Available online at www.eccomasproceedia.org
Eccomas Proceedia UNCECOMP (2025) 453-475

ECCOMAS
Proceedia

UNCECOMP 2025
6th ECCOMAS Thematic Conference on
Uncertainty Quantification in Computational Science and Engineering
M. Papadrakakis, V. Papadopoulos, G. Stefanou (eds.)
Rhodes Island, Greece, 15-18 June 2025

BAYESIAN ESTIMATION OF SEISMIC FRAGILITY CURVES BASED ON VARIATIONAL REFERENCE PRIORS USING NEURAL NETWORKS

Nils Baillie[1,2], Antoine Van Biesbroeck[1,2], Cyril Feau[2] and Clément Gauchy[3]

[1]CMAP, CNRS, École polytechnique, Institut Polytechnique de Paris,
91120 Palaiseau, France
nils.baillie@polytechnique.edu

[2] Université Paris-Saclay, CEA, Service d'Études Mécaniques et Thermiques,
91191 Gif-sur-Yvette, France
{nils.baillie, antoine.vanbiesbroeck, cyril.feau}@cea.fr

[3] Université Paris-Saclay, CEA, Service de Génie Logiciel pour la Simulation,
91191 Gif-sur-Yvette, France
clement.gauchy@cea.fr

Abstract. *Seismic fragility curves are key quantities in seismic probabilistic risk assessment studies. They express the probability of failure of a mechanical structure of interest conditional to a scalar measure derived from the ground motion and called intensity measure (IM). The evaluation of such curves using Monte Carlo methods with artificial seismic signals is common today. However, when they are based on complex high-fidelity numerical modeling or on results from shaking table experiments, the amount of data available is limited. In this context, the Bayesian approach can provide an efficient learning of the parameters that define the fragility curve, often expressed by a probit-lognormal model. Nevertheless, the choice of the prior can have an important influence on the posterior and the parameter estimation. To limit the added subjectivity from a priori information, one can use the framework of reference priors. In this work, we propose a variational inference approach to approximate the prior. It takes the form of the output of a neural network, optimized to solve the maximization problem that defines the reference priors. Our algorithm is flexible: it can retrieve reference priors when constraints are specified in the optimization problem to ensure the solution is proper, and it proposes a simple method to recover a relevant approximation of the parametric posterior distribution using Markov Chain Monte Carlo methods even if the density function of the parametric prior is not known in general. We validate our methodology on a case study taken from the nuclear industry. Our results show the usefulness of this approach in recovering the target distribution.*

Keywords: Reference priors, Variational inference, Neural networks, Fragility curves

ISSN:2623-3339
Peer-review under responsibility of the organizing committee of UNCECOMP 2025.
doi: 10.7712/120225.12370.21225

1 INTRODUCTION

Seismic fragility curves are a fundamental component of the seismic probabilistic risk assessment (SPRA) framework. Originally introduced in the 1980s for seismic risk assessment studies in the context of the nuclear industry [1–5], they have since become a key-tool in performance-based earthquake and engineering (PBEE) [5]. These curves express the probability of failure of a given mechanical equipment as a function of the seismic hazard. The latter being generally reduced to a single intensity measure (IM) that is derived from seismic ground motions, under the so-called "sufficien y assumption" [5, 6]. Common choices for IMs include peak ground acceleration (PGA), pseudo spectral acceleration (PSA), and peak ground velocity (PGV), among others.

Different data sources can be exploited to estimate seismic fragili y curves (e.g. experimental data [4] or analytical results given by numerical models and artificia seismic excitation [7–21]). In this work we address problems within which the observed data are (i) scarce, (ii) independent, and (iii) limited to binary outcomes. They take the form of independent realizations of the tuple $X = (z, a)$ where $z \in \{0, 1\}$ represents the outcome —failure or non-failure— and $a \in \mathcal{A} \subset (0, \infty)$ represents the IM. In such cases, fragility curves are commonly parameterized using the probit-lognormal model [7–9, 11–18, 21–24], which introduces a parameter θ to describe the probabilistic relationship between z and a.

Among the various methods that exist to estimate the parameter θ, the Bayesian approach has gained increasingly popularity in the literature [10, 16, 18, 21, 24–33]. The Bayesian framework offers several advantages over classical frequentist methods, particularly in avoiding the generation of irregular fragility curves —such as unit step functions— that may arise from limited data. The challenge of this approach, however, lies in selecting an appropriate prior distribution as it can significantl influenc the posterior distribution and, consequently, the resulting estimates.

A well-established tool for addressing the prior selection issue is the reference prior theory, firs introduced in [34]. Reference priors are designed to maximize the influenc of the observed data in Bayes ian inference by maximizing the mutual information between the prior and the posterior. This property makes them particularly suitable for fragility curve estimation when data are sparse, as they minimize the influenc of arbitrary subjective assumptions. A notable application of this approach in seismic fragility curve estimation is suggested in [31, 32] where a reference prior is specificall derived for this problem. Their implementation relies on a numerical approximation of the "complex" theoretical form of the prior, which can be cumbersome to evaluate.

In this work, we propose a novel methodology for estimating seismic fragility curves using a variational approximation of the reference prior (VA-RP). The VA-RP, as introduced in [35], parameterizes the reference prior as the output of a neural network, permitting efficien computation of *a priori* samples. The weights of the neural network are then trained to maximize the mutual information. In our study, the neural network architecture is specially tailored to the context of fragility curves estimations. Additionally, because a significan issue in the estimation of fragility curves with reference priors is the yielding of improper posteriors when the data are scarce, we propose the incorporation of constraints to the VA-RP. These problem-specifi constraints ensure that both the prior and the posterior are proper. This additional feature enhances the robustness of Bayesi an inference, making the method applicable even in cases with severely limited observational data, which is common in earthquake engineering.

After reviewing the reference prior theory in Section 2, we introduce the VA-RP methodol-

ogy in Section 3. The tailoring of the framework for fragility curve estimation is described in Section 4. Then, in Section 5, we evaluate our approach on a case study taken from the nuclear industry. Finally, concluding remarks are discussed in Section 6.

2 REFERENCE PRIORS THEORY

We consider the following situation: we observe i.i.d data samples $\mathbf{X} = (X_1, ..., X_N) \in \mathcal{X}^N$ with $\mathcal{X} \subset \mathbb{R}^d$. The likelihood function $L_N(\mathbf{X} \mid \theta) = \prod_{i=1}^N L(X_i \mid \theta)$ is known and $\theta \in \Theta \subset \mathbb{R}^q$ is the parameter to be inferred. For this purpose, θ is considered to be stochastic in the Bayesian framework, so that it admits a prior distribution that we denote by π. The marginal distribution is define as the one whose density is $p_{\pi,N}(\mathbf{X}) = \int_\Theta \pi(\theta) L_N(\mathbf{X} \mid \theta) d\theta$. In those settings, the posterior distribution is define as follows:

$$\pi(\theta \mid \mathbf{X}) = \frac{\pi(\theta) L_N(\mathbf{X} \mid \theta)}{p_{\pi,N}(\mathbf{X})}, \tag{1}$$

it is used to generate *a posteriori* estimates of θ.

The reference prior theory has been introduced in [34], and formalized further in [36]. The theory suggests to select the prior π as one that maximizes the quantity called mutual information and denoted by $I(\pi; L_N)$. It represents the amount of information that differs between the prior and the posterior. Thus, the definitio of the reference prior amounts to solve the following optimization problem:

$$\pi^* \in \arg\max_{\pi \in \mathcal{P}} I(\pi; L_N) \,, \tag{2}$$

where $\mathcal{P}$ is a class of admissible probability distributions. The mutual information itself is define by:

$$I(\pi; L_N) = \int_{\mathcal{X}^N} \mathrm{KL}(\pi(\cdot \mid \mathbf{X}) \,||\, \pi) p_{\pi,N}(\mathbf{X}) d\mathbf{X}, \tag{3}$$

where KL designates the Kullback-Leibler divergence. Therefore, π^* is a prior that maximizes the Kullback-Leibler divergence between itself and the corresponding posterior averaged by the marginal distribution of datasets. The Kullback-Leibler divergence between two probability measures of density p and q define on a generic set Ω writes:

$$\mathrm{KL}(p \,||\, q) = \int_\Omega \log\left(\frac{p(\omega)}{q(\omega)}\right) p(\omega) d\omega. \tag{4}$$

Thus, π^* is the prior that maximizes the influenc of the data on the posterior distribution, justifying its reference (or objective) properties. Using Fubini's theorem and Bayes' theorem, a more practical form for the mutual information can be obtained:

$$I(\pi; L_N) = \int_\Theta \mathrm{KL}(L_N(\cdot \mid \theta) || p_{\pi,N}) \pi(\theta) d\theta, \tag{5}$$

with $p_{\pi,N}$ denoting the distribution of the data $\mathbf{X}$. The above equation expresses $I(\pi; L_N)$ as the amount of information that differs between the random variable the random variable of the parameters $\theta \sim \pi$ and the random variable of the data $\mathbf{X} \sim p_{\pi,N}$. A more generalized definitio of RPs has been proposed in [37]. The generalized mutual information with a f-divergence is define by

$$I_{\mathrm{D}_f}(\pi; L_N) = \int_\Theta \mathrm{D}_f(p_{\pi,N} || L_N(\cdot \mid \theta)) \pi(\theta) d\theta, \tag{6}$$

with

$$\mathrm{D}_f(p\,||\,q) = \int_\Omega f\left(\frac{p(\omega)}{q(\omega)}\right) q(\omega) d\omega, \tag{7}$$

where f is convex and verifie $f(1) = 0$. The classical mutual information corresponds to the special case $f = -\log$, indeed, $\mathrm{D}_{-\log}(p\,||\,q) = \mathrm{KL}(q\,||\,p)$. When $N \to +\infty$, it has been shown in [37, 38] that the solution of this asymptotic problem is the Jeffreys prior when the mutual information is expressed as in Equation (3), or when it is define using an δ-divergence, as in Equation (6) with $f = f_\delta$, where:

$$f_\delta(x) = \frac{x^\delta - \delta x - (1-\delta)}{\delta(\delta-1)}, \quad \delta \in (0,1). \tag{8}$$

The Jeffreys prior is define as: $J(\theta) \propto \det(\mathcal{I}(\theta))^{1/2}$ with $\mathcal{I}$ being the Fisher information matrix.

