Probabilistic TSUnami Hazard MAPS
for the NEAM Region (TSUMAPS-NEAM)
ECHO/SUB/2015/718568/PREV26

PHASE I – STAGE 2: Pool of Expert Kick-off

Author: TSUMAPS-NEAM Technical Integrator (TI) team

Date: 2 June 2017

Version: 1.0

Table of Contents

This page is intentionally left blank.

Executive Summary

In this document we report the initialization of the activity of the Pool of Experts (PoE) in the TSUMAPS-NEAM project.

In ***Section 1***, we shortly describe the PoE and its goal within the project, summarizing the information reported in the document ***Doc_P1_S1_Project_Summary***.

In ***Section 2***, we describe the PoE kick-off meeting, held in Athens on 30 June 2016. The goal of this meeting was to introduce the project to the members of the PoE, and to initialize the experts' weights to be adopted in all the elicitations of the project.

In ***Section 3***, we discuss the evaluation of the experts' weights, as computed from their answers to a seed questionnaire. We evaluated group's answers by adopting the different experts' weights. This analysis has been presented in its preliminary version at the end of the Athens meeting, and it is reported in more details in Section 3.3.

To complete the information of the present document, we attach the following Appendixes:

- ***Appendix A***: Minutes of the Athens meeting;
- ***Appendix B***: Presentations of the PoE kick-off meeting;
- ***Appendix C***: Questionnaire used for the elicitation experiment.

This page is intentionally left blank.

1. Pool of Experts in the TSUMAPS-NEAM project

Fifteen experts were invited to join the Pool of Experts (PoE) of the TSUMAPS-NEAM project. Following the EU project STREST protocol (Selva et al., *in prep*), the PoE assists the Technical Integrator (TI) in taking the critical decisions required for implementing the Seismic Probabilistic Tsunami Hazard Analysis (S-PTHA) for the NEAM region.

The members of the PoE have been selected from inside and outside the project, based on their competences and known field of expertise. The experts from within the project which are members of the PoE are not participating in any activity related to the coordination of the multiple-expert process, that is, they are not part of the Technical Integrator (TI) and Project Manager (PM) teams. The fifteen experts of the PoE are listed in **Table 1**.

Table 1: List of experts of the PoE along with their affiliations.

Expert	Affiliation
F. Romano	INGV
R. Omira	IPMA
F. Lovholt	NGI
A. Babeyko	GFZ
A. Yalciner	METU
G. Papadopoulos	NOA
M. Canals	UB
A. Bouallegue	INM
A. Armigliato	UNIBO
M. Sorensen	UIB
C. Ozer	KOERI
G. Davies	GA
W. Power	GNS
J. Polet	CALTECH
C. Meletti	INGV

The members of the PoE participate in the expert elicitation procedure that the Technical Integrator (TI) organizes during the TSUMAPS-NEAM project. Two elicitations are scheduled, the first one did already take place:

- **Elicitation 1 (conducted during Phase 1, pre-assessment)** Prioritization of the levels at which the analysis of epistemic uncertainty has to be deepened during the pre-assessment phase.
- **Elicitation 2 (conducted during Phase 2, assessment):** Quantification of the weights to be assigned to alternative models for the ensemble model of the SPTHA uncertainty during the assessment phase.

Further elicitation experiments may be organized if required by the TI and the PM teams. More details can be found in ***Doc_P1_S1_Project_Summary.***

2. PoE Kick-off Meeting

The general scheme of the hazard assessment in the TSUMAPS-NEAM project was presented and discussed during the first day of the technical meeting held in Athens (Greece) from 29 June to 1 July, 2016. The members of the PoE were invited to participate to the whole meeting, during which they had the opportunity to start familiarizing with the TSUMAPS-NEAM project approach, goals, and partners.

A half-day PoE kick-off session was then organized in the morning of the second day of the meeting. In this kick-off session, the TI presented the role of the PoE and the whole process of interaction with the TI and the PM teams. This enabled an effective discussion to clarify the scope of the tasks of experts in the elicitation and their specific role.

