Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference
Edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Bouder, Roger Flage, Marja Ylönen

doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P0589-cd

The NASA and DNV Challenge on Optimization Under Uncertainty

Christian Agrell
Group Research and Development, DNV, Norway. E-mail: christian.agrell@dnv.com

Luis G. Crespo
NASA Langley Research Center. E-mail: luis.g.crespo@nasa.gov

Vegard Flovik
Group Research and Development, DNV, Norway. E-mail: vegard.flovik@dnv.com

Erik Vanem
Energy Systems, DNV, Norway. E-mail: erik.vanem@dnv.com

Sean Kenny
NASA Langley Research Center. E-mail: sean.kenny@nasa.gov

This paper presents an Uncertainty Quantification (UQ) challenge focusing on key aspects of uncertainty quantification and optimization-based design in the presence of aleatory and epistemic uncertainty.

Keywords: Uncertainty Quantification, Optimization

1. Introduction

NASA missions often involve the development of new vehicles and systems that must be designed to operate in harsh domains with a wide array of operating conditions. Similarly, DNV works to safeguard life, property, and the environment across a range of industries by ensuring that complex engineered systems can be trusted to operate reliably, even during rare and extreme events. DNV collaborates across various sectors to address the challenges of designing and operating safety-critical systems. Both NASA and DNV deal with high-consequence and safety-critical systems for which quantitative data is either very sparse or prohibitively expensive to collect. Limited heritage data may exist, but is also usually sparse and may not be directly applicable to the system of interest, making UQ extremely challenging.

Recognizing the shared challenges in UQ across different sectors, NASA Langley Research Center and DNV Group Research and Development have developed a challenge problem to unite a community of researchers toward common goals. While the problem formulation is presented in a discipline-independent framework, the underlying application is consistent with the complexities of realistic systems. This collaborative effort aims to advance methodologies in UQ that are broadly applicable, enhancing safety and reliability across various domains.

2. Uncertainty Quantification

This challenge problem adopts the generally accepted classification of uncertainty referred to as aleatory and epistemic Agrell et al. (2023); Crespo and Kenny (2014); Roy and Oberkampf (2011). *Aleatory* uncertainty (also called irreducible or stochastic) is caused by inherent variation or randomness. As such, aleatory parameters are often modeled as random variables. In contrast, *epistemic* uncertainty is caused by lack of knowledge in the true value of a parameter. Therefore, epistemic uncertainty is not an inherent property of the system, but instead it represents the state of knowledge of the analyst. In the context of this challenge, an epistemic variable can take on any fixed value within a bounded set, and an aleatory variable is represented by a probability distribution.

The physical origin of a parameter as well as the knowledge we have about it must be used to create an Uncertainty Model (UM) for it. Intervals, fuzzy sets, random variables, probability boxes Ferson et al. (2003), etc. are commonly used classes of UMs.

Because most models — especially those characterizing uncertainty — are imperfect, the possibility of improving them always exists. A reduction of the uncertainty in an epistemic variable is attained by reducing the size of the set where the true value of such a variable is expected to be. This reduction can be attained by performing additional physical or computational experiments.

3. Framework

The general problem framework is described here, and specific details like the number of dimensions, datapoints, etc. will be available from the simulation model and data.

We consider a physical system whose state is represented by an input vector $\mathbf{X}$ and a corresponding multivariate time series $\mathbf{Y}$. There will be uncertainty associated with both the input $\mathbf{X}$ and the conditional response $\mathbf{Y}|\mathbf{X}$.

3.1. *Input X*

The input $\mathbf{X}$ contains parameters that describe a physical structure and its environment. Typically $\mathbf{X} \in \mathbb{R}^{n_x}$ with dimensionality in the order $n_x \approx 10$. For this challenge problem, the relevant range of $\mathbf{X}$ is standardized and we let $\mathbf{X} \in [0,1]^{n_x}$. The components of $\mathbf{X}$ have different characteristics, and we will write $\mathbf{X} = (\mathbf{X}_a, \mathbf{X}_e, \mathbf{X}_c)$ as the combination of three vectors where

- $\mathbf{X}_a \in \mathbb{R}^{n_a}$ is an aleatory variable whose value varies randomly.
- $\mathbf{X}_e \in \mathbb{R}^{n_e}$ is the epistemic variable whose true value $\mathbf{X}_e^\star$ is unknown.
- $\mathbf{X}_c \in \mathbb{R}^{n_c}$ is the control variable whose value is to be set by the analyst.

