Submitted to Natural Hazards

IMPACT ON LOSS/RISK ASSESSMENTS OF INTER-MODEL VARIABILITY IN VULNERABILITY ANALYSIS

Jacopo Selva [a], Sotiris Argyroudis [b] and Kyriazis Pitilakis[c]

[a] Istituto Nazionale di Geofisica e Vulcanologia, via D. Creti 12, 40128, Bologna, Italy, selva@bo.ingv.it
[b] Aristotle University of Thessaloniki, Dept. of Civil Engineering, P.O.B. 424, 54124, Thessaloniki, Greece, sarg@civil.auth.gr
[c] Aristotle University of Thessaloniki, Dept. of Civil Engineering, P.O.B. 424, 54124, Thessaloniki, Greece, kpitilak@civil.auth.gr

Keywords: Fragility Curves, Bayesian approach, Epistemic Uncertainty, Inter-Model Variability, Seismic Risk

Fragility curves (FCs) constitute an emerging tool for the seismic risk assessment of all elements at risk. They express the probability of a structure being damaged beyond a specific damage state for a given seismic input motion parameter, incorporating the most important sources of uncertainties, i.e., seismic demand, capacity and definition of damage states. Nevertheless, the implementation of FCs in loss/risk assessments introduces other important sources of uncertainty, related to the usually limited knowledge about the elements at risk (e.g. inventory, typology). In this paper, it is developed a general methodology to merge into a single model the information provided by multiple FC models, weighting them according to their credibility/applicability. This combination enables to efficiently capture Inter-Model Variability (IMV) and to propagate it into risk/loss assessments, allowing the treatment of a large spectrum of vulnerability-related uncertainties, usually neglected. As case study, fragility curves for shallow tunnels in alluvial deposits, when subjected to transversal seismic loading, are developed with two conventional procedures, based on a quasi-static numerical approach. Noteworthy, loss/risk assessments resulting from such conventional methods show significant unexpected differences. Conventional fragilities are then combined in a Bayesian *framework, in which also probability values are treated as random variables, characterized by their probability density functions. The results show that this Bayesian Combined Model (BCM) efficiently projects the whole variability of input models into risk/loss estimations. This demonstrates that BCM is a suitable framework to treat IMV in vulnerability assessments, in a straightforward and explicit manner.*

1- INTRODUCTION

In their original form, Fragility Curves (FCs) describe the probability of a structure being damaged beyond a specific damage state for various levels of ground shaking (e.g., ALA 2001; NIBS 2004). They are widely adopted in seismic expected loss and risk assessments, since they are a valuable tool to explicitly evaluate vulnerability of structures (e.g., NIBS 2004; Cornell and Krawinkler 2000; Pitilakis et al 2006). Their applicability in estimating the probability of damage levels of particular element at risk contributes in the retrofitting decisions, emergency response planning and estimation of direct and indirect losses of built environments as well as lifeline systems (Pitilakis et al 2006; Kappos et al 2008; Azevedo et al 2010). Also, the use of fragility curves goes beyond the seismic risk analysis (e.g., Spence et al 2005), and it has been proposed as the general framework for vulnerability assessment in all natural risks (e.g., Douglas 2007; Shmidt et al 2011).

Many methods are used to generate FCs, based either on past recorded damages (e.g., Basöz et al 1999 for bridges; Maruyama et al 2010 for expressway embankments; Rossetto and Elnashai 2003 for buildings), analytical modeling of the structures behavior under input ground motion (e.g., Moschonas et al 2009 for bridges; Akkar et al 2005 for buildings), expert judgment (e.g., ATC-13; ATC-25), as well as, on a combination of such methods (hybrid methods, e.g., Kappos et al 2006 for buildings). All methods are based on the definition of a given set of damage states, possibly defined in terms of measurable quantities (Kappos 1997; Mackie and Stojadinovic 2003) and on an appropriate intensity measure (*IM*) describing ground motion (Pinto 2007; Mackie and Stojadinovic 2003). Commonly, seismic FCs are represented as cumulative log-normal distribution (Shinozuka et al 2000; Wen et al. 2003; Ellingwood and Kinali 2009, NIBS 2004). This assumption reduces the problem into the definition of the several parameters (e.g., Kennedy and Ravindra 1984; Choun and Elnashai 2010). Aleatory and epistemic uncertainties are often not separated and the problem is further reduced to the assessment of the two parameters of the log-normal distribution, that is, median m and logarithmic standard deviation β. Given damage data (real or modeled), m and β are assessed through different methodologies, from purely statistical methods, either classical (e.g., Rossetto and Elnashai 2003; Shinozuka et al 2003) or Bayesian (e.g., Shinghal and Kiremidjian 1996; Straub and Der Kiureghian 2008; Koutsourelakis 2010), to more physical assessment of critical points of structures, where seismic demand overcome structure capacity (e.g., Moschonas et al 2009; Nielson and DesRoches 2007). This procedure provides single best estimate FCs, which considers a composite variability parameter that does not explicitly separate out uncertainties (e.g., Bhargava et al 2002).

All the steps toward the quantitative definition of FCs bring into the estimation procedures many uncertainties, both epistemic and aleatory. The most commonly analyzed uncertainty sources are the demand, capacity and damage state definition’s uncertainties (e.g., NIBS 2004; Pinto 2007). Demand uncertainty reflects the fact that *IM* is not exactly sufficient, so different records of ground motion with equal *IM* may have different effects on the same structure (e.g. Karim and Yamazaki 2001; Nielson and DesRoches 2007). Capacity uncertainty reflects the variability of structure properties as well as the fact that the modeling procedures are not perfect. Damage state definition uncertainties are due to the fact that the thresholds of the damage indexes or parameters used to define damage states are not known. Such uncertainties are usually assumed independent, while in many cases they are quantified through a single parameter based on expert judgement (e.g., NIBS 2004).

The variety of analysis techniques, structural idealizations, seismic hazard and damage models being used, strongly influence the derived vulnerability curve shapes, and different choices have been seen to result in significant discrepancies between the seismic risk assessments made by different authorities for the same location, structure type and seismicity (Rosseto and Elnashai 2005). This inevitably leads to a large availability in literature of different FCs, even for similar or identical structures. As a matter of fact, even fragility curves derived from the same model or the same dataset may show important differences. As an example, Basöz and Kiremidjian (1998) developed different fragility curves for the same bridge damage data after the 1994, Northridge earthquake due to two different available sets of PGA values. Moreover, Shinozuka et al (2003) proposed fragility curves for the same dataset following another statistical analysis procedure. This significant variability has been explicitly shown in the ongoing European project SYNER-G (2010-2013), where different FCs, derived from different approaches, have been collected for several European typologies of buildings in a single tool. However, when FCs are applied in loss/risk assessments, in common practice only one single set of FCs is used (e.g. Kappos et al 2008; Azevedo et al 2010; Pitilakis et al 2010; Bommer et al 2008), usually referred to groups or typologies of structures. Often, if not always, there are no objective reasons to choose one set of curves instead of another, considering the large variety in FC , the variability of structures within each typology and the often inhomogeneous definition of typologies in different studies. In addition, the usually relatively poor knowledge about many or even most of the assets at risk further increases the (epistemic) uncertainty on the selection of one single set of FCs. However, the variability on the results due to different and subjective choices related to the vulnerability assessments is usually not considered at all, even if it may potentially introduce non predictable consequences in loss/risk assessments (e.g., Paté-Cornell 1996; Winkler 1996).

In seismic hazard, as well as in other fields, such Inter-Model Variability (IMV) is often assessed through Logic Trees, by mixing different approaches (Cornell and Merz 1975; McGuire 1977; McGuire and Shedlock 1981; Giner et al 2002; Gruppo di Lavoro MPS 2004; SHARE project, 2009-2012) or fully treating all uncertainties (Kulkarni et al 1984; Coppersmith and Youngs 1986; Electric Power Research Institute 1987; National Research Council 1988). In logic trees, alternative modeling choices are combined together, weighting each choice by its overall applicability/credibility. The result is a discrete set of different hazard curves, which means that, for each single value of the *IM*, a discrete set of "possible" probabilities is assessed, often expressed as hazard maps at different percentiles (e.g., Gruppo di Lavoro MPS 2004). However, Logic Trees have several important drawbacks (e.g., Bommer and Scherbaum 2008). Among them, we mention (i) the fact that each branch duplication largely increases the computational effort, making practically difficult its application in medium/large scale loss/risk assessments (e.g., SYNER-G 2010-2013); (ii) the difficulty in defining a mutually exclusive and collectively exhaustive set of alternative models and (iii) the consequent impossibility to treat uncertainty if only one model is available, even when weakly constrained or not completely applicable. In addition, since Logic Trees consider only a discrete number of alternative models, the whole variability among models is never explored. A more structured method to treat IMV consists of the use of the Bayesian probability concept, which allows us to assign a 'subjective' belief to different hypotheses, thing that is not conceivable in a classical framework (e.g., Lindley 1965; Gelman e al. 1995; Hofer 1996). In practice, this means that future frequencies of events (i.e., their 'probability') can be treated as random variables, characterized by their own probability density functions (Gelman e al. 1995). Such an idea has found lately applicability also in different fields of natural hazard analysis (e.g., Wen et al 2003; Marzocchi et al 2008, 2010; Grezio et al 2010).

