16th World Conference on Earthquake Engineering, 16WCEE 2017

Santiago Chile, January 9th to 13th 2017

Paper N° 2380 (Abstract ID)

Registration Code: S-I1461924602

ROBUSTNESS AND STABILITY OF PSHA RESULTS

P. L.A. Renault[(1)]

[(1)] *Managing director, swissnuclear, philippe.renault@swissnuclear.ch*

Abstract

In the past decades, many probabilistic seismic hazard analyses (PSHA) have been performed on a regional or even site-specific level. Today, some of them have been updates of previous estimates of the seismic hazard. This allows a discussion on the robustness and stability of PSHA results over time and also to reflect on the trend. It has been accepted in the scientific community that it is very difficult, if not impossible, to test PSHA results against data and check their validity through comparison.

The paper discusses aspects of quality assurance as well as established processes of expert elicitation when dealing with large uncertainties. Scientific progress in earth sciences and new data is evaluated with respect to the need for updates and in light of their complexity. In the framework of codes and regulations, the question on robustness and stability – or even longevity – of results becomes very important. The latter can play an important role in the context of acceptability and credibility. With the help of some examples, the issues in evaluating the robustness and stability of PSHA results will be highlighted and future challenges will be shown. The case study used to demonstrate robustness is based on the PEGASOS and PEGASOS Refinement Project (PRP). In the conclusions, the paper makes an attempt to discuss and introduce criteria for the robustness and stability of PSHA results. To date, there is no measure for rationally evaluating the robustness of PSHA results. But there are qualitative means when comparing updated results with previous estimates. It is proposed to use a change of larger than 20% as quantification for a 'significant' change in the results.

Keywords: PSHA, Robustness, Stability, Quality Assurance, Comparison, PEGASOS

1. Introduction

In the past decades, many probabilistic seismic hazard analyses (PSHA) have been carried out. Some have been on a regional level (e.g. for national seismic hazard maps) and others have been site specific (e.g. for critical infrastructures such as important lifelines, bridges or power plants). It is well known that seismic hazard estimates have associated large uncertainties, especially in the case of low to moderate seismicity regions, where data is scarce. Today, some analyses are updates of previous estimates of the seismic hazard. While the median hazard of repeated seismic hazard estimates might have been within the same order of magnitude as in the previous studies, one can observe that there has been a tremendous effort to better estimate the epistemic uncertainties and thus to reduce them. All this available data and knowledge allows a discussion on the robustness and stability of PSHA results over time and also to reflect on the trend.

Scientific progress in earth sciences and new data are evaluated with respect to the need for updates and in light of their complexity. Usually after a new significant earthquake event, a lot of work is carried out to check the previous existing hazard estimates and to compare the prediction with the real data. In some cases, this triggers the need for revisions of the hazard estimates or even the design spectra defined in codes and standards. The new available data allows the refinement of the models and knowledge of the local seismicity, but often also raises more questions than providing answers. In the field of natural hazards, and here especially in the case of seismic hazard, there is an interplay between more knowledge to improve the modelling of the earth's behavior and the robustness and stability of hazard estimates as they are specified in regulatory codes and standards. Those codes and standards are provided as a solid design basis and thus are not expected to change frequently or after each new (relevant) earthquake. In the framework of codes and regulations, the question of robustness and stability – or even longevity – of results becomes very important. In a regulatory framework, the robustness and stability of the results have also a different meaning. The results must be traceable and reproducible, irrespectively of which person or group performs the study. Here the robustness is tied to an invariance with respect to the author of the analysis. In the past, the example of the 'Livermore' and 'EPRI' studies at sites of nuclear power plants for the central and eastern United States led in the USA to the development to the Senior Seismic Hazard Analysis Committee (SSHAC) guidelines [1] and recently to a companion document based on the current experiences with the SSHAC process [2]. The latter is expected to be used for seismic hazard assessment whenever NRC guidance or communications call for use of NUREG/CR-6372 or the SSHAC guidelines.

Longevity plays an important role in the context of acceptability and credibility. Designers and society rely on a robust design, which implies longevity of the design and built system, structure or component. This third aspect might sound trivial from the designer or regulator perspective. But in the framework of evolving knowledge about earthquakes and seismic hazard analyses, this becomes a challenging balance to achieve. In order to improve the existing PSHA results, scientific research is considered and carried out, leading to new insights and thus the potential to impact existing results. In most regulations, the consideration of all technically defensible interpretations and state-of-the-art models is a requirement. This is a closed loop and should over time lead to a robust and stable estimate as well as increase safety. But depending on the defined time frame to close the loop this might be a contradiction to 'longevity'. Furthermore, the SSHAC guidelines do not use the term 'robust' – even if it would seem more adequate as we will see – but claim to ultimately produce stable and long living results by following the proposed methodology.

> The objective of the SSHAC Project is summarized as: "Because PSHA results can be so important for both engineering design and public-policy decision-making, a goal of this project is that the PSHA methodology will ensure the stability of the numerical results for a reasonable period of time (five to ten years) or until significant new technical information presents itself." (Chapter 1.4 and also 1.7 in [1]).
>
> And furthermore: "A desirable outcome of a Level 3 or 4 study is increased longevity and stability of the hazard assessment. This means that the numerical results of the hazard analysis can be expected to remain stable for a reasonable period of time after the completion of the hazard study." (Chapter 3.3 in [2]).

In order to better understand the requirements of robustness and notion of stability, it is important to discuss those two terms and distinguish them. The next section will provide a brief overview of robustness and stability in a more general context in order for the reader to recognize the difference and underlying assumptions and concepts.

