COMPDYN 2019
7th ECCOMAS Thematic Conference on
Computational Methods in Structural Dynamics and Earthquake Engineering
M. Papadrakakis, M. Fragiadakis (eds.)
Crete, Greece, 24–26 June 2019

IDENTIFYING UNCERTAINTY CONTRIBUTIONS TO THE SEISMIC FRAGILITY ASSESSMENT OF A NUCLEAR REACTOR STEAM LINE

Pierre Gehl[1], Marine Marcilhac-Fradin[2], Jeremy Rohmer[1], Yves Guigeno[2], Nadia Rahni[2], Julien Clément[2]

[1] BRGM
Orléans, France
p.gehl@brgm.fr, j.rohmer@brgm.fr

[2] IRSN
Fontenay-aux-Roses, France
marine.marcilhacfradin@irsn.fr, yves.guigueno@irsn.fr, nadia.rahni@irsn.fr, julien.clement@irsn.fr

Abstract

In nuclear applications, fragility curves are an essential element of the seismic probabilistic safety assessment that is performed at the level of the power plant. They are required to account for the aleatory randomness and the epistemic uncertainty generated by various sources of variability, such as the representation of the seismic input by intensity measures, the assumptions in the structural model (e.g., mechanical or geometrical parameters) and the confidence in the statistical estimation of the fragility parameters (i.e., related to number of data points used). Therefore, this study investigates the relative contributions of such variables to the dispersion of the resulting fragility functions, while ensuring the separation between aleatory and epistemic uncertainty sources, as advocated by the standards in effect in the nuclear industry. To this end, vector-valued fragility functions, based on two intensity measures, are also investigated: it appears that they allow for a partial transfer from the record-to-record variability to an epistemic uncertainty component that is related to the description of the seismic loading given the hazard at the studied site.
The proposed uncertainty decomposition is applied to the fragility assessment of the main steam line of a nuclear reactor: the total dispersion of the resulting fragility models is then decomposed into different aleatory and epistemic components. Although it is found that vector-valued intensity measures contribute to a significant part of the total dispersion, the uncertainty due to the variability of mechanical and geometrical parameters appears to be even larger.

Keywords: Fragility curves, intensity measures, dynamic analysis, epistemic uncertainty.

1 INTRODUCTION

The quantification of the vulnerability of structures and equipment constitute a crucial step of the seismic Probability Safety Assessment (PSA) of a Nuclear Power Plant (NPP). To this end, fragility curves are common tools developed in the nuclear industry. The vulnerability of a component may then be represented by the so-called High Confidence Probability of Failure (HCLPF) capacity, which corresponds to the value of the intensity measure (IM) leading to a failure probability of 5% on the 95% confidence interval of the fragility curve (EPRI, 2003). Therefore a rigorous distinction between aleatory and epistemic uncertainty sources is required, following for instance the decomposition proposed by Kennedy et al. (1980): the total dispersion β of the fragility curve is decomposed into a term β_R representing aleatory randomness (i.e., the "slope" of the curve) and a term β_U representing epistemic uncertainty (i.e., the width of the confidence interval).

$$P_f(im) = \Phi\left(\frac{\ln im - \ln \alpha + \beta_U \Phi^{-1}(Q)}{\beta_R}\right) \quad (1)$$

where P_f is the conditional probability of reaching or exceeding a given damage state, Φ is the normal cumulative distribution function, α is the median of the fragility function, and Q is the confidence level

Due to the complexity of a ground-motion time history, the description of the seismic loading through a single IM is acknowledged as a significant source of uncertainty, usually referred to as the record-to-record variability. Several studies have addressed this issue, either by searching for adequate IMs (Luco & Cornell, 2007; Padgett et al., 2008) or by deriving vector-IM fragility functions or surfaces (Baker & Cornell, 2005; Seyedi et al., 2010; Modica & Stafford, 2014). Generally, a decrease in the dispersion of the fragility functions is observed, however its interpretation in terms of uncertainty transfer remains to be clarified. Therefore, the aim of this study is to quantify the various uncertainty sources that may contribute the total dispersion of the derived fragility models. In particular, the comparison between single-IM fragility curves and vector-IM fragility functions should offer insight on the treatment of record-to-record variability and on its impact on the aleatory uncertainty component.

The proposed analysis will be demonstrated on the fragility assessment of the main steam line of a pressurizer water reactor (Rahni et al., 2017), using a set of nonlinear time-history analyses (Section 2). Then, several criteria for the selection of IMs will be investigated (Section 3), before deriving vector-IM fragility functions with multivariate regression models (Section 4). Finally, in Section 5, three main types of uncertainty will be compared thanks to different fragility formulations: *(i)* the aleatory uncertainty due to the record-to-record variability, *(ii)* the epistemic uncertainty due to the number of simulations (data points) used, and *(iii)* the epistemic uncertainty due to the variability of mechanical and geometrical parameters. As a result, this work will allow to decompose the total dispersion into several components, thus identifying the terms that deserve the most attention when computing the HCPLF.

2 NUMERICAL MODEL AND ANALYSES

This section details the modelling assumptions for the main steam line of the reactor, as well as the design of experiment for the dynamic analyses.

2.1 Modelling assumptions

The coupled model of a supporting structure and a secondary system is considered here, corresponding to the main steam line of a pressurised water reactor. The model, built and

computed with the CAST3M finite-element software (Combescure et al., 1982), is taken from Rahni et al. (2017). Structural elements, representing the containment building, are modelled with multi-degree-of-freedom stick formulations. The containment building has a double-wall structure, with an inner reinforced prestressed concrete wall and an outer reinforced concrete wall (see Figure 1). The steel steam line is modelled by means of beam elements, representing pipe segments and elbows, as well as several valves, supporting devices and stops at different elevations of the supporting structure. The stick models, which have the benefit of enabling fast computations, have been calibrated from detailed finite-element 3D models of the containment building (Rahni et al., 2017).

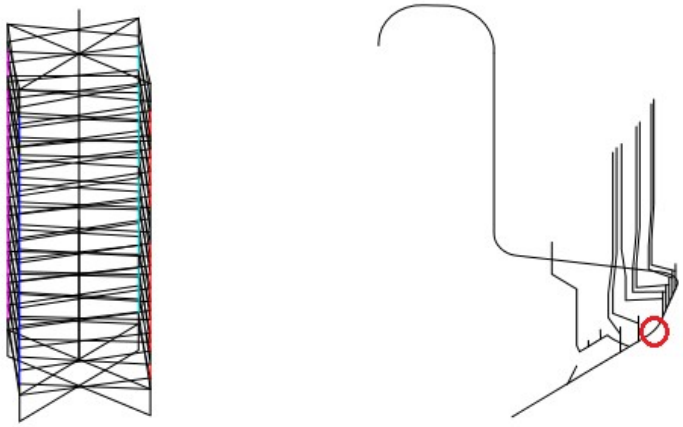

Figure 1:Left: stick model of the containment building (inner wall, outer wall and internal structures); Right: steam line beam model, where the red circle points to the location of the vertical stop.