After the discussion, the fifteen members of the PoE were invited to answer a seed questionnaire in order to assign weights to the experts themselves. In particular, the questionnaire consisted of two parts. In the first part each expert was asked to answer a set of seed questions prepared by the TI. In the second part each expert was asked to acknowledge two other experts of her/his own choice within the PoE. These activities are propaedeutic to the application of two different weighting schemes which are discussed in **Section 3**.

The minutes of the Athens meeting are reported in ***Appendix A***. The PoE participated in the whole meeting (Appendix A, Sections 1.3 and 1.4), but the PoE kick-off took place on 30 June: Morning session (Appendix A, section 2.3). In this session, we specifically discussed i) the overall multiple-expert process followed within the TSUMAPS-NEAM project, the different actors foreseen, and the specific role of the PoE; ii) the elicitation process and the potential techniques to be applied; iii) the goals of the weighting experts and the weighting schemes to be adopted; iv) the seed questionnaire and its use to quantify experts' weights. Three presentations were given, each followed by questions and discussion. These presentations are reported in ***Appendix B***. Note that these presentations were prepared on the basis of an earlier version of the EU project STREST protocol (Selva et al. 2015), whereas the protocol shortly discussed in ***Doc_P1_S1_Project_Summary*** is based on a review of that experimental process in light of the TSUMAPS-NEAM application (Selva et al., *in prep*). After this preliminary part, the experts answered to the seed questionnaire reported in ***Appendix C***.

To complete the explanations regarding the use of experts' weights and their importance to the PoE, the preliminary version of the results reported in **Section 3** were presented at the end of the Athens meeting (see **Section 2.5.3**. Additional presentations of ***Appendix A***).

3. Experts' weights

The aggregation of experts' judgements is often supported by the use of different weighting schemes for the experts themselves. Many different schemes were proposed and adopted in previous analyses, such as equal weighting, self-weighting, peer weighting, performance weighting, acknowledgement weighting, and many others (e.g., Selva et al. 2012; Aspinall and Cooke, 2013). The rationale for adopting experts' weights in the aggregation phase is to favor rational consensus, based on reproducibility, accountability, empirical control, neutrality, and fairness (Cooke 1991; Cooke & Goossens 2000).

In TSUMAPS-NEAM, the idea is to compare the impact on the elicitation results of three alternative schemes that depend on the expertise and acknowledgement of the experts within the tsunami science community. These comparisons also allow for evaluating the impact of the selection of the PoE members, and thus enable a more aware decision making.

3.1. Alternative weighting schemes

Three alternative weighting schemes are taken into account in the elicitation process and their respective results are compared one to another, in order to check their consistency, and all together will provide the input for final decision-making.

In order to assign weights to the PoE members, we consider the following alternative schemes:

- Equal Weighting (EW) scheme,
- Acknowledge-based Weighting (AW) scheme,
- Performance-based Weighting (PW) scheme.

Equal weighting scheme: this weighting scheme is straightforward. Every expert gets the same weight which is obtained by W=1/N, where N is the number of involved experts. In our case, fifteen experts were involved in the PoE, so each expert was assigned a weight of 1/15.

Acknowledgement-based weighting scheme: a weight is assigned to each expert on the basis of mutual recognition among the experts themselves, expressed through a blind procedure (Selva et al. 2012). To quantify such weights, each expert is asked to vote for other members of the PoE.

During the PoE kick-off meeting, the following activity was performed. The list with the names of the 15 experts was distributed. The experts were requested to assign a weight of 1 (one) to themselves and a weight of 1 and 3 to other two colleagues. The weight of 1 and 3 indicates the different levels of confidence that one expert put on his/her colleagues. The weight of each expert is then evaluated by 1) summing the weights that each expert received, and 2) renormalizing to 1 the sum of the weights of all experts.

The resulting weights are presented in anonymized form in **Figure 1**. Note that to preserve the anonymity of experts, **Table 1** and **Figure 1** are intentionally sorted in different ways. The graph in **Figure 1** shows that the experts E6, E8, and E13 received higher acknowledgment from their colleagues, thus obtaining the highest weights. The experts E1, E3, and E4 received the lowest weight.