2. Discussion on robustness and stability

The first question that comes to mind is: "What's the difference between stable and robust?" Answering the question isn't made any easier by the fact that 'robustness' has multiple, sometimes conflicting, interpretations – only a few of which can be stated with any rigor. Researchers who work with quantitative models or mathematical theories have addressed this issue (see e.g. [3] and [4]).

Loosely speaking, a solution (meaning an equilibrium state) of a dynamical system is said to be stable if small perturbations to the solution result in a new solution that stays 'close' to the original solution for all time. Perturbations can be viewed as small differences effected in the actual state of the system: the point of matter of stability is that these differences remain small. A dynamical system is said to be structurally stable if small perturbations to the system itself result in a new dynamical system with qualitatively the same dynamics. Perturbations of this sort might take the form of changes in the external parameters of the system itself, for example.

So what's the difference between stability and robustness? In the literature, it is argued that robustness is broader than stability in two respects. Firstly, robustness addresses behavior in a more varied class of systems and perturbations; and secondly, robustness leads to questions that lie outside the purview of stability theory.

Robustness is a concept appropriate to measuring feature persistence in certain contexts; namely, systems where the features of interest are difficult to parameterize, where the perturbations represent significant changes either in system architecture or in the assumptions built into the system through history or design, or where the system behavior is generated through adaptive dynamics coupled to strong organizational architecture. The study of robustness then naturally prompts questions relating to organization, the role of history, the implications for the future, and other questions even more difficult to formulate relating to creativity, intentionality, and identity.

For instance, there are arguments that much of the complexity in sophisticated engineering systems stems not from the specifications for functionality, but from the exigencies of robustness [5]. The authors argue that in traditional engineering design, regulatory mechanisms for robustness are typically superimposed after the fact on the mechanisms for functionality, and that Rube Goldberg ("doing something simple in a very complicated way that is not necessary") prevails as a consequence.

Intuition tells us for example that there are tradeoffs between robustness and evolvability. Robustness may be seen as insensitivity to external and internal perturbations. Evolvability on the other hand requires that entities alter their structure or function in order to adapt to changing circumstances – in our case e.g. the improvement of data and knowledge through data collection or new research.

The use of robustness as a design principle raises a deep set of questions as to the nature of and the response to assemblages of assaults previously unencountered and in a real sense unforeseeable. What design principles should guide the construction of systems for which there exists an infinite number of possible assaults to which the system may be subjected (e.g. the known unknowns and unknown unknowns in PSHA)? The possibility of using joint probability distributions to estimate the likelihoods and consequences of failure is fairly dim here. It is rather challenging to answer the question how tools can be developed to endow a system with "open-ended robustness".

As a particular example of the challenge of modeling assaults, what is the difference between designing robustness against 'purposeless' versus 'purposeful' perturbations or attacks? To first order, stability theory can be said to address the consequences of unintended perturbations. By contrast, as pointed out by [6] and [7], the robustness of computer network security systems is an example – as is the rule of Cosimo de'Medici (see Section 9) – in which it is necessary to posit instead the existence of attackers intimately familiar with the vulnerabilities of the specific system, and in possession of the expertise and resources needed to mount a coordinated attack explicitly designed to cripple or to destroy. Robustness to this form of attack clearly calls for design that includes the ability to learn, to anticipate, and to innovate.

In its weakest form, the argument for robustness as different from stability can be stated as follows [5]: "Robustness is an approach to feature persistence in systems for which we do not have the mathematical tools to use the approaches of stability theory. The problem could in some cases be reformulated as one of stability theory, but only in a formal sense that would bring little in the way of new insight or control methodologies."

As a last remark it should be noted that a different view of the interplay between organizational architecture and robustness emerges from the study of certain hierarchical systems. As pointed out in the previous section, the discussion of robustness for such systems has meaning only when the level of the system is clearly identified. Robustness may exist on the level of the individual components, or on an intermediate level, or on the level of the whole, or not at all. Robustness on one level need not imply robustness on any other level. Conversely, robustness at one level may – through processes serendipitous or otherwise – confer robustness at another level.

3. Quality Assurance in PSHA

As a matter of fact, no explicit literature exists on how to check seismic hazard analysis results for robustness or stability. In practice the definition of the process employed and associated quality assurance measures are used as a substitute. Lately, a practical example of the application of quality assurance for logic-tree implementations has been published [8], dealing with the state-of-the-art quality assurance procedures and products employed in industrial PSHA projects.

The PEGASOS project introduced for the first time the so called hazard input documents (HID) as a quality assurance measure. The HIDs represent a rigorous collection of expert approved synthesis of parameter and model description which allows an unambiguous hazard computation by the hazard analyst. This step not only improved significantly the documentation of a PSHA, but also allowed a transparent and detailed review of the PSHA. Other international projects have afterwards also adopted the concept of HID and through this improved the state-of-the-art in industrial, commercial PSHA. Today, this is standard practice.

Of course the quality assurance is a prerequisite for correct PSHA results, but it can only be applied to the individual pieces of the PSHA. Input data, models, computations, etc. can be quality assured, but this doesn't guarantee in the end a stable and robust result. The same argument can also be applied to testing the PSHA. Testing PSHA results against data and checking their validity through comparison is very difficult, if not impossible.

4. Comparisons and a case study

To the author's knowledge, only one publication exists which compares and discusses results of various PSHA studies. In [9] the main emphasis is on the comparison of the (epistemic) uncertainties among the different studies, but, nevertheless, it provides an excellent summary and overview of past high quality PSHA studies.