Dynamic analyses are performed in two successive steps: first, the seismic loading is applied at the base of the building; and the resulting time history of structural displacements is then applied to the steam line model in order to estimate the induced strains and stresses on the beam elements. Rahni et al. (2017) have proposed several mechanical failure criteria for the verification of the steam pipe line integrity, such as the equivalent stress at any point of the pipe line (conservative assumption), the total plastic deformation at the pipe location corresponding to the containment penetration (i.e., accounting for the non-linear behaviour of the steel material), and the effort calculated at the beam model's node corresponding to a vertical stop (see Figure 1). For simplicity purposes, the latter failure criterion is considered here for the fragility analysis. The threshold for the occurrence of the damage state considered (i.e., failure at the vertical stop) is arbitrarily set at EDP_{th} = 400 kN for the maximum transient effort: this choice is made in order to obtain a relatively good balance between data points corresponding to intact or damaged states, for demonstration purposes.

As a preliminary to the non-linear time-history analyses, a modal analysis of the supporting structure is carried out, in order to identify the main vibration modes (see Table 1).

Mode #	Period [s]	3D mass participation factor [%]
1	0.38	61
2	0.38	65
3	0.15	35
4	0.14	97

5	0.14	31

Table 1: Modal analysis of the model of the containment building.

Modes #1 and #2 correspond to the excitation of the structure in the *Y*- and *X*-direction, respectively: due the almost symmetrical properties of the building, they are almost identical. Therefore, the fundamental period of the structure is taken as $T_1 = 0.38$s, while the second one is taken as $T_2 = 0.15$s (i.e., cluster of modes #3, #4 and #5).

In order to integrate the epistemic uncertainties due to the identification of some mechanical and geometrical parameters, ten variables are sampled within a Latin Hypercube Sampling design (see Table 2), following the values provided by Rahni et al. (2017).

Variable	Description	Uniform distribution interval
E_{IC}	Young's Modulus – Inner containment	[27700 – 45556] MPa
ξ_{RPC}	Damping ratio – reinforced prestressed concrete	[4 – 6] %
ξ_{RC}	Damping ratio – reinforced concrete	[6 – 8] %
e_1	Pipe thickness – Segment #1	[29.8 – 38.3] mm
e_2	Pipe thickness – Segment #2	[33.3 – 42.8] mm
e_3	Pipe thickness – Segment #3	[34.1 – 43.9] mm
e_4	Pipe thickness – Segment #4	[33.3 – 42.8] mm
e_5	Pipe thickness – Segment #5	[53.4 – 68.6] mm
e_6	Pipe thickness – Segment #6	[34.1 – 43.9] mm
ξ_{SL}	Damping ratio – steam line	[1 – 4] %

Table 2: Range of variation of the ten uncertain parameters considered, based on Rahni et al. (2017). A uniform distribution is assumed.

2.2 Selection of input ground-motion records

The conditional spectrum method (Lin et al., 2013) is used here for the selection of the input ground motions, for subsequent dynamic analyses. This approach has the benefit of enabling a light scaling of a set of natural records, while saving the consistency of the associated response spectra. Therefore, it is especially suited for the use of spectral values, such as *SA* (spectral acceleration) at various periods. The main steps of this procedure are the following:

- *Probabilistic hazard assessment of the studied site*: here, an arbitrary location in Southern Europe is selected. Hazard curves are generated with the OpenQuake platform (www.globalquakemodel.org), accounting for 13 seismogenic areas which have been characterized in the SHARE project (Woessner et al., 2013).

- *Selection of a period of interest and of scaling levels*: here, the response spectrum is chosen to be conditioned on $SA(T_1=0.38\text{s})$, with 6 scaling levels ranging from 0.185g to 3.882g, corresponding to return periods from 20 years to 20 000 years.

- *Identification of reference earthquakes*: for the studied site, the OpenQuake software performs a hazard disaggregation for each scaling level in order to identify a reference earthquake scenario.

• *Generation of the target response spectrum* and *selection of spectrum-compatible ground-motion records* (Jayaram et al., 2011). The final selection from the PEER database (PEER, 2013) consists of 30 records for each of the 6 scaling levels (i.e., 180 ground-motion records in total), as shown in Figure 2.

The ground-motion selection using conditional spectrum implies the evaluation of the seismic hazard at a given site, along with the identification of reference earthquakes at various return periods: as a result, this approach leads to site-specific fragility functions, which are well suited to the context of NPPs.

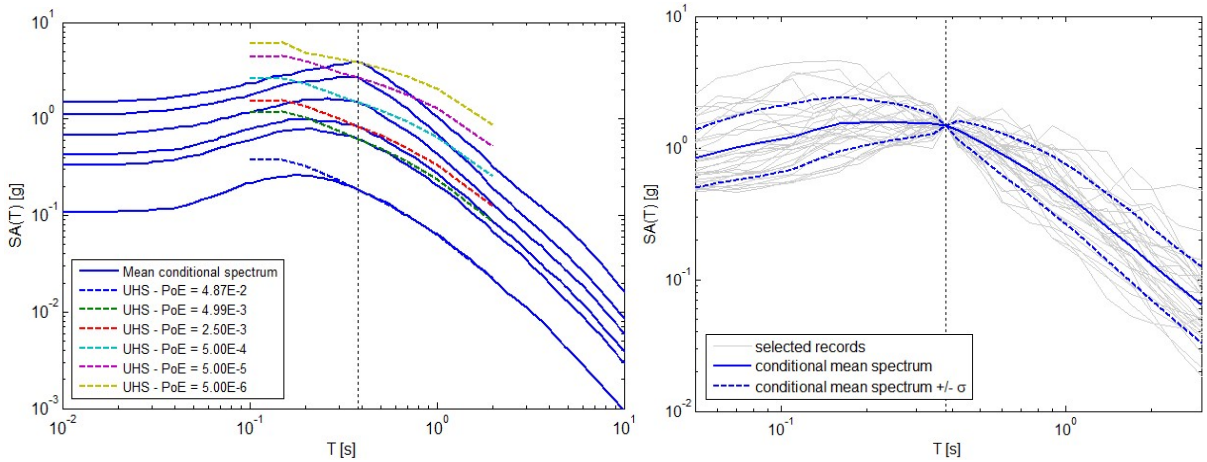

Figure 2: Left: conditional mean spectra and uniform hazard spectra for the 6 scaling levels; Right: conditional spectrum for scaling level #4 and corresponding set of 30 selected ground-motion records.

2.3 Non-linear time-history analyses

The 180 ground-motion records are applied to the base of the 3D model of the containment building, in the form of a 3-component loading. In total, 360 models of the PWR structure are built in the CAST3M environment, so that each ground-motion record may be applied to two different models, with the objective of generating enough data points. As an example, some simulation outcomes are presented in Figure 3, with PGA and SA(0.5s).

