14th International Conference on Structural Safety and Reliability - ICOSSAR'25
1-6 June 2025 - Los Angeles California, USA
R. Ghanem, E. Johnson, S. Masri and A. Olivier (Eds)

A SURROGATE MODEL APPROACH FOR UNCERTAINTY TREATMENT IN PERFORMANCE-BASED ENGINEERING

MATÍAS BIRRELL*[1], RODRIGO ASTROZA[1] AND YONG LI[2]

[1]Universidad de los Andes
Av. Monseñor Alvaro del Portillo 12.455, Santiago, Chile
mbirrell@miuandes.cl
rastroza@miuandes.cl

[2]University of Alberta
9211 116 Street NW, Edmonton, AB, Canada
yong9@ualberta.ca

Key words: Uncertainty quantification and propagation, performance-based engineering, Gaussian process surrogate models, Global sensitivity analysis.

1. INTRODUCTION

Uncertainty quantification and propagation (UQ&P) is a key component of performance-based engineering (PBE). The PEER framework, widely accepted as the standard approach for PBE under earthquake hazards (PBEE), considers the probabilistic characterization of output variables of each of the four sequential analysis blocks (i.e., probabilistic seismic hazard analysis or PSHA, probabilistic seismic demand analysis or PSDA, probabilistic seismic fragility analysis or PSFA, and probabilistic seismic loss analysis or PSLA). Therefore, uncertainty at each stage must be adequately characterized and propagated to the following stages to establish a meaningful risk assessment.

One significant challenge in UQ&P is the high computational cost of extensive simulation performed on structure-level finite element (FE) models. To mitigate this cost, researchers have adopted machine and deep learning surrogate modeling techniques, in which a reduced set of data is used to train a cheap-to-evaluate prediction model. In this area, different approaches have been taken to replace massive FE simulations. These include structure-specific modeling of responses, direct modeling of fragility function (FF) parameters, and portfolio-oriented/regional-scale models. Each approach alludes to a certain scope of application and its associated level of detail modeling. For instance, while portfolio-oriented or regional-scale approaches are useful towards improving disaster management policies, structure-specific models might retain greater detail in their focus on structural dynamics and mechanics.

This work addresses the challenge at a structure-specific scale by proposing a surrogate model-based approach for UQ&P in PBEE which considers (i) probabilistic seismic hazard modeling and consistent ground motion selection, (ii) probabilistic structural modeling by considering parameter uncertainty in nonlinear material and components, (iii) development of surrogate models for model responses of interest, (iv) probabilistic seismic demand modeling including aleatoric and epistemic uncertainty arising from seismic and model parameter variability, as well as from the surrogate models themselves, (v) fragility analysis considering

the uncertainty in limit state and capacity definitions, and (vi) the decomposition of uncertainty in fragility estimates via global sensitivity analysis (GSA).

The approach is then validated through the application on an actual Chilean five-span reinforced concrete (RC) highway bridge featuring a suite of over 200 subductive ground motions selected from earthquakes in Chile and Japan. Probabilistic modeling of the bridge includes Bayesian calibration of nonlinear materials and components models. Gaussian Process (GP) surrogate models for seven structural responses are developed and validated for new earthquakes and model parameter realizations. Cloud fragility analysis including uncertainties in capacity and demand is performed, and the uncertainty propagated to fragility estimates is decomposed through total Sobol sensitivity indices.

2. METHODOLOGY

The proposed methodology can be divided into four main stages, as illustrated in Figure 1. In the first stage, a probabilistic model definition of the structure is established in terms of uncertain parameters as $\mathbf{Y}_{FE}(\boldsymbol{\Theta}) \in \mathbb{R}^{n_s \times n_r}$, with $\boldsymbol{\Theta} \in \mathbb{R}^{n_s \times n_p}$ being the matrix of FE model parameters realizations, for a total number of n_p model parameters and n_s samples; and n_r is the number of recorded model responses. In this scheme, $\boldsymbol{\Theta}$ is defined at a material and component level, where individual constitutive models are calibrated probabilistically beforehand to define their probability distributions (PDF).

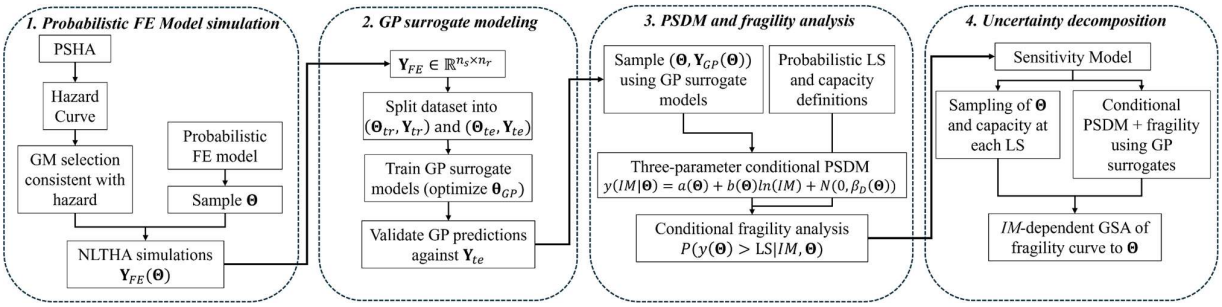


Figure 1. Proposed methodology.

