A global framework for active learning reliability in UQLab

M. Moustapha, S. Marelli and B. Sudret

Chair of Risk, Safety and Uncertainty Quantification
Stefano-Franscini-Platz 5
CH-8093 Zürich

Data Sheet

Journal:	Proc. 13th International Conference on Structural Safety & Reliability, Tongji University, Shanghai (China), September 13-17
Report Ref.:	RSUQ-2022-011
Arxiv Ref.:	
DOI:	-
Date submitted:	December 16, 2021
Date accepted:	September 13, 2022

A global framework for active learning reliability in UQLab

Maliki Moustapha*[1], Stefano Marelli†[1], and Bruno Sudret‡[1]

[1] *Chair of Risk, Safety and Uncertainty Quantification, ETH Zürich, Switzerland*

December 16, 2021

Abstract

Since its introduction in the field of reliability analysis, active learning has been increasingly used for the solution of complex reliability problems at a manageable cost. The basic idea is to adaptively build an accurate approximation of the limit-state surface by sparsely covering the input space. In the early contributions, a metamodel, typically Kriging, was updated through a so-called learning function and then used for the estimation of the failure probability with Monte Carlo simulation. Popular methods include efficient global reliability analysis (EGRA) and active-Kriging Monte Carlo simulation (AK-MCS). More recently, a considerable number of methods that draw on this idea have been proposed by merely modifying one or more of these ingredients. In this contribution, we first conduct a survey of active learning reliability methods available in the literature. We then identify the basic ingredients that make the backbone of these approaches. Drawing on their similarity, we propose a global framework for active learning reliability that combines non-intrusively four different blocks: metamodelling, reliability analysis, learning function and convergence criterion. By wisely choosing each element of the framework, a solution scheme that is tailored to a specific type of problems can be devised, *e.g.* problems with high-dimensional inputs or extremely rare events. Using this framework, an active learning reliability module is implemented in UQLab, a Matlab-based framework for uncertainty quantification. In this paper, we show how such a framework is implemented, thus allowing users to easily build solution strategies by selecting independently each ingredient. The module is tested with 20 different limit state functions, and multiple combinations of the four ingredients, leading to more than $12,000$ reliability analyses. These results are eventually aggregated to provide user-oriented recommendations.

1 Introduction

Uncertainty quantification (UQ) has gained a lot of attention in the past few years thanks to an ever increasing understanding of the role of uncertainties in engineering systems. Both researchers

*moustapha@ibk.baug.ethz.ch

†marelli@ibk.baug.ethz.ch

‡sudret@ethz.ch

and field practitioners are indeed interested in identifying and quantifying the various sources of uncertainty affecting a given system and then in propagating those uncertainties to some quantities of interest through a computational model designed to numerically mimic the behavior of the system [1]. Algorithmic advances on the one hand and the development of dedicated yet general-purpose software on the other hand, have played a crucial role in promoting UQ to a broader use by engineers and researchers. Among those, UQLab is a Matlab-based framework for uncertainty quantification ([2], `www.uqlab.com`). It is a versatile platform offering state-of-the-art UQ tools and, thanks to its modular infrastructure, allows for the rapid integration of new algorithms. Such algorithms are collected into modules whose aim is to carry out a specific UQ task. We focus here on the reliability analysis module.

Generally speaking, reliability analysis primarily aims at calculating the failure probability of a structure given uncertainties associated to its design or environmental parameters. This can be computed as follows:

$$P_f = \int_{\{\mathbf{x}:g(\mathbf{x})\leq 0\}} f_{\mathbf{X}}\left(\mathbf{x}\right) d\mathbf{x}, \tag{1}$$