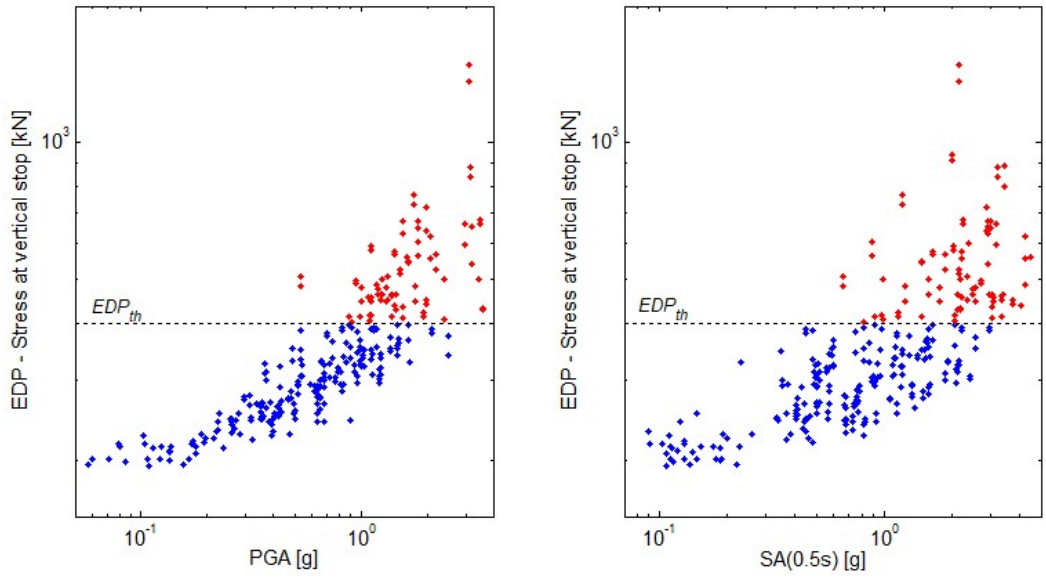

Figure 3: IM-EDP data points, with respect to PGA (left) and SA(0.5s) (right).

3 SELECTION OF SCALAR IMS

From Figure 3, it may be observed that a linear fit between the logarithms of IM and EDP is not justified for this specific case study: it is proposed here to apply the maximum estimation approach (MLE – cf. Shinozuka et al., 2000) approach, which only requires a separation between the intact and damaged states (i.e., respectively blue and red points in Figure 3). To this end, a large number of ground-motion parameters is evaluated as potential IMs:

- Spectral acceleration at various periods: *SA*(*T*)
- Peak parameters (acceleration, velocity and displacement): *PGA*, *PGV*, *PGD*
- Arias Intensity parameters: *AI* (Arias Intensity), *A95*[1], *SL75-95*[2]
- Spectral intensities: *SI* (Housner Intensity), *ASI* (Acceleration Spectral Intensity)
- Duration parameters: *RSD75-95*[3]
- Cyclic parameters: *NCy* (number of effective cycles), *DCy* (cyclic damage parameter)
- Energy-related parameters: *NED* (Normalised Energy Density)
- Parameters related to time-integrated acceleration: *CAV* (Cumulative Absolute Velocity), *ARMS* (Root-Mean-Square Acceleration)

The adequacy of IMs for the derivation of fragility curves has been addressed by many studies, using efficiency and sufficiency indicators (Luco & Cornell, 2007), statistical classifiers (Lancieri et al., 2015) or the concept of hazard compatibility (Hariri-Ardebili & Saouma, 2016). Based on these previous works, it is proposed to use three performance metrics in order to estimate the adequacy of the considered IMs:

1. **Standard-deviation β of the fragility curve:** although the MLE approach is used here, the estimated dispersion parameter β may be interpreted as the quantity described by Padgett et al. (2008) as the *proficiency* measure (i.e. combination of efficiency and practicability measures).
2. **Akaike Information Criterion (AIC):** this criterion quantitatively assesses the goodness-of-fit of a given model. The AIC accounts for the number of parameters used in a model through the variable k, in order to penalize the over-parametrization of some models. In the case of scalar-IM fragility curves, $k = 2$ (i.e., parameters α and β). The AIC is then expressed as follows:

$$AIC = 2k - 2\ln L \tag{2}$$

   where L is the likelihood function of the fragility model, which is computed as a product of the 360 conditional probabilities corresponding to the 360 simulation outputs. Therefore, a small AIC value implies a great statistical fit given the data.
3. **Area under the ROC curve (AUC):** the ROC curve is a possible representation of a ROC analysis, where the ability of a given model to be both specific and sensitive is evaluated by plotting the true positive rate versus the false positive rate. This approach has been applied by Gehl et al. (2013) to the evaluation of fragility curves (i.e. ability of the model to accurately predict the damage state or not, given an IM taken as a predictor). Therefore, the AUC provides a quantification of how well the fragility model

[1] Level of acceleration that contains 95% of the Arias Intensity

[2] Slope of the Husid plot (cumulative *AI* over time) between 5% and 75% (or 95%) of the total *AI*

[3] Relative Significant Duration: length of time interval between when *AI* first exceeds 5% of total value and when *AI* first exceeds 75% (or 95%) of total value

works as a classifier: a large AUC value (i.e., area close to 1) implies a model that works significantly better than a random classifier (i.e., the 1:1 diagonal).

These three criteria are first estimated for *SA* at various periods, ranging from 0.05s to 2s (see Figure 4). The curves reveal an optimum at T = 0.29s, whatever the metric considered. Local optima are also found at periods equal to 0.14s and 0.50s. The three identified periods are close to the periods that corresponding to the first two vibration modes (i.e., T_1 = 0.38s and T_2 = 0.15s). However, they are not exactly identical, and these differences may be explained by two factors, i.e. *(i)* the combination of superior modes that may be excited by some ground motions and *(ii)* the lengthening of the fundamental period due to the loss of elasticity of the structural components.

