UNIVERSIDAD DE LOS ANDES
FACULTAD DE INGENIERÍA Y CIENCIAS APLICADAS

EFFICIENT UNCERTAINTY QUANTIFICATION AND PROPAGATION IN PERFORMANCE-BASED EARTHQUAKE ENGINEERING

MATÍAS BIRRELL ARANGUA

TESIS PARA OPTAR AL GRADO DE

DOCTORADO EN CIENCIAS DE LA INGENIERÍA

Director: RODRIGO ASTROZA EULUFÍ

SANTIAGO, ABRIL DE 2025

Ir Por Más

RESUMEN

En las últimas décadas, el constante deterioro de la infraestructura existente y la creciente exposición a amenazas naturales provocadas por los procesos geológicos y condiciones climáticas cambiantes ha motivado el desarrollo de marcos metodológicos que constituyen una evolución en la filosofía de la ingeniería estructural. Esta nueva corriente filosófica se conoce como ingeniería estructural basada en el desempeño. Tiene como objetivo proporcionar un marco de trabajo riguroso y fundamentado en la ciencia, a través de una evaluación integral del riesgo estructural y entregando como producto final una variable de decisión.

Con este objetivo, la ingeniería basada en el desempeño establece un marco de trabajo probabilístico que busca tratar la incertidumbre acerca de (i) las amenazas a las que se somete la estructura, (ii) el comportamiento real de la estructura versus lo predicho por el modelo ingenieril, y (iii) los daños provocados en la estructura al superarse determinados niveles de intensidad. En cada una de estas etapas es crucial, para el éxito de la evaluación del riesgo, cuantificar de buena manera la incertidumbre existente, y posteriormente propagarla de manera secuencial a las etapas siguientes.

En este sentido, el avance metodológico ha sido progresivo, y acompañado del desarrollo tecnológico que ha permitido la implementación de métodos probabilísticos. El costo asociado a la adopción de un marco probabilístico ha sido alto, especialmente debido a la necesidad de simulación masiva de modelos de elementos finitos, lo cual requiere alta inversión computacional y de tiempo. Por esta razón, el desarrollar metodologías que permitan la cuantificación y propagación de incertidumbre en la ingeniería basada en el desempeño de manera eficiente, es un desafío que sigue abierto y motiva gran parte de la investigación actual en el área.

De esta manera, esta tesis presenta dos enfoques que buscan proporcionar métodos eficientes para la cuantificación y propagación de incertidumbre apoyando la simulación de modelos estructurales con modelos surrogados de machine learning usando procesos gaussianos.

El primero se enfoca en la cuantificación y descomposición de incertidumbre inducida en las respuestas estructurales por los parámetros del modelo, bajo escenarios específicos de amenaza. Este enfoque tiene como objetivo de aplicación facilitar análisis probabilísticos basados en muestreo, como calibración y actualización de modelos, diseño iterativo basado en desempeño, análisis de sensibilidad, entre otros.

El segundo se enfoca en la cuantificación, propagación y descomposición de incertidumbre en la evaluación de la vulnerabilidad estructural frente a un rango amplio de eventos sísmicos. Este enfoque implementa y discute el marco de trabajo de la ingeniería basada en el desempeño desde un punto de vista filosófico, aunque se aplica a un caso de estudio real. Se discuten las definiciones disponibles sobre estados de daño en componentes de puentes y las relaciones que existen entre éstos y sus consecuencias.

Ambos enfoques se desarrollan de manera íntegra incluyendo análisis probabilístico de la amenaza sísmica, modelación estructural probabilística y descomposición de incertidumbre.

ABSTRACT

In recent decades, the constant deterioration of existing infrastructure and the increasing exposure to natural hazards driven by geological processes and changing climate conditions have motivated the development of a new philosophical approach to structural engineering, known as performance-based engineering. Its goal is to provide a rigorous, science-based framework through a comprehensive assessment of structural risk, ultimately delivering a decision variable that is useful for practical decision-making.

To this end, performance-based engineering establishes a probabilistic framework that aims to address uncertainty regarding (i) the hazards to which the structure is exposed, (ii) the actual behavior of the structure versus that predicted by the engineering model, and (iii) the damage caused when certain intensity levels are exceeded. At each of these stages, properly quantifying uncertainty and subsequently propagating it through the following stages is critical for a successful risk assessment.

In this context, methodological progress has been gradual, supported by technological advances that have enabled the implementation of probabilistic methods. However, the cost of adopting a probabilistic framework has been high, especially due to the need for large-scale simulation of finite element models, which requires significant computational and time investment. For this reason, developing methods that enable efficient uncertainty quantification and propagation in performance-based engineering remains an open challenge and a key area of current research.

This thesis presents two approaches aimed at providing efficient methods for uncertainty quantification and propagation by supporting structural simulations with machine learning surrogate models using Gaussian processes.

The first approach focuses on quantifying and decomposing parameter-induced uncertainty in structural responses under specific hazard scenarios. Its goal is to support probabilistic sampling-based analyses, including model calibration and updating, iterative performance-based design, and sensitivity analysis.

The second approach focuses on the quantification, propagation, and decomposition of uncertainty in structural vulnerability assessment under a broad range of seismic events. This approach implements and discusses the performance-based engineering framework from a philosophical standpoint, although applied to a real-world case study. Available definitions of damage states in bridge components and the relationships between these and their consequences are discussed.

Both approaches are developed in a fully probabilistic setting, including probabilistic seismic hazard analysis, probabilistic structural modeling, and uncertainty decomposition.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Professor Rodrigo Astroza, for giving me the opportunity of being a part of his team since my undergraduate research. Within this team, I have had the opportunity to work with many world-class researchers and colleagues. The work we have done together has taken me on adventures I never imagined. The group's work ethic, integrity, and hunger for excellence have shaped me personally and professionally. It is a privilege to take part in Professor Astroza's team, and the formation I have received throughout nearly eight years is something I will always be grateful for.

I would also like to thank Universidad de los Andes and the Faculty of Engineering and Applied Sciences for fostering the thriving environment in which our team is embedded. In 2012, I chose to believe in the educational project of this relatively new institution and join as a freshman in Civil Engineering. Almost thirteen years later, I have come to complete my training as an Engineer and Researcher in the same institution. I am proud to have seen the growth of the university from within, and to embody its values in my future endeavors. I will always look forward to "Ir por más".

I would like to thank my family and friends for always supporting my pursuits, however insane they might sound. Their unconditional support has allowed me to pursue this path. In this sense, completing a doctoral degree is a shared accomplishment. Among colleagues, this support has been no less. The constant discussions and exchange of ideas and visions have been the key to developing new ideas and projects, as well as personal growth.

Finally, I would like to thank the financial support of ANID through Beca de Doctorado Nacional Nr. 21210182 and project FOVI230030, for believing in me and supporting me financially during this project. Going forward, the challenge is to repay this opportunity, even if in a small proportion, by contributing to shape the future of the civil engineering landscape as a researcher, educator and engineer.

GENERAL INDEX

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

1.1 MOTIVATION

Currently, structural engineering deals with infrastructure systems that are constantly deteriorating and subjected to multiple natural hazards, both existing and projected due to climate conditions and geological processes. In recent decades, enabled by technological advancements, probabilistic frameworks have been developed to understand sources of risk in structures and generate solutions to mitigate them. Thus, what was once sufficient for structural engineering (i.e., prescriptive design based on empirically developed codes and collegial opinion) has evolved into an approach that aims to establish concrete structural performance objectives that can be explicitly verified.

This philosophy is known as performance-based structural engineering, and in the particular case of earthquake hazards, performance-based earthquake engineering (PBEE). Its methodological framework aims to assess the risk imposed by seismic events on one or more structures, with the goal of providing a tangible decision variable that distills this risk into an interpretable metric (e.g., expected annual economic or human losses). This approach not only sets the bar for performance objective definitions in future design codes but also facilitates communication between structural engineers and decision-makers, in order to protect the common interest and move toward more resilient societies.

In this context, the technical challenge in advancing performance-based engineering lies in the definition and implementation of methodologies that enable this development. In traditional analysis approaches, linear-elastic constitutive models subject to deterministic analysis scenarios are usually considered. However, in reality, there exists uncertainty about seismic events, material properties, deterioration due to aging or natural/man-made hazards, modeling techniques, limit states defining damage thresholds, etc. Bluntly put, there is hardly any deterministic aspect to structural behavior.

Therefore, probabilistic approaches are taken to each stage of performance and risk assessment, including (i) seismic hazard, (ii) structural response, (iii) damage estimation, and (iv) damage consequences. Nevertheless, the resources required to implement a fully

probabilistic PBEE framework have deterred its adoption by the industry. It is in this context that recent research has focused on leveraging artificial intelligence (AI) to develop efficient probabilistic approaches. One of them is the use of surrogate models, which can be defined as a "model of a model".

Surrogate models are used to predict the response of an original model. They can be formulated as a simplified version of the original one (e.g., reduced-order models) or they can be data-driven predictors of the original model. AI surrogates usually fall in the latter category. Among the available AI surrogate modeling techniques, machine and deep learning algorithms provide the ability to train a numerical model that is able to achieve exceptional accuracy at an equally exceptional computational cost.

At this time, while the PBEE framework seems superior and extensive academic efforts are being made to transfer this knowledge and philosophy to the industry, the current state-of-practice is vastly limited by computational and time resources. Thus, the promised improved outcome of PBEE has not been enough to warrant a widespread desire by engineering firms to invest in the transition. This is precisely what motivates this research: lowering the barrier of entry to fully implement a probabilistic structural risk assessment framework by leveraging AI-based surrogate modeling techniques.

1.2 CENTRAL TOPIC OF THIS THESIS

The work presented in this doctoral thesis details the development of two novel methodological approaches for uncertainty treatment in structural engineering, which leverage Gaussian Process surrogate models to replace the original finite element models in massive computer simulations, and which can be used for future sampling under a variety of seismic scenarios of interest.

The first approach focuses on studying parameter-induced uncertainty in structural responses. To this end, (i) hazard-specific seismic scenarios are selected through probabilistic seismic hazard analysis, (ii) probabilistic characterizations of mechanics-based structural finite element models are developed through Bayesian constitutive model calibrations, (iii) large-scale simulations of the structural models are performed, (iv) scenario-specific

surrogate models are trained, and (v) parameter-induced uncertainty is decomposed through sampling-based sensitivity analysis using the surrogate models.

The second approach aims to provide a framework for uncertainty treatment in performance-based earthquake engineering. In this approach, (i) a comprehensive database of actual ground motions is selected, (ii) the structure is defined probabilistically as in approach 1, (iii) probabilistic damage thresholds and component capacity estimates are derived, (iv) large-scale simulation is used to train surrogate models that can accurately predict responses in a wide range of seismic scenarios, (v) probabilistic parameter-conditional fragility assessment is conducted, and (vi) parameter-induced uncertainty in vulnerability assessments is investigated.

1.3 RESEARCH HYPOTHESIS

This thesis holds that it is feasible to develop and implement efficient approaches to a complete implementation of the PBEE framework, while providing a comprehensive treatment of uncertainties in each stage, including their quantification and propagation, by leveraging currently available surrogate modeling techniques. Furthermore, this thesis holds that further development of said implementations can facilitate the adoption of the PBEE framework by the industry in upcoming years, thus helping develop concrete and explicitly verifiable structural performance objectives and so, design codes and paradigms.

1.4 RESEARCH OBJECTIVES

General Objective

The main objective of this thesis is to develop and validate comprehensive methodologies for the efficient quantification and propagation of uncertainties that allow a practical, although complete, adoption of the PBEE framework. The scope of research of this thesis is defined to be the application of the proposed methodologies to highway bridges in Chile subjected to earthquake hazards.

Specific Objectives

To achieve the main research objective, the following specific milestones are established:

1. Establish the workflow for probabilistic seismic hazard assessment, considering Chilean seismicity and earthquake databases.

2. Establish probabilistic characterizations of bridge structural finite element models, via calibration of their components and materials' nonlinear constitutive models using relevant experimental data.

3. Implement and perform large-scale simulations of the finite element models in a high-performance cluster environment.

4. Select representative structural responses to characterize the structure's dynamic behavior under earthquake excitations.

5. Develop surrogate models to accurately replicate training data and predict new scenarios.

6. Establish an uncertainty decomposition approach leveraging the low cost of surrogate model sampling and variance-based sensitivity analysis techniques.

7. Study parameter-induced uncertainty in structural responses under hazard-specific scenarios, as well as in fragility estimates under general earthquake hazards.

1.5 THESIS STRUCTURE

This thesis is structured as follows:

Chapter 2 presents a literature review and state-of-the-art of the different PBEE-related topics discussed in this thesis.

Chapter 3 describes the mathematical tools used and the methodologies proposed for both approaches.

Chapter 4 presents the results and discussion obtained from each task carried out.

Chapter 5 establishes key conclusions.

2 LITERATURE REVIEW AND STATE OF THE ART

2.1 THE CONTEXT FOR UQ&P IN PBEE AND THE NEED FOR SURROGATE MODELS

Performance-based earthquake engineering (PBEE) (Günay & Mosalam, 2013) provides a probabilistic framework for performance and risk analysis of structures subjected to seismic hazards. This framework consists of four main sequential stages: (i) probabilistic seismic hazard analysis (PSHA); (ii) probabilistic seismic demand analysis (PSDA); (iii) probabilistic seismic damage/fragility analysis (PSFA); and (iv) probabilistic seismic loss analysis (PSLA). Uncertainties involved in these stages need to be quantified and propagated to finally achieve risk-informed decision-making variables (i.e., monetary cost, human casualties, downtime, and environmental impacts).

Therefore, uncertainty quantification and propagation (UQ&P) at each stage play a crucial role in the effective adoption of the PBEE framework and the generation of meaningful outcomes. Both aleatoric and epistemic uncertainties (Kiureghian & Ditlevsen, 2009) are present regarding the occurrence and intensity of seismic events, actual structural properties, accuracy of numerical models representing the structure, capacity and limit state (LS) definitions, and the consequences of reaching such LSs. Developing a comprehensive and practical approach for the treatment of uncertainty in PBEE requires efficient handling of these uncertainties, limiting the computational cost of performing simulations on nonlinear finite element (FE) models. This challenge has motivated the use of surrogate models (SMs) to mitigate the computational costs arising from running FE models. SMs include a variety of formulations, ranging from reduced order models to data-driven approaches (e.g., (Gomez-Cabrera & Escamilla-Ambrosio, 2022; Kudela & Matousek, 2022; Patsialis & Taflanidis, 2020; Soleimani & Hajializadeh, 2022; Spiridonakos & Chatzi, 2015; Thai, 2022; Xie et al., 2020; Yazdani et al., 2020)).

Particularly, SMs based on machine learning and deep learning algorithms have become popular due to their ability to be trained on a reduced dataset and to be used for prediction at virtually no additional computational cost (e.g., (Wang et al., 2023)). In the context of uncertainty quantification, Bayesian methods such as Gaussian Process (GP) models

(Rasmussen & Williams, 2005) emerge as a natural alternative due to their inherently probabilistic formulation. Although Bayesian deep learning algorithms (e.g., (Xie, 2024)) may provide improved efficiency and predictions for diverse datasets, GPs remain attractive due to the straightforward formulation, the high interpretability, and relatively low data volume required for training at a viable computational cost. GPs have been successfully adopted as a surrogate modeling tool in a wide range of applications in structural engineering, including various structure types (e.g., buildings (Gentile & Galasso, 2020; Xiao et al., 2021), wharves (Su et al., 2021), bridges (Castellon et al., 2023; Saida & Nishio, 2023), dams (Kang et al., 2023)), and diverse scopes of applications such as structural components (Mo et al., 2024), simplified dynamic systems (Gidaris et al., 2015; Hung & Thang, 2022; Lu et al., 2019), structure-specific fragility and reliability analysis (Hoang et al., 2021; Su et al., 2017; Xiao et al., 2021; Xing et al., 2022; Zhang & Wu, 2019), life-cycle assessment (Yin et al., 2024), regional fragility assessment (Ning et al., 2024), and structure portfolios (Gentile & Galasso, 2020).

2.2 Parameter-Induced Uncertainty Quantification in Hazard Specific Scenarios

In the probabilistic structural response simulation stage of PBEE, nonlinear FE models of the structure are usually subjected to a suite of ground motions selected to represent a range of seismic hazard levels (Bazzurro & Allin Cornell, 1999). This is usually done by simulating the structural response through nonlinear time history analysis (NLTHA) and extracting relevant engineering demand parameters (EDPs) to quantify structural demands. These EDPs are later incorporated in fragility analysis, associated to damage assessment (Baker, 2015).

It has been shown that the variability in structural responses due to uncertain material and mechanical properties is significant (e.g., (Mangalathu et al., 2018; Xie & DesRoches, 2019)). In this context, parametric uncertainty is commonly addressed by sampling model parameter probability distributions (PDFs) and performing NLTHA on a certain model class numerous times to generate probabilistic models for the responses of the structure. While sampling can produce an improved understanding of the actual expected structural behavior, the computational cost associated with FE simulation can quickly become prohibitive, as a new NLTHA must be executed for each sample realization.

Previous literature has addressed the study of parameter sensitivity in a broader context (e.g., within fragility analysis (Saida et al., 2024; Yan et al., 2024), reliability analysis (Saida & Nishio, 2023; Skandalos et al., 2022), performance assessment (Su et al., 2021)). However, the thorough study of parameter-induced uncertainty in a single structure's model responses is an important first step towards probabilistic surrogate modeling in the context of PBEE. A deep understanding of the composition of model response uncertainty helps in two main aspects. Firstly, to validate the ability of a surrogate model to capture the underlying dynamic structural behavior. This ensures the viability of using surrogate models for further sampling-based applications, such as model updating (e.g., (Li et al., 2024)). Secondly, by identifying the influence of different model parameters in model response uncertainty, data-collection efforts for probabilistic characterization of structural material and components can be informed to allocate resources more efficiently (see for example, the parameter selection through principal component analysis and SM-based Sobol sensitivity analysis in (Nagel et al., 2020)).

Although other statistical machine learning techniques could be used for Sensitivity Analysis (SA) (e.g., PCE (Sudret, 2008), SVR (Cheng et al., 2017), multi-fidelity methods (Shang et al., 2023)). Currently, there is no one-size-fits-all solution for SM choice in view of UQ&P, and the model needs to be chosen carefully for the specific problem at hand (Cheng et al., 2020). GPs remain a top choice in the context of this thesis for two main reasons: (i) the probabilistic formulation of GP can directly translate to further uncertainty propagation in performance assessment, as the predictive variance can be included as a source of epistemic uncertainty in structural demands, and (ii) the nonparametric kernel-based formulation allows for flexibility in fitting non-smooth data and capturing the underlying nonlinear parameter interactions.

2.3 Uncertainty Quantification and Propagation in Performance-Based Earthquake Engineering

In the PBEE framework, the first three stages, from PSHA to PSFA, cover structural performance itself. The sequential framework consists of estimating seismic hazards to which the structure is subjected (PSHA), assessing structural demands imposed by earthquakes at those hazard levels (PSDA), and estimating damage probabilities at different earthquakes

intensities (PSFA). Uncertainties in seismic hazard, structural properties, and damage states, are incorporated through probabilistic characterizations of the relevant parameters. To this extent, fragility functions in PSFA are inherently influenced by uncertainties propagated from previous stages. A strong emphasis has been placed recently on methods for obtaining fragility parameter estimates, including component-based (e.g., (Cao et al., 2023; Jalayer et al., 2017; Jalayer et al., 2015)) and joint-based approaches (e.g., (Conde Bandini et al., 2022; Lan et al., 2024; Yan et al., 2022; Zhou & Li, 2019)). Understanding the sensitivity of fragility estimates to uncertainty in model parameters is crucial for achieving reliable structural performance assessments. In this context, the importance of accounting for parameter uncertainty has been demonstrated across various structure types (e.g., (Padgett & DesRoches, 2007; Pan et al., 2020; Sarkar & Dasgupta, 2024; Xie & DesRoches, 2019)).

GP-based SM approaches have been developed to analyze uncertainty in fragility and reliability estimates. Skandalos et al. (Skandalos et al., 2022) proposed an approach using multi-fidelity GP surrogates for reliability analysis. Saida et al. (Saida et al., 2024) proposed a multi-output GP-based approach for fragility analysis of bridges and investigated the contributions of input and structural parameters to the SM predictions. Ning et al. (Ning et al., 2024) developed an approach combining GP and active learning to study the uncertainty in regional seismic fragility assessments of bridge portfolios. Kim and Kim (Kim & Kim, 2025) proposed a method for probabilistic fragility analysis incorporating epistemic uncertainty in ground motions and structural models, using GP surrogates for fragility curve parameters. Yan et al. (Yan et al., 2024) introduced a GP-based technique for global sensitivity analysis (GSA) of fragility functions, accounting for parameter uncertainty by employing GP surrogates for parameters of fragility functions.

Thus, the study of uncertainty composition in fragility estimates has gained attention among researchers due to the critical role of proper damage characterization in PBEE. However, previous studies have exhibited certain limitations; for example, they (i) focused on regional assessment, addressing a broader scope rather than the physical phenomena involved in the behavior of single structure (e.g., (Gentile & Galasso, 2020; Ning et al., 2024)); (ii) approached surrogate modeling directly for parameterized fragility functions, thereby aggregating uncertainties from PSHA and PSDA to PSFA (e.g., (Kim & Kim, 2025; Yan et

al., 2024)); (iii) applied SMs to simplified structures for academia (e.g., (Kim & Kim, 2025)); or (iv) not focused on establishing a framework for the study of uncertainty composition in fragility functions (e.g., (Saida & Nishio, 2023; Saida et al., 2024; Skandalos et al., 2022)). To the best of the author's knowledge, there remains a need to develop an efficient framework for the comprehensive treatment of uncertainty across all stages of PBEE, explicitly including both aleatoric and epistemic uncertainties.

2.4 Constitutive Modeling of Exterior Shear Keys in Bridges

As an important part of a road network, critical or essential bridges shall remain operative during and after earthquakes. Current efforts in performance-based earthquake engineering (PBEE) are directed at developing methodologies for the assessment and risk-oriented design of bridges, with a strong focus on the quantification of model parameter uncertainty at a local level and its propagation to structural behavior (e.g., (Buckle et al., 2020; Deb, 2021; Deb et al., 2022; Liu et al., 2022; Pastén et al., 2021; Soleimani, 2021)).

Cases of inadequate bridge performance during strong earthquakes have been evidenced in recent events, such as the 2010 Mw=8.8 Maule earthquake in Chile (Buckle et al., 2012; Kawashima et al., 2011) and the 2011 Mw=9.1 Tohoku earthquake in Japan (Kawashima, 2012). Previous research has reported that exterior shear keys play a crucial role in seismic bridge performance (Wilches et al., 2017; Xiang & Li, 2018; Yue et al., 2023), as they restrict lateral superstructure displacements and prevent damage in substructure components, including deck bearings, piers, pier cap beams, pier foundations, and abutments.

For optimal structure performance, shear keys should act as sacrificial elements; thus, they should dissipate energy to prevent damage to those components (Bozorgzadeh, Megally, Restrepo et al., 2006; Megally et al., 2002; Silva et al., 2003; Silva et al., 2009). By accepting a controlled level of damage, a ductile design can be achieved for this purpose. For this reason, it is essential to understand how shear keys fail depending on their design criteria and detailing. The ability to model each shear key in a bridge pier or abutment independently is also important to attain an accurate representation of its behavior and the interaction with adjacent bearings (Wei et al., 2023).

Three main failure modes have been identified for exterior shear keys (Han, Zhou, Ou et al., 2017; Megally et al., 2002). They include (a) diagonal tension (DT), in which the shear failure is governed by a diagonal crack propagating from the base of the shear key to base of the supporting element (i.e., stem wall or pier cap beam, in the cases of abutment or pier shear keys, respectively); (b) sliding shear (SS), in which a horizontal shear failure of the monolithic joint between shear key and supporting element is observed; and (c) sliding friction (SF), in which the shear key slides laterally along a frictionally weak interface which has been intentionally placed between the shear key and the supporting element, where the dowel effect of the vertical reinforcement primarily resists the shear force. The former two modes of failure have been identified in monolithic joint constructions (i.e., non-isolated shear keys), while the latter is prominent in resilient-type and smooth joint constructions (i.e., isolated shear keys) (Han, Zhou, Zhong et al., 2017; Kottari, 2016).

In this context, DT failure has been extensively documented and studied. Megally et al. (Megally et al., 2002) conducted experimental tests and developed an analytical model for shear keys failing in DT, which was later simplified by Goel and Chopra (Goel & Chopra, 2008) to a tri-linear backbone curve which can be easily implemented in finite element (FE) modeling software. Bozorgzadeh et al. (Bozorgzadeh, Megally, Ashford et al., 2006) further extended the battery of experimental tests conducted by Megally et al. (Megally et al., 2002) and further validated the model. Kottari (Kottari, 2016) extended the battery of experimental tests by Megally et al. (Megally et al., 2002) and proposed an equilibrium-based expression to estimate the peak capacity of shear keys in DT. These models have been validated in terms of capacity prediction for DT failure. Han et al. (Han, Zhou, Ou et al., 2017; Han, Zhou, Zhong et al., 2017) conducted experimental testing on shear keys designed under Chinese Guideline for Design of Highway Bridges (Ministry of Transport of the People's Republic of China, 2008) and proposed an analytical model for the force-displacement response of the three main failure modes. Additionally, in (Han, Zhou, Ou et al., 2017; Han, Zhou, Zhong et al., 2017; Kottari, 2016), expressions to determine the failure mechanism of monolithic shear keys based on the distribution of their reinforcement have been proposed.

Comparatively, exterior shear keys failing in SS have been scarcely investigated. In past experimental campaigns, Megally et al. (Megally et al., 2002) conducted six tests, of which

five failed in DT and one showed flexural failure. Bozorgzadeh et al. conducted six tests, of which three failed in DT, two non-monolithic shear keys failed in SF and one failed by bond slip of vertical rebars. Kottari (Kottari, 2016) conducted eight tests on non-skewed shear keys, of which two non-monolithic shear keys failed in SF and six monolithic shear keys failed in SS. Han et al. (Han, Zhou, Ou et al., 2017; Han, Zhou, Zhong et al., 2017) conducted ten tests, of which three monolithic shear keys failed in DT, three in SS and four non-monolithic shear keys failed in SF. Recently, Mei and Guo (Mei & Guo, 2023) reported experimental results from ten tests on monolithic and non-monolithic shear keys, in which three failure modes were identified as diagonal and horizontal shear for monolithic specimens, and sliding shear for non-monolithic specimens. In the experiments conducted by Han et al. (Han, Zhou, Ou et al., 2017; Han, Zhou, Zhong et al., 2017) and Mei and Guo (Mei & Guo, 2023), the sliding shear mechanism described for monolithic keys appear to be a mixture of the DT and SS modes discussed in this paper, characterized by an initial sliding component and a flexure component which led the cracks to propagate diagonally earlier than in the tests reported in (Han, Zhou, Ou et al., 2017; Han, Zhou, Zhong et al., 2017; Kottari, 2016). Additionally, a rise in post-sliding capacity is observed, which further suggests the interaction of shear and flexure mechanisms in those shear keys. This behavior is not reported in the tests performed by Kottari.

To serve as sacrificial elements, shear keys should be designed to withstand a certain degree of damage to dissipate energy instead of transferring it to pier columns. As previous authors have reported, DT failure is observed when peak sliding capacity exceeds the shear capacity of the stem wall (in abutments) or cap beam (in piers). In that case, the shear key does not behave exactly as a sacrificial element, but rather as a stopper for lateral displacement, due to the higher capacity of shear keys failing in DT. A direct consequence of this behavior is the damage observed during strong earthquakes in piers due to vertically transferred loads, as incoming energy not dissipated via damage to shear key is transferred to the cap beam and subsequently to pier columns. To prevent this type of damage, Kottari (Kottari, 2016) recommends that monolithic exterior shear keys should be designed to fail in SS.

In performance-based seismic assessment, it is necessary to model SS-failing keys in a manner that is viable to implement in structure-level FE models. To this extent, the

expressions proposed by Kottari et al. (Kottari, 2016; Kottari et al., 2020) can be used to predict the peak sliding capacity, but do not provide information about the lateral displacement. Additionally, solid tri-dimensional FE models were developed in (Kottari, 2016) to represent the behavior of monolithic and isolated shear keys, along with a phenomenological interface element for dowel action, although a model formulation for use in structure level FE models was not proposed. Han et al. (Han, Zhou, Zhong et al., 2017) proposed an analytical formulation for the lateral force-displacement response of shear keys in the three main failure modes. In that model, however, the response is determined in relation to the vertical rotation of the shear key, which has shown great variability across different studies, and therefore may not be practical to compute for engineering practice. Moreover, model predictions reported by Han et al. (Han, Zhou, Zhong et al., 2017), while capturing capacity to an accurate degree, showed significant discrepancies in terms of displacement predictions when compared against experimental data. Considering future endeavors to define displacement-based limit states for component and system-level fragility analyses of exterior shear keys, the challenge remains to develop a simple, yet effective lateral force-displacement model for SS failure.

3 METHODOLOGY

3.1 SENSITIVITY ANALYSIS

3.1.1 LOCAL SENSITIVITY ANALYSIS THROUGH INPUT/OUTPUT CORRELATION

The Input/Output (I/O) Correlation method (Marelli et al., 2022) for Local Sensitivity Analysis (LSA) investigates the linear correlation between parameter and response values. In this method, correlation coefficients ρ_i are computed for each model parameter x_i. When the model in question presents a vector response, the coefficients are computed for each input value w (e.g., strain history, time history). The method is thus defined by a vector of strain-dependent correlation coefficients as shown in Eq. (3.1.1):

$$\rho_i(w) \equiv \rho(x_i, \boldsymbol{y}(w)) = \frac{E[(x_i - \mu_i)(\boldsymbol{y}(w) - \mu_{\boldsymbol{y}(w)})]}{\sigma_i \sigma_{\boldsymbol{y}(w)}} \tag{3.1.1}$$

Where μ_i is the mean of all samples for parameter x_i, $\mu_{\boldsymbol{y}(w)}$ is the mean of all sampled responses at strain w, and $\sigma_i, \sigma_{\boldsymbol{y}(w)}$ are their respective standard deviations. By definition, the value of correlation coefficients is in the interval $[-1,1]$. Subsequently, $\rho_i(w)$ will be close to 1 when x_i greatly influences $\boldsymbol{y}(w)$ in a positive manner, close to −1 if a greater value of x_i significantly decreases the value of $\boldsymbol{y}(w)$, and close to 0 if a change in x_i does not have a considerable impact on the value of $\boldsymbol{y}(w)$.

3.1.2 GLOBAL SENSITIVITY ANALYSIS BASED ON SOBOL'S INDICES

Global Sensitivity Analysis (GSA) measures the variance induced in model responses by the parameter's variance. Variance-based indices as defined by Sobol (Sobol, 1993) are based on a decomposition of the model output variance into summands of different dimensions. The formulation considers the parameter space to be the unit hypercube $\Omega^n = \{\boldsymbol{x} | 0 \leq x_i \leq 1, i = 1, \dots, n\}$ and assumes f to be square-integrable over Ω^k, with $k = \{1, \dots, n\}$. In (Sobol, 2001), Sobol adopted the term ANOVA-representation, as introduced by Efron and Stein (Efron & Stein, 1981), defined in Eq. (3.1.2), by a total of 2^n summands.

$$f(\boldsymbol{x}) = f_0 + \sum_{s=1}^{n} \sum_{i_1 < \dots < i_s}^{n} f_{i_1 \dots i_s}(x_{i_1}, \dots, x_{i_s}) \tag{3.1.2}$$

Which must satisfy four conditions, as follows:

$$\int_0^1 f_{i_1 \dots i_s} dx_k = 0, \forall k = i_1 \dots i_s \tag{3.1.3}$$

Hence, all components from Eq. (3.1.2) are orthogonal. Moreover, each summand can be written in terms of the integral of f over its domain (see Eqs. (3.1.4) to (3.1.6)).

$$\int f(\boldsymbol{x})\, dx = f_0 \tag{3.1.4}$$

$$\int f(\boldsymbol{x})\, dx_{\sim i} = f_0 + f_i(x_i) \tag{3.1.5}$$

$$\int f(\boldsymbol{x})\, dx_{\sim i,j} = f_0 + f_i(x_i) + f_j(x_j) + f_{ij}(x_i, x_j) \tag{3.1.6}$$

Where sub index $\sim i$ indicates to integrate over Ω^k, except for the *i*-th dimension. Index $\sim i, j$ indicates to integrate over all dimensions except for the *i*-th and *j*-th dimensions, and so on. Squaring Eq. (3.1.2) and integrating over Ω^k yields

$$\int f^2(\boldsymbol{x})\, dx - f_0^2 = \sum_{s=1}^{n} \sum_{i_1 < \dots < i_s}^{n} \int f_{i_1 \dots i_s}^2 dx_{i_1} \dots dx_{i_s} \tag{3.1.7}$$

The terms in Eq. (3.1.7) are called partial variances and are designated as $D = \int f^2(\boldsymbol{x})\, dx - f_0^2$ and $D_{i_1 \dots i_s} = \int f_{i_1 \dots i_s}^2 dx_{i_1} \dots dx_{i_s}$. The global sensitivity index of order $i_1 \dots i_s$ is introduced as the relation in Eq. (3.1.8)

$$S_{i_1 \dots i_s} = \frac{D_{i_1 \dots i_s}}{D} \tag{3.1.8}$$

Where it holds that the sum of all the terms from Eq. (3.1.8) is equal to 1.

It is often impractical, however, to perform direct integration of the total indices, as it would require computing all terms. As an alternative, a Monte Carlo simulation is employed to obtain the indices. Homma and Saltelli (Homma & Saltelli, 1996) and Janon et al. (Janon et al., 2014) have proposed estimators for the computation of indices for a particular set of variables. In this study, Janon's estimator T is used to compute the indices. The simulation-based estimator is derived from Eq. (3.1.9), in which the Sobol index of order p is introduced, where $p = \{i_1, \dots, i_s\}$ denotes the set of parameters for the index to be computed and N is the number of simulations to be performed.

Let $\boldsymbol{Y} = \{\boldsymbol{y}_i\}_{i=1,\dots,N}$ and $\boldsymbol{X} = \{\boldsymbol{x}_i\}_{i=1,\dots,N}$ be the matrices containing all samples for response and parameters, respectively. Parameters can be divided into two groups p and $\sim p$ such that $\boldsymbol{y} = f(\boldsymbol{x}_p, \boldsymbol{x}_{\sim p})$. By also denoting by $\boldsymbol{y}_p = f(\boldsymbol{x}_p, \boldsymbol{x}'_{\sim p})$ the case where a sample $\boldsymbol{x}'_{\sim p}$ is a sample independent from $\boldsymbol{x}_{\sim p}$, taken from the sample set $\boldsymbol{X}_{\sim p}$. It follows that $\boldsymbol{X}'_{\sim p}$ is an independent copy of $\boldsymbol{X}_{\sim p}$, and analogously $\boldsymbol{X}'_p$ is an independent copy of $\boldsymbol{X}_p$. Therein, the sets of responses $\boldsymbol{Y} = f(\boldsymbol{X}_p, \boldsymbol{X}_{\sim p})$ and $\boldsymbol{Y}^p = f(\boldsymbol{X}_p, \boldsymbol{X}'_{\sim p})$ are obtained.

$$S_p = \frac{\mathrm{Var}[E[\boldsymbol{Y}|\boldsymbol{X}_p]]}{\mathrm{Var}[\boldsymbol{Y}]} = \frac{\mathrm{Cov}[\boldsymbol{Y}, \boldsymbol{Y}^p]}{\mathrm{Var}[\boldsymbol{Y}]} \tag{3.1.9}$$

The Janon Estimator T_N^p is then defined by Eq. (3.1.10):

$$T_N^p = \frac{\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{Y}_i \boldsymbol{Y}_i^p - \left(\frac{1}{N}\sum_{i=1}^{N}\left[\frac{\boldsymbol{Y}_i + \boldsymbol{Y}_i^p}{2}\right]\right)^2}{\frac{1}{N}\sum_{i=1}^{N}\left[\frac{\boldsymbol{Y}_i{}^2 + \left(\boldsymbol{Y}_i^p\right)^2}{2}\right] - \left(\frac{1}{N}\sum_{i=1}^{N}\left[\frac{\boldsymbol{Y}_i + \boldsymbol{Y}_i^p}{2}\right]\right)^2} \tag{3.1.10}$$

Which eliminates the need to compute all indices to obtain the effect of a particular group of variables. In the case where $p = \sim i = \{1, .., i-1, i+1, \dots, n\}$, the total effect of the variable x_i is obtained from Eq. (3.1.11):

$$S_i^T = 1 - S_{\sim i} \tag{3.1.11}$$

Expressing Eqs. (3.1.9) and (3.1.11) in terms of the input vector w yields the generic and total order index history as

$$S_{i_1 \dots i_s}(w) = \frac{D_{i_1 \dots i_s}(w)}{D(w)} \tag{3.1.12}$$

$$S_i^T(w) = 1 - S_{\sim i}(w) \tag{3.1.13}$$

The analyses presented in this document consider the first and total order Sobol indices, which enable a better understanding of the composition of response uncertainty induced by parameter variability and interactions among the parameters in each model.

3.2 PROBABILISTIC CONSTITUTIVE MODEL CHARACTERIZATION

3.2.1 PARAMETER ESTIMATION USING BAYESIAN MONTE CARLO ALGORITHMS

The Bayesian approach for parameter estimation allows for a probabilistic characterization upon the values of model parameters. Its objective is to update prior beliefs about parameter values and variability through observed data and conditional probabilities, obtaining as an output the so-called posterior distribution (or posterior PDF) for each parameter.

The general formulation of the Bayesian inversion problem is derived from Bayes' theorem, which expressed in terms of model parameters $\boldsymbol{x}$, measured response $\boldsymbol{y}$ and model class M yields Eq. (3.2.1):

$$p(\boldsymbol{x}|\boldsymbol{y}, M) = \frac{p(\boldsymbol{y}|\boldsymbol{x}, M)\pi(\boldsymbol{x}|M)}{p(\boldsymbol{y}|M)} \tag{3.2.1}$$

Where $p(\boldsymbol{x}|\boldsymbol{y}, M)$ is the updated belief, commonly known as posterior PDF, about the values of parameters after observing data $\boldsymbol{y}$. The likelihood function $p(\boldsymbol{y}|\boldsymbol{x}, M)$ corresponds to the degree of uncertainty associated with model class M predicting response $\boldsymbol{y}$ when using parameters $\boldsymbol{x}$. Such formulation implies that variability $\boldsymbol{\Sigma}$ in the likelihood function accounts for the model predictive capacity, with $\mathbf{0}$ meaning that the correct parameters generate a perfect prediction of the observed data. Prior beliefs are represented by $\pi(\boldsymbol{x}|M)$. This function incorporates subjective interpretations of the model behavior based on previous knowledge and experience. The term $p(\boldsymbol{y}|M) = \int_{\Omega^n} p(\boldsymbol{y}|\boldsymbol{x}, M)\pi(\boldsymbol{x}|M)d\boldsymbol{x}$, known as model evidence, is a multi-dimensional integral over an uncertain parameter space. Computation of this integral is often not practical, and a simulation approach can be taken through the use of Monte-Carlo (MC) (e.g., Markov-Chain MC (MCMC), Sequential MC (SMC)) algorithms to obtain posterior PDFs of parameters (Kroese et al., 2011).

3.2.2 CONVERGENCE CRITERIA FOR MCMC PARAMETER ESTIMATIONS

In Bayesian model calibrations using MC algorithms, such as MCMC and SMC, it is crucial to verify the convergence of the sampled posterior PDF to the target distribution to ensure the parameter estimation is terminated appropriately and reliably.

MCMC algorithms require careful convergence analysis due to the formulation of samplers used to construct the Markov chains. The MCMC applications discussed in this thesis use the Adaptive Metropolis sampler (Haario et al., 2001).

To solve Eq. (3.2.1) numerically by MCMC, multiple Markov Chains are initialized at random points within the prior distribution of the parameters. The likelihood function is evaluated at $\boldsymbol{y}$ and the sample is accepted or rejected based on the improvement of the fit with respect to the current distribution. This process is repeated indefinitely until the MC simulation is terminated. Thus, the convergence of the sampled chains to a stationary (i.e., the target) distribution is not guaranteed. To check the convergence of multiple chains to a single stationary distribution, Gelman and Rubin (Gelman & Rubin, 1992) proposed the Multivariate Potential Scale Reduction Factor (MPSRF, or $\hat{R}$) which measures the convergence of means of independent MCMC chains. This is expressed by Eq. (3.2.2)

$$\hat{R} = \frac{T}{T+1} + \left(\frac{C+1}{C}\right)\lambda_M \tag{3.2.2}$$

Where C is the number of independently run chains, T is the number of sample points per chain, and λ_M is the maximum eigenvalue of the matrix $W^{-1}B$. Herein, W and B are two covariance matrices that reflect the in-chain and inter-chain covariances, respectively. The covariance of the *i*-th chain is given by Eqs. (3.2.3) and (3.2.4)

$$W_i = \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{x}_{i,t} - \boldsymbol{x}_{i,\mu})\,(\boldsymbol{x}_{i,t} - \boldsymbol{x}_{i,\mu})^T \tag{3.2.3}$$

$$W = \frac{1}{C}\sum_{i=1}^{C} W_i \tag{3.2.4}$$

Where $\boldsymbol{x}_{i,t}$ is the *t*-th sample of chain *i*, for the vector of parameters $\boldsymbol{x}$, and $\boldsymbol{x}_{i,\mu}$ is the mean of chain *i*.

Likewise, the inter-chain covariance is given by Eqs. (3.2.5) and (3.2.6).

$$B = \frac{1}{C-1}\sum_{i=1}^{C}(\boldsymbol{x}_{i,\mu} - \boldsymbol{x}_{\mu})\,(\boldsymbol{x}_{i,\mu} - \boldsymbol{x}_{\mu})^T \tag{3.2.5}$$

$$\boldsymbol{x}_{\mu} = \frac{1}{C(T+1)} \sum_{i=1}^{C} \sum_{t=1}^{T} \boldsymbol{x}_{i,t} \tag{3.2.6}$$

Where $\boldsymbol{x}_{\mu}$ is the average of all states of all the sampled chains.

Thus, if all chains converge to the same mean, $\hat{R}$ converges to 1 from above. The threshold for convergence used in this thesis is $\hat{R} < 1.02$.

Furthermore, although convergence might be reached, the quality of the inference is not guaranteed because of possible autocorrelation between samples in a chain. To verify the quality of a multivariate posterior PDF inferred by MCMC, Vats et al. (Vats et al., 2019) proposed a metric called multivariate effective sample size (mESS), which is an extension of the previously available univariate effective sample size.