In this paper, we propose an explorative application of the Bayesian probability concept to deal with IMV in vulnerability analysis in loss/risk assessments. In particular, we adopt a Bayesian framework to merge into a single overall prior model (Bayesian Combined Model, BCM) the information provided by different available methodologies, not necessarily homogeneous in their formulations, to be eventually fit to pertinent independent data, when available. Then, we evaluate the performance of BCM in propagating IMV into loss/risk assessments. Note that the goal of this is not to develop better or more refined FC models for one single asset, but to provide more reliable loss/risk assessments by propagating in them IMV. In fact, the epistemic uncertainty that emerges when FC models are selected in loss/risk assessment is relative to the applicability of FC, more than to FCs by themselves, and it can be modeled accounting for the IMV. Indeed, this uncertainty is usually significant when the vulnerability of a specific target area is modeled starting from generic fragility models and it is not necessarily related to their structural analyses. In this, the presented approach differs significantly from Bayesian FC approaches, since they have different goals (e.g., Straub and Der Kiureghian 2008; Koutsourelakis 2010).

In the followings, we first develop the Bayesian framework in which different and non-homogeneous FC models can be combined in order to propagate IMV in loss/risk assessments (*section 2*). The applicability of this procedure is then discussed (*section 3*), showing that it may potentially treat a broad range of different sources of IMV in loss/risk assessments. Its practical applicability is finally demonstrated and discussed through one realistic, though simplified, case study (*section 4*).

2- BAYESIAN COMBINATION MODEL (BCM) FOR VULNERABILITY ASSESSMENT

The ultimate goal of all Fragility Curves (FC) models is to assess the probability that a given damage state is reached or exceeded, given the occurrence of a certain level of the Intensity Measure (IM) that describes the size of hazardous phenomena (e.g., ground shaking). The punctual probability of each damage state (π_i) can be obtained from the FCs, that is, for m damage states:

$$\begin{cases} \pi_0 = 1 - F_1(IM) \\ \pi_1 = F_1(IM) - F_2(IM) \\ \pi_2 = F_2(IM) - F_3(IM) \\ \dots \\ \pi_m = F_m(IM) \end{cases} \tag{1}$$

where $F_i(IM)$ indicates the FC (exceedance probability) relative to damage state i. In Eq. 1, the index runs from 0 (no damages) to m (collapse) and, by definition, $\pi_i > 0$ and sum to 1.

Given the occurrence of an earthquake, FC models and the assumed repair cost for each damage state, it is possible to assess the expected losses for a given set of elements. In practice, the expected cost to repair the generic k-th element at risk depends on its actual damage state, and can be written as fraction of the total replacement cost ($RC^{(k)}$):

$$C_i^{(k)} = RC^{(k)} \cdot CDF_i \tag{2}$$

where the index i indicates the i-th damage state, and CDF_i (cost damage factor) is the fraction of $RC^{(k)}$ necessary to repair the i-th damage state. Note that, for simplicity, in this application, uncertainty in CDF_i is not included. However, also such uncertainties may be treated by sampling each CDF_i value from specific probability distributions (Stergiou and Kiremidjian 2006).

Assuming a seismic scenario, with a Monte Carlo (MC) simulation, it is possible to obtain a sample of losses for each element at risk (eq. 1), and thus a sample of the expected total losses, for the entire set of elements (eq. 2). Assuming that damages, given an IM scenario, are statistically independent, a single random damage state i^* can be selected for each element by comparing its π_i with a random number in the interval [0,1]. Given the obtained random damage scenarios, the total loss for the whole set of elements, for each realization, can be evaluated by summing over all elements, that is

$$l = \sum_k C_i^{(k)} \tag{3}$$

After repeating this procedure many times, we obtain a sample of possible total losses, given the IM values at each element location provided by a seismic scenario. From this sample of possible losses, we can evaluate the loss curve for a given scenario IM = im, defined as

$$Lc(im;\vec{\pi}) = p(\geq l \mid im;\vec{\pi}) \tag{4}$$

in which its dependence on the fragility model's results (through π_i) is explicitly reported. The average of expected losses l_m reads:

$$l_m(a,\vec{\pi}) = \sum_k \sum_i C_i^{(k)} \pi_i \tag{5}$$

In risk assessments, losses for each im values are combined together. For example, in medium/large areas, the Average Annualized Earthquake Losses (*AEL*) is often adopted as risk index (e.g., FEMA 2008). *AEL* can be assessed as the average of losses $l_m(a,\pi_i)$ over all possible ground motion values IM = *im*, that is

$$AEL(\vec{\pi}) = \int_{IM} l_m(im, \vec{\pi})dh(im) \qquad (6)$$

where *dh(im)* is the non-cumulative hazard, for an exposure time of 1 year.

Different FC models for the same element at risk result in different assessments of the probabilities π_i in eq. 1, and consequently different punctual assessments of losses (eqs. 4 and 5) and risk (eq. 6). Here, we indicate with $\pi_i^{(Mk)}$ the probability values resulting from the FC model M_k. With the goal of combining into a single high-level model the results of distinct standard FC models, without restriction in their formulation, as well as specific pertinent observations, the variability in π_i can be treated using the Bayesian probability concept, in which each probability value may be treated as a variable (e.g., Gelman et al. 1995). A schematic representation of this concept is reported in **Fig. 1**. It is beyond the goals of this paper to discuss the "philosophical" implications of this "probability of probability" assessment, but we stress that (i) this is a quite common procedure in dealing with this type of uncertainties, in many fields of science (starting from Mosimann 1962), including geophysics (e.g., Marzocchi et al. 2008, 2010; Grezio et al. 2010; Selva et al. 2010, 2012) and earthquake engineering (e.g., Paté-Cornell 1996; Wen et al 2003), and (ii) a similar philosophy is implicitly assumed whenever procedures like Logic Tree are adopted (e.g., Bommer and Scherbaum 2008 and references therein).

To assess the distribution $[\vec{\pi}]$, given a certain number of past data *{D}*, that is, a set of observed damage states due to a given IM value, we can make use of Bayes' rule, that is

$$[\vec{\pi} \mid \{D\}] \propto [\{D\} \mid \vec{\pi}][\vec{\pi}] \qquad (7)$$

where $[\vec{\pi}]$ is the prior, $[\{D\} \mid \vec{\pi}]$ is the likelihood, and $[\vec{\pi} \mid \{D\}]$ is the posterior probability distribution, the latter representing the final result of the inference. The choice of the functional form for prior and the likelihood distributions represents the core of the Bayesian inference and necessarily implies several assumptions about the modeled process (e.g., Gelman et al. 1995).

The much more common situation in loss/risk assessmentsis the scarcity or even non-availability of {D} for most of the structures or classes of structures at risk in the target area. If large datasets were available, both the construction of specific fragility curves, and/or the selection of appropriate ones, would be a rather simple task, and thus IMV would be almost negligible. On the other hand, a rather high availability of theoretical models is quite common, at least for many structure typologies (e.g., SYNERG 2010-2013). However, often, there are not objective reasons to select one specific FC model, among the available ones. This uncertainty is not necessarily related to each FC model itself, but it is essentially linked to scarce knowledge about the target stock (see discussion in *section 3*). Hence, this epistemic uncertainty is related to the application of FC in loss/risk assessments, rather than to FCs by themselves, and it can be modeled accounting for the Inter-Model Variability (IMV). Since this variability essentially affects the prior distribution $[\vec{\pi}]$, we concentrate our attention on how IMV can be modeled at this level, to be fitted to pertinent independent (i.e., not used in the prior) data through Eq. 2, when these are available.