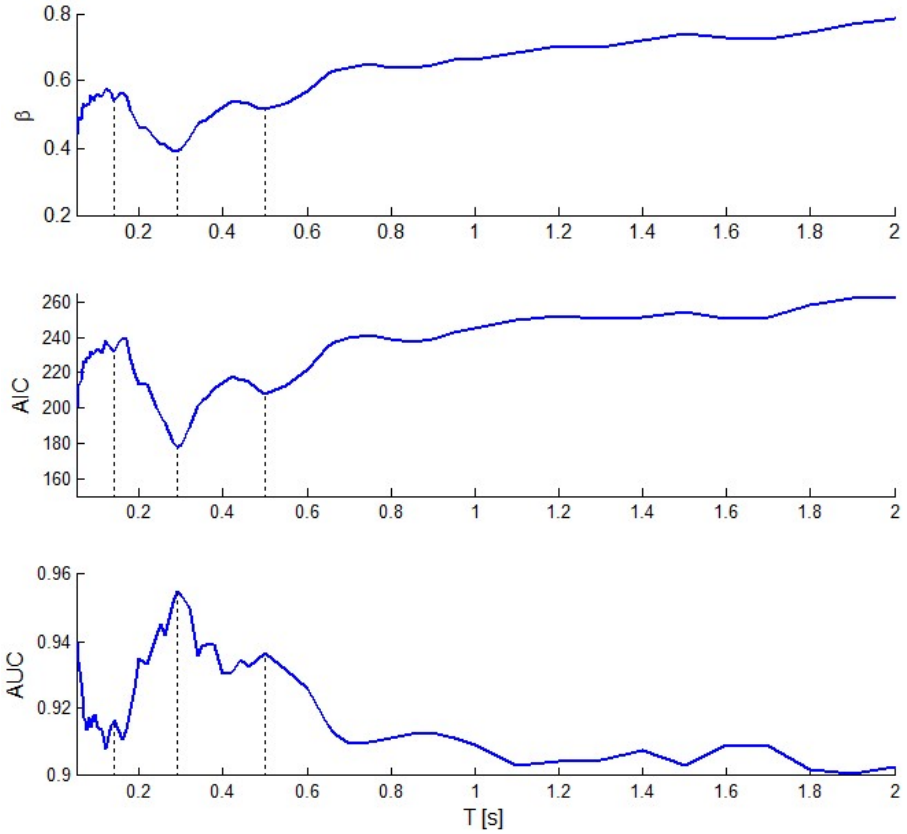

Figure 4: Evolution of the three performance metrics considered, with SA at different periods.

As a result, *SA* at the three identified periods (0.14s, 0.29s and 0.50s) are considered as potential IMs; and their performance is compared to other ground-motion parameters (see Table 3).

IM	β	AIC	AUC
SA(0.14s)	0.5415	229.91	0.9166
SA(0.29s)	0.3898	175.82	0.9547
SA(0.50s)	0.5144	206.12	0.9363
PGA	0.4403	198.45	0.9399
PGV	0.4928	205.73	0.9381
PGD	1.2622	308.93	0.8469

AI	0.7674	182.13	0.9485
A95	0.4041	192.88	0.9389
SL75	0.9471	206.91	0.9325
SL95	0.7681	176.59	0.9508
SI	0.5293	213.94	0.9347
ASI	0.3775	176.23	0.9519
RSD75	-	-	-
RSD95	-	-	-
NCy	-	-	-
DCy	0.8123	191.58	0.9412
NED	1.4284	259.27	0.8968
CAV	0.6012	247.15	0.8951
ARMS	0.5728	240.45	0.9073

Table 3: Estimated values of the three performance metrics, for different potential IMs. The gray cells indicate the five best performing IMs, for each metric.

It is found that the parameters *SA*(0.29s) and *ASI* are the most consistent, since they show a satisfying performance across the three metrics. Other well performing parameters are *PGA*, *PGV*, *AI*, *A95*, *SL95* and *DCy*. However, it should be noted that *DCy*, *SL95* and *A95* may not be easily computed from current GMPEs (see **Erreur ! Source du renvoi introuvable.**). The metrics cannot be evaluated for some parameters (*RSD75*, *RSD95*, *NCy*) because the fragility estimation has not converged due the poor IM-EDP correlation: these parameters may still be used as secondary IMs when deriving fragility surfaces, if the right IM combination is found.

4 DERIVATION OF VECTOR-IM FRAGILITY FUNCTIONS

In order to improve the predictive power of the fragility curves and to reduce the dispersion due to the record-to-record variability, it is proposed to combine two IMs and to use this vector-valued predictor for the derivation of the fragility functions. To this end, the following functional form for the damage probability is assumed:

$$\begin{aligned} P_f(im_1, im_2) &= P(ds \geq DS | IM_1 = im_1, IM_2 = im_2) \\ &= \frac{1}{2}[1 + erf(c_1 + c_2 \ln im_1 + c_3 \ln im_2)] \end{aligned} \quad (3)$$

where *erf* is the error function and c_1, c_2 and c_3 are the coefficients to be estimated (i.e., fragility parameters).

A composite variable im_V may then be introduced as follows:

$$\mathrm{im_V} = \mathrm{im_1}^{\frac{c_2}{c_2+c_3}} \cdot \mathrm{im_2}^{\frac{c_3}{c_2+c_3}} \quad (4)$$

Using im_V as the IM, the functional form in Eq. 3 is expressed as:

$$
\begin{aligned}
P_f(im_V) &= P(ds \geq DS|IM_V = im_V) = P(ds \geq DS|IM_1 = im_1, IM_2 = im_2) \\
&= \frac{1}{2}[1 + \mathrm{erf}(c_1 + c_2 \ln im_1 + c_3 \ln im_2)] \\
&= \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\frac{c_1}{c_2+c_3} + \ln im_1^{\frac{c_2}{c_2+c_3}} + \ln im_2^{\frac{c_3}{c_2+c_3}}}{\frac{1}{c_2+c_3}}\right)\right] \\
&= \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\ln im_V + \frac{c_1}{c_2+c_3}}{\frac{1}{c_2+c_3}}\right)\right] \\
&= \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\ln im_V - \ln \alpha_V}{\beta_V \sqrt{2}}\right)\right]
\end{aligned}
\tag{5}
$$

where α_V and β_V are the "fragility parameters" of the composite IM im_V, which are finally identified as follows:

$$
\begin{cases}
\alpha_V = exp\left(-\frac{c_1}{c_2+c_3}\right) \\
\beta_V = \frac{1}{(c_2+c_3)\sqrt{2}}
\end{cases}
\tag{6}
$$

The coefficients c_1, c_2 and c_3 are estimated with a MLE approach, using the same likelihood function as for scalar-IM fragility curves (expect that there are now three parameters to find, instead of two).

Thanks to the identification of the "composite" dispersion parameter β_V, it is possible to compute the same three of types of performance metrics, as for the case of scalar-IM fragility curves. More than sixty combinations of vector-valued IMs are tested, and the results for the most promising couples of IMs are detailed in Table 4.

IM_1	IM_2	β_V	AIC	AUC
SA(0.14s)	*SA*(0.29s)	0.3724	173.71	0.9568
SA(0.14s)	*SI*	0.3464	167.19	0.9508
SA(0.29s)	*SA*(0.50s)	0.3834	173.74	0.9571
SA(0.29s)	*PGA*	0.3370	161.33	0.9424
SA(0.29s)	*PGV*	0.3718	171.22	0.9591
SA(0.29s)	*AI*	0.4687	171.69	0.9530
SA(0.29s)	*SI*	0.3659	167.43	0.9610
SA(0.29s)	*RSD95*	0.4371	174.11	0.9559
SA(0.50s)	*PGA*	0.3348	158.25	0.9439
PGA	*PGV*	0.3389	166.18	0.9532
PGA	*AI*	0.5027	170.79	0.9468
PGA	*SI*	0.3225	155.43	0.9339
PGA	*ASI*	0.3447	169.06	0.9416
PGV	*ASI*	0.3668	173.61	0.9526

Table 4: Estimation values of the three performance metrics, for different vector-valued IMs. The gray cells indicate the three best performing couples of IMs, for each metric.