The mESS quantifies the loss of information due to autocorrelation. In the case where all samples within a set of MCMC chains have been drawn independently, the mESS is equal to the number of samples drawn. As this is virtually impossible due to the formulation of samplers used in MCMC, a minimum mESS (minESS) can be defined to achieve a relative precision ε_p with a confidence interval of α_p. By this definition, ε_p is the relative precision defined in terms of MC standard error of the chains.

The mESS is defined by Eq. (3.2.7) and the minESS is defined by Eq. (3.2.8), whereby if the mESS obtained is greater than minESS, the MCMC sampling is terminated.

$$mESS(\mathbf{X}, \mathrm{C}, \mathrm{T}) = \mathrm{T}\left(\frac{\det(\Lambda)}{\det(\Sigma_{MC})}\right)^{1/p} \tag{3.2.7}$$

$$minESS\left(\alpha_p, \varepsilon_p, p\right) = \frac{2^{2/p}\pi}{\left(p\Gamma\left(\frac{p}{2}\right)\right)^{2/p}} \frac{\chi^2_{1-\alpha_c,p}}{\varepsilon_p} \tag{3.2.8}$$

Where Λ and Σ_{MC} are the sample and MC covariance matrices of the multivariate chain, respectively, p is the number of parameters, Γ is the gamma function, and χ^2 is the chi-squared distribution.

3.2.3 COMBINATION OF POSTERIOR DISTRIBUTIONS

To combine parameter posteriors, one option is to compute the conflation of their distributions. The conflation of PDFs, proposed by Hill and Miller (Hill & Miller, 2010), allows to combine distributions from multiple datasets for the same quantity. It is defined by Eq. (3.2.9), where each distribution in the combination may be assigned a weight to represent the credibility given to its associated dataset.

$$F(x) = \frac{\prod_{i=1}^{n} f_i^{w_i/w_{max}}(x)}{\int_{-\infty}^{\infty} \prod_{i=1}^{n} f_i^{w_i/w_{max}}(t)dt} \tag{3.2.9}$$

where $F(x)$ is the resulting conflated distribution, and w_i, w_{max} are the individual and maximum weights, respectively.

A notable drawback of PDF combination through conflation, however, is the loss of inter-dataset variability. Due to its formulation, the conflation of PDFs produces a narrow resulting distribution. Therefore, this method should be used with caution, subject to the uniformity of the data from the different sources.

In particular, only some of the parameter PDFs used throughout this thesis have been combined by conflating individual calibrations, namely the PDFs for the constitutive models of elastomeric bearings, vertical anchoring seismic bars, and the numerical parameters in the proposed constitutive model for exterior shear keys in bridges.

A second option for combining PDFs belonging to the same population is to simply pool the data samples of all experiments and compute the statistics (i.e., mean, variance, correlations). This method should also be used carefully, as the resulting PDF might not be representative of any individual dataset if there is considerable inter-dataset variability. This is the method used to combine the constitutive models for reinforcing steel and the mechanical parameters of the proposed model for exterior shear keys in bridges.

3.3 PROBABILISTIC SEISMIC HAZARD ANALYSIS

The first stage of PBEE considers the estimation of seismic hazard to which the structure is exposed. Practically, the goal of this stage is to define a realistic hazard model and select a suite of GMs to be used as input excitations in NLTHA of structures.

To define the hazard model, this study considers the methodology proposed by (Baker et al., 2021) to establish PSHA of the structure. In this approach, defined by Eq. (3.3.1), the mean annual rate of an intensity measure (IM) exceeding a certain value (MRE) is computed numerically by integrating the individual hazard contributions of all considered scenarios. The scenarios are defined according to geometric models that characterize multiple possible sources, seismicity and magnitude-recurrence laws, ground motion models (GMMs), and decision trees that quantify the epistemic uncertainty in GMM predictions.

$$\lambda_r(IM > im) = \sum_{i=1}^{n_{sources}} \lambda_s(M_i > m) \int_{m_{min}}^{m_{max}} \int_{0}^{r_{max}} P(IM > im|m,r) f_{M_i,R_i}(m,r) dr dm \qquad (3.3.1)$$

where $\lambda_r(IM > im)$ is the MRE of a value of *IM*, $\lambda_s(M_i > m)$ is the rate of occurrence of source $i \in \{1, \dots, n_{sources}\}$ producing an earthquake of magnitude $M_i > m$, $f_{M_i,R_i}(m,r)$ is the joint distribution of magnitudes and distances of all considered scenarios in source *i*, m_{min}, m_{max}, and r_{max} are the minimum and maximum magnitudes, and maximum distance considered for source *i,* respectively. Of these terms, λ_s and f_{M_i,R_i} are defined by the seismicity model, while $P(IM > im|m,r)$ is computed by predicting the motion intensity using a GMM. To account for the considerable variability among different GMM predictions, logic trees are commonly used to derive $\lambda(IM > im)$ according to Eq. (3.3.2).

$$\lambda(IM > im) = \sum_{k=1}^{n_{GMM}} w_k \lambda_{r_k}(IM > im) \qquad (3.3.2)$$

where n_{GMM} is the number of GMMs considered in the logic tree, w_k is the weight assigned to the *k*-th GMM, and λ_{r_k} is the MRE estimated by Eq. (3.3.1) using the *k*-th GMM. Additionally, to perform the analysis numerically, Eq. (3.3.1) is solved in discrete form.

The outcome of PSHA is a continuous hazard curve that indicates the MRE of exceeding each *IM* value in a certain range of interest. To select a consistent GM suite, several

approaches have been proposed in the literature. This thesis considers the amplitude scaling method using Conditional Mean Spectra (CMS) for hazard-specific GM selection, and cloud-based GM selection for performance assessment.

3.3.1 Ground Motion Selection Based on Target Spectra using Amplitude Scaling

Hazard Deaggregation

The computation of hazard using Eqs. (3.3.1) and (3.3.2) is the aggregate result of all contributions, including all sources and GMMs considered, and integrating over magnitude and distance combinations. To establish hazard-specific scenarios, it is of interest to find those combinations that are more likely at a certain hazard level. This process is known as hazard deaggregation, and it is commonly used to develop target spectra that serve as a reference for GM selection.

The deaggregation of the hazard curve in terms of joint magnitude-distance scenarios (i.e., M-R deaggregation) consists of finding those scenarios that contribute the most to the aggregate hazard (Baker et al., 2021). Rather than integrating all contributions, the scenario is defined by Eq. (3.3.3)

$$P(M = m, R = r | IM > im) = \frac{\lambda(IM > im, M = m, R = r)}{\lambda(IM > im)} \tag{3.3.3}$$

Where the numerator $\lambda(IM > im, M = m, R = r)$ is obtained from Eq. (3.3.1) by summing over all sources but not integrating over M or R, which is expressed by Eq. (3.3.4)

$$\lambda(IM > im, M = m, R = r) = \sum_{i=1}^{n_{sources}} \lambda_s(M_i > m)\, P(IM > im | m, r) f_{M_i,R_i}(m, r) \tag{3.3.4}$$

The level of deaggregation can be defined by selecting which quantities to add up. For instance, Eq. (3.3.4) yields a deaggregation by GMM, as the logic tree is not applied. Similarly, it would be possible to define a deaggregation by zone contribution by applying the logic tree without summing up over all sources, and so on. With the intent of later developing target spectra, the deaggregation by GMM is considered for this thesis.

UNIFORM HAZARD SPECTRUM

The uniform hazard spectrum (UHS) is defined as the response spectrum which represents an equal hazard in all its coordinates. For instance, a UHS for spectral acceleration (i.e., $S_a(T)$) is defined by those $S_a(T)$ which are equally likely to be exceeded at every structural period T. To build a UHS for a certain hazard level, the hazard curve is obtained for a range of structural periods using Eq. (3.3.1). Then, the UHS is constructed by interpolating between all the curves at the desired MRE.

The resulting UHS, in consequence, does not represent a realistic response spectrum for any specific scenario because the MRE at the desired hazard level might be a consequence of vastly different events depending on T. For this reason, the UHS should be understood as neither a design spectrum nor a GM selection target spectrum, as it is not actually representative of the expected response spectrum of any given structure under any specific earthquake scenario. However, it can be seen as a useful representation of a collection of hazard-consistent IM values in a straightforward graphical setting.

From the construction of a UHS, target IM values at a specific period of interest can be drawn, which can be later used to construct more realistic target spectrum for design or GM selection.

CONDITIONAL MEAN SPECTRUM

One of those approaches is the construction of the Conditional Mean Spectrum (CMS) (Baker, 2011). The goal of the CMS is to provide a spectrum that is representative of an expected response spectrum conditioned to a target hazard at a specific period.

To build a CMS, the UHS is used as a reference point from which the specific target $IM(T = T^*)$ is obtained, but the spectral shape is constructed by using calibrated spectral correlation functions. Thus, the construction of a CMS, like its associated UHS, is GMM-dependent. This is due to the obtention of $IM(T^*)$. For clarity, the following description will be given considering that the IM used throughout this thesis is $S_a(T)$, although any IM of interest could be used.

The first step to construct a CMS is to obtain the expected mean response by evaluating a GMM at the desired target scenario. The target scenario $(M, R, lnS_a(T^*))$ is obtained from

the M-R deaggregation, and its response is given by the GMM predicted mean and standard deviation $\mu_{lnSa}(M,R,T^*), \sigma_{lnSa}(T^*)$, respectively. This is the standard form of a GMM prediction, most of which are defined in log-space and assume a lognormal distribution of the hazard.

The next step is to define the target spectral dispersion ε_{Sa}, which is expressed as the dispersion between the target $lnS_a(T^*)$ and its associated GMM prediction, given by Eq. (3.3.5). By extension,

$$\varepsilon_{Sa}(T^*) = \frac{lnS_a(T^*) - \mu_{lnSa}(M,R,T^*)}{\sigma_{lnSa}(T^*)} \tag{3.3.5}$$

From here, the mean response spectrum is constructed conditioned on $\mu_{lnSa}(M,R,T^*), \sigma_{lnSa}(T^*)$, and $\varepsilon_{lnSa}(T^*)$ by applying the spectral correlation function $\rho\big(\varepsilon_{lnSa}(T^*), \varepsilon_{lnSa}(T)\big) = \rho(T^*,T)$, which is calibrated to represent the expected shape of the target spectrum given T^*. The CMS is then defined by Eq. (3.3.6).

$$\mu_{lnSa(T)|lnSa(T^*)}(M,R,T^*) = \mu_{lnSa}(M,R,T) + \rho(T^*,T)\sigma_{lnSa}(T^*)\varepsilon(T^*) \tag{3.3.6}$$

Likewise, the conditional spectral standard deviation $\sigma_{lnSa(T)|lnSa(T^*)}$ is defined by Eq. (3.3.7)

$$\sigma_{lnSa(T)|lnSa(T^*)} = \sigma_{lnSa}(T)\sqrt{1-\rho^2(T^*,T)} \tag{3.3.7}$$

From Eqs. (3.3.6) and (3.3.7), a complete probabilistic conditional response spectrum is obtained, akin to a GMM prediction consisting of $\mu_{lnSa}, \sigma_{lnSa}$, although conditioned on a target value of $S_a(T^*)$. Termed as a Conditional Spectrum (CS), this provides a log-normally distributed prediction for the response spectrum of the structure of interest, that is grounded in PSHA, thus advancing generic GMM predictions, and solving the issue of UHS not being representative of any specific scenarios.

The final step to obtain a target CMS or CS is to combine all individual GMM-dependent spectra into one. For this purpose, Lin et al. (Lin, Harmsen et al., 2013) proposed four possible methods (i.e., (i) single GMM, (ii) aggregate M-R deaggregation and logic trees for the spectrum, (iii) GMM-dependent M-R deaggregation and spectrum construction, and logic trees for the combination, and (iv) exact CS computation during PSHA).

This thesis uses method 3, for which the procedure described in this section is performed to construct each GMM-dependent CMS, and a logic tree with the same weights as Eq. (3.3.2) is applied to combine them.

Extensive discussion on hazard deaggregation, target spectra, and GM selection based on those spectra can be referred to the work by Baker et al. and Lin et al. as comprehensive references (Baker et al., 2021; Lin, Harmsen et al., 2013; Lin, Haselton, & Baker, 2013a, 2013b).

GM SELECTION USING THE AMPLITUDE SCALING METHOD

Furthermore, target CMS/CS-based GM selection is often conducted for multiple stripes-based PSFA (MSA) (Baker, 2015). To this extent, after constructing the target spectrum, a suite of GMs is either selected from an existing database or simulated using synthetic-GM generation algorithms. This thesis focuses on selecting actual recorded GMs.

The goal of this selection is to obtain a GM suite that is consistent with the expected response spectrum which has been constructed for hazard-specific scenarios. For this purpose, the following procedure is established, known as the amplitude scaling method for GM selection:

1. From the target spectrum, a range of periods is defined. To account for possible elongations of the natural period due to stiffness degradation during the earthquake event, an upper bound is commonly set as $T_2 = 1.5T^*$. Likewise, to account for potential contribution of higher structure modes, a lower bound is commonly taken as $T_1 = 0.2T^*$. The resulting range of $[0.2T^*, 1.5T^*]$ is then taken as a reference range of fit within the target spectrum to measure the relevance of the GMs in the database to the scenario of interest.

2. The first screening of the database is performed, depending on the extent of information available about the events recorded. The goal is to initially retain those records which are consistent with the structure's site conditions. For instance, an initial screening based on soil quality can be considered.

3. The response spectra of the GM database is obtained for the IM used to construct the target spectrum. This thesis uses the median directional spectral acceleration response for this purpose (i.e., RotD50).

4. The difference between each record and the CMS is measured as $e(T) = |\mu_{lnSa(T)|lnSa(T^*)}(M,R,T) - RotD50(T)|$ within $[T_1, T_2]$.

5. The database is sorted in ascending order by $e(T)$ and the desired number of GMs nGM with the lowest discrepancy to the target spectrum are selected.

6. For each selected GM, a scaling factor $FS = \frac{\mu_{lnSa(T)|lnSa(T^*)}(M,R,T^*)}{RotD50(T^*)}$ is applied to the record.

By following this procedure, a suite of nGM records, scaled to match the target $S_a(T^*)$ at each specific hazard, is obtained. These are used as input excitations in model simulations during PSDA. Usually, nGM can be determined through design code requirements or research needs.

GROUND MOTION SELECTION FOR CLOUD-BASED ANALYSIS

Hazard-specific GM selection is useful for certain applications, such as performance-based design, or scenario-specific analysis. However, some applications benefit from considering a massive suite of non-hazard-specific GMs, with the goal of evaluating structural responses in a wider range of hazard scenarios.

The latter approach is known as cloud-based analysis (Jalayer et al., 2017; Jalayer et al., 2015), which aims to evaluate the structure under a "continuum" of hazard levels of interest using real, unscaled GM records. By doing so, a more complete notion of the uncertainty in structural response under different seismic events can be achieved. However, this approach could be less feasible to implement than the amplitude scaling method approach, depending on the availability of a relevant GM database.

In this thesis, the cloud-based analysis approach is taken to select a GM suite for application II. In this case, the selection process is not as algorithmic as the amplitude scaling method.

From the database, the GMs are selected manually based on their relevance to construct an appropriate cloud of points which covers a range of hazard levels of interest.

The criteria used herein was to build an extended database of subductive earthquakes, taken from Chilean and Japanese databases (Castro et al., 2022; Fayaz et al., 2024; National Research Institute for Earth Science and Disaster Resilience). From all records, those with a spectral acceleration $RotD50(T^*) > 0.2g$ were selected. The detail of this selection is described later in the results section in 4.4.2.

3.4 GAUSSIAN PROCESS SURROGATE MODELING

This section provides a general overview of GP as a data modeling method grounded in Bayesian statistics, as well as the development of SMs through GP. The broad topic of GP modeling is discussed at length in the relevant literature (e.g., (Álvarez et al., 2012; Rasmussen & Williams, 2005)).

The general form of a GP is that of a Gaussian probability density function (PDF), which describes the distribution of functions, rather than scalar values. If a certain data point $\{\mathbf{x}, y\}$ is considered, where $\mathbf{x} \in \mathbb{R}^{1\times m}$ is a vector of m parameters and $y \in \mathbb{R}$ is the observed output associated to the input $\mathbf{x}$, the set of all possible functions describing the data as $y = f(\mathbf{x})$ is assumed by the GP to be a Gaussian distribution with the form described by Eqs. (3.4.1) to (3.4.3):

$$y = f(\boldsymbol{x}) \sim GP\big(\mu(\boldsymbol{x}), k(\boldsymbol{x}, \boldsymbol{x}')\big) \tag{3.4.1}$$

$$\mu(\boldsymbol{x}) = E[f(\boldsymbol{x})] \tag{3.4.2}$$

$$k(\boldsymbol{x}, \boldsymbol{x}') = E\big[\big(f(\boldsymbol{x}) - \mu(\boldsymbol{x})\big)\big(f(\boldsymbol{x}') - \mu(\boldsymbol{x}')\big)\big] \tag{3.4.3}$$

where $\mathbf{x}$ and $\mathbf{x}'$ are different realizations of the parameters and $\mu(\mathbf{x})$ and $k(\mathbf{x}, \mathbf{x}') \in \mathbb{R}$ are the mean and covariance functions of the GP, respectively. The function $k(\mathbf{x}, \mathbf{x}')$ describes the covariance of $f(\mathbf{x})$ and $f(\mathbf{x}')$. It is common to assume, without loss of generality, that GPs have a zero-mean distribution (i.e., the data is 'de-meaned'). Then, the problem of fitting a GP to a certain dataset lies in the selection of an appropriate covariance function, also known as kernel, which is defined by its own vector of parameters $\boldsymbol{\theta}$, referred to as the hyperparameters of the GP. This is usually done by regression through likelihood

maximization of the GP over the dataset. The number of components of $\boldsymbol{\theta}$ and their meaning depends on the kernel selection.

In what follows, a dataset of n observations will be considered, in which case the data takes the form $\{\mathbf{X}, \mathbf{y}\}$ with $\mathbf{X} \in \mathbb{R}^{n \times m}$ and $\mathbf{y} \in \mathbb{R}^{n \times 1}$. An independent Gaussian noise term $\varepsilon_n \sim N(0, \sigma_n)$ is typically added to the variance to account for measurement noise and errors. Considering the above, the prior distribution of the GP takes the form of Eq. (3.4.4).

$$\boldsymbol{y}(\boldsymbol{X}) \sim N(\mathbf{0}, \boldsymbol{K}(\boldsymbol{X}, \boldsymbol{X}) + \sigma_n^2 \boldsymbol{I}) \tag{3.4.4}$$

Where $\mathbf{y}(\mathbf{X}) \in \mathbb{R}^{n \times 1}$ is the vector form of Eq. (3.4.1), for n observations (i.e., samples), $\mathbf{K}(\mathbf{X}, \mathbf{X})$ is the $n \times n$ kernel function in which $\mathbf{K}_{i,j} = k(\mathbf{x}_i, \mathbf{x}_j), i, j \in \{1, \dots, n\}$, $\sigma_n^2 \in \mathbb{R}$ is the variance of the noise term, and $\mathbf{I} \in \mathbb{R}^{n \times n}$ is the identity matrix. $\mathbf{K}(\mathbf{X}, \mathbf{X})$ is positive semidefinite.

The log marginal likelihood of the GP predicting $\mathbf{y}(\mathbf{X})$ is expressed as Eq. (3.4.5)

$$L = -\boldsymbol{y}(\boldsymbol{X})^T (\boldsymbol{K}(\boldsymbol{X}, \boldsymbol{X}) + \sigma_n^2 \boldsymbol{I})^{-1} \boldsymbol{y}(\boldsymbol{X}) - \frac{1}{2} \ln(|\boldsymbol{K}(\boldsymbol{X}, \boldsymbol{X}) + \sigma_n^2 \boldsymbol{I}|) - \frac{n}{2} \ln(2\pi) \tag{3.4.5}$$

By maximizing the likelihood (Rasmussen & Williams, 2005), the parameters defining $\mathbf{K}(\mathbf{X}, \mathbf{X})$ (i.e., $\boldsymbol{\theta}$) can be optimized (Snoek et al., 2012; Wu et al., 2019). Then, a new data point $y(\mathbf{x}^*)$ where $\mathbf{x}^* \in \mathbb{R}^{1 \times m}$ and $y(\mathbf{x}^*) \in \mathbb{R}$, can be predicted conditioned on the optimized $\boldsymbol{\theta}$, by computing the joint distribution of $\mathbf{y}(\mathbf{X})$ and $y(\mathbf{x}^*)$ as in Eq. (3.4.6):

$$\begin{bmatrix} \boldsymbol{y}(\boldsymbol{X}) \\ y(\boldsymbol{x}^*) \end{bmatrix} \sim N\left(\mathbf{0}, \begin{bmatrix} \boldsymbol{K}(\boldsymbol{X}, \boldsymbol{X}) + \sigma_n^2 \boldsymbol{I} & \boldsymbol{k}(\boldsymbol{X}, \boldsymbol{x}^*) \\ \boldsymbol{k}(\boldsymbol{x}^*, \boldsymbol{X}) & k(\boldsymbol{x}^*, \boldsymbol{x}^*) \end{bmatrix}\right) \tag{3.4.6}$$

Where $\mathbf{0} \in \mathbb{R}^{(n+1) \times 1}$, $\boldsymbol{k}(\mathbf{X}, \mathbf{x}^*) \in \mathbb{R}^{n \times 1}$ is the kernel evaluated at $\mathbf{X}$ and $\mathbf{x}^*$, and $k(\mathbf{x}^*, \mathbf{x}^*) \in \mathbb{R}$ is the kernel evaluated at $\mathbf{x}^*$.

Eq. (3.4.6) can be marginalized to obtain the predictive distribution of $y(\mathbf{x}^*)$ as Eq. (3.4.7)

$$y(\boldsymbol{x}^*) | \boldsymbol{x}^*, \boldsymbol{X}, \boldsymbol{y}(\boldsymbol{X}) \sim N(\mu(\boldsymbol{x}^*), \sigma^2(\boldsymbol{x}^*)) \tag{3.4.7}$$

In which,

$$\mu(\boldsymbol{x}^*) = \boldsymbol{k}(\boldsymbol{x}^*, \boldsymbol{X})^T (\boldsymbol{K}(\boldsymbol{X}, \boldsymbol{X}) + \sigma_n^2 \boldsymbol{I})^{-1} \boldsymbol{y}(\boldsymbol{X}) \tag{3.4.8}$$

$$\sigma^2(\boldsymbol{x}^*) = k(\boldsymbol{x}^*, \boldsymbol{x}^*) - \boldsymbol{k}(\boldsymbol{x}^*, \boldsymbol{X})(\boldsymbol{K}(\boldsymbol{X}, \boldsymbol{X}) + \sigma_n^2 \boldsymbol{I})^{-1} \boldsymbol{k}(\boldsymbol{X}, \boldsymbol{x}^*) \tag{3.4.9}$$

Where $\mu(\mathbf{x}^*) \in \mathbb{R}$ in Eq. (3.4.8) and $\sigma^2(\mathbf{x}^*) \in \mathbb{R}$ in Eq. are the mean and variance of the posterior predictive distribution of the GP for the new data point $\mathbf{x}^*$ conditioned on $\mathbf{X}$, respectively.

In the context of machine learning and surrogate modeling, the dataset $\{\mathbf{X}_{tr}, \mathbf{y}(\mathbf{X}_{tr})\}$ is labeled as the training dataset consisting of n_{tr} data points, which produces $\mathbf{y}_{tr} = \mathbf{y}(\mathbf{X}_{tr}) \in \mathbb{R}^{n_{tr} \times 1}$. By extension of Eqs. (3.4.7) to (3.4.9), to generate predictions on a set of n_{te} testing data points, the posterior predictive distribution is evaluated at the dataset $\{\mathbf{X}_{te}, \mathbf{y}_{te}\}$, where $\mathbf{X}_{te} \in \mathbb{R}^{n_{te} \times m}$ and $\mathbf{y}_{te} \in \mathbb{R}^{n_{te} \times 1}$. Thus, $\boldsymbol{\mu}(\mathbf{X}_{te})$ and $\boldsymbol{\sigma}^2(\mathbf{X}_{te}) \in \mathbb{R}^{n_{te} \times 1}$ are obtained, which are the posterior mean and variance of the GP over a set of new, unseen data.

APPLICATION OF GP SURROGATE MODELING TO STRUCTURAL FINITE ELEMENT MODELS

Structural FE models can be expressed as the mapping $\mathbf{y}_{\text{FE}} = g(\overline{\boldsymbol{\Theta}})$, where $\overline{\boldsymbol{\Theta}} \in \mathbb{R}^{1 \times d}$ is the input vector of d model parameters and $\mathbf{y}_{\text{FE}} \in \mathbb{R}^{1 \times n_r}$ is the output vector of n_r model responses. Currently, nonlinear FE models are extensively used to simulate the behavior of structures subjected to dynamic excitations, such as earthquakes, for which $\mathbf{y}_{\text{FE}}$ is computed through NLTHA.

In probabilistic structural analysis, it is acknowledged that the model parameters in $\overline{\boldsymbol{\Theta}}$ are uncertain due to natural variability of material properties, environmental conditions, and other sources of uncertainty (i.e., aleatoric variability), as well as a lack of knowledge about the actual phenomena (i.e., epistemic sources) (Kiureghian & Ditlevsen, 2009). To assess structural performance adequately, it is crucial to account for these sources of uncertainty. To this end, probabilistic FE models are established through numerical simulation. Uncertainties in $\overline{\boldsymbol{\Theta}}$ are considered through PDFs for the model parameters, i.e., $\overline{\boldsymbol{\Theta}} \sim \boldsymbol{f}_{\boldsymbol{\Theta}}(\boldsymbol{\mu}_{\boldsymbol{\Theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\Theta}})$, where $\boldsymbol{f}_{\boldsymbol{\Theta}}$ is the multivariate PDF of $\overline{\boldsymbol{\Theta}}$, with mean $\boldsymbol{\mu}_{\boldsymbol{\Theta}} \in \mathbb{R}^{1 \times d}$ and covariance $\boldsymbol{\Sigma}_{\boldsymbol{\Theta}} \in \mathbb{R}^{d \times d}$. Thus, a model sample produces a probabilistic characterization of $\mathbf{y}_{\text{FE}}$ given by $\mathbf{y}_{\text{FE}} \sim \boldsymbol{f}_{\mathbf{y}}(\boldsymbol{\mu}_{\mathbf{y}}, \boldsymbol{\Sigma}_{\mathbf{y}})$, where $\boldsymbol{\mu}_{\mathbf{y}} \in \mathbb{R}^{1 \times n_r}$ and $\boldsymbol{\Sigma}_{\mathbf{y}} \in \mathbb{R}^{n_r \times n_r}$.

The computational cost of such analysis, however, increases proportionally to the number of samples drawn, as the model evaluation of n_s samples generates the matrices $\boldsymbol{\Theta} \in \mathbb{R}^{n_s \times d}$ and $\mathbf{Y}_{\text{FE}} \in \mathbb{R}^{n_s \times n_r}$. For FE models of large civil structures, this rapidly becomes a restriction due

to limited availability of computational resources. To mitigate the cost of a full-blown probabilistic FE simulation framework, SMs such as GP can be trained using a dataset $\{\boldsymbol{\Theta}^{tr}, \mathbf{Y}_{\mathrm{FE}}^{tr}\}$ that is representative of the statistical distribution of $\mathbf{Y}_{\mathrm{FE}}$, yet is practical to produce by simulation. The trained SM can then be used to generate further samples on the new dataset $\{\boldsymbol{\Theta}^{te}, \mathbf{Y}_{\mathrm{FE}}^{te}\}$ for uncertainty analysis.

Developing SMs for real structures, such as the application presented later in this paper, poses an additional challenge compared to simpler academic examples. Due to the added complexity in model inputs (i.e., number of parameters, probabilistic characterization of parameter PDFs and correlations) and outputs (i.e., non-smooth distribution of model responses, increased level of nonlinearity due to component interactions and geometric discontinuities such as gaps), the design of the SMs needs to be tailored carefully to capture the underlying relationships between model parameters and responses.

SETUP OF THE GP MODELS

The main challenge of fitting a GP model to a certain dataset lies in the definition of an appropriate kernel function and its hyperparameters $\boldsymbol{\theta}_{GP}$. The selection of a kernel primarily depends on the characteristics of the dataset, for example, its smoothness or periodicity. Common kernel functions include the squared exponential (also known as RBF), Matérn, rational quadratic, sinusoidal, among others (Rasmussen & Williams, 2005).

For RBF and Matérn kernels, the basic version of $\boldsymbol{\theta}_{GP} = [l, \sigma_f] \in \mathbb{R}^{2\times 1}$ consists of the lengthscale l, which determines the smoothness of the covariance function, and the signal variance σ_f^2, which determines the amplitude. Additionally, it is possible to implement Automatic Relevance Determination (ARD), by assigning each parameter its own length scale. Thus, the vector of hyperparameters to be optimized is $\boldsymbol{\theta}_{\boldsymbol{GP}} = [\boldsymbol{l}, \sigma_f] = [l_1, \dots, l_m, \sigma_f] \in \mathbb{R}^{m+1\times 1}$.

A RBF kernel is appropriate to model relatively smooth data, and is commonly the default choice for a kernel function in structural engineering applications (e.g., (Gentile & Galasso, 2020; Ning et al., 2024; Su et al., 2017)). Its ARD version is defined by Eq. (3.4.10) for two independent realizations of $\boldsymbol{\Theta}$, $\mathbf{x}$ and $\mathbf{x}' \in \mathbb{R}^{1\times m}$.

$$k(\boldsymbol{x}, \boldsymbol{x}') = \sigma_f^2 \exp\left(-\frac{1}{2}\sum_{i=1}^{m}\frac{(x_i - x_i')^2}{l_i^2}\right) \tag{3.4.10}$$

Matérn kernels are suitable for less smooth data, as the smoothness of the kernel is controlled by the parameter ν, as defined in Eq. (3.4.11):

$$k(\boldsymbol{x}, \boldsymbol{x}') = \sigma_f^2 \frac{2^{1-\nu}}{\Gamma(\nu)}\left(\frac{\sqrt{2\nu}}{l}\|(\boldsymbol{x} - \boldsymbol{x}')\|\right)^{\nu} K_{\nu}\left(\frac{\sqrt{2\nu}}{l}\|(\boldsymbol{x} - \boldsymbol{x}')\|\right) \tag{3.4.11}$$

Where $\Gamma(\cdot), \|\cdot\|, K_{\nu}$ are the gamma function, euclidean distance and modified Bessel function, respectively. Usually, ν takes a value of 1/2, 3/2, or 5/2, with the higher value of ν providing smoother kernels. Eq. (3.4.12) defines the Matérn kernel with $\nu = 3/2$ and ARD.

$$k(\boldsymbol{x}, \boldsymbol{x}') = \sigma_f^2\left(1 + \sqrt{3}\sqrt{\sum_{i=1}^{m}\frac{(x_i - x_i')^2}{l_i^2}}\right)\exp\left(-\sqrt{3}\sqrt{\sum_{i=1}^{m}\frac{(x_i - x_i')^2}{l_i^2}}\right) \tag{3.4.12}$$

In this work, RBF and Matérn with $\nu = 3/2$ are used due to their suitability to describe smooth and noisy data, respectively, following preliminary testing of several kernels. It is worth noting that ARD kernels allow for automatic feature selection due to the use of individual length scales. This circumvents potential dimensionality problems associated with GP models with large feature spaces. As a result, no prior assumptions need to be made about the sensitivity of model parameters.

To define the GP priors, a general expression is used for the length scale of each parameter to initialize $l_i, i \in \{1, \ldots, m\}$ defined as the mean Euclidean distance between sample points. The prior signal standard deviation is set for the GP of each response $\mathbf{y}_{\mathrm{FE},i} \in \mathbb{R}^{n_s \times 1}$ as $\sigma_f = \sigma(\mathbf{y}_{\mathrm{FE},i})/\sqrt{2}$. Given that the data modeled in this work comes from the original FE models, the variance due to measurement noise is expected to be close to zero, although a low prior value of $\sigma_n = \sigma_f/10$ is used for numerical stability. Maximum likelihood with the L-BFGS (Liu & Nocedal, 1989) algorithm is used to optimize $\boldsymbol{\theta}_{GP}$. The data is standardized (i.e., $z = (x - mean(x))/std(x)$) during optimization, as is common practice in GP modeling (Rasmussen & Williams, 2005). The weak Wolfe (Wolfe, 1969) line search method is used for L-BFGS, with a Hessian memory size of 15 iterations and a convergence criteria of 10^{-2} for the gradient descent.

After optimization of $\boldsymbol{\theta}_{GP}$ on the training dataset $\{\boldsymbol{\Theta}^{tr}, \mathbf{Y}_{\mathrm{FE}}^{tr}\}$, the GP SMs are validated by generating predictions on the testing dataset $\{\boldsymbol{\Theta}^{te}, \mathbf{Y}_{\mathrm{FE}}^{te}\}$, by evaluating the posterior distribution in Eq. (3.4.7) for each response in $\mathbf{Y}_{\mathrm{FE}}$. This produces the GP predictions on the testing dataset, $\{\boldsymbol{\Theta}^{te}, \mathbf{Y}_{\mathrm{GP}}^{te}\}$. Then, the quality of the predictions is assessed through goodness-of-fit metrics. Throughout section 4, three commonly used metrics are considered, namely coefficient of determination (R^2), relative root mean squared error ($RRMSE$), and mean absolute error (MAE). Chapter 0 considers R^2. The metrics are defined in Eqs. (3.4.13) to (3.4.15), respectively.

$$R^2(\boldsymbol{y}_{FE}^{te}, \boldsymbol{y}_{GP}^{te}) = 1 - \frac{\sum_{i=1}^{n_{te}} \left(y_{FE,i}^{te} - y_{GP,i}^{te}\right)^2}{\sum_{i=1}^{n_{te}} \left(y_{FE,i}^{te} - \mu_{\boldsymbol{y}_{FE}^{te}}\right)^2} \tag{3.4.13}$$

$$RRMSE(\boldsymbol{y}_{FE}^{te}, \boldsymbol{y}_{GP}^{te}) = \frac{\sqrt{\frac{1}{n_{te}} \sum_{i=1}^{n_{te}} \left(y_{FE,i}^{te} - y_{GP,i}^{te}\right)^2}}{\sqrt{\frac{1}{n_{te}} \sum_{i=1}^{n_{te}} \left(y_{FE,i}^{te}\right)^2}} * 100\% \tag{3.4.14}$$

$$MAE(\boldsymbol{y}_{FE}^{te}, \boldsymbol{y}_{GP}^{te}) = \frac{1}{n_{te}} \sum_{i=1}^{n_{te}} |y_{FE,i}^{te} - y_{GP,i}^{te}| \tag{3.4.15}$$

Where $\mu_{\mathbf{y}_{\mathrm{FE}}^{te}} \in \mathbb{R}$ is the mean value of $\mathbf{y}_{\mathrm{FE}}^{te}$.

3.5 PROBABILISTIC SEISMIC DEMAND AND FRAGILITY ANALYSIS

3.5.1 LS AND CAPACITY DEFINITION

Following the development of the SMs for several structural responses of interest, their associated EDPs need to be defined for demand and fragility modeling. This study considers local EDPs, defined at a material or component level, in line with recent recommendations for fragility assessment in PBEE (Deb et al., 2022). To this end, LS for each EDP can be defined either referred from available literature or defined for the specific application.

Regarding component capacity at each LS, probabilistic definitions are established by MC sampling of $\boldsymbol{\Theta}$. A large number of parameter samples is drawn and the capacity at each LS is computed from the constitutive model. From here, a mean (i.e., $\mu_{c,LS}$) and log-standard deviation (i.e., $\beta_{c,LS}$) are obtained for the capacity at each LS for every response of interest.

3.5.2 EDP DEFINITION

EDPs herein are defined as normalized quantities in logarithmic space, by taking the ratio between model responses and the capacity at their first LS. This is expressed as: $\boldsymbol{\delta}_{FE} = \ln(\mathbf{y}_{FE}/\mu_{c,LS1})$ and $\boldsymbol{\delta}_{GP} = \ln(\mathbf{y}_{GP}/\mu_{c,LS1})$.

3.5.3 PSDM

Deterministic log-linear PSDM have been widely used to approximate expected demands on a certain component depending on *IM* values (Shome & Cornell, 1999; Soleimani & Hajializadeh, 2022). The log-linear form is expressed by Eq. (3.5.1). However, a probabilistic form can be established conditioned on parametric uncertainty by performing the linear regression on each sample of the model parameters $\boldsymbol{\Theta}$, producing the form in Eq. (3.5.2).

$$\delta_{PSDM}(IM) = a + b\ln(IM) + N(0, \beta_D^2) \tag{3.5.1}$$

$$\delta_{PSDM}(IM|\boldsymbol{\Theta}) = a(\boldsymbol{\Theta}) + b(\boldsymbol{\Theta})\ln(IM) + N(0, \beta_{D|\boldsymbol{\Theta}}^2) \tag{3.5.2}$$

where a and b are the linear regression coefficients. The variance of residuals is computed as $\beta_D^2 = \mathrm{Var}(\delta_{PSDM}(IM) - \delta)$ in Eq. (3.5.1) and $\beta_{D|\boldsymbol{\Theta}}^2 = \mathrm{Var}(\delta_{PSDM}(IM|\boldsymbol{\Theta}) - \delta(\boldsymbol{\Theta}))$ in Eq. (3.5.2), respectively. In this context, δ and $\delta(\boldsymbol{\Theta})$ represent the "true" model responses (i.e., $\boldsymbol{\delta}_{FE}$ or $\boldsymbol{\delta}_{GP}$, accordingly).

When using SMs, GP prediction uncertainty also needs to be included as a contribution to demand uncertainty. For this purpose, σ_{GP} from Eq. (3.4.9) is also normalized and transformed into the log-space using Eq. (3.5.3).

$$\beta_{GP}^2|\boldsymbol{X} = \ln\left(1 + \frac{\sigma_{GP}^2}{\mu_{GP}^2}\right) \tag{3.5.3}$$

From this expression, $\beta_{GP}^2|\mathbf{X} \in \mathbb{R}^{n_s \times 1}$ yields a measure of prediction uncertainty at every sample of $\mathbf{X}$. When $\mathbf{X}$ includes features of model parameters as well as GMs, this leads to an *IM*-dependent value of $\beta_{D,GP}$. Although this would be the more rigorous approach, to ensure monotonicity in fragility functions in later stages, $\beta_{D,GP}|\mathbf{X} = median(\beta_{GP}|\mathbf{X})$ is used for the PSDM. Then, the expression of demand variance for GP surrogates becomes: $\beta_{D|\boldsymbol{\Theta}}^2 = \mathrm{Var}(\delta_{PSDM}(IM|\boldsymbol{\Theta}) - \delta_{GP}(\boldsymbol{\Theta})) + \beta_{D,GP}^2|\mathbf{X}$.

3.5.4 CONDITIONAL FRAGILITY ANALYSIS

The cloud-based conditional fragility curve can be expressed by Eq. (3.5.4) (Fatemeh Jalayer, 2003; Nielson & DesRoches, 2007).

$$P\left(\delta > \mu_{c,LS} \middle| \boldsymbol{\Theta}\right) = \Phi\left(\frac{\delta_{PSDM}(IM|\boldsymbol{\Theta}) - \log\left(\mu_{c,LS}\right)}{\beta_{tot}|\boldsymbol{\Theta}}\right) \tag{3.5.4}$$

where $\beta_{tot}|\boldsymbol{\Theta} = \sqrt{\beta_{LS}^2 + \beta_{c,LS}^2 + \beta_{D|\boldsymbol{\Theta}}^2}$ is the total combined uncertainty, including uncertainties in LS, capacity, and demand definitions. The epistemic uncertainty from GP surrogates is directly propagated to this stage through $\beta_{D|\boldsymbol{\Theta}}$.

By computing fragility functions in this manner, the uncertainties in input excitations, parametric variability, as well as constitutive and surrogate modeling accuracy, are propagated to probabilistic fragility assessment. Sampling $\boldsymbol{\Theta}$ and performing the conditional fragility analysis on each sample generates a probabilistic description of damage probabilities, from which statistics of interest can be derived, such as mean fragility and 95% confidence intervals (CI).

Finally, a probabilistic characterization of structural risk can be established by integrating fragility and hazard descriptions (Eads et al., 2013), although this does not include loss estimations. In this study, the conditional mean annual exceedance rate of each LS for every component of interest are computed according to Eq. (3.5.5).

$$\lambda(LS|\boldsymbol{\Theta}) = \int_{im \in IM} P\left(\delta > \mu_{c,LS} \middle| \boldsymbol{\Theta}\right) \left|\frac{d\lambda(im)}{dim}\right| dim \tag{3.5.5}$$

Although a complete risk assessment would include the characterization of decision variables, these are beyond the scope of this study. However, establishing a probabilistic measure for $\lambda(LS|\boldsymbol{\Theta})$ enables discussion within PSFA about: (i) what the most critical structural components are, based on their individual LS exceedance rates and their consequences; and (ii) how consistently locally defined LS for different components relate to one another in measuring structural performance.

As shown through the application example in following sections, conditional fragility estimates derived from the original FE model samples (i.e., $\mathbf{Y}_{FE}$) are compared to those

obtained by new sampling of the trained GP surrogates for validation purposes. The ability to generate samples of $\mathbf{Y}_{GP}$, which can correctly interpolate predictions for new model parameters, enables further sampling-based analysis, such as GSA based on Sobol indices obtained by simulation, at a reduced computational cost.

3.6 Proposed Methodology for Application I: Study of Parameter-Induced Uncertainty in Mechanics-Based FE Models leveraging Gaussian Process Surrogate Models

This chapter proposes a methodology for the efficient analysis of parameter-induced uncertainty in structure-level FE models, supported by surrogate models. To this end, the proposed methodology considers (i) the probabilistic characterization of structural material and component constitutive models through Bayesian calibration informed by experimental data, (ii) the selection of single earthquake scenarios representative of the expected hazard, (iii) the development and validation of surrogate models, and (iv) the decomposition of parameter-induced uncertainty via surrogate-based sensitivity analysis (SA).

For this task, GP surrogate models present an idoneous tool due to their inherently probabilistic and straightforward-to-interpret formulation. Although other statistical machine learning techniques could be used for SA (e.g., PCE (Sudret, 2008), SVR (Cheng et al., 2017), multi-fidelity methods (Shang et al., 2023)), GP has been selected for this study for two main reasons: (i) the probabilistic formulation of GP can directly translate to further uncertainty propagation in performance assessment, as the predictive variance can be included as a source of epistemic uncertainty in structural demands, and (ii) the nonparametric kernel-based formulation allows for flexibility in fitting non-smooth data and capturing the underlying nonlinear parameter interactions. The proposed methodology is implemented and later validated through two case study examples. The proposed method is depicted in Figure 3.6-1.

