frontiers

Supplementary Material to Probabilistic Tsunami Hazard and Risk Analysis – A Review on Research Gaps

1 INTRODUCTION

The article *Probabilistic Tsunami Hazard and Risk Analysis – A Review on Research Gaps* covers a large number of currently open research gaps. In 47 sections – about half of them dedicated to hazard analysis, and the second half to risk analysis – diverse aspects of research gaps are described with over 350 literature references as starting points for a deeper study and further information.

This collection was originally intended to be a first milestone for a collaborative project, the COST Action *Accelerating Global Science in Tsunami Hazard and Risk Analysis - AGITHAR* (agi, 2020). While the authors found this unweighted and simple collection already useful and its compilation challenging, the reviewers suggested to add some judgement on the importance or urgency of each of the research gaps.

While a thorough assessment and analysis of the importance of each gap is a research topic in itself and theoretically quite challenging (after all, we identified gaps, in which it is often not clear how a closure would contribute to reduce uncertainty), one feasible methodology to assess the influence of gaps was found to be an expert judgement. The authors are aware that this is an ad hoc assessment and does not replace a formal expert elicitation (Cooke, 1992; Morgan, 2014; Basili et al., 2021, e.g.) or a rigorous sensitivity analysis of each of the gaps. However, it is a good starting point and the results give a first impression of the relative importance of each of the research gaps.

With the number of co-authors exceeding 50 and all of the co-authors being experts in the field, a critical mass of competent participants in this expert judgement exercise was identified. A questionnaire was designed and a poll conducted. With a return rate of 63% (32 out of 51) we think this result is valuable and significant.

Three questions for each gap section of the article were asked to the experts, with a very simple answer scale:

1. How sensitive does the hazard or risk result react to the gap; has this knowledge gap a large or small impact and does it induce a large (additional) uncertainty? (Answer scale 1–3 from low to high);
2. How difficult is it to fill this gap with scientific effort; How much effort is needed to tackle the gap? (Answer scale 1–3 from easy to hard);
3. Is this particular gap more related to a lack of theoretical/conceptual knowledge or a lack of available data (or even both)? (Possible answers: mostly theoretical knowledge gap; both theory and data gap; mostly data gap).

Experts were allowed to leave some of the answers blank, if they felt they were not expert enough to give a qualified answer.

From these questions, we could derive a priority matrix that puts higher priority to those gaps with large influence on the results and low difficulty to be solved scientifically. Of course it turned out that most of

the research gaps are judged difficult to solve but with large influence. This is expected, because otherwise such gaps would have been closed. Nevertheless, there is some significant distinction in the answers.

2 DETAILED DESCRIPTION OF METHODOLOGY

The questionnaire can still be found as a Google Form (Behrens, 2020). As mentioned above, experts were asked three questions for each of the 47 research gaps subsections. While some of the subsections mention several individual aspects of a gap, we found that the ordering in subsections was detailed enough to justify treating each subsection as an individual gap area. Abbreviated titles of the gap sections together with annotation coordinates are given in table S1.

In order to determine the overall sensitivity of the PTHA/PTRA results with respect to the corresponding research gap, we computed an average value for all given answers. This average value must be found between 1 and 3, since these were the margins of the answer scale allowed in the questionnaire. The same procedure was applied to the scientific tractability measure. The coordinate pair (tractability,sensitivity) was then plotted in the matrix. Scientific tractability was chosen to be plotted on the x-axis, with hard to solve problems to the right. Sensitivity was selected to be plotted on the y-axis with higher impact on top.

The assessment of type of gap was treated similarly. An answer *mostly theoretical knowledge gap* was translated into a value 1, whereas *mostly data gap* was given a value 3. Again an average was computed and rounded to the nearest integer for the ability to being back translated into the plain words. So, if half of the participants voted for data gap and the second half for theory gap, it would be marked as *both theory and data gap*. This gives a slight bias towards the judgement that a certain gap is related to both missing data and theory. On the other hand, it appears natural to assume that in most cases there is some theoretical knowledge gap related to missing data (or the other way around). Colors are assigned to each of the three different types of gaps.

