Emulating Stochastic Processes for Efficient Reliability-Based Design Optimisation

A Literature Review and Project Proposal

Course Code: APL871
Course Title: Product Reliability and Maintenence

Department of Applied Mechanics
Indian Institute of Technology Delhi

Submitted by: Karan Noor (2022AM11220)
Rishabh Jain (2022AM11793)

Professor: Prof. Souvik Chakraborty

TA Advisor: Tushar Tyagi

September 5, 2025

Contents

Chapter 1

Background and Motivation

1.1 Introduction to Reliability-Based Design Optimisation (RBDO)

1.1.1 Deterministic vs. Probabilistic Design

In traditional engineering design, systems are often optimized based on deterministic models, where all input parameters such as material properties, geometric dimensions, and applied loads are assumed to be fixed, known values. This approach, while computationally convenient, can lead to designs that are suboptimal or even unsafe in real-world applications [13]. Deterministic optimization tends to push designs to the limits of their performance constraints, leaving little margin for the inevitable variabilities and uncertainties inherent in manufacturing, material composition, and operational environments [8, 13]. A design that is "optimal" under nominal conditions may exhibit a high probability of failure when subjected to even minor, unmodeled fluctuations [13].

Reliability-Based Design Optimisation (RBDO) emerges as a powerful paradigm to address this shortcoming by formally integrating uncertainty into the design process [6, 8]. RBDO acknowledges that many design parameters are not fixed constants but are better described by probability distributions. These uncertainties can be broadly categorized as **aleatory** (inherent randomness, e.g., material property variations) and **epistemic** (lack of knowledge, e.g., model form error). By modeling these uncertainties using probability theory and statistics, RBDO seeks to identify designs that not only meet performance targets but also satisfy a prescribed level of reliability or, equivalently, an acceptably low probability of failure [13]. This probabilistic approach provides a more robust and realistic framework for engineering design, ensuring that structures and systems perform safely and effectively under real-world conditions.

1.1.2 Mathematical Formulation of the RBDO Problem

The RBDO problem is formally expressed as a constrained optimization problem that seeks to minimize an objective function while satisfying one or more probabilistic constraints. The general mathematical formulation is as follows [8, 13]:

$$\min_{\mathbf{d}} F(\mathbf{d}, \boldsymbol{\mu}_{\boldsymbol{X}})$$

subject to:

$$P_f(\mathbf{d}) = P[g(\mathbf{d}, \mathbf{X}) \leq 0] \leq P_{f,target}$$
$$\mathbf{d}_L \leq \mathbf{d} \leq \mathbf{d}_U$$

Here, $F(\mathbf{d}, \boldsymbol{\mu}_{\boldsymbol{X}})$ is the objective function, which typically represents cost, weight, or some other performance metric to be minimized. The vector $\mathbf{d}$ contains the deterministic design variables that the engineer controls (e.g., geometric dimensions). The vector $\mathbf{X}$ represents the set of random variables, which are not under the designer's control and are characterized by their probability distributions (e.g., material strength, applied loads); $\boldsymbol{\mu}_{\boldsymbol{X}}$ denotes their mean values [13].

The core of the RBDO formulation lies in the probabilistic constraint, $P_f(\mathbf{d}) \leq P_{f,target}$. The term $P_f(\mathbf{d})$ is the probability of failure for a given design $\mathbf{d}$. It is defined as the probability that the limit state function, $g(\mathbf{d}, \mathbf{X})$, is less than or equal to zero. The limit state function is a critical mathematical construct that partitions the space of random variables into safe ($g > 0$) and failure ($g \leq 0$) domains [7]. The term $P_{f,target}$ is the maximum allowable probability of failure specified by the designer or by regulatory codes. Finally, $\mathbf{d}_L$ and $\mathbf{d}_U$ represent the lower and upper bounds on the design variables.

1.1.3 The Trade-off between Safety and Cost

At its heart, RBDO is a methodology for navigating the inherent conflict between competing design objectives, most notably the trade-off between economic efficiency (cost, weight) and structural safety (reliability) [13]. A purely deterministic optimization might yield a design that is exceptionally lightweight or inexpensive, but this is often achieved by operating at the very edge of the performance constraints, leaving no buffer against uncertainty [13]. Conversely, achieving a very high level of reliability (a very low $P_{f,target}$) typically requires a more conservative design—for instance, using thicker components or higher-grade materials, which directly increases cost and weight. RBDO provides the formal, quantitative framework to find the "best compromise" by explicitly quantifying the probability of failure, allowing designers to make informed decisions about how much cost they are willing to incur to achieve a specific, measurable target of reliability [6, 8, 13].