The Jeffreys prior is a non-informative prior, known to be improper (i.e. it integrates to infinity in many circumstances. Moreover, its decay rates can produce improper posteriors. Improper posteriors are a problem because they prevent the generation of *a posteriori* estimates by Markov Chain Monte Carlo (MCMC) methods. To handle this issue, we adapted the algorithm to the framework of [39] that suggests solving a constrained version of the initial optimization problem to ensure the solution is proper:

$$\tilde{\pi}^* \in \underset{\substack{\pi\in\mathcal{P}\\ \text{s.t. } \mathcal{C}(\pi)<\infty}}{\arg\max}\, I_{\mathrm{D}_{f_\delta}}(\pi; L_N), \tag{9}$$

where the constraint is $\mathcal{C}(\pi) = \int_\Theta a(\theta)\pi(\theta)d\theta$ with a positive function a. When a δ-divergence is used in the definitio of the mutual information, and the two following integrals are finite

$$\int_\Theta J(\theta)a(\theta)^{1/\delta} d\theta < \infty \quad \text{and} \quad \int_\Theta J(\theta)a(\theta)^{1+1/\delta} d\theta < \infty, \tag{10}$$

a result in [39] states that the proper constrained RP $\tilde{\pi}^*$ has the following asymptotic form:

$$\tilde{\pi}^*(\theta) \propto J(\theta)a(\theta)^{1/\delta}. \tag{11}$$

3 VARIATIONAL APPROXIMATION OF THE REFERENCE PRIOR (VA-RP)

3.1 Definition

Variational inference encompasses a range of techniques designed to approximate probability distributions by solving an optimization problem —such as maximizing the evidence lower bound (ELBO) [40]. Given that our objective involves findin the prior that maximizes the mutual information define in Equation (5), variational inference naturally lends itself to the approximation of RPs.

We present in this section the approach developed in [35], which is inspired by [41, 42] regarding the variational approximation of the RP, referred as the VA-RP. We limit the family of priors to a parametric class $\{\pi_\lambda,\ \lambda \in \Lambda\}$, with $\Lambda \subset \mathbb{R}^L$, thereby reducing the problem to a finite dimensional setting. Consequently, the optimization tasks in equations (2) and (9) translate into findin $\arg\max_{\lambda\in\Lambda} I_{\mathrm{D}_f}(\pi_\lambda; L_N)$. We defin the priors π_λ implicitly through the transformation:

$$\theta \sim \pi_\lambda \iff \theta = g(\lambda, \varepsilon), \quad \text{where} \quad \varepsilon \sim p_\varepsilon. \tag{12}$$

Here, g is a parameterized measurable function —typically a neural network— where λ represents its weights and biases, it has to be differentiable with respect to λ. The variable ε serves as a latent variable, drawn from a simple, easy-to-sample distribution p_ε. In practice, we choose a centered multivariate Gaussian $\mathcal{N}(0, \mathbb{I}_{p\times p})$.

This implicit construction provides considerable flexibility in defining prior distributions. However, except in simpler cases, the density function of π_λ remains unknown and cannot be evaluated directly. Instead, we obtain samples $\theta \sim \pi_\lambda$.

The method in [35] extends the ideas in [41, 42] to higher-dimensional models while incorporating a broader definition of mutual information based on f-divergences, as given in Equation (5).

3.2 Objective function and gradient

The numerical resolution of the optimization problem is based on stochastic gradient ascent with an objective function corresponding to a lower bound of the mutual information. This function was first introduced in [42] and a generalization to f-divergences with f decreasing is given in [35]:

$$B_{\mathrm{D}_f}(\pi; L_N) = \int_\Theta \int_{\mathcal{X}^N} f\left(\frac{L_N(\mathbf{X}\,|\,\hat{\theta}_{MLE})}{L_N(\mathbf{X}\,|\,\theta)}\right) \pi(\theta) L_N(\mathbf{X}\,|\,\theta) d\mathbf{X} d\theta, \tag{13}$$

with $\hat{\theta}_{MLE}$ being the maximum likelihood estimator (MLE). Since the function f_δ used in δ-divergence (Equation (8)) is not decreasing, we replace it with $\hat{f}_\delta$ defined hereafter, because $\mathrm{D}_{f_\delta} = \mathrm{D}_{\hat{f}_\delta}$:

$$\hat{f}_\delta(x) = \frac{x^\delta - 1}{\delta(\delta-1)} = f_\delta(x) + \frac{1}{\delta - 1}(x-1). \tag{14}$$

The gradient of the lower bound can be written as:

$$\frac{\partial B_{\mathrm{D}_f}}{\partial \lambda_l}(\pi_\lambda; L_N) = \mathbb{E}_\varepsilon\left[\sum_{j=1}^{q} \frac{\partial \tilde{B}}{\partial \theta_j}(g(\lambda, \varepsilon)) \frac{\partial g_j}{\partial \lambda_l}(\lambda, \varepsilon)\right], \tag{15}$$

where:

$$\frac{\partial \tilde{B}}{\partial \theta_j}(\theta) = \mathbb{E}_{\mathbf{X}\sim L_N(\cdot\,|\,\theta)}\left[\frac{\partial \log L_N}{\partial \theta_j}(\mathbf{X}\,|\,\theta) F\left(\frac{L_N(\mathbf{X}\,|\,\hat{\theta}_{MLE})}{L_N(\mathbf{X}\,|\,\theta)}\right)\right], \tag{16}$$

with $F(x) = f(x) - xf'(x)$. The previous gradient can be approximated via Monte Carlo using samples $\varepsilon \sim p_\varepsilon$ and $\mathbf{X} \sim L_N(\cdot|\theta)$. For the MLE, we also use samples of ε when no formula is available:

$$L_N(\mathbf{X}\,|\,\hat{\theta}_{MLE}) \approx \max_{t\in\{1,\dots,T\}} L_N(\mathbf{X}\,|\,g(\lambda, \varepsilon_t)) \quad \text{where} \quad \varepsilon_1, \dots, \varepsilon_T \sim p_\varepsilon. \tag{17}$$

The implementation in [35] relies on the PyTorch package [43] and the Adam optimizer [44], where the differentiation operations and the sampling operations are separated when computing the gradient with automatic differentiation.

3.3 Constraints incorporation

The authors in [35] suggest a method to incorporate a constraint in the optimization process. The constraint on the prior must take the form of $\mathbb{E}_{\theta\sim\pi}[a(\theta)] = b$, with $a : \Theta \longrightarrow (0,\infty)$ being a positive function, and $b \in (0,\infty)$ being a positive scalar. The optimization problem becomes the following:

$$\underset{\substack{\lambda\in\Lambda,\\ \text{subject to } \mathcal{C}(\pi_\lambda)=0}}{\arg\max} \quad \mathcal{B}_{\mathrm{D}_{\hat{f}_\delta}}(\pi_\lambda; L_N) \quad \text{with} \quad \mathcal{C}(\pi_\lambda) := \mathbb{E}_{\theta\sim\pi_\lambda}\left[a(\theta)\right] - b. \tag{18}$$

The same algorithm can be used with the augmented Lagrangian as the objective function instead:

$$\mathcal{L}_A(\lambda,\eta,\tilde{\eta}) = \mathcal{B}_{\mathrm{D}_{\hat{f}_\delta}}(\pi_\lambda; L_N) + \eta\mathcal{C}(\pi_\lambda) - \frac{\tilde{\eta}}{2}\mathcal{C}(\pi_\lambda)^2. \tag{19}$$

The right value for b is obtained through the estimation of the two integrals in Equation (10), which requires an approximation of the Jeffreys prior, the details are given in [35].

3.4 Posterior sampling

In order to build *a posteriori* estimates, we propose a sampling method for θ from its posterior distribution $\pi_\lambda(\theta \,|\, \mathbf{X})$. The prior distribution π_λ has an implicit form in general and cannot be evaluated. The samples are obtained from the posterior distribution of the variable ε which has a known density function up to a multiplicative constant:

$$\pi^{\text{post}}_{\varepsilon,\lambda}(\varepsilon \,|\, \mathbf{X}) \propto p_\varepsilon(\varepsilon) L_N(\mathbf{X} \,|\, g(\lambda,\varepsilon)) \; ; \quad \text{if } \varepsilon \sim \pi^{\text{post}}_{\varepsilon,\lambda}(\varepsilon \,|\, \mathbf{X}) \text{ then } g(\lambda,\varepsilon) \sim \pi_\lambda(\theta \,|\, \mathbf{X}). \tag{20}$$

We utilize in practice an adaptive version of the Metropolis-Hastings algorithm with a multivariate Gaussian as the proposition distribution.

4 VA-RP FOR SEISMIC FRAGILITY CURVES ESTIMATION

4.1 Problem statement

A seismic fragility curve is the probability of failure, denoted by $P_f(a)$, of a mechanical structure subjected to a seism as a function of an IM, which is a scalar value denoted by a and derived from the seismic ground motion. We denote by Z the variable that equals 1 if the mechanical structure fails, and 0 otherwise. We introduce the probit-lognormal model used to estimate seismic fragility curves, it is also referred as the lognormal model in the literature. The properties of the Jeffreys prior for this model are highlighted by [31].