where $\mathbf{X} \in \mathcal{D}_{\mathbf{X}} \subset \mathbb{R}^M$ is a set of random variables describing the system and following a joint probability distribution $f_{\mathbf{X}}$ and g is the so-called limit-state function describing the system response. It is conventionally assumed that the system is in a failure state when $g(\mathbf{x}) \leq 0$. Equation (1) is a multi-dimensional integration problem over an implicitly defined domain whose solution is but straightforward. The standard solution approach relies on computationally expensive simulation methods such as crude Monte Carlo or other variance-reduction techniques, *e.g.* importance sampling, subset simulation, line sampling, etc. Each of these methods work well for a large, albeit non-overlapping class of problems. They however share a common drawback, which is their computational cost, *i.e.* a large number of evaluations of the limit-state function is necessary to achieve an accurate estimate of the failure probability. This is a serious impediment to the deployment of reliability analysis in real-world applications. In the past decades, metamodels, which can be used as cheap-to-evaluate surrogates of the limit-state function, have been intensively employed in the field of structural reliability. More specifically, adaptive approaches, a.k.a. active learning, have been developed to accurately approximate the vicinity of the limit-state surface using only a limited number of calls to the limit-state function. This is often achieved by the sequential enrichment of the experimental design (ED) so as to enhance the capability of the metamodel to correctly classify the failure/safety states of the system. One of the most popular techniques is the family of so-called AK (active Kriging) methods [3] which rely on using Kriging together with a properly calibrated learning function. The latter is a mathematical function whose goal is to find the best candidate for the improvement of the ED. This was pioneered in structural reliability by the work of Bichon et al. [4] and further popularized by the AK-MCS (active Kriging - Monte Carlo simulation) method [5]. Since then a plethora of similar methods have been investigated in the literature. At their core, such methods draw on the AK-MCS framework and only differ in the choice of ingredients combined to build the active learning scheme. In this paper, we conduct a short literature review and show that such schemes can be built by choosing different methods from a set of four components: metamodelling, learning function, reliability algorithm and convergence criterion. This modular framework has been implemented in UQLab in version 1.4 of the reliability module. Following a literature review in Section 2, Section 3 explains the module implementation and follows with a simple usage scenario. Finally, Section 4 illustrates the use of this new feature for the realization of a large-scale benchmark involving 20 reliability problems and 39 solution schemes, the results of which have been used to design an extremely efficient solution scheme for the TNO black-box reliability challenge [6].

2 Generalized global framework

2.1 Pseudo-algorithm

The active learning methods as published in the early contributions have sketched an out-line for a general framework, which can be summarized as follows:

1. **Initialization:** Defining the initial experimental $\mathcal{X} = \left\{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(N_0)}\right\}$ and the corresponding limit-state evaluations $\mathcal{G} = \left\{g\left(\mathbf{x}^{(1)}\right), \dots, g\left(\mathbf{x}^{(N_0)}\right)\right\}$;
2. **Metamodel construction/update:** Using the current experimental design $\{\mathcal{X}, \mathcal{G}\}$, a metamodel $\widehat{g}$ is built as an approximation of g;
3. **Reliability run:** The failure probability $\widehat{P}_f$ is estimated using an appropriate reliability estimation method (*e.g.* MCS, importance sampling, etc.) together with the metamodel $\widehat{g}$ in lieu of the original limit-state function g;
4. **Convergence:** One or more convergence criteria on the accuracy of the failure probability estimate are checked. If they are respected, the algorithm goes to Step 6, otherwise it proceeds with Step 5.
5. **Enrichment:**
   (a) Learning function: A learning function is evaluated to find the next candidate $\mathbf{x}^{\text{next}}$ that will best enhance the metamodel in view of improving the accuracy of $\widehat{P}_f$;
   (b) ED update: Once $\mathbf{x}^{\text{next}}$ is found, the actual limit-state function $g\left(\mathbf{x}^{\text{next}}\right)$ is evaluated and the pair is added to the experimental design, *i.e.* $\{\mathcal{X}, \mathcal{G}\} \leftarrow \{\mathcal{X}, \mathcal{G}\} \cup \{\mathbf{x}^{\text{next}}, g\left(\mathbf{x}^{\text{next}}\right)\}$;
6. **End:** The currently estimated failure probability $\widehat{P}_f$ is returned.

The popular AK-MCS [5] method is obtained by using Kriging as metamodel and Monte Carlo simulation as reliability algorithm. The next best point is chosen by minimizing the so-called deviation number learning function while convergence is assumed when the probability of misclassifying any point from a predefined sample set is small enough. The efficient global reliability algorithm (EGRA) [4] is similar but to the noticeable difference that the expected feasibility function (EFF) is used as learning function, while the convergence is tailored to attaining a sufficiently small value of EFF. As we show in the sequel, various other methods have been derived by simply modifying any one of the four elements used in Steps 2 to 5.

2.2 The various ingredients of the generalized framework

We argue in this paper that, up to a few exceptions, most of the recent active learning reliability contributions spring from a mere recombination of the four ingredients introduced in EGRA and AK-MCS. These four ingredients are listed in this section and exemplified.