The specific goal is merging the information brought by a given number of starting models (M_k, $k=1,2, \ldots,N_M$) into a single prior probabilistic model $[\vec{\pi}]$ that accounts for the variability on their results, that is, the Inter-Model Variability (IMV), allowing the whole variability around the input model results to be explored. The distributions $[\vec{\pi}]$ are, by definition, subjective (e.g., Gelman et al 1995). However, it is definitely more subjective to assume that only one of the available models is correct and/or applicable (Paté-Cornell 1996; Marzocchi et al. 2008). As a matter of fact, to assume one specific model means that the other FCs are assumed as wrong/non-applicable, even when they are almost equally acceptable. In addition, this assumption also implies that we do not distinguish at all between well-accepted and consolidated in literature models and less constrained ones. This may undoubtlylead to uncontrolled biases in the final loss/risk assessments and to wrong conclusions/decisions (e.g., Paté-Cornell 1996; Woo 1999).

The first step to set BCM consists in assigning a specific functional form to the prior probability density function $[\vec{\pi}]$ in Eq. 7. Given an IM value, the probabilities π_i for all damage states represent a partition of the event ‘damage’. In other words, damage states form a set of exhaustive and mutually exclusive events, that is, for each model and each IM value, the punctual probability of damage states (e.g., π_0 for no damage, π_1 for minor damages, π_2 for moderate damages, π_3 for extensive and π_4 for complete damages) sum to 1. In this case, a common choice in statistics is an *m*-dimensional Dirichlet distribution (e.g., Mosimann 1962; Gelman et al. 1995; in natural hazards: Marzocchi et al. 2008, 2010; Selva et al. 2010, 2012):

$$[\vec{\pi}] = Dir_m(\vec{\pi}; \vec{a}) \tag{8}$$

where the vector $[\vec{\pi}]$ = (π_0, π_1, …, π_m) contains all punctual probabilities, and the vector $\vec{a}$ = (a_1, a_2, ..., a_{m+1}) contains the parameters of the Dirichlet distribution, that is, the hyper-parameters. Note that (i) the total variability of this prior distribution can be fully obtained by setting the hyper-parameters $\vec{a}$ for all IM values, and that this is independent from the number of considered models, and (ii) the Dirichlet distribution automatically accounts for the correlation among the probabilities relative to the different damage states. Since the marginal distribution of the Dirichlet is a Beta distribution, which is unimodal, this functional choice implies the assumption that the transition among different FCs is expected to be soft, in other words, intermediate FCs are expected to exist and be applicable. The consequence of this on the applicability of BCM is discussed in *section 3*. The sum $\sum_i a_i$ is inversely proportional to the total variance and thus represents a prompt of the global estimated IMV (e.g., Marzocchi et al. 2008).

The second step of BCM is to set the prior distribution, starting from the models results. Different procedures may be adopted, in which different levels of control of the average and/or the variance of the distribution $[\vec{\pi}]$ are set. For example, the means of $[\vec{\pi}]$ can be set as the (weighted) average of models' $\vec{\pi}^{(Mk)}$, and variance according to the (subjective) credibility of each model (e.g., Marzocchi et al. 2010). Here, we prefer a procedure in which both means and variance are controlled by the input model, since we want to investigate the whole IMV. In this case, the probability assessments of different input FC models can be treated as independent samples from the unknown prior $[\vec{\pi}]$. Bayes' rule on $\vec{a}$ reads:

$$[\vec{a} \mid \{M\}] \propto [\{M\} \mid \vec{a}][\vec{a}] \qquad (9)$$

where *{M}* stands for the models' $\pi_i^{(Mk)}$. Since in common practice the probabilities π_i are treated as perfectly known, instead of adding a further layer to the Bayesian model, we prefer to keep the model simple. Hence, we infer the best guess values of the hyper-parameters $\vec{a}$, starting from the results obtained by the set of input models M_k. To assess the best guess a^*, we make use of a Maximum A Posteriori (MAP) estimation, that is, we select the parameters that maximize $[\vec{a} \mid \{M\}]$. The simplest choice for the prior $[\vec{a}]$ is an improper non-informative uniform distribution, choice that makes MAP equivalent to a standard Maximum Likelihood (ML) method. To consider models with different credibility, the likelihood function $[\{M\} \mid \vec{a}]$ can be weighted by the credibility/applicability of each model (e.g., Wang et al. 2004; Ahmed et al. 2005). In this case, the best guess hyper-parameters $\vec{a}^*$ are selected by maximizing the weighted likelihood

$$[\vec{\pi}] \equiv [\vec{\pi} \mid \{M\}] = Dir_m(\vec{\pi}; \vec{a}^*) \quad \leftarrow \quad \vec{a}^* = \text{argmax}_a\left(\prod_k \left[Dir_m(\vec{\pi}^{(Mk)}; \vec{a})\right]^{w_k}\right) \qquad (10)$$

where k runs over the models; the weight w_k represents the subjective credibility of the k-th model, its actual values matter in a relative, more than absolute, sense; $Dir_m(\vec{\pi}; \vec{a})$ represents the m-dimensional Dirichlet probability density function with parameters $\vec{a}$; $\vec{\pi}^{(Mk)}$ is a vector containing the guessed probabilities from the FC model M_k for a given value of IM (from Eq. 1). Both mean and variance of $[\vec{\pi} \mid \{M\}]$ are controlled by the input models. The underlying assumption is that such input models well represent the whole IMV. In particular, since the variance of $[\vec{\pi}]$ is controlled by the models, it represents the *a posteriori* estimation of IMV, and it is small only when the input model are in agreement. In addition, since $\sum_i a^*_i$ changes at each IM level, the model permits different levels of IMV to be considered.

The obtained prior $[\vec{\pi}] \equiv [\vec{\pi} \mid \{M\}]$ can be input in Eq. 7, and updated in light of new pertinent data, if any, in which case the IMV it will reshaped in agreement with new observations, that is

$$\left[\vec{\pi} \mid \{D\},\{M\}\right] \propto [\{D\} \mid \vec{\pi}][\vec{\pi} \mid \{M\}] = [\{D\} \mid \vec{\pi}] \cdot Dir_{m}(\vec{\pi}; \vec{a}^{*}, \vec{\pi}^{(Mk)}) \qquad (11)$$

where a standard choice for the functional form of the likelihood $[\{D\} \mid \vec{\pi}]$ is a Multinomial distribution (from Mosimann 1962). In the followings, to simplify the notations, we will always refer to the final result of BCM as $[\vec{\pi}]$, noted that this symbol may represent either the prior, or the posterior distribution.

It is worth to stress that, with this parameterization, peaks on specific probability values may arise only by a convergence of the input models, or by a large set of coherent observation *{D}*. Note also that BCM does not assume any functional form for FC models *{M}*, in order to extend its applicability to all non log-normal FC methodologies. Indeed, this is rather common both for seismic (e.g., Basöz and Kiremidjian 1998; Dueñas-Osorio et al. 2007) and non-seismic (e.g., Spence et al. 2005) vulnerability assessments. This possibility enables to make use as potential input models of the large set of studies available in literature, which is particularly important whenever uncertainty of epistemic type is treated (e.g., Marzocchi et al 2008, 2010).

The last step of BCM model is to propagate IMV in loss/risk assessment. Indeed, BCM models the variability in the probability assessments provided by different FC input models, which is the input for loss (eqs. 4 and 5) and risk (eq. 6) assessments. Hence, the IMV on π_i propagates in loss/risk assessments, providing variability in their numerical assessments. In other words, instead of single punctual assessments, BCM provides an estimate on the uncertainty on those values, since the probabilities π_i are not assumed as perfectly known.

In particular, for each single sample of π_i, different loss curves *Lc* can be evaluated through eq. 4. As a consequence, *Lc* will follow, at all IM levels, a probability density function *[Lc]*. The variability on *Lc* can be visualized by assessing expected losses at different levels of confidence, for all IM levels, that is:

$$Lc^{(x)}(im) \leftarrow p(\le Lc^{(x)} \mid im; [\vec{\pi}]) = x \qquad (12)$$

and the best guess estimation of the loss curve Lc^{*} can be obtained averaging over all possible π_i:

$$Lc^{*}(im) = \int_{\pi} Lc(im;\vec{\pi})d[\vec{\pi}]d\pi_{i} \approx \sum_{i} Lc^{(x_i)}(x_i - x_{i-1}) \qquad (13)$$

The approximation is valid for an adequate selection of percentiles x_i (e.g., Choun and Elnashai 2010). On the other hand, the variability on $Lc^{(x)}$, at different confidence level x, represents the variation induced by IMV in loss assessments.