3.2. *Response Y*

The time-varying response of the system $\mathbf{Y}$ depends on the input $\mathbf{X}$ and an additional aleatory variable ω. Thus for any fixed $\mathbf{X} = \mathbf{x}$, the response $\mathbf{Y}_t(\mathbf{x}) = \mathbf{Y}(\mathbf{x}, t)$ is a stochastic process. In particular, $\mathbf{Y}$ for any given input $\mathbf{x}$ and ω is a multivariate time series. Formally, we can write the response model as

$$\mathbf{Y}_t(\mathbf{X}, \omega) = \mathbf{Y}(\mathbf{X}, \omega, t) = [Y_1(\mathbf{X}, \omega, t), \ldots, Y_{n_y}(\mathbf{X}, \omega, t)], \tag{1}$$

where $Y_i \in \mathbb{R}$ for $i = 1, \ldots, n_y$, $\mathbf{X} \in [0,1]^{n_x}$, $t \in [0,1]$, and ω belongs to a probability space $(\Omega, \mathcal{F}, P_\omega)$. We assume that ω and $\mathbf{X}_a$ are independent. In the computational model, the dependency of the response on ω is controlled by a seed parameter s. Users will be able to specify both $\mathbf{X}$ and s when evaluating such a model.

The data obtained as output from the computational model, as well as data obtained from the real system, will be n_y time series evaluated at m discrete time steps t_i where $t_1 = 0$ and $t_m = 1$. This data takes the form $\mathcal{Y} = \{y^{(i)}\}_{i=1}^m$, where $y^{(i)} = y(t_i) \in \mathbb{R}^{n_y}$. Thus, $y_j^{(i)}$ for $i = 1, \ldots m$ is the univariate series corresponding to the j-th response.

3.3. *Computational model*

A computational model for simulating the system response having any given $\mathbf{x}$ and s as inputs will be available to the respondents. This model will yield the multivariate time series $\mathcal{Y}(\mathbf{X}_a, \mathbf{X}_e, \mathbf{X}_c, s)$. A software implementation of this model in a compiled binary and instructions on how to run it in Python and MATLAB will be provided.

3.4. *Data from the real system*

The response of the true system is obtained by specifying the control variable $\mathbf{X}_c$ and observing the response $\mathcal{Y}(\mathbf{X}_c)$. In this process, $\mathbf{X}_a$ and s are sampled at random from a distribution that is unknown to the observer, whereas $\mathbf{X}_e$ takes on the true (but unknown) value $\mathbf{X}_e^*$. One observation from the real system therefore corresponds to a random sample $\mathcal{Y}(\mathbf{X}_c) \sim \mathcal{Y}(\mathbf{X}_a, \mathbf{X}_e^*, \mathbf{X}_c, s)$ where $\mathbf{X}_a$ and s are realizations of a random process.

4. Problem formulation

The challenge problem consists of two parts. Problem 1 is about modeling uncertainty in the input $\mathbf{X}$ by using the computational model together with data from the real system. Problem 2 is about optimizing the control of the physical system under this uncertainty, and balancing performance and risk.

4.1. *Tools*

For this challenge the following will be made available:

- The computational model described in 3.3, which will be downloaded and run locally.
- Data from the true system described in 3.4. This will be available online for the respondents who register for the challenge.

The participants will have a total budget of N observations of the real system corresponding to a chosen $\mathbf{X}_c$. The output corresponding to $\mathbf{X}_c$ is given by the K i.i.d. observations in $\mathcal{D}(\mathbf{X}_c) = \{\mathcal{Y}_1(\mathbf{X}_c), \ldots, \mathcal{Y}_K(\mathbf{X}_c)\}$. Therefore, a maximum of $N \times K$ observations will be available. Further information about the models and data are found on the UQ Challenge GitHub page DNV (2025).

4.2. *Problem 1: Uncertainty quantification*

We will provide an initial estimate of the support set of $\mathbf{X}_a$ and of E.

(1) This subproblem seeks to create an UM of $\mathbf{X}_a$ and $\mathbf{X}_e$ from data for a baseline design $\mathbf{X}_c'$. This UM is given by the pair (f_a, E), where

- f_a is a probability density for $\mathbf{X}_a$.
- E is an arbitrarily shaped set satisfying $\mathbf{X}_e^\star \in E$.