The model is define by the observation of a vector of i.i.d. samples $\mathbf{X} = (X_1,\dots,X_N)$ where for any index i, $X_i \sim (Z,a) \in \mathcal{X} = \{0,1\}\times(0,\infty)$. The random variable (Z,a) has the following distribution that depends on the parameters $\theta = (\alpha,\beta) \in \Theta = (0,\infty)^2$:

$$\begin{cases} a \sim p_a \\ P_f(a) = \Phi\left(\dfrac{\log a - \log\alpha}{\beta}\right) \\ Z \sim \text{Bernoulli}(P_f(a)), \end{cases} \tag{21}$$

where Φ designates the cumulative distribution function of the standard Gaussian. The distribution p_a is supposed to be known, it is common to associate it to a log-normal distribution:

$\log(a) \sim \mathcal{N}(\mu_a, \sigma_a^2)$ (see [31–33] and Section 5). The likelihood writes:

$$L_N(\mathbf{X} \mid \theta) = \prod_{i=1}^{N} p_a(a_i) \prod_{i=1}^{N} P_f(a_i)^{Z_i}(1 - P_f(a_i))^{1-Z_i} \propto \prod_{i=1}^{N} P_f(a_i)^{Z_i}(1 - P_f(a_i))^{1-Z_i} \tag{22}$$

For simplicity, we denote : $\gamma_i = \dfrac{\log a_i - \log \alpha}{\beta} = \Phi^{-1}(P_f(a_i)) = \text{probit}(P_f(a_i))$, the gradient of the log-likelihood has the following expression:

$$\begin{cases} \dfrac{\partial \log L_N}{\partial \alpha}(\mathbf{X} \mid \theta) = \displaystyle\sum_{i=1}^{N} \frac{1}{\alpha\beta} \left((-Z_i) \frac{\Phi'(\gamma_i)}{\Phi(\gamma_i)} + (1 - Z_i) \frac{\Phi'(\gamma_i)}{1 - \Phi(\gamma_i)} \right) \\ \dfrac{\partial \log L_N}{\partial \beta}(\mathbf{X} \mid \theta) = \displaystyle\sum_{i=1}^{N} \frac{\gamma_i}{\beta} \left((-Z_i) \frac{\Phi'(\gamma_i)}{\Phi(\gamma_i)} + (1 - Z_i) \frac{\Phi'(\gamma_i)}{1 - \Phi(\gamma_i)} \right). \end{cases} \tag{23}$$

We do not dispose of an explicit expression of the MLE. Consequently, we approximate it using samples in the variational inference process.

In [31], the authors study the asymptotical behaviors of the likelihood and of the Jeffreys prior. They prove that the Jeffreys prior is improper w.r.t. β, namely, when $\alpha > 0$ is fi ed,

$$\begin{cases} J(\theta) \propto 1/\beta & \text{as} \quad \beta \longrightarrow 0 \\ J(\theta) \propto 1/\beta^3 & \text{as} \quad \beta \longrightarrow +\infty. \end{cases} \tag{24}$$

If $\beta > 0$ is fi ed, the prior is proper on the firs component α:

$$J(\theta) \propto \frac{|\log \alpha|}{\alpha} \exp\left(-\frac{(\log \alpha - \mu_a)^2}{2\beta^2 + 2\sigma_a^2} \right) \quad \text{when} \quad |\log \alpha| \longrightarrow +\infty. \tag{25}$$

It is proven in [31] that the improper rates of the Jeffreys prior w.r.t. β can yield improper posteriors. Based on this work and taking into account additional constraints to make it proper (see Section 2), the authors in [33] suggest, in order to have a more in-depth look at the Jeffreys prior, to approximate it from its decay rates. The resulting prior π_a is then given by:

$$\pi_a(\theta) = \frac{K}{\beta + \beta^3} \frac{|\log \alpha|}{\alpha} \exp\left(-\frac{(\log \alpha - \mu_a)^2}{2\beta^2 + 2\sigma_a^2} \right) \tag{26}$$

K represents a positive normalization constant. Furthermore, it is proven in Appendix B that, $\mathbb{E}_{\theta \sim \pi_a} |\log \alpha|^2 = \infty$. This result will be very useful in the sequel to interpret the results of the VA-RP training.

4.2 Neural network architecture

A single layer neural network is implemented for this problem. It takes the following form

$$\varepsilon \mapsto \begin{pmatrix} \mathbf{h}_1(\varepsilon) \\ \mathbf{h}_2(\varepsilon) \end{pmatrix} \mapsto \begin{pmatrix} \tilde{\varepsilon}_1 = \mathbf{w}_1^T \mathbf{h}_1(\varepsilon) + b_1 \\ \tilde{\varepsilon}_2 = \mathbf{w}_2^T \mathbf{h}_2(\varepsilon) + b_2 \end{pmatrix} \mapsto \begin{pmatrix} \alpha = u_1(\tilde{\varepsilon}_1) \\ \beta = u_2(\tilde{\varepsilon}_2) \end{pmatrix}. \tag{27}$$

The coordinates of the vectors $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{R}^p$ and $\mathbf{b} = (b_1, b_2) \in \mathbb{R}^2$ constitute the weights of the neural network, they are initialized independently and randomly w.r.t. a Gaussian distribution.

$\mathbf{h}_1$, $\mathbf{h}_2$, u_1, u_2 are activation functions. They are specificall chosen in order to ensure the VA-RP's decay rates match with the ones of the target prior: $u_1 = u_2 = \exp$ and for $\mathbf{x} \in \mathbb{R}^p$, $\mathbf{h}_1(\mathbf{x}) = (h_1(x_i))_i$ and $\mathbf{h}_2(\mathbf{x}) = (h_2(x_i))_i$ with

$$h_1(x) = x; \quad h_2(x) = \log(1 - \Phi(x)), \tag{28}$$

where Φ designates the c.d.f. of a standard Gaussian. In Appendix A, we prove that under those settings the VA-RP yields marginal distributions p_α, p_β w.r.t. α, β, that take the following form:

$$\begin{aligned} p_\alpha(\alpha) &= \frac{1}{\alpha\sqrt{2\pi}\|\mathbf{w}_1\|_2} \exp\left(-\frac{(\log\alpha - b_1)^2}{2\|\mathbf{w}_1\|_2^2}\right) \\ p_\beta(\beta) &= \sum_{i,\, w_{2i}>0} K_i \beta^{-\frac{1}{w_{2i}}-1} \mathbb{1}_{\beta > e^{b_2}} + \sum_{i,\, w_{2i}<0} K_i \beta^{-\frac{1}{w_{2i}}-1} \mathbb{1}_{\beta < e^{b_2}}, \end{aligned} \tag{29}$$

where $(w_{2i})_{i=1}^p$ denote the coordinates of $\mathbf{w}_2$, and the K_i are constants that depend on $\mathbf{w}_2$ and b_2. The marginal distribution p_α is a log-normal distribution, which is similar, up to the $\log\alpha$ term, to the Jeffreys prior decay rates w.r.t. α (Equation (25)). Concerning p_β, calling $w_{2j}^{-1} = \min\{w_{2i}^{-1},\, w_{2i} > 0\}$ and $w_{2l}^{-1} = \max\{w_{2i}^{-1},\, w_{2i} < 0\}$, it admits the following decay rates:

$$p_\beta(\beta) \underset{\beta\to 0}{\sim} K_j \beta^{-\frac{1}{w_{2l}}-1}; \quad p_\beta(\beta) \underset{\beta\to\infty}{\sim} K_l \beta^{-\frac{1}{w_{2j}}-1}. \tag{30}$$

They are consistent with the ones of the Jeffreys prior.

4.3 Numerical results

Training of the unconstrained VA-RP In this section, we present the training process of the VA-RP. We have implemented the probit-lognormal modeling with $\mu_a = 0$ and $\sigma_a = 1$. These parameters, which defin the distribution of a, are the only elements that have a theoretical impact on the reference prior. Their values are derived from the case study which is considered in Section 5.

As for the different parameters for the training, we take $N = 50$, $T = 50$, $p = 50$. The training consists in the maximization of $\mathcal{B}_{\mathrm{D}_{\hat{f}_\delta}}$ with $\delta = 0.5$. The optimization has been done through a number of 9000 epochs, with a learning rate of 0.005.

Figure 1 shows the evolutions of the weights of the neural network that defin the marginal distributions of p_α and p_β (see Equation (29)), as a function of the number of epochs. Thus, the bias b_1 is constant and equal to 0, which is consistent with the asymptotic distribution of α. $1/w_{2l}$ and $1/w_{2j}$ both tend to 0. As a result, the marginal distribution of β tends to $1/\beta$. This is an improper distribution which does not exactly match the expected improper Jeffreys prior (see Equation (24)): it has a heavier tail as $\beta \longrightarrow \infty$ than the original Jeffreys. However, we will see in the end of this section that the distribution is not limited to its tails, and that this prior actually ascribes huge weights to small values of β (see the paragraph "Posterior evaluation"). Finally, $\|\mathbf{w}_1\|_2$ tends to $+\infty$. This means that the variance of $\log\alpha$ tends to $+\infty$ as it is expected (see Appendix B), even though this does not lead to the target distribution. This last result is nevertheless not surprising. It is the consequence of the chosen architecture of the VA-RP: p_α is a lognormal distribution (Equation (30)), which is not exactly the case for the Jeffreys prior (Equation (25)). What fundamentally distinguishes these two distributions is that (i) in the firs case, that of Jeffreys, the variance of $\log\alpha$ is finit conditionally to β, and is infinit otherwise because of the heavy tail of the distribution w.r.t. β while (ii) in the second case, the variance

becomes infinit to correspond, via the training process, to the target variance of the Jeffreys prior. In doing so, the marginal distribution of α tends towards the improper distribution $1/\alpha$. Although this result is open to criticism, it appears to be an "acceptable" compromise insofar as the model's likelihood tends to substantially attenuate the prior's tails with respect to α, thereby limiting the influenc of the prior's decay rates on the posterior. An alternative approach could involve modifying the neural network architecture to better capture the correlation between α and β. However, such adjustments would inevitably increase the computational complexity of the training process.

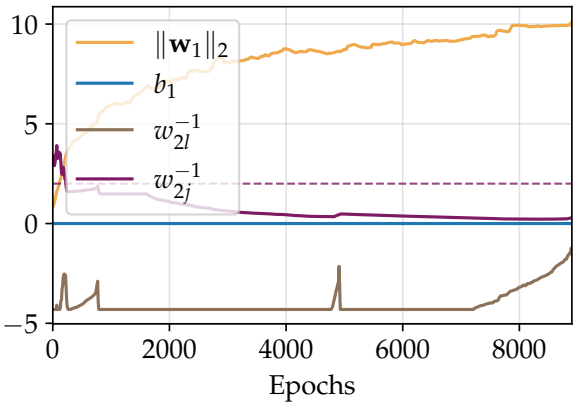


Figure 1: Weights of interest of the neural network during the unconstrained optimization. The brown and violet curves represent respectively $1/w_{2l}$ and $1/w_{2j}$ that are define in Section 4.2. In violet dashed line is plotted the target value of $1/w_{2j}$.