Chapter 2

A Review of Stochastic Emulation Methods

2.1 Parametric Emulation: Generalized Lambda Models (GLaM)

2.1.1 The Generalized Lambda Distribution (GLD)

The Generalized Lambda Model (GLaM) approach, developed by Zhu and Sudret, provides a powerful parametric framework for emulating stochastic simulators [14]. The foundational assumption of GLaM is that the conditional response distribution of the simulator, $Y|X = \mathbf{x}$, can be well-approximated by a member of the Generalized Lambda Distribution (GLD) family. The GLD is a highly flexible, four-parameter family of distributions defined not by its PDF, but by its quantile function, $Q(u; \boldsymbol{\lambda})$, where $u \in [0, 1]$ is a probability level and $\boldsymbol{\lambda} = \{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$ is the vector of parameters [14]. A common form is the Freimer-Mudholkar-Kollia-Lin (FMKL) parameterization:

$$Q(u; \boldsymbol{\lambda}) = \lambda_1 + \frac{1}{\lambda_2} \left(\frac{u^{\lambda_3} - 1}{\lambda_3} - \frac{(1-u)^{\lambda_4} - 1}{\lambda_4} \right)$$

The four parameters govern the distribution's characteristics: λ_1 is a location parameter, $\lambda_2 > 0$ is a scale parameter, and λ_3 and λ_4 are shape parameters that control the skewness and kurtosis of the distribution [14]. The great advantage of the GLD is its ability to represent a wide variety of unimodal distribution shapes.

2.1.2 Parameter Dependence using PCE

The core innovation of the GLaM framework is to model the dependence of the output distribution on the simulator's input variables, $\mathbf{x}$. This is achieved by treating the four GLD parameters not as constants, but as functions of $\mathbf{x}$: $\boldsymbol{\lambda}(\mathbf{x}) = \{\lambda_1(\mathbf{x}), \lambda_2(\mathbf{x}), \lambda_3(\mathbf{x}), \lambda_4(\mathbf{x})\}$ [14]. To approximate these unknown functions, GLaM employs Polynomial Chaos Expansions (PCE). Each parameter function $\lambda_l(\mathbf{x})$ is thus represented as:

$$\lambda_l(\mathbf{x}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}_l} c_{l,\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\mathbf{x})$$

where $\Psi_{\boldsymbol{\alpha}}(\mathbf{x})$ are multivariate orthogonal polynomials, $c_{l,\boldsymbol{\alpha}}$ are the unknown coefficients, and $\mathcal{A}_l$ is a set of multi-indices defining the basis [14].

2.1.3 The Replication-Free Approach

Perhaps the most significant contribution of the GLaM methodology is its ability to be trained without requiring replicated runs of the simulator at each point in the experimental design [14]. A novel statistical fitting procedure allows the coefficients of the PCEs for all four λ parameters to be estimated simultaneously using a dataset where each input point $\mathbf{x}^{(i)}$ has only a single corresponding output realization $y^{(i)}$. This replication-free paradigm represents a breakthrough in efficiency, as for a fixed total number of simulator runs, one can sample the input space much more broadly and sparsely.

2.1.4 Performance and Limitations

The GLaM approach has been shown to be highly effective and data-efficient [14]. It can provide semi-analytical expressions for the PDF and CDF, which is valuable for reliability analysis. However, its primary limitation stems from its foundational assumption. The standard four-parameter GLD is inherently unimodal. Therefore, the GLaM framework is not well-suited for emulating simulators that produce multimodal or other topologically complex response distributions.

2.2 Non-Parametric Emulation with Gaussian Processes

An alternative to parametric models is the use of Gaussian Processes (GPs), also known as Kriging, which offer a non-parametric and flexible framework. For stochastic simulators, standard Kriging is insufficient as it assumes a deterministic response. An extension, known as **heteroscedastic GP modeling**, addresses this by creating two separate GPs: one to model the mean of the stochastic response and another to model its variance. Another advanced technique is **Polynomial-Chaos-based Kriging (PC-Kriging)**, which combines the spectral decomposition of PCE with the local interpolation accuracy of Kriging [11]. This hybrid method builds a GP model on the residual of a PCE approximation, effectively capturing both global trends and local variations in the response with high efficiency.

2.3 Non-Parametric Emulation: Deep Generative Models

2.3.1 Generative Neural Networks

A more recent approach to stochastic emulation is to use the power of deep learning. In the work by Thakura and Chakraborty, a generative neural network is proposed to directly approximate the stochastic response of the simulator [12]. Instead of assuming a parametric form, this method leverages the universal approximation capabilities of neural networks to learn the complex, non-linear mapping from the input parameters $\mathbf{x}$ to the output distribution $P(Y|X = \mathbf{x})$.