It is found that the vector-valued IMs tend to perform slightly better than the scalar IMs, judging from the values of the three metrics. Some scalar IMs that were not identified as ade-

quate (e.g., *RSD95*) become much more efficient when combined together. Some examples of vector-valued fragility functions are displayed in Figure 5, for selected combinations of IMs.

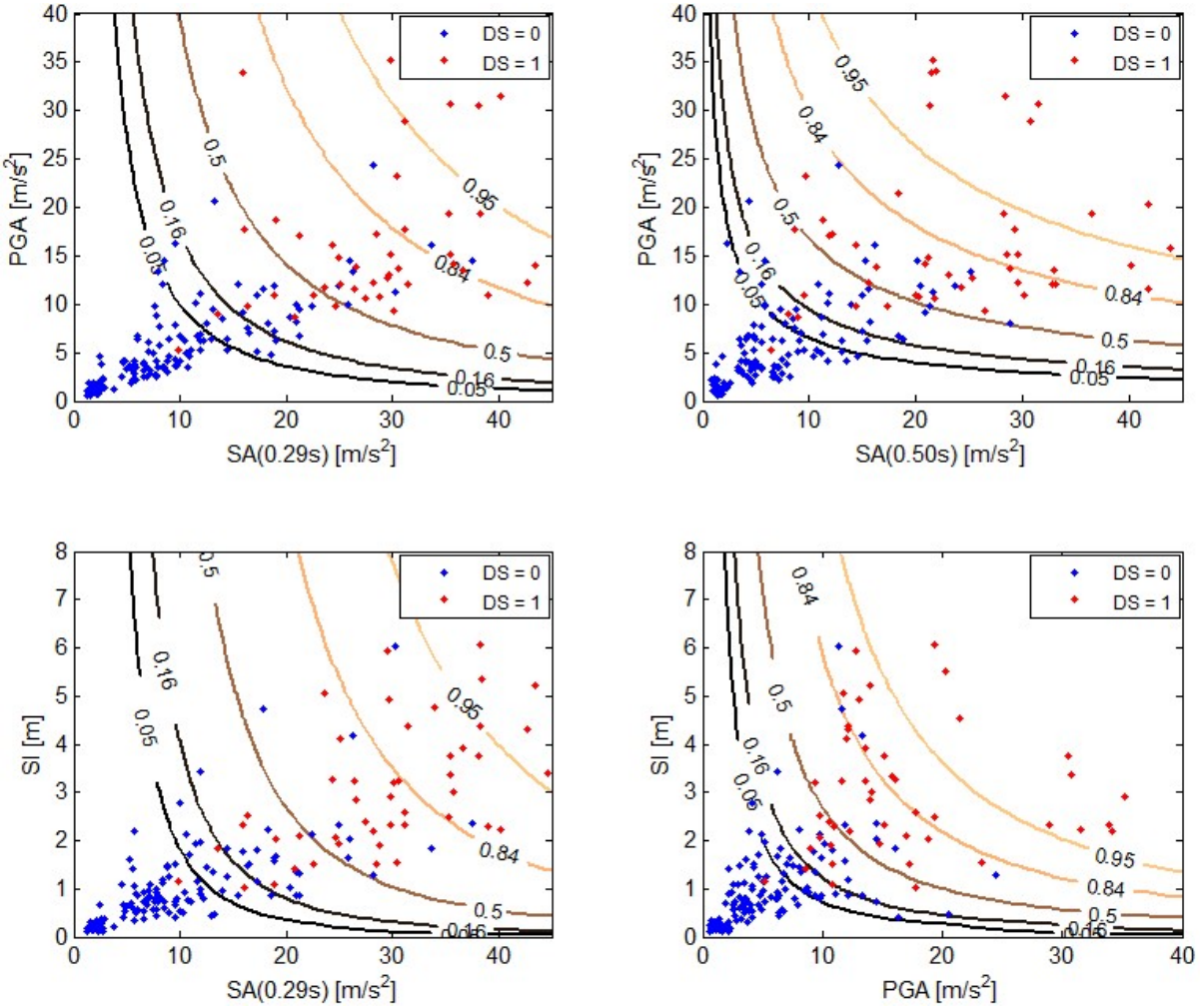

Figure 5: Iso-probability lines corresponding to some examples of vector-valued fragility function (i.e., probabilities of 0.05, 0.16, 0.5 0.84 and 0.95). The dots represent the outcomes of the NLTHAs in the vector-IM space.

5 COMPARATIVE ANALYSIS OF UNCERTAINTY SOURCES

This section discusses a set of procedures for the quantitative estimation of some uncertain components, namely the record-to-record variability, the uncertainty due to the number of data points, and the in-situ variability due to the variability of mechanical and geometrical parameters

5.1 Decomposition of the record-to-record variability

The contribution of the record-to-record variability to the global uncertainty structure may be estimated thanks to the comparison between scalar-IM fragility curves and vector-IM fragility surfaces. To this end, as an example, it is proposed to reduce the fragility surface w.r.t. [*PGA* ; *SA*(0.29s)] (see top left plot in Figure 5) into a fragility curve w.r.t. *SA*(0.29s) only. This operation should consider the correlation between the two IMs, in order to preserve the hazard consistency of the applied loading. Therefore, a first step consists in estimating the distribution of the secondary IM (i.e, *PGA*) w.r.t. *SA*(0.29s), using the dataset of the input ground-motion records: a median line and its 16%-84% confidence intervals are then plotted

(see Figure 6, left). The space delimited by this interval provides also practical guidance on the validity domain of the fragility surface, in the sense that it identifies the IM combinations that are very unlikely.