A part of the scientific community does not believe in PSHA and rather votes for a deterministic or neo-deterministic perspective and approach, respectively. These two concepts are complementary and not exclusive, the point being that, in the context of politics and society, an informed decision making is necessary which tries to handle various risks and their associated probabilities. Of course, it is not directly possible to check earthquake probabilities for specific or occurred events, but recent work has shown the benefits of testing PSHA results and using Bayesian techniques to update hazard estimates [10]. Some authors claim that PSHA results are directly testable by existing data [11][12]. This is not as trivial as it sounds at a first glance. Today, if a hazard study – regardless of whether it is deterministic or probabilistic – is done properly, all available data and its associated uncertainties are considered in the development of the hazard estimate. As seismic data for large events is rare in regions of low to moderate seismicity, all of the data must be used to get a reasonable estimate. Thus, in the end there is no independent data set left to test with, otherwise it would be a circular argument. Furthermore, comparing a specific local event to PSHA results is pointless, as nobody can assign an annual probability of exceedance to the new event. A comparison can only help to evaluate whether the existing models are able to explain the specific new earthquake and how it compares to the predicted average or how 'unusual' it was if it is well beyond the mean hazard – but still within the probability distribution.

In the following section robustness and stability of PSHA results will be discussed with the help of an example. The aim is to see if the two terms apply and if there is a quantitative measure for it. As mentioned above, by following a SSHAC process one should obtain stable and long-lasting results. Yet, it is not clear how one stand-alone result can be checked for robustness or stability – apart from relying on the SSHAC promise that the process will lead to such. However, here the adopted approach is to use site-specific PSHA results which were carried out for a given site multiple times in order to evaluate their relative change from one study to another, independently of the process which was followed to obtain them. In theory, if a PSHA is performed honestly and correctly, all unknowns and thus, uncertainties are properly captured, leading to a large uncertainty of the overall result if there is lack of data. Assuming that the lack of data and knowledge decreased over time, the next PSHA for the same site / region should be able to make use of the new data and models in order to improve the PSHA results. Here, the hypothesis is that this is measurable by checking the change in median hazard and the change in the uncertainty range by looking at the outer fractiles – which ideally should decrease within the range of the previous range.

4.1 Case study

The PEGASOS Project, a new state-of-the-art probabilistic seismic hazard assessment for the nuclear power plant sites in Switzerland, was carried out from 2000 to 2004 [13]. The quantification of the epistemic uncertainty and aleatory variability in seismic hazard at the four Swiss nuclear power plant sites was the key aspect of the PEGASOS Project. After the completion of the project, the Swiss utilities decided to perform a refinement of the study by collecting additional data and using new advances in earthquake science, especially in the field of ground motion modelling, to further reduce the identified epistemic uncertainties by their more realistic quantification. The PEGASOS Refinement Project (PRP) started in 2008 and was completed in 2013. An overview of the different components of the project and the approaches used are documented in the project reports [14]. Some new challenges for the earthquake science community were also identified. An important aspect of the ground motion characterization that had a significant impact on the PSHA results and associated uncertainties is the adjustment of ground motion prediction equations to make them applicable to the site conditions in the study region.

The PEGASOS Project and PRP studied the seismic hazard for four sites twice within a decade. Furthermore, an independent national seismic hazard update was completed in 2015 by the Swiss Seismological Service (SED) [15], providing the unique opportunity to compare the evolution of the results for the same site. PEGASOS used a generic rock with a presumed V_{S30} = 2000 m/s. In PRP, site-specific rock V_{S30} values were defined, which makes it more difficult to compare results at the rock directly. Thus, for a true comparison on rock (and soil respectively) the SED rock hazard results (based on a generic rock with V_{S30} = 1100 m/s) were converted to the corresponding NPP site-specific rock and soil definition.

For the subsequent comparison, only one site out of the four investigated is shown. It is assumed that the provided return period of 9975 years (≡1.0025E-4) by SED is equivalent to the annual probability of exceedance of 1E-4 defined in PRP.

While it is true that the median curve is often preferred to the mean curve as a central measure of hazard due to its stability, a clear rationale for this practice or, more generally, a procedure for dealing with epistemic uncertainty in decision making is not presented in the SSHAC report and is not the subject of this paper. Here the focus is on the range of the results which lies within the range of the previous hazard estimates when doing a comparable elaborated study. Considering this, one can claim that there is robustness at least in the new results, as they are not outside the range of the previous ones. If the latter should be the case, then this would indicate that fundamental aspects of hazard and associated models were not well captured in one or the other study.

The robustness and stability of the new PRP results compared to the old PEGASOS results can best be assessed when comparing the median and range of variability of the results. As can be seen from the illustrative example in Fig. 1, the median hazard results are very consistent and the range given by the fractiles demonstrate the achieved reduction of uncertainties. In the view of the author, the median can be used as a measure for stability and would fit within its definition, while the overall results – including the range of fractiles – are reflecting robustness, as they are influenced by 'external and internal perturbations'.

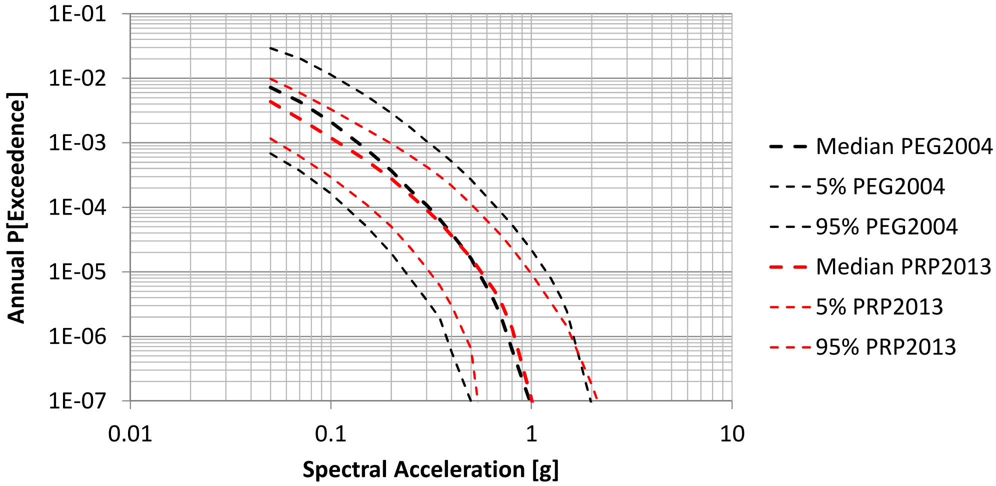

Fig. 1 – Soil surface hazard comparison for a specific site at PGA, horizontal ground motion, for PEGASOS 2004 and PRP 2013. Median and fractile curves.