As for the loss curve Lc, BCM estimates an entire distribution also for the mean loss assessment $[l_m]$. Also in this case, at each level of $IM = im$, we can define the mean loss at different level of confidence $l_m^{(x)}$ as the quintiles of the distribution $[l_m]$ as in Eq. 12, and the best guess value l_m* as in eq. 13. As a consequence, this variability is transferred to *AEL* assessment. A proxy of the variability of *AEL* can be assessed obtained by assessing it with different levels of confidence on l_m, that is

$$AEL^{(x)} = \int_{IM} l_m^{(x)}(im)dh(im) \qquad (14)$$

and, again, the best guess *AEL*, indicated as *AEL**, can be obtained as

$$AEL^{*} = \int_{\pi} AEL(\vec{\pi})d[\vec{\pi}] \approx \sum AEL^{(x_i)}(x_i - x_{i-1}) \qquad (15)$$

As for the loss curves *Ls*, also in this case $AEL^{(x)}$, when plotted as a function of x, shows how likely is it that a given *AEL* results an underestimation of the true one, and thus it represents a prompt of the variability induced in *AEL* by IMV. Similar considerations can be extended to all possible risk indexes.

In summary, BCM allows us to assess both best guess values and confidence on the estimation of losses and risk, propagating the IMV on vulnerability to the final results of loss/risk assessments. Noteworthy, such estimates have a lower likelihood of being biased than single models' results, since they account for more information (e.g., Woo 1999). On the other hand, the assessment of confidence on best guess values is of major importance (e.g., Paté-Cornell 1996), since it enables meaningful comparisons among losses/risks in different areas, as well as different losses/risks in the same area, in a multi-risk perspective (e.g., Grüntal et al 2006).

It is worth noting that BCM strongly differs from Bayesian inference procedures for fragility assessment proposed in literature (e.g., Shinghal and Kiremidjian 1996; Straub and Der Kiureghian 2008; Koutsourelakis 2010), as well as from FCs evaluated at different confidence levels (without composite β-values, e.g., Kennedy and Ravindra 1984), where a distribution form is assumed (log-normal) and distributions' parameters are inferred in order to assess the 'best' curve for a given structure. On the opposite, BCM is targeted to produce more accurate loss/risk assessment by including IMV on FCs. Indeed, having different goals, BCM is not in alternative to such approaches, since they simply focus on different and complementary issues. This is clearly demonstrated by the fact that the results of one (or more) of these models may be input to BCM, by randomly drawing *N* probability assessments from the model $\{\vec{\pi}^{(M)}\}$, and use each sample as single estimation with weight $w_i = w/N$ in Eq. 5, where *w* represents the weight of the overall model *M*. Of course, if large dataset of pertinent past data are available (same structures, large range of IMs), all models should lead to the same results, since IMV would be negligible in this case. Such data would enable us also to discriminate among different FC models, rejecting the ones that cannot 'explain' them (probabilities too far away from observed frequencies). This results also in BCM, since the application of Eq. 11 (Bayes' rule) with a large dataset would lead to posterior distributions highly peaked (very small variability) on the observed frequencies (Gelman et al 1995).

3. APPLICABILITY OF BCM

The uncertainty modeled by BCM essentially corresponds to the practical impossibility, common in many applications, to unequivocally select one specific FC model for a given structure (or typology of structures), because of the lack of background for one specific selection and the lack of resources to produce structure specific FCs (one for each element in the analyzed area).The presented procedure may be applied virtually to most of the sources of IMV, that is, whenever the selection of one specific FC model is highly disputable, for example:

1. FC models developed for similar configurations, but with different procedures that imply significantly different results. For example, FCs obtained by different statistical procedures of the same empirical or numerical damage data, FCs obtained by different numerical procedures (e.g. dynamic or quasi-static analysis) or slightly different characteristics of the same structure, or FCs referred to the same structural class which were derived based on analytical or empirical procedures.

2. Lack of structure-specific FC models, leading to select nonspecific or generalized FC models from literature (e.g., FCs developed in different areas, with different construction practices)
3. FC models developed for slightly different input IMs, among which it is difficult to distinguish in long-term aggregated hazard assessments (e.g., different incidence angle of seismic waves for bridges)
4. Rough description of structures in the application area, leading to difficulty in classifying them into a well-defined taxonomy or, in the opposite, a rough taxonomy leading to broad classes (e.g., different number of floors in a generic typology of buildings)

Note that cases 1 and 2 are somehow different from cases 3 and 4. Indeed, in cases 1 and 2, one 'true' model equal for all elements is expected to exist, and thus π_i should be sampled at once for all identical elements. On the opposite, in cases 3 and 4, the variability is expected within the target stock, and the 'true' FC is expected to be different from element to element. Consequently, the probability π_i should be sampled independently for each element.

In *section 2*, we discussed that the choice of a Dirichlet distribution implies a rather soft transition in the set of applicable FCs, implying a limitation on the definition of broad/mixed typologies in the taxonomy. This limitation applies for the IMV described in cases 3 and 4, above, since only there, 'different' typologies are mixed up. For example, one typology can mix up structures with different number of floors, but cannot mix up masonry and RC structures, or RC structures designed with or without seismic code.

4- CASE STUDY: LOSS/RISK ASSESSMENT FOR TUNNELS

To show the applicability of the BCM, a case study is considered. It has the goal of showing in details how BCM models IMV and it propagates this uncertainty into loss/risk assessments. To control all the parameters, we select a relatively simple application, that is, two input FC models and no past data. This configuration permits a simpler check of all steps, but any more complicated application do not introduce further either technical, or theoretical issues. This application is related to the IMV described in case 1 in *section 3*, that is, FCs obtained by different statistical approaches: due to the lack of adequate pertinent past data, these two approaches and the derived FCs can be considered to be equally applicable.

The final goal of this application is to assess the expected seismic losses and risk for a segment of bored tunnel (metro line) with a length of 1 km. Such a segment is assumed to be composed by 10 elements with a length of 100 m, each one of them laying in a specific soil type, as shown in **Fig. 2A**. The length of such segments is set so that the occurrence of damages in each element can be reasonably considered independent. The RC (repair cost) for each segment is set to 0.5 million euro, while the value CDF_i (cost damage factor) for each damage state is reported in Table 1, col. 8, based on the repair model that is proposed by Werner et al (2006) for drilled tunnels in California. Such assumptions and values are indicative, but they are realistic for a preliminary application. For the application area, we consider the hazard curve in **Fig. 2B**, which is a reasonable hazard for the city of Thessaloniki, Greece (Pitilakis et al 2007). To concentrate on the effects of IMV in vulnerability assessment, we assume the hazard perfectly known, i.e., not affected by epistemic uncertainty.

Two different procedures to develop FCs for shallow tunnels in alluvial are then considered, based on the same modeling procedure. The vulnerability assessment is based on a quasi-static numerical analysis (Argyroudis and Pitilakis 2012), and the dataset of damages produced by this model are then used to estimate two different sets of log-normal FCs, through two quite common approaches, that is, linear regression method (M1, *appendix A*) and maximum likelihood method

(M2, *appendix A*). Such FC models represent a set of two equally acceptable procedures to derive FCs, and both could be independently selected to perform loss/risk assessments. Both procedures are repeated for two different tunnel typologies, differentiated by the soil conditions in which tunnels are built, i.e., soil C and D according to Eurocode 8 classification. All the obtained FC models (M1 and M2, for both typologies) consider Peak Ground Acceleration (PGA) as IM, and use 3 damage states (minor, moderate, and extensive-to-complete). The parameters are reported in **Table 1**.

The FCs of all models are reported in **Fig. 3**. We can note that M1 and M2 provide quite different results, in both soils. Given a PGA value, the punctual probabilities of the damage states (π_i, Eq. 1), as assessed by such models, are quite unlike, for both soils C and D. In *sections 4.1* and *4.2*, we will show that such differences lead to significantly different loss/risk estimations.

The results of these–models are used to analyze the capability of BCM to combine and propagate IMV in loss/risk assessments. The analysis is divided in two parts. In *section 4.1*, we apply the BCM to one specific segment of tunnel built in soil C, in order to show how IMV propagates for one element and how different choices influence the results. In *section 4.2*, the preferred BCM model is applied to the schematic tunnel (metro line), analyzing the effect of IMV on the loss/risk assessments in a larger area, with different soil characterizations and with more than one element at risk.