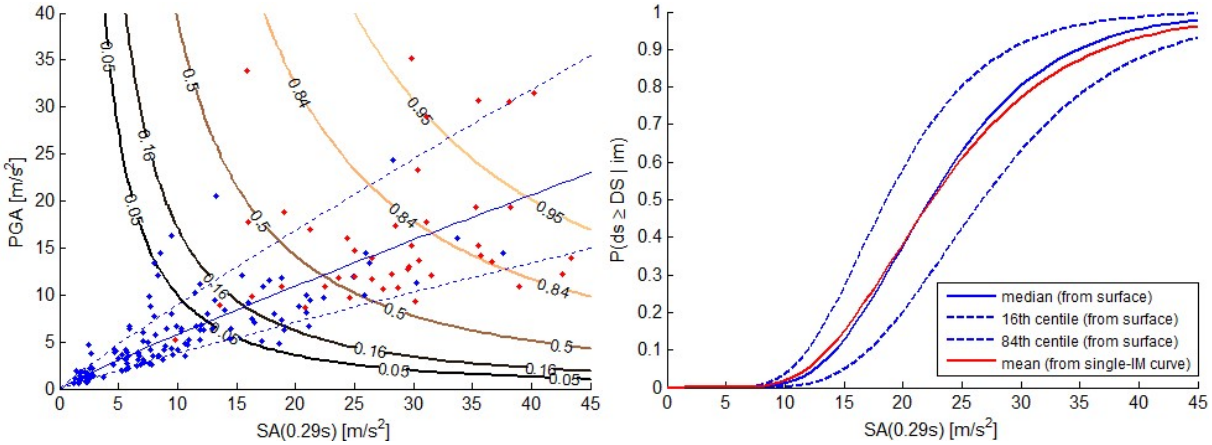

Figure 6: Left: fragility surface w.r.t. PGA and SA(0.29s), the solid blue line represents the median of the PGA-SA(0.29s) distribution and the dashed blue lines the 16%-84% confidence intervals; Right: equivalent fragility curves w.r.t. SA(0.29s).

It is then proposed to generate "slices" of the fragility surfaces by following the distribution of *PGA* as a function of *SA*(0.29s) (i.e., solid and dashed lines in Figure 6, left). As a result, the "slices", now represented as a function of the single IM *SA*(0.29s), may be compared to the original scalar-IM fragility curve. The fragility curves in Figure 6, right, are identified as follows:

- *Solid red line*: "mean" fragility curve corresponding to the scalar-IM fragility curve, derived w.r.t. *SA*(0.29s) only;
- *Solid blue line*: median fragility curve corresponding to the "median" slice of the fragility surface;
- *Dashed blue lines*: 16%-84% confidence intervals around the median fragility, corresponding to the lower and upper bounds of the slices of the fragility surfaces.

Finally, this family of fragility functions corresponds to the probabilistic framework by Kennedy et al. (1980), where the identification of aleatory and epistemic uncertainties (β_R and β_U, respectively), as introduced in Eq. 1. The mean fragility curve, w.r.t. to *SA*(0.29s), has a total standard deviation $\beta_{tot} = (\beta_R^2 + \beta_U^2)^{1/2} = 0.390$. Meanwhile, the median fragility curve, obtained from the fragility surface w.r.t. *PGA* and *SA*(0.29s), has a standard deviation of 0.349, which actually corresponds to the aleatory randomness term only (i.e., β_R). The confidence intervals obtained from the graphical construction in Figure 6 are then used to estimate the epistemic uncertainty term, i.e. $\beta_U \approx 0.174$.

It may be concluded that the vector-IM fragility functions lead to the transfer of a part of the record-to-record variability into a form of epistemic uncertainty, which is related to the description of the seismic loading given the hazard at the specific site.

5.2 Uncertainty due to the number of data points

The epistemic uncertainty due to the number of data points in the simulations, i.e. related to the quality of the statistical estimation of the fragility parameters, may also be evaluated in the case of vector-IM fragility functions. To this end, a bootstrap sampling approach is ap-

plied to the fragility surface w.r.t. of *PGA* and *SA*(0.29s). The outcomes of the bootstrap sampling are displayed in Figure 7, in the case of 16%-84% confidence intervals.

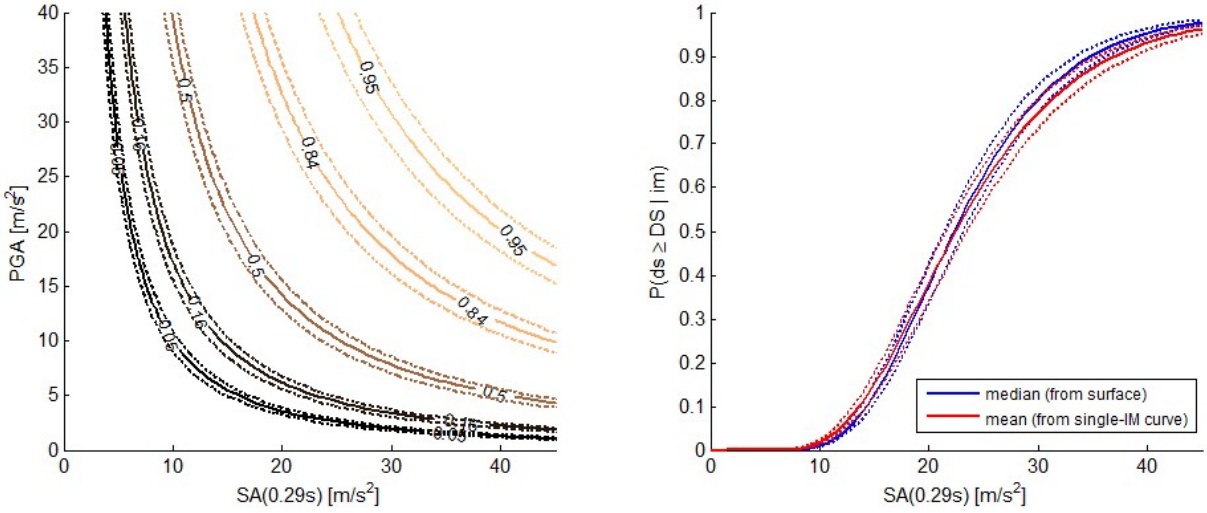

Figure 7: Left: fragility surface w.r.t. PGA and SA(0.29s), with the 16%-84% confidence intervals (dashed lines) due to the statistical estimation; Right: equivalent fragility curves w.r.t. SA(0.29s) and related 16%-84% confidence intervals due to the statistical estimation.

In order to compare these confidence intervals with the ones estimated for the corresponding scalar-IM fragility curves, the fragility surface is reduced to a scalar case w.r.t. *SA*(0.29s) only, following the same approach as before (i.e., use of a "median" slice of the fragility surface). The curves in Figure 7, right, reveal a similar order of magnitude for the confidence intervals due to the amount of data points, both for the single-IM fragility curve ($\beta_U = 0.043$) and for the vector-IM fragility function ($\beta_U = 0.048$).

5.3 Uncertainty due to the variability of mechanical and geometrical parameters

Based on the 360 randomly generated structures, a new vector-IM fragility curve is derived by incorporating the effect of the epistemic uncertainties related to the $p = 10$ mechanical and geometrical parameters (described in Table 2). This is done by adding linear terms in Eq. 3 as:

$$\frac{1}{2}\left[1 + \mathrm{erf}\left(c_1 + c_2 ln\, im_1 + c_3 ln\, im_2 + \sum_{i=1}^{p} c_{x_i} x_i\right)\right] \tag{7}$$

where x_i is the i^{th} uncertain parameter and c_{x_i} the corresponding regression coefficient.