Once the FE model has been defined, PSHA is performed for the structure with the purpose of establishing a target hazard curve (HC) or demand spectrum to serve as the reference for ground motion (GM) selection. This work considers a cloud-based selection, thus computing an aggregated HC according to the methodology proposed in [1], through Eqs. (1) and (2).

$$\lambda_r(IM > im) = \sum_{i=1}^{n_{sources}} \lambda_s(M_i > m) \int_{m_{min}}^{m_{max}} \int_0^{r_{max}} P(IM > im|m,r) f_{M_i R_i}(m,r) dr dm \tag{1}$$

$$\lambda(IM > im) = \sum_{k=1}^{n_{GMM}} w_k \lambda_{r_k}(IM > im) \tag{2}$$

Where $\lambda_r(IM > im)$ is the mean annual rate (MRE) of an intensity measure (IM) exceeding a certain value im; $n_{sources}$ and n_{GMM} are the number of seismic sources and GMMs considered, respectively; $\lambda_s(M_i > m)$ is the MRE of a seismic event of magnitude M_i exceeding a minimum considered magnitude m occurring in source i; m_{min} and m_{max} are the minimum

and maximum considered magnitudes for source i; f_{M_i,R_i} is the joint PDF of an event of magnitude M_i occurring at a distance R_i; and $\lambda(IM > im)$ is the resulting HC.

Eq. (1) is used to compute λ_r for a single ground motion model (GMM) in terms of geometrical and seismicity parameters (i.e., λ_s, $P(IM > im)$, f_{M_i,R_i}). The aggregate HC is obtained by applying a logic tree with weights w as per Eq. (2).

In stage 2, $\mathbf{Y}_{FE}(\boldsymbol{\Theta})$ is split into training $(\boldsymbol{\Theta}_{tr}, \mathbf{Y}_{tr})$ and testing $(\boldsymbol{\Theta}_{te}, \mathbf{Y}_{te})$ datasets. The selection of the training data is a crucial step to ensure adequate predictions on new scenarios (i.e., model parameter samples and GMs). Thus, a stratified selection approach is taken to produce an even coverage of all relevant data distributions. At this stage, the goal is to produce a training dataset of as few data points as possible while maintaining prediction accuracy on new data, thus minimizing the training time of the surrogate models and avoiding overfitting issues.

Then, independent GP [2] surrogate models are trained for each model response of interest. A GP model can be represented as a Gaussian regression as in Eq. (3), where a data point $\mathrm{y}(\mathbf{x})$ defined by a set of features $\mathbf{x} \in \mathbb{R}^{1\times n_f}$ is modeled by a GP of mean $\mu(\mathbf{x})$ and a correlation function (i.e., kernel) $k(\mathbf{x}, \mathbf{x}')$, as well as a Gaussian discrepancy term ε. Predictions on new data, after the model has been trained with a set of features $\mathbf{X}_{tr}$, are obtained by marginalization of the distribution where the posterior conditional prediction is given by Eqs. (4) to (6). As GPs are commonly defined, without loss of generality, to be zero-meaned (i.e., $\mu_{GP}(\mathbf{x}) = 0$), the issue of training a GP surrogate model lies on defining an appropriate kernel. This is done by log-likelihood minimization using the L-BFGS [3] algorithm.

In Eqs. (4) to (6), $\mathrm{y}(\mathbf{x}^*) \in \mathbb{R}$ and $\mathbf{y}(\mathbf{X}_{tr}) \in \mathbb{R}^{n_{tr}\times 1}$ are the GP prediction on a new data point $\mathbf{x}^* \in \mathbb{R}^{1\times n_f}$, and $\mathbf{X}_{tr} \in \mathbb{R}^{n_{tr}\times n_f}$, respectively, with n_{tr} the number of training data points. Similarly, $\mathbf{k}(\mathbf{x}^*, \mathbf{X}_{tr}) \in \mathbb{R}^{1\times n_{tr}}$ and $\mathbf{K}(\mathbf{X}_{tr}, \mathbf{X}_{tr}) \in \mathbb{R}^{n_{tr}\times n_{tr}}$ are the corresponding kernel evaluations derived during the marginalization, and $\mathbf{I} \in \mathbb{R}^{n_{tr}\times n_{tr}}$ is the identity matrix. As a result of this process, the mean prediction $\mu_{GP}(\mathbf{x}^*) \in \mathbb{R}$ and its associated variance $\sigma^2_{GP}(\mathbf{x}^*) \in \mathbb{R}$ are obtained.

$$\mathrm{y}(\mathbf{x}) = f(\mathbf{x}) \sim GP\big(\mu(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')\big) + \varepsilon \tag{3}$$

$$\mathrm{y}(\mathbf{x}^*)|\mathbf{x}^*, \mathbf{X}_{tr}, \mathbf{y}(\mathbf{X}_{tr}) \sim N(\mu_{GP}(\mathbf{x}^*), \sigma^2_{GP}(\mathbf{x}^*)) \tag{4}$$

$$\mu_{GP}(\mathbf{x}^*) = \boldsymbol{k}(\mathbf{x}^*, \mathbf{X}_{tr})^{\mathrm{T}}(\mathbf{K}(\mathbf{X}_{tr}, \mathbf{X}_{tr}) + \sigma_n^2\mathbf{I})^{-1}\mathbf{y}(\mathbf{X}_{tr}) \tag{5}$$

$$\sigma^2_{GP}(\mathbf{x}^*) = k(\mathbf{x}^*, \mathbf{x}^*) - \boldsymbol{k}(\mathbf{x}^*, \mathbf{X}_{tr})(\mathbf{K}(\mathbf{X}_{tr}, \mathbf{X}_{tr}) + \sigma_n^2\mathbf{I})^{-1}\boldsymbol{k}(\mathbf{X}_{tr}, \mathbf{x}^*) \tag{6}$$

To validate the accuracy of the trained GP surrogates, the coefficient of determination R^2 is computed on the testing dataset of n_{te} points, according to Eq. (7), where the predictions $\mathbf{y}^{te}_{\mathrm{GP}}$ are compared to the FE model responses $\mathbf{y}^{te}_{\mathrm{FE}}$. It is noted that the notation $\mathbf{X}$ is used for the set of features used to define the GP models, as these are not necessarily the same as the FE model parameters $\boldsymbol{\Theta}$.