Training of the constrained VA-RP As for the constrained case, we use a moment constraint of the form $a(\alpha, \beta) = \beta^\kappa$ for the second component with $\kappa \in \left(0, \frac{2}{1+1/\delta}\right)$, to ensure the proper character of the resulting constrained prior. We defin $\gamma = \kappa/\delta \in (0, \frac{2}{1+\delta}) \subset (0, 2)$, the target proper prior is : $\pi^*(\theta) \propto J(\theta)\beta^\gamma$. Its asymptotic rates are identical to those of the Jeffreys prior for α. If we fi α, we have then:

$$\begin{cases} \pi^*(\theta) \propto \beta^{\gamma-1} & \text{as} \quad \beta \longrightarrow 0 \\ \pi^*(\theta) \propto \beta^{\gamma-3} & \text{as} \quad \beta \longrightarrow +\infty. \end{cases} \tag{31}$$

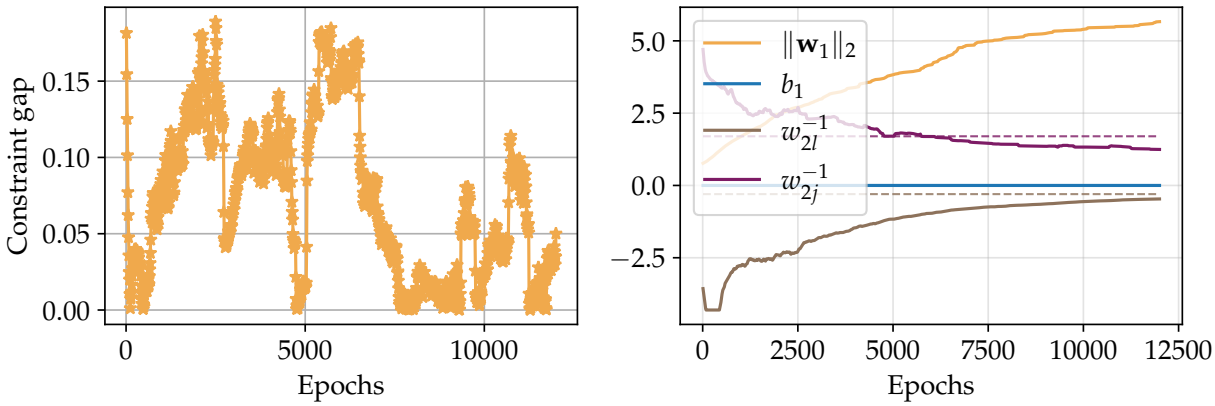
In this case, the approximation of the target prior, as suggested in [33] and expressed in Equation (26), becomes:

$$\pi_a^\gamma(\theta) = \frac{K_\gamma}{\beta^{1-\gamma} + \beta^{3-\gamma}} \frac{|\log \alpha|}{\alpha} \exp\left(-\frac{(\log \alpha - \mu)^2}{2\beta^2 + 2\sigma^2}\right), \tag{32}$$

with K_γ depending only on γ. In Appendix B, it is proven that for any $\gamma > 0$ this prior verifie $\mathbb{E}_{\theta\sim\pi_a^\gamma} |\log \alpha|^2 = \infty$.

The divergence is still set with $\delta = 0.5$, and we fi $\kappa = 0.15$, so that $\gamma = 0.3$. The training process of the constrained VA-RP is done during 12000 epochs with a learning rate of 0.001. We take $N = 50$, $T = 50$ and $p = 50$. the parameter η in the augmented Lagrangian is updated every 10 epochs.

Figure 2a suggests that the constraint seems reasonably verifie during the whole optimization as the constraint gap is close to zero and appears to be decreasing. Figure 2b shows that the inverse weights $1/w_{2l}$ and $1/w_{2j}$ appear to steadily approach their theoretical values, respectively $-\gamma = -0.3$ and $2 - \gamma = 1.7$. As in the unconstrained case, we observe that $\|\mathbf{w}_1\|_2$ tends to $+\infty$. This is consistent with the theoretical results of the Appendix B which show, recall, that for all $\gamma \geq 0$, $\mathbb{E}[|\log \alpha|^2] = \infty$. So, we verify that adding constraints therefore allows, as expected, to "control" the distribution of β. Concerning α, we fin the same trend as in the unconstrained case, which is theoretically expected from the point of view of the variance of the marginal distribution.



(a) Constraint value gap.

(b) Weights of interest of the neural network.

Figure 2: (a) Evolution of the constraint value gap during training. It corresponds to the difference between the target and current values for the constraint (in absolute value). (b) Weights of interest of the neural network during the constrained optimization. The yellow and blue curves represent respectively $\|\mathbf{w}_1\|_2$ and b_1, where $\|\mathbf{w}_1\|_2^2$ is the variance of the log of the firs component, and b_1 its mean. The brown and violet curves represent respectively $1/w_{2l}$ and $1/w_{2j}$ that are define in Section 4.2. In same color dashed line are plotted their target value.

Posterior evaluation We suggest to evaluate the posterior that the VA-RPs derived in this section yield. For this purpose, we take as dataset 50 samples that are generated given a probit-lognormal model with θ_{true} close to $(3.37, 0.43)$. The distribution of a considered is still a log-normal with $\mu_a = 0$ and $\sigma_a = 1$.

A posteriori samples are generated using the adaptive Metropolis-Hasting on ε as described in Section 3.4. A number of 5000 samples of θ are generated this way. For the sake of completeness, they are compared with samples generated considering an other prior: the actual reference prior derived by approximating numerically its theoretical expression, as depicted in [31]. The samples are generated by an other adaptive Metropolis-Hastings algorithm, which iteratively evaluates the prior. In the following, we refer to the latter as AJ.

In this example, we have not retained the samples for which the data are dissociated into two disjoint subsets when they are classifie according to the values of a. In this case, indeed, the posterior distribution is improper because the likelihood is constant when β tends to 0 [31]. We then say that the dataset is degenerate [33]. The probability of occurrence of such datasets is not negligible if the dataset has a small size. The following plots are obtained with non-degenerate

datasets.

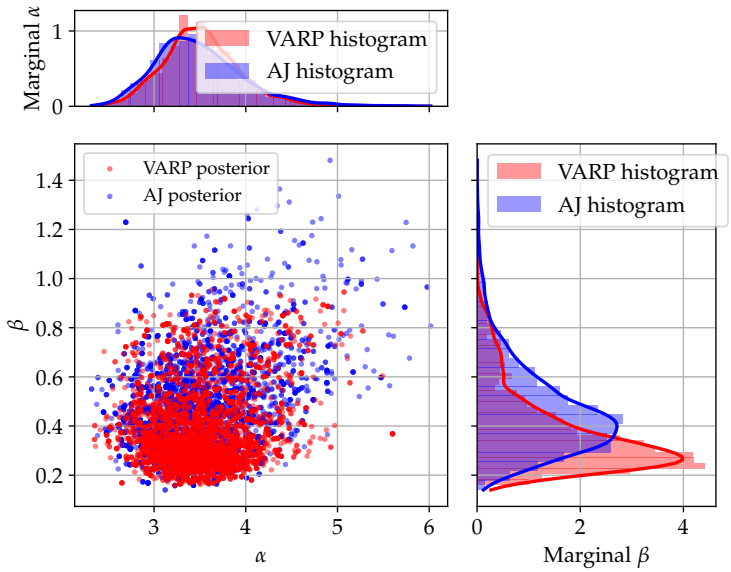


Figure 3: Scatter histogram of the unconstrained VA-RP posterior and the approximated Jeffreys (AJ) posterior distributions obtained from 5000 samples. Kernel density estimation is used on the marginal distributions in order to approximate their density functions with Gaussian kernels.

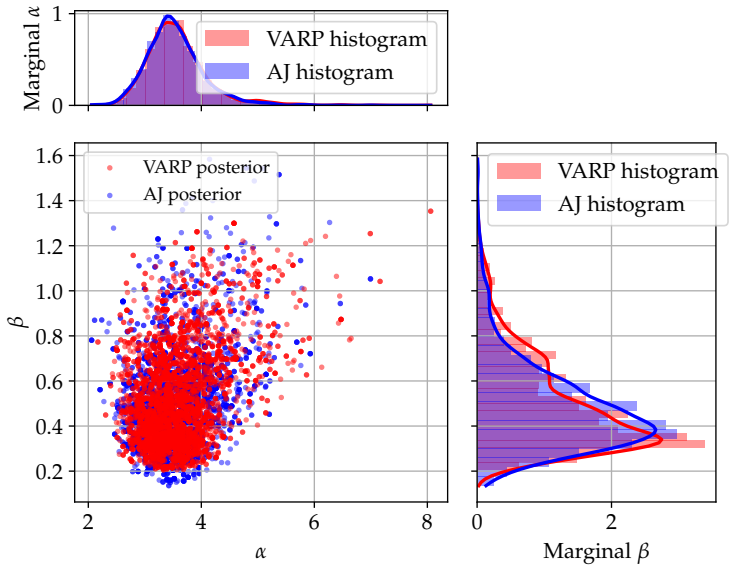


Figure 4: Scatter histogram of the constrained VA-RP posterior and the approximated Jeffreys (AJ) posterior distributions obtained from 5000 samples. Kernel density estimation is used on the marginal distributions in order to approximate their density functions with Gaussian kernels.

Figures 3 and 4 show that the posterior distribution obtained from the VA-RP is more like Jeffreys prior in the constrained case than in the unconstrained one. More specificall , w.r.t. α, the VA-RP is similar to the Jeffreys prior in the both cases. That suggests that the limitation of

our neural network architecture discussed earlier has little impact on the posterior distribution. Regarding β, in the unconstrained case, the VA-RP ascribes predominant weight to smaller values of β compared to the Jeffreys prior, with a mode that is closer to 0. In terms of fragility curve estimation that could represent an issue, as it may lead to estimates with thinner credibility intervals around a fragility curve.