4.1- ONE ELEMENT: SINGLE TUNNEL ELEMENT IN SOIL C

We first consider one single tunnel element in soil C. In **Fig. 4**, we report the loss/risk assessment results for each single model. In panels A1, we report the results of the loss assessment for one specific scenario, in this case set to PGA = 0.6 g, in terms of the loss curve *Lc*, as assessed by the input models M1 and M2. The difference between the expected losses is significant, and it results in quite different probability estimations, being M1 results significantly larger than the corresponding values for M2. It is important to note that these considerations are not a specific characteristic of the selected scenario: in panel A2, $l_m(PGA)$ for all PGA values are reported as assessed by both models. In Fig. 4, panel A3, we report the risk index *AEL* for both models, considering the hazard curve in Fig. 2B.

To model IMV among such models, we first have to set their (subjective) credibility. To do so, we consider that models M1 and M2 have equal credibility, since based on the same data and on equally credible statistical procedures. Therefore, our best guess weighting scheme is $w_1=0.5$, $w_2=0.5$. The sensitivity in this choice is then tested.

In **Fig. 5**, panels A1 to A3, we report the results of the loss/risk assessments for the best guess weighting scheme. In particular, we report best guess values and confidence intervals for all the assessments reported in Fig. 4, that is *Lc*, l_m and *AEL*. $AEL^{(x)}$ as obtained by model BCM is plotted as a function of *x*, indicating the confidence at which the true unknown *AEL* value is smaller than the various *AEL* values. For comparison, the punctual losses and risk index evaluated by M1 and M2, and best guess for BCM, are reported. Noteworthy, this variability in both loss and risk assessments cannot be dealt by variations of the β-value (e.g., Ferson and Ginzburg 1996) since, whatever β-value is used, any single choice provides only punctual probabilities (Eq. 1) and does not model the variability on such probabilities (eq. 1), and consequently cannot propagate it in risk/loss assessments (eqs. 2, 3 and 4).

In **Fig. 6**, we report the same results as above, with other weighting schemes. In particular, we select three further weighting schemes: 0.7 and 0.3; 0.9 and 0.1 and 0.1 and 0.9. The results are essentially the same, but here the distributions tend to move toward the model with greater weight. However, it is important to note that, also in this case, both mean values and confidence intervals still preserve memory of the less weighted model, and this memory tends to decrease for increasing difference on weights of models.

4.2- MANY ELEMENTS AND TYPOLOGIES: SEGMENT OF METROLINE IN SOILS C & D

In this application, one seismic scenario consists of a PGA value for each one of the segments and, for simplicity, (i) all the sites within the same soil type are assumed with equal PGA, and (ii) the PGA in soil type D is assumed equal to 1.3 times the PGA in soil type C. With these simplifications, the seismic scenario is completely defined by the selection of one PGA value for soil C. As exemplificative scenario, we select again a PGA value in soil C of 0.60 g. Also the comparison between BCM and 'standard' procedures results more complicated. Indeed, we have two models (M1 and M2) related to two typologies of tunnel (built in soil C and D), thus we must consider 4 possible combinations for standard FCs: *M1* in Soil C and *M1* in Soil D is indicated as M11; *M1* in soil C and *M2* in soil D as M12, and so on.

In **Fig. 7**, we report the same results that we reported in *section 4.*2, obtained by the best guess BCM model (equal weights) and compared with M1 and M2 results. In particular, the loss curve *Lc* for a scenario of $PGA = 0.6\ g$ (panel A1), the mean loss for all PGA levels (panel A2) and the risk index AEL (panel A3) are reported. Interestingly, the IMV is only very slightly reduced by staking a larger number of elements, and confidence intervals well describe the variability among the four possible combinations. In addition, in panel B, we report the distribution of losses for a scenario of $PGA = 0.6\ g$.

Note that, to produce these results, an unknown unique 'true' model for each tunnel typology (soil) is assumed to exist, since all the elements of the same type are assumed identical and BCM variability represents alternative models for such a typology. In practice, this means that the distribution $[\pi_i]$, at each run of the model, it is sampled only once for all identical elements (i.e., in the same soil). As discussed in *section 3*, this is not always the case for all types of IMV.

To show the potentiality of BCM, we consider a further application, adding a third input FC model. As third FC model we consider the one proposed by ALA (2001) for alluvial (all soil types) tunnels with good construction. As first assumption, definition of minor and moderate damage states of M1/M2 is assumed equal to the one of M3. Since M3 does not include extensive-to-complete damages, this damage state is assumed not possible ($\pi_3^{(M3)} \approx 0$, *i.e.*, $\pi_3^{(M3)} = \pi_3^{(M2)} \cdot 10^{-5}$, for numerical reasons). Note that this addition implies an abrupt increase on the possible combinations (i.e. eight: M111, M112, M121, M122, M211, M212, M221 and M222). On the contrary, with BCM this addition implies only the setting of a further weighting factor. In **Fig. 8**, we report the same results of Fig. 7, using as input the three models with weighting factors $w_1=0.45$, $w_2=0.45$, and $w_3=0.10$. As expected, the distributions of losses and risk again well represent the input variability, and large tail toward smaller values of loss is present, since M3 estimations forecast significantly smaller losses.

4.3- DISCUSSION OF RESULTS

The results in Figs. 4 to 8 clearly show that the expected risk/losses are essentially 'fragility model'-dependent. The loss curves *Lc* for a given scenario, as well as mean losses l_m and the risk index *AEL* show significant differences, when M1 and M2 are applied. For example, in Fig. 4, it is shown that, for one tunnel segment built in Soil C, the estimations of M2 are systematically greater than the M1 ones. Such differences are even more evident when the effects are stacked over a larger set of elements (Fig. 7). It is also evident that such differences are quite unreasonable, considering that M1 and M2 can be considered equally acceptable, but their results are highly incompatible. Note, for example, that the mean expected loss for M1, combination M11, results in the tail of the expected loss distribution of M2, combination M22 (see Fig. 7B). Such differences lead to the conclusion that at least one of the models M1 or M2 is significantly biased.

This apparent paradox is related to a lack in uncertainty evaluation, even though all the principal sources of uncertainty (demand, capacity and damage state definition) have been formally introduced in both M1 and M2 FCs. Interestingly, this uncertainty cannot be modeled simply by increasing the β-value, which formally describes only the error in the position of the medians m_i, since whatever parametrical choice is adopted, any single fragility cannot neither model nor propagate IMV in loss/risk assessments.

On the contrary, BCM allows us to describe and propagate the uncertainty related to the impossibility to choose among single FC models (IMV). As expected, BCM distributions generally include all the values that either M1 and/or M2 produce, when equal applicability is assumed (Figs. 5 and 7). The losses forecasted by BCM cover the whole variability previewed by both input models M1 and M2 (Fig. 7B), and do not simply average them. Of course, the frequency of each single loss depends on how likely it is for all input models (Figs. 7 and 8).

The most likely area for expected losses and risk lays between the ones of the input models, i.e., in the area in which all models provide likely values. In our opinion, this is highly reasonable, given the assumption that M1 and M2 are equally acceptable and likely (equal weights). In the Soil C case (Fig. 5), the best guess estimates of BCM are close to the mean of input models. On the opposite, the non-compatibility between FCs for soil D, moderate and extensive-to-complete damages (Fig. 3), leads to best guess estimates slightly shifted toward smaller values (Fig. 7A).

If equal acceptability is not assumed, BCM adapts its behavior to this information, provided by the models' weights. In practice, BCM's loss/risk estimates move toward the most likely model's ones, preserving in its variability memory of the less likely model's ones (Fig. 6, upper panels). This variability tends to disappear only when the difference in acceptability is rather high (Fig. 6, lower panel).

Noteworthy, the addition of further input models does not imply any supplementary neither theoretical nor computational effort, as it is demonstrated by the application in Fig. 8, where a third model is considered.

5. FINAL REMARKS

The choice of one single set of FCs is often largely subjective, and different fragility may lead to significantly different expected loss and risk assessments. Hence, this uncertainty, of epistemic type, strongly increases the possibility of biased loss/risk estimations and consequently weakens their practical usability (Fournier d'Albe 1979; Paté-Cornell 1996).

We have developed a Bayesian methodology (Section 2) that allows us to account and propagate into loss/risk assessments a large spectrum of uncertainties related to the application of

FC models in vulnerability assessments (Section 3), essentially linked to scarce knowledge about the target stock. This epistemic uncertainty is relative to the application of FC in loss/risk assessments, more than to FCs by themselves, and it can be modeled accounting for the Inter-Model Variability (IMV). This kind of variability cannot be treated by the standard uncertainty treatment (e.g., Ferson and Ginzburg 1996) and it is usually neglected. On the other hand, we have shown that the Bayesian Composition Model (BCM), explicitly modeling the variability in probability, appropriately and efficiently describes IMV by combining the results of different standard fragility analyses and pertinent data, explicitly quantifying the influence of such an uncertainty in loss/risk assessments. BCM considers, eventually with different weights, many inhomogeneous sources of information, independently from their formulation and their statistical representation. In addition, BCM does not involve a dramatic increase of the computation effort.

