A generalized framework for active learning reliability analysis in UQLab

Presentation

Author(s):
Moustapha, Maliki; Marelli, Stefano; Sudret, Bruno

Publication date:
2020-02-25

Permanent link:
https://doi.org/10.3929/ethz-b-000466665

A generalized framework for active active learning reliability analysis in UQLab

Maliki Moustapha

Chair of Risk, Safety and Uncertainty Quantification – ETH Zürich

TNO 2nd workshop: Computational challenges in the reliability assessment of engineering structures
Delft, The Netherlands , 25/02/2020

Reliability analysis

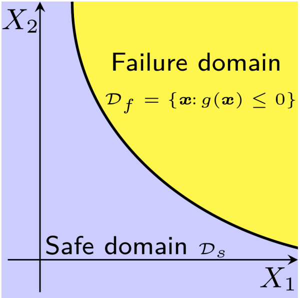
- Estimate the probability of occurence of an adverse event

$$P_f = \int_{\mathcal{D}_f} f_{\boldsymbol{X}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}$$

$f_{\boldsymbol{X}}(\boldsymbol{x})$: Joint distribution of the random vector $\boldsymbol{X}$

$\mathcal{D}_f = \{\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{X}} : g(\boldsymbol{x}, \mathcal{M}(\boldsymbol{x}) \leq 0)\}$: Failure domain

- Failure is assessed by a limit-state function $g : \boldsymbol{x} \in \mathcal{D}_{\boldsymbol{X}} \mapsto \mathbb{R}$
- The limit-state function is based on a computational model $\mathcal{M}$ describing the performance of the system



Reliability analysis (2/2)

Principal challenges

- Integration on an implicit domain
  - Analytical solutions only in a few cases *e.g.* Gaussian inputs + linear limit-state
  - In general problem solved using:
    - Approximation methods *e.g.* FORM, SORM
    - Simulation methods *e.g.* Monte Carlo, Importance sampling, Subset simulation
- High-dimensionality
  - Large number of parameters are often needed to characterize the system
  - Often inputs are stochastic processes
- Rare events
  - Failure are often rare events with low rates of occurence *e.g.* 10^{-5} to 10^{-8}

These challenges boil down to computational cost/time

Reliability analysis (2/2)

Principal challenges

- Integration on an implicit domain
  - Analytical solutions only in a few cases *e.g.* Gaussian inputs + linear limit-state
  - In general problem solved using:
    - Approximation methods *e.g.* FORM, SORM
    - Simulation methods *e.g.* Monte Carlo, Importance sampling, Subset simulation
- High-dimensionality
  - Large number of parameters are often needed to characterize the system
  - Often inputs are stochastic processes
- Rare events
  - Failure are often rare events with low rates of occurence *e.g.* 10^{-5} to 10^{-8}

These challenges boil down to computational cost/time

⬇

Surrogate models can be used to alleviate the computational burden

Surrogate models for uncertainty quantification

A surrogate model $\tilde{\mathcal{M}}$ is an approximation of the original computational model $\mathcal{M}$ with the following features:

- It is built from a limited set of runs of the original model $\mathcal{M}$ called the experimental design $\mathcal{X} = \left\{ \boldsymbol{x}^{(i)},\, i = 1, \ldots, N \right\}$
- It assumes some regularity of the model $\mathcal{M}$ and some general functional shape