$$R^2(\mathbf{y}^{te}_{\mathrm{FE}}, \mathbf{y}^{te}_{\mathrm{GP}}) = 1 - \frac{\sum_{i=1}^{n_{te}}\left(\mathrm{y}^{te}_{\mathrm{FE},i} - \mathrm{y}^{te}_{\mathrm{GP},i}\right)^2}{\sum_{i=1}^{n_{te}}\left(\mathrm{y}^{te}_{\mathrm{FE},i} - \mu_{\mathbf{y}^{te}_{\mathrm{FE}}}\right)^2} \tag{7}$$

After validation of the GP surrogate models, these can be used for sampling in place of the FE model. In stage 3, $\boldsymbol{\Theta}$ is sampled for a total of n_n new points, and the GPs are evaluated. In parallel, for each sample, the corresponding constitutive models are evaluated to determine the distribution of component capacities in log-space (i.e., $\mu_{c,LS},\ \beta_{c,LS}$) at each constitutive-level-defined limit states (LS). Engineering demand parameters (EDPs) are computed as normalized quantities relative to these LS (e.g., drift ductility ratio). Using Eq. (8), the three-parameter linear PSDM in log-space can be defined probabilistically, conditioned to $\boldsymbol{\Theta}$ for cloud-based PSFA [4].

$$\delta_{PSDM}(IM|\boldsymbol{\Theta}) = a(\boldsymbol{\Theta}) + b(\boldsymbol{\Theta})\log(IM) + N(0, \beta^2_{D|\boldsymbol{\Theta}}) \tag{8}$$

Where $a(\boldsymbol{\Theta})$ and $b(\boldsymbol{\Theta})$ are the linear regression coefficients and $\beta^2_{D|\boldsymbol{\Theta}} = \mathrm{Var}(\delta_{PSDM}(IM|\boldsymbol{\Theta}) - \delta(\boldsymbol{\Theta}))$ is the variance of residuals with respect to the actual EDP value $\delta(\boldsymbol{\Theta})$. In the case of GPs, the predictive variance needs to be accounted for as an epistemic source of demand uncertainty, thus contributing to $\beta^2_{D|\boldsymbol{\Theta}}$. The transformation into log-space and applied to the PSDM is defined as $\beta^2_{GP}|\mathbf{X} = \mathrm{median}\left(\log\left(1 + \frac{\sigma^2_{GP}}{\mu^2_{GP}}\right)\right)$. This term is added to $\beta^2_{D|\boldsymbol{\Theta}}$ in GP-based PSDMs.

Having defined the PDFs for demand, capacity and LS, the conditional fragility curves can be computed by evaluating Eq. (9).

$$P(\delta > \mu_{c,LS}|\boldsymbol{\Theta}) = \Phi\left(\frac{\delta_{PSDM}(IM|\boldsymbol{\Theta}) - \log(\mu_{c,LS})}{\beta_{tot}|\boldsymbol{\Theta}}\right) \tag{9}$$

where $\beta_{tot}|\boldsymbol{\Theta} = \sqrt{\beta^2_{LS} + \beta^2_{c,LS} + \beta^2_{D|\boldsymbol{\Theta}}}$ is the total combined uncertainty, including LS, capacity, and demand definitions.

In stage 4, the uncertainty in FFs is decomposed in terms of $\boldsymbol{\Theta}$ by performing global sensitivity analysis (GSA) of each EDP to each parameter using sample-based total Sobol indices (S_i^T) [5]. This study considers *T*-estimator proposed by Janon et al. [6], which is given by Eq. (10), to compute the sampling-based indices using GP surrogates.

$$T^p_{N_s}(\boldsymbol{y}) = \frac{\frac{1}{N_s}\sum_{s=1}^{N_s}\boldsymbol{y}_s\boldsymbol{y}_s^p - \left(\frac{1}{N_s}\sum_{s=1}^{N_s}\left[\frac{\boldsymbol{y}_s + \boldsymbol{y}_s^p}{2}\right]\right)^2}{\frac{1}{N_s}\sum_{s=1}^{N_s}\left[\frac{\boldsymbol{y}_s{}^2 + (\boldsymbol{y}_s^p)^2}{2}\right] - \left(\frac{1}{N_s}\sum_{s=1}^{N_s}\left[\frac{\boldsymbol{y}_s + \boldsymbol{y}_s^p}{2}\right]\right)^2} \tag{10}$$

Where N_s is the total number of Monte-Carlo (MC) samples taken to compute the indices and p is the index order to be computed. For S_i^T, $p = {\sim}i$ corresponds to every model parameter other than i. It is noted that when $\boldsymbol{y}$ is a vector response, such as a fragility curve evaluated at n_{im} points where $\boldsymbol{y} \in \mathbb{R}^{1\times n_{im}}$, a "history" of indices is obtained such that $S_i^T(im) \in \mathbb{R}^{1\times n_{im}}$.

3. APPLICATION EXAMPLE: AGUILA NORTE BRIDGE

The methodology is validated through the study of the Aguila Norte (AN) bridge. It is a five-span RC highway bridge which has been studied extensively due to its representativity of typical Chile bridge design practices. In this work, a fixed-base version of the AN model, developed in

OpenSees, is used (Figure 2). The application of the proposed methodology and the results obtained are reported in the following sections.