5 APPLICATION OF THE METHOD ON A CASE STUDY TAKEN FROM THE NUCLEAR INDUSTRY

5.1 Artificial seismic signals

In this study, we rely on a large database of 10^5 artificia seismic signals that were generated beforehand. The generation of them is based on the stochastic generator suggested in [45]. It has been calibrated as in [46] from 97 real accelerograms taken from the European strong motion database for a magnitude that lives between 5.5 and 6.5, and a source-to-site distance that is lower than 20 km [47].

The IM that we have derived for this study is the peak ground acceleration (PGA). In Figure 5, we compare the empirical distribution of the PGA given from that database with an approximation of the one that comes from the 97 real seismic signals. In the same figur is also plotted the p.d.f. of a log-normal distribution with same median and same log deviation than the empirical distribution of the IM. This distribution corresponds to a log-normal with parameters (μ_a, σ_a) that are approximately equal to $(0, 1)$.

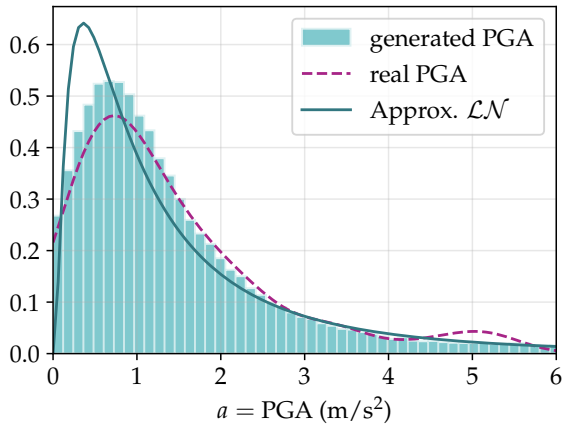


Figure 5: Histogram of the PGA of 10^5 generated signals compared with: 1. an approximation of the distribution of 97 PGA derived from real seismic signals using kernel density estimation (dashed line), 2. the p.d.f. of a log-normal distribution with same median and same log deviation than the generated signals (solid line).

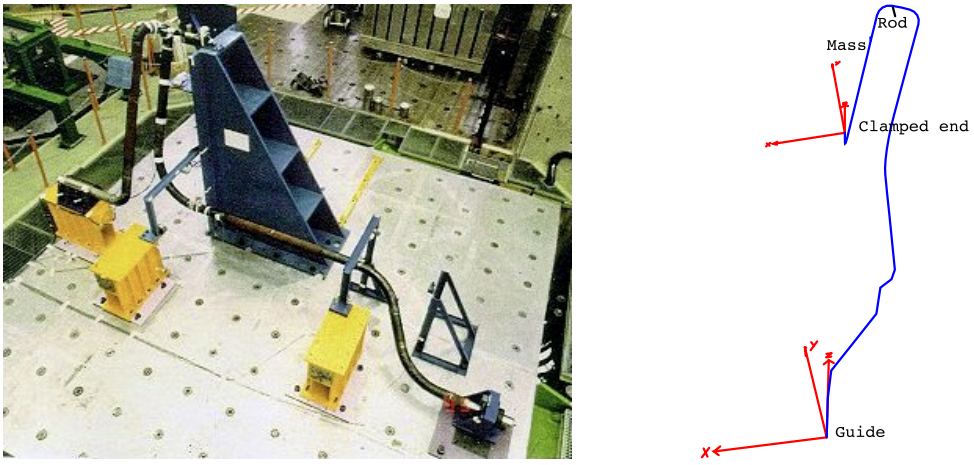
5.2 Description of the case study

This case study examines a piping system that constitutes a secondary line within a French pressurized water reactor. The system was investigated both experimentally and numerically [48]. Figure 6a presents an overview of the experimental mock-up, which was mounted on the Azalee shaking table at the EMSI laboratory of CEA. The corresponding finit element model (FEM), constructed using beam elements, is depicted in Figure 6b. This numerical model was

implemented using the in-house finit element code CAST3M [49] and validated through an extensive experimental campaign.

The mock-up is fi ed at one end, while the opposite end is supported by a guide to restrict displacement in the X and Y directions. Additionally, a rod is positioned atop the specimen to constrain mass displacement in the Z direction (see Figure 6b). During the experimental tests, excitation was applied exclusively in the X direction. The artificia input signals were filtere using a fictitiou 2% damped linear single-mode system at 5 Hz, which corresponds to the firs eigenfrequency of the piping system with a 1% damping ratio.

The engineering demand parameter (EDP) adopted in this study is the out-of-plane rotation of the elbow located near the clamped end of the mock-up, as recommended in [50]. The failure is define as an excessive value of that EDP. The critical rotation threshold is set at $3.8°$, representing the 90% quantile of a sample consisting of $8 \cdot 10^4$ numerical simulations (see Figure 7b).



(a) (b)

Figure 6: (a) Overiew of the mock-up of the piping system on the Azalee shaking table in CEA and (b) associated FEM.

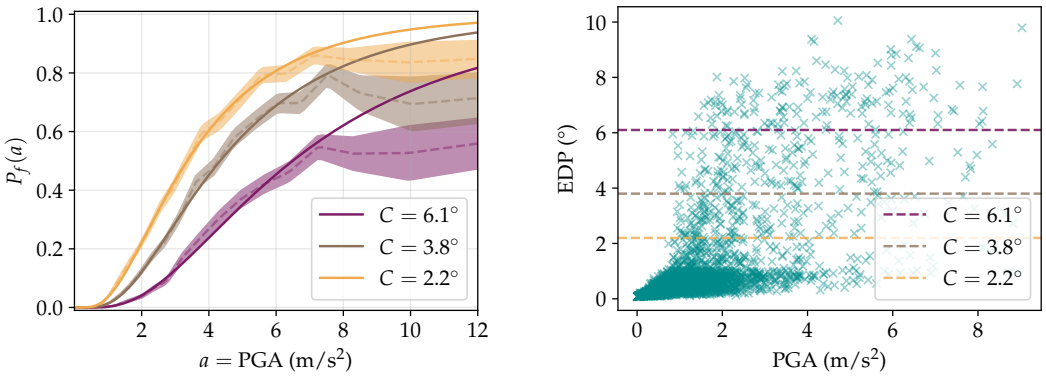
5.3 Reference fragility curve and benchmarking metrics

To evaluate our estimates, we compare them to a reference fragility curve, derived for this case study using the full batch of the $8 \cdot 10^4$ data. That reference fragility curve, denoted by P_f^{ref} consists in a non-parametricestimate based on Monte-Carlo estimations of local probability of failures w.r.t. the IM, as suggested in [12].

Different fragility curves, derived considering different excessive rotation threshold for defin ing the equipment's failure are elucidated in Figure 7a. Each is compared to a probit-lognormal curve P_f^{MLE} whose parameters are estimated by maximum likelihood estimation using the $8 \cdot 10^4$ results presented in Figure 7b. That comparison demonstrates the adequacy of such modeling of the fragility curve. However, it also emphasizes an existing bias that the model has compared with the non-parametricestimate. In the following, we refer to this difference between P_f^{MLE} and P_f^{ref} as the square model bias, denoted by $\mathcal{MB}$:

$$\mathcal{MB} = \| P_f^{\text{ref}} - P_f^{\text{MLE}} \|_{L^2}^2; \tag{33}$$

where the norm $\| \cdot \|_{L^2}$ is define by $\|P\|_{L^2}^2 = \int_0^\infty P(a)^2 da$. It is derived numerically from Simpson's interpolation on a sub-division of $(0, \infty)$.



(a) (b)

Figure 7: (a) Non-parametric reference fragility curves (dashed lines) surrounded by their 95% confidenc intervals for different critical rotation threshold C. The thresholds yield different proportion of failures in the database, respectively 95% (purple), 90% (brown), and 85% (yellow). For each value of C are plotted (same color, solid line) the corresponding probit-lognormal MLE. (b) Results of the $8 \cdot 10^4$ numerical simulations : each cross is an element of the database (IM, EDP).

Given an observed sample of size N: $\mathbf{X} = (X_1, \ldots, X_N)$, we denote by $P_f^N : a \mapsto \Phi(\beta^{-1} \log(a/\alpha))$ whose distribution inherits from the posterior distribution $\pi(\cdot \,|\, \mathbf{X})$ of θ. For any value of a, we note $q_r^N(a)$ the r-quantile of $P_f^N(a)$, and $m^N(a)$ its median. We propose to evaluate our estimates using the following metrics:

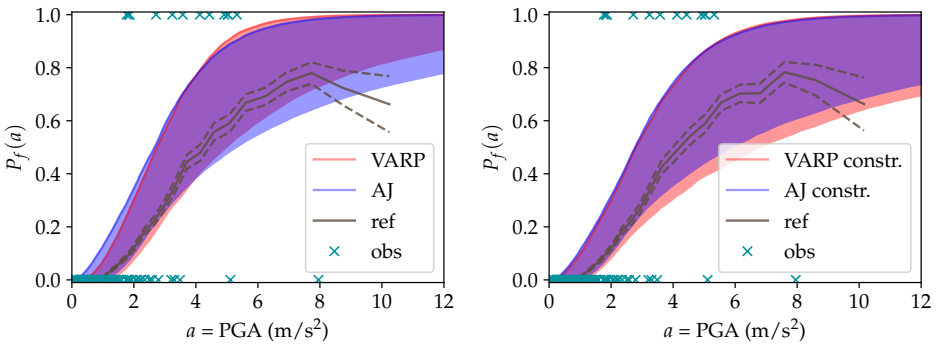
- The square bias to the median $\mathcal{B}^N = \|m^N - P_f^{\text{ref}}\|_{L^2}^2$.
- The credibility width $\mathcal{W}^N = \|q_{1-r/2}^N - q_{r/2}^N\|_{L^2}^2$, with $r = 0.05$.
- The upper quantile error $\mathcal{Q}^N = \|q_{1-r/2}^N - P_f^{\text{ref}}\|_{L^2}^2$.

The quantiles and the medians are approximated using generated samples.

5.4 Results and discussion

In this section, we compare the evaluation of our estimates that results from using the VA-RP with the evaluation of the ones that result from using an approximation via numerical derivation of the Jeffreys prior. In the following, we refer to the latter as AJ. Both are implemented with and without the incorporation of a constraint that is tailored to ensure they are proper. The considered constraint is detailed in Section 4.3.