When the ten parameters are considered at the same time, it is found that the MLE-based regression cannot converge, due to the limited number of data points. Therefore, partial combinations of parameters are tested until the most influent parameters are identified. As a result, the MLE regression on Eq. 7 has been carried out with three geometrical parameters, namely the thicknesses e_2, e_4 and e_6 (see Table 5). The analysis by the *p*-value of the Wald test has been used to decide whether the corresponding parameter is significant or not, using a significance threshold at 5%. Interestingly, the mechanical parameters barely affect the mean of the vector-IM fragility function (at 5% significance).

Parameter	Regression coefficient	Std. Error	*p*-value (Wald statistic)
c_1 - Intercept	-22.008	4.030	4.73e-8
c_2 - *SA*(0.29s)	1.991	0.359	2.98e-8

c_3 - *PGA*	1.284	0.342	1.75e-4
c_4 - Thickness e_2	95.504	46.303	0.0392
c_5 - Thickness e_4	135.553	46.184	0.0033
c_6 - Thickness e_6	98.819	44.217	0.0254

Table 5: Regression coefficients, standard error and *p*-value of the Wald statistic of the vector-IM fragility function.

On this basis, a new vector-IM fragility function is plotted in Figure 8, left, by incorporating only the significant geometrical parameters. The median surface is obtained by setting the thickness parameters at the median of their uniform distribution interval (cf. Table 2), while its 16%-84% confidence bounds account for the corresponding variations of these parameters. The same approach is applied to the construction of the single-IM fragility curve, enabling the comparison between the two approaches (see Figure 8, right).

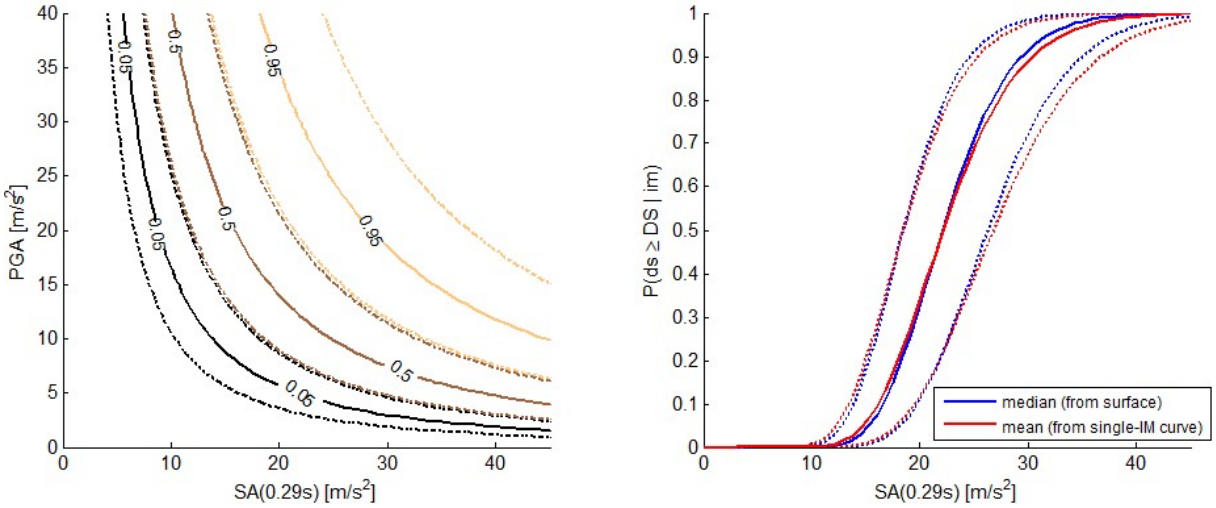

Figure 8: Left: fragility surface w.r.t. PGA and SA(0.29s) with the 16%-84% confidence intervals (dashed lines) due to variability in the parameters; Right: equivalent fragility curves w.r.t. SA(0.29s) and related 16%-84% confidence intervals due to variability in the parameters.

These fragility models, based on the disaggregation of the variability of some mechanical or geometrical, show a significantly reduced dispersion, when compared to the ones that do not use these parameters as input variables (i.e., models in Figures 6 and 7). However, it should be noted that some regression coefficients have been estimated with a significant standard error (see Table 5), which implies that much more data points from simulations should be necessary in order to get stable estimates.

5.4 Discussion on the respective contribution of uncertainty sources

The previous statistical analyses have helped identifying the uncertainty sources at play in different fragility modelling strategies (see Figure 9):

- *Single-IM fragility curve*: the aleatory dispersion β_R includes most of the variability, with a very small contribution of the epistemic uncertainty due to the number of data points.

- *Vector-IM fragility function*: a part of the record-to-record variability may be interpreted as an epistemic uncertainty term. The uncertainty term due to the number of data points is slightly larger, without becoming significant.
- *Vector-IM fragility function accounting for mechanical and geometrical parameters*: a multivariate MLE-based regression has allowed to explicitly account for the most influent mechanical and geometrical, further reducing the aleatory dispersion.

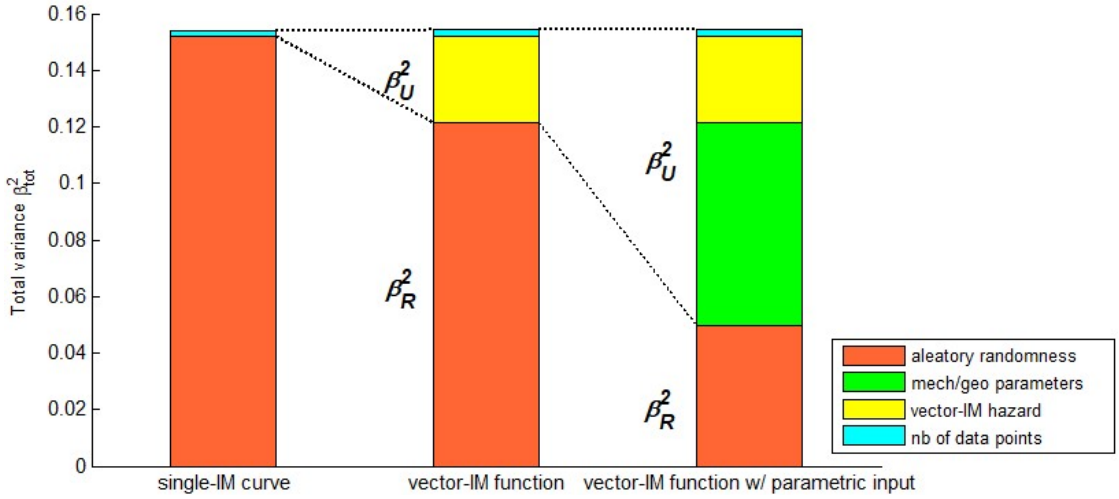

Figure 9: Decomposition of the uncertainty sources in terms β_R (aleatory) and β_U (epistemic), depending on the modelling strategy used.