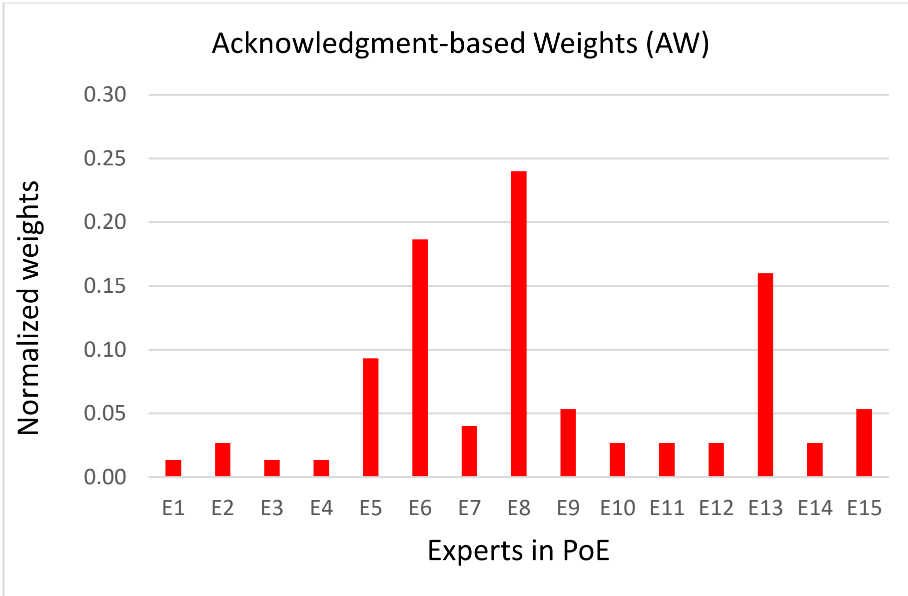

Figure 1: The weights assigned to experts according to the Acknowledgement-based Weighting (AW) scheme. Note that the experts' numbers are assigned by randomly shuffling the list in **Table 1**.

Performance-based weighting scheme: This scheme is based on a classical method developed by Cooke (1991), in which the weights on experts' opinion are assigned through experts' relative performance in answering a set of seed questions.

During the PoE kick-off meeting, a questionnaire related to tsunami hazard was given to the experts and they were asked to express their best guess (numerical value) and related confidence intervals (5th, 50th, and 95th percentiles) to each question. Their assessments were used to obtain weights using Excalibur, a software package for structured expert judgement elicitation using Cooke's (1991) model.

The questionnaire was prepared observing a balance between questions belonging to various aspects:

- 2 main categories: Earthquake (E) and Tsunami (T) science;
- 3 sub-categories: Phenomenology (Phe, observations from past events), Probability (Prb, hazard analysis), Modelling (Mod, physical or numerical);
- 2 typologies for Phe and Prb: Local (L, target area) or Global (G, any area in the world).

The questionnaire comprised 14 questions, in order to have enough (> 5) seed questions in each category, as reported in **Table 2**.

Table 2: Balance of expertise in preparing the seed questions.

Q	DESCRIPTION	E	T	Phe	Prb	Mod
1	Displacement of one historical earthquake in the Atlantic	x		L		
2	Length of fault of Chile 1960 earthquakes	x		G		
3	M > 6.0 earthquakes in Europe (from ANSS)	x			L	

4	Change in Mc of tapered Pareto, worldwide	x			G	
5	Vertical displacement, Java earthquake	x				x
6	Tsunami wave-height, Greece		x	L		
7	Tsunami inundation height, Japan		x	G		
8	Messina Strait Area, percentage of seismic tsunamis in the last 500 yr		x		L	
9	NOAA/WDS worldwide, n of tsunamis > 10 m in 1970-2005		x		G	
10	Run-up along inundation with MOST		x			x
11	Amplification Synolakis		x			x
12	Mmax close to Gibraltar	x	x		L	
13	Return period Cascadia	x	x		G	
14	ARP for 2% in 50 yr	x	x		G	

The resulting normalized weights are presented in anonymized form in **Figure 2**. Note that the expert ordering in **Figure 1** and **Figure 2** is the same, but it does not correspond to the ordering in **Table 1**. Most of the weight on experts' opinions is assigned to four experts (E1, E4, E9, and E13). Among those, E9 received the highest weight. This shows that the relative performance of four experts in answering the questions was significantly better than that of all other members of the PoE.