Fig. 2 shows the comparison at the specific example NPP site of the rock uniform hazard spectrum (UHS) for PRP and SED 2015 in terms of mean, median, 5% and 95% fractiles. The SED rock UHS consists of 11 spectral acceleration amplitudes for the given 11 frequencies. The PRP results are based on 9 project specific frequency points. It needs to be noted that the SED UHS has no intermediate points between 20 and 100 Hz, while PRP uses additionally 33 and 50 Hz in this range. Thus, the shape of the two curves between 20 and 100 Hz is not really comparable even if here for the graphical representation it has been connected with a straight line. Furthermore, it must be stressed, that the SED 2015 hazard was developed independently of the PRP results and is based on assumptions and decisions taken by independent SED experts; even though of course the SED hazard might have profited from work done in PRP, as it was only completed two years after the PRP.

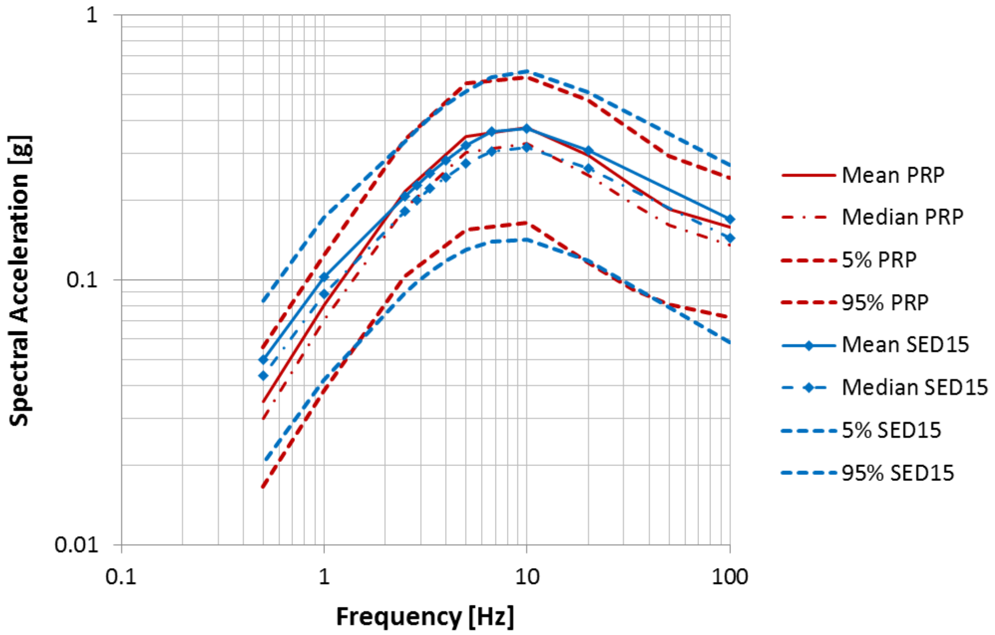

Fig. 2 – Rock hazard comparison for a specific site at an annual probability of exceedance of 1E-4, horizontal ground motion, for PRP 2013 and SED 2015.

The comparison at rock (when converted to the same rock reference) provides remarkable consistency between SED and PRP results in terms of mean, median and fractiles. It is worthwhile mentioning that there are fundamental differences when developing site specific hazard (PRP model) and regional hazard results at the scale of country (SED model). Nevertheless, what can be seen in this specific comparison is that, despite the different model assumptions used to compute the hazard, the resulting hazard for the same geographical location (here shown in terms of UHS at 1E-4) is comparable and consistent. The latter is also an indication for robustness, as two different parties came to almost the same result, using the same available data but different approaches.

Fig. 3 and Fig. 4 show the comparison of the soil surface hazard UHS for PEGASOS 2004, PRP 2013 and SED 2015 in terms of mean, median, 5% and 95% fractiles. The SED rock UHS consists of only 7 spectral acceleration amplitudes, as not the same frequencies between PRP and SED are available. On the other hand, the PRP results on soil are based on 57 frequency points. This leads to a very smooth but detailed curve, where the site specific resonance peaks are clearly visible. It needs to be noted again that the SED UHS has no intermediate points between 20 and 100 Hz and the curve shown here in the graph is simply a straight line connection. The V_{S30} values on soil surface within PRP are not fully identical to the soil surface V_{S30} values defined in PEGASOS, but consistent. As a fair comparison with PEGASOS is only possible at the surface (and not on the rock level), the PEGASOS UHS were omitted in the rock comparisons above.

As can be seen from the comparisons in Fig. 3 the main difference between the old PEGASOS and the new PRP mean hazard is a reduced amplitude in the low frequency range (≤5 Hz). This is mainly due to the spectral shape of the new and modern ground motion predication equations which have been used. The median hazard between the old PEGASOS and the new PRP are consistent and even the uncertainty ranges match quite well over a broad frequency range (see Fig. 4).