In order to assess the reliability of the judgement the spread of answers for the first two questions was computed in terms of the mean normal deviation (σ) of the sample. Since these values can differ for the sensitivity and the scientific tractability, a root mean square value ($s = \sqrt{\sigma(q1)^2 + \sigma(q2)^2}$, $q1$: question 1, $q2$: question 2) is used for visualization. Each point in the matrix is scaled by a multiple of s such that again three categories are visible.

3 PRIORITIZATION

The expert judgement and the priority matrix in figure S1 allow for a prioritization. However, a number of metrics for prioritization are possible. We just describe how we decided to prioritize.

It appears natural to rank those gaps that are of high impact (high sensitivity) and relatively easy to close scientifically. These are the gaps that fall into the upper left quadrant of our matrix. However, a number of research gaps in the upper right quadrant may also be of high priority, since their impact is unanimously accepted and their tractability is still moderate.

Research gaps that are of low impact but hard to solve (lower right quadrant) should probably be ranked further down in the priority of research activities.

It should be noted that other metrics are also possible and should be considered. For example a high priority could be assigned to those research gaps that directly influence the well-being of individuals or societies or that could immediately reduce number of human losses. Other selection of priorities are also possible, when considering different types of applications. An insurance company may for example prioritize those research gaps that reduce uncertainty in determining economic losses.

4 SUPPLEMENTARY TABLES AND FIGURES

For convenience, we plot the figure from the article here again (S1), so that the reference to the data underlying the visualization can be made more conveniently.