2.2.1 Component #1: Metamodel

A wide variety of metamodels has been used for active learning. At the onset, they can be distinguished in two classes: classification and regression/interpolation, with the latter being by far more prevalent. Borrowing from the machine learning community, artificial neural networks [7], support vector machines for classification and regression [8, 9] have been introduced in active learning reliability schemes. Polynomial chaos expansions [10] and PC-Kriging [11] are another commonly used regression methods in this framework. The most popular approach is certainly Kriging (a.k.a. Gaussian process models) which thanks to its built-in error measure, has lead to the proliferation of AK methods, a number of which has been reviewed in Teixeira et al. [12].

2.2.2 Component #2: Reliability

Reliability algorithms are crucial because they allow for the estimation of the failure probability. In numerous cases, they also direct the exploration of the random input space by providing the candidate samples for ED enrichment. While Monte Carlo simulation is still widely used, various contributions aim at by-passing its limitations (*i.e.* unmanageable costs for $P_f < 10^{-8}$) by considering advanced variance-reduction techniques. These include for instance subset simulation [13], importance sampling [14] or line sampling [15].

2.2.3 Component #3: Enrichment

The learning function aims at finding the best candidate for the enrichment of the ED considering various factors, including the current ED, the approximated limit-state surface and the metamodel properties. Teixeira et al. [12] extensively review the learning functions used in active learning reliability. Most are based on the Kriging variance, *e.g.* the deviation number, the expected feasibility function, the least improvement function, the surrogate uncertainty reduction or the information entropy LF. When a built-in error similar to the Kriging variance is not available, other error measures derived from statistical methods such as cross-validation or bootstrap are used, *e.g.* the fraction of bootstrap replicates (FBR) [10]. Finally, distance-based learning functions have also widely been used [8].

2.2.4 Component #4: Convergence

Convergence criteria are necessary in stop-ping the enrichment scheme when the failure probability estimate is judged accurate enough. They are an important part of the framework as they directly affect the efficiency and robustness of any developed active learning scheme. They can generally be classified into two groups. The first consists in direct measures of the ability of the learning function to produce meaningful new samples. In the original AK-MCS for instance, enrichment is stopped when $U > 2$ for all candidate samples. Similarly in EGRA, the criterion is translated into $\max(EFF) \leq 10^{-3}$. Such criteria have been shown to be very conservative [11]. Alternatively, direct measures of the accuracy of the estimated failure probability can be considered. Such measures may either be based on the epistemic uncertainty due to replacing g by $\widehat{g}$ or on the stability of the $\widehat{P}_f$ within enrichment iterations. More advanced techniques combining one or more of these approaches have been discussed in various contributions focusing on the stopping criterion [16].

3 Active learning reliability in UQLab

3.1 Implementation

An implementation of the active learning re-liability (ALR) framework described in the previous section is provided within UQLab. Basically, the new ALR module is a wrapper that combines various other existing modules: **METAMODEL**, **RELIABILITY** and, to some extent, **INPUT**. Figure 1 shows the structure of the ALR module and how it is connected to other UQLab modules. The **INPUT** module is used in the step 1 of the pseudo-algorithm described in Section 2.1 to draw the initial experimental design. Available methods include Monte Carlo simulation, latin hypercube sampling and quasi-Monte Carlo sequences. The **METAMODEL** and **RELIABILITY** modules lie at the core of the framework. The methods within these two modules are used as black-box, *i.e.* no modifications to the underlying algorithms was necessary. All of their pre-existing tuning options can be directly accessed and set within the ALR module. Finally, the two lower blocks as illustrated in Figure 1, *i.e.* Enrichment and Convergence criterion, are new and specifically developed for the active learning module. Currently, only a selection of the best known ones is available but the list can be easily extended. They are also implemented in a black-box mindset making it easy to add new ones and at the same time to provide gateways for user-defined custom learning functions and convergence criteria.