For any $\mathbf{X}_c'$ chosen by the respondents, we will provide the dataset $\mathcal{D}(\mathbf{X}_c')$ described above. Estimate the UM based on these data, and evaluate the extent by which the number of data points $N \times K$ impacts your answer. The resulting UM will be denoted as UM1.

(2) **[Optional]** Repeat subproblem one as if you were only going to use the "most informative" 95% of the data. Compare and contrast the resulting UM, denoted as UM2, with UM1.

(3) Denote as $y(\mathbf{X}_a, \mathbf{X}_e, \mathbf{X}_c, s, t)$ a component of $\mathcal{Y}(\mathbf{X}_a, \mathbf{X}_e, \mathbf{X}_c, s)$. The probability of the greatest y exceeding the upper limit u is

$$p_u(\mathbf{X}_e, \mathbf{X}_c, u) = \mathbb{P}\left[\max_{0\le t\le 1} y(\mathbf{X}_a, \mathbf{X}_e, \mathbf{X}_c, s, t) > u\right], \tag{2}$$

whereas the probability of the lowest y falling below the lower limit l is

$$p_l(\mathbf{X}_e, \mathbf{X}_c, l) = \mathbb{P}\left[\min_{0\le t\le 1} y(\mathbf{X}_a, \mathbf{X}_e, \mathbf{X}_c, s, t) < l\right], \tag{3}$$

where $\mathbb{P}[\cdot]$ is the probability with respect to $\mathbf{X}_a$ and s. The limits of the prediction interval of the response[a] are

$$\bar{y}(\mathbf{X}_c, \alpha) = \inf\left\{ u \mid \max_{\mathbf{X}_e \in E} p_u(\mathbf{X}_e, \mathbf{X}_c, u) \le 1-\alpha \right\}, \tag{4}$$

$$\underline{y}(\mathbf{X}_c, \alpha) = \sup\left\{ l \mid \max_{\mathbf{X}_e \in E} p_l(\mathbf{X}_e, \mathbf{X}_c, l) \le 1-\alpha \right\}. \tag{5}$$

Hence, for any $\mathbf{X}_e \in E$ and any $t \in [0,1]$, the probability that y falling above $\bar{y}(\mathbf{X}_c, \alpha)$ or below $\underline{y}(\mathbf{X}_c, \alpha)$ are both bounded by α. For a given baseline design $\mathbf{X}_c$, the goal of this subproblem is to determine the tightest prediction intervals of all n_y outputs corresponding to the identified uncertainty model(s) for both $\alpha = 0.999$ and $\alpha = 0.95$. Comment on your results.

4.3. *Problem 2: Design optimization*

The index i of the n_y system responses can either be in the set I_1, which prescribes the cost function, or the set I_2, which prescribes the constraints. In this section we seek a control parameter $\mathbf{X}_c$ that is optimal with respect to different objective functions and constraints.

Consider the objective function

$$J(\mathbf{X}_e, \mathbf{X}_c) = \int_0^1 \sum_{i \in I_1} \mathbb{E}\left[y_i(\mathbf{X}_a, \mathbf{X}_e, \mathbf{X}_c, s, t)\right] \, dt, \tag{6}$$

[a] Please note that $\mathbf{X}_e$ can only take the value of a *single* element of E. Take this into consideration when applying the constraint $\mathbf{X}_e \in E$ hereafter.

where $\mathbb{E}[\cdot]$ is the expected value operator with respect to $\mathbf{X}_a$ and s. Hence, this objective function is the expected total contribution from the system outputs in I_1 for the chosen $\mathbf{X}_c$.