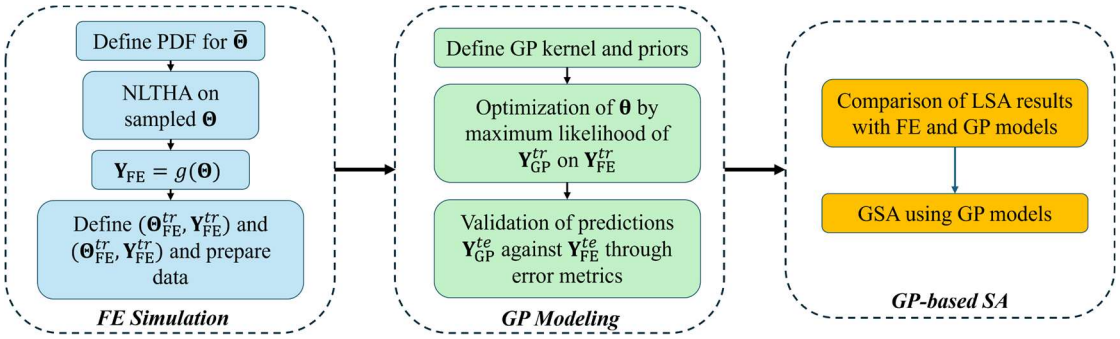


Figure 3.6-1. Flowchart of the proposed methodology.

The workflow follows five main steps, which can be summarized as follows:

1) FE models are simulated by sampling the joint distribution $\bar{\boldsymbol{\Theta}}$ of the input parameters, thus generating the sample $\boldsymbol{\Theta}$ and performing NLTHA to obtain the responses of interest $\mathbf{Y}_{\mathrm{FE}}$ (e.g., the peak nodal displacement from each sample). The PDF which is sampled to generate $\boldsymbol{\Theta}$ is adopted from probabilistic characterizations of the model parameters.

2) The dataset is separated into training and testing datasets, $\{\boldsymbol{\Theta}^{tr}, \mathbf{Y}_{\mathrm{FE}}^{tr}\}$ and $\{\boldsymbol{\Theta}^{te}, \mathbf{Y}_{\mathrm{FE}}^{te}\}$, respectively. It is noted that later in the application examples, in the first case study, the data is kept directly as response values, while in the second case, it is further transformed to demand parameters.

3) A GP SM is developed for each response independently. The GP prior for each response is established by defining a kernel function and prior distributions for $\boldsymbol{\theta}_{GP}$, which is then optimized by maximizing the likelihood on the training dataset. Independent GPs are chosen over a multivariate GP approach under the premise that independent models with a sufficiently good fit implicitly maintain the correlation structure between model responses.

4) The optimized GP is tested against the testing dataset and the predictive ability of each SM is measured with three commonly used goodness-of-fit metrics, which will be defined later.

5) Finally, the GP SMs are used to perform LSA and GSA of each response to all parameters. LSA results are presented to validate the ability of GP surrogates to capture underlying relationships between parameters and peak responses, by comparing the indices obtained from the GPs and FE model. Then, GSA is carried out using GP SMs.

3.7 PROPOSED METHODOLOGY FOR APPLICATION II: UQ&P AND ITS DECOMPOSITION IN PBEE SUPPORTED BY GP SURROGATE MODELS

This chapter proposes a novel methodology to quantify and propagate uncertainty in: (i) hazard estimates and ground motion (GM) selection; (ii) structural model parameter uncertainty (arising from material and component properties and modeling); (iii) capacity and LS definitions; (iv) demand models; and (v) fragility estimates. GP surrogates are developed to directly model probabilistic structural responses, enabling their use for further sampling of a structure with a new set of parameter values subjected to new GMs. The reduced computational cost of the GP surrogates is further exploited by performing Sobol-based GSA of structural demands with respect to model parameters. A flowchart of the proposed methodology is illustrated in Figure 3.7-1.

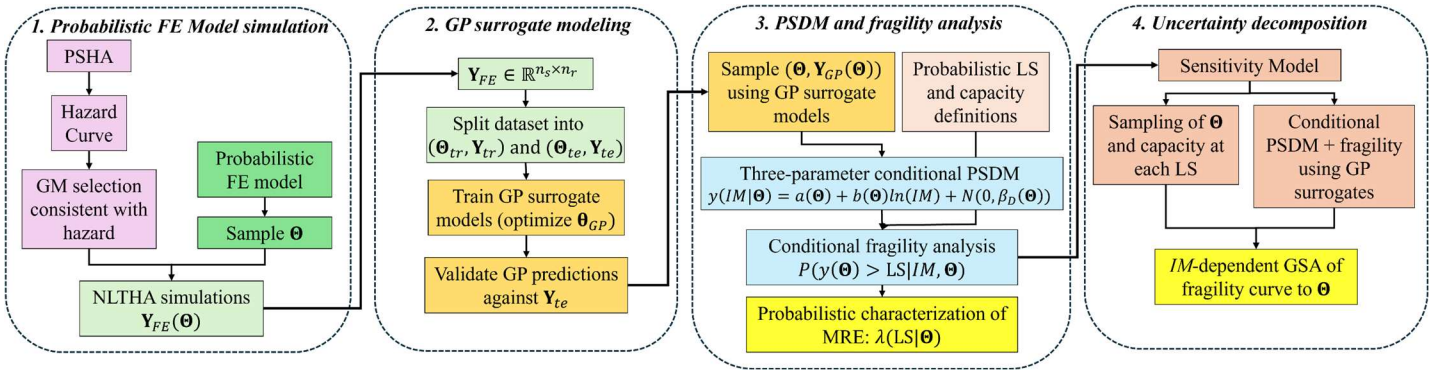


Figure 3.7-1. Flowchart of the proposed methodology.

In stage 1, seismic hazard uncertainty is quantified by PSHA, which considers a vast range of realistic earthquake scenarios and subsequently selects the suite of actual recorded GMs consistent with the estimated hazard. In parallel, a probabilistic FE model of the structure is developed through Bayesian calibration of its nonlinear components and material constitutive models using experimental data. Then, selected GMs are used as inputs to perform nonlinear time-history analysis (NLTHA) using the probabilistic FE model, where the scheme accounts for the combination of GMs and FE model parameter realizations. Thus, the first stage focuses on quantifying and propagating uncertainty about input excitations and model behavior by obtaining $\mathbf{Y}_{FE}(\boldsymbol{\Theta}) \in \mathbb{R}^{n_s \times n_r}$, where $\boldsymbol{\Theta}$ is the matrix containing all model parameter samples, n_s is the number of samples or simulations, n_r is the number of recorded model responses, and $\mathbf{Y}_{FE}$ is the matrix of all recorded model responses.

In stage 2, the goal is to develop and validate SMs to enable further sampling of the model while achieving a drastically lower computational cost. To this end, the dataset $(\boldsymbol{\Theta}, \mathbf{Y}_{FE})$ is split into training $(\boldsymbol{\Theta}_{tr}, \mathbf{Y}_{tr})$ and testing $(\boldsymbol{\Theta}_{te}, \mathbf{Y}_{te})$ datasets. GP surrogates are selected due to their probabilistic nature as a Bayesian method and nonparametric formulation. This approach provides a natural continuation of the Bayesian framework adopted for model calibration in stage 1, in addition to the ease of interpretation of its predictions. GP surrogates are trained by optimizing their hyperparameters $\boldsymbol{\theta}_{GP}$ through maximum likelihood estimation on $(\boldsymbol{\Theta}_{tr}, \mathbf{Y}_{tr})$. Subsequently, the quality of the SMs and their ability to predict unobserved cases in new data are validated by testing their predictions on $(\boldsymbol{\Theta}_{te}, \mathbf{Y}_{te})$, which includes model parameter realizations and GMs unseen during the training process.

Once the GP surrogates have been validated, they can be used in place of the original FE model for subsequent response predictions. In stage 3, three-parameter log-linear probabilistic seismic demand models (PSDM) are defined for each component of interest by sampling the GP surrogates on a large set of new $\boldsymbol{\Theta}$ realizations. The conditional PSDM can be established as a function of IM conditioned on $\boldsymbol{\Theta}$ as $\delta_{PSDM}(IM|\boldsymbol{\Theta}) = a(\boldsymbol{\Theta}) + b(\boldsymbol{\Theta}) \ln(IM) + N\left(0, \beta_{D|\boldsymbol{\Theta}}\right)$ (Jalayer et al., 2017; Yan et al., 2024), where a and b are linear regression coefficients, and $\beta_{D|\boldsymbol{\Theta}}$ represents the conditional uncertainty in the demand as a Gaussian error term. Through this approach, the uncertainty in excitation and structural behavior is propagated to fragility analysis. On the other hand, the uncertainty in component capacity is quantified via sampling at the constitutive model level. Additionally, for each EDP, uncertainties in LS definitions can be obtained from the literature and past statistical data, if available. Conditional fragility analysis for each component is then conducted by incorporating the conditional PSDM and the uncertainties in demand, capacity, LS, and GP model predictions. The fragility function is then sampled as $P(\delta_{PSDM}(IM|\boldsymbol{\Theta}) > \mathrm{LS}|IM, \boldsymbol{\Theta})$. Finally, by integrating the fragility and hazard curves, a probabilistic definition of the MRE of each LS can be obtained (i.e., $\lambda(LS|\boldsymbol{\Theta})$).

In stage 4, the reduced cost of sampling GP surrogates is further leveraged to perform simulation-based GSA with total Sobol indices. During this stage, model parameters are extensively sampled along with their respective component capacities. Subsequently, the uncertainty induced by each model parameter in the fragility estimates for every component

is quantified. As a result, uncertainty decomposition in all fragility estimates is achieved as a function of *IM*.

4 RESULTS AND DISCUSSION

4.1 A SIMPLIFIED TRI-LINEAR CONSTITUTIVE MODEL FOR EXTERIOR SHEAR KEYS

4.1.1 CONTEXT TO THE TRI-LINEAR CONSTITUTIVE MODEL FOR EXTERIOR SHEAR KEYS

This section addresses the gap highlighted in section 2.4 by proposing a simplified tri-linear formulation that can be directly combined with the tri-linear DT model proposed by Goel and Chopra (Goel & Chopra, 2008) to cover both major failure modes (i.e., DT and SS) for the FE modeling of bridge components and systems including exterior monolithic shear keys.

The proposed model adopts the expressions of Kottari et al. (Kottari et al., 2020) to determine the failure mode. The equations used to model SS capacity are based on those proposed by Kottari et al. (Kottari et al., 2020) and Bozorgzadeh et al. (Bozorgzadeh, Megally, Ashford et al., 2006). New expressions are developed for displacement predictions based on experimental data reported by Kottari (Kottari, 2016). The proposed model is calibrated and validated in a probabilistic setting, adopting the Bayesian framework previously used in the calibration of constitutive models for steel reinforcement and bridge component response (Birrell, Astroza, Carreño et al., 2021; Birrell, Astroza, Restrepo et al., 2021; Pinto et al., 2023), which has shown good results. The proposed model and its probabilistic calibration allow to conduct nonlinear FE simulations of bridges accounting for the main failure modes in shear keys (i.e., DT and SS), which is crucial for fragility/reliability analysis and performance assessment of these structures when subjected to seismic loading.

The goal of this model is to provide a simplified formulation to be implemented alongside DT failure in time-history response simulations of structure-level finite element models of bridges, thus covering both major failure mechanisms in monolithic reinforced concrete shear keys.

4.1.2 EXPERIMENTAL DATA ON MONOLITHIC KEYS FAILING IN SLIDING SHEAR

Available experimental data from lateral force-displacement testing conducted on monolithic exterior shear keys failing in SS is scarce. Kottari (Kottari, 2016) studied results from previous experimental campaigns and reported test data for six new test specimens, carried out as part of an experimental program at the University of California, San Diego (UCSD). Specifically, their study considered six tests by Megally et al. (Megally et al., 2002) (specimens 1A-3B) and six tests by Bozorgzadeh et al. (Bozorgzadeh, Megally, Ashford et al., 2006) (specimens 4A-6B) on monolithic (i.e., non-isolated) and isolated shear keys. Kottari (Kottari, 2016) extended the database by testing twelve new monolithic and isolated exterior shear keys, both non-skewed and skewed (specimens 7A-12B). Out of the 24 experiments, only six specimens failed in SS (8A-10B).

Kottari (Kottari, 2016) extended the existing experimental campaign on bridge shear keys carried out at University of California, San Diego (UCSD). Previously, Megally et al. (Megally et al., 2002) and Bozorgzadeh et al. (Bozorgzadeh, Megally, Ashford et al., 2006) had conducted tests on twelve shear keys of both monolithic and isolated construction which failed in DT and SF (with specimen ID's 1A, 1B,…, 6A, 6B). In (Kottari, 2016), twelve new specimens were tested (ID's 7A-12B), of which six failed in SS (ID's 8A-10B).

The experimental lateral force-displacement responses reported by Kottari (Kottari, 2016), particularly for test specimens 8A to 10B related to SS failure, were considered in this study. Experimental data from Han et al. (Han, Zhou, Ou et al., 2017; Han, Zhou, Zhong et al., 2017) and Mei & Guo (Mei & Guo, 2023) were not considered due to the unclear interaction of shear and flexure mechanisms that led to a post sliding peak in capacity in those tests. Specimens 8A-10B were built to the same geometrical dimensions and test setup that were previously employed for the specimens tested by Megally et al. (Megally et al., 2002) and Bozorgzadeh (Bozorgzadeh, Megally, Ashford et al., 2006). Two target strengths of concrete were considered, and the main detailing difference between specimens was the amount and distribution of vertical side reinforcement (i.e., dowel bars). All test series consisted of two shear keys built monolithically at a scale of 1:2.5. The cap beam had a height of 77.5 cm, a total length of 350.5 cm, and a width of 42.6 cm. All shear keys had a total height of 61.0 cm

and the same width as the abutment wall (42.6 cm). One side of each specimen (series A) was tested using a shear key with an inclined face according to the geometry commonly employed for girders in California. The shear key in the opposite side of the specimen (series B) was built with a vertical face towards the inner side. Both shear keys (series A and series B) had the same base length of 61.0 cm, but the inclined shear key (series A) had a top length of 43.2 cm. Series 8 and 9 had a target concrete strength of 27.6 MPa, while series 10 of 41.4 MPa. In all cases, five #8 (25 mm diameter) grade 60 (yield strength of 414 MPa) headed bars were placed near the top of the cap beam to provide shear reinforcement. Horizontally and vertically distributed reinforcement steel bars were also installed, consisting of #3 (10 mm diameter) grade 60 (420 MPa) bars at 11.4 cm and 10.2 cm spacing, respectively. Detailed schematics of the tested specimens were reported by Kottari (Kottari, 2016). Table 4.1-1summarizes measured material properties and relevant design specifications employed in this study.

Shear key tests were performed following a mixed loading protocol, which started with load control until a specific load value was reached; after that, the test was switched to displacement control until failure. Three main phases (Figure 4.1-4) describe the force-displacement curve for the lateral response of shear keys with SS failure, including: (a) elastic and cracked behavior before reaching the peak sliding resistance, (b) progressive degradation of concrete contribution following the initial sliding, and (c) complete failure of the shear key following the fracture of dowel bars. These are illustrated by photos taken during the tests, as shown in Figure 4.1-1. Lateral and vertical displacements and in-plane and out-of-plane rotations were measured and reported by Kottari (Kottari, 2016). Figure 4.1-2 shows the lateral force-displacement curves recorded from each test, which are consequently employed for calibration in section 4.

Table 4.1-1. Material properties for shear key test specimens.

Test ID	f_c' (MPa)	f_y/f_u #8 bars (MPa)	f_y/f_u #3 bars (MPa)	A_{vf} (cm^2)	# of dowel bars per face
8A	32.5	483/648	462/717	4.28	3
8B					
9A	35.2	469/614		7.12	5
9B					
10A	46.5			4.28	3
10B					

f_c': Compressive strength of concrete. f_y and f_u: yield and tensile strength of steel, respectively. A_{vf}: Total area of vertical side steel reinforcement crossing the joint surface.

Figure 4.1-1. Damage progression for specimen 8A: (a) before sliding, (b) after concrete strength degradation, (c) after failure. (Courtesy of A. Kottari)

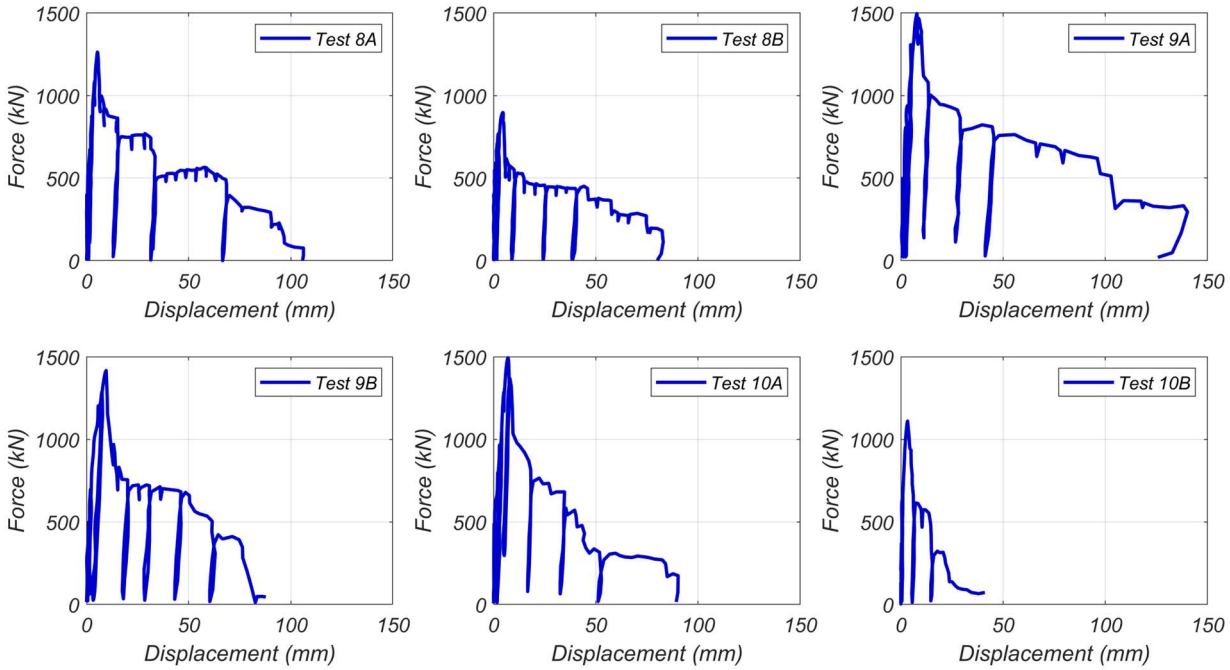


Figure 4.1-2. Experimental force-displacement curves for different test specimens. (Adapted from (Kottari, 2016)).

4.1.3 FORCE-DISPLACEMENT MODEL FOR SHEAR KEYS

The main goal of this study is to propose a simplified model to describe the lateral force-displacement response of exterior shear keys failing due to SS. Previous studies have used detailed three-dimensional FE models to simulate the response of shear keys failing in SS (Kottari, 2016) or proposed analytical formulations to derive force-displacement relations which are not practical for their implementation in available FE software (Han, Zhou, Zhong et al., 2017). Herein, a simplified tri-linear force-displacement model for SS is proposed to

be implemented in conjunction with the currently available model suggested for DT failing shear keys, initially developed by Megally et al. (Megally et al., 2002) and later simplified by Goel & Chopra (Goel & Chopra, 2008). This tri-linear model formulation has been implemented in the OpenSees simulation platform (Mckenna et al.).

The lateral force-displacement response of a shear key depends on the governing failure mode. The shear key will fail in DT if the diagonal shear strength of the supporting element (i.e., stem wall or cap beam) is lower than the sliding shear strength provided by the interface ($V_w < V_{slid}$). In such case, the tri-linear model proposed by Goel and Chopra is used to model the response. In contrast, the shear key will fail according to a SS failure pattern if $V_w > V_{slid}$, in which case the tri-linear model proposed in this study is used.

The DT shear strength (V_w) can be computed according to the equilibrium-derived expression (Eq. (4.1.1)) proposed by Kottari et al. (Kottari et al., 2020) according to the free body diagram shown in Figure 4.1-3a.

$$\sum M_A: V_w = \frac{\sum_i^{\#hor\ bars} F_{i,h} x_{i,h} + \sum_j^{\#vert\ bars} F_{j,v} x_{j,v} + \sum_t^{\#hor.shear\ bars} F_{t,s} x_{t,s} - 0.5 C_c \alpha_c l}{h - L\ tan(\beta)} \qquad (4.1.1)$$

where M_A is the total moment acting at the inferior corner of the cap beam, V_w is the shear strength of the shear key, $F_{i,h}, F_{j,v}$ are the tensile forces in the horizontal and vertical reinforcement bars, respectively, $F_{t,s}$ is the tensile force in the horizontal shear reinforcement bars placed on both lateral sides of the supporting element, $x_{i,h}, x_{j,v}$ and $x_{t,s}$ are the distances from $F_{i,h}, F_{j,v}$, and $F_{t,s}$ to point A, respectively, C_c and $\alpha_c l$ are the compressive force and length of the compressed concrete block, respectively, h and L are the vertical and horizontal distances from the point of application of the load to point A, respectively, l is the length of the shear key, and β is the angle of inclination of the loaded face of the shear key.

Figure 4.1-3b displays the SS failure mechanism, where T is the cohesive strength of concrete, F_s is the clamping force produced by the tension yielding of vertical side steel reinforcement, and V_{slid}, P_{slid} are the horizontal and vertical components of the peak sliding capacity of the shear key, respectively.

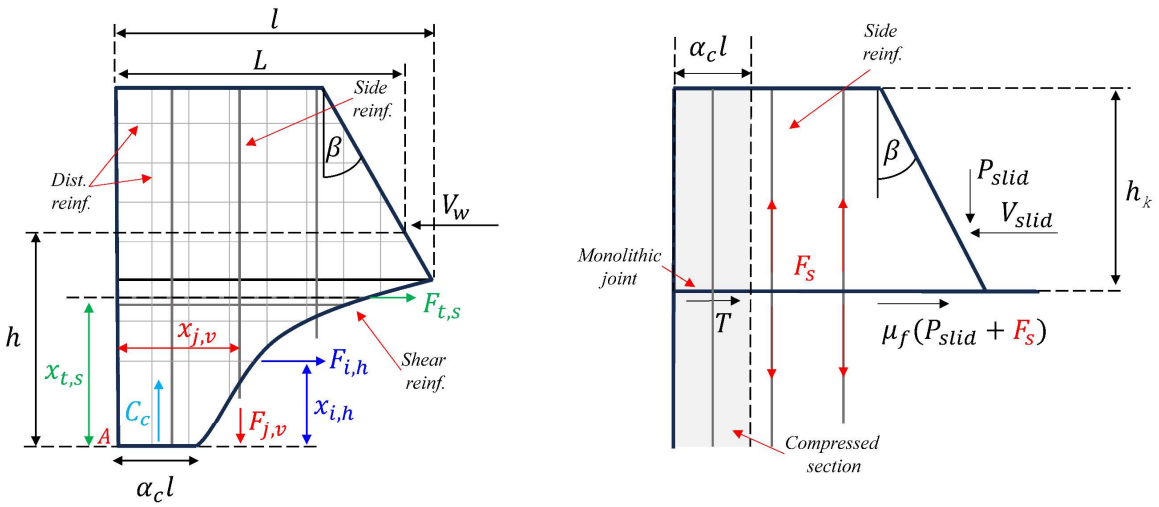


Figure 4.1-3. Free body diagrams of failure mechanism (a) DT, (b) SS. (Adapted from (Kottari, 2016))

Regardless of the failure mechanism (DT or SS), the shear key can be modeled in OpenSees considering a tri-linear material model consisting of a *uniaxialMaterial Hysteretic*, in series with a *uniaxialMaterial ElasticPPGap* and bounded by a *uniaxialMaterial MinMax* for both directions. The shear key can be included in FE models using a zero-length element characterized by the corresponding force-displacement tri-linear model. The model proposed for SS is defined by the three points described in Figure 4.1-4, and implemented in OpenSees as previously described. The first point defines the peak sliding capacity (D2, V2). The second one represents the post-sliding ultimate capacity (D3, V3). The last point is defined according to the displacement that generates the complete failure of the key (D4), i.e., where the residual capacity (V4) is assumed to be zero.

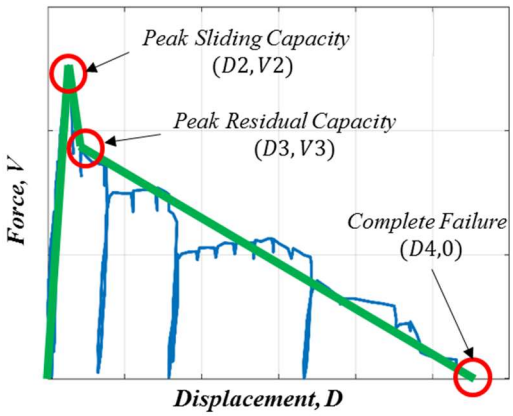


Figure 4.1-4. Diagram of the proposed model.

PEAK AND ULTIMATE SLIDING CAPACITY

To define the peak sliding and peak residual shear capacities, V2 is assumed to be equal to V_{slid} according to the expressions proposed by Kottari et al. (Kottari et al., 2020) (Eq. (4.1.2)). V3 is determined using Bozorgzadeh et al. (Bozorgzadeh, Megally, Ashford et al., 2006) expression for V_u in Eq. (4.1.5). These expressions are derived from the equilibrium of forces and moments and depend directly on the design of the shear key and its geometry. For the former, Kottari et al. (Kottari et al., 2020) established the equilibrium of a shear sliding shear key (Figure 4.1-3b), assuming that all vertical reinforcement has yielded and differentiating between rough and smooth joint construction (using different friction coefficient). In the case of monolithic shear keys, rough joint construction expressions apply. The peak sliding capacity of the key is then defined in terms of the contribution of concrete cohesive forces, friction forces at the joint, and the contribution of vertical dowel bars. Peak capacity is then obtained by establishing the equilibrium between these, defined by Eq. (4.1.2).

$$V2 = V_{slid} = \frac{T + \mu_f F_s}{1 - \mu_f \tan(\beta)} \tag{4.1.2}$$

$$T = c\ d\ \alpha_c l \tag{4.1.3}$$

$$c = \frac{0.15 f_c'}{\sqrt{0.0099\ X + 0.3659}}, X = 1.5\frac{l}{d_a} \tag{4.1.4}$$

In Eq. (4.1.2), *T* is the cohesive force of concrete, proposed by Bažant and Pfeiffer (Bažant & Pfeiffer, 1986) based on shear fracture tests carried out with plain concrete specimens. The friction coefficient μ_f is taken as 1.4 for monolithic concrete joints, as recommended by ACI 318 (American Concrete Institute, 2019). The force of steel dowel bars is denoted by $F_s = A_{vf} f_y$. The inclination of the loaded face is included through $\tan(\beta)$. Cohesive concrete forces in Eq. (4.1.3) depend on the cohesion parameter *c* and the width *d* and base length *l* of the joint. Kottari et al. (Kottari et al., 2020) recommend that the coefficient α_c should be set at a value of 0.25 based on experimental data, as the actual contact area is less than the base length once the joint has cracked. To compute *c*, Eq. (4.1.4) is used, which is an empirical formula proposed in (Bažant & Pfeiffer, 1986), wherein d_a is the maximum diameter of aggregates, included to account for aggregate interlock.

Kottari et al. (Kottari et al., 2020) validated the expression by predicting peak sliding capacity for all test specimens 8A-10B, observing a good agreement between analytical predictions and test values, with relative errors of less than 10% for all cases except 9B, for which a 19% underprediction was obtained. Thus, this expression for the peak sliding capacity of monolithic keys should provide an adequate starting point for the simplified model.

On the other hand, Bozorgzadeh et al. (Bozorgzadeh, Megally, Ashford et al., 2006) proposed an expression for the ultimate capacity of isolated shear keys (with a construction joint) that fail in SF mode. This expression is also based on the equilibrium of forces and moments, as shown in Figure 4.1-5. In this case, capacity is provided by the shear resistance of vertical dowel bars, which are assumed to have reached their ultimate tensile capacity, as well as friction forces at the joint. The formula for the ultimate capacity is given in Eq. (4.1.5).

$$V3 = V_u = \frac{\mu_f \cos(\alpha_k) + \sin(\alpha_k)}{1 - \mu_f \tan(\beta)} A_{vf} f_u \qquad (4.1.5)$$

where α_k is the kink angle of dowel bars at their ultimate state, and A_{vf} and f_u are the total area and ultimate strength of dowel bars, respectively. Similar to Eq. (4.1.2), μ_f is the coefficient of friction at the joint and $\tan(\beta)$ accounts for the inclination of the loaded face.

For isolated keys, Bozorgzadeh et al. (Bozorgzadeh, Megally, Ashford et al., 2006) validated this expression against test results from series 3A, 3B, and 5B and Kottari (Kottari, 2016) for test series 7A and 7B, with prediction errors ranging from 5% to 33%. Both authors further studied the kinking of dowel bars under sliding, providing statistical analysis of α_k concerning different concrete and steel grades, which will be discussed later in this paper.

Although Eq. (4.1.5) formula was initially devised for isolated shear keys, examination of experimental tests reported by Kottari (Kottari, 2016), such as those depicted in Figure 4.1-1b, has revealed that after initial sliding of a monolithic joint, the concrete around the joint gradually ceases to contribute. Cohesive forces are no longer present at that point, and lateral sliding of the key induces kinking in dowel bars. At the point where all concrete around the joint has spalled, a residual peak capacity is reached, which can be identified from experimental test results in Figure 4.1-2. The fracture of dowel bars follows due to shear action until the complete failure is observed. In this context, the proposed model for monolithic joints also uses Eq. (4.1.5) as the residual capacity point for the tri-linear curve.

In summary, the two force parameters of the tri-linear model are defined from equations obtained from equilibrium which were previously proposed and validated against experimental results by their respective authors.

DISPLACEMENT AT PEAK CAPACITY

To estimate the displacement at the peak sliding capacity, $D2 = \Delta_{slid}$, an effective linear behavior is assumed, defined by the shear stiffness of the cracked section. The stiffness of this initial linear branch is given by Eq. (4.1.6).

$$K_{shear} = \frac{k_{eff} f_{red} E_c dl}{2(1+\nu) h_k} \tag{4.1.6}$$

Where d and l are the width and base length of the joint, respectively, h_k is the height of the shear key, E_c is the modulus of elasticity of concrete, ν is the of Poisson ratio of concrete, taken as 0.2 for this study. The factor f_{red} is the shear stiffness reduction factor for rectangular sections which is equal to 0.8. Finally, an effective stiffness reduction due to concrete cracking is considered through the factor k_{eff}.

The effective stiffness of cracked concrete sections has been widely studied for different concrete members. While ACI 318 (American Concrete Institute, 2019) recommends a factor of 0.35 for shear walls, previous research has shown that this value severely underestimates the loss of stiffness in other reinforced concrete members. Based on experimental data, Sezen and Moehle (Sezen & Moehle, 2004) show a factor of about 0.2 for lightly reinforced concrete columns. Nabilah and Koh (Nabilah & Koh, 2017) recommend a factor of 0.1 for coupling beams, while Sharifi and Shafieian (Sharifi & Shafieian, 2018) developed an empirical formula based on extensive data on squat and slender walls, which results in factors of 0.14-0.17 when applied to the specimens 8A-10B.

It should also be noted that previous analytical formulations for lateral response of monolithic shear keys, namely the continuum FE model developed by Kottari (Kottari, 2016), and the analytical model developed by Han et al. (Han, Zhou, Zhong et al., 2017), severely overestimated the actual stiffness of experimental tests. This observation motivated the use of a reduction factor for the computation of D2. As mentioned above, the empirical formula by Sharifi and Shafieian (Sharifi & Shafieian, 2018) resulted in a factor of 0.14-0.17 for specimens 8A-10B, but this also overestimates the observed stiffness from test results. In the absence of experimentally derived factors for the value of k_{eff} for monolithic shear keys, an approximate value of 0.1 is initially recommended, but the actual value of k_{eff} for the proposed model is to be calibrated in section 4.1.5.

DISPLACEMENT AT POST-PEAK RESIDUAL CAPACITY

After the sliding plane is formed, the cohesive component is rapidly lost, and resistance is provided by friction forces generated at the interface and the horizontal projection of axial forces in the vertical bars that bend due to the lateral displacement. At (D3, V3), it is assumed that rotational plastic hinges are developed in the steel dowels on both sides of the sliding plane, (i.e., in the shear key and supporting element, see Figure 4.1-5). Then, the lateral displacement required to reach the dowel´s ultimate capacity is defined as the horizontal component of an axial elongation of ϵ_u between those hinges. After reaching their ultimate capacity, the resistance of steel rebars degrades and eventually the fracture point is reached. To this end, the length of the dowel bars between plastic hinges will be estimated based on the analysis of the dowel action, as described next.

To study dowel action, a phenomenological model proposed by Kottari (Kottari, 2016) is employed. This model is based on previous analytical formulations (Vintzeleou & Tassios (Vintzēleou & Tassios, 1986), He & Kwan (He & Kwan, 2001)), which consider a beam-over-elastic-foundation analogy to represent the steel bar embedded in concrete. The model by Kottari (Kottari, 2016) establishes equilibrium of forces and moments for dowel bars considering the plastic moment capacity of dowel bars (M_{pl}) and the dowel action of steel bars (F_d). These are computed in the case of zero eccentricity (i.e., the force is applied at the sliding surface), according to Eqs. (4.1.7) and (4.1.8) .

$$M_{pl} = \frac{f_{cb} d_b l_y^2}{2} \tag{4.1.7}$$

$$F_d = f_{cb} d_b l_y \tag{4.1.8}$$

$$f_{cb} = 0.68 \left(2 + \frac{12.7}{d_b}\right) f_c'^{1.2} \tag{4.1.9}$$

where f_{cb} is the bearing strength of concrete, d_b is the diameter of dowel bars, and l_y is half the effective length of dowel bars between plastic hinges. Eq. (4.1.9) was proposed by Kottari (Kottari, 2016) based on empirical results from experimental tests and statistical analysis, which establishes the bearing strength of concrete around dowel bars in terms of bar diameter and f_c'. In Eq. (4.1.9), the formula proposed in (Kottari, 2016) has been adjusted for SI units, where d_b is in mm, f_c' is in MPa, and f_{cb} is MPa.

Recalling that plastic moment of a circular section can also be computed as $M_{pl} = f_y d_b^3/6$, Eqs. (4.1.7) and (4.1.8) can be substituted to obtain the expression for l_y using Eqs. (4.1.10) and (4.1.11).

$$F_d = \sqrt{2 M_{pl} d_b f_{cb}} \tag{4.1.10}$$

$$l_y = \frac{F_d}{f_{cb} d_b} \tag{4.1.11}$$

This procedure for estimating l_y in monolithic shear keys assumes that the uplifting of the shear key is negligible so that the dowel action model developed for horizontal sliding of the key is still valid, hence the case of zero eccentricity. Kottari (Kottari, 2016) reported maximum uplifts which ranged from 1.25 to 3.6 cm, compared to the shear key length of 60.96 cm. Hence, this simplification is deemed acceptable. Out-of-plane rotations were reported as negligible in (Kottari, 2016). It is also assumed that all dowel bars reach α_k and ϵ_u at the point (D3, V3).

As shown in Figure 4.1-5, D3 $= \Delta_u$ can be easily defined by the geometry of kinked bars once l_y has been obtained from Eq. (4.1.11). Nevertheless, considering that several assumptions have been made to establish the effective length of the bars, a coefficient is introduced to account for the imperfect estimation of l_y. The proposed expression for D3 is thus defined by Eq. (4.1.12).

$$\mathrm{D3} = \Delta_u = k_{d3} l_y \tan(\alpha_k) \tag{4.1.12}$$

where k_{d3} is the coefficient that accounts for the effective length of dowel bars between plastic hinges, and is initially proposed with a value of 2, due to the trigonometrical relationship in Figure 4.1-5. This parameter, analogous to k_{eff}, is to be later estimated during model calibration.

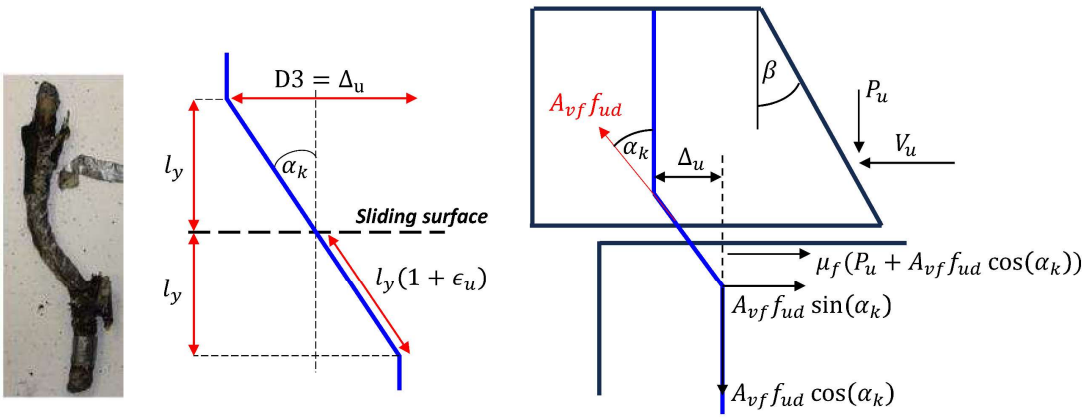


Figure 4.1-5. Deformed bar from experimental tests, diagram of dowel bars at ultimate strain and free body diagram of equilibrium at peak residual capacity.

DISPLACEMENT AT FAILURE

The final point of the proposed model is associated to the displacement at which complete failure of the shear key is observed. At this point, a residual capacity of V4=0 is assumed. The displacement $D4 = \Delta_f$ is characterized by the fracture of all dowel bars previously holding the shear key. To the best of the authors' knowledge, there are no previous studies on the kinematics of rebar fracture for monolithic shear keys. In the proposed model, a simplified approach is taken, based on the assumption of a uniformly distributed dowel bar area over the remaining contact surface after D3.

The approach taken is illustrated by the diagram in Figure 4.1-6. The dowel bar area is assumed to be uniformly distributed over the shear key's surface and to linearly decrease as displacement *x* increases. In this scheme, the "remaining" residual strength *V(x)* decreases proportionally to *x*, as the area of dowel bars, $A_s(x)$, linearly decreases. By this idealization, it is assumed that sliding of the key remains perfectly parallel to the joint surface. Thus, *V(x)* decreases linearly between V3 and V4 = 0, causing an effect of "peeling" as dowel bars fracture progressively, until no vertical reinforcement is left to support the shear key and D4 is reached. The residual displacement between D3 and D4 is represented as a fraction of the shear key's base length and defined as $k_{d4}l$, as shown in Figure 4.1-6. This approach seeks to represent a complex and highly uncertain phenomenon in a simplified, although practical manner.

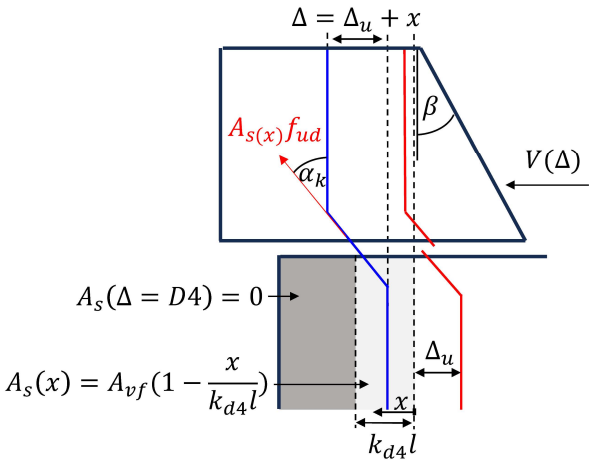


Figure 4.1-6. Diagram of the proposed simplification for dowel bar fracture mechanism.

The actual mechanism is defined by a number of factors that are not trivially identifiable, including the uncertainty in actual dowel bar properties, in- and out-of-plane rotation of the keys, distribution of dowel bars, actual effective length of dowel bars, and concrete bearing strength, among other potential factors. However, the approach taken is motivated by the observation that experimental data from tests 8A-10B in Figure 4.1-2 show a practically linear decrease in residual capacity between the residual peak and complete failure of the keys.

With these considerations, complete failure of the key at D4 is defined by Eqs. (4.1.13) and (4.1.14), in which, the expression for D4 can be interpreted as that the shear key will fail completely when the lateral displacement is enough to fracture all dowel bars, thus bringing the residual capacity to zero. In the same fashion as both previous displacement points defined for the proposed model, D4 includes the numerical coefficient k_{d4}, which is initially proposed to be 0.25, with the same value suggested in (Kottari, 2016) for the contact surface length at the computation of V2. The corresponding value will be estimated during model calibration.

$$V(x) = \frac{\mu_f \cos(\alpha_k) + \sin(\alpha_k)}{1 - \mu_f \tan(\beta)} A_s(x) f_u = V_u\left(1 - \frac{x}{k_{d4} l}\right) \tag{4.1.13}$$

$$\mathrm{D4} = \Delta_f = \Delta_u + k_{d4} l \tag{4.1.14}$$

SUMMARY AND MODEL EQUATIONS

Having defined the three main points, a simplified tri-linear backbone curve for the behavior of SS-failing monolithic shear keys has been proposed in this section. The formulation of the model is composed of two capacity points, which have been previously developed based on the equilibrium of forces and moments and validated through comparison against experimental data; and three displacement points, which are based on the shear stiffness of a concrete member, the geometry of a steel bar in its ultimate state, and the observations made on experimental data, respectively. The resulting backbone curve for sliding shear, of which a diagram is shown in Figure 4.1-4, is compatible to be implemented as a uniaxial material constitutive model in the context of a FE software, and to be used in conjunction with other

tri-linear formulations, such as the simplified Goel and Chopra tri-linear model (Goel & Chopra, 2008) for DT-failing shear keys.

The expressions for each point of the proposed model are summarized in Table 4.1-2. Model parameters are defined by mechanical properties of concrete and steel (i.e., $E_c, \nu, l_y, T, \alpha_k, \mu_f, F_s, f_{ud}$), geometric properties of the shear key (i.e., $d, l, h_k, A_{vf}, \beta, f_{red}$), and numerical coefficients (i.e., k_{eff}, k_{d3}, k_{d4}). The following sections of this paper focus on studying the sensitivity of the proposed model to these parameters and the calibration of the model, to later evaluate the quality of its predictions compared to experimental data from tests 8A-10B.

Table 4.1-2. Model equations.