2.3.2 The CMMD Loss Function

A central challenge in training such a model is defining an appropriate loss function. Thakura and Chakraborty use the **Conditional Maximum Mean Discrepancy (CMMD)**, a non-parametric metric that measures the distance between two probability distributions [12]. The CMMD loss allows the training algorithm to directly minimize the discrepancy between the distribution of samples generated by the neural network and the true samples from the simulator, without making any assumptions about their shape [12].

2.4 Application: Surrogate-Assisted RBDO

The emulation techniques described previously are not merely academic exercises; they are crucial enablers for complex engineering tasks. A prime example is their application in **Reliability-Based Design Optimization (RBDO)**, a field dedicated to finding optimal engineering designs while explicitly accounting for real-world uncertainties [10].

2.4.1 The Challenge of RBDO

The core RBDO problem is to minimize a cost function (e.g., material usage, weight) subject to constraints on the probability of failure. This creates a computationally demanding, nested problem: an outer loop searches for optimal design parameters, and for each candidate design, an inner loop must perform a full reliability analysis to compute its failure probability [10]. Classical approaches that rely on direct simulation or first-order reliability methods (FORM) are often too slow for complex, real-world models or can be inaccurate for highly non-linear systems [10].

2.4.2 A Unified, Modular Framework

To address these challenges, Moustapha and Sudret proposed a unified and modular framework for surrogate-assisted RBDO [10]. Their key insight was to structure the problem into three independent, non-intrusive blocks:

1. **Surrogate Modeling:** The expensive-to-evaluate performance function of the system is replaced by a cheap-to-evaluate surrogate model, such as Kriging or Support Vector Regression (SVR) [10]. This surrogate is built adaptively over an "augmented space" that spans both the design parameters and the random environmental variables.

2. **Reliability Analysis:** With an efficient surrogate in place, the failure probability for any given design can be rapidly estimated using simulation-based methods like crude Monte Carlo or more advanced techniques like Subset Simulation [10].

3. **Optimization:** A suitable optimization algorithm, whether a global search method like CMA-ES or a local gradient-based method like SQP, can then efficiently explore the design space using the fast reliability estimates from the surrogate [10].

This modularity is a significant advantage, as it allows practitioners to mix and match the best available techniques for each part of the problem without needing to alter the other components [10]. The framework demonstrates how advanced surrogate modeling is not just a tool for analysis, but a pivotal technology for enabling robust and efficient engineering design under uncertainty.

2.5 Comparison of Emulation Strategies

The primary methods for stochastic emulation embody a fundamental trade-off between embedded structural knowledge and data-driven flexibility. Parametric methods like **GLaM** embed a strong structural assumption (e.g., unimodality via the GLD), which acts as a powerful regularizer, making them exceptionally data-efficient and suitable for training on sparse, replication-free datasets [14]. In contrast, non-parametric approaches make minimal assumptions about the output distribution's shape. Methods based on **Gaussian Processes**, such as PC-Kriging, are highly effective but traditionally require specialized techniques to handle the intrinsic variance of the simulator [11]. **Deep generative models** trained with a distribution-agnostic metric like CMMD offer the greatest flexibility, capable of capturing highly complex phenomena like multimodality that would challenge other methods [12]. This flexibility, however, typically comes at the cost of requiring larger datasets, often with replications, to effectively train the network and avoid overfitting.

The application of these surrogates in a field like **Reliability-Based Design Optimization (RBDO)** highlights their practical importance. The framework proposed by Moustapha and Sudret demonstrates that by replacing the true, computationally expensive model with a surrogate (like Kriging or SVR), the otherwise intractable nested loop of optimization and reliability analysis becomes feasible [**?**]. This shows that the choice of emulator ultimately depends on a balance between the available computational budget, prior knowledge of the simulator's behavior, and the demands of the engineering task at hand.

Chapter 3

Research Gap and Project Proposal

3.1 Research Gap: RBDO with Stochastic Performance Functions

3.1.1 The Untapped Potential of Stochastic Emulators in RBDO

A thorough review of the literature reveals two parallel but largely disconnected streams of research. The RBDO community has developed sophisticated algorithms and surrogate modeling techniques for computationally expensive but deterministic performance functions [1, 2, 10]. Concurrently, the uncertainty quantification community has developed powerful stochastic emulators for sensitivity analysis and forward UQ [14]. The clear research gap lies at the intersection of these two fields: the application of modern stochastic emulators to solve RBDO problems where the limit state function itself is defined by a stochastic simulator.