Name	Shape	Parameters
Polynomial chaos expansions	$\tilde{\mathcal{M}}(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} a_{\boldsymbol{\alpha}} \, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{x})$	$\boldsymbol{a}_{\boldsymbol{\alpha}}$
Low-rank tensor approximations	$\tilde{\mathcal{M}}(\boldsymbol{x}) = \sum_{l=1}^{R} b_l \left(\prod_{i=1}^{M} v_l^{(i)}(x_i) \right)$	$b_l,\, z_{k,l}^{(i)}$
Kriging (a.k.a Gaussian processes)	$\tilde{\mathcal{M}}(\boldsymbol{x}) = \boldsymbol{\beta}^{\mathsf{T}} \cdot \boldsymbol{f}(\boldsymbol{x}) + Z(\boldsymbol{x}, \omega)$	$\boldsymbol{\beta}\,,\, \sigma_Z^2\,,\, \boldsymbol{\theta}$
Support vector machines	$\tilde{\mathcal{M}}(\boldsymbol{x}) = \sum_{i=1}^{m} a_i \, K(\boldsymbol{x}_i, \boldsymbol{x}) + b$	$\boldsymbol{a}\,,\, b$

A surrogate model is fast to evaluate

Outline

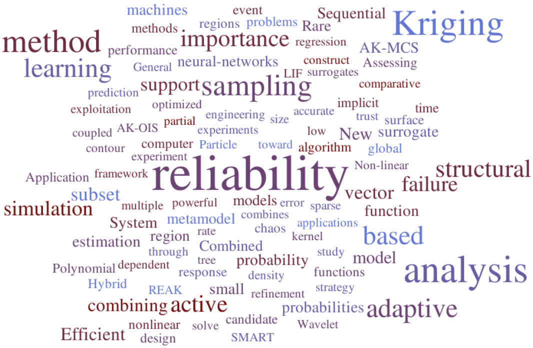
The basic idea

Enrich an initially poor experimental design using a learning function to improve the accuracy of the model in the vicinity of the limit-state surface

1. Generate an experimental design $\{\boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{Y}}\} = \left\{\left(\boldsymbol{x}^{(i)}, g\left(\boldsymbol{x}^{(i)}\right)\right), i = 1, \ldots, N_0\right\}$
2. Train a surrogate model $\widetilde{g}$ using $\{\mathcal{X}, \mathcal{Y}\}$
3. Compute the failure probability $\widehat{P}_f$ using $\widetilde{g}$
4. Check whether some convergence criteria are met. If they are, stop, otherwise go to **step 5**
5. Choose the best next sample $\boldsymbol{x}^*$ to be added to $\boldsymbol{\mathcal{X}}$ based on an appropriate learning function
6. Add $\boldsymbol{x}^*$ and the corresponding response $g\left(\boldsymbol{x}^*\right)$ to the experimental design
7. Return to **step 2**

Survey

- AK-MCS is one of the most popular active learning methods
- Other variants have been proposed by modifying:
  - the surrogate model
  - the learning function
  - the reliability method



Active learning reliability framework

Generalized framework

Reliability	Surrogate model	Learning function	Stopping criterion
Monte Carlo	Kriging	U	LF-based
Subset simulation	PCE	EFF	Stability of β
Importance sampling	SVR	FBR	Stability of P_f
Line sampling	PC-Kriging	CMM	Bounds on β
Directional sampling	Neural networks	SUR	Bounds on P_f
...	...	...	...

UQLAB active learning reliability module

A framework where customized schemes can be built by combining non-intrusively different elements in each block

- Possibility of adding custom/user-defined methods in each block

UQLab active learning module

Important features

- Reliability method
  - Any simulation method
  - Points sampled at each iteration are used as candidates for enrichment
- Surrogate models
  - Any surrogate
  - Not necessarily with embedded error measure
- Learning function
  - Some learning functions are method-specific
  - Allow for multiple points enrichment (adaptive)
- Stopping criteria
  - Some stopping criteria are surrogate-model-specific
  - It is possible to combine various stopping criteria
  - Control of the short-term history of the stopping criteria

Example in UQLab

Preliminary syntax

```
% Select active learning method
ALROptions.Type = 'Reliability';
ALROptions.Method = 'ALR';

% Surrogate model
ALROptions.ALR.Metamodel = 'PCK' ;
ALROptions.ALR.IExpDesign.NSamples = 10 ;


% Reliability analysis
ALROptions.ALR.Reliability = 'subset' ;

% Learning function
ALROptions.ALR.LearningFunction = 'U' ;
ALROptions.ALR.NumOfPoints = [3 1] ;

% Convergence
ALROptions.ALR.Convergence = {'stopBetaBound','stopBeta'} ;
ALROptions.ALR.ConvThres =  [0.01 0.005] ;
ALROptions.ALR.MaxAddedED = 500 ;
```