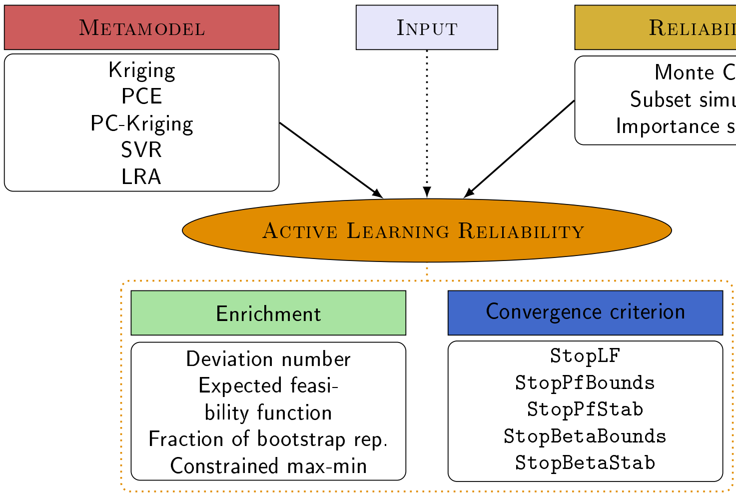

Figure 1: Active learning reliability framework in UQLab. The boxes in small caps represent existing modules.

3.2 Usage

To illustrate the usage of the ALR module, we consider a basic structural reliability problem, namely the R-S case. This problem represents an abstract system subjected to two input random variables: a resistance R and a stress S. Failure occurs when the stress is higher than the resistance, hence leading to the following limit-state function:

$$g(\mathbf{X}) = R - S \quad \text{with} \quad \mathbf{X} = \{R, S\}\,. \tag{2}$$

The probabilistic input is defined such that $R \sim \mathcal{N}(5,\,0.8)$ and $S \sim \mathcal{N}(2,\,0.6)$, where $\mathcal{N}$ denotes the Gaussian distribution.

Figure 2 shows the UQLab code to solve this problem using the ALR module. The sequence of commands is as follows:

- Initialize the UQLab framework;
- Define an **INPUT** object corresponding to the probabilistic input model;
- Define a **MODEL** object which corresponds to the limit-state function in Equation (2);
- Perform the **ANALYSIS** after specifying ALR as method to be used.

By default UQLab will run the following combination: PC-Kriging as metamodel, subset simulation as reliability algorithm, deviation number as learning function and finally `StopBetaBounds` (a criterion related to the bounds of the estimated reliability index w.r.t. Kriging variance) as convergence criterion. This default setting has been defined following an extensive benchmark as briefly discussed in the next section. The user can however combine their methods of choice according to their efficiency w.r.t. the characteristics of the problem at hand. To define AK-MCS for instance, the options shown in Figure 3 need to be additionally specified as they depart from the defaults.

```
% Initialize the UQLab framework
uqlab ;
% Define the input
IOpts.Marginals(1).Name = 'R';
IOpts.Marginals(1).Type = 'Gaussian';
IOpts.Marginals(1).Moments = [5 0.8];
...
myInput = uq_createInput(IOpts);
% Define the (original) model
MOpts.mString = 'X(:,1) - X(:,2)';
myModel = uq_createModel(MOpts) ;
% Define the analysis
AOpts.Type = 'Reliability';
AOpts.Method = 'ALR';
myAnalysis = uq_createAnalysis(AOpts);
```

Figure 2: Basic usage of active learning reliability in UQLab.

```
AOpts.ALR.Metamodel='Kriging';
AOpts.ALR.Reliability= 'MCS';
AOpts.ALR.LearningFunction = 'U';
AOpts.ALR.Convergence = 'StopLF';
```

Figure 3: Settings of the blocks to form AK-MCS.

4 Applications

4.1 Application of 20 reliability problems

To show the effectiveness of the proposed framework, we consider a benchmark on the 20 reliability problems summarized in Table 1. They have been selected so as to cover a wide range of case studies both in terms of dimensionality (from $M = 2$ to $M = 100$) and failure probability magnitude (from $P_{f,\text{ref}} = 3.14 \cdot 10^{-2}$ to $P_{f,\text{ref}} = 1.31 \cdot 10^{-7}$).

To solve these problems, we set up various solution schemes by simply combining methods from the four blocks previously defined, hence leading to a total of 39 solution strategies (See Table 2). To provide reference solutions without surrogates, and therefore a benchmark baseline, subset simulation (SuS) and importance sampling (IS) are added. For each problem and each solution scheme, the analysis is repeated 15 times to assess the robustness of the approach w.r.t. statistical uncertainty by changing the random seed. Ultimately, 12,300 reliability analyses were run. The efficiency of each of these schemes was assessed by first evaluating an accuracy index computed

Table 1: Summary of the benchmark problems (Problems 01 to 11 are from [6]. Problems 19 & 20 are based on finite element models.)