For the soil surface hazard comparisons, PRP and SED results are also consistent, but noticeable is that the PEGASOS median results are in general consistent with the PRP and SED results and that the PRP 2013 fractiles fall within the range of the PEGASOS 2004 distribution. The mean hazard of course changed between PEGASOS 2004 and PRP 2013, SED 2015 respectively, as the upper bound fractiles have been reduced.

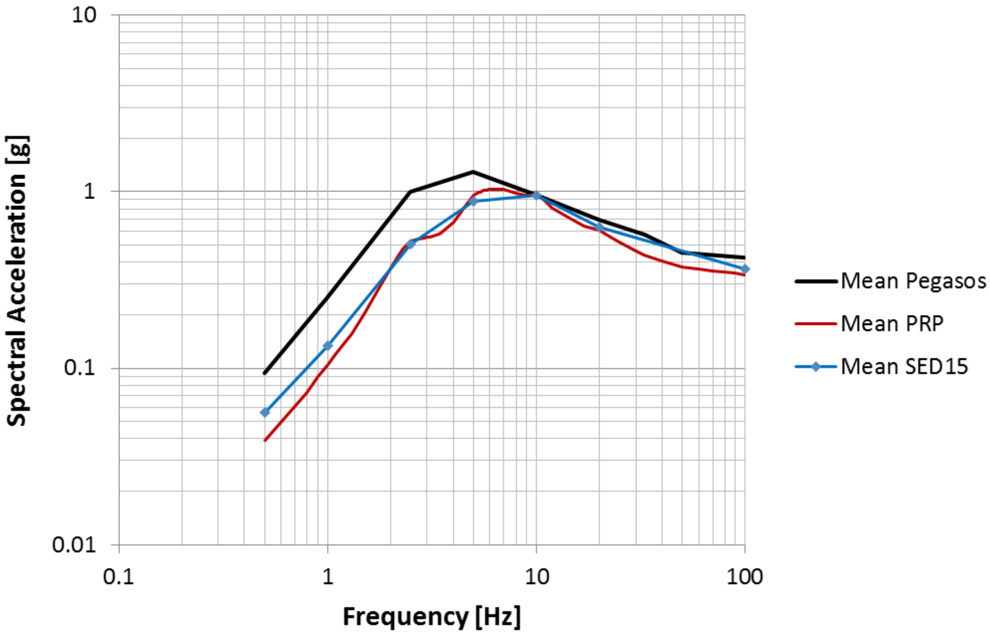

Fig. 3 – Soil surface hazard comparison for a specific site at an annual probability of exceedance of 1E-4, horizontal ground motion, for PEGASOS 2004, PRP 2013 and SED 2015. Mean curves only.

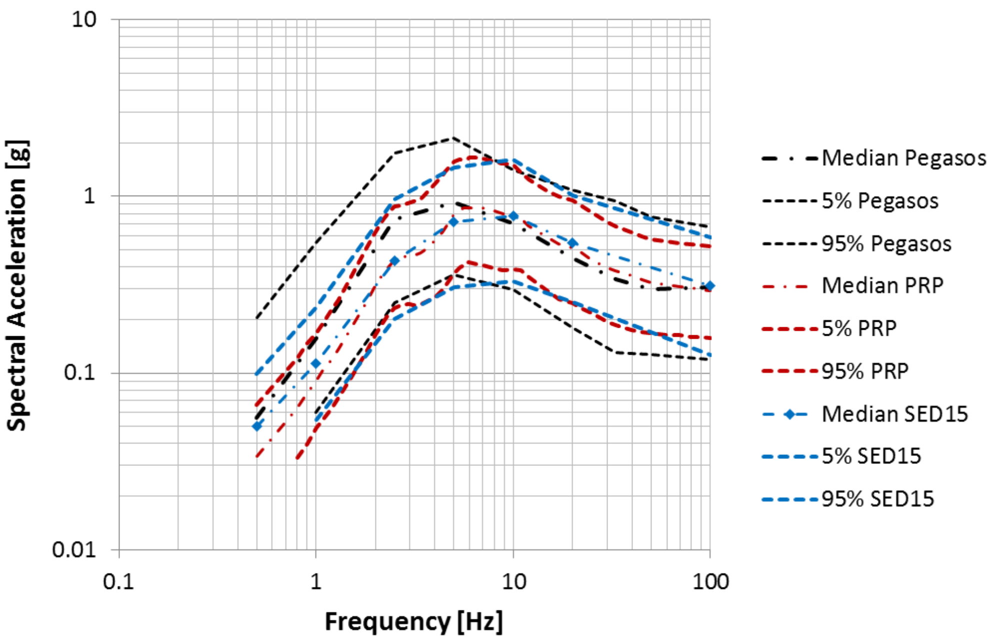

Fig. 4 – Soil surface hazard comparison for a specific site at an annual probability of exceedance of 1E-4, horizontal ground motion, for PEGASOS 2004, PRP 2013 and SED 2015. Median and fractile curves.

5. Discussion of regulatory aspects

It is the declared goal of the SSHAC process to significantly increase the probability of the regulatory acceptance of a PSHA through application of its process. However, it lies within the nature of the issue that the regulator has to evaluate the PSHA and its results independently from the author of the study and he also may come to different conclusions. The independence of the regulator is explicitly endorsed by the SSHAC procedure, as documented in the original report of the Senior Seismic Hazard Analysis Committee [1] and the Practical Implementation Guidelines by NRC [2]:

> "… Regulatory applications: (…) In particular, it is not known to us whether the results of a given PSHA performed using the SSHAC guidance will be useful for nuclear-power-plant regulatory purposes. …" (Chapter 1.9 in [1]).

> "This increased regulatory assurance is the primary benefit obtained by conducting a Level 3 or Level 4 study. However, it is very important to emphasize that adoption of a Level 3 or 4 process does not guarantee regulatory acceptance even if the project fully conforms to the procedural requirements." (Chapter 3.3 in [2]).