Note that the weights assigned to the experts using the PW scheme are here estimated based on all fourteen questions, but they can also be estimated category-wise and sub-category-wise. This was used to check the stability of the results.

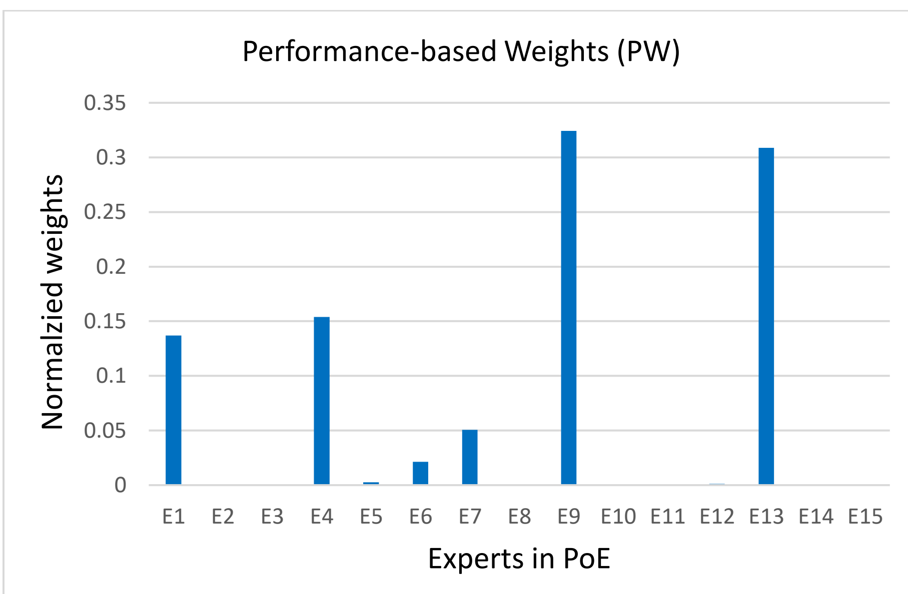

Figure 2: The weights assigned to experts according to the Performance-based Weighting (PW) scheme. Note that the experts' numbers are assigned by randomly shuffling the list in **Table 1** (but same ordering as in Figure 1).

3.2 Comparison of expert weights under alternative weighting schemes

Experts were assigned different weights based on alternative weighting schemes. The comparison for each expert is shown in **Figure 3**.

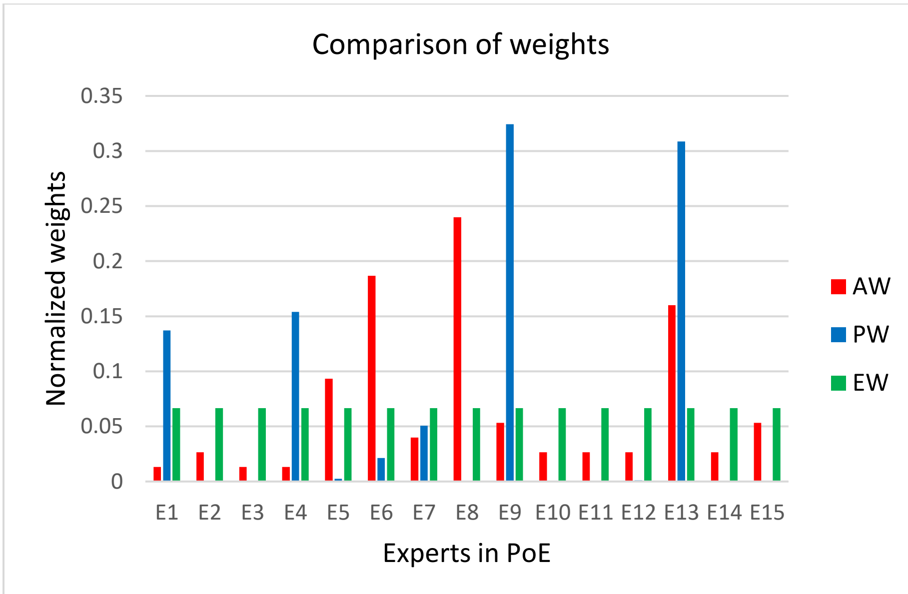

Figure 3: Comparison of weights assigned to experts based on alternative weighting schemes.