Problem	Dimension	Reference
01 (TNO RP1)	5	$7.69 \cdot 10^{-4}$
02 (TNO RP2)	2	$2.90 \cdot 10^{-3}$
03 (TNO RP3)	2	$1.31 \cdot 10^{-7}$
04 (TNO RP4)	2	$3.20 \cdot 10^{-3}$
05 (TNO RP5)	7	$8.20 \cdot 10^{-3}$
06 (TNO RP6)	2	$3.14 \cdot 10^{-2}$
07 (TNO RP7)	20	$9.79 \cdot 10^{-4}$
08 (TNO RP8)	100	$3.77 \cdot 10^{-4}$
09 (TNO RP9)	2	$9.80 \cdot 10^{-3}$
10 (TNO RP10)	10	$2.85 \cdot 10^{-7}$
11 (TNO RP11)	2	$7.83 \cdot 10^{-7}$
12 (4-branch series)	2	$3.85 \cdot 10^{-4}$
13 (Hat function)	2	$4.40 \cdot 10^{-3}$
14 (Damped oscillator)	8	$4.80 \cdot 10^{-3}$
15 (Non-linear oscillator)	6	$3.47 \cdot 10^{-7}$
16 (Frame)	21	$2.25.10^{-4}$
17 (HD function)	40	$2.00 \cdot 10^{-3}$
18 (VNL function)	40	$1.40 \cdot 10^{-3}$
19 (Transmission tower 1)	11	$5.76 \cdot 10^{-4}$
20 (Transmission tower 2)	9	$6.27 \cdot 10^{-4}$

as follows:

$$\varepsilon_{ij}^{(k)} = \left|\beta_{ij}^{(k)} - \bar{\beta}_j\right| / \bar{\beta}_j, \tag{3}$$

where $\beta_{ij}^{(k)}$ is the reliability index obtained by the k-th repetition of the i-th strategy on the j-th problem and $\bar{\beta}_j$ is the reference solution for the j-th problem, obtained using a direct (no surrogate) Monte Carlo simulation with a sufficiently large sample set. For each problem and repetition, we rank the strategies w.r.t. this index and count the number of times each strategy holds a given position. The resulting ranking is shown in Figure 4, where the best solutions appear at the bottom.

We also assess the robustness of the strategies by estimating the relative frequency that sees them within 2, 5 or 10 times the overall best obtained solution. This percentage is represented by the triangles in Figure 4. The more the triangles fall on the right, the more robust the strategy is. The overall ranking of the strategies is actually performed using the position of the outmost triangle. The actual position with each repetition is color-coded with the dark blue representing the first position and the dark red the last (41-st). It can be seen that the combination of PC-Kriging, subset simulation, expected feasibility function and the combined convergence criteria on β (both the bounds and stability) is overall best ranked w.r.t. this criterion.

Another interesting result is that even when considering just the accuracy, and regardless of the number of model evaluations needed to converge, surrogate-based approaches can outperform pure simulation schemes, because direct IS and SuS are ranked among the last. This can be explained by the fact that the reliability algorithms in the active learning schemes were significantly over-calibrated. For importance sampling, we used a sample set of size 10^5 whereas in the direct solution the sample set is traditionally set to 10^3. Similarly for subset simulation, the conditional failure probability p_0 and the sample set size are respectively to 0.25 and 10^5. In the direct case, these values were set as in traditional usage, *i.e.* 0.10 and 10^3. The rationale of using such over-calibrated settings is that we can decrease the stochastic error of the simulation algorithms at a lower cost since we are using an inexpensive metamodel. The dominating error in the estimate is then that of the metamodel. Hence if the active learning scheme can accurately describe the limit-state surface, the estimated error is expected to be smaller than that of the corresponding direct solution, as shown in Figure 4. The second measure of efficiency considered

Table 2: Methods combined to build 39 active learning strategies. PCE is only associated to FBR.

Metamodel	Reliability	Enrichment	Convergence
PC-Kriging	SuS	U	β bounds
Kriging	IS	EFF	β stability
PCE	MCS	FBR	β combined

is the number of model evaluations needed to reach a solution. To avoid counting also premature convergence, we only considered solutions whose relative error is $\varepsilon_{ij}^{(k)} \leq 10^{-2}$. The solutions that did not satisfy this condition were ranked last by default. Figure 5 shows the final ranking. With respect to this criterion, the combination of PCE, SuS, FBR and β stability shows the best performance. Obviously, the direct solutions consistently rank last in this case.