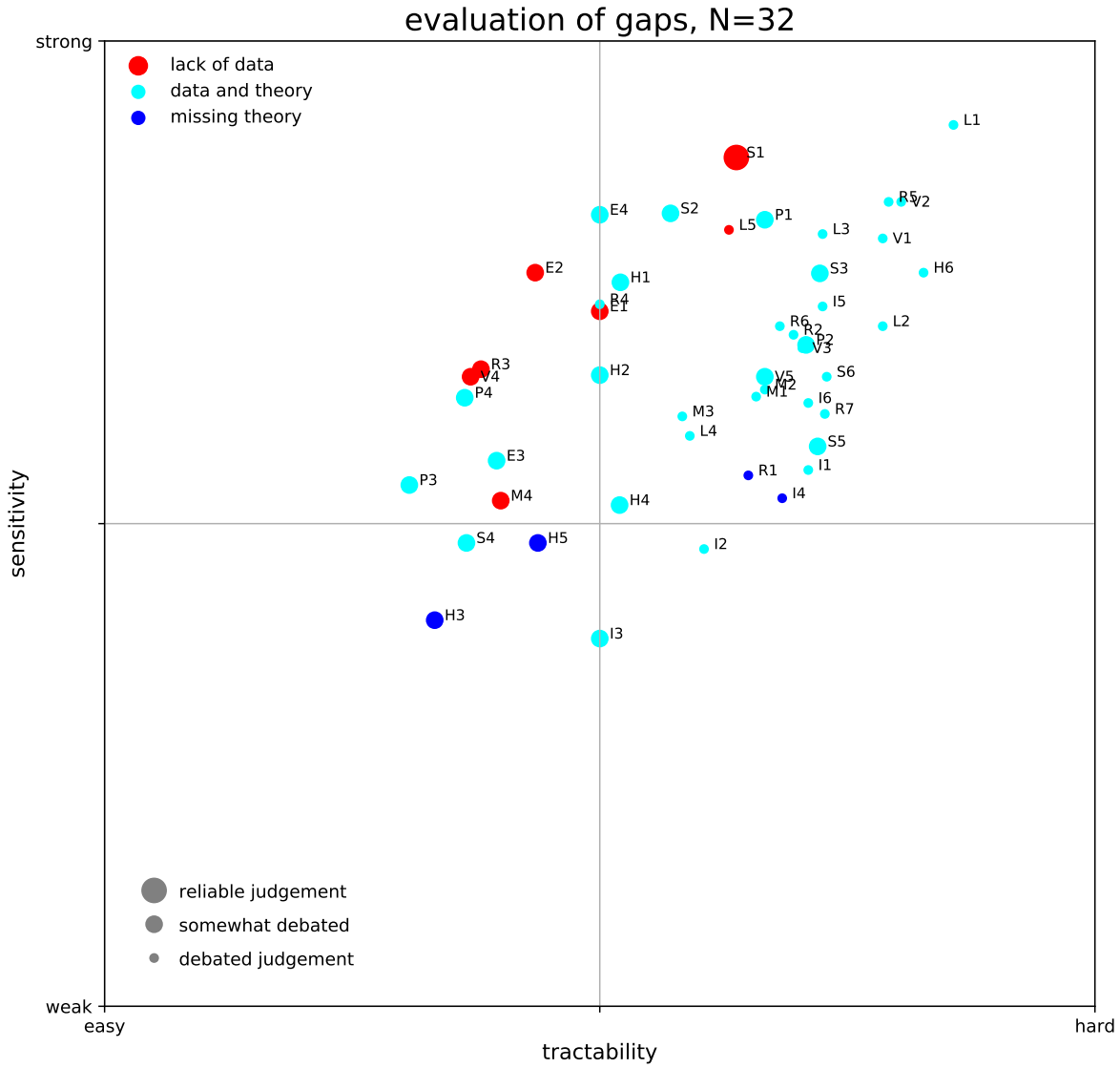

Figure S1. Priority matrix for research gaps. For the short lables of each gap, see table S1. Size of markers relates to the spread of answers (reliability). Color of markers relates to the type of research gap. N is the number of samples (participants in questionnaire).

We list the data used in our priority matrix in table S1. Please note that we also list the values of $\sigma(q1)$ and $\sigma(q2)$ as well as the total number of answers for each of the gaps and questions. We omit the same

analysis for the type of research gap (theoretical, data, or both), since we regard this additional information of minor importance for prioritization.

Table S1. Research gaps, sensitivities, tractability and corresponding spread, number of answers and relation to data or theory