Point	Displacement	Capacity
D2, V2	$D2 = \frac{V2}{K_{shear}}$	$V2 = \frac{T + \mu_f F_s}{1 - \mu_f \tan(\beta)}$
D3, V3	$D3 = k_{d3} l_y \tan(\alpha_k)$	$V3 = \frac{\mu_f \cos(\alpha_k) + \sin(\alpha_k)}{1 - \mu_f \tan(\beta)} A_{vf} f_{ud}$
D4, V4	$D4 = D3 + k_{d4} l$	$V4 = 0$

4.1.4 SENSITIVITY ANALYSIS

In the first stage, LSA and GSA are performed on the model. LSA is measured through correlation indices (Marelli et al., 2022) between parameter samples and model responses, followed by GSA using Sobol indices of first and total orders (S1 and ST, respectively) (Saltelli et al., 2008; Sobol, 2001). LSA provides information about the relationship between the absolute values of parameters and model predictions, including the sign of correlation indices, which gives an idea of whether a parameter has a positive or negative impact on the model response. On the other hand, GSA measures the relationship between variances of parameters and model predictions, which helps to understand the composition of uncertainty in the model in terms of the variance in model parameters. Both aim to gain a better understanding of the model behavior to select the most influential parameters to calibrate. Additionally, they provide a means to explore the prior distributions to be used in the second stage of calibration. Correlation indices and Sobol indices computed from simulation using

the *T*-estimator proposed by Janon et al. (Janon et al., 2014) are obtained through Eqs. (3.1.1) and (3.1.10), respectively.

For the analysis, a set of distributions was defined for each parameter based on typical values reported in the literature and test results reported during UC San Diego campaigns. Model parameters, their nomenclature and distributions are summarized in Table 4.1-3.

Table 4.1-3. Prior distributions for model parameters

Parameter	Symbol	Distribution	Mean	C.O.V. (%)	Reference
Concrete strength	f_c' (MPa)	Lognormal	32 / 35 / 46*	5	(Mander & Matamoros, 2019; Mirza et al., 1979a; S. A. Mirza & J. G. MacGregor, 1976)
Yield strength (dowel)	f_{yd} (MPa)		462	5	
Tensile strength (dowel)	f_{ud} (MPa)		717	8	
Effective stiffness coefficient	k_{eff} (-)		0.1	15	(Nabilah & Koh, 2017; Sezen & Moehle, 2004; Sharifi & Shafieian, 2018)
Kink angle of rebars	α_k (°)		37	5	(Bozorgzadeh, Megally, Ashford et al., 2006)
Friction coefficient	μ_f (-)		1.4	8	(American Concrete Institute, 2019)
Correction coefficient for D3	k_{d3} (-)		1.5	20	
Correction coefficient for D4	k_{d4} (-)		0.25	20	(Kottari et al., 2020)

*For each series, the concrete strength measured at test day was used, as reported in Table 4.1-1.

From these distributions, a total of 10,000 samples were drawn and the proposed model was evaluated against the test protocol from test 9A, which was randomly selected among the available data.

A summary of correlation and Sobol indices is shown in Figure 4.1-7, where grouped bars indicate the indices for each parameter at each of the calculated points of the model. Given the model formulation, where each expression is computed with a subset of parameters, some parameters have a strong influence at some point while no influence is observed at other points. For example, at V3, which is defined by the area of dowel bars, their ultimate strength, angle of kink, and the coefficient of friction, it can be seen that parameters f_{ud} and μ_f were the most influential in the prediction, with less influence from α_k. During initial model explorations, α_c had been included in SA as well. However, due to its low influence on D2 and V2 and the low amount of information for the definition of a prior distribution, it was left out of the calibration presented in this paper. This is consistent with Eq. (4.1.5), as V3 is directly proportional to the ultimate strength of dowel bars. Notably, for the computation of displacement points in the model, the proposed numerical parameters (i.e., k_{eff}, k_{d3}, k_{d4}) had the most influence in their respective points, which could be attributed to the wide prior distributions defined for the analysis. Analogously, similar observations were made about other parameters, and the sensitivity of model response to each of them was considered significant enough to be calibrated in stage 2. As a result, the set of distributions in Table 4.1-3 was kept as prior distributions for the second stage.

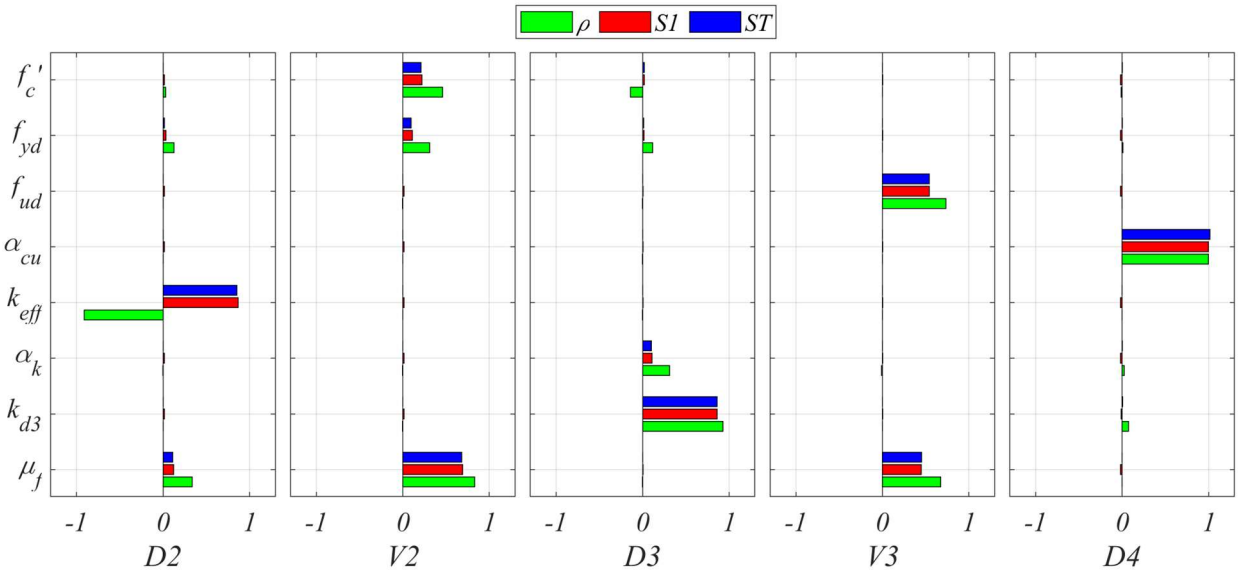


Figure 4.1-7. Sensitivity analysis results.

4.1.5 BAYESIAN MODEL CALIBRATION

In this case, the MCMC algorithm was used. The sampling was done through the adaptive metropolis sampler (Haario et al., 2001). As termination and convergence criteria, the MPSRF and the mESS criteria were used.

The parameter estimation scheme was performed individually for each of the six available test data. In each case, the points associated with each of the model's points (D2, V2, D3, V3, D4) were extracted from the data. The points are shown in Table 4.1-4.

For each estimation, four independent chains were run to a total of 100,000 samples each. Model discrepancy was modelled as a constant, in the form of the likelihood function's standard deviation σ. Given the good agreement that had been previously reported by Kottari (Kottari, 2016) between the expressions for capacity and the data, σ was set at a 1% of the observed capacity for V2 and V3. For D2, D3, and D4, where most of the newly proposed parameters were to be estimated, σ was defined as a 3% of the observed values.

As a starting point, the prior distributions previously explored in SA were used with one exception. The mean compressive strength of concrete was replaced to the measured values, which resulted in prior means of 32.5 MPa for series 8, 35.2 MPa for series 9, and 46.5 MPa for series 10. Finally, the first 30% of samples were discarded as a burn-in period.

Table 4.1-4. Observed data considered for calibration.

Test ID	D2 (mm)	V2 (kN)	D3 (mm)	V3 (kN)	D4 (mm)
8A	5.31	1262.8	9.12	920.7	106.20
8B	4.47	897.2	6.30	616.9	82.98
9A	7.84	1496.8	11.08	1117.8	137.10
9B	9.61	1417.1	11.89	968.8	87.78
10A	7.08	1494.1	11.35	979.4	90.28
10B	3.28	1109.8	5.53	732.6	41.07

The procedure was carried out for each of the six test ID's. After discarding burn-in samples, a total of 280,000 (i.e., 4x70,000) samples were obtained for the posterior of each test. Their statistics are summarized in Table 4.1-5. All estimations met the Gelman-Rubin convergence criteria of MPSRF < 1.02, with values lower than 1.005 in all cases. For the mESS criteria, the minESS needed to satisfy convergence within a 5% relative error and 95% confidence

interval with nine parameters is 8,823 mESS. All estimations exceeded this threshold, at over 9,000 mESS.

In general, mechanical parameters did not show a significant shift in distribution from their priors, while the newly defined numerical parameters showed a considerable update. The coefficient associated with the reduction of shear stiffness under cracking, k_{eff}, resulted in values of 0.04 to 0.07 and C.O.V. close to 3.8%. These means are the lowest among stiffness reduction coefficients proposed for other concrete members, which was consistent with those overestimating the actual stiffness of monolithic shear keys.

Parameter k_{d3} showed more variability among means, consistent with the high sensitivity of D3 to its value. The values obtained for this parameter in series 9 were higher than those of series 8, indicating a longer effective dowel bar length, which could be attributed to the higher number of dowel bars, as well as stronger concrete bearing. In series 10, which had the same design as series 8, although with a stronger concrete, the value of 10A was the highest among all tests, while for 10B it was similar to that of 8B. Despite the stronger concrete, test 10B was the only one that experienced noticeable rotation, as reported by Kottari (Kottari, 2016). As seen in Figure 4.1-2, this rotation led to premature bar fracture, and the lowest displacement capacity among all tests.

Parameter k_{d4} was updated significantly, with a reduction from the prior C.O.V. of 20% to the order of 3%, while the means for each test resulted in residual displacements shorter than the initial contact surface assumed for the calculation of V2. For tests 8A, 8B, 9B and 10A, posterior means were close to 0.13. Test 9A showed the highest value for this parameter at 0.21, while test 10B expectedly showed the lowest, due to premature failure of the specimen.

Despite the number of simplifications taken in the formulation of the proposed model, parameter posteriors from different tests had similar distributions for the numerical coefficients, while meeting all convergence criteria and achieving accurate predictions on experimental data. Posterior samples and their predictions will be discussed in the following section alongside predictions from the combined distribution.

Table 4.1-5. Marginal posterior distributions from MCMC.

Test		f_c' (MPa)	f_{yd} (MPa)	f_{ud} (MPa)	k_{d4} (-)	k_{eff} (-)	α_k (°)	k_{d3} (-)	μ_f (-)	MPSRF	mESS
8A	μ	32.81	469.71	675.87	0.16	0.06	36.55	1.27	1.50	1.002	10949
	C.O.V. (%)	4.48	4.75	3.78	3.20	3.78	4.76	7.66	2.56		
8B	μ	33.29	468.55	649.28	0.13	0.05	35.58	0.92	1.44	1.002	10873
	C.O.V. (%)	4.20	4.60	3.85	3.13	3.75	4.75	7.61	3.99		
9A	μ	35.27	484.22	568.69	0.21	0.05	37.15	1.54	1.35	1.003	10667
	C.O.V. (%)	4.63	4.64	3.51	3.24	3.87	4.87	7.78	2.65		
9B	μ	36.17	519.24	536.76	0.13	0.04	37.54	1.60	1.65	1.002	9071
	C.O.V. (%)	4.79	4.45	3.51	3.33	3.98	4.89	7.83	3.35		
10A	μ	44.58	466.54	733.20	0.13	0.05	37.66	1.81	1.48	1.002	9819
	C.O.V. (%)	4.17	4.76	3.95	3.32	3.74	4.81	7.78	2.73		
10B	μ	46.43	455.62	753.45	0.06	0.07	35.78	0.99	1.47	1.002	9087
	C.O.V. (%)	3.82	4.80	4.16	3.30	3.59	4.74	7.55	4.30		
Prior	μ	33, 35, 47	461	717	0.25	0.10	37.00	1.50	1.40		
	C.O.V. (%)	5	5	8	20	15	5	20	8		

COMBINATION OF POSTERIOR DISTRIBUTIONS

For this case, given the experimental evidence that test 10B failed prematurely, besides being the only test to show significant rotation during testing, it was assigned a weight 0. In other words, test 10B was not included in the combination of PDF, although the combined PDF was evaluated against test 10B for completeness.

To obtain a single distribution, concrete and steel properties of each test were kept to their respective posteriors, as it would have made no physical sense to combine them. The remaining parameters were combined through conflation (see Eq. (3.2.9)). Additionally, correlation matrices from each posterior were consistent with each other. For this reason, a single correlation matrix was also used for simulation. The obtained multivariate normal distribution statistics are summarized in Table 4.1-6 and Table 4.1-7.

Table 4.1-6. Combined marginal parameter distributions.

	k_{d4}	k_{eff}	α_k	k_{d3}	μ_f
$\boldsymbol{\mu}$	0.171	0.051	36.80	1.47	1.53
C.O.V. (%)	4.95	5.29	2.55	7.69	2.79

Table 4.1-7. Estimated correlation matrix for the combined distribution.

	f'_c	f_{yd}	f_{ud}	k_{d4}	k_{eff}	α_k	k_{d3}	μ_f
f'_c	1	-0.20	0.66	0.01	-0.53	0.04	0.36	-0.68
f_{yd}	-0.20	1	0.47	0.00	0.12	-0.05	-0.31	-0.49
f_{ud}	0.66	0.47	1	0.00	-0.41	0.03	0.05	-0.96
k_{d4}	0.01	0.00	0.00	1	0.00	0.00	-0.03	0.00
k_{eff}	-0.53	0.12	-0.41	0.00	1	-0.02	-0.20	0.43
α_k	0.04	-0.05	0.03	0.00	-0.02	1	-0.78	0.01
k_{d3}	0.36	-0.31	0.05	-0.03	-0.20	-0.78	1.00	-0.09
μ_f	-0.68	-0.49	-0.96	0.00	0.43	0.01	0	1

4.1.6 MODEL VALIDATION

COMPARISON OF PREDICTED MODEL POINTS

From both sets of distributions (i.e., individual posteriors and combined by conflation), 300 samples were generated to evaluate the model against experimental data. Predictions were measured against the observed data used for calibration (i.e., D2, V2, D3, V3, D4) and their accuracy was measured through relative mean errors. The results of this step are summarized in Table 4.1-8.

The proposed model achieved excellent accuracy when sampling each posterior to predict their respective dataset. For all model points, errors in the order of 0% to 4.2% were obtained. For capacity points, previously validated expressions were improved by updating parameter PDF. For displacement points, the proposed formulation was validated by the accuracy obtained. Specifically, for D2, prediction errors ranged from 1.3% to 4.2%, thus confirming that the use of a lower effective stiffness to model shear keys in SS was needed in comparison to other concrete members. For D3, prediction errors ranged from 0.1% to 1.2%, suggesting that the proposed geometric formulation was appropriate to represent the ultimate tensile state of dowel bars in these shear keys. Finally, D4 was predicted with errors between 0.4% and 3.5%, with good approximations having been achieved by the assumption of a uniformly distributed dowel bar area.

In terms of the combined distribution, a trade-off between accuracy and convenience was expected. Predictions confirmed such a trade-off, as accuracy was noticeably affected in this approach. For capacity points, errors ranged from 2.6% to 16.8% in the case of V2, and from

3.2% to 19.2% in the case of V3. However, these values still retained good accuracy given that strength parameters were not combined. For displacement points, the loss of accuracy was more significant, with D2 errors ranging from 4.4% to 46.6%, D3 errors ranging from 5.7% to 68.8% and D4 errors ranging from 8.2% to 38.4%, excluding 10B which was severely over-predicted with an error of 174.6% with the combined distribution.

Overall, individual posteriors were able to accurately predict experimental data in this comparison, while the combined distribution lost accuracy. However, given the simplicity of the formulation and the absence of alternatives, the proposed model could prove useful for simulation of SS failing keys. Furthermore, the second comparison was made by combining posteriors from only five datasets. With more experimental data available, a single distribution obtained through conflation could be significantly more representative of a wider range of shear keys.

Table 4.1-8. Comparison of mean predictions from individual posteriors and combined distribution samples.

		8A	**8B**	**9A**	**9B**	**10A**	**10B**
D2 (mm)	**Observed**	**5.31**	**4.47**	**7.85**	**9.60**	**7.09**	**3.28**
	Posterior	**5.20**	**4.34**	**7.59**	**9.20**	**6.85**	**3.23**
	Rel. error	2.0%	2.8%	3.3%	4.2%	3.3%	1.3%
	Combined	**6.49**	**4.67**	**8.45**	**6.31**	**6.69**	**4.80**
	Rel. error	22.3%	4.4%	7.7%	34.3%	5.5%	46.6%
V2 (kN)	**Observed**	**1263**	**897**	**1497**	**1417**	**1494**	**1110**
	Posterior	**1263**	**897**	**1496**	**1412**	**1498**	**1111**
	Rel. error	0.0%	0.0%	0.0%	0.4%	0.2%	0.1%
	Combined	**1295**	**938**	**1748**	**1321**	**1557**	**1140**
	Rel. error	2.6%	4.5%	16.8%	6.8%	4.2%	2.8%
D3 (mm)	**Observed**	**9.12**	**6.30**	**11.07**	**11.89**	**11.35**	**5.54**
	Posterior	**9.16**	**6.37**	**11.08**	**11.86**	**11.29**	**5.56**
	Rel. error	0.4%	1.2%	0.1%	0.2%	0.5%	0.4%
	Combined	**10.74**	**10.63**	**10.44**	**10.65**	**8.90**	**8.59**
	Rel. error	17.7%	68.8%	5.7%	10.4%	21.6%	55.1%
V3 (kN)	**Observed**	**921**	**617**	**1118**	**969**	**979**	**733**
	Posterior	**922**	**618**	**1122**	**974**	**980**	**733**
	Rel. error	0.1%	0.2%	0.4%	0.5%	0.0%	0.0%
	Combined	**950**	**656**	**1333**	**904**	**1031**	**762**
	Rel. error	3.2%	6.4%	19.2%	6.7%	5.3%	4.0%

D4 (mm)	**Observed**	**106.20**	**82.98**	**137.11**	**87.78**	**90.27**	**41.07**
	Posterior	**107.18**	**84.28**	**137.59**	**89.19**	**91.70**	**42.51**
	Rel. error	0.9%	1.6%	0.4%	1.6%	1.6%	3.5%
	Combined	**114.94**	**114.84**	**114.64**	**114.85**	**113.11**	**112.79**
	Rel. error	8.2%	38.4%	16.4%	30.8%	25.3%	174.6%

COMPARISON OF PREDICTED FORCE-DISPLACEMENT CURVES

The purpose of the proposed model was not to predict three points alone, but to provide a practical formulation for complete force-displacement history predictions. To compare the complete experimental and simulated responses, the model was implemented in OpenSees as a uniaxial material, as previously described, and the force-displacement responses from each test were used to produce 300 model simulations. In this case, the accuracy of the predictions can be assessed through visual comparison of tests and model predictions in Figure 4.1-8 and Figure 4.1-9. Note that in the latter, samples of the combined distribution to predict test 10B are shown only for completeness and comparison to the former, although data from this test was not included in the conflation of posteriors.

In Figure 4.1-8, the goodness of fit between each posterior and its respective test became evident, as the defining points of the proposed model were closely predicted. The actual behavior between those points was only approximate, given the simplified formulation. For example, in test 8A, the staircase-like progression of dowel bar fracture could not be captured. Moreover, unloading cycles followed the initial reduced stiffness formulation, managing to capture experimentally obtained unloading cycles. The displacement at final failure of each shear key was also properly captured by the model.

In Figure 4.1-9, estimation errors by the combined distribution were apparent at each test, where, for example, V2 and D3 were clearly underestimated for test 9B, V3 and D4 were overestimated for test 8B. In general, unloading cycle stiffness was captured by the combined PDF, although the model defining points were expectedly less accurate than from individual posterior predictions. However, the general behavior was successfully approximated by the simplified tri-linear formulation. In view of potential displacement-based damage state definitions based on exceedance of points D2, D3 and D4, the proposed model could provide an efficient way to develop fragility curves for SS failing keys. Additionally, energy

dissipation by the shear keys could also be approximated by this formulation. Both results would be of practical use in PBEE. For deterministic use of the proposed model, single predictions can be made using the mean values of the marginal distributions in Table 4.1-6. In both figures, the model is also evaluated in this manner, making a single prediction using parameter means from each distribution. These are shown in a green dash line for reference.

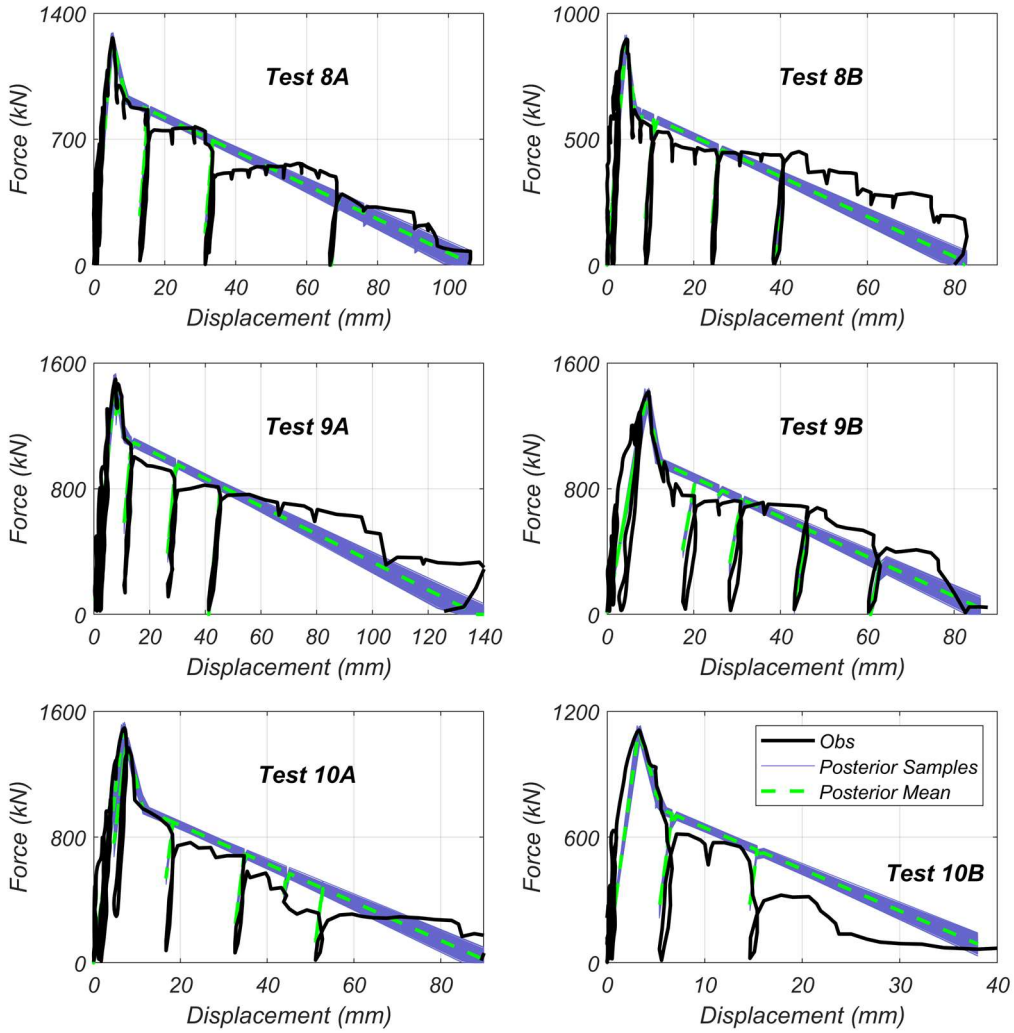


Figure 4.1-8. Model predictions with MCMC posterior distributions.

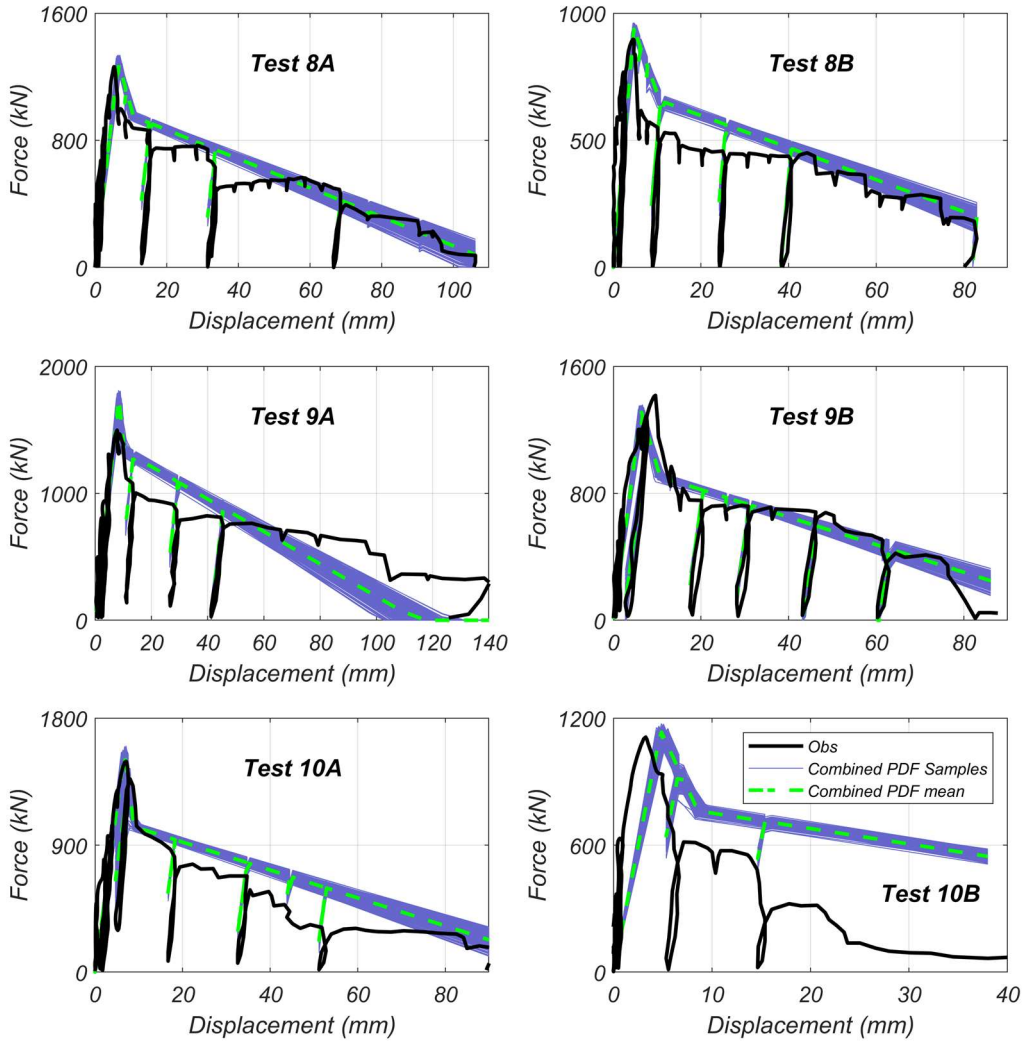


Figure 4.1-9. Model predictions with the proposed distribution.

4.1.7 CLOSING REMARKS ON THE SIMPLIFIED TRI-LINEAR MODEL

A simplified tri-linear model for monolithic shear keys failing in sliding shear has been proposed, calibrated, and validated against experimental data. The model formulation is based on previous studies on monolithic and isolated shear keys in bridges, as it combines previously developed expressions for shear key capacity with newly developed expressions for their associated displacements. With a tri-linear formulation, the proposed model serves as a generalization to a widely available simplified model for shear keys failing in diagonal tension.

The proposed model has been probabilistically calibrated and validated against experimental data in a Bayesian framework through sensitivity analysis and parameter estimation using MCMC. The calibrated distributions have been combined through their conflation to obtain a single multivariate distribution for model simulation. Predictions generated in both approaches have shown good agreement with the data.

To the best of the author's knowledge, the model proposed in this study is the first attempt to define a backbone force-displacement curve for shear keys failing in sliding shear using closed-form equations, which is very convenient for FE simulations of entire structural systems. To develop a tri-linear formulation suitable for structure-level FE simulations, the proposed model provides a simplified understanding of complex phenomena. In this regard, one main limitation is identified. While the adopted approach provides a rigorous methodology to incorporate experimental data and quantify parameter uncertainty, the limited dataset available for shear keys failing in sliding shear challenges the extrapolation of values obtained herein to a wide range of shear key designs.

Finally, the proposed model is defined in direct relation to potential damage states observed in shear keys. In this context, it fills an important gap in performance-based seismic design and fragility assessment of bridges that include this type of component. Moreover, the model is more efficient than previously proposed strategies to capture sliding shear, while achieving more accurate predictions.

4.2 UNIAXIAL CONSTITUTIVE MODELS FOR BRIDGE COMPONENTS

In order to adequately quantify and propagate uncertainty in PBEE, numerical models capable of capturing the uncertainty about the actual behavior of structural components are required. In this regard, probabilistic analysis techniques are currently available that allow rigorous incorporation of experimental data for the probabilistic calibration of these models. It is within this context that the work presented in this section is developed. Bayesian estimation methods are used to calibrate nonlinear constitutive models of bridge components, incorporating data obtained experimentally from a variety of nonlinear tests on reinforcing steel bars, elastomeric bearing plates, and vertical seismic bars.

The probability distributions proposed for the calibrated models are intended to be incorporated in PSDA through NLTHA and structural model simulation. Thus, this work falls within the second stage of the PBEE framework, quantifying the uncertainty in the mechanical properties of the structure at the local level. These properties can then be incorporated to obtain a probabilistic model of the structural response, along with the exceedance rates of the EDPs.

4.2.1 CONSTITUTIVE MODELS

It is noted that the work presented in this section was partially developed prior to the current doctoral research, by the author (Birrell, Astroza, Carreño et al., 2021) and in collaboration within the research group (Pinto et al., 2023). Nevertheless, it is presented in this document for completeness of the proposed approach and to provide the necessary context to understand the parameter PDFs used in the case studies presented later.

The constitutive models considered and their formulation are briefly described below. Further details can be found in the respective works cited in the text.

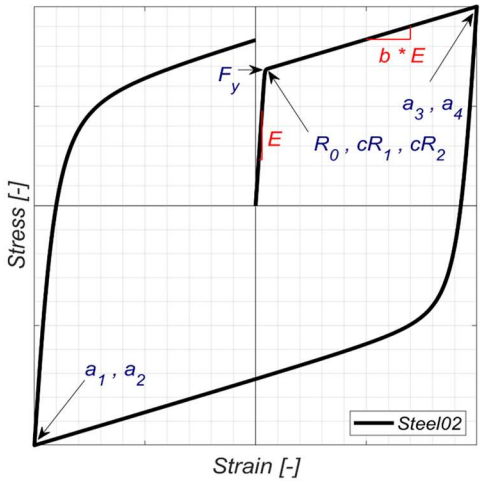
4.2.1.1 REINFORCING STEEL BARS

In the case of reinforcing steel bars, the work carried out by Birrell et al.(Birrell, Astroza, Carreño et al., 2021; Birrell, Astroza, Restrepo et al., 2021) incorporated the data obtained in a battery of experimental tests on bars manufactured by three manufacturers, according to

two ASTM standards: A615 and A706, Gr. 60. The tests, carried out by Carreño et al. (Carreño et al., 2020) considered six cyclic loading protocols with two loading rates.

These data were used for the calibration of three steel models available in the OpenSees software (Mckenna et al.), which numerically implement the popular constitutive models of Menegotto-Pinto and Dodd-Restrepo modified by Carreño (Carreño et al., 2020; Dodd & Restrepo-Posada, 1995). This section presents the calibration results of the uniaxial material model *Steel02*, widely used for nonlinear analysis.

The parameters and the constitutive curve of the model are shown in Figure 4.2-1



$$\sigma* = \left(b + \frac{1-b}{(1+|\varepsilon*|^R)^{\frac{1}{R}}}\right)\varepsilon*$$

$$\varepsilon* = \frac{\varepsilon-\varepsilon_r}{\varepsilon_0-\varepsilon_r} \quad , \quad \sigma* = \frac{\sigma-\sigma_r}{\sigma_0-\sigma_r}$$

$$R(\xi) = R_0\left(1 - \frac{cR_1\xi}{cR_2+\xi}\right) > 0$$

$$\xi = \left|\frac{\varepsilon_p - \varepsilon_0}{\varepsilon_y}\right|$$

$$F_{sh}^{t(c)} = F_y \cdot \frac{a_{3(1)}}{a_{4(2)}{}^{0.8}} \cdot \left(\frac{\Delta\varepsilon_p}{2\varepsilon_y}\right)^{0.8}$$

Figure 4.2-1. Constitutive model and equations for the Steel02 uniaxial material.

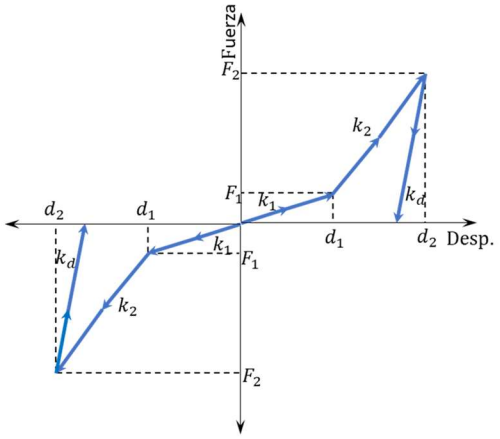
Where $\sigma*, \varepsilon*$ are the stress and strain normalized with respect to the curve's reversal point and to the yield updated in each cycle; $R(\xi)$ is the yield radius of curvature as a function of ξ and the dimensionless parameters R_0, cR_1, cR_2; ξ is the plastic strain range, normalized according to the initial yield, updated yield and the maximum recorded strain; $F_{sh}^{t(c)}$ defines the asymptotes to update the yield conditions in tension (t) and compression (c). The dimensionless parameters $a_{3(1)}, a_{4(2)}$ modify the hardening of the steel in tension and compression, respectively; $\Delta\varepsilon_p, \varepsilon_y$ are the maximum difference between tensile and compressive strain, and initial yield strain, respectively.

4.2.1.2 VERTICAL ANCHORING ANTI-SEISMIC BARS

For anti-seismic vertical bars (SB), the work carried out by Pinto et al. (Pinto et al., 2023) considered the model proposed by Martínez et al. (Martínez et al., 2017) for vertical steel bars that serve the purpose of controlling the lateral and vertical response of the deck. This model considers configurations with and without diaphragms (CD and SD, respectively). Additionally, the calibrations performed by Pinto et al. are informed by the experimental tests conducted by Martínez et al., which considered bars with a diameter of φ16 mm, and grade A440-280H.

Regarding the numerical model, the one proposed by Martínez et al. was implemented in OpenSees through a *zeroLength* element and a combination of the uniaxialMaterial *Hysteretic* and *MinMax* materials, where the former provides the bilinear envelope and the latter provides the bar activation threshold.

Figure 4.2-2 shows the diagram that defines the bilinear constitutive model.



$$d_1 = g_1 * h_1$$

$$d_2 = \begin{cases} h_1, CD \\ g_2 * h_1, SD \end{cases}$$

$$F_k = \gamma_k * f_{y,e} * A_{sb}$$

$$\gamma_1 = \begin{cases} 0.04, CD \\ 0.07, SD \end{cases}, \gamma_2 = \begin{cases} 0.71, CD \\ 0.31, SD \end{cases}$$

$$k_d = \begin{cases} 20 * k_2, CD \\ 15 * k_2, SD \end{cases}$$

Figure 4.2-2. Constitutive model and equations for SB.

Where d_1 is the point at which the behavior changes from bending to tension; d_2 is the maximum lateral displacement; F_k, γ_k are the lateral load and coefficient associated with the points d_1, d_2, respectively; $f_{y,e}$ is the expected yield strength of steel; A_{sb} is the bar area; k_d is the unloading slope; and k_1, k_2 are the loading slopes.

4.2.1.3 ELASTOMERIC BEARINGS

In the case of elastomeric bearings (EB), the data that informed the calibration were those obtained by Hube et al. (Hube et al., 2017), in an experimental campaign that considered cyclic and monotonic tests on unanchored elastomeric bearings (UEB)

For the UEBs, this campaign included six identical specimens measuring 400 mm x 500 mm x 90 mm, reinforced with 3 mm thick steel plates and 72 mm thick rubber. During the tests, they were subjected to axial loads of 400 kN and 600 kN, with shear strains ranging from 10% to 160%, at various loading speeds between 10 and 100 mm/s.

The UEB calibration, presented by Pinto et al., considers the numerical model proposed by Rubilar (Rubilar, 2016), based on an elastic-perfectly-plastic bilinear curve. This model was implemented in OpenSees using a *zeroLength* element, and the constitutive model was implemented using the uniaxial materials *Steel01* and *MinMax*. Similar to SBs, the former provides the bilinear envelope, while the latter defines the behavioral range of the supports.

Figure 4.2-3 shows the diagram that defines the constitutive model considered for the UEB.

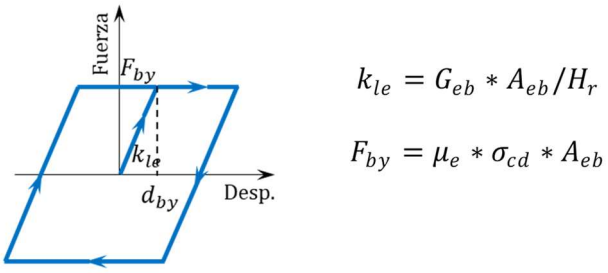


Figure 4.2-3. Constitutive model and equations for UEBs.

Where k_{le} is the lateral stiffness of the bearing; G_{eb} is the shear modulus of the bearing; A_{eb} is the area of the bearing; H_r is the rubber height; σ_{cd} is the design compressive stress; F_{by} is the yield load; and d_{by} is the yield displacement.

4.2.2 PARAMETER ESTIMATION RESULTS

The estimation of parameters was carried out according to section 3.2 by defining the prior distributions of the parameters through the available experimental data in the literature (Carreño et al., 2020; Hube et al., 2017; Martínez et al., 2017; Martínez, 2016; Rubilar, 2016)

and sensitivity analysis based on Sobol and correlation indices to identify key parameters and appropriate PDF ranges.

The calibration of *Steel02* used MCMC with the adaptive metropolis algorithm (Haario et al., 2001), in which the MPSRF and mESS criteria were used as convergence and termination criteria. The calibration of the UEB and SB models used SMC.

The parameters of each model and their prior distributions are shown in Table 4.2-1.

Table 4.2-1. Prior PDFs for model parameters.

Component	Parameter	Distribution	Mean (μ)	COV [%]	Reference
Reinforcing steel bars	F_y[MPa]	Lognormal	480	15	(Carreño et al., 2020; Filippou et al., 1983)
	E[GPa]		205	15	
	b[-]		0.02	15	
	cR_1, cR_2[-]		0.915,0.15	5, 25	
	a_1, a_3[-]		0.04, 0.03	375, 500	
	$R_0, a_2. a_4$[-]	Constant	20, 1, 1	-	
SB	f_y[MPa]	Gaussian	266.24	8.4	(Martínez et al., 2017)
	g_1[-]		0.09	22.22	
	g_2[-]		0.39	10.26	
UEB	G_{eb}[kN/m2]	Gaussian	1083.6	3.6	(Rubilar, 2016)
	μ[-]		0.31	2.59	

After estimating the parameters for each of the data sets available for each model, the convergence of the posteriors was verified through the convergence criteria detailed in section 3.2. The posteriors obtained were then used to generate parameter samples and establish predictions about the models' responses, which were compared against the experimental data. The calibrated models achieved high accuracy in their ability to predict the data, considerably improving the fit with respect to the prior distributions found in the literature. On the one hand, the predictions obtained by evaluating the models at the mean of the posteriors were compared, thereby generating deterministic predictions. On the other hand, sampling the posteriors allowed the probabilistic characterization of parameter-induced uncertainty in the constitutive models, which can be included in future probabilistic structure-level analyses of bridge models. As an example, a summary of calibration results and convergence criteria verification for one of the reinforcing steel coupons is shown in Table 4.2-2. This verification procedure was repeated in all cases (i.e., each dataset and every model discussed).

Table 4.2-2. Summary of estimation and convergence results for Coupon C67 (Birrell, Astroza, Carreño et al., 2021).

C67	μ	σ	MCSE	quantile 2.5	quantile 97.5	$\hat{R}$
F_y (MPa)	473	3.7	0.192	466	480	1.001
E (MPa)	216184	2184.5	66.77	211931	220550	1.002
b	0.014	0.001	3.25e-05	0.012	0.015	1.001
cR_1	0.902	0.002	9.60e-05	0.897	0.906	1.002
cR_2	0.092	0.010	2.93e-04	0.073	0.112	1.000
a_1	0.036	0.001	7.32e-05	0.033	0.039	1.000
a_3	0.018	0.002	9.76e-05	0.015	0.022	1.000

The resulting posterior distributions for each model are summarized in Table 4.2-3, Table 4.2-4, and

Table 4.2-5, although full detail and discussion of the results can be found in the relevant publications (Birrell, Astroza, Carreño et al., 2021; Birrell, Astroza, Restrepo et al., 2021; Pinto et al., 2023).

Table 4.2-3. Posterior PDF for Steel02.

Parameter	**Steel Grade**	**Mean**	**C.O.V (%)**
$F_y\ [MPa]$	A615	512	5
$E\ [GPa]$		206	4
$b\ [-]$		0.023	27
$cR_1\ [-]$		0.893	1
$cR_2\ [-]$		0.093	31
$a_1\ [-]$		0.035	45
$a_3\ [-]$		0.022	51
$F_y\ [MPa]$	A706	467	10
$E\ [GPa]$		206	4
$b\ [-]$		0.018	18
$cR_1\ [-]$		0.889	2
$cR_2\ [-]$		0.095	27
$a_1\ [-]$		0.039	28
$a_3\ [-]$		0.029	28

$$\begin{array}{c} \\ F_y \\ E \\ b \\ cR_1 \\ cR_2 \\ a_1 \\ a_3 \end{array} \begin{bmatrix} 1 & 0.3 & -0.2 & 0.2 & -0.4 & -0.9 & -0.7 \\ & 1 & -0.4 & 0.7 & 0 & -0.2 & -0.2 \\ & & 1 & -0.8 & -0.3 & 0 & 0 \\ & & & 1 & 0.3 & 0 & 0 \\ & & sym. & & 1 & 0.4 & 0.4 \\ & & & & & 1 & 0.8 \\ & & & & & & 1 \end{bmatrix}$$

Correlation matrix for *Steel02*.

Table 4.2-4. Posterior PDF for SB.