Further exploring the results of this benchmark, we can also aggregate the results by best option within each block. The following strategies could be singled out as they appeared most often in the best ranks:

- PC-Kriging as metamodel;
- Subset simulation for reliability;
- Deviation number as learning function;
- β bounds as convergence criterion.

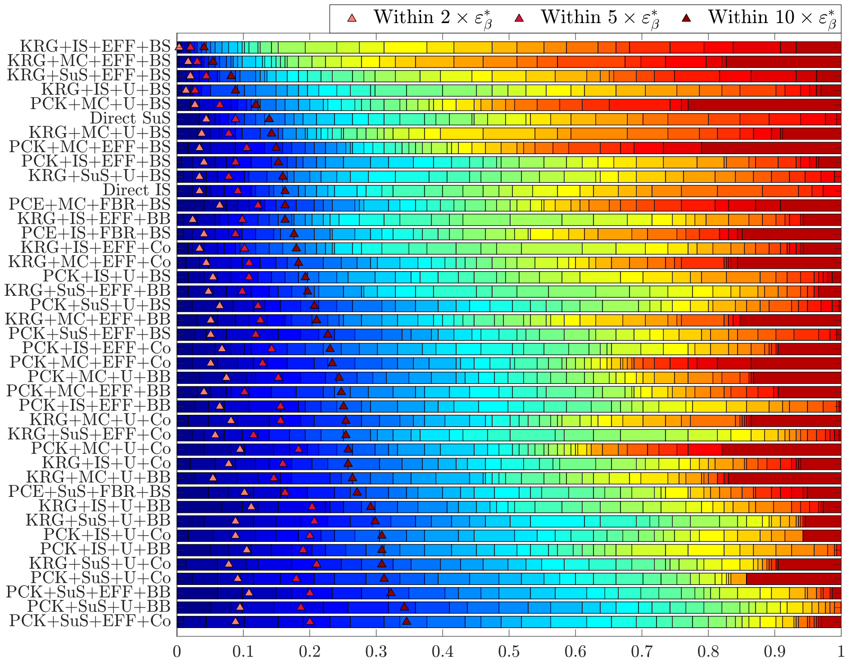

Figure 4: Ranking of the strategies w.r.t. relative error.

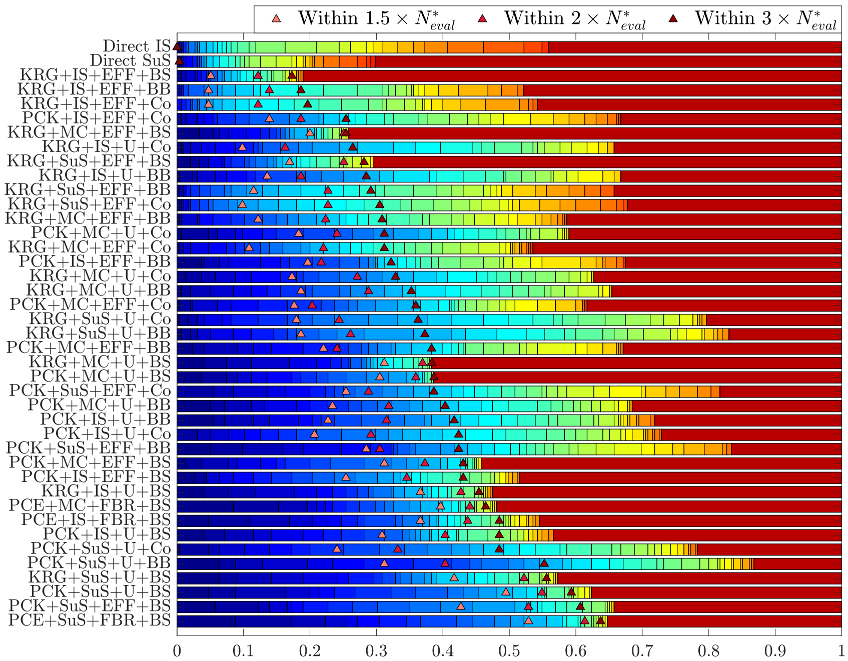

Figure 5: Ranking of the strategies w.r.t. the number of model evaluations.