The citation of and reference to the SSHAC process does by no means imply that the author advertises the SSHAC guidelines as being the only and best solution for performing PSHA. However, it provides a framework which in theory should lead to a robust hazard estimate, but without any guarantee for acceptability by a regulatory body.

The aim of state-of-the-art hazard assessments is to account for the "center, body and range of the technically defensible interpretations" [2] concerning inputs to the analysis. For the 'center', the best-estimate model (e.g. the ground-motion model that is thought to best represent the median ground motions in the region) or parameter (e.g. the best estimate for the *b* value in the Gutenberg-Richter relation) should be used. The 'body' refers to the shape of the alternative interpretations of the available data (e.g. accounting for the uncertainty in the *b* estimate based on its standard deviation) and the 'range' refers to the tails of the interpretations and limiting credible values (e.g. considering analogs to similar regions). In a broader sense, respecting the center, body and range should lead to robust and stable results, as all credible models and parameters have been considered and injecting new data should not threaten the main results and their associated uncertainty.

"According to SSHAC, if a project sponsor and the analysts choose to do a probabilistic hazard analysis, its procedures will yield stable results" (Summary and Conclusions of the review report by the National Academy of Sciences/National Research Council in [1]).

The main results and associated uncertainty are understood here as the mean, median hazard estimate and the fractiles defining the uncertainty distribution around (e.g. 5, 16, 84, 95% fractiles). After an update of a study with new data the mean, median and fractiles will be modified in a mathematical sense as the input parameters and maybe models have changed. For areas with limited data and knowledge (high uncertainty) the 'body' and 'range' dominate the logic tree approach results, but these are more difficult to capture (and potentially more subjective) whereas areas for which data are abundant the 'center' is the most important. Nevertheless, if the PSHA is robust, the new results should lie within the range of the previous ones, assuming that the uncertainties could be reduced with new knowledge and refined models.

The increased regulatory assurance is the primary benefit obtained by conducting a Level 3 or 4 study:

> "We define regulatory assurance to mean confidence (…) that the data, models, and methods of the larger technical community have been properly considered and that the center, body, and range of technically defensible interpretations have been appropriately represented and documented. In other words, it is increased confidence that the basic objectives of a SSHAC process have been met." (Executive Summary and Chapter 3.3 in [2]).

It is not viable to expect that updates of the PSHA for a same region or site will not modify the existing results. Nevertheless, from the perspective of maintaining longevity and robustness, any change compared to the existing results can be seen as perturbation or challenge. Thus, the key for evaluating whether a result is robust or not should be anchored to robustness as fitness of strategic options or if results from one study to the other significantly differ.

In the framework of the above introduced case study, the regulator developed its own PSHA by combining the seismic source characterization model of SED 2015 with the ground motion and site response characterization of PRP 2013. The resulting hazard can be called a hybrid model (ENSI hybrid model 2015 [16]) and was finally imposed to the NPP operators as the basis for performing the safety and risk verifications of the plants. As example, Fig. 5 shows a comparison of the PRP 2013 and the ENSI hybrid model 2015 for PGA.

As can be seen the mean hazard has the same shape and is only 10% higher than the original PRP at an annual probability of exceedance of 1E-4, for this example. Compared to the overall uncertainty of around 340% it seems very ambiguous to challenge any of the two results. It is more a matter of significant or relevant change in the eyes of the beholder. Although not shown in Fig. 5, it must be mentioned that the uncertainty range of the hybrid result is consistent with the range of the original results. This is of course understood when knowing that the main source of epistemic uncertainty stems from the choice and treatment of the ground motion prediction equations and soil profiles and material parameters for the site response evaluation. Thus, using an alternative seismic source characterization was not expected to significantly change the main results.

For the sake of completeness, it must be mentioned that after the Fukushima-Daiichi accident in 2011, a so called ‘intermediate hazard’ estimate had to be developed by the utilities in order to be used in the EU-Stress Test and Post-Fukushima action plan. As the PRP was at that time only half way completed, a simplified and conservative hazard had to be built up, based on the existing and newly developed information within the PRP. Thus, those results were developed outside the SSHAC framework, but are worth to mention in this context. At that time the regulator accepted this intermediate result as basis for the safety verifications, but of course under the premise that it will be replaced later on by the final hazard results, defined by the regulator. The mean hazard curve for a specific NPP site on soil is shown in Fig. 6 and compared with the other hazard curves at the site. Given the fact that the intermediate hazard 2011 was a crude estimate it is already pretty close to the final PRP hazard for this site delivered in 2013.