3.1 FE MODEL

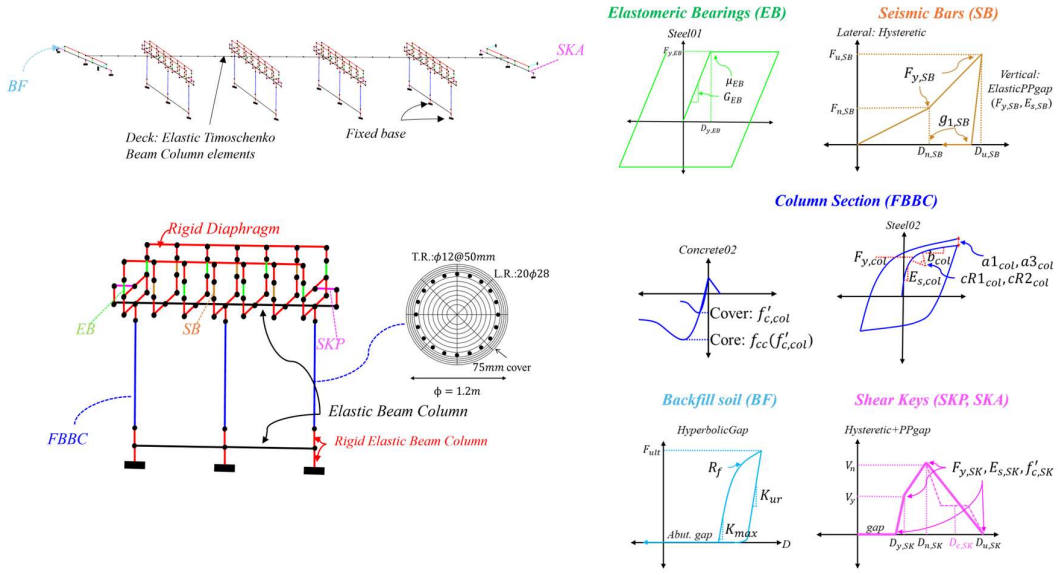


Figure 2. FE model of the AN bridge.

The modeling strategy for AN consists of elastic elements for the deck (elastic Timoschenko beam columns), as well as cap and foundation beams at each pier and abutments (elastic beam columns). Nonlinear constitutive models are used for pier columns (i.e., *Concrete02* and *Steel02* materials and force-based beam-column or *FBBC* elements), elastomeric bearings (i.e., EB; *Steel01* material and *twoNodeLink* elements), vertical anchoring seismic bars (i.e., SB; *Hysteretic* material and *twoNodeLink* elements), shear keys (i.e., SKP, SKA; *Hysteretic* + *elasticPPgap* materials and *zeroLength* elements), and backfill soil (i.e., BF; H*yperbolicGap* material and *zeroLength* elements). Further detail on the bridge model can be found in [7].

3.2 MODEL PARAMETERS

As part of stage 1, parametric uncertainty was quantified for the nonlinear constitutive models through Bayesian calibration using representative experimental data. From the Bayesian framework, a multivariate PDF for each model was obtained. The marginal PDF for each of the 20 parameters in $\boldsymbol{\Theta}$ are summarized in Table 1, although the correlation matrices are omitted for brevity. Complete detail of each model and their calibration can be found in the following publications: [8] for *Steel02*, [9] for *Concrete02*, [10] for EB and SB, [11, 12] for SKP and SKA. The constitutive model of BF was not sampled due to lack of relevant experimental data.

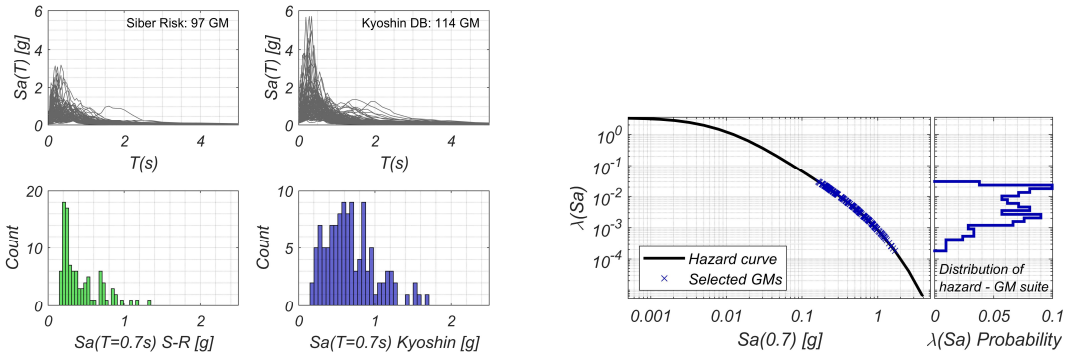
Table 1. Model parameter PDFs.

Parameter	Mean (μ)	C.O.V. (%)	Distribution	Component
$F_{y,col}$ (MPa)	467	10	Lognormal	Columns (*Steel02*)
$F_{u,col}$ (MPa)	703	4.3	Lognormal	

$E_{s,col}$ (GPa)	206	4	Lognormal	
b_{col}	0.018	18	Lognormal	
$cR1_{col}$	0.889	2	Lognormal	
$cR2_{col}$	0.095	27	Lognormal	
$a1_{col}$	0.039	28	Lognormal	
$a3_{col}$	0.029	28	Lognormal	
$f'_{c,col}$ (MPa)	25	20	Lognormal	Columns (*Concrete02*)
G_{EB} (MPa)	1.176	3	Lognormal	EB
μ_{EB}	0.23	2.6	Lognormal	
$F_{y,SB}$ (MPa)	206.4	6.8	Lognormal	SB
$g_{1,SB}$	0.104	27.6	Lognormal	
$E_{s,SB}$ (GPa)	206	4	Lognormal	
$F_{y,SKP}$ (MPa)	467	10	Lognormal	SKP
$E_{s,SKP}$ (GPa)	206	4	Lognormal	
$f'_{c,SKP}$ (MPa)	25	20	Lognormal	
$F_{y,SKA}$ (MPa)	467	10	Lognormal	SKA
$E_{s,SKA}$ (GPa)	206	4	Lognormal	
$f'_{c,SKA}$ (MPa)	25	20	Lognormal	

3.3 PSHA AND GM SELECTION

PSHA is conducted based on the Chilean subduction zone geometry defined by the *Slab*2 model [13], seismicity and GMMs defined for Chile by Poulos et al. [14], Montalva et al. [15], and Idini et al. [16] using a 50/50 logic tree. The hazard curve shown in Figure 3b is then obtained using Eqs. (1) and (2).