It may be noted that, for the considered application, the use of vector-IMs contributes to a significant part of the total dispersion, although the uncertainty due to the variability of mechanical and geometrical parameters appears to be much larger. When computing the HCLPF, the uncertainty decomposition from Figure 9 is preserved, according to the following equation derived from Eq. 1:

$$SA(0.29s)_{HCLPF} = \alpha_{SA} \cdot exp[-(\beta_R + \beta_U) \cdot \Phi^{-1}(0.95)] \quad (8)$$

Therefore, when assessing a given element from an NPP, its HPCLPF may be gradually reduced if the following measures are taken to lower the epistemic uncertainty term:

- Reduction of β_U due to the number of data points, by performing more simulations (i.e., increased confidence in the statistical estimation);
- Reduction of β_U due to the description of the seismic loading, by performing a vector probabilistic hazard assessment of the considered site (i.e., more accurate knowledge of the excepted vector-IM distribution);
- Reduction of β_U due to the variability of mechanical and geometrical parameters, by testing or qualifying the materials used (i.e., increase of knowledge).

6 CONCLUSIONS

This study has confronted the concept of vector-IM fragility functions to the probabilistic framework commonly employed in nuclear applications. This exercise has then allowed for a systematic analysis of various sources of aleatory or epistemic uncertainty.

Regarding the selection of IMs, carefully selected vector-IMs make excellent candidates in terms of IM sufficiency and efficiency, when compared to scalar IMs. As a result, vector-IM

fragility functions tend to generate less dispersion (i.e., aleatory uncertainty due to record-to-record variability) than single-IM fragility curves: this difference may be interpreted as a partial transfer from the record-to-record variability to an epistemic uncertainty component that is related to the description of the seismic loading given the hazard at the studied site. However, in the present example, it appears that the epistemic uncertainty due to the variability of mechanical and geometrical parameters is still much larger.

Although a wide range of statistical tools are available for the quantification and propagation of sources of uncertainties, it appears that all the epistemic uncertainties usually cannot be adequately covered and accounted for (e.g., much more simulations would be required in order to accurately model the variability of the mechanical and geometrical parameters). In most cases, expert judgment would be necessary in order to constrain the assumptions and to interpret the simulation results.

Finally, the present study has followed the lognormal assumption for the functional form of the fragility functions. This constraint, while convenient for the combination of nested uncertainty terms, is bound to introduce significant biases in the statistical estimates (i.e., the outcomes of the MLE-based regression). A similar analysis based on undefined functional forms would also be able to deliver valuable lessons.

7 ACKNOWLEDGEMENT

This study has been carried out within the NARSIS project, which has received funding from the European Union's H2020-Euratom Program under grant agreement N° 755439. The Pacific Earthquake Engineering Research Center is gratefully acknowledged for providing access to their database of ground-motion records.

REFERENCES

[1] EPRI, *Seismic Probabilistic Risk Assessment Implementation Guide*. EPRI Report TR-1002989, Palo Alto, CA, 2003.

[2] R.P. Kennedy, C.A. Cornell, R.D. Campbell, S. Kaplan, H.F. Perla, Probabilistic seismic safety study of an existing nuclear power plant. *Nuclear Engineering and Design*, **59**(2), 305-338, 1980.

[3] N. Luco, C.A. Cornell, Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. *Earthquake Spectra*, **23**(2), 357-392, 2007.

[4] J.E. Padgett, B.G. Nielson, R. DesRoches, Selection of optimal intensity measures in probabilistic seismic demand models of highway bridge portfolios. *Earthquake Engineering & Structural Dynamics*, **37**(5), 711-725, 2008.

[5] J.W. Baker, C.A. Cornell, A vector-valued ground motion intensity measure consisting of spectral acceleration and epsilon. *Earthquake Engineering & Structural Dynamics,* **34**, 1193-1217, 2005.

[6] D.M. Seyedi, P. Gehl, J. Douglas, L. Davenne, N. Mezher, S. Ghavamian, Development of seismic fragility surfaces for reinforced concrete buildings by means of nonlinear time-history analysis. *Earthquake Engineering & Structural Dynamics*, **39**(1), 91-108, 2010.

[7] A. Modica, P.J. Stafford, Vector fragility surfaces for reinforced concrete frames in Europe. *Bulletin of Earthquake Engineering,* **12**, 1725-1753, 2014.

[8] N. Rahni., M. Lancieri, C. Clement, G. Nahas, J. Clement, L. Vivan, Y. Guigueno, E. Raimond, An original approach to derived seismic fragility curves – Application to a PWR main steam line. *Proceedings of the International Topical Meeting on Probabilistic Safety Assessment and Analysis (PSA2017)*, Pittsburgh, PA, 2017.

[9] A. Combescure, A. Hoffmann, P. Pasquet, The CASTEM finite element system. Brebbia C.A. eds. *Finite Element Systems*, Springer, Berlin, Heidelberg, 1982.

[10] T. Lin, C.B. Haselton, J.W. Baker, Conditional spectrum-based ground motion selection. Part I: hazard consistency for risk-based assessments. *Earthquake Engineering & Structural Dynamics*, **42**(12), 1847-1865, 2013.

[11] J. Woessner, L. Danciu, P. Kaestli, D. Monelli, *Database of seismogenic zones, Mmax, earthquake activity rates, ground motion attenuation relations and associated logic trees*. FP7 SHARE Deliverable Report D6.6, 2013.

[12] N. Jayaram, T. Lin, J.W. Baker, A computationally efficient ground-motion selection algorithm for matching a target response spectrum mean and variance. *Earthquake Spectra*, **27**(3), 797-815, 2011.

[13] PEER, *PEER NGA-West2 Database*. Pacific Earthquake Engineering Research Center, https://ngawest2.berkeley.edu, 2013.

[14] M. Shinozuka, M. Feng, J. Lee, T. Naganuma, Statistical analysis of fragility curves. *Journal of Engineering Mechanics*, **126**(12), 1224-1231, 2000.

[15] M. Lancieri, M. Renault, C. Berge-Thierry, P. Gueguen, D. Baumont, M. Perrault, Strategy for the selection of input ground motion for inelastic structural response analysis based on naïve Bayesian classifier. *Bulletin of Earthquake Engineering*, **13**(9), 2517-2546, 2015.

[16] M.A. Hariri-Ardebili, V.E. Saouma, Probabilistic seismic demand model and optimal intensity measure for concrete dams. *Structural Safety*, **59**, 67-85, 2016.

[17] P. Gehl, D.M. Seyedi, J. Douglas, Vector-valued fragility functions for seismic risk evaluation. *Bulletin of Earthquake Engineering*, **11**(2), 365-384, 2013.