3.1.2 A New Framework using Stochastic Emulators

The integration of a stochastic emulator into the RBDO workflow represents a paradigm shift. Instead of repeatedly performing an expensive reliability analysis inside the optimization loop, the new approach consists of two phases:

1. **Offline Training Phase:** A stochastic emulator (e.g., a GLaM) is trained once on data generated from the expensive stochastic simulator.

2. **Online Optimisation Phase:** The trained, fast-to-evaluate emulator is used inside the RBDO loop. When queried with a design point $\mathbf{d}$, the emulator acts as a "probabilistic oracle," returning the parameters that define the entire conditional probability distribution of the performance function, $P(g|\mathbf{d}, \mathbf{X})$ [14]. From these parameters, the conditional CDF, $F_g(y|\mathbf{d}, \mathbf{x})$, can be computed almost instantaneously. The failure probability for that design is then simply $F_g(0|\mathbf{d}, \mathbf{x})$. This replaces the entire costly inner reliability analysis loop with a single, fast function call.

3.2 Project Proposal: Emulator-Assisted RBDO

3.2.1 Project Objectives and High-Level Workflow

Primary Objective: To develop and demonstrate a proof-of-concept RBDO framework where a system's performance is described by a stochastic simulator, with the computational cost managed by a GLaM-based stochastic emulator.

High-Level Workflow:

1. **Emulator Implementation and Validation:** Implement a GLaM emulator and validate its performance against a benchmark stochastic process with a known analytical solution.

2. **Stochastic Structural Simulator:** Define and implement a stochastic simulator for a representative structural component where the limit state function includes intrinsic randomness.

3. **Emulator Training & RBDO Integration:** Train the validated GLaM emulator on data from the simulator and integrate it into a single-loop RBDO algorithm.

4. **Demonstration:** Solve a set of benchmark RBDO problems using the emulator-assisted framework to demonstrate its efficacy and analyze the results.

3.2.2 Phase 1: Emulator Implementation and Initial Validation

The Black-Scholes model for option pricing provides an ideal initial benchmark for validating the emulator [3]. The price of the underlying asset, S_t, follows a geometric Brownian motion, a stochastic process whose output distribution is known analytically to be lognormal. A numerical solver for the governing SDE will serve as the "stochastic simulator." The trained GLaM's predictions for the mean, variance, and CDF will be compared against the analytical lognormal solution to confirm the correct implementation and quantify its accuracy.

3.2.3 Phase 2: Application to a Benchmark Engineering Problem

As a representative engineering problem, the project will address the reliability of a simply supported I-beam of length L subjected to a point load P [4]. To make the problem stochastic, we introduce an intrinsic model uncertainty term, $\epsilon(\omega)$, into the limit state function to represent unmodeled effects or numerical noise. The stochastic limit state function is then defined as:

$$g(\mathbf{d}, \mathbf{X}, \omega) = \sigma_y - \frac{P \cdot L \cdot h}{8 \cdot I(\mathbf{d})} + \epsilon(\omega)$$

The simulator for this phase will compute the deterministic part of g and add a random realization of $\epsilon(\omega)$ (e.g., sampled from a zero-mean normal distribution, $\mathcal{N}(0, \sigma^2_{\text{model}})$). The detailed problem parameters are specified in Table 3.1.

3.2.4 Phase 3: Framework Integration and Demonstration

A concrete RBDO problem will be formulated for the beam: minimize the cross-sectional area, $A(h, b)$, subject to a probabilistic constraint on failure. The formal problem is:

$$\min_{h,b} A(h, b)$$

subject to:

$$P_f(h,b) = P[g(h,b,P,\sigma_y,\omega) \le 0] \le 10^{-3}$$
$$h_{min} \le h \le h_{max}, \quad b_{min} \le b \le b_{max}$$

The GLaM emulator will be used to evaluate the probabilistic constraint. The overall failure probability for a given design $\mathbf{d} = \{h, b\}$ is the expectation of the conditional failure probability (provided by the emulator) over the distribution of the random variables $\mathbf{X} = \{P, \sigma_y\}$:

$$P_f(\mathbf{d}) = \mathbb{E}_{\mathbf{X}}[P(g \le 0|\mathbf{d},\mathbf{X})] = \int P(g \le 0|\mathbf{d},\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}$$

The term $P(g \le 0|\mathbf{d},\mathbf{x})$ is obtained directly from the emulator's conditional CDF, $F_{\text{GLaM}}(0|\mathbf{d},\mathbf{x})$. Because the GLaM evaluation is computationally trivial, this integral can be efficiently estimated using Monte Carlo simulation. A standard gradient-based optimization algorithm (e.g., SQP) will then be used to solve the RBDO problem.

