Reply to comments R1

Dear Graeme,

Thank you very much for careful and helpful revision. We answered positively to all the questions you raised, and hope that the manuscript is now better than before.

Here in the followings the detailed list of your comments is given with our replies, a zip attached contains the revised manuscript, with tracked changes and some figures that have been modified; the quality of other pictures is fixed in the original ones. On behalf of all the authors

Laura Peruzza

R1. Abstract:

Line 16: Change "more standard PSHA" to just "standard PSHA", and change "which are most broadly due to" to either "which are mostly due to" or "which are broadly due to".

Done

R1. Introduction:

Page 2, Line 1: This opening statement might be helped by giving examples of two or three damaging earthquakes that are volcanic in origin (can be in Italy or elsewhere).

We modified the text to account for this comment, referring more clearly to the issue of seismic hazard in volcanic zones and providing some references of major case-studies worldwide.

Page 2, Line 19: "poissonian" needs an upper case "P"

Page 2, Line 30: "effects" should be "affects".

Done

R1. Seismic Source Model

Page 3, Line 13: "engine of earthquake occurrences" seems an unusual phrase. Suggest ". . . basic assumptions behind the physical processes driving earthquake occurrences"

Done

Page 3, Lines 14 – 16: The mention of volcanic tremor here seems to be quite pertinent, yet this is the only place in either of the two papers where this gets mentioned. Volcanic tremor events are usually quite different in their spectral characteristics then those arising due to brittle failure (or those due to general tectonic loading on existing faults). The ground motion model subsequently in this paper does not seem to distinguish between the different types, suggesting that both are present in the database use to fit the GMM. This would inevitably influence the total variability of the GMM. The consideration of tremor events, which one would assume might be more prevalent during periods of raised volcanic activity, may warrant more discussion as possible caveats of the approach presented in these two papers.

The referee's comment is right, but the sentence on the "volcanic tremor" is a mistype which had to be deleted. Actually the events we used to fit the ground motion model are all volcanic-tectonic (VT) events caused by double-couple fracturing and fault slip. In our study, no volcanic tremor events have been taking into account and the database used to estimate the ground motion model, and then the GMPE, include just VT earthquakes. We now modified the text accordingly.

Page3, Line 18: "high-quality instrumental network which geometry and characteristics are essentially remained unchanged" – Change "which" to "whose", and "are" to "have".

Page 3, Line 27: Change "independency" to "independence".

Done

R1. GMPE at MT Etna

1. The choice of the functional form of the GMPE here is not well justified given the description of some of the phenomena observed in the ground motion dataset. The most important inconsistency is in the treatment of focal depth. In the first paragraph of the section the authors indicate that the strong motions have a clear dependence on hypocentral depth, with shallower events richer in frequency content than deeper ones. However in their functional form it appears that they are adopting a model with a fixed pseudo-depth, and an explicit dependence on hypocentral distance. This largely negates the depth dependence as deep events at close epicentral distances can produce similar levels of shaking at shallow events at larger epicentral distances. Some justification is needed by the authors as to why this is preferable to a GMPE in which the hypocentral depth is considered as a separate term in the model.

We did not enter into much details with respect to the empirical GMPE as they can be found in the paper by Tusa & Langer (2016). The model presented here is close to a standard formulation "ITA10" (by Bindi et al, 2010) widely used in the Italian Territory. Nonetheless we now added some more explanation which may match the reviewer's concern. We point out that the "pseudo-focal depth h" was not fixed a-priori but identified during the inversion process in the same way as the other coefficients of the model. Besides, please note that we focus on the shallow events, i. e., events having a focal depth less than 5 km, and about the 95% of the data we used have focal depth less than 2.5 km. Depth dependence for the given data is therefore of minor importance, besides the area very close of the epicenter – for which the data coverage is indeed poor. This problem is not neglected, however, as we propose a specific treatment for this case on the base of synthetic simulation.

2. The coefficient of the anelastic attenuation term (c3) is positive but very close to zero. This is problematic as it means that at longer distances the attenuation trend will reverse and ground motions will increase with distance. This flattening at longer distances is already visible in Figure 3 at about 100 km and will reverse the attenuation at greater distances. Given this unphysical behavior and the fact that the coefficient is barely significant there is little justification for including an anelastic term in the ground motion model.