(a) (b)

Figure 3. Selected GMs from the S-R and Kyoshin databases and associated hazard levels.

The GM selection for later cloud-based PSFA considered the Siber-Risk Chilean [17] and the Kyoshin Japanese [18] databases. From each, subductive records with a spectral acceleration at the fundamental period, $Sa(T_1)$ with T_1=0.7s, and an average shear wave velocity in the upper 30 m (Vs_{30}) between 180 and 760 m/s were selected. A total of 97 and 114 GMs were selected from each database, respectively. Their RotD50 spectra and histograms of their $Sa(T_1)$ are shown in Figure 3a, while their associated hazard level in relation to the PSHA is shown along

the hazard curve in Figure 3b. The selected suite of unscaled GMs covered a range of return periods (T_r) between T_r=30 years and T_r=5,000 years, thus addressing service, design, severe and MCE hazard levels.

3.4 GP SURROGATE MODELS

A total of 10,550 FE model runs were executed, consisting of all combinations of 50 samples from $\boldsymbol{\Theta}$ drawn using Sobol sequences and all 211 GMs. Seven peak time-history responses were recorded, corresponding to the maximum among peak responses from all elements of the same type (e.g., maximum among the peaks of all 12 pier columns). The recorded responses were longitudinal and lateral curvature at column bases (i.e., $K_{col,lon}, K_{col,lat}$), shear deformation in EB and SB (i.e., Δ_{EB}, Δ_{SB}), lateral displacement in SKP and SKA (i.e., $\Delta_{SKP}, \Delta_{SKA}$) and longitudinal displacement at BF (i.e., Δ_{BF}). These 7 responses captured the parametric uncertainty quantified for all nonlinear constitutive models and propagated to nonlinear time-history analysis (NLTHA).

Subsequently, the data was split into training and testing sets using a two-stage stratified approach. In the first step, 33 out of the 50 samples from $\boldsymbol{\Theta}$ were used for training. From the GM suite, the maximum between 25 and the number of GMs at each bin of 0.1g increments in $Sa(T_1)$ were taken. This produced the first subset of 4,653 samples. In the second step, 20% of the data points were selected randomly, producing a final training set of n_{tr}=987 samples. The remaining n_{te}=9,563 samples were kept for testing.

To define the GP surrogate models, the kernel was defined using a Matérn 3/2 kernel with ARD, which can be referred to in detail in [2]. To capture parametric and input excitation variability, the set of features in $\mathbf{X}$ was defined by $\boldsymbol{\Theta}$ and 50 values of $Sa(T)$ from the RotD50 spectra, producing a set of 70 features. The models were trained using Eqs. (3) to (6). Subsequently, the R^2 of $\mathbf{Y}_{tr}$ and $\mathbf{Y}_{te}$ were computed to validate the prediction accuracy of the trained GPs. Regression plots for the testing data are shown in Figure 4, where $R^2(\mathbf{Y}_{te}) > 0.9$ was obtained for every response, while obtaining perfect fits for $\mathbf{Y}_{tr}$. Such fits on training data could be explained by the use of ARD kernels, and overfitting concerns were discarded due to the excellent predictions on the testing set.

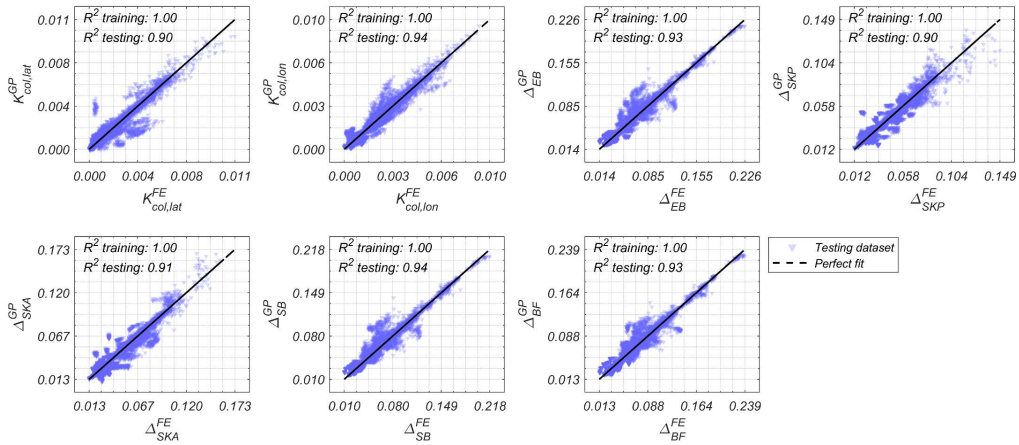


Figure 4. GP model predictions on testing dataset.

4. FRAGILITY ANALYSIS

Stage 3 of the proposed approach aims to propagate uncertainties in structural responses into FFs. To this end, cloud-based PSFA is established in terms of the log-linear three-parameter PSDM defined in Eq. (8) and computing the conditional FF from Eq. (9). The formulation includes uncertainty in EDPs, LSs, and component capacity at each LS.

4.1 LS AND CAPACITY DEFINITIONS

Regarding LS for each EDP, the definitions from [19] were adopted for columns, EB, and BF. For SK and SB, they were defined in terms of critical points from their respective constitutive models depicted in Figure 2. LS definitions are summarized in Table 2.

Table 2. LS definitions.