The comparison shows that the variability under the different schemes is rather large. The resulting weights are indeed often inconsistent, with some extreme case where highly acknowledged experts have very small performance weights (e.g., expert E8) and vice versa (e.g., expert E4). One exception is expert E13 who performed very well (second best) in answering the questions and was also significantly acknowledged by other colleagues.

3.3. Performance of the PoE on seed questions

Here we present how weights would have performed, if applied in aggregating the answers of experts on the 14 seed questions used for the PW scheme. We note that, since the PW are obtained by maximizing the performance on exactly these 14 questions, the results are automatically biased in favor of this weighting scheme. Conversely, EW and AW are computed independently from these 14 questions. For this reason, the effective performance of the three weighting schemes cannot be judged by these results only.

To each of the 14 seed questions, the experts provided their assessment in terms of best guesses and associated confidence intervals. What would be the best estimate and uncertainty ranges for each question that takes into account weighted assessments from all experts? This solution is presented in **Tables 3**, **4,** and **5,** based on each weighting scheme along with the real value of the answer. In these tables, the columns with headers 5%, 50%, and 95% represent, as a suggested solution, the minimum, the best, and the maximum value, respectively. The column named "realization" provides all the real values, i.e. the expected correct answer. The same results are reported in **Figure 4** normalized to 1 by dividing each value (5th, 50th and 95th percentiles of the group result) by the realization (correct answer) to the same question. The results show a good agreement between best guess and true answers, with uncertainty bounds very variable but in general within the 10% relative variation bounds. This demonstrate that the group collectively perform well over the seed questions.

Table 3: Results for Equal Weighting (EW) scheme.

Solution for : EW

Resulting solution (combined DM distribution of values assessed by experts)
Bayesian Updates: no Weights: equal DM Optimisation: no
Significance Level: 0.0000 Calibration Power: 1.0000

Nr.	Id	Scale	5%	50%	95%	Realization	Full Name
1	Q1	UNI	0.7519	9.928	34.01	11	Displacement of one historical earthquake in the Atlantic (Earthquake)
2	Q2	UNI	295.4	797.4	2353	920	Length of fault of Chile 1960 earthquakes (Earthquake)
3	Q3	UNI	8.64	69.55	295	75	M > 6.0 earthquakes in Europe (from ANSS) (Earthquake)
4	Q4	UNI	7.32	8.418	19.5	8.76	Change in Mc of tapered Pareto, worldwide (Earthquake)
5	Q5	UNI	0.05969	0.2662	0.7032	0.43	Vertical displacement, Java earthquake (Earthquake)
6	Q6	UNI	2.95	10.87	31.97	20	Tsunami wave-height, Greece (Tsunami)
7	Q7	UNI	4.117	16.14	45.4	19.5	Tsunami inundation height, Japan (Tsunami)
8	Q8	UNI	2.755	12.28	18.91	13	Messina Strait Area, percentage of seismic tsunamis in the last 500 yr (tsunami)
9	Q9	UNI	1.418	9.162	60.82	15	NOAA/WDS worldwide, n of tsunamis > 10 m in 1970-2005 (Tsunami)
10	Q10	UNI	1.055	2.689	14.4	3.01	Run-up along inundation with MOST (Tsunami)
11	Q11	UNI	0.2968	3.015	59.83	4	Amplification Synolakis (Tsunami)
12	Q12	UNI	5.083	5.968	8.223	6.4	Mmax close to Gibiltar (Earthquake)
13	Q13	UNI	218.9	672.3	1.331E004	512	Return period Cascadia (Earthquake)
14	Q14	UNI	28.12	2438	2708	2475	ARP for 2% in 50 yr (Earthquake)

Table 4: Results for Acknowledgement-based Weighting (AW) scheme.