Again, the coefficient c3 was "inherited" from the form of ITA10. We agree, being it close to 0 it can be probably neglected, but then a reader might ask why we did not consider it as it makes part of the standard formulations. We now state that is parameter could be in reality neglected, as we found it indeed being close to 0. The" unphysical behavior" at larger distance (>100 km) is a result of extrapolation of the model to a distance range not covered by the data. Problems of extrapolating non-linear models outside the parameter range of the data set are common, and must be strictly avoided. We stress this now. By the way, from a physical point of view we must be aware that mechanisms of attenuation may change, for instance for the role of reflections at deeper discontinuities, Q-factors/velocities not being constant along the path or other.

3. The use of nonlinear least square fitting accompanied by bootstrapping is unconventional compared to the more common nonlinear mixed effects regression approach. Could the authors comment on why they adopted this approach rather than mixed effects regression?

We used the non-linear least-squares (NLLS) Marquardt-Levenberg algorithm (Press et al. 1992) because the random-effects regression technique proposed by Brillinger and Preisler (1984) does not converge to a unique solution and depends on the initial model parameters. Additionally, it tends to give approximately the same coefficient values founded through the more classical non-linear least-squares technique. We also added a few sentences to the Boot-strap techniques, which simply consists of resampling the data and repeating the inversion many times. The Boot-strap comes with the advantage not to make any a-priori assumption on the distribution of the data set (commonly Gaussian) and allows a robust estimation of the statistics of the coefficients of the model, i.e., their mean and their dispersion. Here we see that the estimation of the coefficients and their uncertainties are rather stable.

Page 8, Lines 29-30: I'm not sure that the reference to the filename of the tusa_lager_2016.py file is relevant (and it is always possible that in the future the location of the code archive can change). I suggest

to remove this particular mention and simply refer the URL to the main OpenQuake archive or repository from which the software can be accessed.

Combining R1 and R2 suggestions, we modified the sentence in lines 27-30 page 8.

R1. Accounting for topography

Page 9, Line 5: Replace "we have done" with just "introduced"

Done

This section is an interesting development on the conventional assumptions made in PSHA. There are two issues though that the authors may wish to comment upon. The first is the potential influence of topographic amplification effects that may mean that two sites at similar elevations and source-to-site distance may experience different amplitudes of shaking (and different frequency content) depending upon the local gradient of the slope or the proximity to steep edges. Is there evidence for topographic effects in the residuals of the ground motion model?

Secondly, the assumption of hypocentral distance, though simple and practical, has its own inconstancies when the hypocentres are located above the reference surface. If the source is located above sea level within the volcanic edifice yet the site is located away from the base then a linear path from the source to site cannot be constructed as it would be for events below the surface, as it implies travel of the waves above the free surface. This is perhaps another argument as to why hypocentral distance may be an unsuitable metric in the present case and an explicit hypocentral depth parameter may be able to better account for changes in elevation

We modified the text to address these two issues related to accounting for the topographic surface in PSHA (see Page 9, Lines 13 – 23). We clarify that we are not accounting for topographic site effects and justify the use of a linear source-to-site path approximation.

R1. Accounting for Site-Specific Response

Generally a good section, although some comment from the authors as to whether the local site conditions of the volcanic edifice are consistent with the typical conditions needed in order for HVSR to be a good indicator of amplification (i.e. strong impedence contrasts, minimal lateral heterogeneities etc.) is needed.

We added a new sentence aiming at supporting the use of spectral ratios methods in a volcanic environment.

R1 Results

Pages 12 – 13 (Lines 31 – 3): There is some information that is unclear here, possibly suggesting a slight inconsistency in the methodology. The authors describe the use of floating ruptures on fault surface, yet they don't indicate the choice of magnitude scaling relation or how they distribute the location of the hypocenter within the rupture. Given the magnitude frequency distribution is characteristic then one would expect that the finite rupture dimensions would be close to the total area of the fault. Thus, floating ruptures would not necessarily move the rupture to a great extent. If the hypocenter is located mainly at the centre of the rupture then the resulting sources will be distributed very closely around the centroid of the full fault surface. If the intent is for the hypocentres of ruptures on the fault source to be constrained to a small area within the middle of the source volume, as one might do if one wished to infer bilateral propagation of ruptures on the fault, then the results make sense given that from the GMPE a more uniformly radial pattern is expected. However, as GMPE itself does not require information regarding the rupture finiteness, it may be preferable to consider aleatory uncertainty in distance by considering the hypocentres as points on the fault plane. This would mean that hypocentres, rather than ruptures, can be floated across the fault plane and the pattern of the map should look less radial (though the peak level of the hazard in the centre of the fault will be lower). Some clarification is needed from the authors as to how they have generated the finite ruptures being floated, the distribution of hypocentres within the rupture plane and whether the resulting distribution of sources conforms with the modelling intention (e.g. whether bilateral propagation of ruptures is desired).