Model	Parameter	Mean	C.O.V. (%)
CD	$F_1\ [MPa]$	206.4	6.76
	$g_1\ [-]$	0.104	27.6
SD	$F_1\ [MPa]$	264.4	7.6
	$g_1\ [-]$	0.091	13.8
	$g_2\ [-]$	0.395	6.25
$\begin{matrix} \\ F_1 \\ g_1 \end{matrix}\begin{bmatrix} 1 & 0.26 \\ 0.26 & 1 \end{bmatrix}$ Correlation matrix for the model with diaphragm (CD).			

Table 4.2-5. Posterior PDF for UEB.

Parameter	Mean	C.O.V (%).
$G_{eb}\ [kN/m^2]$	1176	1.15
$\mu\ [-]$	0.23	0.8
$\begin{matrix} \\ G_{eb} \\ \mu \end{matrix}\begin{bmatrix} 1 & 0.31 \\ 0.31 & 1 \end{bmatrix}$ Correlation matrix.		

4.2.3 CLOSING REMARKS ON THE CONSTITUTIVE MODEL CALIBRATIONS

In general, it can be observed that the incorporation of experimental data for the calibration of nonlinear constitutive models of bridge components can allow for a robust probabilistic characterization of these components. With regard to the structure-level probabilistic simulation required in PBEE, a robust local characterization allows for the rigorous identification and propagation of the various sources of parameter-induced uncertainty in structural responses. For this reason, there is a need to conduct more experimental campaigns and expand efforts to collect data on materials and structures, thereby advancing the understanding of the dynamic phenomenon of structural response to seismic events.

Likewise, the use of Bayesian tools for parameter estimation allows for the rigorous incorporation of such experimental information. In this section, different simulation-based methods are used for this purpose, achieving good fits and a probabilistic characterization that appears to adequately represent the complex behavior of nonlinear components present in bridges. The calibrated models use popular and widely available formulations (e.g., in OpenSees or equivalent formulations in commercial software), making their implementation in probabilistic simulation at the structural level straightforward.

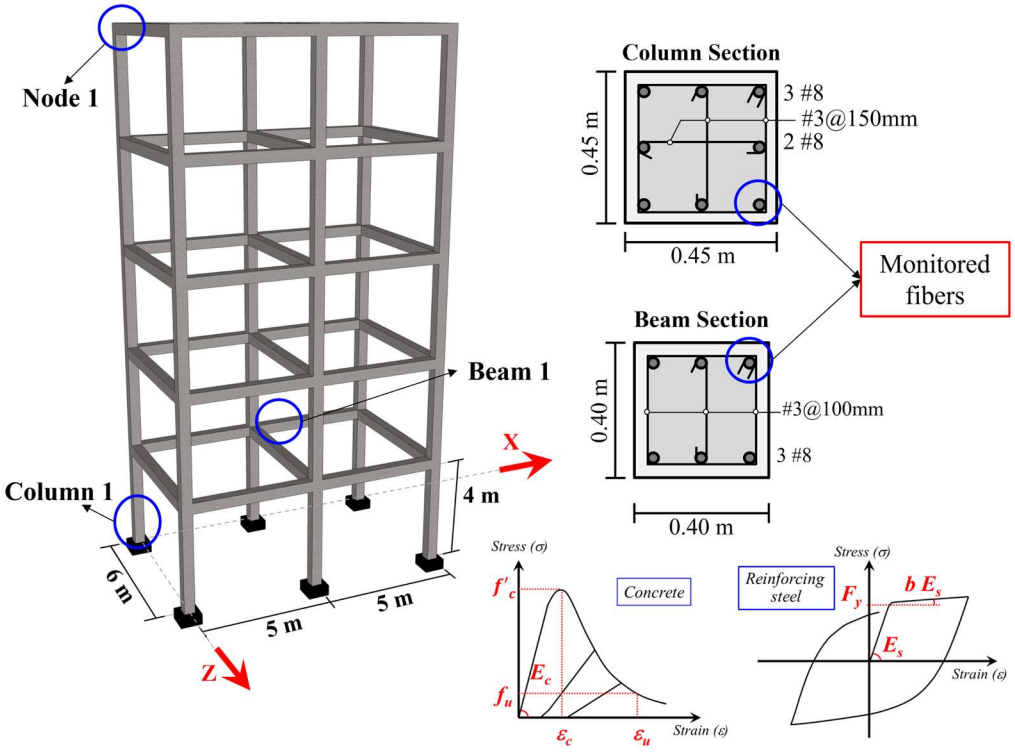
4.3 CASE STUDY I: APPLICATION I TO A REINFORCED CONCRETE BUILDING

The first case study is the application of the methodology proposed in section 3.6 to a 5-story RC moment resisting frame building (termed RCB herein), with one bay in the z-direction and two bays in the x-direction. The typical story height is 4m, while the z-bay width is 6m and the x-bay length is 5m. The building was presented originally by Astroza et al. (Astroza et al., 2017), where the nonlinear 3D-model was updated using Bayesian methods. The building has 6 columns with a 45cm×45cm cross-section and reinforced with 8 #8 Grade 75 longitudinal steel rebars and #3@150mm closed hoop and stirrups. The longitudinal beams are defined by a 40cm×40cm cross-section with 6 #8 Grade 60 longitudinal rebars and #3@100mm closed hoop and stirrups. The transverse beams have a 40cm×45cm cross-section with 6 #8 Grade 60 longitudinal rebars and #3@100mm closed hoop and stirrups. A diagram of the RCB design and modeling approach is shown in Figure 4.3-1.

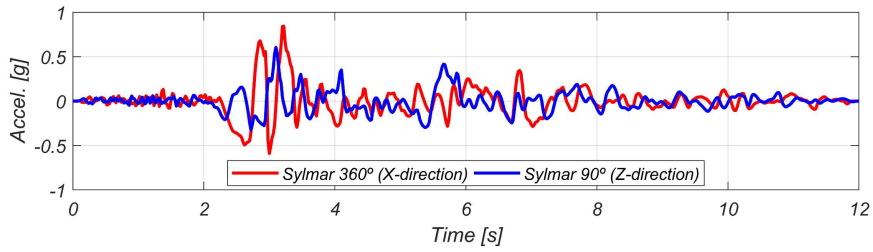
This section presents the GP SM developed for RCB to capture six relevant building responses from NLTHA under the 1994 Northridge earthquake recorded at the Sylmar station, shown in Figure 4.3-1b.

4.3.1 FINITE ELEMENT MODEL

The OpenSees platform (Mckenna et al.) was used to develop a model for the RCB, which uses displacement-based beam column elements, with a fiber section that considers concrete core, concrete cover, and reinforcing steel layers. Concrete and reinforcing steel fibers are characterized by the Popovics-Saenz (Balan et al., 2001; Popovics, 1973; Saenz, 1964) and the modified Giuffre-Menegotto-Pinto (GMP) (Filippou et al., 1983; Menegotto & Pinto, 1973) materials, respectively. For the GMP model, the yield strength ($F_{y,col}, F_{y,beam}$), elastic modulus ($E_{s,col}, E_{s,beam}$), and post-yield hardening coefficient (b_{col}, b_{beam}) are sampled for this analysis. For concrete, a common parameter was used for compressive strength of both columns and beams (f_c').



(a)



(b)

Figure 4.3-1. (a) 3D RC Building FE Model. (b) Components of the Sylmar ground motion

4.3.1.1 MODEL SAMPLING

The parameters for constitutive models used for the RCB were sampled according to the distributions shown in Table 4.3-1, in which the mean values are taken from (Astroza et al., 2017), according to the different steel grades used for each section, and the coefficient of variation (C.O.V) for each parameter was defined according to previous probabilistic model calibrations of the GMP model (Birrell, Astroza, Carreño et al., 2021) and reported statistical

data for concrete (Mirza et al., 1979b). The natural period of the model was close to 2.0s in each direction after applying gravity loads, subject to variability due to parameter sampling.

A total of 2,000 parameter samples were drawn using Sobol sequences (Sobol, 1967) and NLTHA of the RCB model under the 90° and 360° horizontal components of the Sylmar record from the 1994 Northridge earthquake were conducted for each sample. The seismic records had a duration of 12s, sampled at 50Hz. The time history analysis was conducted with a time step of 0.02s (i.e., same as the sampling time of the seismic records) using the Newton-Raphson algorithm with the Newmark average acceleration method. Six responses of interest were tracked from the time-history (TH) analysis. Specifically, the peak values of lateral displacements of node 1 (i.e., Δ_{x1}, Δ_{z1}), stress and strain at the monitored fiber in the base section of column 1 (i.e., $\sigma_{s2}, \varepsilon_{s2}$), and stress and strain at the monitored fiber at the end of beam 1 (i.e., $\sigma_{s3}, \varepsilon_{s3}$) were recorded. Node 1, column 1 and beam 1, along with the monitored fibers are shown in Figure 4.3-1. These responses are indicative of the global and local responses of the RCB.

Table 4.3-1. Parameter distributions for the RC building model.

Parameter	Mean	C.O.V. (%)	Distribution	Reference
$F_{y,col}$ (MPa)	517	10	Lognormal	(Astroza et al., 2017; Birrell, Astroza, Carreño et al., 2021)
$E_{s,col}$ (GPa)	200	4	Lognormal	
b_{col}	0.01	18	Lognormal	
$F_{y,beam}$ (MPa)	414	10	Lognormal	
$E_{s,beam}$ (GPa)	200	4	Lognormal	
b_{beam}	0.05	18	Lognormal	
f_c' (MPa)	40	20	Lognormal	(Astroza et al., 2017; Mirza et al., 1979b)

Histograms of all 2,000 samples from each response are shown in Figure 4.3-2, where considerable variability can be observed including, for example, a range between 17 cm and 38 cm displacement for Δ_{z1} and between 1% and 4.14% strain for ε_{s2}. Thus, parameter-induced uncertainty in RCB is significant. Roof displacements in the x-direction are larger than in the z-direction, although considerably larger variability is observed for Δ_{z1} than for Δ_{x1}. *Larger roof displacements in Δx1 than in Δz1 were obtained, while greater variability in Δz1 than in Δx1 was observed. This was most likely due to the greater spectral demands*

in the X-direction and higher spectral variability in the Z-direction of the Sylmar record, respectively. Similarly, the ranges of variation in stress and strain responses for the monitored column and beam fibers are significant, implying that key behavior change points (e.g., yielding of steel rebars) could have been reached for some samples.

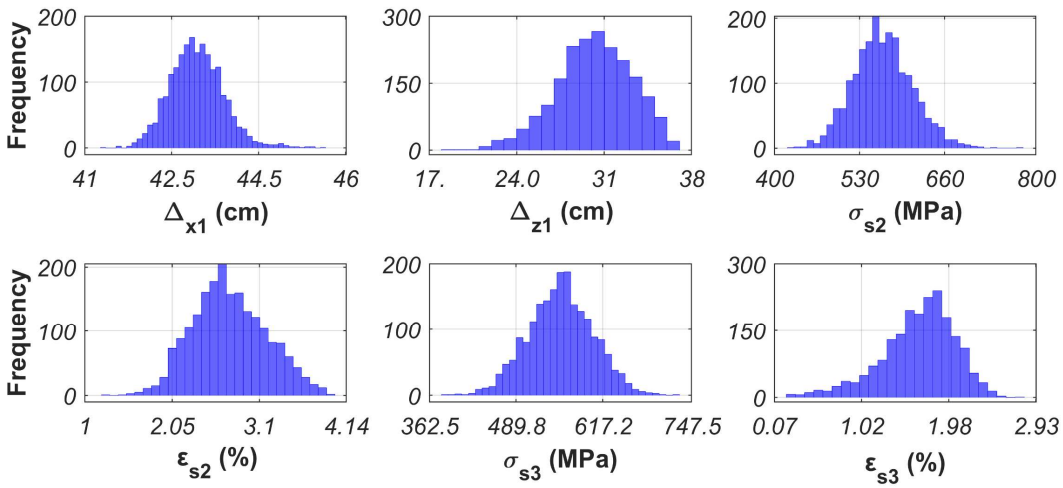


Figure 4.3-2. Histograms of building responses of interest.

4.3.1.2 GAUSSIAN PROCESS SURROGATE MODEL

For each of the six responses, a GP SM was developed independently. The GP prior was defined using a squared exponential kernel (see Eq. (3.4.10)) and the dataset obtained from the 2,000 NLTHA samples was split into 500 training and 1,500 testing data points. After the optimization of $\boldsymbol{\theta}$, the GP SM was validated by generating predictions on the remaining 1,500 data points, by evaluating the posterior distribution in Eq. (3.4.7). Then, the quality of the predictions was assessed through the three metrics defined in Eqs. (3.4.13) to (3.4.15) ($R^2, RRMSE, MAE$, respectively). It is common practice in machine learning to use an 80/20 training/testing split. In this work, however, the total of 2,000 samples were generated a priori to ensure a large enough size of data for the scope of this research. The 25/75 split in case study 1 was selected after iteratively increasing the training dataset size and inspecting the goodness of fit of the GP models. Although this convergence analysis is omitted from case study 1 for brevity, the same procedure was followed for case study 2 for which the convergence analysis is presented later in section 4.4.3.7. In a practical use scenario, where computing time could be a major constraint in generating the complete dataset, the

convergence analysis and sample generation should be interconnected to achieve a good enough fit for the desired application with a dataset as small as possible.

The resulting GP surrogate achieved great accuracy for both training and testing data sets, as depicted in Figure 4.3-3. In the figure, each subplot shows the GP predictions on training data in blue triangles, GP predictions on testing data in red dots, and the R^2 value of GP predictions on the testing data. In black, the identity (i.e., $x = y$) represents the FE model responses. Predictions which lie on the black line achieved a perfect fit with respect to the FE model. Likewise, values of R^2 close to 1.0 indicate a near perfect fit between GP predictions and test data.

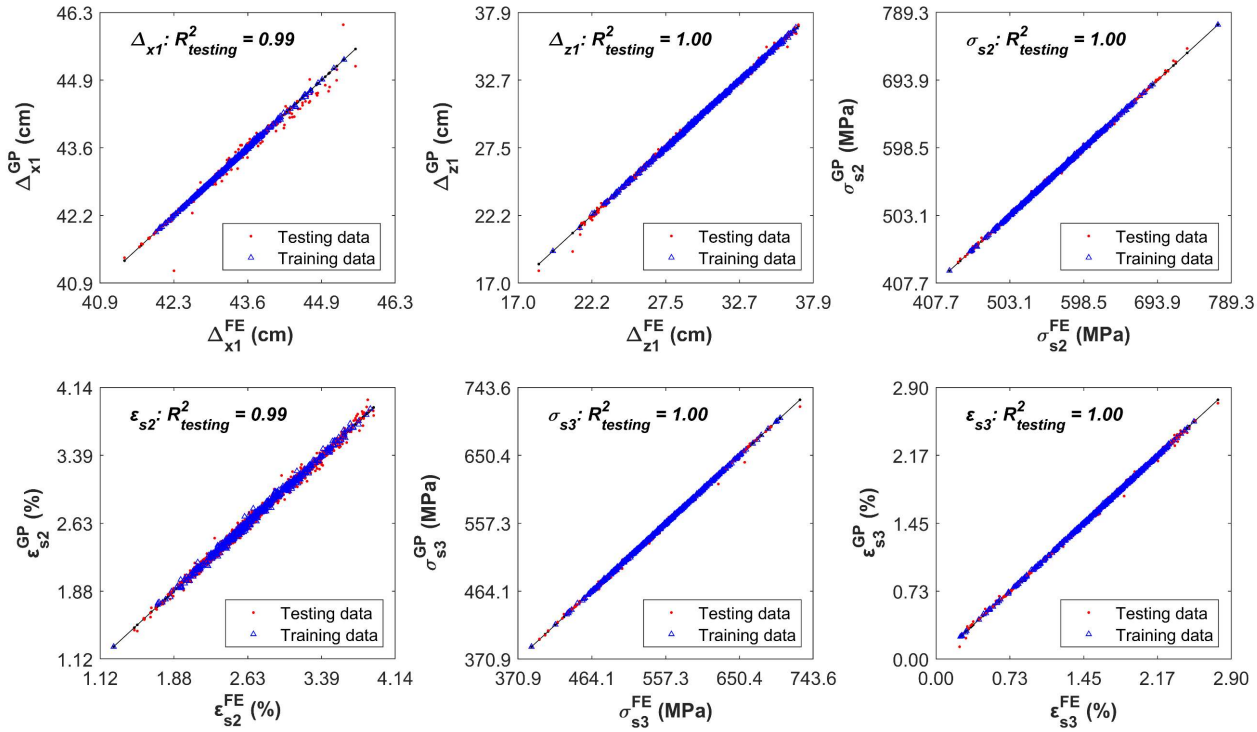


Figure 4.3-3. GP surrogate model predictions for building responses.

Error metrics are summarized in Table 4.3-2. A value of R^2 of over 0.99 was achieved for all responses, along with $RRMSE$ lower than 1.3% and MAE lower than 0.86. As a result, GP surrogates were able to explain over 99% of model response variance, while maintaining accurate predictions over the range of model responses. Additionally, MAE values account for units (e.g., the MAE for ε_{s2} was 0.027% in strain units, where ε_{s2} ranged between 1.12% and 4.14%, and for σ_{s3} it was 0.852 MPa with σ_{s3} in a range between 370 MPa and 740

MPa). Therefore, the GP surrogate presented in this section could be used reliably to generate samples of the RCB under the Sylmar record.

Table 4.3-2. Error metrics of the GP regression fit for the building model taken on the testing data.

	Δ_{x1} (cm)	Δ_{z1} (cm)	σ_{s2} (MPa)	ε_{s2} (%)	σ_{s3} (MPa)	ε_{s3} (%)
R^2	0.991	0.999	0.99	0.994	0.999	0.999
$RRMSE$ (%)	0.12	0.34	0.20	1.26	0.23	0.68
MAE	0.017	0.070	0.760	0.027	0.852	0.008

4.3.1.3 LSA COMPARISON

To validate the possibility of replacing the original FE model with the GP surrogate for further sampling, LSA was conducted on each of them to obtain correlation indices using Eq. (3.1.1). The indices were computed for the FE model using the 2,000 samples originally obtained from the NLTHA. For the GP SMs, the indices were computed from 2,000, 10,000, and 50,000 new samples, with the purpose of verifying that the correlation structure between model parameters and responses was correctly captured by the GP at different sample sizes. The indices are shown in Figure 4.3-4, where it can be seen that ρ^{FE} and ρ^{GP} match closely, for all GP sample sizes. The difference between ρ^{FE} and ρ^{GP} is lower than 0.1 for the different model parameters, confirming that the GP model can effectively be used instead of the FE model for sampling in the GSA stage.

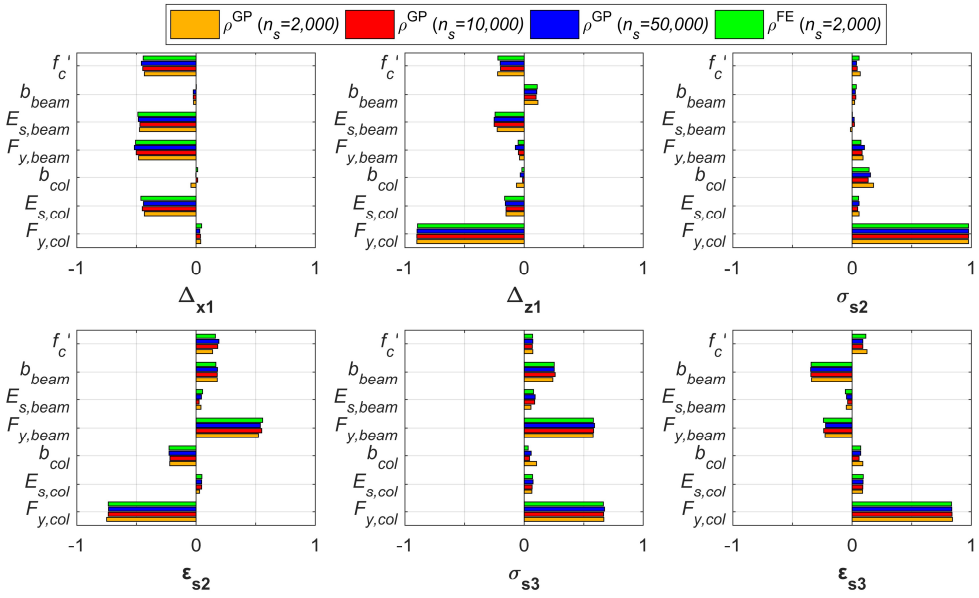


Figure 4.3-4. Comparison of correlation indices from the FE and GP models.

4.3.1.4 GP-BASED GSA

Using the GP SMs, GSA was carried out for the RCB model by computing Sobol indices via simulation using the T-estimator from Eq. (3.1.10). A total of $N = 10{,}000$ samples were taken, requiring 90,000 model evaluations. The resulting indices are shown in Figure 4.3-5, where the first order (S^1) and total order (S^T) indices are shown in blue and red, respectively.

These indices measure the variability induced in model responses by each parameter, where, for example, the uncertainty about peak roof-level drift is dependent on beams and columns' properties on the x-direction (Δ_{x1}), but heavily dependent on $F_{y,col}$ on the z-direction (Δ_{z1}), in which there is only one bay. Likewise, it could be of interest to observe that both $F_{y,col}$ and $F_{y,beam}$ contribute significantly to the uncertainty in peak strain at the base of column 1 (ε_{s2}) and stress at the joint-end of beam 1 (σ_{s3}). Such results reflect the inter-connection between building components during a dynamic excitation. Furthermore, interactions between parameter groups can be quantified by the difference between S^T and S^1, which were not significant.

Although in this case, where the structure is a RC frame, it might be straightforward to pinpoint the key interactions, this study case serves as an example of a SM being effectively used to perform GSA through Sobol indices efficiently. On the contrary, even for a simple nonlinear structure, the computational cost of performing $n_s(m+2)$ model evaluations could be prohibitive.

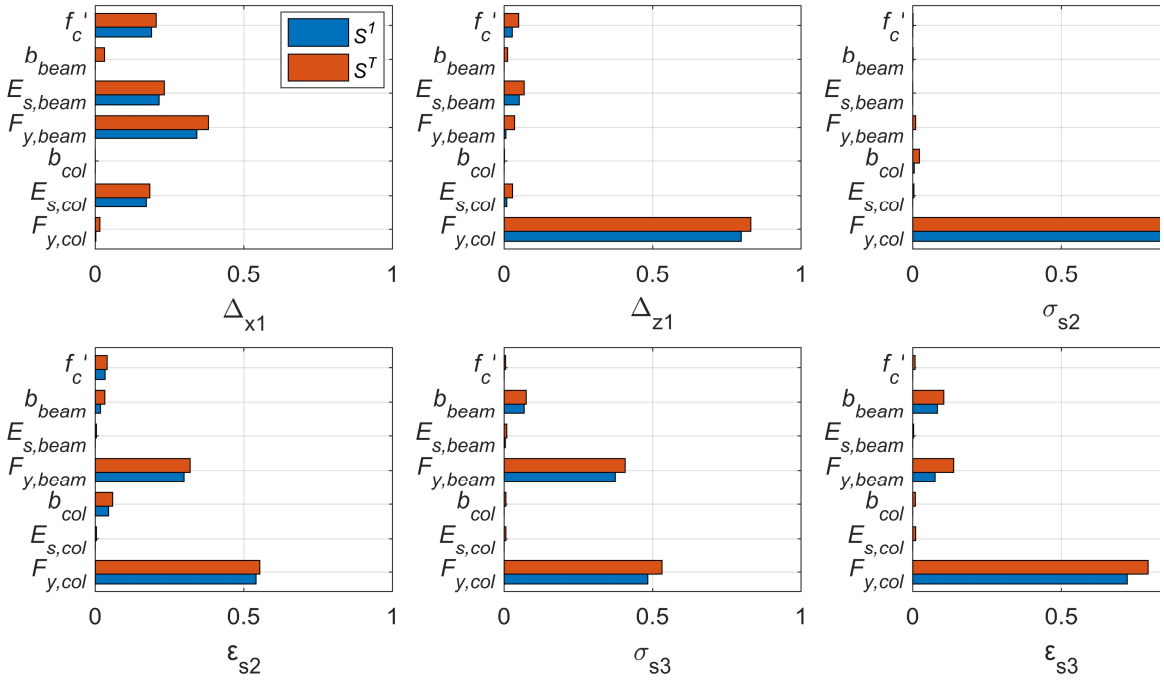


Figure 4.3-5. Sobol indices for the RCB case using the GP models.

4.4 MAIN CASE STUDY: AGUILA NORTE BRIDGE MODEL

A fixed-base version of the AN's structure level FE model is considered. A schematic of the actual bridge and the OpenSees (Mckenna et al.) model is shown in Figure 4.4-1. The model incorporated elastic Timoschenko beam-column elements for the deck. For the foundation and cap beams in each pier, elastic beam-column elements were used. The columns were connected to both beams using rigid elastic beam-column elements, and a rigid diaphragm was applied at the top end of the bent.

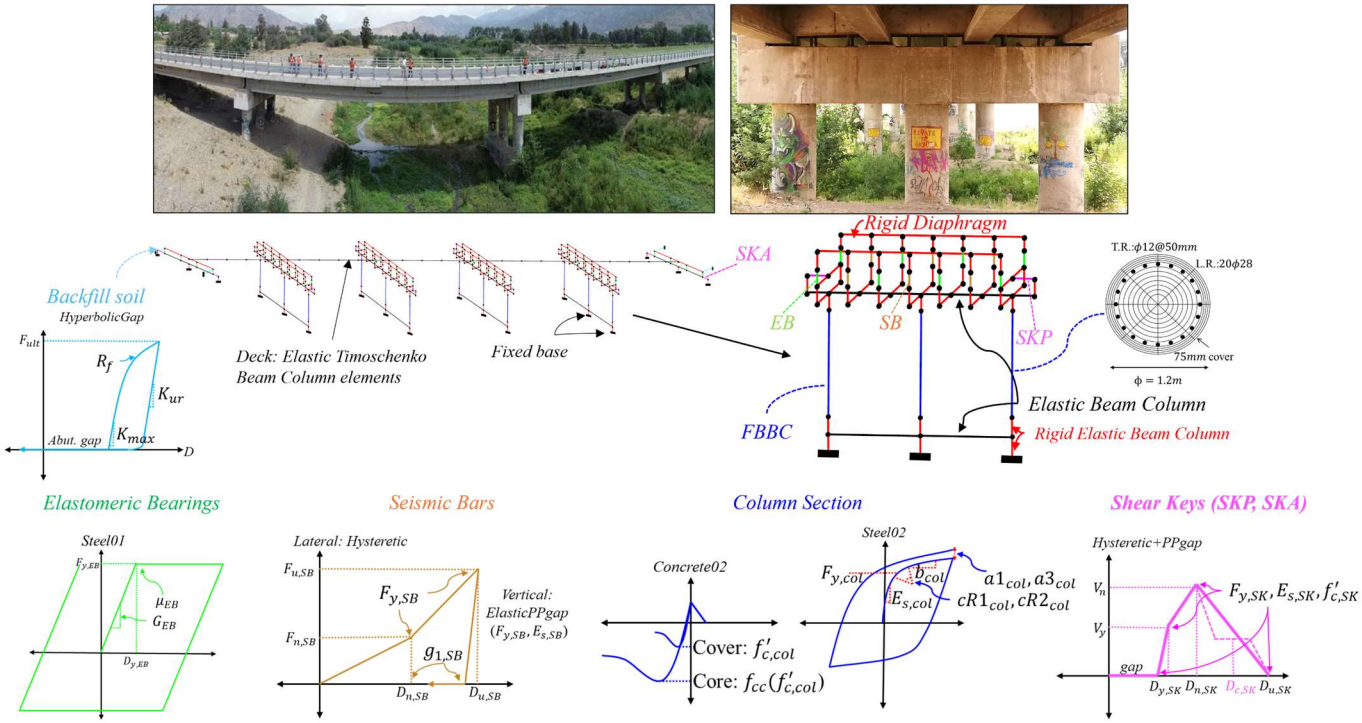


Figure 4.4-1. AN bridge FE model.

Force-based beam-column elements (FBBC) were used for pier columns. The circular column section, with a diameter of 1.2m, was designed with 20 ϕ28 longitudinal rebars and a transverse reinforcement of ϕ12@50mm. *Concrete02* and *Steel02* material models were used for concrete (i.e., core and cover) and reinforcing steel, respectively. The lateral force-deformation of unanchored elastomeric bearings (EB) were represented by the *Steel01* material model proposed by Rubilar (Rubilar, 2016), while the vertical and rotational behaviors are assumed elastic. Vertical seismic bars (SB), commonly used in Chilean bridges to prevent vertical uplift of the deck due to vertical earthquake motion, were modeled with a *Hysteretic* material, as proposed by Martínez (Martínez, 2016), for lateral response and an

ElasticPPGap material for the vertical response. Monolithic exterior shear keys (SK) in piers (SKP) and abutments (SKA) were modeled using a *Hysteretic* model in series with an *ElasticPPGap* material (gap = 0.075m for SKP and 0.06m for SKA), based on the simplified tri-linear model proposed by Goel and Chopra (Goel & Chopra, 2008). The implementation of the extended model proposed in (Birrell et al., 2024) was used, although due to the SK design of AN, the failure mode is always expected to be in diagonal tension rather than sliding shear. Passive pressure of the backfill soil (BF) in bridge abutments was modeled using a *HyperbolicGap* material (gap = 0.125m) with the definition proposed by Duncan and Mokwa (Duncan & Mokwa, 2001).

From the FE model, seven responses were recorded: lateral and longitudinal section curvature at the bottom of the columns (i.e., $K_{col,lat}$, $K_{col,lon}$), longitudinal relative displacement of EB, SB, and BF (i.e., Δ_{EB}, Δ_{SB}, Δ_{BF}), and lateral displacement in SKP and SKA (i.e., Δ_{SKP}, Δ_{SKA}). For every component, the maximum response was taken (e.g., $K_{col,lon} = \max_{j\in\{1,\dots,12\}}(\max_{TH}(K^{j}_{col,lon}))$ for the maximum longitudinal curvature among the 12 pier columns). Similarly, the maximum response for the 32 pier EB, the 24 SB, the 8 SKP, the 4 SKA, and 2 BF were taken.

The time history analyses in the following sections utilized the Newton-Raphson algorithm with the Newmark average acceleration method and a time step of 0.005s.

The AN bridge has been extensively investigated in different applications, including soil structure interaction effects, soil spatial variability, effects of multiple hazards considering corrosion and local pier scour, and the development of guidelines for the nonlinear analysis of Chilean bridges, among others (Guajardo et al., 2024; Pastén et al., 2021; Pinto, Astroza, Bazáez et al., 2024; Pinto, Astroza, Pizarro et al., 2024). Therefore, the validation of the methodologies proposed in sections 4.4.3 and 0 is an idoneous choice to extend the research on bridge performance in the national context, as it is representative of a typical RC highway bridge in Chile.

4.4.1 MODEL PARAMETER SAMPLING

Constitutive model parameters for columns, EB, SB, SKP, and SKA were defined as uncertain quantities for sampling, while the BF response was defined as deterministic. Bayesian calibration of *Steel02* (i.e., $F_{y,col}, F_{u,col}, E_{s,col}, b_{col}, cR1_{col}, cR2_{col}, a1_{col}, a3_{col}$) (Birrell, Astroza, Carreño et al., 2021), EB (i.e., G_{EB}, μ_{EB}), and SB parameters (i.e., $F_{y,SB}, g_{1,SB}, E_{s,SB}$) (Pinto et al., 2023) have been previously carried out using experimental data representative of the AN bridge design. Uncertainty in the strength of concrete (i.e., $f'_{c,col}, f'_{c,SKP}, f'_{c,SKA}$) was obtained from extensive statistical data by Mirza et al. (Mirza et al., 1979b) and uncertainty in SK reinforcing steel parameters (i.e., $E_{s,SKP}, F_{y,SKP}, E_{s,SKA}, F_{y,SKA}$) was obtained from (Birrell, Astroza, Carreño et al., 2021), as described in section 4.1. Deterministic parameter values, as well as detailed modeling of the AN bridge, can be found in (Pastén et al., 2021). All PDFs for uncertain parameters are summarized in Table 4.4-1, and parameter correlations from Bayesian calibrations of *Steel02*, EB, and SB parameters are summarized in

Table 4.4-2.

To generate the initial FE model simulation, parameter samples were drawn using Sobol sequences (Sobol, 1967).

Table 4.4-1. Model parameter PDF.

Parameter	Mean (μ)	C.O.V. (%)	Distribution	Component	Reference
$F_{y,col}$ (MPa)	467	10	Lognormal	Columns (*Steel02*)	(Birrell, Astroza, Carreño et al., 2021)
$F_{u,col}$ (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*)	(Mirza et al., 1979b)
G_{EB} (MPa)	1.176	3	Lognormal	EB	(Pinto et al., 2023)
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	(Birrell et al., 2024; Mirza et al., 1979b)
$E_{s,SKP}$ (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		

Table 4.4-2. Model parameter correlation matrices.

		$F_{y,col}$	$E_{s,col}$	b_{col}	$cR1_{col}$	$cR2_{col}$	$a1_{col}$	$a3_{col}$
Steel02	$F_{y,col}$	1	0.3	-0.2	0.2	0.4	-0.8	-0.7
	$E_{s,col}$	0.3	1	-0.4	0.7	0	-0.2	0
	b_{col}	-0.2	-0.4	1	-0.8	-0.3	0	0
	$cR1_{col}$	0.2	0.7	-0.8	1	0.3	0	0
	$cR2_{col}$	0.4	0	-0.3	0.3	1	0.2	0.4
	$a1_{col}$	-0.8	-0.2	0	0	0.2	1	0.8
	$a3_{col}$	-0.7	0	0	0	0.4	0.8	1

		G_{EB}	μ_{EB}
EB	G_{EB}	1	0.31
	μ_{EB}	0.31	1

		$F_{y,SB}$	$g_{1,SB}$
SB	$F_{y,SB}$	1	0.26
	$g_{1,SB}$	0.26	1

4.4.2 PROBABILISTIC SEISMIC HAZARD ANALYSIS OF THE AN BRIDGE

PSHA was conducted numerically on the AN bridge's site following the methodology described in section 3.3. For this purpose, a platform was developed, which builds upon the work by Candia (Candia et al., 2019). The numerical implementation was defined as follows:

The geometry of Chilean subduction was adopted from the Slab2 model (Hayes et al., 2018). This model provides a 3-D surface of the subduction zone for Chile, which was later meshed into individual triangular sources. The mesh, along with the location of the AN bridge in Chile's central zone is shown in Figure 4.4-2, where the site is marked as a yellow star. For the purposes of the analysis, each triangle was considered as a possible seismic source, of which their distance to the site was taken from the triangle's centroid.

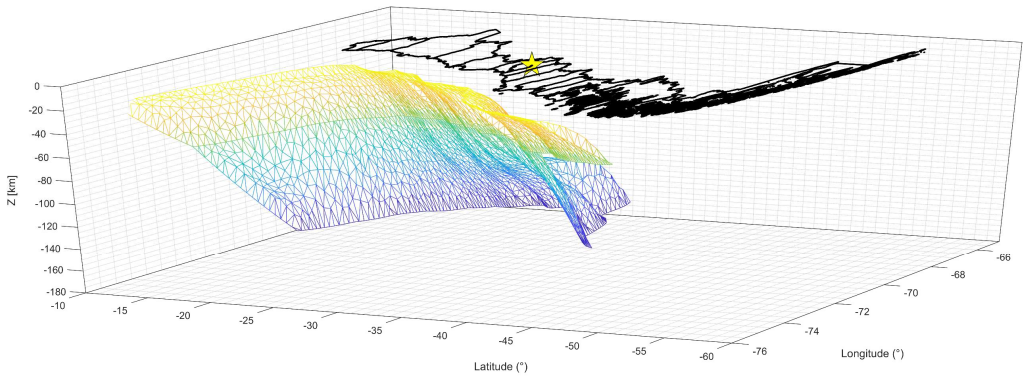


Figure 4.4-2. Mesh of the subduction geometry and location of the AN's site.

The seismic recurrence and zonation model proposed by Poulos et al. (Poulos et al., 2019) was used to define the λ_s and $f_M(m)$ for each zone, under the common assumption that $f_{M,R}(m,r) = f_M(m)f_R(r)$. From the seismicity model, $f_M(m)$ is defined as a truncated Gutenberg-Richter law using the coefficients calibrated by Poulos et al. for each of seven zones defined in that model. The coefficients and zonation can be referred to in (Poulos et al., 2019). The distribution of distances $f_R(r)$ was defined as a uniform distribution, in which the probability of each source was computed as the area of the corresponding mesh triangle divided by the total area.

To compute the hazard curve in Eq. (3.3.1), the GMM proposed by Idini et al. (Idini et al., 2017) and the "HQ" version of the GMM proposed by Montalva et al. (Montalva et al., 2017) were considered, referred to as Idi16 and Mon17HQ, respectively. Both GMMs were developed and calibrated to predict spectral acceleration (S_a) in Chile, accounting for interface and intraslab mechanisms. In Eq. (3.3.2), the logic tree was computed with $w_k = 0.5$ for both models. The resulting hazard curves and the weighted curve are shown in Figure 4.4-3.

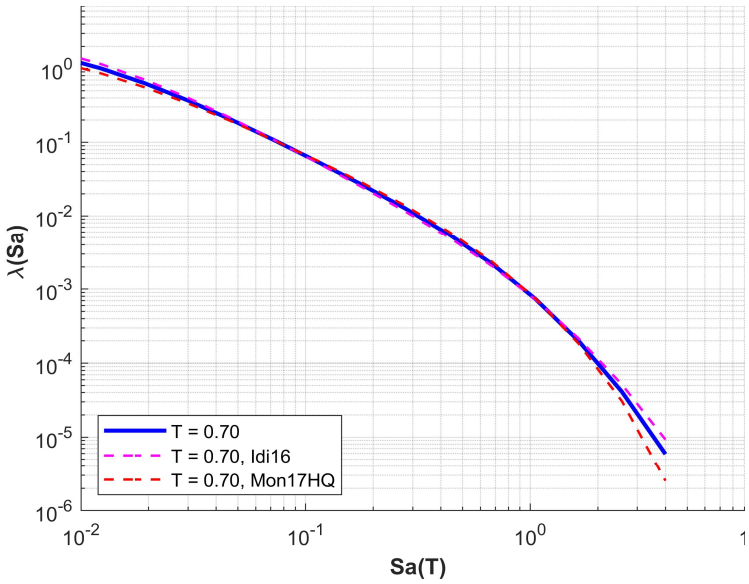


Figure 4.4-3. Hazard Curve for Aguila Norte.

HAZARD-SPECIFIC GROUND MOTION SELECTION

For hazard-specific scenarios, such as the analysis presented later in section 4.4.3, it is key to select GMs that are consistent with the hazard levels of interest. To this end, it is common to use hazard deaggregation methods to derive target spectra for GM selection.

For the purpose of application I in section 4.4.3, the selection of GM consistent with the hazard estimated for the AN site was done with CMS spectra (Baker, 2011) as computed according to section 3.3 and considering the SIBER-RISK (S-R) strong motion database published by Castro et al. (Castro et al., 2022), which includes 7948 GMs recorded in Chile between 1985 and 2023. Multiple records from several stations are included in this database from the major earthquakes observed in Chile during this time.

Magnitude-Distance (M-R) deaggregation of the hazard curve was used to define the CMS scenarios of interest, which are summarized in Table 4.4-3. In the table, T_1 is the fundamental period of the AN bridge, T_r is the return period for DLE and MCE, respectively, $S_a(T_1)$ is the spectral acceleration for each scenario, obtained from each GMM's hazard curve, and Mw, Rrup, Rhyp, Z, Rjb are the mean magnitude, distance to rupture, hypocentral distance, hypocentral depth and Joyner-Boore distance, respectively, obtained from each scenario's M-R deaggregation.

Table 4.4-3. M-R Deaggregation scenarios for each hazard level and GMM.

T_1 (s)	GMM	Tr (yrs)	S_a(T1) (g)	Mw	Rrup (km)	Rhyp (km)	Z (km)	Rjb (km)
0.7	Mon17HQ	461	0.715	8.04	104.1	152.1	73.0	68.6
0.7	Idi16	461	0.723	7.92	93.5	138.5	51.0	66.6
0.7	Mon17HQ	2475	1.434	8.22	89.7	142.5	75.0	48.9
0.7	Idi16	2475	1.332	8.21	84.0	142.9	49.4	54.5

Based on the deaggregation scenarios, the CMS was computed using Lin et al.'s (Lin, Haselton, & Baker, 2013a, 2013b) third method, in which the CMS for each scenario is computed independently, and then weighted by a logic tree (in this case, 50% for each GMM) to obtain the CMS for each hazard level.

Afterwards, the GM selection for application I considered one ground motion per hazard level, with the purpose of focusing on parameter-induced variability in bridge response, rather than earthquake-induced. As such, the procedure to select GMs using the amplitude scaling method consisted of the following: (i) the S-R catalog was filtered to consider only records with a PGA greater than 0.1g and a Vs_{30} between 180 m/s and 760 m/s, thus keeping the records with similar characteristics to the AN site of Vs_{30} = 275 m/s. After this filter, a subset of 229 records remained. (ii) The RotD50 spectrum of each GM was extracted to consider directionality effects. (iii) The mean quadratic error between each record and the target CMS was computed and the record with the lowest discrepancy in the range of periods [$0.2T_1$, $1.5T_1$] was selected. (iv) The selected record for each hazard level was scaled by a scaling factor $SF = S_a(T_1)_{CMS}/S_a(T_1)_{RotD50}$ to match the target $S_a(T_1)$ from the respective CMS.

The selected GMs for each hazard level are summarized in Table 4.4-4, where a record from the 2014 Iquique M_w=7.6 earthquake and a record from the 2010 Maule M_w=8.8 were selected for DLE and MCE, respectively. Both events correspond to interface mechanisms. The selected GMs closely matched the CMS in the range of periods of interest, as shown in Figure 4.4-4, and especially at T_1 = 0.7s, as evidenced by their SF of 0.9356 and 0.9615 for DLE and MCE, respectively.