LS	$\boldsymbol{K_{col}}$	$\boldsymbol{\Delta_{EB}}$	$\boldsymbol{\Delta_{SK}}$	$\boldsymbol{\Delta_{SB}}$	$\boldsymbol{\Delta_{BF}}$
1	K_y K_y: yield curvature.	$0.2H_{EB}$ H_{EB}: height of the bearing	$D_{y,SK}$	$D_{y,SB}$	gap gap: between abutment and deck
2	$\min(K: \varepsilon_c > 0.004, K: \varepsilon_s > 0.015)$ ε_c: concrete strain. ε_s: steel strain.	H_{EB}	$D_{n,SK}$	$D_{u,SB}$	$0.0093H_w$ H_w: Height of the backwall
3	$\min(K: \varepsilon_c > 0.004 + 1.4\frac{\rho_w f_{yw}}{f'_{cc}}, K: \varepsilon_s > 0.06)$ ρ_w, f_{yw}: reinforcement ratio and yield strength of transverse hoops. f'_{cc}: confined concrete strength.	$2H_{EB}$	$D_{c,SK}$	-	$0.037H_w$
4	$\min(K: M < 0.9M_{max}, K: \varepsilon_s > 0.075)$ M, M_{max}: moment and maximum moment from sectional analysis.	$3H_{EB}$	-	-	$0.081H_w$

For columns, EB and BF responses, locally defined EDPs have been discussed in previously published literature. Relative-deformation metrics are commonly used. This work adopted the definitions proposed by Stefanidou et al. [19] for those components. As EDPS for SK and SB have not been widely discussed, they have been defined for this work in a similar fashion to those of EB, as drift ductility metrics. All EDP definitions are summarized in Table 3. For convenience, the EDPs have also been defined herein as a normalized quantity by dividing by their LS1.

Table 3. EDP definitions.

EDP	Columns	EB	SK	SB	BF
Definition	Curvature ductility	Shear deformation ductility	Drift ductility	Drift ductility	Drift
EDP	$\phi = \frac{K_{col}}{K_{y,col}}$	$\delta_{EB} = \frac{\Delta_{EB}}{0.2H_{EB}}$	$\delta_{SK} = \frac{\Delta_{SK}}{D_{y,SK}}$	$\delta_{SB} = \frac{\Delta_{SB}}{D_{y,SB}}$	$\delta_{BF} = \frac{\Delta_{BF}}{H_w}$

Capacity points at each EDP were obtained by evaluating the respective constitutive models on 10,000 new samples of $\boldsymbol{\Theta}$, and transformed to log-space to obtain μ_c and β_c for each case. For the LSs adopted from [19], β_{LS} was readily available. In contrast, for the LS defined in this work there was no uncertainty in LS definition, thus $\beta_{LS} = 0$ is assumed in those cases. Moreover, $\beta_c = 0$ for EB and BF, as the former is defined in terms of the unchanging bearing

height, and the latter's constitutive model parameters were constant. The resulting values for $\mu_c, \beta_c, \beta_{LS}$ are summarized in Table 4.

Table 4. Sampled capacity parameters for each EDP.

	$\boldsymbol{\phi_{col}}$			$\boldsymbol{\delta_{EB}}$			$\boldsymbol{\delta_{SKP}}$			$\boldsymbol{\delta_{SKA}}$			$\boldsymbol{\delta_{SB}}$			$\boldsymbol{\delta_{BF}}$		
	μ_C	β_{LS}	β_C	μ_C	β_{LS}	β_C	μ_C	β_{LS}	β_C	μ_C	β_{LS}	β_C	μ_C	β_{LS}	β_C	μ_C	β_{LS}	β_C
LS1	0	0.6	0.083	0	0.61	0	0	0	0.16	0	0	0.13	0	0	0.18	0	0.53	0
LS2	1.72	0.53	0.0102	1.61	0.22	0	2.12	0	0.16	2.33	0	0.13	2.28	0	0	0.10	0.22	0
LS3	3.07	0.6	0.0065	2.30	0.22	0	3.08	0	0.05	3.30	0	0.04	-	-	-	0.34	0.22	0
LS4	3.29	0.5	0.0075	2.71	0.32	0	-	-	-	-	-	-	-	-	-	0.64	0.32	0

4.2 CONDITIONAL FRAGILITY FUNCTIONS

Following the definition of PSDM quantities, the conditional FFs were computed using Eq. (9). Figure 5 shows the plots for all cases, where the mean and 95% confidence interval (CI) curves obtained from the GPs evaluated on 1,000 new parameter samples are compared to those obtained from the original FE model evaluated on $\boldsymbol{\Theta}$ and the 211 GMs. This comparison served as further validation for the feasibility of using the GPs for sampling-based analysis, including GSA in the following section. The lines in gray show the 1,000 individual curves from the GPs (i.e., $P(\delta > \mu_{c,LS} | \mathbf{X}_{new})$). It is worth noting the overlap between the FF for LS1 and LS2 of δ_{BF}, which was due to the noticeable difference in β_{LS1} and β_{LS2} for BF.

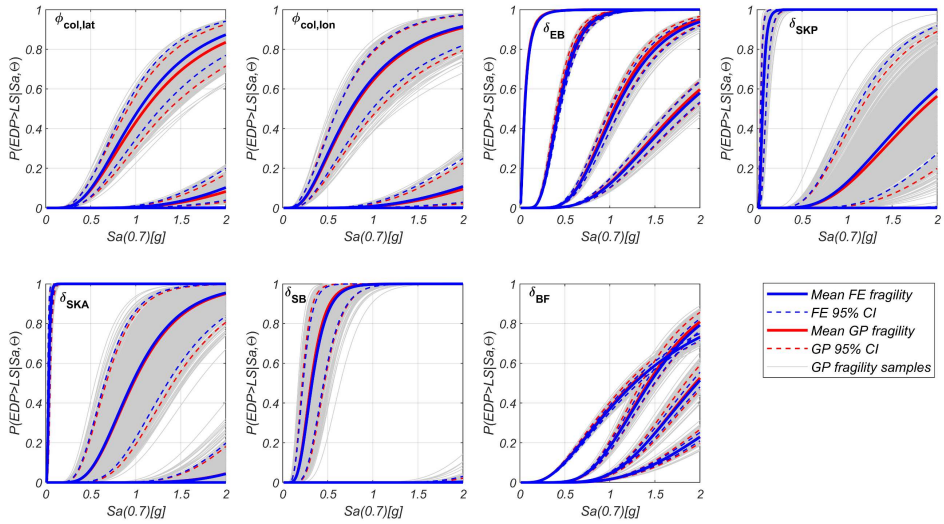


Figure 5. Conditional fragility curves for each EDP.