We modified the text to account for these comments, describing more clearly the MSR, rupture floating, hypocentre location, and the observed pattern (see Page 13, Lines 17 – 25).

Page 13, Line 2: "In this cause" – should be "In this case"

Done

R1. Conclusions

Page 14, Line 18: "strongly spatially uncorrelated" – Not sure how one can discern a strong uncorrelation from a weak uncorrelation? Simply "spatially uncorrelated" is sufficient.

Figures/Tables:

Table 2: Exponential notation shown in the probabilities columns is broken over two lines. Try to avoid this (or fix in the final typeset).

All Done

Figure 1: Image quality is low. Please use a higher resolution in the final submission.

Figure 3: Image quality is very low. Please use a higher resolution in the final submission.

Figure 5: Screenshots are usually of too low a quality for journal publication. Text is not legible. Higher resolution maybe difficult to obtain and it is not obvious that these figures add specific value to the manuscript. If quality cannot be improved consider removing this figure.

The original images have higher resolution, we slightly modified figure 5 as given in the attached zip file of figures

Reply to comments R2

Dear Anonymous Reviewer,

Thank you very much for careful and helpful revision. We hope we have satisfactorily answered to all the questions you raised, and appreciate your comments, as the manuscript in our opinion is now better than before.

Here in the followings the detailed list of your comments is given with our replies, a zip attached contains the revised manuscript, with tracked changes and some figures that have been modified; the quality of other pictures is fixed in the original ones. On behalf of all the authors

Laura Peruzza

R2. Abstract

Reference of Ordaz et al., 2013 is cited in order to justified the used of software CRISIS 2008. Nevertheless the software package has been upgraded since this publication with CRISIS 2012 and more recently CRISIS 2015. Why the last version of CRISIS has not been used? => https://sites.google.com/site/codecrisis2015/home

We modified the text to make clear that the most recent release available has been taken into account, with updates that we developed that are now available for users. The references are old, as not all the new versions of software packages have an accompanying paper.

It is well explained in the text of the paper why M>6 regional seismogenic sources have not been taken into account. But in the abstract it is not clear why these faults have not been considered, as it is standard in PSHA studies, even if calculation are for short return periods (5-30 years). This fact could be repeated because it is an important limit of the study to focus on short return periods.

We acknowledged the limit in the abstract, and briefly comment a simple disaggregation tests that we performed (Figure A1-4) for the benefit of the discussion only.

R2 Introduction

I have no comment on this section.

R2. Seismic Source Model

Difference between level 1, level 2 and level 3 are not very clear in the text. Nevertheless the figure 2 is remarkable and explains well these differences. Figure 2 should be largely cited in order in order to make sure that this paragraph is more readable. It is not clear if maximal magnitudes used for Source Model level 1 are for the 4 zones Mmax 5.2 or they are the maximal magnitudes indicated in Table 1 as in Source Model Level 2. It should be specified. Even if it is written in line 30-31 of page 4 "that the Pernicana Fault is always modelled as GR, not having historical observations supporting the choice of a characteristic earthquake model", characteristics of historical and geological-kinematic models are indicated in Table 2 that is confusing.

We modified the text to eliminate duplications and hopefully to increase the readability of this chapter.

Page 4, line 19: XIX century => XIXth century.