The quantification of IMV in loss/risk assessments is important, since it (i) significantly reduces the likelihood of biased cost/risk assessments, increasing their usability in practical applications, and (ii) explicitly assesses the confidence on loss/risk results. This permits meaningful and robust comparisons among losses/risks in different areas, as well as different losses/risks in the same area, in a multi-risk perspective (e.g., Grüntal et al 2006). Indeed, risk hierarchization is ultimately one of the most important goals of any loss/risk assessment. The probability that BCM results are biased is lower than the ones based on single models, since it is based on more information (e.g., Woo 1999). Obviously, to achieve the goal of an unbiased estimation in a real application, any pertinent information should be included, opportunely weighting models according to their different reliability/applicability.

AKNOWLEDGMENTS

The work described in this paper was supported by the projects ByMuR “Bayesian Multi-Risk Assessment: A case study for Natural Risks in the city of Naples” ("Quantificazione del Multi-Rischio con approccio Bayesiano: un caso studio per i rischi naturali della città di Napoli", http://bymur.bo.ingv.it/) in the “Futuro in Ricerca 2008” call of MIUR (Italian Ministry of Education, University and Research), and SYNER-G “Systemic seismic vulnerability and risk analysis for buildings, lifeline networks and infrastructures safety gain” under Grant Agreement No. 244061 in the 7th Framework Program of the European Commission. This support is gratefully acknowledged. We finally thanks the anonymous reviewers for the useful and appropriate comments, which largely helped in strengthening the results of this paper.

REFERENCES

Ahmed, E.S., Volodin, A.I., Hussein, A.A. (2005). “Robust Weighted Likelihood Estimation of Exponential Parameters”, IEEE Transactions on Reliability, 54(3): 289-395.

Akkar, S., Sucuoglu, H., Yakut, A. (2005). “Displacement-based fragility functions for low- and mid-rise ordinary concrete buildings”. Earthquake Spectra 21(4): 901-927.

ALA, American Lifelines Alliance (2001). “Seismic Fragility Formulations for Water Systems”. Part 1 – Guideline. ASCE-FEMA, 104 pp.

Argyroudis, S., Pitilakis, K. (2012). “Seismic fragility curves of shallow tunnels in alluvial deposits”. Journal of Soil Dynamics and Earthquake Engineering, (35): 1–12..

ATC-13, Applied Technology Council Report (1985). “Earthquake damage evaluation data for California.” Palo Alto.

ATC-25, Applied Technology Council Report (1991). “Seismic vulnerability and impact of disruption of lifelines in the conterminous United States”. Redwood City.

Azevedo, J., Guerreiro, L., Bento, R., Lopes, M., Proença, J. (2010). “Seismic vulnerability of lifelines in the greater Lisbon area”."Bulletin of Earthquake Engineering, 8:157–180.

Bardet, J.P., Ichii, K., Lin, C.H., (2000). “EERA: a computer program for equivalent-linear earthquake site response analyses of layered soil deposits”, University of Southern California, Department of Civil Engineering, 40 p.

Basöz, N.I., Kiremidjian, A.S. (1998). “Evaluation of bridge damage data from the Loma Prieta and Northridge, California earthquake”. Technical Report MCEER-98-0004, State University of New York, Buffalo.

Basöz, N.I., Kiremidjian A.S., King S.A., Law K.H. (1999). “Statistical analysis of bridge damage data from the 1994 Northridge, CA, Earthquake”, Earthquake Spectra 15(1): 25-54.

Bernardo, J.M. and Smith, A.F.M. (1994). “Bayesian Theory”, Chichester: Wiley.

Bhargava, K., Ghosh, A.K., Agrawal, M.K., Patnaik, R., Ramanujam, S., Kushwaha, H.S. (2002), “Evaluation of seismic fragility of structures – a case study”, Nucl. Eng. And Des 212: 253-272.

Bommer, J.J., Spence, R., Erdik, M., Tabuchi, S., Aydinoglu, N., Booth, E., del Re, D., Peterken, O., (2002), “Development of an earthquake loss model for Turkish catastrophe insurance”, J. of Seism. 6: 431-446.

Bommer, J.J., and Scherbaum, F. (2008). “”The Use and Misuse of Logic Trees in Probabilistic Seismic Hazard Analysis”, Earthquake Engineering Practice, 24(4): 997-1009.

Choun, Y.S., and Elnashai, A.S. (2010). “A simplified framework for probabilistic earthquake loss estimation”, Prob. Eng. Mec. 25: 355-364.

Coppersmith, K.J., and Youngs, R.R. (1986). “Capturing uncertainty in probabilistic seismic hazard assessments within intraplate environments”, Proceedings of the Third U.S. National Conference on Earthquake Engineering, August 24-28, 1986, Charleston, SC, Earthquake Engineering Research Institute, Berkeley, CA, Vol I, 301-312.

Cornell, C.A., Krawinkler, H. (2000). “Progress and challenges in seismic performance assessment”, PEER Center News, 3(2).

Cornell, C. A., and Merz, H. A. (1975). “Seismic risk analysis of Boston”, Journal of the Structural Engineering Division, ASCE, Vol 101, pp 2027-2043..

Douglas, J. (2007). “Physical vulnerability modelling in natural hazard risk assessment”, Nat. Hazards Earth Syst. Sci., 7, 283-288,

Dueñas-Osorio, L., Craig, J.I., Goodno, B.J. (2007). “Seismic response of critical interdependent networks”, Earthquake Eng. Struct. Dyn., 36(2): 285-306.

Electric Power Research Institute (1987). “Seismic hazard methodology for the central and Eastern United States” Volume 1, Part 1, Theory, and Part 2, Methodology, Prepared for Seismicity Owners Group under Research Projects P101-38, -45, -46, and 2256-14, EPRI Report NP-4726, Revised, February, Palo Alto, CA.

Ellingwood, B., Kinali, K. (2009). “Quantifying and communicating uncertainty in seismic risk assessment”. Structural Safety, 31(2): 179–187.

EC8 (2004). “Eurocode 8: Design of structures for earthquake resistance”. European Committee for Standardization: Brussels, Belgium, The European Standard EN 1998-1.

Fournier d’Albe, E.M. (1979). “Objectives of volcanic monitoring and prediction”, J Geol Soc London 54: 57-67.

FEMA, Federal Emergency Management Agency (2008). “Estimated Annualized Earthquake Losses for the United States”, FEMA 366.

Ferson, S., Ginzburg, L.R. (1996). “Different methods are needed to propagate ignorance and variability”, Reliability Engineering and System Safety 54: 133-144.

Gruppo di Lavoro MPS (2004). R”edazione della mappa di pericolosità sismica prevista dall’Ordinaza PCM 3274 del 20 Marzo 2003”. Rapporto Conclusivo per il Dipartimento della Protezione Civile, INGV, Milano – Roma, aprile 2004, 65 pp + 5 Appendici.

Gelman, A., Carlin, J., Stern, H., Rubin, D. (1995). “Bayesian data analysis”, Chapman and Hall/CRC.

Giner, J.J., Molina, S., Delgado, J., Jauregui (2002). “Mixing methodologies in seismic hazard assessment via a logic tree procedure: An application for eastern Spain”, Natural Hazards 25, 59–81.

Grezio, A., Marzocchi, W., Sandri, L., Gasparini, P., (2010). “A Bayesian procedure for probabilistic tsunami hazard assessment”, Natural Hazards 53, 159-174, DOI 10.1007/s11069-009-9418-8.

Grüntal, G., Thieken, A.H., Shwarz, J., Radtke, K.S., Smolka, A., MMerz, B. (2006). "Comparative Risk Assessments fo the City of Cologne – Storms, Foods, Earthquakes", Nat. Haz. 28: 21-44.

Hofer, E. (1996). "When to separate uncertainties and when not to separate", Reliability Engineering and System Safety 54: 113-118.

Kalbfleisch, J.G. (1977). "Probability and Statistical Inference", voll.2, pp. 658, Springer, Berlin

Kappos, A.J. (1997). "Seismic damage indices for R/C buildings: Evaluation of concepts and procedures", Progress in Structural Engineering and Materials, 1(1): 78-87.