Overall, the considerable variability in FFs across LSs and components evidence the impact of parametric uncertainty in PSFA, and consequently in performance assessment. Hence, parametric uncertainty in structural materials and components should be accounted for carefully in PBE, both from the perspective of structural responses and LS/capacity estimates. This served as motivation to study the composition of these uncertainties through GSA, leveraging the low sampling cost of GP surrogates.

5. IM-DEPENDENT GSA OF MODEL PARAMETERS

The resulting FFs from stage 3 of the methodology exhibited considerable variability due to parametric uncertainty. However, this uncertainty has been aggregated by the propagation through NLTHA and capacity estimates. Thus, the purpose of stage 4 is to disaggregate the uncertainty in FFs in terms of parametric variability. This is achieved through sampling-based GSA of each FF using GP surrogate models.

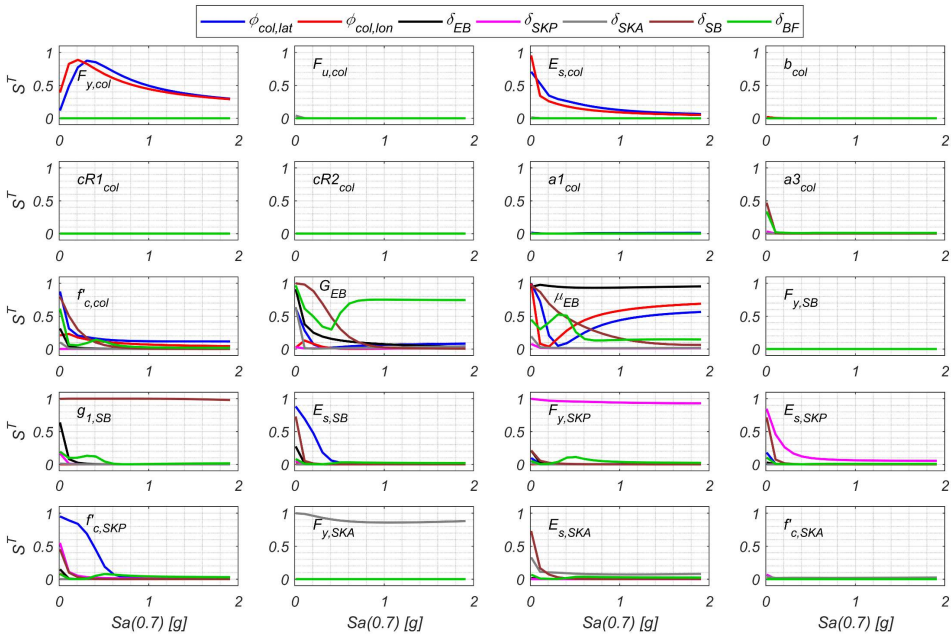


Figure 6. GP-based Sobol total indices for each EDP.

To obtain S^T for each EDP numerically, Eq. (10) was used considering the following GSA sampling procedure to obtain $\boldsymbol{y}_s$: (i) generate a new sample of $\boldsymbol{\Theta}$ using Sobol sequences, (ii) evaluate μ_c for the sample, (iii) obtain the predicted response from the corresponding GP for all GMs, and (iv) evaluate the conditional FF on the sample. The resulting $S^T(S_a) \in \mathbb{R}^{1\times 211}$ for each parameter reflects the total portion of variance in the FF that can be attributed to the parameter's variance, at each level of $Sa(T_1)$. As a reference, the computation of S^T in this work required 21×10^6 GP evaluations, with a total computation time of about nine hours on a laptop with a 12-core AMD Ryzen 9 7845HX CPU and 32GB of RAM. This number of simulations on the FE model would not have been feasible within a reasonable time.

The indices obtained are shown in Figure 6, from which key takeaways can be made. For instance, due to Chilean design practices, pier columns are rarely demanded beyond their yield curvature, hence the sensitivity of ϕ_{col} (in both directions) to the nonlinear parameters of *Steel02* is negligible. In turn, the parameters controlling the elastic column response (i.e., F_y, E_s, f_c') control the variance of ϕ_{col} at low S_a values. As S_a increases, the demand on pier columns is increasingly controlled by the top displacement induced by the deck, at which point the influence of other components' parameters on ϕ_{col} becomes apparent (most notably μ_{EB}).

Similar inference can be made about parameter-induced uncertainty on other EDPs, although omitted herein due to space limitations.

6. CONCLUSIONS

Parameter-induced uncertainty is significant in the context of PBE. Thus, careful consideration should be given to (i) gathering relevant experimental data, (ii) properly characterizing probabilistic structural models, (iii) defining LSs and considering parametric uncertainty in capacity computations, and (iv) selecting the most relevant model parameters in terms of performance assessment. The proposed approach leverages GP surrogate models to establish an efficient workflow for this purpose.

Beyond the specific methods used in this work (i.e., Bayesian calibration, GP surrogate modeling, site-specific PSHA, cloud-based PSFA), different methods could be adopted successfully, as well as extended for other hazards and structure types. The methods used were selected due to their inherent probabilistic definitions, in direct relationship with a probabilistic framework for PBE.