Selected strategies

- Combination of various methods of each block
- Combination with Kriging and PCK: 36 strategies
- Combination with PCE: 3 strategies

Reliability	Metamodelling	Learning function	Stopping criterion
Monte Carlo	Kriging	U	LF-based
Subset simulation	PCE	EFF	Stability of β
Importance sampling	SVR	FBR	Stability of P_f
Line sampling	PC-Kriging	CMM	Bounds on β
Directional sampling	Neural networks	SUR	Bounds on P_f
...	...	...	...

Selected strategies

- Combination of various methods of each block
- Combination with Kriging and PCK: 36 strategies
- Combination with PCE: 3 strategies

Reliability	Metamodelling	Learning function	Stopping criterion
Monte Carlo	Kriging	U	LF-based
Subset simulation	PCE	EFF	Stability of β
Importance sampling	SVR	FBR	Stability of P_f
Line sampling	PC-Kriging	CMM	Bounds on β
Directional sampling	Neural networks	SUR	Bounds on P_f
...	...	...	...

Selected strategies

- Combination of various methods of each block
- Combination with Kriging and PCK: 36 strategies
- Combination with PCE: 3 strategies

Reliability	Metamodelling	Learning function	Stopping criterion
Monte Carlo	Kriging	U	LF-based
Subset simulation	PCE	EFF	Stability of β
Importance sampling	SVR	FBR	Stability of P_f
Line sampling	PC-Kriging	CMM	Bounds on β
Directional sampling	Neural networks	SUR	Bounds on P_f
...	...	...	...

Benchmark problems

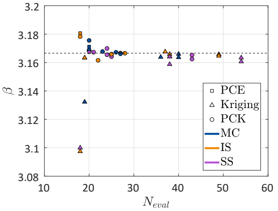
- Aiming at a set of 30 different problems
- Various features: low- to high-dimensions, rare events, complex limit-state surfaces

ID	Problem	Dimension	Reference solution	Remarks
01	TNO benchmark RP1	5	$7.61 \cdot 10^{-3}$	
02	TNO benchmark RP2	2	$2.03 \cdot 10^{-3}$	
03	TNO benchmark RP3	2	$1.21 \cdot 10^{-7}$	
04	TNO benchmark RP4	2	$1.83 \cdot 10^{-4}$	
05	TNO benchmark RP5	7	$6.36 \cdot 10^{-3}$	
06	TNO benchmark RP6	2	$2.33 \cdot 10^{-2}$	
07	TNO benchmark RP7	20	$9.88 \cdot 10^{-4}$	
08	TNO benchmark RP8	100	$3.85 \cdot 10^{-4}$	
09	TNO benchmark RP9	2	$9.68 \cdot 10^{-3}$	
10	TNO benchmark RP10	10	$2.70 \cdot 10^{-7}$	
11	TNO benchmark RP11	2	$4.19 \cdot 10^{-7}$	
12	Four branch series	2	$4.45 \cdot 10^{-3}$	
13	Hat function	2	$3.96 \cdot 10^{-4}$	
14	Damped oscillator	8	$7.59 \cdot 10^{-4}$	
15	Non-linear oscillator	6	$5 \cdot 10^{-8}$	AK-SS paper; varying Pf; highly-nonlinear
16	HD function	50	$2 \cdot 10^{-3}$	AK-SS paper; varying Pf and dimension
17	Modified Rastrigin	50	$2 \cdot 10^{-3}$	varying Pf and dimension
...	...	...	...	...