For the responses in I_2, define the limit state functions

$$g_i(\mathbf{X}, s) = c_i - \max_{0\le t\le 1} |y_i(\mathbf{X}, s, t)|, \text{ for all } i \in I_2, \tag{7}$$

where the constants c_i are fixed. The event $g_i(\mathbf{X}, s) < 0$ represents a failure in the i-th physical response. The individual failure probability is defined as

$$\mathrm{pof}_i(\mathbf{X}_e, \mathbf{X}_c) = \mathbb{P}[\, g_i(\mathbf{X}_a, \mathbf{X}_e, \mathbf{X}_c, s) < 0\,], \tag{8}$$

where $\mathbb{P}[\cdot]$ is the probability with respect to $\mathbf{X}_a$ and s. The system failure probability is

$$\mathrm{pof}_{\mathrm{sys}}(\mathbf{X}_e, \mathbf{X}_c) = \mathbb{P}[\, \min_{i\in I_2} g_i(\mathbf{X}_a, \mathbf{X}_e, \mathbf{X}_c, s) < 0\,]. \tag{9}$$

Consider the following design optimization problems:

(1) Find the control variable $\mathbf{X}_c$ that maximizes

$$\min_{\mathbf{X}_e\in E} J(\mathbf{X}_e, \mathbf{X}_c). \tag{10}$$

We refer to the resulting $\mathbf{X}_c^\star$ as the performance-based design.

(2) Find the control variable $\mathbf{X}_c$ that minimizes the greatest value that (9) might take as $\mathbf{X}_e$ varies in E. We refer to the resulting $\mathbf{X}_c^\star$ as the reliability-based design.

(3) Find the control variable $\mathbf{X}_c$ that maximizes (10) under the constraint

$$\max_{\mathbf{X}_e\in E} \mathrm{pof}_{sys}(\mathbf{X}_e, \mathbf{X}_c) \le \epsilon, \tag{11}$$

for a specified $\epsilon \in [0, 1]$. We refer to the resulting $\mathbf{X}_c^\star$ as the ϵ-constrained design.

(4) **[Optional]** Find the control variable $\mathbf{X}_c$ that maximizes (10) such that

$$\max_{\mathbf{X}_e\in E} \mathbb{E}[\, h \mid h < 0\,] \ge \beta, \tag{12}$$

where $h = \min_{i\in I_2} g_i(\mathbf{X}_a, \mathbf{X}_e, \mathbf{X}_c, s)$. We refer to the resulting $\mathbf{X}_c^\star$ as the risk-based design.

Solve the above problems for the UMs identified in Problem 1, by using $\epsilon = 10^{-3}$ and $\epsilon = 10^{-4}$ in (11) and $\beta = -300$ in (12). For each of these cases, please provide $\mathbf{X}_c^\star$, the range of the objective function, and the range of the failure probability corresponding to all epistemic values in E.

5. Actions and deliverables

(1) Each group should first register by sending an email to *uqchallenge@dnv.com* that includes the name of all the participants and their affiliation. Details on how to download the model and how to make the submissions below will be given after registration.

(2) Submit the uncertainty models created in Problem 1, the best guess of the true value of $\mathbf{X}_e^*$ (if any), and the prediction intervals for both values of α. To submit an UM, we ask for 10^4 samples from f_a and 10^3 samples generated uniformly within E as arrays stored in separate csv files.

(3) Submit the optimal controls obtained in Problem 2. For each case, report also the minimum and maximum value of the objective function (6) and the system failure probability (9) as $\mathbf{X}_e$ varies in E.

(4) Write a paper for the ESREL 2025 conference summarizing your response to the challenge. The deadline for this submission will be at the end of March 2025. The paper should summarize the response to the above requests, and carefully describe:

- The process used to obtain the UMs in Problem 1. How was the design $\mathbf{X}_c'$ used to collect data from the real system chosen? Can the uncertainty models for the epistemic and aleatory variables be identified separately? Was the amount of data available sufficient?
- The process used to derive the optimal controls in Problem 2. Compare and contrast all such controls. What control should be used in a safety-critical application?

References

Agrell, C., K. R. Dahl, and A. Hafver (2023). Optimal sequential decision making with probabilistic digital twins - theoretical foundations. *SN Applied Sciences 5*(4), 114.

Crespo, L. G. and S. P. Kenny (2014). The NASA Langley multidisciplinary uncertainty quantification challenge. *AIAA 2014-1347*.

DNV (2025). Uq challenge 2025. `https://dnv-opensource.github.io/UQ-Challenge-2025/`.

Ferson, S., V. Kreinovich, L. Ginzburg, D. S. Myers, and K. Sentz (2003). Constructing probability boxes and Dempster-Shafer structures. Technical Report SAND2002-4015, Sandia National Laboratories.

Roy, C. and W. Oberkampf (2011). A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing. *Computer methods in applied mechanics and engineering 200*, 2131–2144.