Finally, establishing a comprehensive approach for UQ&P in PBE and its decomposition in terms of model parameters can improve resource allocation for data collection campaigns, either through inspection or experimental testing. For instance, the results presented in this paper suggest that a thorough characterization of nonlinear *Steel02* parameters might not be necessary in the context of typical Chilean highway bridges.

Acknowledgements

This work was funded by the Chilean National Agency for Research and Development (ANID) through the projects FOVI230030 and National Doctoral Scholarship 21210182.

References

[1] J. W. Baker, B. A. Bradley, and P. J. Stafford, Probabilistic seismic hazard and risk analysis. Cambridge: Cambridge University Press, 2021. doi: 10.1017/9781108425056.

[2] C. E. Rasmussen and C. K. I. Williams, Gaussian Processes for Machine Learning. The MIT Press, 2005. doi: 10.7551/mitpress/3206.001.0001.

[3] D. C. Liu and J. Nocedal. On the limited memory BFGS method for large scale optimization, *Mathematical Programming*, vol. 45, 1-3, pp. 503–528, 1989, doi: 10.1007/BF01589116.

[4] F. Jalayer, R. de Risi, and G. Manfredi. Bayesian Cloud Analysis: efficient structural fragility assessment using linear regression, *Bulletin of Earthquake Engineering*, vol. 13, no. 4, pp. 1183–1203, 2015, doi: 10.1007/s10518-014-9692-z.

[5] I. M. Sobol. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, *Mathematics and Computers in Simulation*, vol. 55, 1-3, pp. 271–280, 2001.

[6] A. Janon, T. Klein, A. Lagnoux, M. Nodet, and C. Prieur. Asymptotic normality and efficiency of two Sobol index estimators, *ESAIM: Probability and Statistics*, vol. 18, pp. 342–364, 2014, doi: 10.1051/ps/2013040.

[7] C. Pastén, R. Astroza, R. Bazáez, N. Contreras, J. Grand, F. Hernández, and F. Ochoa, Guia para el análisis sísmico no lineal de puentes chilenos (In Spanish). Santiago, Chile: U. de los Andes; U. de Chile; UTSFM, 2021.
[8] M. Birrell, R. Astroza, R. Carreño, J. I. Restrepo, and G. Araya-Letelier. Bayesian parameter and joint probability distribution estimation for a hysteretic constitutive model of reinforcing steel, *Structural Safety*, vol. 90, no. 16, 2021, doi: 10.1016/j.strusafe.2020.102062.
[9] S. A. Mirza, J. G. MacGregor, and M. Hatzinikolas. Statistical Descriptions of Strength of Concrete, *Journal of the Structural Division*, vol. 105, no. 6, pp. 1021–1037, 1979, doi: 10.1061/JSDEAG.0005161.
[10] F. J. Pinto, J. Toledo, M. Birrell, R. Bazáez, F. Hernández, and R. Astroza. Uncertainty Quantification in Constitutive Models of Highway Bridge Components: Seismic Bars and Elastomeric Bearings, *Materials*, vol. 16, no. 5, 2023, doi: 10.3390/ma16051792.
[11] Rakesh K. Goel and Anil K. Chopra. Role of Shear Keys in Seismic Behavior of Bridges Crossing Fault-Rupture Zones, *Journal of Bridge Engineering*, vol. 13, pp. 398–408, 2008.
[12] M. Birrell, R. Astroza, J. Murcia-Delso, F. Hernández, and R. Bazáez. A simplified tri-linear model for monolithic exterior shear keys failing in sliding shear, *Structures*, vol. 61, no. 14, p. 105934, 2024, doi: 10.1016/j.istruc.2024.105934.
[13] G. P. Hayes, G. L. Moore, D. E. Portner, M. Hearne, H. Flamme, M. Furtney, and G. M. Smoczyk. Slab2, a comprehensive subduction zone geometry model, *Science (New York, N.Y.)*, early access. doi: 10.1126/science.aat4723.
[14] A. Poulos, M. Monsalve, N. Zamora, and J. C. de La Llera. An Updated Recurrence Model for Chilean Subduction Seismicity and Statistical Validation of Its Poisson Nature, *Bulletin of the Seismological Society of America*, vol. 109, no. 1, pp. 66–74, 2019, doi: 10.1785/0120170160.
[15] G. Montalva, N. Bastias, and A. Rodriguex-Marck. Ground motion prediction equation for the Chilean subduction zone, *Bulletin of the Seismological Society of America*, vol. 107, no. 2, pp. 901–911, 2017.
[16] B. Idini, F. Rojas, S. Ruiz, and C. Pastén. Ground motion prediction equation for the Chilean subduction zone, *Bulletin of Earthquake Engineering*, vol. 15, pp. 1853–1880, 2017.
[17] S. Castro, R. Benavente, J. G. F. Crempien, G. Candia, and J. C. de La Llera. A Consistently Processed Strong-Motion Database for Chilean Earthquakes, *Seismological Research Letters*, vol. 93, no. 5, pp. 2700–2718, 2022, doi: 10.1785/0220200336.
[18] National Research Institute for Earth Science and Disaster Resilience, NIED K-NET, KiK-net. National Research Institute for Earth Science and Disaster Resilience, doi: 10.17598/NIED.0004. [Online]. Available: www.kyoshin.bosai.go.jp
[19] S. P. Stefanidou, E. A. Paraskevopoulos, V. K. Papanikolaou, and A. J. Kappos. An online platform for bridge-specific fragility analysis of as-built and retrofitted bridges, *Bulletin of Earthquake Engineering*, vol. 20, no. 3, pp. 1717–1737, 2022, doi: 10.1007/s10518-021-01299-3.