Kappos, A., Panagopoulos, G., Panagiotopoulos, Ch., Penelis, G. (2006). "A hybrid method for the vulnerability assessment of R/C and URM buildings". Bull. Earthquake Engineering, 4: 391–413.

Kappos, A.J., Panagopoulos, G., Penelis, G. (2008). "Development of a seismic damage and loss scenario for contemporary and historical buildings in Thessaloniki, Greece", Soil Dynamics and Earthquake Engineering, 28(10-11): 836-850.

Karim, K.R., Yamazaki, F. (2001). "A simplified method of constructing fragility curves for highway bridges", Earthquake Engineering and Structural Dynamics, 32(10): 1603-1626.

Kennedy, R.P., and Ravindra, M.K. (1984). "Seismic fragilities for nuclear power plant risk studies", Nucl. Eng. Des., 79(1): 47-68.

Kim, S., Feng, M.Q. (2003). "Fragility analysis of bridges under ground motion with spatial variation." International Journal of Non- Linear Mechanics, 38: 705-721.

Koutsourelakis, P. (2010). "Assessing structural vulnerability against earthquakes using multi-dimensional fragility surfaces: A Bayesian framework", Probabilistic Engineering Mechanics, 25, 49–60, doi:10.1016/j.probengmech.2009.05.005.

Kulkarni, R.B., Youngs, R.R., and Coppersmith, K.J. (1984). "Assessment of confidence intervals for results of seismic hazard analysis", Proceedings of the Eighth World Conference on Earthquake Engineering, San Francisco, CA, July 21-28, 1984, Vol 1, pp 263-270.

Lindley, D.V. (1965). "Introduction to Probability and Statistics from a Bayesian Viewpoint". Cambridge University Press, Cambridge, UK.

Mackie, K., Stojadinovic, B. (2003). "Seismic Demands for Performance-Based Design of Bridges". PEER Report 2003/16. Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Marzocchi, W., Sandri, L, Selva, J. (2008). "BET_EF: a probabilistic tool for long- and short-term eruption forecasting", Bull. Volcanol., 70, 623-632, doi:10.1007/s00445-007-0157-y

Marzocchi, W., Sandri, L., Selva, J., (2010). "BET_VH: a probabilistic tool for long-term volcanic hazard assessment", Bull. Volcanol., 72, 705-716, DOI 10.1007/s00445-010-0357-8.

Maruyama Y., Yamazaki F., Mizuno K., Tsuchiya Y., Yogai H. (2010). "Fragility curves for expressway embankments based on damage datasets after recent earthquakes in Japan". Soil Dynamics and Earthquake Engineering, 30: 1158-1167.

McGuire, R.K. (1977). "Effects of uncertainty in seismicity on estimates of seismic hazard for the east coast of the United States", Bull Seism. Soc. Am. 67,827-848.

McGuire, R.K., Shedlock, K.M. (1981). "Statistical uncertainties in seismic hazard evaluations in the United States", Bull Seism. Soc. Am., 71, 1287-1308.

Moschonas, I., Kappos, A., Panetsos, P., Papadopoulos, V., Makarios, T., Thanopoulos, P. (2009). "Seismic fragility curves for Greek bridges: Methodology and Case Studies", Bulletin of Earthquake Engineering, 7(2): 439-468.

Mosimann,. J.E. (1962). "On the compound multinomial distribution, the multivariate fl-distribution and correlation among proportions". Biometrika 49: 65-82.

National Research Council (1988). "Probabilistic seismic hazard analysis", National Academy Press, Washington, DC.

National Institute of Building Sciences (NIBS) (2004). "HAZUS-MH: Users's Manual and Technical Manuals". Report prepared for the Federal Emergency Management Agency, Washington, D.C.

Nielson, B.G., DesRoches, R (2007). "Seismic fragility methodology for highway bridges using a component level approach". Earthquake Engineering and Structural Dymamics, 36: 823–839.

Paté-Cornell, M.E. (1996). "Uncertainties in risk analysis: Six levels of treatment", Reliability Engineering and System Safety 54: 95-111.

Pinto, P.E. (editor) (2007). "Probabilistic methods for seismic assessment of existing structures". LESSLOSS Report No. 2007/06, Istituto Universitario di Studi Superiori di Pavia, IUSS Press, ISBN: 978-88-6198-010-5.

Pitilakis, K., Alexoudi, A., Argyroudis, S., Monge, O., Martin, C. (2006). "Chapter 9: Vulnerability and risk assessment of lifelines". In Assessing and Managing Earthquake Risk, Geo-scientific and Engineering Knowledge for Earthquake Risk Mitigation: developments, tools, techniques, Eds. X. Goula, C. S. Oliveira, A. Roca, Springer, ISBN 1-4020-3524-1. pp. 185-211.

Pitilakis, K, Cultrera, G., Margaris, B., Ameri, G., Anastasiadis, A., Franceschina, G., Koutrakis, S. (2007). "Thessaloniki seismic hazard assessment: probabilistic and deterministic approach for rock site conditions", 4th International Conference on Earthquake Geotechnical Engineering, June 2007, Paper No 1701.

Pitilakis, K., Anastasiadis, A., Kakderi, K., Alexoudi, M., Argyroudis, S. (2010). "The role of soil and site conditions in the vulnerability and risk assessment of lifelines and infrastructures. The Case of Thessaloniki (Greece)". 5th

International Conference on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics, San Diego, California, May, 24-29.

Plaxis 2D (1998), Reference Manual, version 7.

Rossetto, T., Elnashai, A. (2003). "Derivation of vulnerability functions for European-type RC structures based on observational data". Engineering Structures, 25: 1241–1263.

Rossetto, T., Elnashai, A. (2005). "A new analytical procedure for the derivation of displacement-based vulnerability curves for populations of RC structures". Engineering Structures, 27(3): 397-409.

Salmon, M., Wang, J., Jones, D., Wu, Ch., (2003), "Fragility formulations for the BART system", in: Proceedings of the 6th U.S. Conference on Lifeline Earthquake Engineering, TCLEE, Long Beach.

Saxena, V., Deodatis, G., Shinozuka, M., Feng, M.Q. (2000). "Development of fragility curves for multi-span reinforced concrete bridges." Proceedings of the International Conference on Monte Carlo Simulation, Principality of Monaco, Balkema Publishers.

Schmidt, J., Marcham, I., Reese, S., King, A., Bell, R., Henderson, R., Smart, G., Cousins, J., Smith, W., Heron, D., (2011). "Quantitative multi-risk analysis for natural hazards: a framework for multi-risk modeling", Natural Hazards, DOI:10.1007/s11069-011-9721-z.

Selva, J., A. Costa, W. Marzocchi, L. Sandri, (2010). "BET_VH: Exploring the influence of natural uncertainties on long-term hazard from tephra fallout at Campi Flegrei (Italy)", Bull. Volcanol., DOI:10.1007/s00445-010-0358-7.

Selva, J., Orsi, G., Di Vito, M.A., Marzocchi, W., Sandri, L. (2012). "Probability hazard map for future vent opening at the Campi Flegrei caldera, Italy". Bull. Volcanol. 74, 497-510.

SHARE (2009-2012). Seismic Hazard Harmonization in Europe, European Research Project funded by FP7, http://www.share-eu.org/

Shinghal, A., Kiremidjian, A.S. (1996). "Bayesian updating of fragilities with application to RC frames", J. of Structural Engineering 122 (12): 1459-1467.

Shinozuka, M., Feng, M.Q., Lee, J., Naganuma, T. (2000). "Statistical analysis of fragility curves." J. Engrg. Mech. ASCE 126(12): 1224-1231.

Shinozuka, M., Feng, M.Q., Kim, H.K., Uzawa, T., Ueda, T. (2003). "Statistical analysis of fragility curves", Technical Report MCEER-03-0002, State University of New York, Buffalo.

Spence, R.J.S., Kelman, I., Baxter, P.J., Zuccaro, G., Petrazzuoli, S. (2005). "Residential building and occupant vulnerability to tephra fall", Nat. Hazards Earth Syst. Sci., 5(4): 477-494.

Stergiou, E., Kiremidjian, A.S. (2006). "Treatment of uncertainties in seismic risk analysis of transportation systems". Technical Report No. 154, John A. Blume Earthquake Engineering Center, Civil Engineering Department, Stanford University, Stanford, CA.

Straub, D., Der Kiureghian, A, (2008). "Improved seismic fragility modeling from empirical data", J. Structural Safety 30, 320-336.