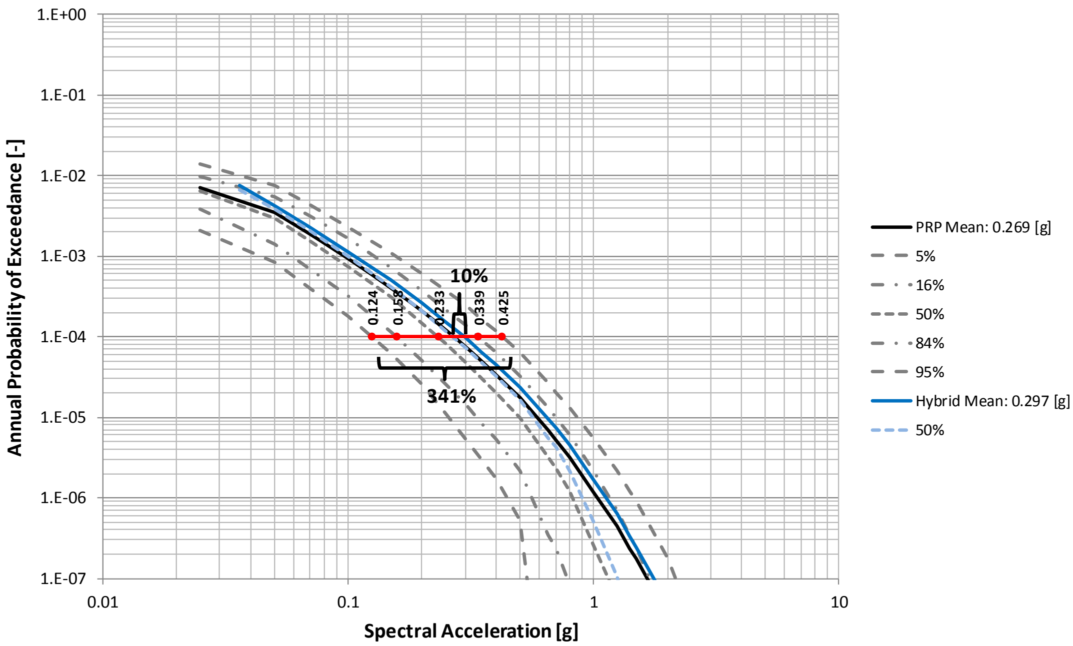

Fig. 5 – Soil sub-surface hazard curve comparison for a specific site at 100 Hz (PGA), horizontal ground motion, for PRP 2013 and ENSI hybrid model 2015. For the ENSI hybrid model only the mean and median hazard is plotted. The values shown in the graph (red intersection points) correspond to the annual probability of exceedance of 1E-4 in order to have a quantitative reference.

Closely related to the issue of hazard significance is the issue of the precision (or imprecision) of hazard estimates (e.g., [17]). If a different group of equally qualified experts were given the same fundamental seismic data for a region (e.g., the same historical earthquake catalog, tectonic information, ground-motion data, site profile information, etc.), that group would derive a slightly different set of inputs and epistemic uncertainties. This would result in a slightly different estimate of mean hazard. Thus, any estimate of hazard has some associated imprecision regardless of how many experts are used in the assessment and how qualified they are. It is important to recognize this imprecision, to attempt to quantify it, and to evaluate the significance of possible future changes caused by new hypotheses or new data. Given this fact, the main guidance should be based on the followed process and its transparency, as well as documentation to be able to rely on credibility.

As a general principle, [18] conjectures the existence of strict 'conservation laws' for robust systems that require that high tolerance to certain assaults be accompanied necessarily by low tolerance to other assaults. For example, an aircraft can either be very light to save fuel and be highly maneuverable or be extremely resistant to anti-aircraft fire, but probably not both. Other consequences might relate to the number and type of options opened up or closed off by the particular form of robustness. The political regime in Renaissance Florence established by Cosimo de'Medici (1389-1464), for example, is analyzed by [19] as deriving its robustness from a strategy of flexible opportunism that permits actions to be "interpreted coherently from multiple perspectives simultaneously", with the consequence of maintaining what Padgett calls "discretionary options across unforeseeable futures in the face of hostile attempts by others to narrow those options."

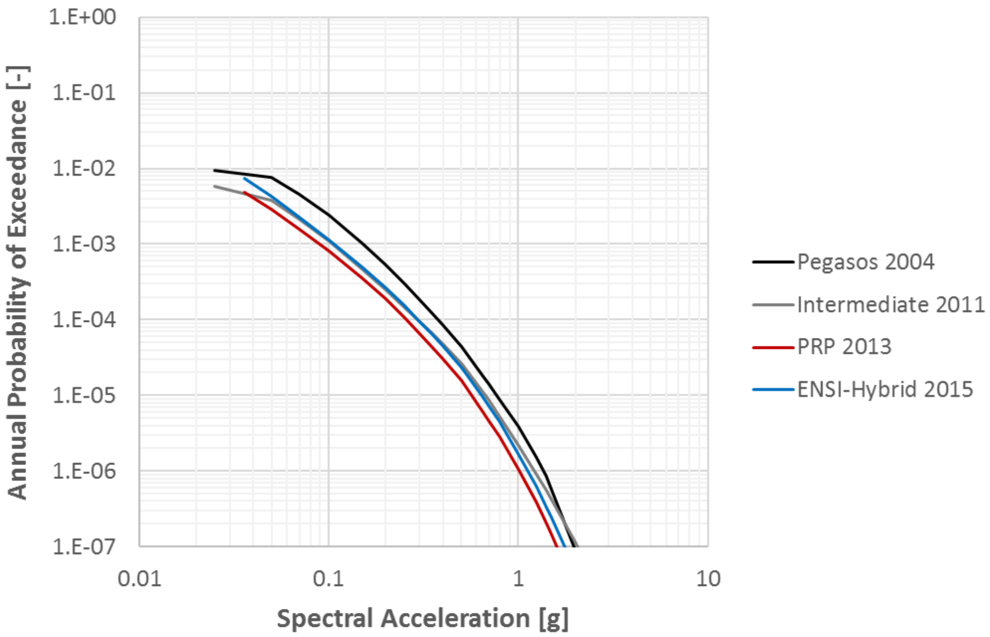

Fig. 6 – Soil sub-surface mean hazard curve comparison for a specific site at 100 Hz (PGA), horizontal ground motion. The four mean curves show in chronological order the hazard curves for PEGASOS 2004, Intermediate hazard 2011, PRP 2013 and ENSI's hybrid 2015.

6. Conclusions and outlook

PSHA means dealing with large uncertainties, at least in regions of low to moderate seismicity – as e.g. in Europe. Scientific progress in earth sciences is achieved on a regular basis and new data collection is ongoing. In the light of the complexity of the problem and lack of data, there is still a need for updates of existing seismic hazard studies. Especially after large earthquakes, there is an increased public awareness and demand which needs to be satisfied by the earth science and engineering community. Furthermore, in the framework of codes and regulations, the question of robustness and stability – or even longevity – of results becomes very important. The latter can play an important role in the context of acceptability and credibility.