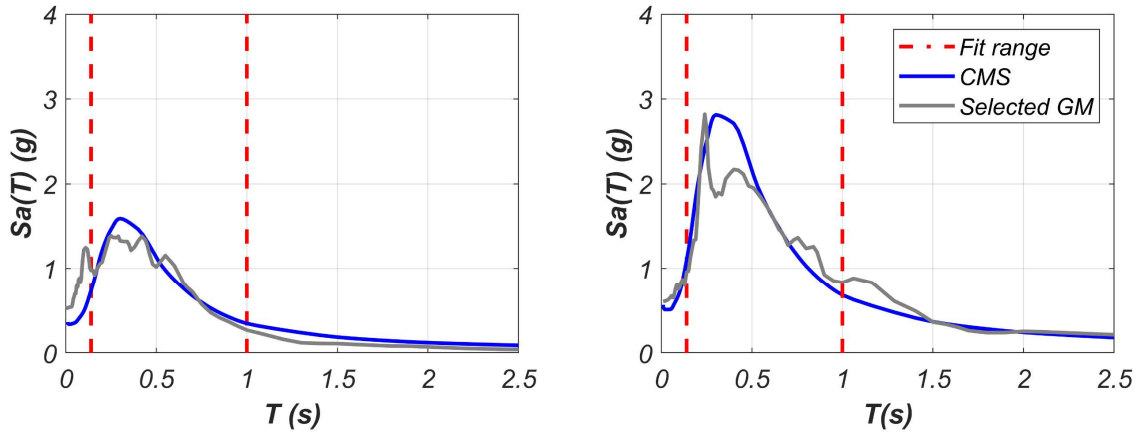


Figure 4.4-4. CMS and ground motion selection for (a) DLE and (b) MCE.

Table 4.4-4. Selected ground motions for Aguila Norte.

T_r (yrs)	Event	M_w	Type	SF
461	Iquique – April 3rd, 2014	7.6	Interface	0.9356
2475	Maule – February 27th, 2010	8.8	Interface	0.9615

CLOUD-BASED GROUND MOTION SELECTION

For PBEE applications, such as application II presented in section 4.4.4, it is important to consider the uncertainty in input excitations (i.e., GMs) as they might be the most significant source of uncertainty in structural demands. With the purpose of developing a framework for the comprehensive treatment of uncertainties in PBEE, including the UQ&P in the stages of PSHA, PSDA, and PSFA, application II considers cloud-based analysis including an extensive database of GMs.

To perform cloud-based PSFA in section 4.4.4, a suite of 211 unscaled subductive GMs, including 97 records from the S-R Chilean database (Castro et al., 2022) and 114 records from the Japanese database (Fayaz et al., 2024; National Research Institute for Earth Science and Disaster Resilience), were selected for NLTHA. Records from interface and intraplate earthquakes were included, and all records had a RotD50 spectral acceleration at 0.7s (i.e., $S_a(0.7)$) of at least 0.2g. In subsequent analyses, $S_a(0.7)$ is used as the *IM*.

Figure 4.4-5 shows the RotD50 spectra of selected GMs and histograms of the *IM* from each source. It can be noted that the spectra for Chilean and Japanese GMs are similar in shape,

with strong peaks at about $T \approx 0.5s$ and a rapid decay, wherein a second peak can sometimes be observed at $T \approx 2s$. The decision to include Japanese GMs was motivated by the lack of abundant strong (i.e., *IM*>1g) demands in the Chilean database alone and justified by the similarities in spectra from both databases.

By interpolating the *IM* of each selected GM in the hazard curve obtained from PSHA, it was verified that the selected suite covered a range of return periods from $T_r = 30$ to $T_r = 5{,}000$ years, approximately, as illustrated by Figure 4.4-6. In the figure, the hazard curve is plotted in blue, and the hazard of each GM is shown as a red dot, accompanied by a histogram of return periods for all GMs on the right-hand side.

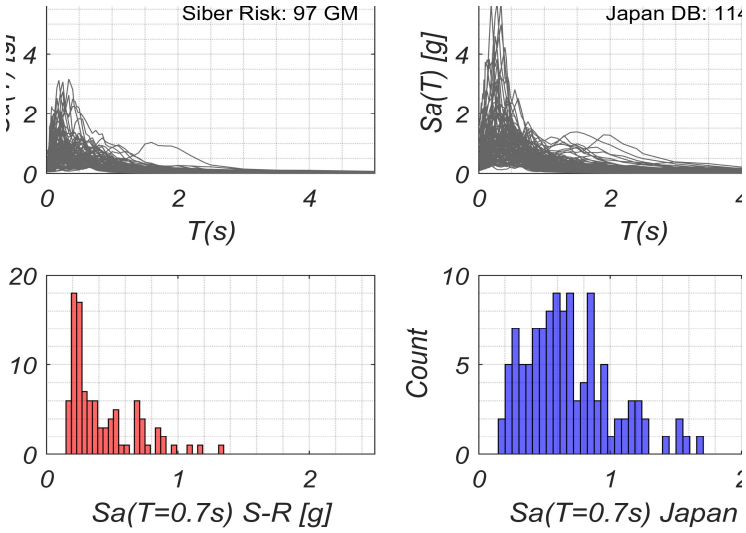


Figure 4.4-5. Selected GMs.

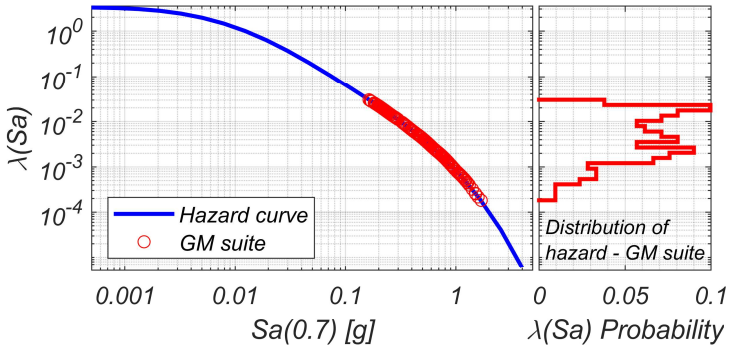


Figure 4.4-6. Hazard curve for the AN bridge and hazard levels of selected GMs.

4.4.3 APPLICATION I: GP-BASED APPROACH FOR UNCERTAINTY QUANTIFICATION IN HAZARD-SPECIFIC SCENARIOS

4.4.3.3 MODEL SIMULATIONS

For NLTHA of the AN bridge, 10,000 parameter samples were generated from the PDFs in Table 4.4-1 using Sobol sequences (Sobol, 1967). It was verified that the sample distribution converged to the target distribution of each parameter. Two hazard levels were considered for simulation. Termed as Design Basis Earthquake (DBE) and Maximum Considered Earthquake (MCE), they were defined as the hazards with a probability of occurrence of 15% in 75 years and 2% in 50 years, respectively. These represent return periods of 461 and 2475 years, respectively. PSHA was conducted to select one GM for each hazard level, as described in section 4.4.2. Among other alternatives, such as spectral matching, amplitude scaling was selected due to the availability of actual ground motions recorded in Chile that are representative of the expected seismic hazards to AN. The selected records were taken from the 2014 Iquique Mw=7.6 earthquake for DBE and from the 2010 Maule Mw=8.8 earthquake for MCE. The DBE records had a duration of 116s, and the MCE records had a duration of 143s, both at a sampling rate of 200Hz.

For each hazard level (i.e., DBE and MCE), 10,000 OpenSees FE models were executed in a high-performance computing cluster at University of los Andes, using five nodes equipped with AMD Ryzen 9 5950X CPUs, with 64gb RAM memory each. With this configuration, 150 concurrent models ran in parallel, with one model per thread. The wall-clock computing time in this setup was about 100 hours for DBE and 120 hours for MCE.

From these samples, the TH maximum responses of all components were recorded using envelope recorders. Afterwards, the maximum envelopes among components of the same type were extracted (e.g., $\Delta_{EB}^{k} = \max_{k}(\max_{i} \Delta_{EB,i}^{k}))$, $k = \{1, \dots, 10000\}, i = \{1, \dots, 32\}$ for EB). This strategy aligns with state-of-the-art limit state definitions that focus on material level metrics, such as the framework for performance analysis of bridges proposed by Deb et al. (Deb, 2021; Deb et al., 2022).

Figure 4.4-7 shows histograms of the seven model responses obtained for DBE (a) and for MCE (b) hazard levels. It was apparent that parameter uncertainty induced significant

variability in model responses. For example, at DBE a range between 0.0023 to 0.0053 was observed for $K_{col,lon}$, which could lead to different decisions in terms of damage state determination in fragility analysis depending on limit state definitions, which will be discussed in the following section. At MCE, an even larger variability was observed in model responses, where for $K_{col,lon}$ a range between 0.003 and 0.009 was observed. In this case, as many as two or three different damage states could have been covered only due to parameter uncertainty.

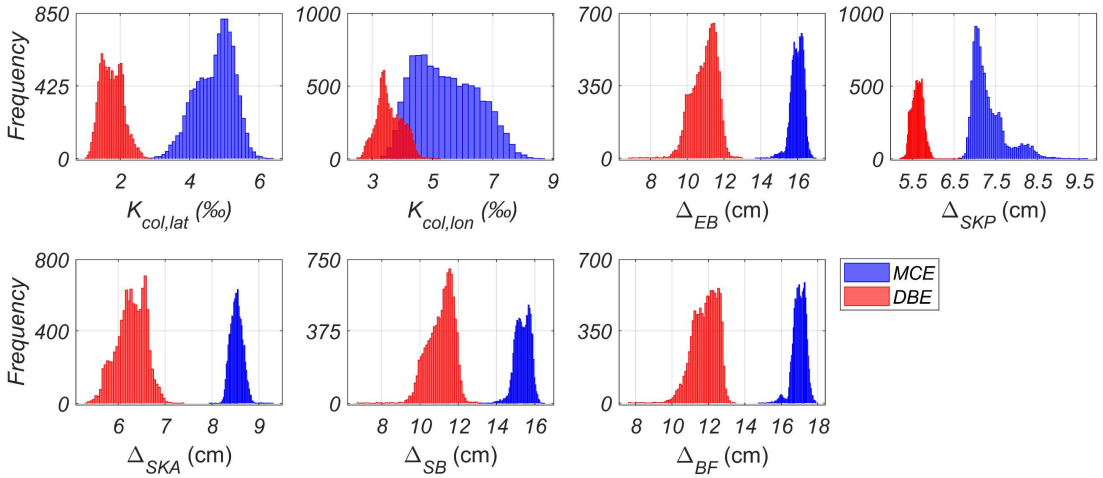


Figure 4.4-7. Histograms of maximum responses from each bridge component.

4.4.3.4 LIMIT STATE DEFINITIONS FOR BRIDGE COMPONENTS

Component responses shown in Figure 4.4-7 present variability induced by the uncertainty in model input parameters. To correctly define seismic demands, this uncertainty must be accounted for in the capacity definition of each component. To this end, and within the scope of application I, the first limit state (LS) is defined for each component in terms of local damage metrics. For columns, elastomeric bearings, and abutment backfill, the recommendations by Stefanidou and Kappos (Stefanidou & Kappos, 2017) and Stefanidou et al. (Stefanidou et al., 2022) were adopted. For columns, the first LS (LS1), is taken to be the yield curvature of the section ($K_{y,col}$), defined as the curvature at which reinforcing steel bars yield. For elastomeric bearings, LS1 is defined as the yield displacement of the component ($\Delta_{y,EB}$). The same principle was followed to define LS1 for shear keys and seismic bars. For shear keys, the nominal displacement ($\Delta_{n,SKP}, \Delta_{n,SKA}$), defined as the displacement at which the maximum capacity is reached, is considered (Goel & Chopra,

2008). For seismic bars, the yield displacement ($\Delta_{y,SB}$) is considered (Hube et al., 2017). For backfill soil, LS1 was defined as $\Delta_{LS1,BF} = 1.1\delta_{gap}$ where the gap for AN is 12.5 cm, as recommended by (Stefanidou & Kappos, 2017).

Table 4.4-5. LS statistics for bridge components.

	$K_{y,col}$ $(-)$	$\Delta_{y,EB}$(m)	$\Delta_{n,SKP}$ (m)	$\Delta_{n,SKA}$ (m)	$\Delta_{y,SB}$(m)
$\boldsymbol{\mu}$	0.00325	0.0304	0.10	0.076	0.035
C.O.V.	8%	3%	16%	13%	24%

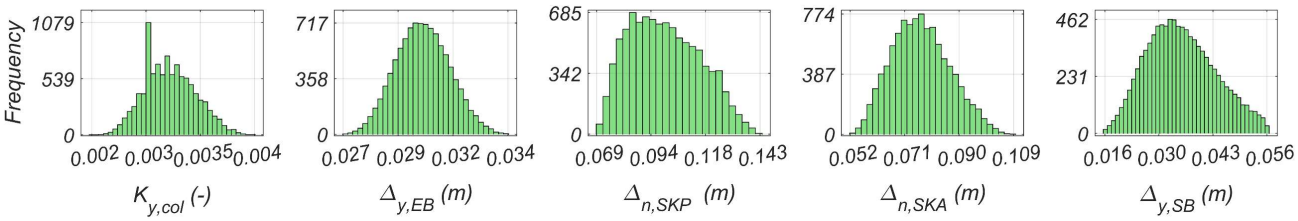


Figure 4.4-8. Sampled LS1 for each component.

4.4.3.5 GAUSSIAN PROCESS MODELING OF SEISMIC DEMANDS

To develop the seven independent SMs for the AN bridge, one for each EDP, the FE model responses were normalized by their corresponding LS samples from Figure 4.4-8. In this manner, the EDPs were defined as follows: lateral and longitudinal curvature ductility demand at columns ($\phi_{col,lat} = \frac{K_{col,lat}}{K_{y,col}}$ and $\phi_{col,lon} = \frac{K_{col,lon}}{K_{y,col}}$, respectively), displacement ductility demand at elastomeric bearings, shear keys, and seismic bars ($\delta_{EB} = \frac{\Delta_{EB}}{\Delta_{y,EB}}$, $\delta_{SKP} = \frac{\Delta_{SKP}}{\Delta_{n,SKP}}$, $\delta_{SKA} = \frac{\Delta_{SKA}}{\Delta_{n,SKA}}$, and $\delta_{SB} = \frac{\Delta_{SB}}{\Delta_{y.SB}}$, respectively), and displacement at abutment backfill soil (δ_{BF}). A close to perfect correlation ($\rho = 0.99$) between Δ_{BF} and Δ_{EB} was observed. Due to the lack of sampling of BF model parameters, and to mitigate the potential effect of low value outliers in Δ_{BF} on the assumptions of homoscedasticity in GP modeling, δ_{BF} was also defined in a normalized expression as $\delta_{BF} = \Delta_{BF}/\Delta_{y,EB}$. The EDPs for each component are shown in Figure 4.4-9, where it can be seen that by normalizing each response by its corresponding LS sample, smoother distributions were obtained than those in Figure 4.4-7.

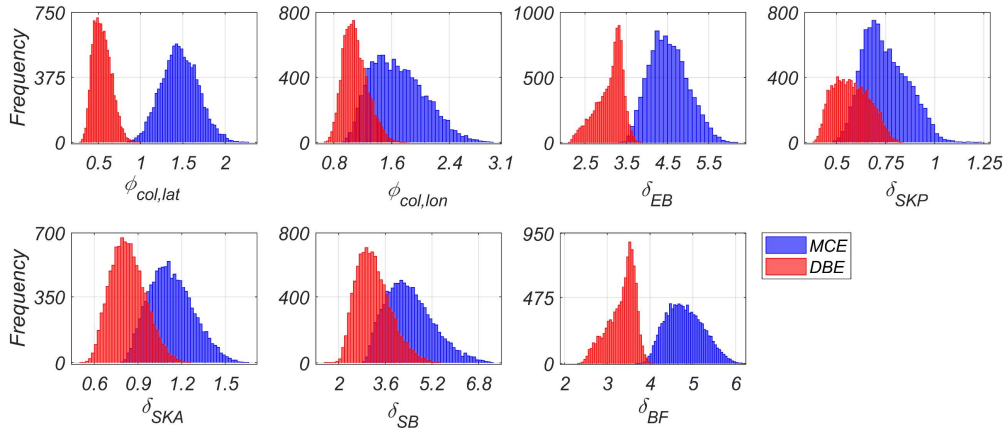


Figure 4.4-9. Histograms of EDPs for all components at both hazard levels.

4.4.3.6 MODEL FITTING

Seven GP models were defined independently, one for each EDP (i.e., $\phi_{col,lat}, \phi_{col,lon}, \delta_{EB}, \delta_{SKP}, \delta_{SKA}, \delta_{SB}, \delta_{BF}$). The models were defined using a combination of Matern kernel with $\nu = 3/2$ and ARD and a signal noise term (see Eq. (3.4.12)) as this kernel provides more flexibility than the squared exponential in terms of modeling the less-smooth data from AN than the RCB case. This kernel was selected following a series of preliminary tests where multiple kernel families were considered, and Matérn with $\nu = 3/2$ provided the best fit for the dataset.

The data was split into training and testing sets (i.e., $\mathbf{Y_{tr}}, \mathbf{Y_{te}}$), and the number of training points was determined following a convergence analysis which is shown in section 4.4.3.7 of this paper. A total of 2,000 points were randomly selected for training, while GP predictions were tested for the remaining 8,000 points. The goodness of fit of each SM was measured through R^2, $RRMSE$, and MAE error metrics, taken on the testing data.

The resulting GP SMs for each EDP are shown in Figure 4.4-10 and Figure 4.4-11 for the DBE and MCE hazard levels, respectively. Figure 4.4-10 and Figure 4.4-11 show the regressions for all EDPs at both hazard levels. At both hazard levels, R^2 values of over 0.95 were achieved for all testing data.

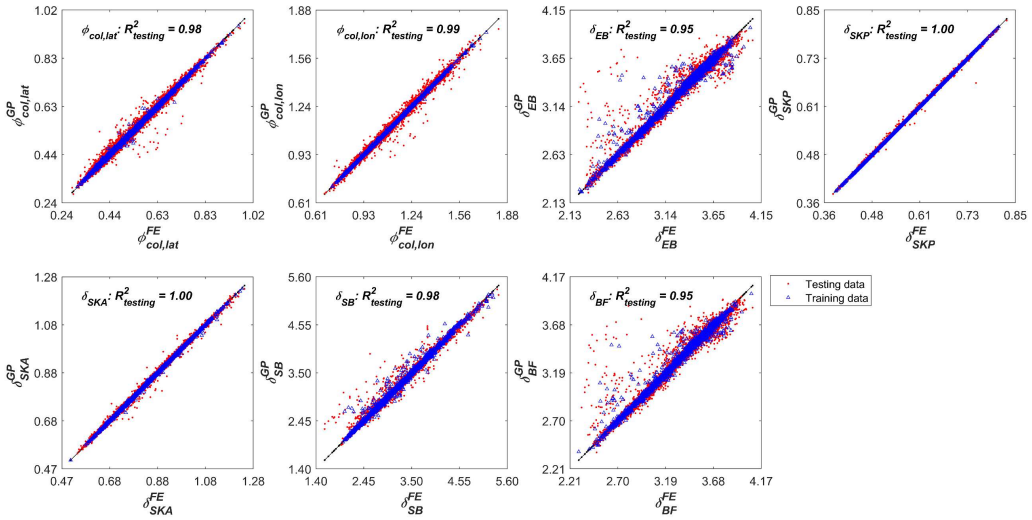


Figure 4.4-10. GP Regressions for all components at DBE level.

Error metrics for all GP regressions are summarized in Table 4.4-6 and Table 4.4-7 for DBE and MCE, respectively. With values of R^2 over 0.95, $RRMSE$ between 0.38% and 3.46%, and MAE between 0.0041 and 0.0356 for all models, the generated GP SMs proved to be able to accurately predict 8,000 new samples covering all regions of the input parameter space. Therefore, they could be reliably used as a substitute for the original FE model for further sampling-based analysis of the AN bridge model.

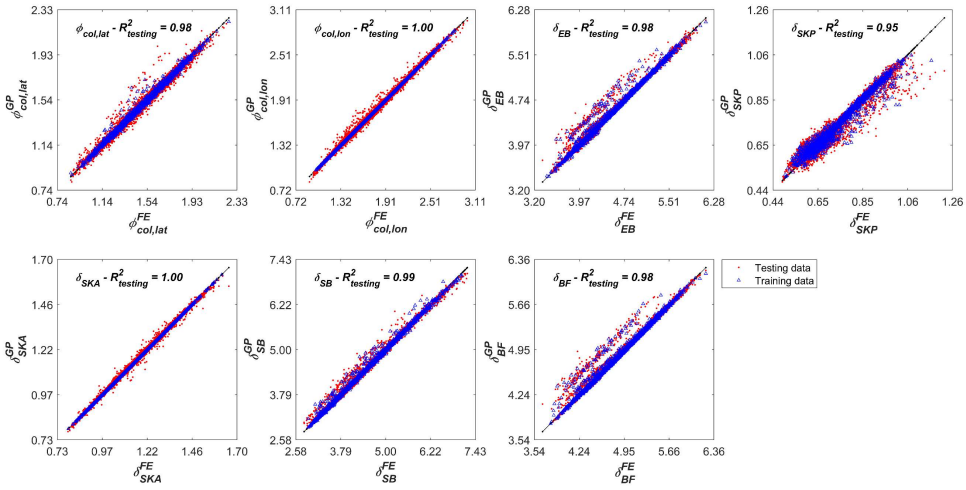


Figure 4.4-11. GP Regressions for all components at MCE level.

Table 4.4-6. Error metrics for the GP fit on testing data of the Aguila Norte Bridge at DBE level.

	$\phi_{col,lat}$	$\phi_{col,lon}$	δ_{EB}	δ_{SKP}	δ_{SKA}	δ_{SB}	δ_{BF}

R^2	0.99	0.99	0.95	0.99	0.99	0.98	0.95
$RRMSE$ (%)	2.35	1.48	2.18	0.38	0.76	2.23	2.08
MAE	0.0085	0.0101	0.0322	0.0014	0.0041	0.0356	0.0311

Table 4.4-7. Error metrics for the GP fit on testing data of the Aguila Norte Bridge at MCE level.

	$\phi_{col,lat}$	$\phi_{col,lon}$	δ_{EB}	δ_{SKP}	δ_{SKA}	δ_{SB}	δ_{BF}
R^2	0.98	1.00	0.98	0.95	1.00	0.99	0.98
$RRMSE$ (%)	2.02	1.46	1.34	3.46	0.70	1.43	1.31
MAE	0.0196	0.0169	0.0297	0.0158	0.0051	0.0340	0.0299

4.4.3.7 CONVERGENCE ANALYSIS

To select an appropriate number of training data points for the GP models presented in section 4.4.3.6, the convergence of the R^2 metric was assessed. To this end, the R^2 between the testing dataset and model predictions was computed for several number of training data points. The R^2 values obtained for each case and for each EDP are summarized in Table 4.4-8 (DBE) and Table 4.4-9 (MCE). No significant variation was observed beyond 1,000 points. At 2,000 points, all models achieved a value of R^2 = 0.95 or higher. Based on this, 2,000 training points were considered for each GP model. It is noted that even for only 200 training points, R^2 values greater than 0.9 were achieved for all the GPs.

Table 4.4-8. Convergence analysis for DBE surrogate model.

	R^2 of testing dataset						
Training points	$\phi_{Col,lat}$	$\phi_{Col,lon}$	δ_{EB}	δ_{SKP}	δ_{SKA}	δ_{BF}	δ_{SB}
50	0.17	0.57	0.62	0.97	0.89	0.6	0.58
100	0.89	0.94	0.89	0.98	0.96	0.92	0.89
150	0.92	0.98	0.93	0.99	0.98	0.97	0.92
200	0.95	0.98	0.91	1.00	0.99	0.97	0.9
250	0.96	0.98	0.92	1.00	0.99	0.96	0.9
300	0.97	0.98	0.93	1.00	0.99	0.98	0.92

400	0.97	0.98	0.94	1.00	0.99	0.98	0.91
500	0.98	0.98	0.94	1.00	0.99	0.98	0.92
750	0.98	0.99	0.93	1.00	1.00	0.98	0.93
1000	0.98	0.99	0.95	1.00	1.00	0.98	0.95
2000	0.98	0.99	0.95	1.00	1.00	0.98	0.95

Table 4.4-9. Convergence analysis for MCE surrogate model.

	R^2 of testing dataset						
Training points	$\phi_{col,lat}$	$\phi_{col,lon}$	δ_{EB}	δ_{SKP}	δ_{SKA}	δ_{BF}	δ_{SB}
50	0.85	0.94	0.96	0.8	0.97	0.97	0.95
100	0.92	0.99	0.97	0.88	0.99	0.99	0.97
150	0.92	0.99	0.97	0.89	0.99	0.98	0.96
200	0.95	0.99	0.97	0.92	0.99	0.98	0.97
250	0.96	0.99	0.98	0.92	0.99	0.99	0.96
300	0.96	0.99	0.97	0.92	0.99	0.99	0.96
400	0.97	0.99	0.98	0.93	1.00	0.99	0.98
500	0.97	0.99	0.98	0.94	1.00	0.99	0.98
750	0.97	1.00	0.98	0.94	1.00	0.99	0.98
1000	0.97	1.00	0.98	0.94	1.00	0.99	0.98
2000	0.98	1.00	0.98	0.95	1.00	0.99	0.98

4.4.3.8 LSA COMPARISON

Similarly to case study 1, LSA was conducted on 50,000 new samples from the GP surrogates for the AN bridge, as well as on the 10,000 FE model samples. The resulting ρ^{GP} and ρ^{FE} are shown in Figure 4.4-12, where a close match between them can be observed for all EDPs. For brevity, only the LSA results from the MCE case are shown herein, but consistent results were observed for DBE. Thus, it was deemed appropriate to replace the FE model with the GP surrogates for GSA in this case study.

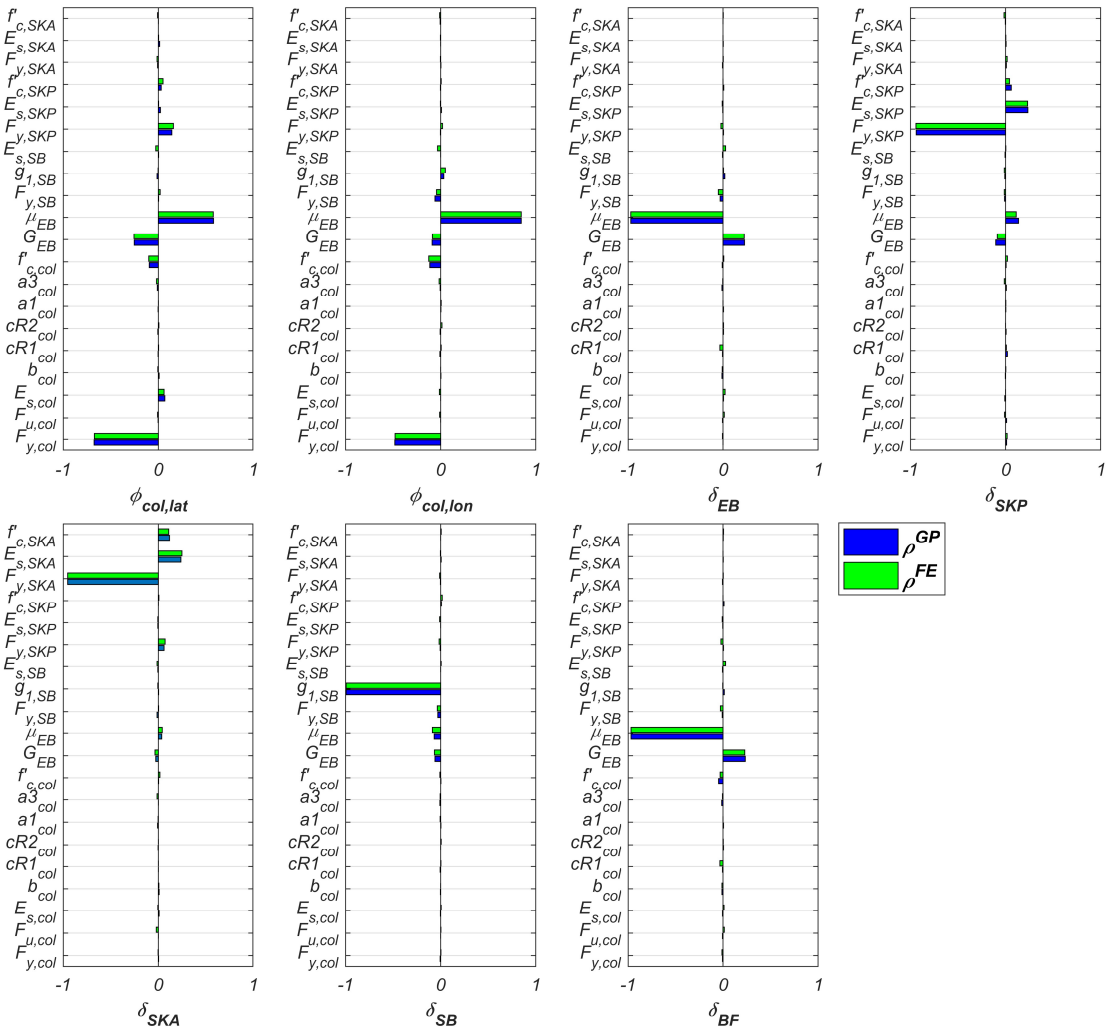
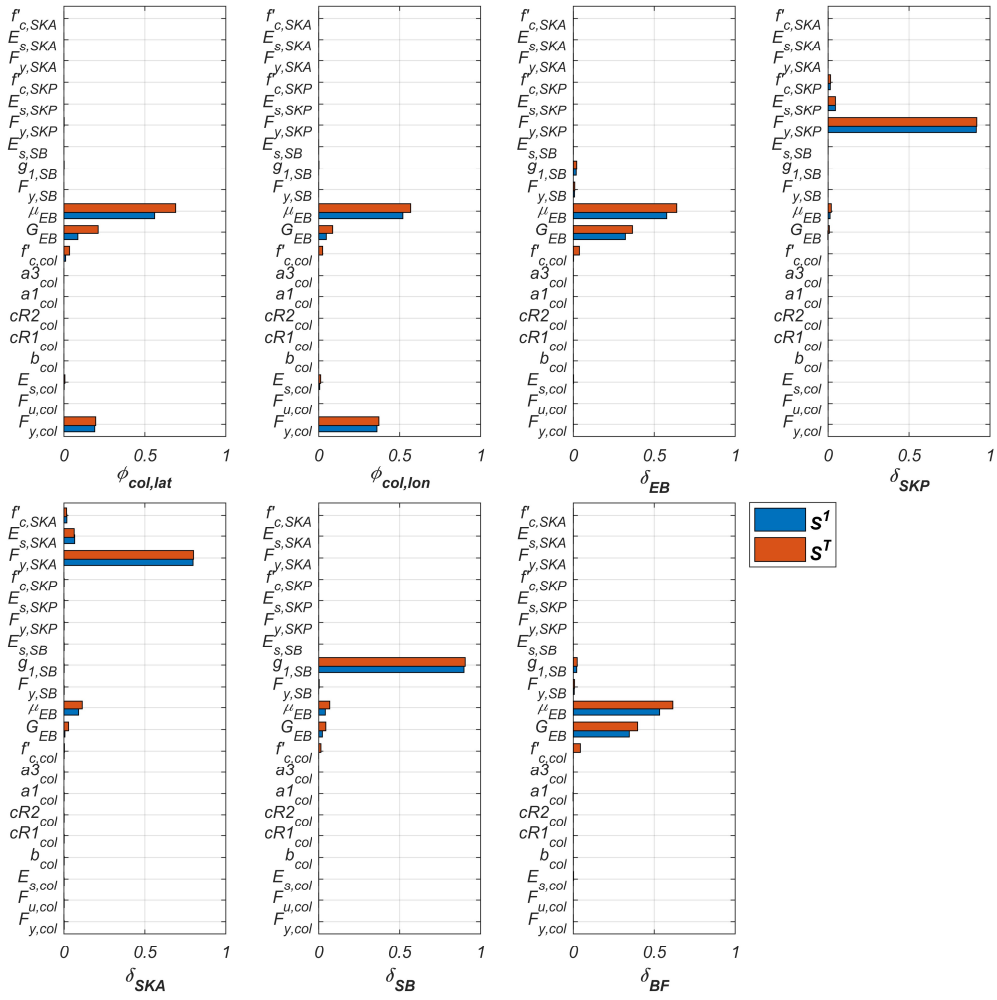


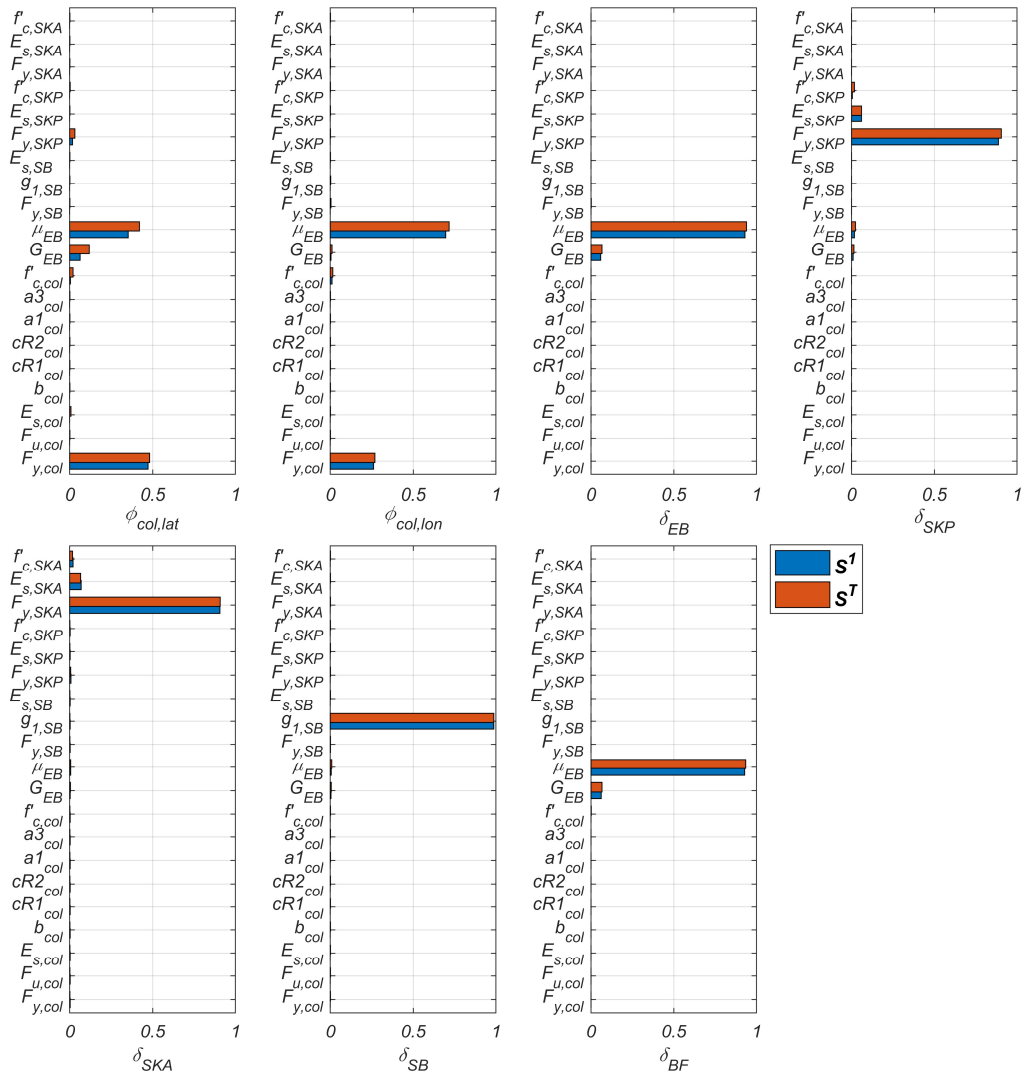
Figure 4.4-12. Comparison of local sensitivity indices from GP and FE models.

4.4.3.9 GP-BASED GSA

GSA of the seven defined EDPs to all input parameters were computed through S_1 and S_T indices, using the T estimator from Eq. (3.1.10). A total of N = 100,000 samples were used, which for the case of d = 20 parameters, resulted in a total of 2,200,000 model evaluations. The number of samples was chosen following a convergence analysis of the indices. Each parameter distribution was defined according to those in Table 4.2-3, Table 4.2-4, and Table 4.2-5. The process was repeated for both hazard levels. The computed indices for each of the AN EDPs are shown in Figure 4.4-13 for DBE and MCE hazard levels.


f'c,SKA
Es,SKA
Fy,SKA
f'c,SKP
Es,SKP
Fy,SKP
Es,SB
g1,SB
Fy,SB
μEB
GEB
f'c,col
a3col
a1col
cR2col
cR1col
bcol
Es,col
Fu,col
Fy,col
0
0.5
1
ϕcol,lat
ϕcol,lon
δEB
δSKP
δSKA
δSB
δBF
S1
ST

(a) DBE



(b) MCE

Figure 4.4-13. Sobol indices from surrogate-based GSA. (a) DBE, (b) MCE.

It can be seen in Figure 4.4-13 that each EDP is mostly governed by its own constitutive model parameters. This is a consequence of the response normalization, by which the FE model response is directly related to the sampling of each component's LS. Nonetheless, the impact of EB constitutive model parameters on global behavior was evident. This is evident from the curvature ductility demands on pier columns, which showed significant dependence on μ_{EB} and G_{EB} despite the normalization. This could be explained by the bridge's dynamic, by which the rigid deck imposes displacement on the piers thus generating strong interactions between EB and columns, especially at MCE hazard. It should be noted that besides the

influence of EB parameters, ϕ_{col} in both directions was mostly governed by $F_{y,col}$ despite the uncertainty in $f'_{c,col}$ and other reinforcing steel parameters. While previous results obtained by the author (Birrell, Astroza, Carreño et al., 2021) showed that *Steel02* parameters controlling the nonlinear behavior of the model, such as $cR1_{col}$ and $a1_{col}$ were influential at the material level, that was not the case for the AN bridge model. This might be due to Chilean design practices, which seek to avoid damage and nonlinear behavior in pier columns, even at the MCE level. Such practice is reflected in the column demands obtained from FE simulations, lower than 1.89 for DBE and 3.1 for MCE, respectively, which would not reach the second limit state as defined by (Stefanidou & Kappos, 2017), hence not incurring in significant nonlinear behavior for reinforcing steel.

Other components, including EB, SKP, SKA, and SB, showed significantly higher demands compared to pier columns, in relation to their respective LS1, which translated into a more prevalent influence from their respective nonlinear constitutive parameters. As such, the consideration of uncertainty in these should be maintained for future fragility analysis. For BF, the sensitivity was expectedly controlled by EB parameters, due to the normalization and the lack of sampling of BF parameters, although at the DBE hazard level the influence of column stiffness on the longitudinal pounding can be seen through the influence of $f'_{c,col}$ on δ_{BF}.

Finally, the GP surrogate developed in this study enabled the otherwise cost-prohibitive generation of 2,200,000 model evaluations for the execution of GSA presented in this section. Due to the SM-based GSA, it was possible to gain insight on the composition of seismic demand uncertainty and select relevant parameters for future analysis on the AN bridge or other large and costly FE models of civil structures.

4.4.4 APPLICATION II: GP-BASED APPROACH FOR UNCERTAINTY QUANTIFICATION AND PROPAGATION IN PBEE

The proposed methodology from section 3.7 is illustrated and validated through the study of the AN bridge. In the first stage of validation, the structure-level FE model is defined, including probabilistic characterizations of its nonlinear components' constitutive models, which were previously calibrated using experimental data in a Bayesian framework (Birrell, Astroza, Carreño et al., 2021; Birrell et al., 2024; Pinto et al., 2023). Subsequently, a suite of over 200 unscaled GMs from subduction earthquakes in Chile and Japan is selected. The suite covers seismic hazard levels with return periods (T_r) between 30 and 5000 years, consistent with PSHA of the bridge's site. A database of 10,000 FE model simulations is generated through nonlinear time-history analysis (NLTHA), which is used to train and test GP SMs for seven engineering demand parameters (EDPs) of interest. A cloud-based fragility analysis (Jalayer et al., 2017) is then conducted, incorporating uncertainties due to model parameters, LS definitions, and component capacity at each LS, as well as epistemic uncertainty induced by the inaccuracy of GP surrogates. The GP-based approach for PSFA is validated by comparing the fragility functions computed from further GP sampling to those obtained from the initial FE model simulations. As part of PSFA, the MRE of LSs for each EDP is discussed, by integrating fragility functions with the hazard. This step provides insight into what the most critical components are in terms of damage recurrence, as well as opening a discussion for component-level LS definitions. Finally, the composition of uncertainty in seismic fragility functions is studied through GSA using Monte Carlo (MC)-based Sobol indices (Sobol, 2001), enabled by the computationally efficient GP surrogates for further sampling. Through GSA, meaningful insights are obtained into the impact of model parameter uncertainty on the bridge's EDPs at each hazard level. By applying the proposed methodology, future decisions on data collection for model calibration, LS definitions, inspection planning, and other critical activities, could be better informed.

4.4.4.3 FE MODEL SIMULATION

A total of 10,550 (i.e., 50×211) unique combinations of model parameters (i.e., 50 samples from the PDFs in Table 4.4-1 were drawn and checked for convergence to the target distributions) and GMs were configured and NLTHA was performed, considering the GM

suite selected for application II, presented in section 4.4.2. As a result, 10,550 data points for each of the seven model responses were generated, which are summarized as histograms in Figure 4.4-14. A slight degree of right skewness was observed for all responses, consistent with having fewer earthquakes available at $S_a(0.7)>1g$. A wide range of values was covered for each response (e.g., 0.01m to about 0.22m for Δ_{EB} and 0.01m to about 0.16m for Δ_{SKA}). The responses followed non-smooth, multimodal distributions. Thus, this data set serves to extensively test the ability of the SMs developed later, in section 4.4.4.4, to predict model responses at a wide range of hazard levels over the complete domain of each parameter.

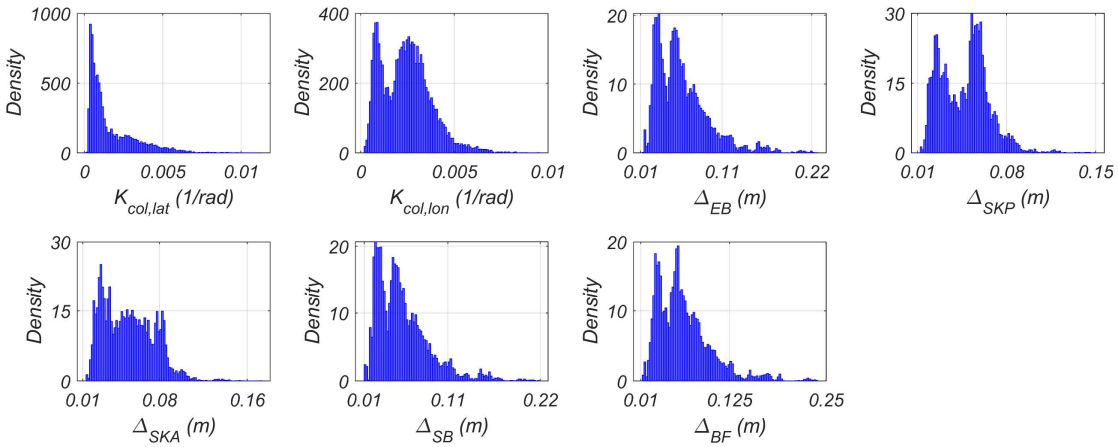


Figure 4.4-14. Histograms of FE model responses.