Table 3.1: Definition of the Simply Supported Beam Reliability Problem

Category	**Parameter Description and Distribution**
Problem Definition **Failure**	Simply supported I-beam, length L, point load P. Bending stress $>$ Yield strength.
Design Variables (d)	h: Height of the beam's web (to be optimized). b: Width of the beam's flanges (to be optimized).
Random Variables (X)	P: Point Load. Lognormal distribution. σ_y: Material Yield Strength. Lognormal distribution.
Deterministic Params	L: Length of the beam. t_w, t_f: Thickness of web and flanges. $I(\mathbf{d})$: Moment of inertia, a function of design variables.
Limit State Function	$g(\mathbf{d},\mathbf{X},\omega) = \sigma_y - \frac{M(\mathbf{d},P)}{S(\mathbf{d})} + \epsilon(\omega) = \sigma_y - \frac{P \cdot L \cdot h}{8 \cdot I(\mathbf{d})} + \epsilon(\omega)$ $\epsilon(\omega)$: Stochastic noise term, $\mathcal{N}(0, \sigma^2_{\text{model}})$.

3.2.5 Scope, Simplifications, and Expected Outcomes

Scope: The project is constrained to time-independent, component-level reliability to ensure feasibility. The focus is on the successful proof-of-concept integration of the emulator into the RBDO loop.

Expected Outcomes: The project will deliver:

1. A functional codebase implementing the entire workflow.
2. A final report detailing the results of each phase, including the quantitative validation of the GLaM emulator and characterization of its performance on the engineering problem.
3. The final optimal, reliable designs obtained from the emulator-assisted RBDO framework for one or more benchmark problems, highlighting the value of the probabilistic approach compared to a deterministic one.

Bibliography

[1] Y. Aoues and A. Chateauneuf. Benchmark study of numerical methods for reliability-based design optimization. *Structural and Multidisciplinary Optimization*, 41:277-294, 2010.

[2] B. J. Bichon, M. S. Eldred, L. P. Swiler, S. Mahadevan, and J. M. McFarland. Reliability-based design optimization using kriging surrogates and subset simulation. *Structural and Multidisciplinary Optimization*, 44:699-714, 2011.

[3] F. Black and M. S. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637-654, 1973.

[4] M. T. Devi, T. S. U. Maheshwari, and A. Satyanarayana. Reliability index of simply supported beam based on hl method. *Journal of Reliability and Statistical Studies*, 15(2):745-758, 2022.

[5] X. Du and W. Chen. Sequential optimization and reliability assessment method for efficient probabilistic design. *ASME Journal of Mechanical Design*, 126(2):225-233, 2004.

[6] B. Guan, M. Wan, and X. Wu. Reliability-based design optimization of structures with disturbed design variables and dependent uncertainties. *Journal of Physics: Conference Series*, 2239, 012018, 2022.

[7] A. M. Hasofer and N. C. Lind. An exact and invariant first order reliability format. *Journal of the Engineering Mechanics Division*, 100(1):111-121, 1974.

[8] X. Lai, J. Huang, Y. Zhang, C. Wang, and X. Zhang. A general methodology for reliability-based robust design optimization of computation-intensive engineering problems. *Journal of Computational Design and Engineering*, 9(5):2151-2169, 2022.

[9] X. Li, Q. Yang, Y. Wang, X. Han, Y. Cao, L. Fan, and J. Ma. Development of surrogate models in reliability-based design optimization: A review. *Mathematical Biosciences and Engineering*, 18(5):6386-6409, 2021.

[10] M. Moustapha and B. Sudret. Surrogate-assisted reliability-based design optimization: a survey and a unified modular framework. *Structural and Multidisciplinary Optimization*, 60(5):2157-2176, 2019.

[11] R. Schöbi, B. Sudret, and J. Wiart. Polynomial-chaos-based Kriging. *International Journal for Uncertainty Quantification*, 5(5):443-469, 2015.

[12] A. Thakura and S. Chakraborty. A deep learning based surrogate model for stochastic simulators. *arXiv preprint arXiv:2110.13809*, 2021.

[13] D. Zhu, J. Zhou, C. Liu, and Z. Wang. A short review of reliability-based design optimization. *IOP Conference Series: Materials Science and Engineering*, 1043, 032041, 2021.

[14] X. Zhu and B. Sudret. Emulation of stochastic simulators using generalized lambda models. *SIAM/ASA Journal on Uncertainty Quantification*, 9(4):1345-1380, 2021.