In Figure 8 we provide examples of *a posteriori* fragility curves credibility intervals yielded by the two priors, with and without incorporating a constraint on the decay rates. These examples arise from given samples of $N = 80$ observations. They illustrate that the VA-RP is capable of providing fragility curve estimates that are equivalent of the ones provided by the AJ, which is an approximation of the exact expression of the reference prior.



(a) (b)

Figure 8: Example of *a posteriori* 95% credibility intervals yielded by the posterior distribution resulting from a sample of $N = 80$ data and from the VA-RP (in red) or the AJ (in blue) : (a) unconstrained priors and (b) constrained priors. The reference curve P_f^{ref} (solide line) is surrounded by its 95% confidenc interval (dashed line) in each figure The cyan crosses represent the observed sample.

To deepen the comparison, we implemented the methods multiple times for values of N varying in $[10, 250]$. More precisely, for each $N \in \{10, 20, \ldots, 250\}$, 10 random data samples of size N have been drawn. Each of these samples were then used to compute the metrics $\mathcal{B}^N$, $\mathcal{W}^N$ and $\mathcal{Q}^N$.

The average values of these metrics along with their 95% confidenc intervals are presented in Figure 9 (for the unconstrained case) and Figure 10 (for the constrained case) as functions of N.

Figures 9 and 10 confir previous results on a larger scale, namely that, with the current architecture, the constrained VA-RP outperforms the unconstrained VA-RP. Regarding bias, for example, it produces less erratic results. Compared to the AJ approximation, the VA-RP approximations exhibit slightly larger biases. The gap decreases as the number of data increases as expected, but only in the constrained case. Regarding the sizes of the credibility intervals (the values of $\mathcal{W}^N$), they are generally smaller in the unconstrained case. This result is consistent with the observation made in Figure 3 concerning the marginal distribution of β. This shows smaller values than with the AJ approximation of the Jeffreys prior. It is important to emphasize that it is primarily the lower bounds of the credibility intervals that are "penalized" with the constrained VA-RP approximation. The upper bound remains consistent with that of the AJ, as reflecte in the values of $\mathcal{Q}^N$. This result illustrates the relevance of using the constrained VA-RP approach for safety study. Indeed, in such a context, focusing on the upper quantile is consistent with a conservative approach.

6 CONCLUSION

Estimating seismic fragility curves when the available data are scarce is a daunting task. While the performances of a Bayesian framework in such settings are appealing, its implementation cannot overlook the prior selection process. Previous works have proven the strengths and highlighted the weaknesses of the reference prior theory incorporation in this context, with one major drawback being the computational complexity of deriving the reference prior.

In this work, we introduced a novel framework for estimating fragility curves using a variational approximation of the reference prior (VA-RP). The variational approximation that we

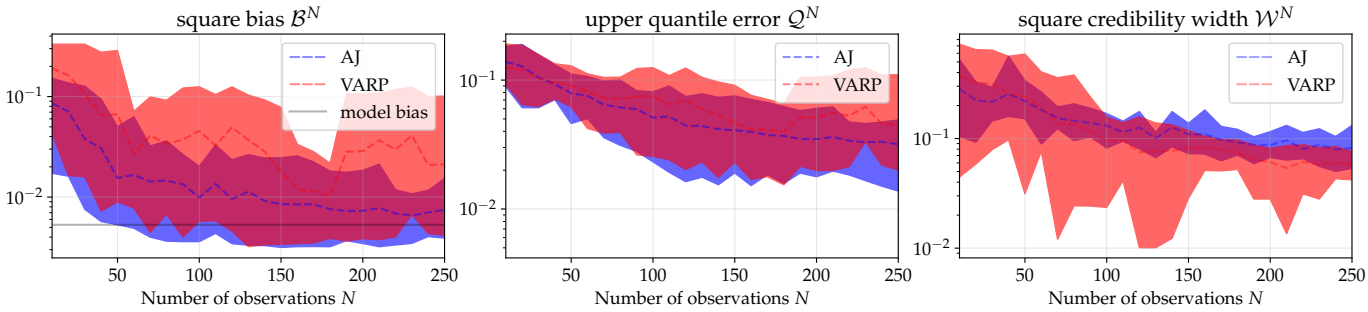


Figure 9: Average errors (dashed line) $\mathcal{B}^N$, $\mathcal{Q}^N$, $\mathcal{W}^N$ and their 95% confidenc intervals, derived from numerical replication of the unconstrained methods for each value of N in $\{10, 20, \cdots, 250\}$. The values derived using the VA-RP (red) are compared with the ones using a numerical derivation to approximate the Jeffreys prior (blue). In the left figure the square model bias $\mathcal{MB}$ is plotted (solid line).

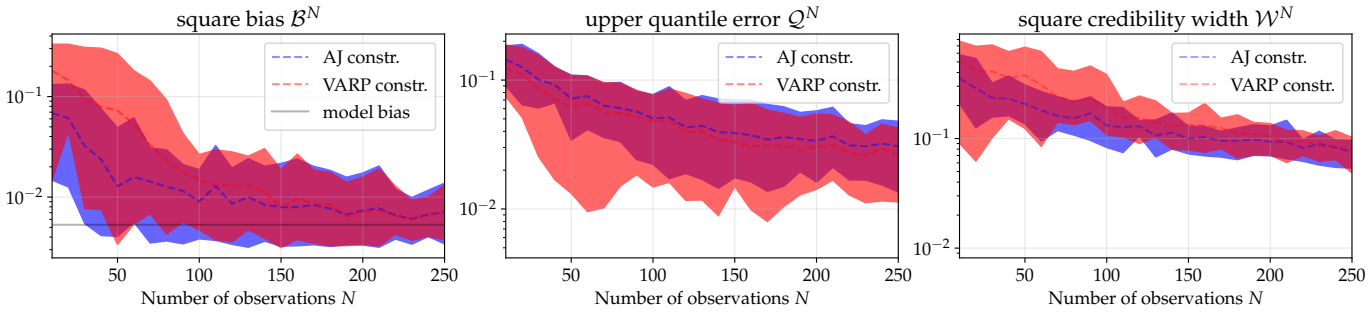


Figure 10: Same as Figure 9, but when a constraint is incorporated in the priors.

propose results from a neural network architecture specificall tailored for this application. We proved that the VA-RP accurately captures the asymptotical behavior of the reference prior through the training process. Additionally, we addressed the issue of improper posteriors by incorporating an adequate constraint into the prior. This framework facilitates an objective Bayesian estimation of seismic fragility curves, with the provision of estimates without requiring to evaluate a complex prior.

Eventually, our case study taken from the nuclear industry validated the effectiveness of the VA-RP in producing fragility curves estimates comparable to those obtained with the actual reference prior.

ACKNOWLEDGMENT

This research was supported by the CEA (French Alternative Energies and Atomic Energy Commission) and the SEISM Institute (https://www.institut-seism.fr/en/).

A APPENDIX: DECAY RATES OF THE VA-RP

In this appendix we elucidate the form taken by the marginal distribution of the VA-RP w.r.t. β. The VA-RP is assumed to be expressed as presented in Section 4.2 (Equation (27)): the random variable β is define as

$$\beta = \exp\left[\sum_{i,\, w_i \neq 0} w_{2i} \log(1 - \Phi(\varepsilon_i)) + b_i\right], \tag{34}$$

where the $(\varepsilon_i)_i$ are independent standard Gaussian variables. We recall that the weights $(w_{2i})_i$ are initialized randomly w.r.t. a Gaussian distribution. Therefore, we reasonably suppose that they have distinct absolute values. We rely on the two following lemmas.

Lemma 1. *Let X be a random variable following a standard Gaussian distribution, and $\mu > 0$. Then*

- *the r.v. $Z = -\mu \log(1 - \Phi(X))$ admits the density f_Z defined by $f_Z(z) = \frac{e^{-\frac{z}{\mu}}}{\mu} \mathbb{1}_{z>0}$, it is an exponential distribution with parameter μ^{-1};*
- *the r.v $\tilde{Z} = \mu \log(1 - \Phi(X))$ admits the density $f_{\tilde{Z}}$ defined by $f_{\tilde{Z}}(z) = \frac{e^{\frac{z}{\mu}}}{\mu} \mathbb{1}_{z<0}$, we call it an opposite exponential distribution with parameter $-\mu^{-1}$.*

Lemma 2. *Let $Z_1, \dots, Z_n$ be independent exponential or opposite exponential distribution with parameters $(\mu_i)_{i=1}^n$, distinct in absolute value. Their sum $\overline{Z}$ admits a density $f_{\overline{Z}}$ that takes the form*

$$f_{\overline{Z}}(z) = \sum_{i,\, \mu_i>0} K_i e^{-\mu_i z} \mathbb{1}_{z>0} + \sum_{i,\, \mu_i<0} K_i e^{-\mu_i z} \mathbb{1}_{z<0}, \tag{35}$$

for some constants $K_1, \dots, K_n$.

The expression of β as written in Equation (34) expresses $\log(\beta - b_2)$ as a sum of independent exponential or opposite exponential distributions with parameters $(|w_{2i}|^{-1})_i$, which are distinct. Thus, the r.v. $\log(\beta - b_2)$ admits a density $\tilde{p}$ that is define as in Equation (35):

$$\tilde{p}(z) = \sum_{i,\, w_{2i}>0} K_i e^{-\frac{z}{w_{2i}}} \mathbb{1}_{z>0} + \sum_{i,\, w_{2i}<0} K_i e^{-\frac{z}{w_{2i}}} \mathbb{1}_{z<0}. \tag{36}$$

Given that $p_\beta(\beta e^{b_2}) e^{b_2} \propto \tilde{p}(\log \beta)/\beta$ there exists some constants $(\tilde{K}_i)_{i=1}^n$ such that

$$p_\beta(\beta) = \sum_{i,\, w_{2i}>0} \tilde{K}_i \beta^{-\frac{1}{w_{2i}}-1} \mathbb{1}_{\beta>e^{b_2}} + \sum_{i,\, w_{2i}<0} \tilde{K}_i \beta^{-\frac{1}{w_{2i}}-1} \mathbb{1}_{\beta<e^{b_2}}. \tag{37}$$

Proof of Lemma 1 First, Φ being the c.d.f. of a standard Gaussian distribution, $\Phi(X)$ follows an uniform distribution in $(0, 1)$. The density $f_Z(z) = \frac{e^{-\frac{z}{\mu}}}{\mu}$ involved in the firs statement of the lemma is the p.d.f. of an exponential distribution with parameter μ^{-1}. The c.d.f. of that distribution is define by $F_Z(z) = -\mu \log(1 - z)$. Thus, $-\mu \log(1 - \Phi(X))$ is an exponential distribution with parameter μ^{-1}.
To conclude, notice that the second statement of the lemma simply results by elucidating the p.d.f. of $\tilde{Z}$ from the one of Z given that $\tilde{Z} = -Z$.