4.2 TNO black-box reliability challenge

The methods highlighted in the previous section were used to build an active learning scheme for the solution of the problems proposed within the black-box reliability challenge organized by TNO, Netherlands [6]. This challenge aimed at testing the performance of structural reliability methods while providing a framework where the limit-state functions are truly black-box, *i.e.*

they were not known by the participants and were only accessible via an anonymous server API. The participants could only submit a set of input arguments and would receive in return the corresponding limit-state evaluations.

Using the four methods previously highlighted, the framework presented in this paper was used to participate in the challenge. The results were disclosed in terms of accuracy and number of model evaluations as illustrated in Figures 6 and 7, taken from [6]. The ones obtained using the UQLab ALR module are highlighted with the black arrow. Except for three (out of 26) cases, the submitted results were both the most accurate and efficient compared to other participants.

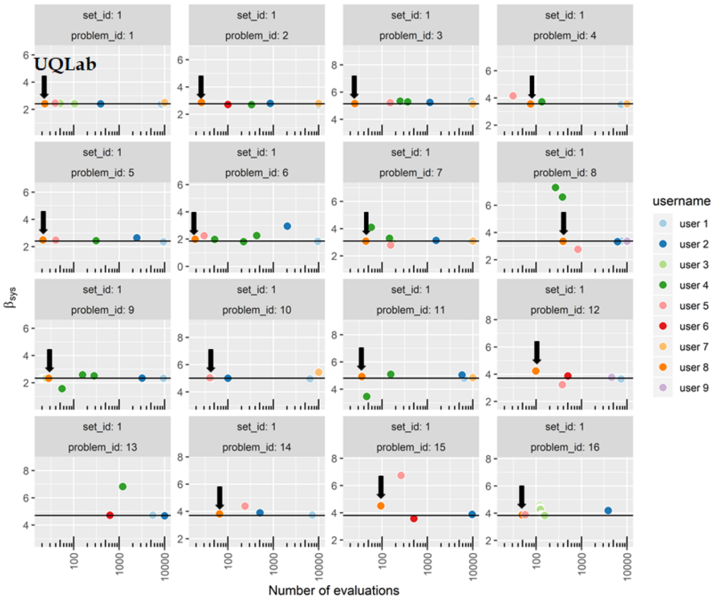

Figure 6: Results of the black-box reliability challenge (part 1) as disclosed by [6]. Results submitted using the UQLab ALR module corresponds to User 8.

5 Conclusion

This paper presents a generalized framework for active learning reliability together with its implementation in UQLab. The framework consists of four independent blocks, namely meta-modelling, reliability estimation algorithm, enrichment scheme and convergence criteria. Each of these blocks themselves are associated to existing UQLab modules. The active learning module therefore allows the user to build custom active learning schemes by non-intrusively combining methods appropriately selected from each of the four blocks. To further help in the choice of methods, a large-scale benchmark comprising 20 reliability problems and 39 selected strategies was carried out. This has shown that the use of metamodels outperforms a direct simulation method thanks to the over-calibration of the reliability estimation algorithm within the active learning scheme, hence using surrogates is always advantageous. The benchmark further showed that no single method consistently outperforms others and the final choice needs to be driven by

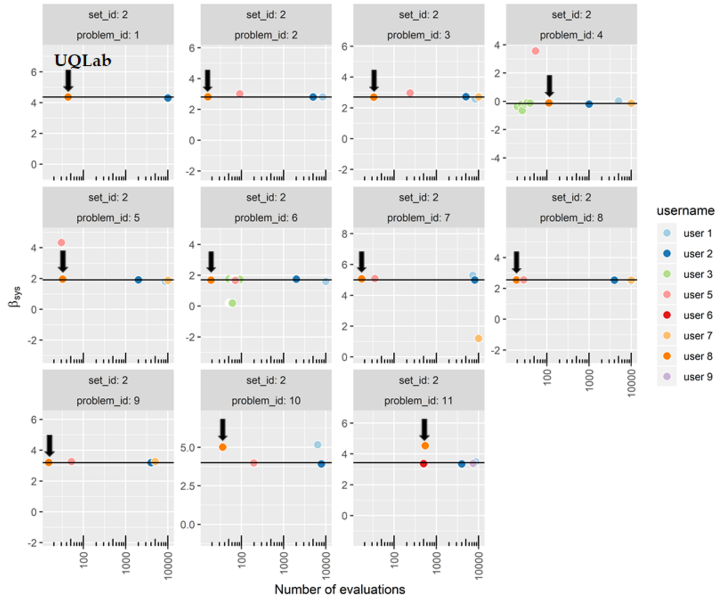

Figure 7: Results of the black-box reliability challenge (part 2) as disclosed by [6]. Results submitted using the UQLab ALR module corresponds to User 8.