gap (short)	sens. mean	sens. σ	count	trac. mean	trac. σ	count	theor./data
Limited past events (S1)	2.759	0.901	29	2.276	0.864	29	data gap
Fault identification (S2)	2.643	1.014	28	2.143	0.857	28	data+theory
Variety, complexity (S3)	2.519	1.053	27	2.444	0.998	27	data+theory
Empirical scaling (S4)	1.960	1.089	25	1.731	0.861	26	data+theory
Complex, non-stationary (S5)	2.160	1.073	25	2.440	1.182	25	data+theory
Other Constraints (S6)	2.304	1.162	23	2.458	1.202	24	data+theory
Lack of Understanding (L1)	2.826	1.311	23	2.714	1.340	21	data+theory
Difference of Onshore (L2)	2.409	1.215	22	2.571	1.285	21	data+theory
Limited Constraints (L3)	2.600	1.340	20	2.450	1.274	20	data+theory
Limited Availability (L4)	2.182	1.199	22	2.182	1.173	22	data+theory
Limited Past Events (L5)	2.609	1.244	23	2.261	1.166	23	data gap
Variety of Potential (V1)	2.591	1.293	22	2.571	1.310	21	data+theory
Difficulties in Constraining (V2)	2.667	1.225	24	2.609	1.244	23	data+theory
Gaps in Modelling (V3)	2.364	1.218	22	2.409	1.215	22	data+theory
Lack of Data (V4)	2.304	1.162	23	1.739	0.901	23	data gap
Limited Availability (V5)	2.304	1.135	23	2.333	1.173	24	data+theory
Lack of Understanding (M1)	2.263	1.240	19	2.316	1.244	19	data+theory
High Sensitivity (M2)	2.278	1.205	18	2.333	1.261	18	data+theory
Limited Availability (M3)	2.222	1.225	18	2.167	1.218	18	data+theory
Limited Past Events (M4)	2.048	1.107	21	1.800	0.992	20	data gap
PTHA Uncertainty (H1)	2.500	1.159	26	2.042	1.030	24	data+theory
Tsunami Generation (H2)	2.308	1.139	26	2.000	1.029	25	data+theory
Uncertainty and Variability (H3)	1.800	0.964	25	1.667	0.866	24	theory gap
Nonlinearity and (H4)	2.038	1.019	26	2.040	1.057	25	data+theory
Quantifying the Influence (H5)	1.960	1.000	25	1.875	1.057	24	theory gap
Modelling Situations (H6)	2.520	1.212	25	2.654	1.149	26	data+theory
Lack of Detail (E1)	2.440	1.208	25	2.000	1.088	23	data gap
Lack of Exposure Data (E2)	2.520	1.212	25	1.870	1.078	23	data gap
Lack of Tsunami (E3)	2.130	1.118	23	1.792	0.922	24	data+theory
Spatio-Temporal (E4)	2.640	1.171	25	2.000	1.023	26	data+theory
Limitation in Asset (P1)	2.630	1.138	27	2.333	1.015	27	data+theory
Effect of Multiple (P2)	2.370	1.031	27	2.417	1.210	24	data+theory
Lack of Consensus (P3)	2.080	1.053	25	1.615	0.845	26	data+theory
Gaps in Building (P4)	2.261	1.166	23	1.727	0.950	22	data+theory
Gaps Related (R1)	2.100	1.158	20	2.300	1.223	20	theory gap
Challenges in Characterizing (R2)	2.391	1.205	23	2.391	1.205	23	data+theory
Lack of a Tsunami (R3)	2.320	1.102	25	1.760	0.960	25	data gap
General Lack of Risk (R4)	2.455	1.261	22	2.000	1.111	22	data+theory
Assessing Tsunami (R5)	2.667	1.250	24	2.583	1.223	24	data+theory
Lack of Understanding (R6)	2.409	1.240	22	2.364	1.218	22	data+theory
The Weakness of (R7)	2.227	1.199	22	2.455	1.261	22	data+theory
The Difficulty of (I1)	2.111	1.184	18	2.421	1.298	19	data+theory
Ambiguities in Definition (I2)	1.947	1.149	19	2.211	1.236	19	data+theory
Lack of Tsunami (I3)	1.762	1.093	21	2.000	1.088	23	data+theory
Integrated Approaches (I4)	2.053	1.192	19	2.368	1.296	19	theory gap
Considering Community (I5)	2.450	1.299	20	2.450	1.274	20	data+theory
Incorporating Risk (I6)	2.250	1.221	20	2.421	1.273	19	data+theory

REFERENCES

[Dataset] (2020). Accelerating global science in tsunami hazard and risk analysis. https://agithar.uni-hamburg.de

Basili, R., Brizuela, B., Herrero, A., Iqbal, S., Lorito, S., Maesano, F., et al. (2021). The making of the neam tsunami hazard model 2018 (neamthm18). *Front. Earth Sci.* doi:10.3389/feart.2020.616594

[Dataset] Behrens, J. (2020). Questionnaire for ptha/ptra paper. https://forms.gle/hx11w4HpzQKGYvg4A

Cooke, R. M. (1992). *Experts in Uncertainty, Opinion and Subjective Probability in Science* (Oxford University Press)

Morgan, M. G. (2014). Use (and abuse) of expert elicitation in support of decision making for public policy. *Proceedings of the National Academy of Sciences* 111, 7176–7184. doi:10.1073/pnas.1319946111