Proof of Lemma 2 We prove this result by induction over n. When $n = 1$, the form in Equation (35) is consistent with the p.d.f. of an exponential or opposite exponential distribution.

We suppose this statement true for $n - 1$ such r.v., we denote by $\tilde{f}$ the p.d.f. of $\sum_{i=1}^{n-1} Z_i$, and by f_{Z_n} the one of Z_n. As $\sum_{i=1}^{n-1} Z_i$ and Z_n are independent, $f_{\overline{Z}}$ equals the convolution between $\tilde{f}$ and f_{Z_n}: $f_{\overline{Z}} = \tilde{f} * f_{Z_n}$.

Before going further, let us derive the following integrals, for any $0 < \nu_1 < \nu_2$:

$$\begin{aligned}
&\int_{\mathbb{R}} e^{-\nu_1(y-x)} e^{-\nu_2 x} \mathbb{1}_{y-x>0} \mathbb{1}_{x>0} dx = \frac{e^{-\nu_2 y} - e^{-\nu_1 y}}{\nu_2 - \nu_1} \mathbb{1}_{y>0}, \\
&\int_{\mathbb{R}} e^{-\nu_1(y-x)} e^{\nu_2 x} \mathbb{1}_{y-x>0} \mathbb{1}_{x<0} dx = \frac{e^{\nu_2 y}}{\nu_1 + \nu_2} \mathbb{1}_{y<0} + \frac{e^{-\nu_1 y}}{\nu_1 + \nu_2} \mathbb{1}_{y>0}, \\
&\int_{\mathbb{R}} e^{\nu_1(y-x)} e^{\nu_2 x} \mathbb{1}_{y-x<0} \mathbb{1}_{x<0} dx = \frac{e^{\nu_1 y} - e^{\nu_2 y}}{\nu_2 - \nu_1} \mathbb{1}_{y<0}.
\end{aligned} \tag{38}$$

Thus, if $f_{Z_n}(x) = \mu_n e^{-\mu_n x} \mathbb{1}_{x>0}$ for $\mu_n > 0$ we get

$$\begin{aligned}
f_{\overline{Z}}(z) = &\sum_{i<n,\, \mu_i>0} \frac{K_i \mu_n}{\mu_i - \mu_n} e^{-\mu_i z} \mathbb{1}_{z>0} + \sum_{i<n,\, \mu_i<0} \frac{K_i \mu_n}{\mu_i + \mu_n} e^{-\mu_i z} \mathbb{1}_{z<0} \\
&- e^{-\mu_n z} \left[\sum_{i,\, \mu_i>0} \frac{K_i \mu_n}{\mu_i - \mu_n} + \sum_{i,\, \mu_i<0} \frac{K_i \mu_n}{\mu_i + \mu_n} \right] \mathbb{1}_{z>0};
\end{aligned} \tag{39}$$

and if $f_{Z_n}(x) = \mu_n e^{\mu_n x} \mathbb{1}_{x<0}$ for $\mu_n > 0$ then

$$\begin{aligned}
f_{\overline{Z}}(z) = &\sum_{i,\, \mu_i>0} \frac{K_i \mu_n}{\mu_i + \mu_n} e^{-\mu_i z} \mathbb{1}_{z>0} + \sum_{i,\, \mu_i<0} \frac{K_i \mu_n}{\rho - \mu_n} e^{-\mu_i z} \mathbb{1}_{z<0} \\
&+ e^{\mu_n z} \left[\sum_{i,\, \mu_i>0} \frac{K_i \mu_n}{\mu_i + \mu_n} + \sum_{i,\, \mu_i<0} \frac{K_i \mu_n}{\mu_i - \mu_n} \right] \mathbb{1}_{z<0}.
\end{aligned} \tag{40}$$

In any case, it fit the expected form.

B APPENDIX: VARIANCE OF α A PRIORI

In this section we consider the asymptotical version of the reference prior, that we denote by π_a^γ:

$$\pi_a^\gamma(\theta) = \frac{K}{\beta^{1-\gamma} + \beta^{3-\gamma}} \frac{|\log \alpha|}{\alpha} \exp\left(-\frac{(\log \alpha - \mu)^2}{2\beta^2 + 2\sigma^2} \right), \tag{41}$$

where $\mu \in \mathbb{R}$, $\sigma > 0$. The parameter γ is non null if the reference prior is not constrained, it lives in $(0, 2)$ if a constraint is incorporated as elucidated in Section 4.3. K represents a normalization constant that depends only on γ.

We fi $\beta > 0$, let us consider a random variable $X \sim \mathcal{N}(\mu, \Sigma^2)$, where $\Sigma = \Sigma(\beta) = \sqrt{\sigma^2 + \beta^2}$. We can write

$$\mathbb{E}|X| = \int_{-\infty}^{\infty} \frac{|x|}{\sqrt{2\pi\Sigma^2}} e^{-\frac{(x-\mu)^2}{2\Sigma^2}} dx = \frac{1}{\sqrt{2\pi\Sigma^2}} \int_0^{\infty} \frac{|\log y|}{y} e^{-\frac{(\log y - \mu)^2}{2\Sigma^2}} dy \tag{42}$$

$$\text{and} \quad \mathbb{E}|X|^3 = \int_{-\infty}^{\infty} \frac{|x|^3}{\sqrt{2\pi\Sigma^2}} e^{-\frac{(x-\mu)^2}{2\Sigma^2}} dx = \frac{1}{\sqrt{2\pi\Sigma^2}} \int_0^{\infty} \frac{|\log y|^3}{y} e^{-\frac{(\log y - \mu)^2}{2\Sigma^2}} dy. \tag{43}$$

Equation (42) elucidates the conditional distribution of $\alpha|\beta$ as the one whose density is

$$p_a^\gamma(\alpha|\beta) = (\mathbb{E}|X|\sqrt{2\pi\Sigma^2})^{-1} \frac{|\log \alpha|}{\alpha} \exp\left(-\frac{(\log \alpha - \mu)^2}{2\beta^2 + 2\sigma^2} \right); \tag{44}$$

and the marginal distribution of β as the one whose density is

$$p_a^\gamma(\beta) = \frac{K(\mathbb{E}|X|\sqrt{2\pi\Sigma^2})}{\beta^{1-\gamma} + \beta^{3-\gamma}}. \tag{45}$$

This way,

$$\mathbb{E}[|\log\alpha|^2 \,|\, \beta] = \frac{\mathbb{E}|X|^3}{\mathbb{E}|X|}. \tag{46}$$

The right hand term of the above equation equals the following (see [51]):

$$\begin{aligned} \mathbb{E}|X|^3 &= 2^{3/2}\Sigma^3 \frac{\Gamma(2)}{\Gamma(-\frac{3}{2})} \sum_{k=0}^{\infty} \frac{\Gamma(-\frac{3}{2}+k)}{\Gamma(\frac{1}{2}+k)} \frac{\left(-\frac{\mu}{2\Sigma^2}\right)^k}{k!}, \\ \text{and} \quad \mathbb{E}|X| &= 2^{1/2}\Sigma \frac{\Gamma(1)}{\Gamma(-\frac{1}{2})} \sum_{k=0}^{\infty} \frac{\Gamma(-\frac{1}{2}+k)}{\Gamma(\frac{1}{2}+k)} \frac{\left(-\frac{\mu}{2\Sigma^2}\right)^k}{k!}. \end{aligned} \tag{47}$$

Using the identity $z\Gamma(z) = \Gamma(z+1)$, we derive

$$\frac{\Gamma(-\frac{3}{2}+k)}{\Gamma(\frac{1}{2}+k)} = \frac{\Gamma(-\frac{1}{2}+k)}{(-\frac{3}{2}+k)\Gamma(\frac{1}{2}+k)} = \frac{1}{(-\frac{3}{2}+k)(-\frac{1}{2}+k)}. \tag{48}$$

Going back to Equation (46), we deduce:

$$\mathbb{E}|\log\alpha|^2 = \mathbb{E}\left[\mathbb{E}[|\log\alpha|^2 \,|\, \beta]\right] = K' \int_0^\infty \frac{(\beta^2+\sigma^2)^2}{\beta^{1-\gamma}+\beta^{3-\gamma}} \sum_{k=0}^{\infty} \frac{(-\mu(2\sigma^2+2\beta^2)^{-1})^k}{(2k-3)(2k-1)k!} d\beta, \tag{49}$$

with $K' > 0$ being a normalization constant. To conclude, we perform an asymptotical analysis of the integrated function in the above equation as $\beta \longrightarrow \infty$. The power series

$$\sum_{k=0}^{\infty} \frac{z^k}{(2k-3)(2k-1)k!} \tag{50}$$

has an infinit radius of convergence, so that when $z \longrightarrow 0$, it converges toward $\frac{1}{3}$. Also, while z is a negative real number, the power series is positive.

Therefore, the term in the integral of Equation (49) is positive and asymptotically equivalent to $\beta^{1+\gamma}$ as $\beta \longrightarrow \infty$. We deduce that for any $\gamma \geq 0$,

$$\mathbb{E}[|\log\alpha|^2] = \infty. \tag{51}$$

REFERENCES

[1] R. P. Kennedy, C. A. Cornell, R. D. Campbell, S. J. Kaplan, F. Harold, Probabilistic seismic safety study of an existing nuclear power plant, *Nuclear Engineering and Design,* **59**, 315–338, 1980.

[2] R. P. Kennedy, M. K. Ravindra, Seismic fragilities for nuclear power plant risk studies, *Nuclear Engineering and Design,* **79**, 47–68, 1984.