DATA PREPARATION FOR GP TRAINING

To develop GP surrogates for each model response, the 10,550 data were split into training (i.e., $\{\mathbf{X}_{tr}, \mathbf{Y}_{tr}^{FE}\}$) and testing (i.e., $\{\mathbf{X}_{te}, \mathbf{Y}_{te}^{FE}\}$) datasets. All model parameters from Table 4.4-1 (twenty in total) and 50 spectral accelerations taken from each GM's RotD50 (i.e., Sa(T) at T = {0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.10, 0.13, 0.16, 0.19, 0.22, 0.25, 0.28, 0.34, 0.40, 0.42, 0.44, 0.46, 0.48, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95, 1.00, 1.10, 1.20, 1.30, 1.40, 1.50, 1.60, 1.70, 1.80, 1.90, 2.00, 2.20, 2.40, 2.50, 3.00, 3.50, 4.00, 5.00} s) were considered as features for the GPs. Thus $\mathbf{X}_{tr} \in \mathbb{R}^{n_{tr}\times 70}$ and $\mathbf{X}_{te} \in \mathbb{R}^{n_{te}\times 70}$, respectively.

The training set was selected by a stratified sampling strategy, with the following algorithm:

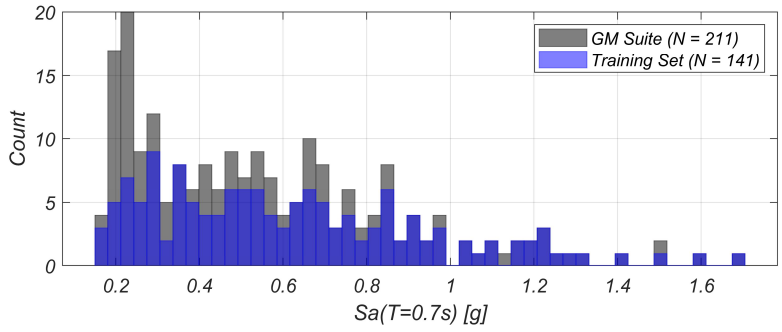
1. For each bin of 0.1g from $S_a(0.7)$ values, randomly select the maximum number between 25 GMs and the number of GMs in the bin. This step selected 141 GM out of the 211.

2. Select parameter samples as *round(1:1.5:50)*, i.e., {1, 3, 4, 6, 7, 9, …, 49}, for a total of 33 out of 50 samples.

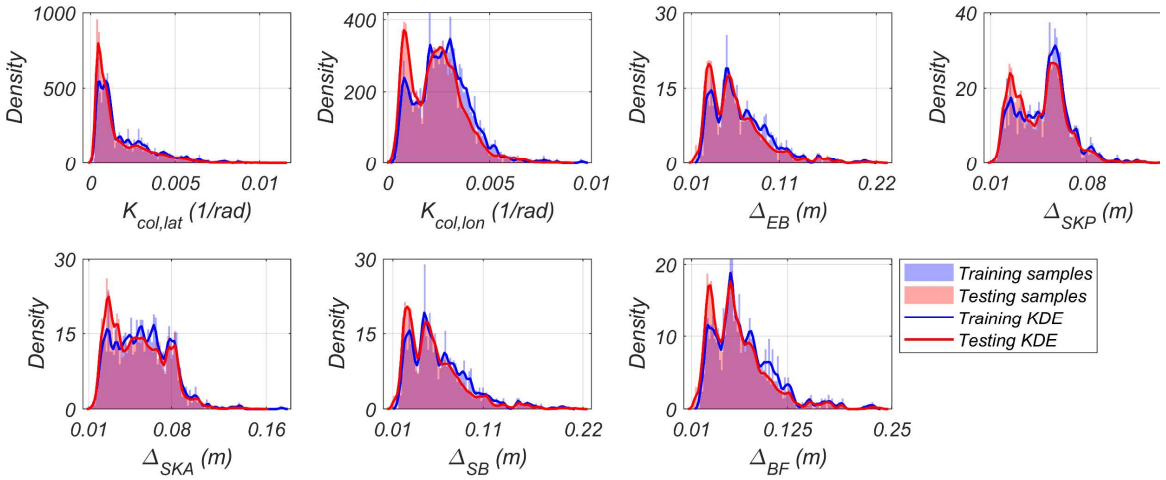
3. From all combinations of the selected GMs and parameter samples, randomly select 20% of the combinations.

Following this selection method, a total of n_{tr} = 987 FE model samples were selected for training, thus n_{te} = 9,563. The number of parameter samples, bin width, and number of GMs per bin were selected iteratively to ensure that every hazard level was appropriately represented and the distribution of $\mathbf{Y}_{tr}$ was consistent with the distribution of $\mathbf{Y}_{te}$. Additionally, three GMs initially selected at $S_a(0.7) > 0.98$g were manually excluded for later validation of the SMs' interpolation ability at higher *IM* values, where less GM training data were available. The stratified approach to selecting training and testing sets ensured that the SMs could be trained with fewer data points than a traditional data split scheme (e.g., 80/20), while maintaining an adequate representation of every region in the data PDFs.

Selected GMs are shown in Figure 4.4-15a, while the kernel densities of $\mathbf{Y}_{tr}$ and $\mathbf{Y}_{te}$ are shown in Figure 4.4-15b. The selected GMs demonstrate an even distribution throughout *IM* values, preventing GP bias towards weaker earthquakes. Consequently, the distribution of $\mathbf{Y}_{tr}$ is representative of the whole sample, despite containing less than 10% of the total data.



(a)



(b)

Figure 4.4-15. Stratified selection of training data. (a) GMs (b) Model responses.

4.4.4.4 GP SURROGATE MODELS

Independent GP surrogates were trained using Matérn 3/2 kernels with ARD, as per Eq. (3.4.12), for the seven responses considered. Consequently, the prior for $\boldsymbol{\theta}_{GP} = [l_1, \dots, l_{70}, \sigma_f, \sigma_n]$ was defined as described in section 3.4 for each response. $\boldsymbol{\theta}_{GP}$ was then optimized through minimization of the log-marginal likelihood in Eq. (3.4.5) using the L-BFGS (Liu & Nocedal, 1989) algorithm. Finally, the accuracy of the trained SMs was assessed through the R^2 (see Eq. (3.4.13)) of GP predictions (i.e., $\mathbf{Y}_{te}^{GP}$, obtained for each response from Eq. (3.4.7)) with respect to $\mathbf{Y}_{te}^{FE}$).

Figure 4.4-16 shows GP predictions on the training (in blue) and the testing (in red) data for each model response on the *y*-axis, plotted against their respective FE model data on the *x*-axis. The identity line (i.e., $y = x$) indicates a perfect fit. Predictions on the training data achieved near-perfect fits with $R^2 > 0.995$ in all cases. While such accuracy on training data could suggest overfitting, this outcome can be expected when using ARD kernels with a large number of features due to the kernel's flexibility to fit non-smooth data.

When tested against $\mathbf{Y}_{te}^{FE}$, which includes GMs and parameter samples previously unseen during training, the SMs retained an excellent performance, achieving $R^2 > 0.9$ in all cases. Moreover, their ability to generalize and interpolate, even at higher *IM* values where fewer

training data points were available, was also confirmed. This result established that the GP surrogates can be used to reliably generate further sampling of model responses for later fragility and sensitivity analyses.

Additionally, GP prediction uncertainty (see Eq. (3.4.9)) is shown in blue and red patches, respectively, in Figure 4.4-16. Certain zones with higher uncertainty can be identified in the figure, particularly for Δ_{EB}, Δ_{SB}, and Δ_{BF}. This uncertainty is taken into consideration in the following sections by including GP prediction uncertainty as part of the total demand uncertainty in GP PSDMs.

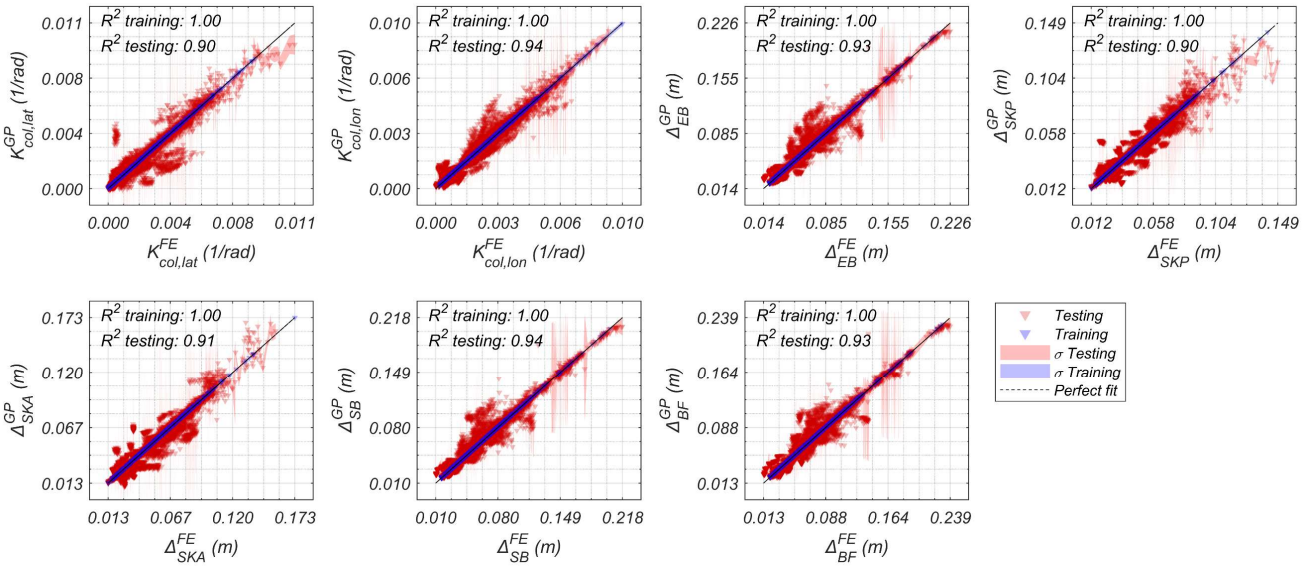


Figure 4.4-16. GP regressions for all model responses.

4.4.4.5 CLOUD-BASED FRAGILITY ANALYSIS

In this section, surrogate-based fragility estimates are obtained by drawing 1,000 new parameter samples (i.e., $\boldsymbol{\Theta}_{fr}$) and generating predictions for each GM (i.e., 211,000 total GP model evaluations $\{\mathbf{X}_{fr}, \mathbf{Y}^{GP}_{fr}\}$).

The procedure to obtain conditional fragility functions as described in section 3.5 using GP surrogates is as follows:

1. Generate GP model predictions for the new set of samples $\{\mathbf{X}_{fr}, \mathbf{Y}^{GP}_{fr}\}$.

2. Evaluate the capacity of components at each LS $\mu_{c,LS}(\boldsymbol{\Theta}_{fr}), \beta_{c,LS}(\boldsymbol{\Theta}_{fr})$ by sampling their respective constitutive model, considering LS definitions from section 4.4.4.6.

3. Compute EDP samples for each component from EDP and capacity definitions in section 4.4.4.7.

4. Compute conditional PSDMs using Eq. (3.5.2), including epistemic uncertainty from Eq. (3.5.3).

5. Compute conditional fragility functions using Eq. (3.5.4).

The resulting fragility functions are then compared to those obtained from the original 10,550 FE model samples for validation. The procedure is analogous for $\{\boldsymbol{\Theta}, \mathbf{Y}^{FE}\}$ by computing component capacities and EDPs, conditional PSDM, and fragility functions, although only the original 50 parameter samples are considered.

4.4.4.6 LS DEFINITIONS

For column curvature, EB shear deformation, and BF deformation, LS definitions proposed by Stefanidou and Kappos (Stefanidou & Kappos, 2017; Stefanidou et al., 2022) were adopted. For SK and SB, no LS definitions are available in published literature to the best of the authors' knowledge. Constitutive model-based LS were defined and used for those components. All LS definitions are summarized in Table 4.4-10.

Table 4.4-10. LS definitions for each component.

Limit state	Expression	Description
	Columns (section curvature)	
LS1	$K_{col} = K_y$	Yield curvature
LS2	$K_{col} = \min\left(K{:}\ \varepsilon_c > 0.004, K{:}\ \varepsilon_s > 0.015\right)$	Spalling of cover concrete, reinforcing steel hardening
LS3	$K_{col} = \min\left(K{:}\ \varepsilon_c > 0.004 + 1.4\frac{\rho_w f_{yw}}{f'_{cc}}, K{:}\ \varepsilon_s > 0.06\right)$	First hoop fracture, concrete core crushing, reinforcing steel buckling
LS4	$K_{col} = \min\left(K{:}\ M < 0.9M_{max}, K{:}\ \varepsilon_s > 0.075\right)$	Loss of carrying capacity
K_y: yield curvature. ε_c: concrete strain. ε_s: steel strain. ρ_w, f_{yw}: reinforcement ratio and yield strength of transverse hoops. f'_{cc}: confined concrete strength. M, M_{max}: moment and maximum moment from sectional analysis.		
	Elastomeric Bearings (shear deformation)	
LS1	$\Delta_{EB} = 0.2H_{EB}$	Initiation of nonlinear behavior, yield of the rubber

LS2	$\Delta_{EB} = H_{EB}$	Yield of steel plates
LS3	$\Delta_{EB} = 2H_{EB}$	Uplifting, delamination, unbonding between rubber and steel plates
LS4	$\Delta_{EB} = 3H_{EB}$	Rotation, unseating, lift-off.
H_{EB}: height of the rubber.		
Shear Keys (lateral displacement)		
LS1	$\Delta_{SK} = D_{y,SK}$	Yield of reinforcing steel, concrete cracking
LS2	$\Delta_{SK} = D_{n,SK}$	Maximum capacity
LS3	$\Delta_{SK} = D_{c,SK}$	Loss of concrete contribution, residual capacity, loss of lateral deck restraining ability
$D_{y,SK}, D_{n,SK}, D_{c,SK}$: yield displacement, displacement at maximum capacity, ultimate displacement.		
Vertical Seismic Bars (longitudinal displacement)		
LS1	$\Delta_{SB} = D_{y,SB}$	Yield of seismic bars
LS2	$\Delta_{SB} = D_{u,SB}$	Maximum capacity, loss of vertical deck anchoring contribution
$D_{y,SB}, D_{u,SB}$: yield and ultimate displacement.		
Abutment Backfill (longitudinal displacement as a proportion of backwall height)		
LS1	$\Delta_{BF} = gap$	Gap closure
LS2	$\Delta_{BF} = 0.0093H_w$	First yield of abutment backfill soil
LS3	$\Delta_{BF} = 0.037H_w$	Major deformation of abutment backfill soil
LS4	$\Delta_{BF} = 0.081H_w$	Ultimate deformation of abutment backfill soil
gap: longitudinal gap between the deck and abutment. H_w: height of abutment backwall.		

Detailed descriptions of LS definitions for column, EB, and BF can be found in (Stefanidou & Kappos, 2017; Stefanidou et al., 2022). The four LSs defined for these components indicate slight, moderate, extensive, and collapse damage states (DSs), respectively.

For SK, the constitutive model shown in Figure 4.4-1 and defined in (Goel & Chopra, 2008) is defined by three points, namely $D_{y,SK}, D_{n,SK}$, and $D_{u,SK}$, which indicate the yield of steel reinforcement, maximum capacity, and fracture displacements, respectively. The first two points were used as LS, as they are directly defined by DSs. The third LS was defined as an intermediate point between $D_{n,SK}$ and $D_{u,SK}$, based on the dotted purple line in Figure 4.4-1, which comes from the original version of the model proposed by Megally et al. (Megally et al., 2002). The point was defined as $D_{c,SK}$, given by Eq. (4.4.1), at which the contribution of

concrete has completely degraded, and the SK is close to complete failure. SKs are designed to be sacrificial components that provide lateral restraint to the deck. They play a critical role in preventing the unseating of the deck during major earthquakes (Buckle et al., 2012). For this reason, complete failure should be avoided, hence the decision of defining SK failure at an earlier displacement than $D_{u,SK}$. The three LS defined for SK indicate moderate, extensive and collapse DSs, respectively.

$$D_{c,SK} = \sqrt{2} \times 0.006 \times (l_d + l_a) \times \frac{h + d}{s_e} \tag{4.4.1}$$

where l_d is the development length of the shear reinforcement in the SK, l_a is the length of the cracked section along the monolithic joint, h is the height of the SK, d is the width of the SK, and s_e is the reinforcing steel spacing.

For SB, the constitutive model shown in Figure 4.4-1 and defined in (Pinto et al., 2023) is defined by two displacement points: $D_{y,SB}$ and $D_{u,SB}$, which indicate displacements at the yield and ultimate strength, respectively. SBs have not been identified as a critical component for bridge performance, and their main role is to provide additional vertical deck anchoring. Thus, two LS were defined for SB, indicating slight and moderate damage, respectively.

4.4.4.7 EDP AND CAPACITY DEFINITIONS

EDPs were defined as normalized quantities by taking the ratio between model responses and the first LS of each component, as indicated in Table 4.4-11. Thus, column curvature ductility (i.e., $\phi_{col,lon}, \phi_{col,lat}$), EB shear deformation ductility (i.e., δ_{EB}), SK drift ductility (i.e., $\delta_{SKP}, \delta_{SKA}$), SB drift ductility (i.e., δ_{SB}), and BF drift (i.e., δ_{BF}) demands are considered for PSFA.

Table 4.4-11. EDP definitions for all components.

EDP	Columns	EB	SK	SB	BF
Definition	Curvature ductility	Shear deformation ductility	Drift ductility	Drift ductility	Drift

EDP	$\phi = \frac{K_{col}}{K_{y,col}}$	$\delta_{EB} = \frac{\Delta_{EB}}{0.2H_{EB}}$	$\delta_{SK} = \frac{\Delta_{SK}}{D_{y,SK}}$	$\delta_{SB} = \frac{\Delta_{SB}}{D_{y,SB}}$	$\delta_{BF} = \frac{\Delta_{BF}}{H_w}$

Subsequently, component capacities at each LS were sampled by computing the corresponding points from the respective constitutive models for each sample in $\boldsymbol{\Theta}_{fr}$ and normalizing them by LS1. Then, their logarithmic means and standard deviations were computed to obtain $\mu_{c,LS}(\boldsymbol{\Theta}_{fr})$ and $\beta_{c,LS}(\boldsymbol{\Theta}_{fr})$. The corresponding values are summarized in Table 4.4-12, along with LS definition uncertainties (i.e., β_{LS}). Note that for EB and BF, $\beta_c = 0$ in all cases because their LS were defined in terms of geometrical properties, which remain constant. Conversely, for SK and SB, $\beta_{LS} = 0$ in all cases, as only the LS definitions from this study were considered.

GP predictions were transformed into the EDP and log-EDP spaces for the upcoming analyses. These data transformations can introduce distortion in regression fits as well as data distributions. The accuracy of GP surrogates in the transformed spaces was verified by re-computing the R^2 indices, which are summarized in Table 4.4-13. The models retained excellent fits in the log-EDP space, with R^2 over 0.87 for all cases.

Table 4.4-12. Capacity PDF for all components.

	ϕ_{col}			δ_{EB}			δ_{SKP}			δ_{SKA}			δ_{SB}			δ_{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

Table 4.4-13. R2 fits on responses, EDP, and log-EDP spaces.

	Col, lat	Col, lon	EB	SKP	SKA	SB	BF
R^2 response	0.90	0.94	0.93	0.90	0.91	0.94	0.93
R^2 EDP	0.90	0.94	0.93	0.91	0.91	0.94	0.93
R^2 log-EDP	0.89	0.89	0.91	0.89	0.87	0.92	0.90

4.4.4.8 PSDM AND CONDITIONAL FRAGILITY ANALYSIS

A three-parameter linear PSDM in the logarithmic space was fit by linear regression for each EDP, conditioned on $\boldsymbol{\Theta}_{fr}$, using Eq. (3.5.2). Figure 4.4-17 shows each of the 1,000 samples in blue, while the mean PSDM and the 95% CI are shown in solid and dash black lines,

respectively. FE model responses are shown as grey dots for reference. The simplification of a constant $\beta_{GP}|\mathbf{X}$ by taking the median value can be seen from the constant slope of the CI. The effect of $\beta_{c,LS} = 0$ for EB and BF can also be noted in the figure, where the variability in $\boldsymbol{\Theta}_{fr}$ did not induce noticeable variance in the EDPs. Distribution statistics for $a(\boldsymbol{\Theta}_{fr})$, $b(\boldsymbol{\Theta}_{fr})$, and $\beta_{D|\boldsymbol{\Theta}_{fr}}$ are summarized in Table 4.4-14.

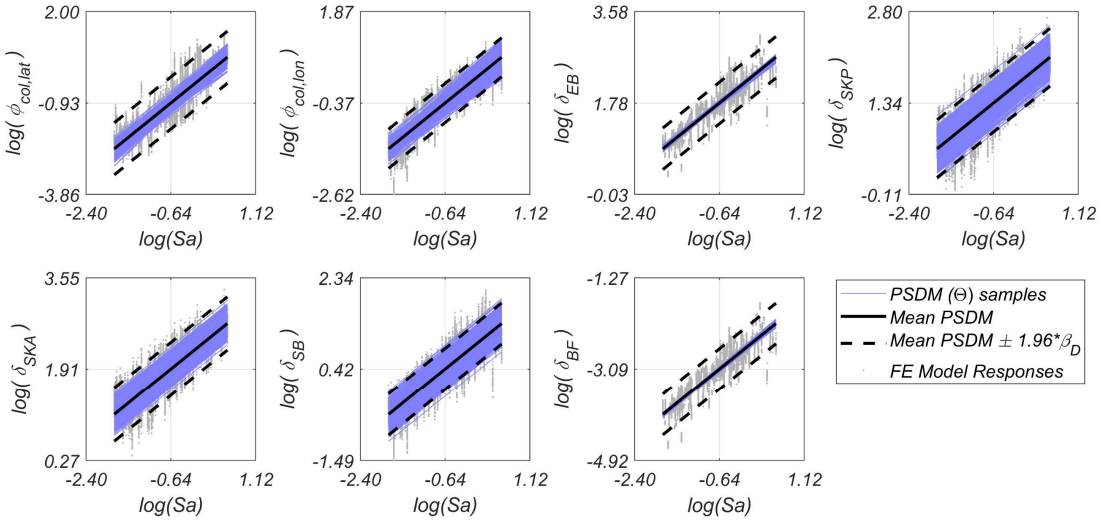


Figure 4.4-17. Sampled PSDM.

Table 4.4-14. PSDM Parameter statistics from FE and sampled GP models.

	PSDM parameter	$\phi_{col,lat}$	$\phi_{col,lon}$	δ_{EB}	δ_{SKP}	δ_{SKA}	δ_{SB}	δ_{BF}
FE	$\mu(a)$	-0.053	0.268	2.252	1.766	2.363	0.933	-2.605
	$\sigma(a)$	0.130	0.147	0.021	0.148	0.129	0.160	0.015
	$\mu(b)$	1.300	0.972	0.772	0.636	0.695	0.832	0.785
	$\sigma(b)$	0.068	0.071	0.010	0.011	0.008	0.013	0.006
	$\mu(\beta_D)$	0.397	0.263	0.222	0.232	0.239	0.240	0.224
	$\sigma(\beta_D)$	0.014	0.014	0.006	0.006	0.005	0.006	0.005
GP	$\mu(a)$	-0.129	0.242	2.268	1.743	2.354	0.945	-2.592
	$\sigma(a)$	0.132	0.146	0.023	0.147	0.131	0.166	0.020
	$\mu(b)$	1.252	0.959	0.771	0.623	0.700	0.818	0.780
	$\sigma(b)$	0.077	0.074	0.013	0.011	0.009	0.015	0.009
	$\mu(\beta_D)$	0.425	0.244	0.209	0.235	0.244	0.220	0.207
	$\sigma(\beta_D)$	0.026	0.015	0.009	0.005	0.009	0.010	0.008

For each sample, conditional fragility functions were computed using Eq. (3.5.4), which are shown in Figure 4.4-18. The mean and 95% CI curves computed from the FE model considering the original 50 parameter samples, are plotted in blue, while the mean and 95%

CI curves computed from the GP surrogates, considering 1,000 new parameter samples, are plotted in red. Individual samples from GP surrogates are shown in grey.

Fragility function estimates obtained from the FE and GP models are consistent. While FE fragility functions only considered 50 parameter samples, the distribution of those samples was already representative of all parameter PDF from Table 4.4-1. However, the benefit of using GPs for sampling is apparent, as 1,000 new samples could be reliably drawn covering PDFs more thoroughly. Specifically, this result validated the idea of relying on GP surrogates for further sampling of the model for GSA.

With respect to the fragility functions themselves, it was evident that parameter uncertainty had a considerable impact on fragility estimates, both due to uncertain demand and capacity measures. As an extreme example, a 95% CI of about 0.2 to 0.9 probability of surpassing LS2 for SKP at higher levels of *IM* was estimated. While most EDPs showed less variability, the uncertainty in fragility estimates was noticeable. Therefore, it is crucial to account for parameter induced uncertainty in PBEE, as its impact on practical decisions should not be ignored.

As an additional observation, the fragility functions for LS1 and LS2 of BF were found to overlap at higher *IM*. This was due to a combination of a higher β_{LS} and the small difference between $\mu_{c,LS1}$ and $\mu_{c,LS2}$.

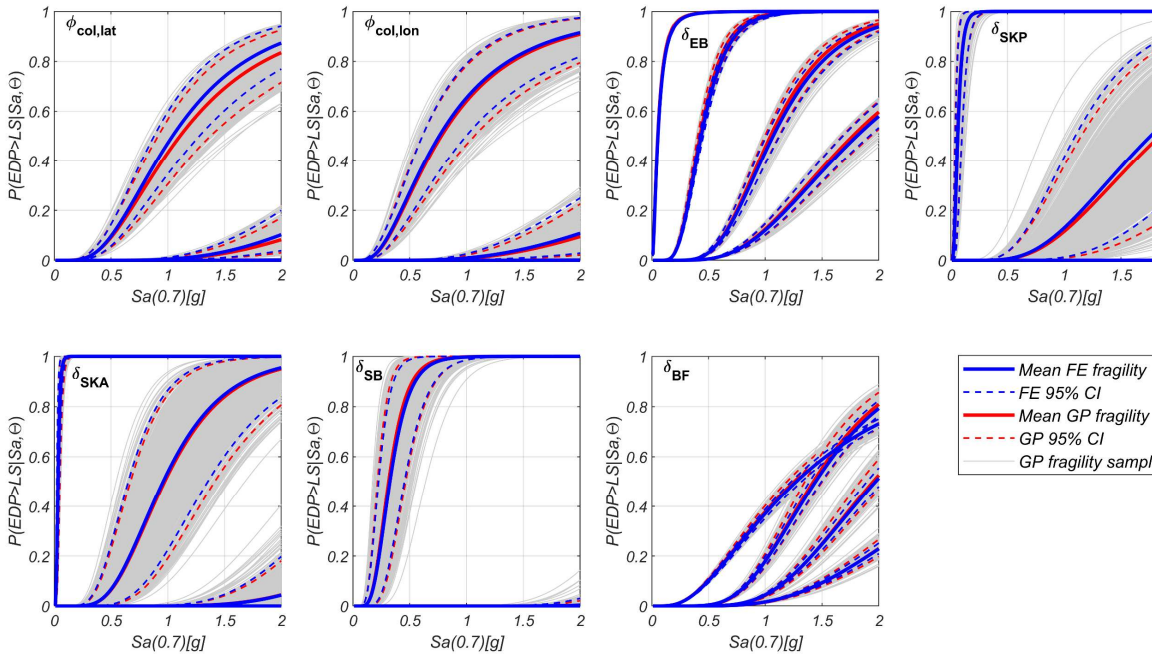


Figure 4.4-18. Comparison of fragility functions derived from FE models and sampled GP SMs.

Conditional fragility functions can be further translated into MRE for each LS by integrating the fragility and hazard curves using Eq. (3.5.5). The resulting functions represent a notion of a damage hazard, which depends on the estimation of hazard, GM selection used in structural simulation, LS and capacity definitions, and fragility curve estimates. In the absence of a damage-to-cost decision variable, these can provide a probabilistic measure of structural performance. Figure 4.4-19 shows the mean and 95% CI MRE curves for each EDP. By taking $T_r(\lambda_{LS}) = 1/\lambda(LS|\boldsymbol{\Theta}_{fr})$, the return period of exceedance of each LS is obtained, which are summarized in Table 4.4-15.

The difference in T_r among different EDPs is evident. For instance, columns would be expected to reach LS1 = K_y in the longitudinal direction with a $T_r = 636$ years, while EB and SK would be expected to surpass their yield-equivalent LS1 with a $T_r = 2$ years. Although the number $T_r = 2$ years seems unusually low, it should be noted that this was obtained due to the high fragility of EB and SK at LS1 (see fragility curves in Figure 4.4-18).

Specifically, LS1 for EB is defined as a 20% relative deformation, which corresponds to 10mm in the case of the AN bridge, where H_{EB} = 50mm. In the case of SK, LS1 was defined

herein as $D_{y,SK}$, which corresponds to a mean value of 11mm for SKP and 7mm for SKA. These levels of longitudinal and lateral deck displacement are easily surpassed during the subduction events considered in this study.

In contrast, local design practices in Chile aim for mostly elastic response of pier columns. This is consistent with the obtained MREs and fragility functions, which only showed a clear probability of exceedance for LS1 in a GM suite that covered up to T_r=5,000 years of seismic hazard. However, the global displacement required to reach LS1 in columns (i.e., K_y with a mean value of 0.003) is considerably larger than to reach LS1 for other components.

Despite the results shown in this section being subject to multiple sources of uncertainty and several assumptions about LS definitions and modeling (i.e., fixed base), the inconsistency in damage estimates among components is evident. To further illustrate the issue, SK are commonly designed as fuse elements, yet the probability of SKP exceeding their maximum capacity is lower than the probability of the bridge collapsing due to unseating of the deck, as indicated by EB performance. This highlights the need for future research to define performance-oriented local and global LS that consider Chilean bridge design practices and its seismic context, in order to ensure consistency among damage estimates for multiple interacting components.

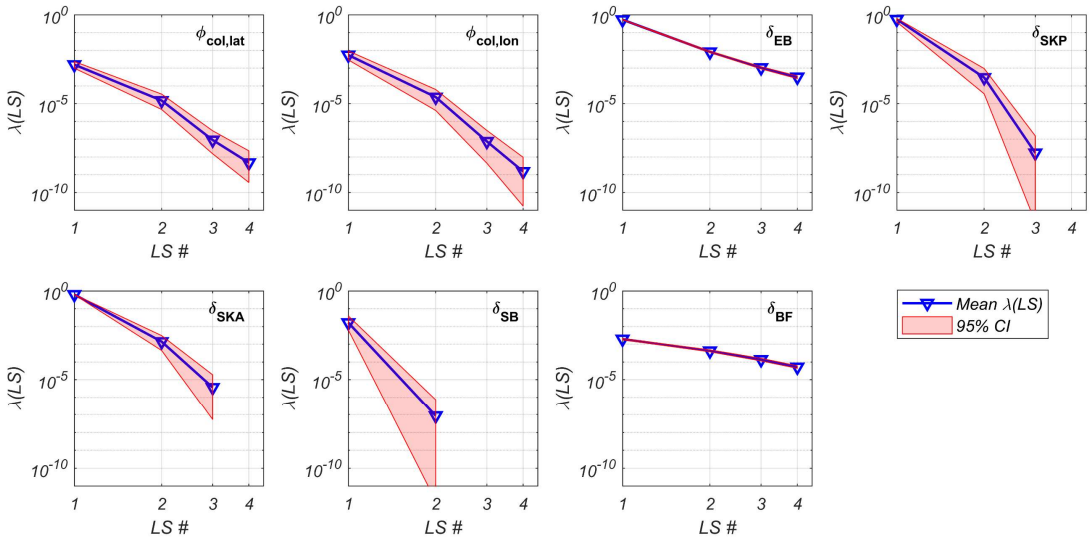


Figure 4.4-19. Mean rate of exceedance of damage states related to the EDPs considered.

Table 4.4-15. Mean return periods of exceedance of each LS by component.

$\mu(T_r)$	$\phi_{col,lat}$	$\phi_{col,lon}$	δ_{EB}	δ_{SKP}	δ_{SKA}	δ_{SB}	δ_{BF}
LS1	636	184	2	2	2	61	513
LS2	6.7×10^4	4.3×10^4	122	3,405	714	1.16×10^7	2,358
LS3	1.1×10^7	1.3×10^7	945	5.9×10^7	2.7×10^5	N/A	7,323
LS4	2.1×10^8	6.5×10^8	3,173	N/A	N/A	N/A	1.9×10^4

4.4.4.9 GLOBAL SENSITIVITY ANALYSIS

The importance of quantifying and propagating uncertainty throughout each stage of PBEE was evidenced by conditional fragility results. In this section, the composition of such uncertainty is investigated through GSA using total Sobol indices. Total Sobol indices were computed using Eq. (3.1.13) at 20 *IM* levels, to identify the most influential parameters at both lower and higher hazard levels. The analysis considered 5,000 samples following a convergence analysis of S^T, resulting in a total of 110,000 fragility function evaluations (i.e., 23.21×10^6 GP model evaluations, due to the 211 GMs). Evidently, this analysis was made possible by the reduced computational cost of GP sampling. The model sampling procedure described in section 4.4.4.5 was applied, consisting of the following steps: (i) draw parameter samples; (ii) generate GP predictions; (iii) compute capacities and EDP; and (iv) compute conditional PSDM and fragility functions. For each EDP, *IM*-dependent S^T were computed for the resulting fragility functions.

The resulting indices are shown in Figure 4.4-20. In each subplot, the indices obtained for each model parameter on every EDP are displayed. Some key results can be observed:

1. The uncertainty in column-related parameters primarily affected column responses. Specifically, $F_{y,col}$ and $E_{s,col}$ had a noticeable impact on the uncertainty of $\phi_{col,lon}$ and $\phi_{col,lat}$, although this impact was more pronounced at lower *IM* levels. Conversely, parameters governing the nonlinear response of reinforcing steel did not have a clear influence. This lack of influence can be attributed to the columns not sustaining significant demands due to their conservative design. As a consequence, the uncertainty in the nonlinear parameters of *Steel02* could be considered constant for future performance analysis of typical Chilean highway bridges.

2. Uncertainty in the fragility of EB could be mostly attributed to μ_{EB}, although G_{EB} had noticeable impact at lower *IM* levels. This reflects the impact of shear stiffness on the elastic behavior of EB, while the nonlinear behavior is primarily controlled by the friction coefficient.

3. Similarly to EB, the constitutive parameters of SB and SK showed that stiffness-related parameters influenced fragility at lower *IM* levels, while the parameters controlling nonlinear behavior became more significant at higher *IM* levels, when demands increased.

4. Most noticeably, EB parameters had the greatest impact on global longitudinal responses at higher *IM* levels, which is consistent with the dynamic behavior of the bridge and the capacity definitions, as discussed in section 4.4.4.8. The acceleration induced by the GM on the rigid deck is transferred to the bents by the shear deformation of the bearings. At low *IM* levels, where local elastic behaviors control the response, the impact is smaller. However, at higher *IM* levels, where the global nonlinear response becomes significant, the uncertainty induced by EB parameters on the response of columns and BF is greater.

5. Other cross-effects include the influence of $f'_{c,SKP}$ and $E_{s,SB}$ on the lateral response of columns at low *IM* levels, consistent with the lateral restraining function of SKP and the contribution of lateral stiffness by SB.

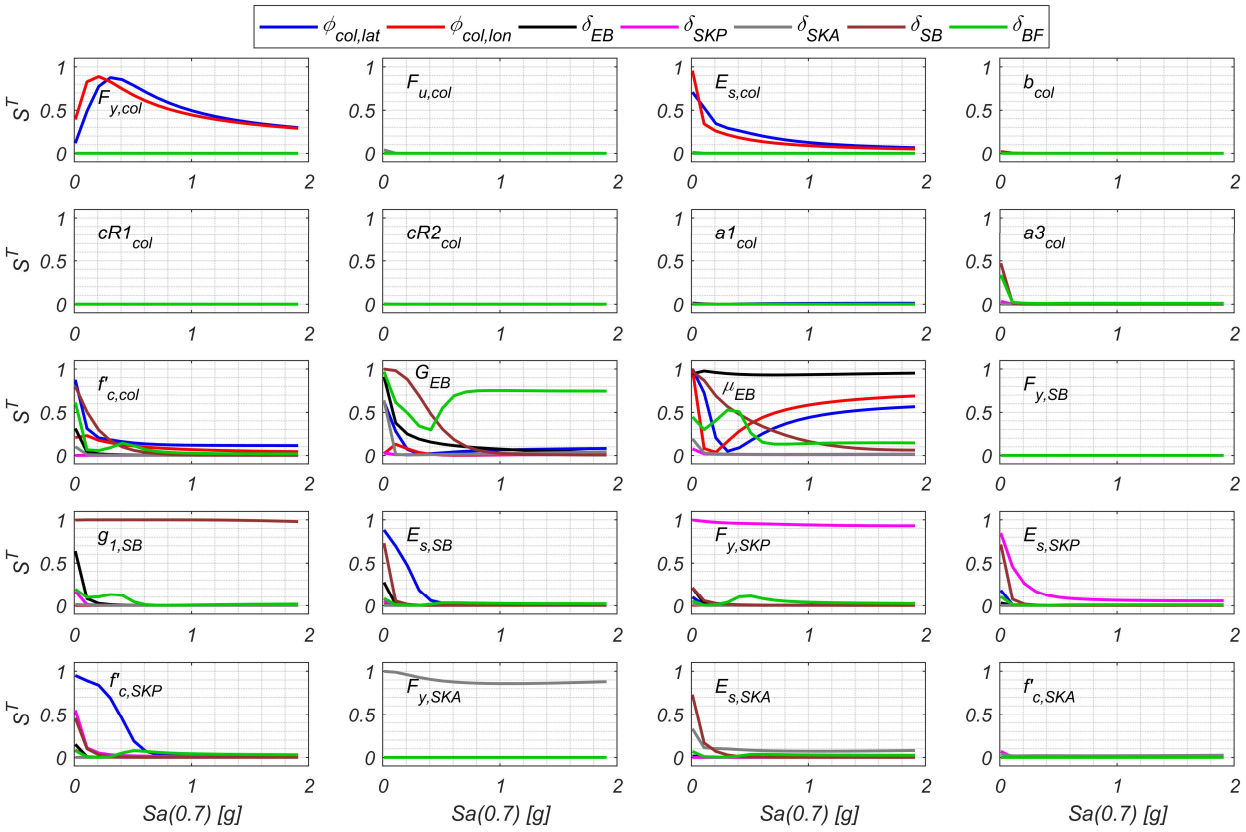


Figure 4.4-20. Total Sobol indices for all parameters.

In general, performing GSA using the proposed methodology can help provide a deeper understanding of the composition of uncertainty in the performance of complex structures at a drastically reduced computational cost.

For reference, the 10,550 original FE model simulations were performed on a high-performance computing cluster, using five nodes equipped with 32-core AMD Ryzen 9 5950X CPUs and 64GB of RAM, allowing for 160 concurrent NLTHA. The total wall-clock time of the simulations was approximately 250 hours. In contrast, the 211,000 GP evaluations for fragility analysis and 23.21×10^6 GP evaluations for GSA had a computational cost of about 5 minutes and 9 hours, respectively, on a laptop equipped with a 12-core AMD Ryzen 9 7845HX CPU and 32GB of RAM. Executing both analyses would have been unfeasible relying on FE models alone.

4.5 CLOSING REMARKS ON THE PROPOSED METHODOLOGIES

4.5.1 DISCUSSION AND CLOSING REMARKS ON APPLICATION I

Sections 4.3 and 4.4.3 apply and validate an efficient approach to study the uncertainty induced by model parameters in system-level structural models using GP SMs, as proposed in section 3.6. Through sampling uncertain parameters and NLTHA simulation, the approach uses a reduced subset of FE simulation data to train GPs. With as few as 200 training points, the models achieved R^2 values above 0.9 when predicting 9,800 new samples in 4.4.3.

The method was validated with two case studies: a five-story reinforced concrete (RC) frame building and a five-span highway bridge. In both cases, the GPs accurately modeled several key responses. Comparisons of LSA results confirmed that the SMs preserved the relationships between input parameters and structural responses. The low cost of sampling GPs enabled efficient GSA using Sobol indices. This approach can also support other computationally intensive tasks like model updating and probabilistic fragility analysis, drastically reducing the required time and computational resources.

Although extensive research has focused on surrogate modeling of structure portfolios and regional approaches, the focus on a single scenario allows for a deeper understanding of the dynamic phenomena governing structural behavior. In this context, the proposed approach can help establish more efficient analysis for such studies, which could aid decisions on experimental data campaigns for different materials and components. For example, preliminary SA could be carried out to identify the most influential parameters on a certain structural model, for which experimental data could be gathered to probabilistically calibrate constitutive models, thus improving the fidelity of the model to the actual structure. A case-specific example of this would be the calibration of *Steel02* for AN. Gathering experimental data from cyclic uniaxial testing of rebars would not improve model fidelity significantly due to the low sensitivity of column responses to nonlinear *Steel02* parameters even at MCE hazard. In this case, calibration of linear model parameters against experimental data from monotonic uniaxial tests would suffice, while drastically reducing the cost of experimental testing (i.e., cyclic loading-capable universal testing machine and measuring equipment).

Among other machine learning alternatives, GPs were employed for this application due to their inherently probabilistic definition and straightforward interpretation. In the context of uncertainty quantification, GP directly provides a measure of epistemic uncertainty through the posterior variance. Further research in this line could extend the proposed approach by evaluating different algorithms, as well as multivariate approaches to surrogate modeling.

4.5.2 DISCUSSION AND CLOSING REMARKS ON APPLICATION II

Section 4.4.4 applies and validates an efficient approach for the quantification and propagation of uncertainty to support practical implementation of PBEE by using GP SMs, as proposed in section 3.7. The proposed approach incorporates both aleatoric and epistemic sources of uncertainty in hazard, structural response, and damage assessment, providing a complete probabilistic framework for structural performance assessment.