Solution for : DMaker1

Resulting solution (combined DM distribution of values assessed by experts)
Bayesian Updates: no Weights: user DM Optimisation: no
Significance Level: 0.0000 Calibration Power: 1.0000

Nr.	Id	Scale	5%	50%	95%	Realization	Full Name
1	Q1	UNI	0.6548	8.66	29.55	11	Displacement of one historical earthquake in the Atlantic (Earthquake)
2	Q2	UNI	303	754.2	1954	920	Length of fault of Chile 1960 earthquakes (Earthquake)
3	Q3	UNI	18.31	75.61	240.3	75	M > 6.0 earthquakes in Europe (from ANSS) (Earthquake)
4	Q4	UNI	7.486	8.418	17.89	8.76	Change in Mc of tapered Pareto, worldwide (Earthquake)
5	Q5	UNI	0.03961	0.2257	0.6551	0.43	Vertical displacement, Java earthquake (Earthquake)
6	Q6	UNI	2.43	9.735	28.84	20	Tsunami wave-height, Greece (Tsunami)
7	Q7	UNI	2.65	14.73	46.17	19.5	Tsunami inundation height, Japan (Tsunami)
8	Q8	UNI	3.973	13.56	18.94	13	Messina Strait Area, percentage of seismic tsunamis in the last 500 yr (tsunami)
9	Q9	UNI	2.318	11.75	49.39	15	NOAA/WDS worldwide, n of tsunamis > 10 m in 1970-2005 (Tsunami)
10	Q10	UNI	1.166	2.639	11.02	3.01	Run-up along inundation with MOST (Tsunami)
11	Q11	UNI	0.2702	3.198	41.1	4	Amplification Synolakis (Tsunami)
12	Q12	UNI	5.059	5.797	7.626	6.4	Mmax close to Gibiltar (Earthquake)
13	Q13	UNI	222.7	660	3899	512	Return period Cascadia (Earthquake)
14	Q14	UNI	36.96	2468	2687	2475	ARP for 2% in 50 yr (Earthquake)

Table 5: Results for Performance-based Weighting (PW) scheme.

Solution for : PW

Resulting solution (combined DM distribution of values assessed by experts)
Bayesian Updates: no   Weights: global   DM Optimisation: no
Significance Level: 0.0000   Calibration Power: 1.0000

Nr.	Id	Scale	5%	50%	95%	Realization	Full Name
1	Q1	UNI	1.224	11.96	35.35	11	Displacement of one historical earthquake in the Atlantic (Earthquake)
2	Q2	UNI	221.4	827	3003	920	Length of fault of Chile 1960 earthquakes (Earthquake)
3	Q3	UNI	22.69	72.49	170.9	75	M > 6.0 earthquakes in Europe (from ANSS) (Earthquake)
4	Q4	UNI	5.912	9.094	27.31	8.76	Change in Mc of tapered Pareto, worldwide (Earthquake)
5	Q5	UNI	0.02973	0.2515	0.7012	0.43	Vertical displacement, Java earthquake (Earthquake)
6	Q6	UNI	3.025	19.65	35.81	20	Tsunami wave-height, Greece (Tsunami)
7	Q7	UNI	5.521	20.43	40.82	19.5	Tsunami inundation height, Japan (Tsunami)
8	Q8	UNI	7.989	15.28	19.23	13	Messina Strait Area, percentage of seismic tsunamis in the last 500 yr (tsunami)
9	Q9	UNI	2.093	11.42	36.35	15	NOAA/WDS worldwide, n of tsunamis > 10 m in 1970-2005 (Tsunami)
10	Q10	UNI	0.5977	2.748	12.71	3.01	Run-up along inundation with MOST (Tsunami)
11	Q11	UNI	1.414	3.814	14.65	4	Amplification Synolakis (Tsunami)
12	Q12	UNI	5.083	5.985	8.237	6.4	Mmax close to Gibiltar (Earthquake)
13	Q13	UNI	203	687	1.993E004	512	Return period Cascadia (Earthquake)
14	Q14	UNI	26.13	2445	2578	2475	ARP for 2% in 50 yr (Earthquake)