Partial results (per problem)

Problem 01

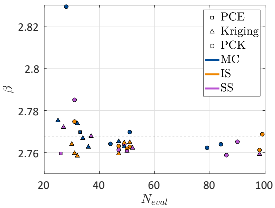
$M = 5\ ;\ \beta_{\text{ref}} \approx 3.17$



Strategy	N_{eval}	$\varepsilon_{\text{rel.}}$
PCK+IS+U+BS	18	0.0037
PCE+IS+FBR+BS	18	0.0044
KRG+IS+U+BS	19	0.0010

Problem 02

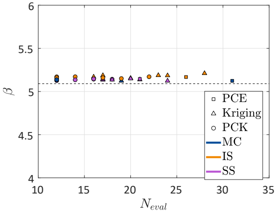
$M = 2\ ;\ \beta_{\text{ref}} \approx 2.77$



Strategy	N_{eval}	$\varepsilon_{\text{rel.}}$
KRG+MCS+U+BS	25	0.0026
PCE+SuS+FBR+BS	26	0.0030
KRG+SuS+U+BS	27	0.0019

Problem 03

$M = 2\ ;\ \beta_{\text{ref}} \approx 5.09$



Strategy	N_{eval}	$\varepsilon_{\text{rel.}}$
PCK+SuS+U+BB	12	0.0070
PCK+MC+U/EFF+BB	12	0.0090
PCK+SuS+EFF+BB	12	0.0086

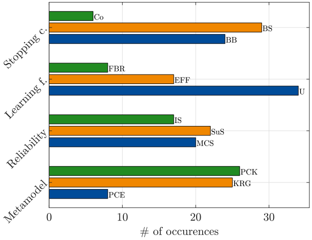
$$\varepsilon_{\text{rel.}} = \frac{\left|\text{median}\left(\beta\right) - \beta_{\text{ref}}\right|}{\beta_{\text{ref}}}$$

- All solutions with $\varepsilon_{\text{rel.}} > 0.01$ are discarded

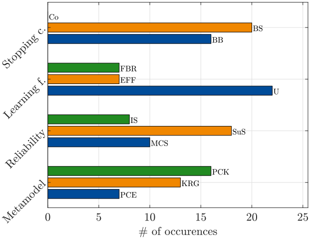
Partial results (aggregate)

- Number of times a given method is within the K best results

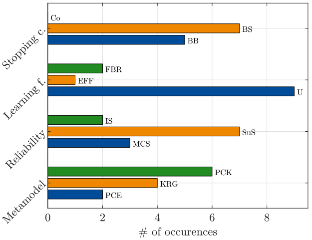
$K = 5$



$K = 3$



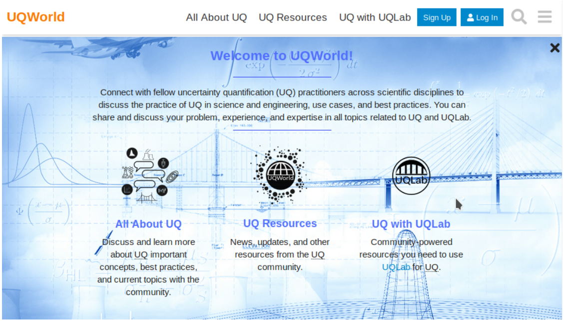
- Number of times a method is the best
- Multiple counts in case of equality



Concluding remarks

- Global framework for active learning
- Methods can be combined non-intrusively according to the specifics of the problem at hand
- Large benchmark currently running
- Statistics on different methods in each block w.r.t. problems features
- Module to be released in UQLab 1.4

UQWorld



uqworld.org

- Share your work and research to the community; discuss them with the community
- Help others by sharing your own best practices in reliability analysis
- Get news and updates from the UQ community at large

Questions ?

Chair of Risk, Safety & Uncertainty Quantification

www.rsuq.ethz.ch

The Uncertainty Quantification Laboratory

www.uqlab.com

Thank you very much for your attention !