Done

Definition of Figure 1 must be improved

The original image has higher resolution, some minor bugs fixed (coordinates were wrong, now corrected, surface Etna outline on the top)

R2. GMPE at Mt Etna

This section is very interesting but it is in my opinion unbalanced: it is too developed or, on the contrary, too short and too superficial. A distinct paper could maybe favorably describe the procedure and the results to develop a specific GMPE for this Etna case study. Developing a new GMPE is always touchy and must be carefully studied. So this part should be develop if it serves as reference for a new GMPE in

a volcanic context or it should be simplified if the GMPE is just an element of the PSHA, as in this study. For instance lines 27-29 page 8, lines 1-3 page 9 lines 18-30 are not indispensable. In particular the lines about the standard deviations are not useful for the study because it is not used in the following of the paper. The conclusion "higher inter-station variability than the inter-event component, suggesting that the local site effects are the main source of ground motion variability for our data" is not used in the following, mainly in the dedicated part on site-specific response. Nevertheless, it could be interesting to compare the variability estimated here with the differences in the site-specific response part.

We tried to keep it short as many of the issues are already explained in a previous paper by Tusa and Langer, 2016. Nonetheless we keep the information regarding the statistical parameters, as the GMPE presented here uses hypocenter distances instead of epicenter distance, so reporting the statistics is mandatory. Moreover, we removed the lines that you defined "not indispensable" from the original version of the manuscript.

Line 11, page 8 it is written that "about the same values" are predicted by ITA10 and this model near the source. On the Figure 19 we note that the difference for 'soil A' could be about a factor of 3 for distance less than 1 km (about 30 cm/s/s for this study and more than 100 cm/s/s for ITA10) and so these are not about the same values. Moreover there are no data to check this observation because on the figure 3 we note that there are no data with d<3 km, for 3.9<M<4.1. This paragraph could be formulated with other words even if the idea is well explained.

Only the number total of data (1'200) used for the calculation of this GMPE is indicated and not the repartition of these events. On the figure 3, only 3.9<M<4.1 events are represented. Moreover on the figure ESM1 no data are represented. As usually done in articles which deal with a new GMPE it could be useful to indicate the repartition of the data used in terms of magnitude, distance, focal mechanism, depth, soil classification.

We rephrased the sentence in p8 / line 11, following your suggestion. Concerning the figure 3 on page 19 we should remark that ITA and our results converge ± for soil B and D at low distances, but strongly diverge at larger distance for all soil classes. The nationwide used ITA10 turns out to be conservative close to the epicenter, and overly conservative at larger distances.

The black dots are shown in order to show some real data, certainly the cover is poor for M=3.9…4.1 as only few records are available for this magnitude class and soils A or D. We do not see a way to fix this, like many other papers on empirical GMPE, where certain parameter ranges are often poorly covered.

Concerning a figure ESM1, we can refer to Tusa & Langer, 2016, where the magnitude-distance coverage is shown in all detail. In fact, in the original version of the manuscript, page 6 and lines 28-29, we wrote "For an extensive description of statistical characteristics of the data set, for example, in terms of number of data–magnitude, number of data-recording site class, and distance-magnitude distributions, we direct the readers to TL16". Therefore, we wonder whether it is worthwhile to add figure regarding the partition of the events, which would equal to the ones in Tusa and Langer.

R2. Accounting for topography

This section is very interesting because it is a configuration that is not described and taken into account usually. It is quite well explain in the text and particularly on Figure 4.

Lines 14 and 17, Page 9: functionalities and version of CRISIS are cited but no reference are given.

The new functionalities of DEM/topographic surfaces are available in the actual release of both software, but not documented by specific publications.

Definition of Figure 5 must be improved: screenshots are usually of too low quality for publication.

The original images have higher resolution, we slightly modified figure 5 as given in the attached zip file of figures

Legend of Figure 6 must be completed with indications for a), b), c), d): for instance, a) Etna DEM, b) results without DEM, c) results with DEM and d) difference between b) and c).

Done

R2. Accounting for Site-Specific Response

This section is very interesting. It is developed and well-argued while that is not taken into account in lots of PSHA studies.

Lines 7 and 12, page 10: no reference are given for HVSR method or Konno-Ohmachi filter

Following the reviewer's advice we added some references for HVSR, HVNR and Konno-Ohmachi filter

Line 23, Page 10: "the incoming waves are assumed to be travelling vertically": this strong hypothesis should be checked and at least justified. Seeing Figure 4 (that is nevertheless not at scale 1:1) it is not obvious that incoming waves could be assumed as travelling vertically. Moreover this hypothesis, and the fact that horizontal and vertical motions at bedrock have no amplification, that are usually assumed in order to determined site-specific response, should be justified in this very particular case of volcanic context. In our opinion only shallow events may be considered in such ideal conditions.