Robustness and stability of results in the given context are demanding and very ambiguous. With the help of some examples, the issues in evaluating the robustness and stability of PSHA results were highlighted. Yet, there is no measure for rationally evaluating the robustness of PSHA results, but there are qualitative means when comparing updated results with previous estimates. That results will change over time is obviously implicit to the problem. The question is only: are those changes significant enough to trigger changes in a regulatory framework and/or design basis? This is almost a philosophical and subjective question, but needs to be addressed in order to make progress. For example, one could propose to look at the changes in mean or median hazard and the ratio of the 5 and 95% fractiles (describing the range of uncertainty). A change in the mean is maybe relevant for a regulator, but on the other hand over simplifies the criteria to a single curve or number, respectively. PSHA results are a distribution of probabilistic estimates and thus, should be treated as such and consequently a new input to PSHA should not be overweighed, unless it leads to a significant change in the overall results. A quantification of 'significant' could, for example, mean here larger than 20% change. Below that it is very difficult to argue and almost becomes debating about the thickness of the line which covers that range.

In the case of the Swiss example, the regulator was responsible for the review of the SSHAC Level 4 study. The regulator conducted participatory and late stage reviews and in the end, imposed a hazard estimate that was higher than the original studies results. As sponsors it is practically very challenging to defend such costly and time consuming projects, if in the end there is not a certain degree of assurance of higher regulatory acceptance. Of course, the author is by no means saying that just by performing an expert elicitation process it must be acceptable to the regulator. But on the other hand, new rules and measures must be developed in order to evaluate the robustness and stability of results and not only rely on the process, which is not unique and flawless. The author hopes that the scientific community can support industry and regulatory bodies in this respect to develop such quantitative criteria in the near future.

7. References

[1] US NREC (1997): Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts, *Report NUREG/CR-6372*.

[2] US NRC (2012): Practical Implementation Guidelines for SSHAC Level 3 and 4 Hazard Studies, *Report NUREG-2117*, Rev. 1, April.

[3] Soler L, Trizio E, Nickles T, Wimstatt WC (2012): *Characterizing the Robustness of Science - After the Practice Turn in Philosophy of Science*, Springer, ISBN 978-94-007-2758-8.

[4] Jen E (2005): *Stable or Robust? What's the Difference?*, Robust Design: A Repertoire of Biological, Ecological, and Engineering Case Studies, E. Jen (ed), Oxford University Press, ISBN 13 978-0-19-516532-6.

[5] Doyle J, Carlson J (2005): *Robustness and complexity*, Robust Design: A Repertoire of Biological, Ecological, and Engineering Case Studies, E. Jen (ed), Oxford University Press, ISBN 13 978-0-19-516532-6.

[6] Schneider F (2001): *Robustness in computing systems*, RS-2001-036, In Robustness in Natural, Engineering and Social Systems, Santa Fe Institute, http://discuss.santafe.edu/robustness/stories/StoryReader$36.

[7] Kearns M: *Network security* (in preparation).

[8] Bommer JJ, Strasser FO, Pagani M, Monelli D (2013): Quality Assurance for Logic-Tree Implementation in Probabilistic Seismic-Hazard Analysis for Nuclear Applications: A Practical Example, Seismological Research Letters, **84**(6), 938-945.

[9] Douglas J, Ulrich T. Bertil D, Reymond J. (2014): Comparison of the ranges of uncertainty captured in different seismic hazard studies, *Seismological Research Letters*, **85**(5), 977-985.

[10] OECD/NEA (2015): Workshop on Testing Probabilistic Seismic Hazard Analysis Results and the Benefits of Bayesian Techniques, *Technical Report NEA/CSNI/R(2015)15*.

[11] Kossobokov V, Peresan A, Panza GF (2015) Reality Check: Seismic Hazard Models You Can Trust. *EOS*, **96**(13), 9-11

[12] Wyss M, Nekrasova A, Kossobokov V (2012) Errors in expected human losses due to incorrect seismic hazard estimates. *Natural Hazards*, **62** (3), 927-935.

[13] NAGRA (2004): Probabilistic Seismic Hazards Analysis for Swiss Nuclear Power Plant Sites (PEGASOS Project). Final Report, Vol. 1-6. Prepared for Unterausschuss Kernenergie der Überlandwerke (UAK). http://www.swissnuclear.ch/en/downloads.html.

[14] Swissnuclear (2013): Probabilistic Seismic Hazard Analysis for Swiss Nuclear Power Plant Sites - PEGASOS Refinement Project. Final Technical Report, Vol. 1-6. http://www.swissnuclear.ch/en/downloads.html.

[15] Swiss Seismological Service (2015): Seismic Hazard Switzerland, www.seismo.ethz.ch/en/knowledge/seismic-hazard/for-professionals.

[16] ENSI (2016): ENSI-Hybrid Modell – Hazard Figures, http://www.ensi.ch/de/wp-content/uploads/sites/2/2016/05/ENSI_hybrid_model_hazard_figures_2016.pdf

[17] McGuire R (2009): Issues in Probabilistic Seismic Hazard Analysis for Nuclear Facilities in the US, *Nuclear Engineering and Technology*, **41**(10), 1235-1242.

[18] Carlson JM, Doyle J (1999): Highly Optimized Tolerance: A Mechanism for Power Laws in Designed Systems, *Phys Rev E* 60, 1412.

[19] Padgett J, Ansell C (1993): Robust action and the rise of the Medici, 1400-1434, *American J. Sociology*, **98(**6), 1259-1319.