Cloud fragility analysis is conducted based on probabilistic seismic hazard analysis (PSHA), probabilistic seismic demand analysis (PSDA), and probabilistic seismic fragility analysis (PSFA). To this end, uncertainties at each stage are accounted for through state-of-the-art seismicity models and ground motion models (GMMs), an extensive ground motion (GM) database covering Chilean and Japanese subduction earthquakes, Bayesian calibration using experimental data for nonlinear bridge components' models, and probabilistic limit state (LS) and capacity definitions. Subsequently, the uncertainty in expected damage state recurrence is discussed, and the composition of parameter-induced uncertainty in fragility estimates is investigated through GSA based on total Sobol indices. Through the presented application example, the use of GP SMs to predict structural responses is validated. The approach presented herein uses seismic and structural model parameters as predictors, achieving accurate model predictions on new earthquakes and parameter samples. Towards SM training, the stratified approach to selecting training data provided an effective method for training the models with few data points while maintaining a representative sample of the data distributions of interest.

The comparison of fragility functions from the original FE model and a new sample generated from the validated SMs confirmed the viability of further sampling of the GPs for uncertainty decomposition in parameter-conditioned fragility functions. In this context, the

importance of accounting for model parameter uncertainty in PBEE is confirmed, as evidenced by the considerable variability in conditional fragility functions. To achieve a realistic and scientifically robust framework for structural performance assessment, a comprehensive treatment of uncertainties is required. Therefore, future efforts should focus on developing efficient methodologies that address this need, leveraging the growing availability of advanced computational tools.

Additionally, damage state recurrence investigated as part of PSFA revealed inconsistencies among component-based LS definitions, in terms of their consequences in the structure's performance. Thus, further investigation is required to tackle this issue. Although the scope of the application presented in this manuscript is limited to component-level PSFA, the resulting damage state recurrence rates can be extended to system-level PSFA.

It is noted that the proposed methodology can be extended to account for multiple sources of hazard, aging-induced deterioration, and other structure types. Similarly, alternative surrogate modeling and uncertainty quantification techniques could be employed, although Bayesian methods (i.e., Bayesian model calibration, GP) are recommended as an inherent fit to a probabilistic framework such as PBEE.

5 CONCLUSIONS

5.1 THEORETICAL CONCLUSIONS

The main and specific objectives of this doctoral research were fulfilled by developing and validating two approaches for the efficient treatment of uncertainties within different aspects of PBEE.

In terms of parameter-induced uncertainty quantification in hazard-specific scenarios, the proposed methodology was able to efficiently capture the relationships between model parameters and structural responses, as well as identify the most influential constitutive model parameters from the standpoint of structural responses. This was validated by comparing local sensitivity analysis results from the original FE model and the GP surrogates and extended to global sensitivity analysis enabled by the low cost of GP surrogate sampling.

Regarding UQ&P in PBEE and its decomposition in terms of model parameters, the developed GP surrogates were able to accurately predict responses under new parameter samples and GMs, thus effectively replacing the original FE model for performance assessment. This is critical, as the final sampling-based sensitivity analysis would not be feasible using the original FE model due to sampling times. At the same time, said sensitivity analysis revealed the impact of certain model parameters on multiple fragility functions, while other parameters were unexpectedly less impactful.

The GP surrogates in application II were trained under a wide range of hazard levels, parameter distributions, and limit state and capacity samples, while compounding all sources of uncertainty in fragility analysis. To the best of the author's knowledge, such application to a real-world case study is novel and provides evidence that a comprehensive approach to UQ&P and a full implementation of the PBEE framework can be achieved.

5.2 PRACTICAL CONCLUSIONS

The computational cost of generating further sampling of structural responses was reduced by hundreds to thousands-fold by using the trained GP surrogate models in place of the original FE models, while retaining excellent accuracy. In view of further research using the

RCB or the AN bridge as study cases, the development of GP surrogates will prove useful in drastically reducing the computational cost future sampling-based analysis under a variety of earthquake scenarios. This conclusion can be extended to the feasibility of developing probabilistic surrogates for other structures.

In terms of initial investment, it remains true that data collection for constitutive model calibration and surrogate model training can be expensive. However, constitutive model calibration at a material level can be shared between different structures, thus lowering subsequent needs for experimental campaigns. For instance, by calibrating reinforcing steel constitutive models, the calibrated material can be used in probabilistic characterizations of different structure types featuring reinforced concrete.

It is crucial to conduct sensitivity analysis within a complete performance assessment context, as model parameters may gain or lose impact on the final decision variable depending on structural component interactions and hazard scenarios. Thus, sensitivity analysis for parameter selection can help design more effective experimental data collection campaigns and thus improve resource allocation in data collection.

Finally, building and publishing databases of probabilistic characterizations of commonly used structural materials and components can drastically facilitate the implementation and adoption of a robust and fully probabilistic PBEE framework by the industry.

6 REFERENCES

Álvarez, M. A., Rosasco, L., & Lawrence, N. D. (2012). Kernels for Vector-Valued Functions: A Review. *Foundations and Trends in Machine Learning*, *4*(3), 195–266. https://doi.org/10.1561/2200000036

American Concrete Institute. (2019). *Building code requirements for structural concrete (ACI 318-19): An ACI standard; commentary on building code requirements for structural concrete (ACI 318R-19).*

Astroza, R., Ebrahimian, H., Li, Y., & Conte, J. P. (2017). Bayesian nonlinear structural FE model and seismic input identification for damage assessment of civil structures. *Mechanical Systems and Signal Processing*, *93*(2), 661–687. https://doi.org/10.1016/j.ymssp.2017.01.040

Baker, J. W. (2011). Conditional Mean Spectrum: Tool for Ground-Motion Selection. *Journal of Structural Engineering*, *137*(3), 322–331. https://doi.org/10.1061/(ASCE)ST.1943-541X.0000215

Baker, J. W. (2015). Efficient Analytical Fragility Function Fitting Using Dynamic Structural Analysis. *Earthquake Spectra*, *31*(1), 579–599. https://doi.org/10.1193/021113EQS025M

Baker, J. W., Bradley, B. A., & Stafford, P. J. (2021). *Probabilistic seismic hazard and risk analysis*. Cambridge University Press. https://doi.org/10.1017/9781108425056

Balan, T. A., Spacone, E., & Kwon, M. (2001). A 3D hypoplastic model for cyclic analysis of concrete structures. *Engineering Structures*, *23*(4), 333–342. https://doi.org/10.1016/S0141-0296(00)00048-1

Bažant, Z. P., & Pfeiffer, P. A. (1986). Shear fracture tests of concrete. *Materials and Structures*, *19*(2), 111–121. https://doi.org/10.1007/BF02481755

Bazzurro, P., & Allin Cornell, C. (1999). Disaggregation of seismic hazard. *Bulletin of the Seismological Society of America*, *89*(2), 501–520. https://doi.org/10.1785/BSSA0890020501

Birrell, M., Astroza, R., Carreño, R., Restrepo, J. I., & Araya-Letelier, G. (2021). Bayesian parameter and joint probability distribution estimation for a hysteretic constitutive model of reinforcing steel. *Structural Safety*, *90*(16). https://doi.org/10.1016/j.strusafe.2020.102062

Birrell, M., Astroza, R., Murcia-Delso, J., Hernández, F., & Bazáez, R. (2024). A simplified tri-linear model for monolithic exterior shear keys failing in sliding shear. *Structures*, *61*(14), 105934. https://doi.org/10.1016/j.istruc.2024.105934

Birrell, M., Astroza, R., Restrepo, J. I., Lotfizadeh, K., Carreño, R., Bazáez, R., & Hernández, F. (2021). Bayesian inference for calibration and validation of uniaxial

reinforcing steel models. *Engineering Structures*, *243*(1). https://doi.org/10.1016/j.engstruct.2021.112386

Bozorgzadeh, A., Megally, S., Ashford, S., & Restrepo, J. (2006). *Seismic response of sacrificial exterior keys in bridge abutments*. Report No. SSRP–04/14. University of California San Diego.

Bozorgzadeh, A., Megally, S., Restrepo, J., & Ashford, S. (2006). Capacity Evaluation of Exterior Sacrificial Shear Keys of Bridge Abutments. *Journal of Bridge Engineering*, *11*(5).

Buckle, I., Hube, M., Chen, G., Yen, W.-H., & Arias, J. (2012). Structural Performance of Bridges in the Offshore Maule Earthquake of 27 February 2010. *Earthquake Spectra*, *28*(1_suppl1), 533–552. https://doi.org/10.1193/1.4000031

Buckle, I., Luco, N., & Kowalsky, M. (2020). *Proposed AASHTO Guidelines for Performance-Based Seismic Bridge Design*. Transportation Research Board. https://doi.org/10.17226/25913

Candia, G., Macedo, J., Jaimes, M. A., & Magna-Verdugo, C. (2019). A New State-of-the-Art Platform for Probabilistic and Deterministic Seismic Hazard Assessment. *Seismological Research Letters*, *90*(6), 2262–2275. https://doi.org/10.1785/0220190025

Cao, X.-Y., Feng, D.-C., & Li, Y. (2023). Assessment of various seismic fragility analysis approaches for structures excited by non-stationary stochastic ground motions. *Mechanical Systems and Signal Processing*, *186*(4), 109838. https://doi.org/10.1016/j.ymssp.2022.109838

Carreño, R., Lotfizadeh, K. H., Conte, J. P., & Restrepo, J. I. (2020). Material Model Parameters for the Giuffrè-Menegotto-Pinto Uniaxial Steel Stress-Strain Model. *Journal of Structural Engineering*, *146*(2). https://doi.org/10.1061/(ASCE)ST.1943-541X.0002505

Castellon, D. F., Fenerci, A., Petersen, Ø. W., & Øiseth, O. (2023). Full long-term buffeting analysis of suspension bridges using Gaussian process surrogate modelling and importance sampling Monte Carlo simulations. *Reliability Engineering & System Safety*, *235*, 109211. https://doi.org/10.1016/j.ress.2023.109211

Castro, S., Benavente, R., Crempien, J. G. F., Candia, G., & La Llera, J. C. de (2022). A Consistently Processed Strong-Motion Database for Chilean Earthquakes. *Seismological Research Letters*, *93*(5), 2700–2718. https://doi.org/10.1785/0220200336

Cheng, K., Lu, Z., Ling, C., & Zhou, S. (2020). Surrogate-assisted global sensitivity analysis: an overview. *Structural and Multidisciplinary Optimization*, *61*(3), 1187–1213. https://doi.org/10.1007/s00158-019-02413-5

Cheng, K., Lu, Z., Zhou, Y., Shi, Y., & Wei, Y. (2017). Global sensitivity analysis using support vector regression. *Applied Mathematical Modelling*, *49*(4), 587–598. https://doi.org/10.1016/j.apm.2017.05.026

Conde Bandini, P. A., Padgett, J. E., Paultre, P., & Siqueira, G. H. (2022). Seismic fragility of bridges: An approach coupling multiple-stripe analysis and Gaussian mixture for multicomponent structures. *Earthquake Spectra*, *38*(1), 254–282. https://doi.org/10.1177/87552930211036164

Deb, A. (2021). *Risk-targeted performance-based seismic assessment and design of bridges* [PhD dissertation]. University of California San Diego. https://escholarship.org/uc/item/57x2t962

Deb, A., Conte, J. P., & Restrepo, J. I. (2022). Comprehensive treatment of uncertainties in risk-targeted performance-based seismic design and assessment of bridges. *Earthquake Engineering & Structural Dynamics*, *51*(14), 3272–3295. https://doi.org/10.1002/eqe.3722

Dodd, L. L., & Restrepo-Posada, J. I. (1995). Model for Predicting Cyclic Behavior of Reinforcing Steel. *Journal of Structural Engineering,* 121(3), pp. 433–445.

Duncan, J. M., & Mokwa, R. L. (2001). Passive Earth Pressures: Theories and Tests. *Journal of Geotechnical and Geoenvironmental Engineering*, *127*(3), 248–257. https://doi.org/10.1061/(ASCE)1090-0241(2001)127:3(248)

Eads, L., Miranda, E., Krawinkler, H., & Lignos, D. G. (2013). An efficient method for estimating the collapse risk of structures in seismic regions. *Earthquake Engineering & Structural Dynamics*, *42*(1), 25–41. https://doi.org/10.1002/eqe.2191

Efron, B., & Stein, C. (1981). The Jackknife Estimate of Variance. *The Annals of Statistics*, *9*(3). https://doi.org/10.1214/aos/1176345462

Fatemeh Jalayer. (2003). *Direct Probabilistic Seismic Analysis: Implementation Non-Linear Dynamic Assessments*. Ph. D. Dissertation. Department of Civil and Environmental Engineering, University of Stanford.

Fayaz, J., Astroza, R., Angione, C., & Medalla, M. (2024). Data-driven analysis of crustal and subduction seismic environments using interpretation of deep learning-based generalized ground motion models. *Expert Systems with Applications*, *238*, 121731. https://doi.org/10.1016/j.eswa.2023.121731

Filippou, F. C., Popov, E. P., & Bertero, V. V. (1983). *Effects of bond deterioration on hysteretic behavior of reinforced concrete joints*. no. UCB/EERC 83-19.

Gelman, A., & Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, *7*(4). https://doi.org/10.1214/ss/1177011136

Gentile, R., & Galasso, C. (2020). Gaussian process regression for seismic fragility assessment of building portfolios. *Structural Safety*, *87*, 101980. https://doi.org/10.1016/j.strusafe.2020.101980

Gidaris, I., Taflanidis, A. A., & Mavroeidis, G. P. (2015). Kriging metamodeling in seismic risk assessment based on stochastic ground motion models. *Earthquake Engineering & Structural Dynamics*, *44*(14), 2377–2399. https://doi.org/10.1002/eqe.2586

Goel, R. K., & Chopra, A. K. (2008). Role of Shear Keys in Seismic Behavior of Bridges Crossing Fault-Rupture Zones. *Journal of Bridge Engineering*, *13*(4), 398–408. https://doi.org/10.1061/(ASCE)1084-0702(2008)13:4(398)

Gomez-Cabrera, A., & Escamilla-Ambrosio, P. J. (2022). Review of Machine-Learning Techniques Applied to Structural Health Monitoring Systems for Building and Bridge Structures. *Applied Sciences*, *12*(21), 10754. https://doi.org/10.3390/app122110754

Guajardo, B., Pinto, F., & Astroza, R. (2024). Effects of soil spatial variability on the seismic response of multi-span simply-supported highway bridges. *Bulletin of Earthquake Engineering*, *22*(5), 2643–2675. https://doi.org/10.1007/s10518-024-01872-6

Günay, S., & Mosalam, K. M. (2013). PEER Performance-Based Earthquake Engineering Methodology, Revisited. *Journal of Earthquake Engineering*, *17*(6), 829–858. https://doi.org/10.1080/13632469.2013.787377

Haario, H., Saksman, E., & Tamminen, J. (2001). An Adaptive Metropolis Algorithm. *Bernoulli*, *7*(2), 223–242. https://doi.org/10.2307/3318737

Han, Q., Zhou, Y., Ou, Y., & Du, X. (2017). Seismic behavior of reinforced concrete sacrificial exterior shear keys of highway bridges. *Engineering Structures*, *139*, 59–70. https://doi.org/10.1016/j.engstruct.2017.02.034

Han, Q., Zhou, Y., Zhong, Z., & Du, X. (2017). Seismic Capacity Evaluation of Exterior Shear Keys of Highway Bridges. *Journal of Bridge Engineering*, *22*(2), Article 04016119, 3. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000978

Hayes, G. P., Moore, G. L., Portner, D. E., Hearne, M., Flamme, H., Furtney, M., & Smoczyk, G. M. (2018). Slab2, a comprehensive subduction zone geometry model. *Science (New York, N.Y.)*, *362*(6410), 58–61. https://doi.org/10.1126/science.aat4723

He, X. G., & Kwan, A.K.H. (2001). Modeling dowel action of reinforcement bars for finite element analysis of concrete structures. *Computers & Structures*, *79*(6), 595–604. https://doi.org/10.1016/S0045-7949(00)00158-9

Hill, T. P., & Miller, J. (2010). *How to Combine Independent Data Sets for the Same Quantity*. https://arxiv.org/pdf/1005.4978.pdf https://doi.org/10.1063/1.3593373

Hoang, P. H., Phan, H. N., Nguyen, D. T., & Paolacci, F. (2021). Kriging Metamodel-Based Seismic Fragility Analysis of Single-Bent Reinforced Concrete Highway Bridges. *Buildings*, *11*(6), 238. https://doi.org/10.3390/buildings11060238

Homma, T., & Saltelli, A. (1996). Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*, *52*(1), 1–17. https://doi.org/10.1016/0951-8320(96)00002-6

Hube, M., Martínez, A., & Rubilar, F. (Eds.) (2017). *Experimental behavior of elastomeric bearing and seismic bars of simply supported Chilean bridges*. Paper Nr. 4682.

Hung, D. V., & Thang, N. T. (2022). Predicting dynamic responses of frame structures subjected to stochastic wind loads using temporal surrogate model. *Journal of Science and Technology in Civil Engineering*, *16*(2), 106–116. https://doi.org/10.31814/stce.huce(nuce)2022-16(2)-09

Idini, B., Rojas, F., Ruiz, S., & Pastén, C. (2017). Ground motion prediction equation for the Chilean subduction zone. *Bulletin of Earthquake Engineering*, *15*, 1853–1880.

Jalayer, F., Ebrahimian, H., Miano, A., Manfredi, G., & Sezen, H. (2017). Analytical fragility assessment using unscaled ground motion records. *Earthquake Engineering & Structural Dynamics*, *46*(15), 2639–2663. https://doi.org/10.1002/eqe.2922

Jalayer, F., Risi, R. de, & Manfredi, G. (2015). Bayesian Cloud Analysis: efficient structural fragility assessment using linear regression. *Bulletin of Earthquake Engineering*, *13*(4), 1183–1203. https://doi.org/10.1007/s10518-014-9692-z

Janon, A., Klein, T., Lagnoux, A., Nodet, M., & Prieur, C. (2014). Asymptotic normality and efficiency of two Sobol index estimators. *ESAIM: Probability and Statistics*, *18*, 342–364. https://doi.org/10.1051/ps/2013040

Kang, F., Wu, Y., Ma, J., & Li, J. (2023). Structural identification of super high arch dams using Gaussian process regression with improved salp swarm algorithm. *Engineering Structures*, *286*, 116150. https://doi.org/10.1016/j.engstruct.2023.116150

Kawashima, K. (2012). Damage of bridges due to the 2011 Great East Japan Earthquake. *Journal of Japan Association for Earthquake Engineering*, *12*, 4_319-4_338. https://doi.org/10.5610/jaee.12.4_319

Kawashima, K., Unjoh, S., Hoshikuma, J.-I., & Kosa, K. (2011). Damage of Bridges due to the 2010 Maule, Chile, Earthquake. *Journal of Earthquake Engineering*, *15*(7), 1036–1068. https://doi.org/10.1080/13632469.2011.575531

Kim, J., & Kim, T. (2025). Efficient seismic fragility analysis considering uncertainties in structural systems and ground motions. *Earthquake Engineering & Structural Dynamics*, *54*(1), 206–226. https://doi.org/10.1002/eqe.4254

Kiureghian, A. D., & Ditlevsen, O. (2009). Aleatory or epistemic? Does it matter? *Structural Safety*, *31*(2), 105–112. https://doi.org/10.1016/j.strusafe.2008.06.020

Kottari, A. (2016). *Horizontal Load Resisting Mechanisms of External Shear Keys in Bridge Abutments: PhD Thesis.* University of California San Diego.

Kottari, A., Shing, P. B., & Bromenschenkel, R. (2020). Shear Behavior of Exterior Non-Isolated Shear Keys in Bridge Abutments. *ACI Structural Journal*, *117*(2). https://doi.org/10.14359/51721317

Kroese, D. P., Taimre, T., & Botev, Z. I. (2011). *Handbook of Monte Carlo methods*. *Wiley series in probability and statistics*. Wiley. http://onlinelibrary.wiley.com/book/10.1002/9781118014967 https://doi.org/10.1002/9781118014967

Kudela, J., & Matousek, R. (2022). Recent advances and applications of surrogate models for finite element method computations: a review. *Soft Computing*, *26*(24), 13709–13733. https://doi.org/10.1007/s00500-022-07362-8

Lan, Y., Xu, J., & Pinnola, F. (2024). Seismic fragility analysis of structures via an Adaptive Gaussian Mixture Model and its application to resilience assessment. *Mechanical Systems and Signal Processing*, *212*(8), 111332. https://doi.org/10.1016/j.ymssp.2024.111332

Li, Q., Du, X., Ni, P., Han, Q., Xu, K., & Yuan, Z. (2024). Efficient Bayesian inference for finite element model updating with surrogate modeling techniques. *Journal of Civil Structural Health Monitoring*, *14*(4), 997–1015. https://doi.org/10.1007/s13349-024-00768-y

Lin, T., Harmsen, S. C., Baker, J. W., & Luco, N. (2013). Conditional Spectrum Computation Incorporating Multiple Causal Earthquakes and Ground-Motion Prediction Models. *Bulletin of the Seismological Society of America*, *103*(2A), 1103–1116. https://doi.org/10.1785/0120110293

Lin, T., Haselton, C. B., & Baker, J. W. (2013a). Conditional spectrum-based ground motion selection. Part I: Hazard consistency for risk-based assessments. *Earthquake Engineering & Structural Dynamics*, *42*(12), 1847–1865. https://doi.org/10.1002/eqe.2301

Lin, T., Haselton, C. B., & Baker, J. W. (2013b). Conditional spectrum-based ground motion selection. Part II: Intensity-based assessments and evaluation of alternative target spectra. *Earthquake Engineering & Structural Dynamics*, *42*(12), 1867–1884. https://doi.org/10.1002/eqe.2303

Liu, D. C., & Nocedal, J. (1989). On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, *45*(1-3), 503–528. https://doi.org/10.1007/BF01589116

Liu, Z., Sextos, A., Guo, A., & Zhao, W. (2022). ANN-based rapid seismic fragility analysis for multi-span concrete bridges. *Structures*, *41*, 804–817. https://doi.org/10.1016/j.istruc.2022.05.063

Lu, J., Zhan, Z., Apley, D. W., & Chen, W. (2019). Uncertainty propagation of frequency response functions using a multi-output Gaussian Process model. *Computers & Structures*, *217*, 1–17. https://doi.org/10.1016/j.compstruc.2019.03.009

Mander, T. J., & Matamoros, A. B. (2019). Constitutive modeling and overstrength factors for reinforcing steel. *ACI Structural Journal,* 116(3), pp. 219–232.

Mangalathu, S., Jeon, J.-S., & DesRoches, R. (2018). Critical uncertainty parameters influencing seismic performance of bridges using Lasso regression. *Earthquake Engineering & Structural Dynamics*, *47*(3), 784–801. https://doi.org/10.1002/eqe.2991

Marelli, S., Lamas, C., Konakli, K., Mylonas, C., Wiederkehr, P., & Sudret, B. (2022). *UQLab user manual - Sensitivity analysis.* Chair of Risk, Safety and Uncertainty Quantification, ETH Zurich, Switzerland.

Martínez, A., Hube, M. A., & Rollins, K. M. (2017). Analytical fragility curves for non-skewed highway bridges in Chile. *Engineering Structures*, *141*, 530–542. https://doi.org/10.1016/j.engstruct.2017.03.041

Martínez, A. (2016). *Efecto de las barras sísmicas en el comportamiento sísmico transversal de puentes de hormigón armado (In Spanish)*. Msc. Thesis. https://doi.org/10.7764/tesisUC/ING/15703

Mckenna, F., Fenves, G. L., & Scott., M. H. *Open system for earthquake engineering simulation.* Pacific Earthquake Engineering Research, University of California. https://opensees.berkeley.edu

Megally, S., Silva, P., & Seible, F. (2002). *Seismic response of sacrificial shear keys in bridge abutments: Report No. SSRP-2001/23.*

Mei, H., & Guo, A. (2023). Quasi-static experimental study on seismic performance of exterior shear key with different failure modes. *Engineering Structures*, *287*, 116173. https://doi.org/10.1016/j.engstruct.2023.116173

Menegotto, M., & Pinto, P. (1973). Method of Analysis for Cyclically Loaded R.C. Plane Frames Including Changes in Geometry and Non-Elastic Behavior of Elements under Combined Normal Force and Bending. In *IABSE Symposium of Resistance and Ultimate Deformability of Structures Acted on by Well Defined Repeated Loads*.

Ministry of Transport of the People's Republic of China. (2008). *Guideline for seismic design of highway bridges.*

Mirza, S. A., MacGregor, J. G., & Hatzinikolas, M. (1979a). Statistical Descriptions of Strength of Concrete. *Journal of the Structural Division*, *105*(6), 1021–1037. https://doi.org/10.1061/JSDEAG.0005161

Mirza, S. A., MacGregor, J. G., & Hatzinikolas, M. (1979b). Statistical Descriptions of Strength of Concrete. *Journal of the Structural Division*, *105*(6), 1021–1037. https://doi.org/10.1061/JSDEAG.0005161

Mo, R., Chen, L., Chen, Y., Xiong, C., Zhang, C., Chen, Z., & Lin, E. (2024). Prediction and correlations estimation of seismic capacities of pier columns: Extended Gaussian process regression models. *Structural Safety*, *109*(2), 102457. https://doi.org/10.1016/j.strusafe.2024.102457

Montalva, G., Bastias, N., & Rodriguex-Marck, A. (2017). Ground motion prediction equation for the Chilean subduction zone. *Bulletin of the Seismological Society of America,* 107(2), pp. 901–911.

Nabilah, A. B., & Koh, C. G. (2017). Experimental study of intermediate length coupling beams subjected to monotonic load. *KSCE Journal of Civil Engineering*, *21*(7), 2807–2813. https://doi.org/10.1007/s12205-017-1185-6

Nagel, J. B., Rieckermann, J., & Sudret, B. (2020). Principal component analysis and sparse polynomial chaos expansions for global sensitivity analysis and model calibration: Application to urban drainage simulation. *Reliability Engineering & System Safety*, *195*(11), 106737. https://doi.org/10.1016/j.ress.2019.106737

National Research Institute for Earth Science and Disaster Resilience. *NIED K-NET, KiK-net.* National Research Institute for Earth Science and Disaster Resilience. www.kyoshin.bosai.go.jp https://doi.org/10.17598/NIED.0004

Nielson, B. G., & DesRoches, R. (2007). Seismic fragility methodology for highway bridges using a component level approach. *Earthquake Engineering & Structural Dynamics*, *36*(6), 823–839. https://doi.org/10.1002/eqe.655

Ning, C., Xie, Y., Burton, H., & Padgett, J. E. (2024). Enabling efficient regional seismic fragility assessment of multi-component bridge portfolios through Gaussian process regression and active learning. *Earthquake Engineering & Structural Dynamics*, *53*(9), 2929–2949. https://doi.org/10.1002/eqe.4144

Padgett, J. E., & DesRoches, R. (2007). Sensitivity of Seismic Response and Fragility to Parameter Uncertainty. *Journal of Structural Engineering*, *133*(12), 1710–1718. https://doi.org/10.1061/(ASCE)0733-9445(2007)133:12(1710)

Pan, H., Tian, L., Fu, X., & Li, H. (2020). Sensitivities of the seismic response and fragility estimate of a transmission tower to structural and ground motion uncertainties.

Journal of Constructional Steel Research, *167*(10), 105941. https://doi.org/10.1016/j.jcsr.2020.105941

Pastén, C., Astroza, R., Bazáez, R., Contreras, N., Grand, J., Hernández, F., & Ochoa, F. (2021). *Guia para el análisis sísmico no lineal de puentes chilenos*. U. de los Andes; U. de Chile; UTSFM. https://drive.google.com/drive/folders/1sW4VPqxU2sXLErOYE1mMABAuP8wNa-EL

Patsialis, D., & Taflanidis, A. A. (2020). Reduced order modeling of hysteretic structural response and applications to seismic risk assessment. *Engineering Structures*, *209*(7), 110135. https://doi.org/10.1016/j.engstruct.2019.110135

Pinto, F., Astroza, R., Bazáez, R., Hernández, F., & Navarro, N. (2024). Probabilistic seismic assessment of multispan RC highway bridges considering soil-structure interaction and chloride-induced corrosion. *Engineering Structures*, *301*(11), 117257. https://doi.org/10.1016/j.engstruct.2023.117257

Pinto, F., Astroza, R., Pizarro, A., Bazáez, R., & Hernández, F. (2024). Seismic fragility analysis of simply supported bridges considering uncertainty in scour condition. *Structures*, *64*(11), 106570. https://doi.org/10.1016/j.istruc.2024.106570

Pinto, F. J., Toledo, J., Birrell, M., Bazáez, R., Hernández, F., & Astroza, R. (2023). Uncertainty Quantification in Constitutive Models of Highway Bridge Components: Seismic Bars and Elastomeric Bearings. *Materials*, *16*(5). https://doi.org/10.3390/ma16051792

Popovics, S. (1973). A numerical approach to the complete stress-strain curve of concrete. *Cement and Concrete Research*, *3*(5), 583–599. https://doi.org/10.1016/0008-8846(73)90096-3

Poulos, A., Monsalve, M., Zamora, N., & La Llera, J. C. de (2019). An Updated Recurrence Model for Chilean Subduction Seismicity and Statistical Validation of Its Poisson Nature. *Bulletin of the Seismological Society of America*, *109*(1), 66–74. https://doi.org/10.1785/0120170160

Rasmussen, C. E., & Williams, C. K. I. (2005). *Gaussian Processes for Machine Learning*. The MIT Press. https://direct.mit.edu/books/monograph/2320/Gaussian-Processes-for-Machine-Learning https://doi.org/10.7551/mitpress/3206.001.0001

Rubilar, F. (2016). *Modelo no lineal para predecir la respuesta sísmica de pasos superiores (In Spanish)*. Msc Thesis. https://doi.org/10.7764/tesisUC/ING/15609

S. A. Mirza, & J. G. MacGregor. (1976). *A Statistical Study of Variables affecting the Strength of Reinforced Normal Weight Concrete Members*. Structural Engineering Report No. 58. Department of Civil Engineering, University of Alberta.

Saenz, L. P. (1964). Discussion of 'Equation for the stress-strain curve of concrete' by Desayi/Krishnan. *Journal of American Concrete Institute*, *61*(9), 1229–1236.

Saida, T., & Nishio, M. (2023). Transfer learning Gaussian process regression surrogate model with explainability for structural reliability analysis under variation in uncertainties. *Computers & Structures*, *281*, 107014. https://doi.org/10.1016/j.compstruc.2023.107014

Saida, T., Rashid, M., & Nishio, M. (2024). System fragility analysis of highway bridge using multi-output Gaussian process regression surrogate model. *Advances in Structural Engineering*, *27*(16), 2803–2822. https://doi.org/10.1177/13694332241291255

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008). *Global sensitivity analysis: The primer*. John Wiley. http://dx.doi.org/10.1002/9780470725184 https://doi.org/10.1002/9780470725184

Sarkar, N., & Dasgupta, K. (2024). Machine learning-based sensitivity analysis of engineering demand parameters for a reinforced concrete wall-frame building. *Structures*, *70*, 107477. https://doi.org/10.1016/j.istruc.2024.107477

Sezen, H., & Moehle, J. P. (2004). Shear Strength Model for Lightly Reinforced Concrete Columns. *Journal of Structural Engineering*, *130*(11), 1692–1703. https://doi.org/10.1061/(ASCE)0733-9445(2004)130:11(1692)

Shang, X., Su, L., Fang, H., Zeng, B., & Zhang, Z. (2023). An efficient multi-fidelity Kriging surrogate model-based method for global sensitivity analysis. *Reliability Engineering & System Safety*, *229*, 108858. https://doi.org/10.1016/j.ress.2022.108858

Sharifi, M., & Shafieian, M. (2018). Effective stiffness of concrete shear walls based on statistical analysis. *Structural Concrete*, *19*(6), 1560–1576. https://doi.org/10.1002/suco.201600202

Shome, N., & Cornell, C. A. (1999). *Probabilistic Seismic Demand Analysis of Nonlinear Structures. Reliability of Marine Structures*. Program Report No. RMS-35. Department of Civil and Environmental Engineering, University of Stanford.

Silva, P., Megally, S., & Seible, F. (2003). Seismic Performance of Sacrificial Interior Shear Keys. *ACI Structural Journal*, *100*(2).

Silva, P. F., Megally, S., & Seible, F. (2009). Seismic Performance of Sacrificial Exterior Shear Keys in Bridge Abutments. *Earthquake Spectra*, *25*(3), 643–664. https://doi.org/10.1193/1.3155405

Skandalos, K., Chakraborty, S., & Tesfamariam, S. (2022). Seismic reliability analysis using a multi-fidelity surrogate model: Example of base-isolated buildings. *Structural Safety*, *97*(1), 102222. https://doi.org/10.1016/j.strusafe.2022.102222

Snoek, J., Larochelle, H., & Adams, R. P. (2012). *Practical Bayesian Optimization of Machine Learning Algorithms*. https://doi.org/10.48550/arXiv.1206.2944

Sobol, I. M. (1993). Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiments*, *1*(4), 407–414.

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

Sobol, I.M. (1967). On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics*, *7*(4), 86–112. https://doi.org/10.1016/0041-5553(67)90144-9

Soleimani, F. (2021). Analytical seismic performance and sensitivity evaluation of bridges based on random decision forest framework. *Structures*, *32*, 329–341. https://doi.org/10.1016/j.istruc.2021.02.049

Soleimani, F., & Hajializadeh, D. (2022). State-of-the-Art Review on Probabilistic Seismic Demand Models of Bridges: Machine-Learning Application. *Infrastructures*, *7*(5), 64. https://doi.org/10.3390/infrastructures7050064

Spiridonakos, M. D., & Chatzi, E. N. (2015). Metamodeling of dynamic nonlinear structural systems through polynomial chaos NARX models. *Computers & Structures*, *157*(9), 99–113. https://doi.org/10.1016/j.compstruc.2015.05.002

Stefanidou, S. P., & Kappos, A. J. (2017). Methodology for the development of bridge-specific fragility curves. *Earthquake Engineering & Structural Dynamics*, *46*(1), 73–93. https://doi.org/10.1002/eqe.2774

Stefanidou, S. P., Paraskevopoulos, E. A., Papanikolaou, V. K., & Kappos, A. J. (2022). An online platform for bridge-specific fragility analysis of as-built and retrofitted bridges. *Bulletin of Earthquake Engineering*, *20*(3), 1717–1737. https://doi.org/10.1007/s10518-021-01299-3

Su, G., Peng, L., & Hu, L. (2017). A Gaussian process-based dynamic surrogate model for complex engineering structural reliability analysis. *Structural Safety*, *68*, 97–109. https://doi.org/10.1016/j.strusafe.2017.06.003

Su, L., Wan, H.-P., Dong, Y., Frangopol, D. M., & Ling, X.-Z. (2021). Efficient Uncertainty Quantification of Wharf Structures under Seismic Scenarios Using Gaussian Process Surrogate Model. *Journal of Earthquake Engineering*, *25*(1), 117–138. https://doi.org/10.1080/13632469.2018.1507955

Sudret, B. (2008). Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety*, *93*(7), 964–979. https://doi.org/10.1016/j.ress.2007.04.002

Thai, H.-T. (2022). Machine learning for structural engineering: A state-of-the-art review. *Structures*, *38*, 448–491. https://doi.org/10.1016/j.istruc.2022.02.003

Vats, D., Flegal, J. M., & Jones, G. L. (2019). Multivariate output analysis for Markov chain Monte Carlo. *Biometrika*, *106*(2), 321–337. https://doi.org/10.1093/biomet/asz002

Vintzēleou, E. N., & Tassios, T. P. (1986). Mathematical models for dowel action under monotonic and cyclic conditions. *Magazine of Concrete Research*, *38*(134), 13–22. https://doi.org/10.1680/macr.1986.38.134.13

Wang, Y., Zheng, Z., Ji, D., Pan, X., & Tian, A. (2023). Machine learning-driven probabilistic seismic demand model with multiple intensity measures and applicability in seismic fragility analysis for nuclear power plants. *Soil Dynamics and Earthquake Engineering*, *171*(1), 107966. https://doi.org/10.1016/j.soildyn.2023.107966

Wei, B., Tan, H., Fu, Y., & Jiang, L. (2023). Influence of collision between Friction Pendulum System and shear key on displacement response. *Engineering Structures*, *296*(2), 116964. https://doi.org/10.1016/j.engstruct.2023.116964

Wilches, J., Santa María, H., Riddell, R., & Arrate, C. (2017). Influence of the use of external shear keys on the seismic behavior of Chilean highway bridges. *Engineering Structures*, *147*, 613–624. https://doi.org/10.1016/j.engstruct.2017.06.015

Wolfe, P. (1969). Convergence Conditions for Ascent Methods. *SIAM Review*, *11*(2), 226–235. https://doi.org/10.1137/1011036

Wu, J., Chen, X.-Y., Zhang, H., Xiong, L.-D., Lei, H., & Deng, S.-H. (2019). Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimization. *Journal of Electronic Science and Technology*, *17*(1), 26–40. https://doi.org/10.11989/JEST.1674-862X.80904120

Xiang, N., & Li, J. (2018). Effect of exterior concrete shear keys on the seismic performance of laminated rubber bearing-supported highway bridges in China. *Soil Dynamics and Earthquake Engineering*, *112*, 185–197. https://doi.org/10.1016/j.soildyn.2018.04.033

Xiao, Y., Yue, F., & Zhang, X. (2021). Seismic Fragility Analysis of Structures Based on Adaptive Gaussian Process Regression Metamodel. *Shock and Vibration*, *2021*(1), Article 7622130. https://doi.org/10.1155/2021/7622130

Xie, Y. (2024). *Deep Learning in Earthquake Engineering: A Comprehensive Review*. http://arxiv.org/pdf/2405.09021

Xie, Y., & DesRoches, R. (2019). Sensitivity of seismic demands and fragility estimates of a typical California highway bridge to uncertainties in its soil-structure interaction

modeling. *Engineering Structures*, *189*, 605–617. https://doi.org/10.1016/j.engstruct.2019.03.115

Xie, Y., Ebad Sichani, M., Padgett, J. E., & DesRoches, R. (2020). The promise of implementing machine learning in earthquake engineering: A state-of-the-art review. *Earthquake Spectra*, *36*(4), 1769–1801. https://doi.org/10.1177/8755293020919419

Xing, L., Gardoni, P., & Zhou, Y. (2022). Kriging metamodels for the dynamic response of high-rise buildings with outrigger systems and fragility estimates for seismic and wind loads. *Resilient Cities and Structures*, *1*(1), 110–122. https://doi.org/10.1016/j.rcns.2022.04.003

Yan, Y., Huang, H., & Sun, L. (2022). Multivariate structural seismic fragility analysis and comparative study based on moment estimation surrogate model and Gaussian copula function. *Engineering Structures*, *262*(2), 114324. https://doi.org/10.1016/j.engstruct.2022.114324

Yan, Y., Xia, Y., & Sun, L. (2024). Efficient sensitivity analysis for structural seismic fragility assessment based on surrogate models. *Structures*, *69*(3), 107299. https://doi.org/10.1016/j.istruc.2024.107299

Yazdani, A., Shahidzadeh, M.-S., & Takada, T. (2020). Bayesian networks for disaggregation of structural reliability. *Structural Safety*, *82*(4), 101892. https://doi.org/10.1016/j.strusafe.2019.101892

Yin, K., Xu, Y., Li, J., & Zhou, X. (2024). Gaussian process regression driven rapid life-cycle based seismic fragility and risk assessment of laminated rubber bearings supported highway bridges subjected to multiple uncertainty sources. *Engineering Structures*, *316*(1), 118615. https://doi.org/10.1016/j.engstruct.2024.118615

Yue, K., Xu, L., Fan, L., Liu, J., & Luo, H. (2023). Effect of Shear Keys on the Quasi-Isolated Behavior of Small-to-Medium-Span Girder Bridges. *Buildings*, *13*(9), 2246. https://doi.org/10.3390/buildings13092246

Zhang, Y., & Wu, G. (2019). Seismic Vulnerability Analysis of RC Bridges Based on Kriging Model. *Journal of Earthquake Engineering*, *23*(2), 242–260. https://doi.org/10.1080/13632469.2017.1323040

Zhou, T., & Li, A.-Q. (2019). Seismic fragility assessment of highway bridges using D-vine copulas. *Bulletin of Earthquake Engineering*, *17*(2), 927–955. https://doi.org/10.1007/s10518-018-0474-x

7 APPENDIX: LIST OF PUBLICATIONS

The research presented in this thesis led to the following Web of Science (WoS) publications:

Pinto, F., Toledo, J., **Birrell, M.**, Bazáez, R., Hernández, F., Astroza, R. (2023). "Uncertainty quantification in constitutive models of highway bridge components: seismic bars and elastomeric bearings". *Materials, 16(5)*. doi: 10.3390/ma16051792

Birrell, M., Astroza, R., Murcia-Delso, J., Hernández, F., Bazáez, R. (2024). "A simplified tri-linear model for monolithic exterior shear keys failing in sliding shear". *Structures, 61(14).* doi: 10.1016/j.istruc.2024.105934

Birrell, M., Li, Y., Astroza, R. (2025). "A Gaussian Process surrogate approach for analyzing parameter uncertainty in mechanics-based structural finite element models". *Engineering Structures, 336(6).* doi: 10.1016/j.engstruct.2025.120435

The following article was written before the start of this doctoral research. However, it was published during the first year of this project, and is included in this list for references because the calibrated constitutive model was used throughout this research:

Birrell, M., Astroza, R., Carreño, R., Restrepo, J., Araya-Letelier, G. (2021). "Bayesian parameter and joint probability distribution estimation for a hysteretic constitutive model of reinforcing steel". *Structural Safety, 90(16)*. doi: 10.1016/j.strusafe.2020.102062

Additionally, the following articles were the product of collaborations carried out during the doctoral program related to probabilistic model characterizations, although not included in this document:

Birrell, M., Pastén, C., Abell, J., Astroza, R. (2022). "Probabilistic characterization of a high-cycle accumulation model for sands". *Computers and Geotechnics, 147(6).* doi: 10.1016/j.compgeo.2022.104798

Murcia-Delso, J., **Birrell, M.**, Astroza, R., Carrillo, J. (2024). "Tensile strength and slip model for steel reinforcement anchorages and lap splices". *Structures, 66(4).* doi: 10.1016/j.istruc.2024.106808

Pinto, F., Torres, C., **Birrell., M.**, Li, Y., Fayaz, J., Astroza, R. (2025). "Probabilistic characterization of inherent and epistemic geotechnical uncertainty in soil constitutive models using polynomial chaos expansion and monotonic drained triaxial tests". *Computers and Geotechnics*, *(Accepted)*.

Finally, a paper detailing the work related to Application II (section 4.4.4) has been submitted to Structural Safety and is currently under review:

Birrell, M., Li, Y., Astroza, R., Araya-Letelier, G. (2025). "A Gaussian Process surrogate approach for uncertainty propagation and sensitivity analysis of seismic fragility". *Structural Safety, (Under Review).*