SYNER-G, (2009-2012). "Systemic seismic vulnerability and risk analysis for buildings, lifeline networks and infrastructures safety gain", European Research Project funded by FP7, http://www.syner-g.eu.

Wang, X., van Eeden, C., Zidek, J.V. (2004). "Asymptotic properties of macimum weighted likelihood estimators", J. Stat. Plan Inference, 119: 37-54.

Wen, Y.K., Ellingwood, B.R., Veneziano, D., Bracci, J. (2003). "Uncertainty modeling in earthquake engineering", MAE Center Project FD-2 Report, Illinois.

Werner, S.D., Taylor, C.E., Cho, S., Lavoie, J-P., Huyck, C., Eitzel, C., Chung, H., Eguchi, R.T. (2006). "REDARS 2: methodology and software for seismic risk analysis of highway systems" , Technical report, MCEER-06-SP08.

Winkler, R.L. (1996). "Uncertainty in probabilistic risk assessment", Reliability Engineering and System Safety 54: 127-132

Woo, G. (1999). "The mathematics of Natural Catastrophes", Imperial College Press, London

Figure 1: The Inter-Model Variability (IMV) is assessed on the punctual probability of the i-th damage state, i.e., π_i, is represented through the probability density function $[\pi_i]$. In this case, as IM we used the Peak Ground Acceleration (PGA).

Figure 2. *Panel A:* Hypothetical metro line with circular cross section and total length 1000m, passing through alluvial deposits of soil type C and D (EC8), divided in ten equal segments of 100m. *Panel B:* Hazard curve for soil C.

Figure 3. Set of fragility curves for circular tunnel following model *M1* and *M2*, (a) for soil C and (b) for soil D. The considered damage states are minor, moderate and extensive-to-complete.

Figure 4. M1 and M2 loss/risk assessments. In particular, we report: in panel A1 the loss curve *Lc* for a scenario *PGA=0.6 g*; in panel A2 the mean loss curve l_m; in panel A3, the risk index *AEL*.

Figure 5. BCM loss/risk assessments with the best guess BCM model ($w_1=w_2=0.5$), compared with input models M1 and M2. In particular, we report in panel A1 the loss curve *Lc* for a scenario *PGA=0.6 g* and in panel A2 the mean loss curve l_m. In here, the estimates for models M1 and M2 are compared with BCM's best guess (*Lc** and l_m*) and confidence intervals. In panel A3, the risk index *AEL* for models M1 and M2 are compared with BCM's best guess *AEL** and $AEL^{(x)}$.

Figure 6: BCM loss/risk assessments, compared with M1 and M2 assessments, for one tunnel element in soil C, with different weighting schemes: BCM with $w_1= 0.7$, $w_2 = 0.3$ is reported in panels A; BCM with $w_1= 0.9$, $w_2 = 0.1$ in panels B; BCM with $w_1= 0.1$, $w_2 = 0.9$ in panels C. In all panels, as in Fig. 5, we report the results for (i) the loss curve *Lc* for a scenario *PGA=0.6 g*, (ii) the mean loss curve l_m, and (iii) the risk index *AEL*.

Figure 7: Results of the loss/risk assessment for the metro line in Fig. 2. We report, as in Figs. 5 and 6, the results for (i) the loss curve *Lc* for a scenario PGA*=0.6 g* in panel A1; the mean loss l_m as function of PGA in panel A2; the risk index *AEL* in panel A3. In panel B, it is reported the distribution of losses for the same scenario as for *Lc* (PGA*=0.6 g*). For comparison, we report as vertical lines the average losses *lm* for M1 and M2 (configurations M11 and M22) for the scenario.

Figure 8: Same as Fig. 7, but with the addition of M3. The weighting scheme is $w_1=w_2=0.45$, $w_3=0.10$.

Table 1. Definition of damage states for the development of analytical fragility curves for tunnels and estimated parameters of the fragility curves based on different methods

Damage State (ds_i)	**Range of damage index (DI)**	**Central value of DI**	**M1 - SOIL C**		**M2 - SOIL C**		**M1 - SOIL D**		**M2 - SOIL D**		**CDF**
			m_i **(g)**	β	m_i **(g)**	β	m_i **(g)**	β	m_i **(g)**	β	
0. No damage	$M/M_{Rd} \leq 1.0$	-	-	-	-	-	-	-	-	-	0
1. Minor	$1.0 < M/M_{Rd} \leq 1.5$	1.25	0.55		0.52		0.47		0.41		0.10
2. Moderate	$1.5 < M/M_{Rd} \leq 2.5$	2.00	0.82	0.70	0.80	0.55	0.66	0.75	0.82	0.60	0.25
3. Extensive-to-Complete	$2.5 < M/M_{Rd} \leq 3.5$	3.00	1.05		1.39		0.83		1.91		0.75

APPENDIX A: *VULNERABILITY ASSESSMENT THROUGH FRAGILITY MODELS*

Recently, new analytical fragility curves for shallow metro tunnels have been proposed based on numerical simulation, considering both structural parameters, local soil conditions and variation of input ground motion (Argyroudis and Pitilakis, 2012). The quantification of the damage states is based on a damage index (DI) that is defined as the exceedance of strength capacity of the most critical sections of the tunnel (i.e. ratio of the developing moment (M) to the moment resistance (MRd) of the tunnel lining). The definition of damage states is then based on the range of damage index values (**Table 1, col. 1-3**). From the evaluated damage index, as a function of the *PGA* at the ground surface, the set of fragility curves relative to a discrete number of Damage States can be derived. Three different damage states are considered due to ground shaking: minor, moderate and extensive-to-complete damage (d_1, d_2, and d_3 respectively). Fragility curves (FC) are usually represented as a two-parameter (median and log-standard deviation) lognormal cumulative distribution functions. The development of FCs requires the definition of 4 parameters, 3 medians m_i and 1 value of β, which are estimated in literature following different procedures.

Two procedures are adopted here: (i) a linear regression method (e.g. Nielson and DesRoches 2007; Pinto 2007), herein referred to as M1, and (ii) a maximum likelihood method (ML, e.g. Saxena et al. 2000; Shinozuka et al. 2000, 2003; Kim and Feng 2003; Straub and Der Kiureghian 2008), herein referred to as M2.

M1 has been recently published in Argyroudis and Pitilakis (2012). Such fragility functions are reported in **Table 1, col 4 5**, and plotted in **Figure 2 (light blue)** for the case of circular (bored) tunnel in soil type C and D. As regards M2, while ML is normally used starting from real data (Kalbfleish 1977), with the same philosophy it is here used with synthetic data produced by a model. In particular, as for M1, the starting database for M2 consists of the result of the coupled numerical analysis, i.e., the earthquake parameter (PGA) and the consequent damage index for the modeled tunnel (PGA_i, DI_i). By defining one threshold in *DI* for each damage state (t_1,t_2, and t_3), the data can be transformed as the result of a Bernoulli trial experiment, associating each PGA to the consequent expected damage state, i.e., (PGA_i, y_i), where y_i is equal to 1 or 0 depending on whether or not the tunnel section sustains the damage state, that is equal to 1 if it is observed the *i*-th damage state, 0 otherwise. To account for the uncertainty on damage state definition, for each starting datum (PGA_i, DI_i), a Monte Carlo simulation is performed, by producing N = 500 couples of (PGA_i, y_i) data, each one obtained by comparing out the observed value for the damage index (DI_i) with randomly sampled thresholds. The thresholds are sampled from uniform distributions in their confidence intervals (**Table 1, col. 2**). The fragility curves are assumed to be log-normally distributed, with different medians m_j and equal β-value. The best guess values for the parameters (m_i' and β') are obtained by numerically maximizing, as a function of m_j, and β, the likelihood function L. The obtained values (m' and β') account for the demand uncertainty, since different seismic records as input for the coupled numerical analysis are used, and the damage state definition uncertainty by randomly selecting the DI thresholds (Eq. *A.2*). Among the principal sources of uncertainties, only the capacity uncertainty is not yet considered. Thus, it is added to the results of the analysis as the square root of the sum of squares of β' and 0.3 (e.g., NIBS 2004).

The obtained β''-value, which includes also capacity uncertainty, is then put into the likelihood function that, this time, is a function of the medians m_j only. The best guess medians (m_i'') are obtained by numerically maximizing $ln(L')$ and, together with the total β''-value, represent the best guess parameters for the log-normal distribution. From this analysis, we obtain the final parameters for M2, as reported in **Table 1, col. 6-7** and plotted in **Figure 2 (dark blue)**.