the knowledge of the problem at hand. However when such prior knowledge is not available, the methods that are most likely to work well are a combination of PC-Kriging and over-calibrated subset simulation. This combination was actually selected for participating to the TNO black-box reliability challenge, and demonstrated to be extremely efficient.

References

[1] B. Sudret. *Uncertainty propagation and sensitivity analysis in mechanical models – Contributions to structural reliability and stochastic spectral methods.* Université Blaise Pascal, Clermont-Ferrand, France, 2007. Habilitation à diriger des recherches, 173 pages.

[2] S. Marelli and B. Sudret. UQLab: A framework for uncertainty quantification in Matlab. In *Vulnerability, Uncertainty, and Risk (Proc. 2nd Int. Conf. on Vulnerability, Risk Analysis and Management (ICVRAM2014), Liverpool, United Kingdom)*, pages 2554–2563, 2014.

[3] N. Lelièvre, P. Beaurepaire, C. Mattrand, and N. Gayton. AK-MCSi: A Kriging-based method to deal with small failure probabilities and time-consuming models. *Struct. Saf.*, 73:1–11, 2018.

[4] B. J. Bichon, M. S. Eldred, L. Swiler, S. Mahadevan, and J. McFarland. Efficient global reliability analysis for nonlinear implicit performance functions. *AIAA Journal*, 46(10):2459–2468, 2008.

[5] B. Echard, N. Gayton, and M. Lemaire. AK-MCS: an active learning reliability method combining Kriging and Monte Carlo simulation. *Struct. Saf.*, 33(2):145–154, 2011.

[6] A. Rozsas and A. Slobbe. Repository and black-box reliability challenge 2019. https://gitlab.com/rozsasarpi/rprepo/, 2019. Accessed: 2020-08-03.

[7] V.S. Sundar and M. D. Shields. Surrogate-enhanced stochastic search algorithms to identify implicitly defined functions for reliability analysis. *Struct. Saf.*, 62:1–11, 2016.

[8] S. Lacaze and S. Missoum. A generalized "max-min" sample for surrogate update. *Struct. Multidisc. Optim.*, 49:683–687, 2014.

[9] Jean-Marc Bourinet. *Reliability analysis and optimal design under uncertainty - Focus on adaptive surrogate-based approaches*. Université Blaise Pascal, Clermont-Ferrand, France, 2018. Habilitation à diriger des recherches, 243 pages.

[10] S. Marelli and B. Sudret. An active-learning algorithm that combines sparse polynomial chaos expansions and bootstrap for structural reliability analysis. *Struct. Saf.*, 75:67–74, 2018.

[11] R. Schöbi, B. Sudret, and S. Marelli. Rare event estimation using Polynomial-Chaos-Kriging. *ASCE-ASME J. Risk Uncertainty Eng. Syst., Part A: Civ. Eng*, 3(2), 2017. D4016002.

[12] R. Teixeira, M. Nogal, and A. O"Connor. Adaptive approaches in metamodel-based reliability analysis: a review. *Struct. Saf.*, 89:102019, 2021.

[13] X. Huang, J. Chen, and H. Zhu. Assessing small failure probabilities by AK–SS: An active learning method combining Kriging and subset simulation. *Struct. Saf.*, 59:86–95, 2016.

[14] B Echard, N Gayton, M Lemaire, and N Relun. A combined importance sampling and Kriging reliability method for small failure probabilities with time-demanding numerical models. *Reliab. Eng. Syst. Safety*, 111:232–240, 2013.

[15] Z. Lv, Z. Lu, and P. Wang. A new learning function for Kriging and its applications to solve reliability problems in engineering. *Comput. Math. Appl.*, 70:1182–1197, 2015.

[16] W. Jian, S. Zhili, Y. Qiang, and L. Rui. Two accuracy measures of the Kriging model for structural reliability analysis. *Reliab. Eng. Syst. Safety*, 167:494–505, 2017.