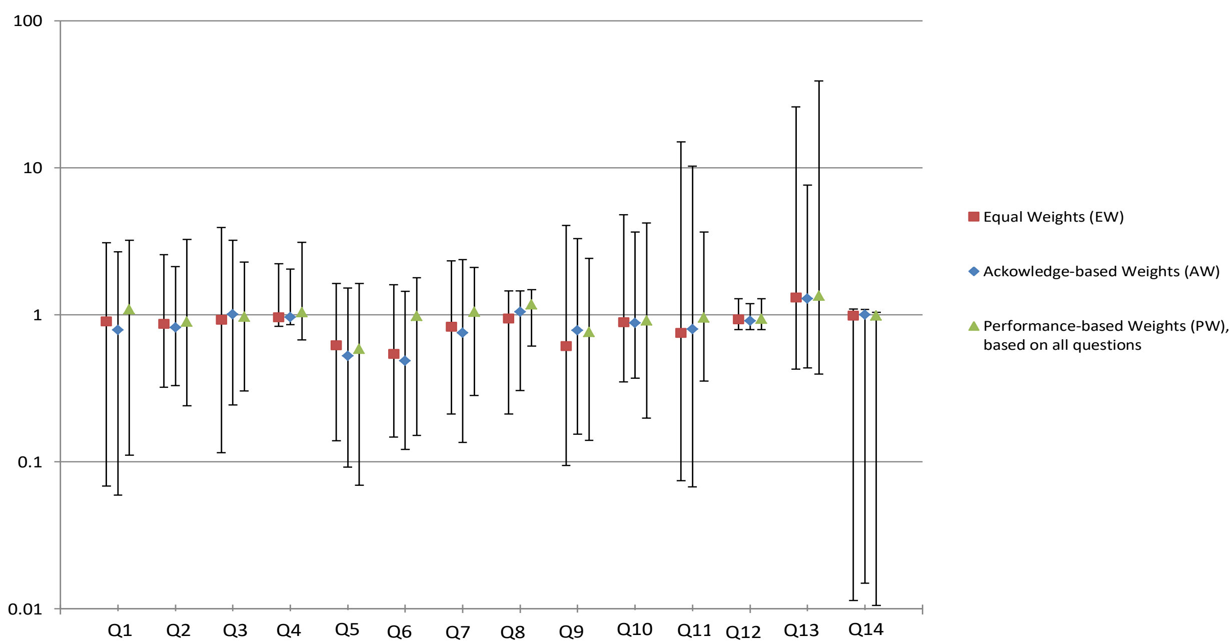

Figure 4: Comparison of normalized solutions for the three weighting schemes. The position of the symbols represent the median, and the bars the interval between the 5th and 95th percentiles.

References

Aspinall, W.P., Cooke, R.M., 2013. Quantifying scientific uncertainty from expert judgement elicitation, in "Risk and Uncertainty Assessment for Natural Hazards" (Eds J. Rougier, L. Hill, R.S.J. Sparks), Cambridge University Press, Cambridge, UK, ISBN 978-1-107-00619-5.

Cooke, R. M. (1991). Experts in Uncertainty: Opinion and Subjective Probability in Science," Oxford University Press, Oxford.

Cooke, R.M. and Goossens, L.H.J. (2000), "A Procedures Guide for Structured Expert Judgment," EUR 18820, European Commission Report.

Selva, J., Marzocchi, W., Papale, P., & Sandri, L. (2012). Operational eruption forecasting at high-risk volcanoes: the case of Campi Flegrei, Naples. Journal of Applied Volcanology, 1(1), 1.

Selva, J. et al. (2015). Report on the effects of epistemic uncertainties on the definition of LP-HC events. Deliverable 3.1 STREST Project.

Selva J., S. Iqbal, F. Cotton, D. Giardini, S. Esposito, B. Stojadinovic, S. Argyroudis, K. Pitilakis, A. Mignan, and S. Lorito (in prep). Management of subjectivity in probabilistic single/multi-hazard/risk assessments: a Multiple-Expert Management Protocol (MEM-Pr).