change the analysis as I suggested above then this section is probably fine –but otherwise, you should treat these 'distinct fault/region b values'

The uncertainty analysis focuses on how the input data (seismic catalog and fault database) affect the uncertainty in the end result.

Concretely, table 1 shows how the magnitude interval, the number of records and the b-value of the seismic catalog affects the uncertainty in the seismic moment rate of the region, a parameter that will have a strong influence on the final result for all sources. Thus, table 1 only considers the uncertainties of the seismic catalog, which is used to model the seismic potential of the region.

The uncertainty related to the input parameters of faults is tackled in the last paragraph of the section.

- P7, line 4: Need a reference for the shortening rate.

- P 7, line 14: Reference for the QAFI database?

The text is changed and the references included

- P 7, line 16: Not clear to me how you estimate M-max if there are multiple faults per region. Do you assume they all rupture at once?

For each fault we obtain a different Mmax, as a function of the length of the fault plane

- P 7, line 29: '.. lacks a COV estimate because the sample of records is not significant ..' – I think you need to re-word this. What do you mean by 'not significant'? It's unclear – the word 'significant' is suggestive of 'statistical significance', but that does not appear to be what you mean. I think you just mean 'there is very little data'. However, I don't see why you can't calculate a COV coefficient (albeit a very uncertain one).

Zone 30 only has 7 seismic records. Any statistical analysis with this sample size is not representative. The text is changed to clarify this point.

- P 8, line 13: 'This result is characteristic of this method ...' – sounds like there must be other studies that show this, please add references.

The text is changed

- P8, line 20: "These results agree with real observations" – can you cite the study?

A reference is included

Technical corrections

- P1, Line 6: 'estimated' should be 'estimates'

- P2, Line 18: suggest changing 'a part of' to 'some of'

The text is changed

P2, Line 18: 'must be linked to faults' – is this always true? What about if there are not catalogue events on the faults? Suggest rewording.

We understand that is that fault is active and generates earthquakes with not very long recurrence periods, the events associated to the fault are contained in the seismic catalog. We have rewritten the text to clarify this point.

- P 2, line 26: 'fixing a Mc value results certainly complicated' - This doesn't make sense, suggest rewording. I'm not 100% sure what you want to say – do you mean 'it is difficult to choose Mc non-arbitrarily'?

Yes, it is complicated to assign a Mc value non-arbitrarily'. We have changed the text

- P2, Line 28: ** The approach presented in this paper, as all probabilistic seismic hazard models, face the challenging question of estimating the expected ground motions with the basis of a short period of observations of earthquake occurrences and limited geological data (with significant uncertainty) to construct recurrence models.** This needs to be re-worded to correct the grammar. One suggestion: ** The approach presented in this paper faces the challenging question of how to estimate the expected ground motion exceedance rate, using a short period of earthquake observations and limited geological data (with significant uncertainties). This challenge is faced by all probabilistic seismic hazard models. **

- P 3, L5: **The zone is defined as the source which seismic potential is residual, excluding the seismic potential of faults **. There are grammatical issues here, as stated it does not make sense. I think you mean 'The zone is the same as the region with fault seismicity removed'. However, you may choose to make different edits (considering also repetition in the subsequent sentence).

- P3, Line 10: This sentence needs rewording (grammar). What about: "The seismicity rate of the region is derived from the seismic catalog using events that are contained in the period for which ...."

- P 3, Line 11: You use 'CP' here, but on line 13 you use 'PC'. Please double check for consistency throughout the paper.

- P 3, Line 13: **The complete periods PC(m) (for different magnitude up to a maximum magnitude of completeness value, MMaxC.) are included in the observation period (OP) of the catalog. ** Is it really correct to say they are 'included'. I suspect you mean to say they are 'inferred using' or are 'less than'?

- P 4, line 1: I suggest you replace the word 'cumulative' with 'total'.

- P 4, line 20: 'establish' instead of 'stablish'

- P 4, line 20 and below. You refer to $ndot{Nmin}$ – should it have a subscript 'fault', considering later notation (Equation 8)?

- P5, line 9: 'An new' should be 'A new'

- P5, line 19: I think this should mention 'faults', e.g. 'by extrapolation of the faults recurrence model'

- P6, line 11: should say 'moment rates for different magnitude values, ...'

- P6, line 12: similar problem as above

- P 7, line 26: 'It is appreciated that ..' – this sounds strange – suggest you change to 'Note that ..'. Also, it's not clear

We have changed the text