Figure 4 is not used in the manuscript for the assumptions concerning the 1D site response modelling. The hypothesis that "the incoming waves are assumed to be travelling vertically" is stated by the author of the 1D modelling software as we now better specify in the manuscript adding also the relative reference. Moreover, we agree that volcanic areas are a complex geologic environment, however the adopted techniques showed appreciable results (see also answer to Reviewer 1).

Line 30, page 10: "The input parameters are the Vs, the density (), and Q": it is not useful because it is already indicated in lines 22.

The sentence was removed following reviewer's suggestions

Line 10, Page 11: "the good matching" of the comparison could be specified using criteria as following SESAME criteria.

Following the reviewer's suggestion we modified the sentence specifying that, as stated by SESAME, ambient vibration method is a reliable alternative approach to earthquake records especially in estimating the fundamental frequencies of a site.

Lines 18-20, Page 11 are not useful, it is already clear with the sentence that SESAME criteria are used

The sentence was removed following reviewer's suggestion

Line 28, Page 11: algoritm => algorithm

Done

Resolution of Figure 7 is very low and it is not possible to read neither the name of the stations (7a) that amplification function plots (7b).

Figure 7 was improved in quality by increasing the label of the station names in (7a) and the amplification levels in (7b): see attached zip file of figures

R2. Results

Presented results are interesting and well explained. However no reference is done about uncertainties despite the fact that it is the principal interest of a PSHA study with reference to a classical DSHA.

For instance in the text aleatory uncertainties are pointed on parameters of "historical" or "geological" models, on the GMPE (inter-event, intra-event, inter-station, intra-station). It is written that GR methodology is used, we can assumed that there are also uncertainties on α, β, or $\lambda(5)$ parameters, ... Moreover epistemic uncertainties could be associated to other parameters (Mmax, Mmin, depth, choice of source model, choice of GMPE, choice of the software CRISIS or Open-Quake, . . .). In our opinion it seems very important to underlined these uncertainties and indicate if results that are presented are close to the median, the mean, or other percentile of the set of simulations that could done using the same parameters.

The main aim of this paper is to account some computational implementations done for PSHA in volcanic areas. In particular, epistemic uncertainties due to different modelling choices are considered via the logic tree. We made clear that the results of the logic tree correspond to the mean. Aleatory uncertainties in source modelling are accounted for and described as much as we can, in the companion paper PART I.

Moreover a logic tree that summarizes all the epistemic uncertainties that are taken into account for the study seems to be indispensable, because of the lot of hypothesis that are taken into considerations. For instance in line 26, Page 12 it is indicated that "the four models in level 3". It may significate Historical-Poisson + Historical-Time dependant +Geological-Poisson + Geological-Time dependant hypothesis? The weights of each branch of the tree are often not very clearly explained in this study and this figure of the logic tree could permit to clarify that point.

We now state explicitly that the "the four models in level 3" correspond to Historical-Poisson + Historical-Time dependant +Geological-Poisson + Geological-Time dependant hypothesis, and each are assigned equal weights of 0.25.

Finally a desagregation of the results could be assumed in order to justify some results. For instance Etna-specific PSHA desagragation could be compare with desagregation of the results of national hazard map in order to show that M>6 are not significant in the estimation of the hazard in this region.