[3] B. Ellingwood, Validation studies of seismic PRAs, *Nuclear Engineering and Design,* **123**, 189–196, 1990.

[4] Y.-J. Park, C. H. Hofmayer, N. C. Chokshi, Survey of seismic fragilities used in PRA studies of nuclear power plants, *Reliability Engineering & System Safety,* **62**, 185–195, 1998.

[5] C. A. Cornell, Hazard, ground motions and probabilistic assessments for PBSD, in: *Proceedings of the International Workshop on Performance-Based Seismic Design - Concepts and Implementation, PEER Center, University of California, Berkeley, 39–52,* BLED, SLOVENIA, 2004.

[6] N. Luco, C. A. Cornell, Structure-specifi scalar intensity measures for near-source and ordinary earthquake ground motions, *Earthquake Spectra,* **23**, 357–392, 2007.

[7] B. R. Ellingwood, Earthquake risk assessment of building structures, *Reliability Engineering & System Safety,* **74**, 251–262, 2001.

[8] S.-H. Kim, M. Shinozuka, Development of fragility curves of bridges retrofitte by column jacketing, *Probabilistic Engineering Mechanics,* **19**, 105–112, 2004.

[9] I. Zentner, Numerical computation of fragility curves for NPP equipment, *Nuclear Engineering and Design,* **240**, 1614–1621, 2010.

[10] P. S. Koutsourelakis, Assessing structural vulnerability against earthquakes using multidimensional fragility surfaces: A Bayesian framework, *Probabilistic Engineering Mechanics,* **25**, 49–60, 2010.

[11] C. Mai, K. Konakli, B. Sudret, Seismic fragility curves for structures using non-parametric representations, *Frontiers of Structural and Civil Engineering,* **11**, 169–186, 2017.

[12] K. Trevlopoulos, C. Feau, I. Zentner, Parametric models averaging for optimized non-parametric fragility curve estimation based on intensity measure data clustering, *Structural Safety,* **81**, 101865, 2019.

[13] F. Wang, C. Feau, Influenc of input motion's control point location in nonlinear SSI analysis of equipment seismic fragilities: Case study on the Kashiwazaki-Kariwa NPP, *Pure and Applied Geophysics,* **177**, 2391–2409, 2020.

[14] T. K. Mandal, S. Ghosh, N. N. Pujari, Seismic fragility analysis of a typical indian PHWR containment: Comparison of fragility models, *Structural Safety,* **58**, 11–19, 2016.

[15] Z. Wang, N. Pedroni, I. Zentner, E. Zio, Seismic fragility analysis with artificia neural networks: Application to nuclear power plant equipment, *Engineering Structures,* **162**, 213–225, 2018.

[16] Z. Wang, I. Zentner, E. Zio, A Bayesian framework for estimating fragility curves based on seismic damage data and numerical simulations by adaptive neural networks, *Nuclear Engineering and Design,* **338**, 232–246, 2018.

[17] C. Zhao, N. Yu, Y. Mo, Seismic fragility analysis of AP1000 SB considering fluid structure interaction effects, *Structures,* **23**, 103–110, 2020.

[18] Y. Katayama, Y. Ohtori, T. Sakai, H. Muta, Bayesian-estimation-based method for generating fragility curves for high-fideli y seismic probability risk assessment, *Journal of Nuclear Science and Technology,* **58**, 1220–1234, 2021.

[19] C. Gauchy, C. Feau, J. Garnier, Importance sampling based active learning for parametric seismic fragility curve estimation, 2021. URL: `https://arxiv.org/abs/2109.04323`, arXiv.

[20] A. Khansefid S. M. Yadollahi, F. Taddei, G. Müller, Fragility and comfortability curves development and seismic risk assessment of a masonry building under earthquakes induced by geothermal power plants operation, *Structural Safety,* **103**, 102343, 2023.

[21] S. Lee, S. Kwag, B.-s. Ju, On efficien seismic fragility assessment using sequential bayesian inference and truncation scheme: A case study of shear wall structure, *Computers & Structures,* **289**, 107150, 2023.

[22] M. Shinozuka, M. Q. Feng, J. Lee, T. Naganuma, Statistical analysis of fragility curves, *Journal of Engineering Mechanics,* **126**, 1224–1231, 2000.

[23] D. Lallemant, A. Kiremidjian, H. Burton, Statistical procedures for developing earthquake damage fragility curves, *Earthquake Engineering & Structural Dynamics,* **44**, 1373–1389, 2015.

[24] D. Straub, A. Der Kiureghian, Improved seismic fragility modeling from empirical data, *Structural Safety,* **30**, 320–336, 2008.

[25] P. Gardoni, A. Der Kiureghian, K. M. Mosalam, Probabilistic capacity models and fragility estimates for reinforced concrete columns based on experimental observations, *Journal of Engineering Mechanics,* **128**, 1024–1038, 2002.

[26] G. Damblin, M. Keller, A. Pasanisi, P. Barbillon, É. Parent, Approche décisionnelle bayésienne pour estimer une courbe de fragilité, *Journal de la Societe Française de Statistique,* **155**, 78–103, 2014.

[27] S. K. Tadinada, A. Gupta, Structural fragility of t-joint connections in large-scale piping systems using equivalent elastic time-history simulations, *Structural Safety,* **65**, 49–59, 2017.

[28] S. Kwag, A. Gupta, Computationally efficien fragility assessment using equivalent elastic limit state and Bayesian updating, *Computers & Structures,* **197**, 1–11, 2018.

[29] J.-S. Jeon, S. Mangalathu, J. Song, R. Desroches, Parameterized seismic fragility curves for curved multi-frame concrete box-girder bridges using bayesian parameter estimation, *Journal of Earthquake Engineering,* **23**, 954–979, 2019.

[30] A. Tabandeh, P. Asem, P. Gardoni, Physics-based probabilistic models: Integrating differential equations and observational data, *Structural Safety,* **87**, 101981, 2020.

[31] A. Van Biesbroeck, C. Gauchy, C. Feau, J. Garnier, Reference prior for bayesian estimation of seismic fragility curves, *Probabilistic Engineering Mechanics,* **76**, 103622, 2024.

[32] A. Van Biesbroeck, C. Gauchy, C. Feau, J. Garnier, Influenc of the choice of the seismic intensity measure on fragility curves estimation in a bayesian framework based on reference prior, in: *Proceedings of the 5th UNCECOMP Conference, 94–111,* Athens, Greece, 2023.

[33] A. V. Biesbroeck, C. Gauchy, C. Feau, J. Garnier, Robust a posteriori estimation of probit-lognormal seismic fragility curves via sequential design of experiments and constrained reference prior, 2025. URL: https://arxiv.org/abs/2503.07343. arXiv:2503.07343.

[34] J. M. Bernardo, Expected Information as Expected Utility, *The Annals of Statistics,* **7**, 686–690, 1979.

[35] N. Baillie, A. Van Biesbroeck, C. Gauchy, Variational inference for approximate reference priors using neural networks, 2025. URL: https://arxiv.org/abs/2502.02364. arXiv:2502.02364.

[36] J. O. Berger, J. M. Bernardo, D. Sun, The formal definitio of reference priors, *The Annals of Statistics,* **37**, 905–938, 2009.

[37] A. Van Biesbroeck, Generalized mutual information and their reference priors under csizar f-divergence, 2024. URL: https://arxiv.org/abs/2310.10530. arXiv:2310.10530, arXiv.

[38] B. S. Clarke, A. R. Barron, Jeffreys' prior is asymptotically least favorable under entropy risk, *Journal of Statistical Planning and Inference,* **41**, 37–60, 1994.

[39] A. Van Biesbroeck, Properly constrained reference priors decay rates for efficien and robust posterior inference, 2024. URL: https://arxiv.org/abs/2409.13041. arXiv:2409.13041, arXiv.

[40] D. P. Kingma, M. Welling, Auto-Encoding Variational Bayes, in: *Proceedings of the 2nd International Conference on Learning Representations (ICLR),* Banff, Canada, 2014.

[41] C. Gauchy, A. Van Biesbroeck, C. Feau, J. Garnier, Inférence variationnelle de lois a priori de référence, in: *Proceedings des 54èmes Journées de Statistiques (JdS),* SFDS, 2023.

[42] E. Nalisnick, P. Smyth, Learning approximately objective priors, in: *Proceedings of the 33rd Conference on Uncertainty in Artificial Intelligence (UAI), Association for Uncertainty in Artificial Intelligence (AUAI),* Sydney, Australia, 2017.

[43] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, Pytorch: An imperative style, high-performance deep learning library, in: *H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, R. Garnett (Eds.), Advances in Neural Information Processing Systems, volume 32,* Curran Associates, Inc., 2019.

[44] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, in: *Proceedings of the 3rd International Conference on Learning Representations (ICLR),* San Diego, USA, 2015.

[45] S. Rezaeian, A. Der Kiureghian, Simulation of synthetic ground motions for specifie earthquake and site characteristics, *Earthquake Engineering & Structural Dynamics,* **39**, 1155–1180, 2010.

[46] R. Sainct, C. Feau, J.-M. Martinez, J. Garnier, Efficien methodology for seismic fragility curves estimation by active learning on support vector machines, *Structural Safety,* **86**, 101972, 2020.

[47] N. N. Ambraseys, P. Smit, R. Berardi, D. Rinaldis, F. Cotton, C. Berge, Dissemination of european strongmotion data, 2000. CD-ROM collection. European Commission, Directorate-General XII, Environmental and Climate Programme, ENV4-CT97-0397, Brussels, Belgium.

[48] F. Touboul, P. Sollogoub, N. Blay, Seismic behaviour of piping systems with and without defects: experimental and numerical evaluations, *Nuclear Engineering and Design,* **192**, 243–260, 1999.

[49] CEA, CAST3M, 2019. URL: http://www-cast3m.cea.fr/.

[50] F. Touboul, N. Blay, P. Sollogoub, S. Chapuliot, Enhanced seismic criteria for piping, *Nuclear Engineering and Design,* **236**, 1–9, 2006.

[51] A. Winkelbauer, Moments and absolute moments of the normal distribution, 2014. URL: https://arxiv.org/abs/1209.4340. arXiv:1209.4340.