This is a good suggestion, in order to assess the contribution of regional M6+ events to the hazard, as a justification for not including them in our model. We performed a simple disaggregation in terms of magnitude and distance considering only the area sources of Level 1 (time independent model) and 4 area sources from the MPS04 national hazard model (those closest to Mt Etna). The results confirm that the hazard at Mt Etna is dominated by the local volcano-tectonic events for exposure periods of 5 and 30 years, and not the larger (M6+) events in the region. For T=5 years, the contribution of regional events is almost negligible, while for T=50 years their contribution is larger, although still less than the local volcano-tectonic events. Some images (Figure A) for the benefit of the discussion are added here, but not in the manuscript/electronic supplementary material, as we believe they are very preliminar, and this point is out of the scope of this paper
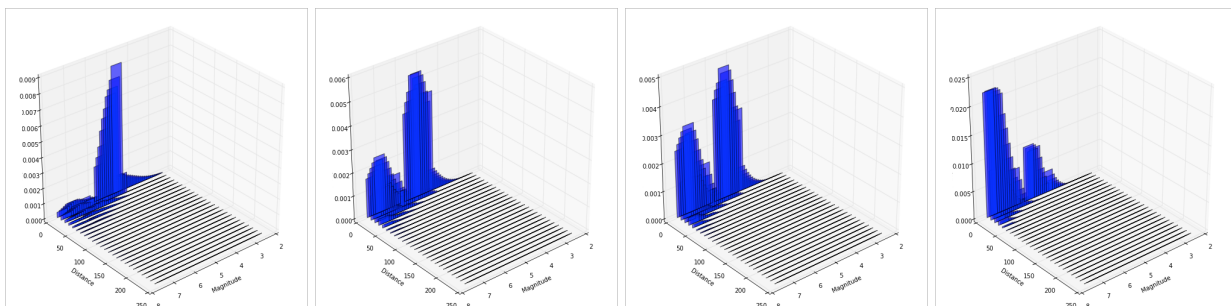
Figure A: disaggregation plots (magnitude, distance, percentage) at 10% exceedence probability for 4 exposure periods, from left to right respectively 5, 30, 50 and 1000 years.

It is noticed in Lines 4-8, Page 13 the main results of different tests, in terms of maximal accelerations. It could be specified the type of difference that are observed in terms of geographical representation, frequency content, velocity, displacement, . . .

We are not sure we understood the reviewer's comment, that part refers to results given in Spectral Acceleration at 2 periods (0.2 and 1.0s) instead of PGA (theoretically SA at 0.0s). They are motivated as they are relevant for site-specific studies and risk analyses.

The justification of using CRISIS and not Openquake for site-specific test is not given.

The quantitative difference between the two software's is not given. What is the percentage of difference between the results respectively? It is rare to use in a PSHA study two software's even if we know that the choice of the software is a source of epistemic uncertainties and it is badly not exploited in the results of the study.

In this paper we made a conscious choice NOT to focus on the differences in the software results (notice that we did not show any of the same results using the two software). Rather, we intended the focus to be more about the new functionalities we developed for both the software (topography, customized GMPE and MSR) and to the impact that some model components may have on final results. This was for instance the reason to explore some other functionalities such as the site effects, even if they are still "immature" (they are mapped accordingly to how they were available at the time of the end of the project that promoted and funded this paper) for a complete analysis of the study area.

The observed intensities (Azzaro et al., 2008) are only cited in introduction and in the last sentence of the results. It could be interesting to indicate main results of this study of 2008 and compare to intensities that could be derived from PSHA accelerations in the same places.

In previous papers we demonstrated the impact of the local volcano-tectonic sources by means of macroseismic data. This paper roughly confirms the previous results for the inhabited areas (where macroseismic data can be assessed), but new sectors of high hazard appear at Mt Etna most deserted areas. We add a brief comment on that in the results, but we did not perform a PGA/SA – Intensity comparison, as in our opinion this is a very "delicate" and "uncertain" procedure.

Results on Figure 9 are presented for mean values? Median values? Mean plus/minus one standard deviation values?

They are mean values, more clearly stated in the caption now.

The figures on Figure 9 and 10 are shown horizontally (with legend at the top) while on the figure 11 the figures are show vertically, with the legend in the right. Also it is far more difficult to make comparisons and it could lead to errors in the interpretation.

The pictures are complex but we hope they will be properly published without rotation. The quality of figures is increased in the original picture.

R2. Conclusions

The main conclusions and limits of the study are well summarized in the conclusions.

The exploitation that could be done of these results is also presented.

However it is unfortunate that the comparisons between CRISIS and OpenQuake are not properly exploited.

We will consider these suggestions for a future analysis, and invite the reviewer to post additional comments if interested.

Line 25, Page 13: "47-284" => "47-284 years"

Line 20, Page 14: "DPC" => "Dipartimento della Protezione Civile".

Done

Bibliography

The same typo must be respected for all the references. For instance:

Line 22, Page 15: => Azzaro R, D'Amico S., Tuvè T:

Line 4 to line 15, Page 17: alphabetical order must be respected

Partially fixed, to be checked at the final submission with the Editorial Staff.