Kriging Adaptive Learning for High Dimensional Reliability Assessment with a Variance-based Learning and Stopping criterion

G. Capasso [1,2], C. Gogu [1], C. Bes [1], J. P. Navarro [2], M. Kempeneers [2]
[1] *Université Paul Sabatier Toulouse III, Institut Clément Ader*
[2] *Airbus Operations SAS*

Presented at MascotNum 2022, Clermont Ferrand, 7-9 June 2022

Introduction

- Probabilistic approaches received growing interest in research community and industry during last decades
- Increasing complexity of industrial numerical model lead to higher and higher computational burden
- Machine Learning (ML) techniques are proven to ease the computational burden of reliability analyses
- Real-world applications often imply a high number of input variables, making the ML-based approaches more challenging

Challenges

- Determine failure probability P_f
- Reference: Monte Carlo (MC)
  - Pros: no curse of dimensionality
  - Cons: requires too many simulations of (often) computationally expensive models
- Alternative: replace model with Kriging
  - Pros: optimal adaptive construction
  - Cons: curse of dimensionality
- Objectives: cost saving & accuracy
  - Capability to efficiently determine P_f
  - Handle high input dimensionality
  - Provide a global measure of accuracy
  - Balance accuracy/ computational burden

Surrogate model [2]

Kriging

+ Good exploration features
+ Proper error structure
- Curse of dimensionality

Training/ prediction based on covariance kernel definition:

$$k(\mathbf{x},\mathbf{x'}) = \sigma^2 \prod_{i=1}^{d} \exp\left(-\theta_i \left(\mathbf{x}_i - \mathbf{x'}_i\right)^2\right)$$

Partial Least Square

+ Allows dimensionality reduction
- Limited accuracy

Projection of input variables:

$$t^{(h)} = XW_* \qquad y \approx ct^{(h)}$$

Kriging Partial Least Square

+ Good exploration features
+ Proper error structure
+ Scalable for high dimensional inputs

Modified covariance kernel definition:

$$k(\mathbf{x},\mathbf{x'}) = \sigma^2 \prod_{l=1}^{h}\prod_{i=1}^{d} \exp\left[-\theta_l \left(\mathbf{w}_{*i}^{(l)}\mathbf{x}_i - \mathbf{w}_{*i}^{(l)}\mathbf{x'}_i\right)^2\right]$$

References

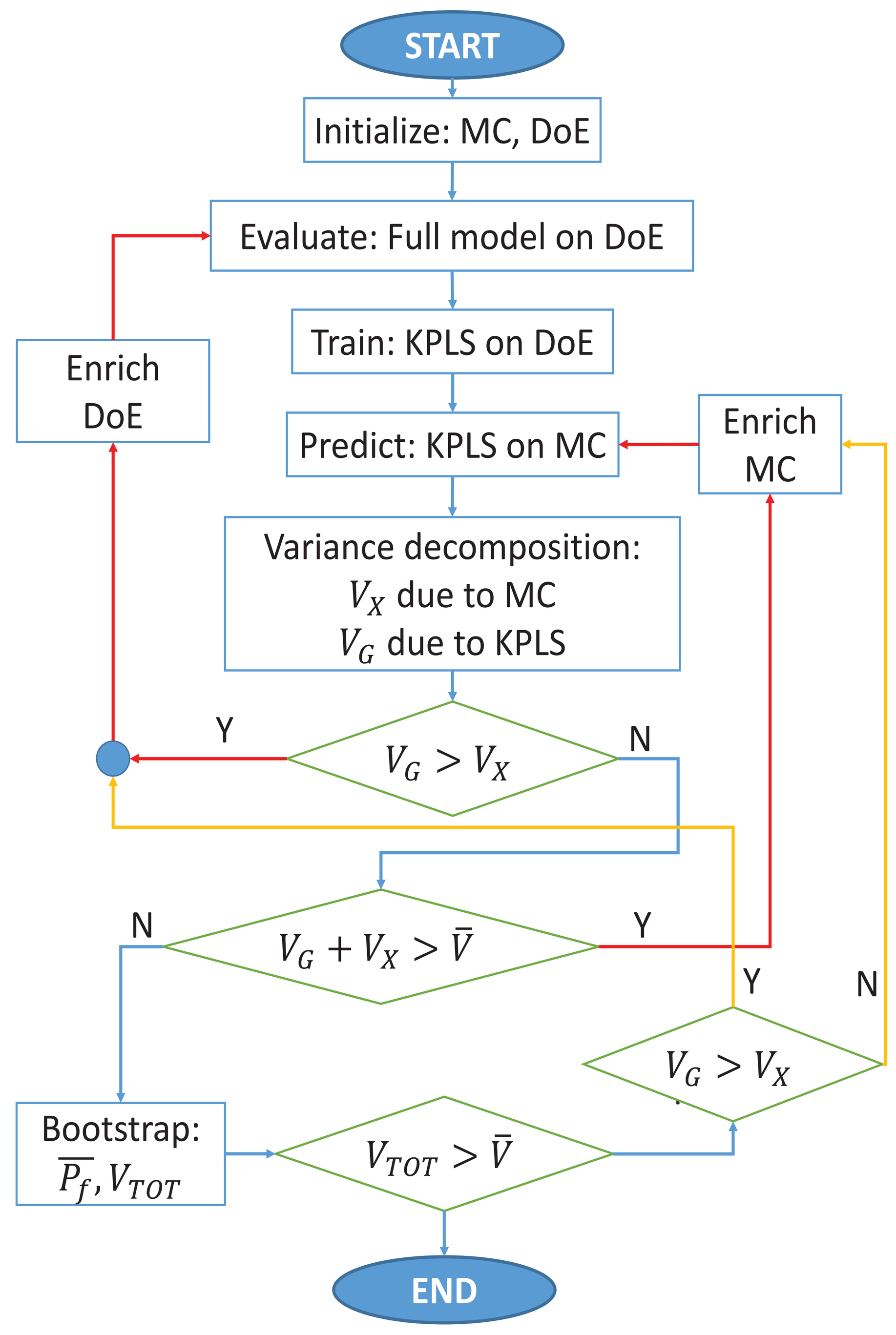
[1] B. J. Bichon et al. Efficient global reliability analysis for nonlinear implicit performance functions. *AIAA journal*, 2008.

[2] M. A. Bouhlel et al. Improving kriging surrogates of high-dimensional design models by partial least squares dimension reduction. *Structural and Multidisciplinary Optimization*, 2016.

[3] B. Echard et al. Ak-mcs: an active learning reliability method combining kriging and monte carlo simulation. *Structural Safety*, 2011.

[4] M. Menz et al. Variance based sensitivity analysis for monte carlo and importance sampling reliability assessment with gaussian processes. *Structural Safety*, 2021.

Proposed Mathematical Framework



The global algorithm architecture is inspired from [4].

Scalable surrogate model
- KPLS [2] → extend Kriging advantages to high-dimensional problems

Fully-adaptive algorithm
- Adaptive learning → iteratively enrich DoE
- Adaptive sampling → iteratively enrich MC

Variance decomposition
- V_G → variance due to KPLS
- V_X → variance due to MC
- V_{GX} → covariance term (assumed negligible)

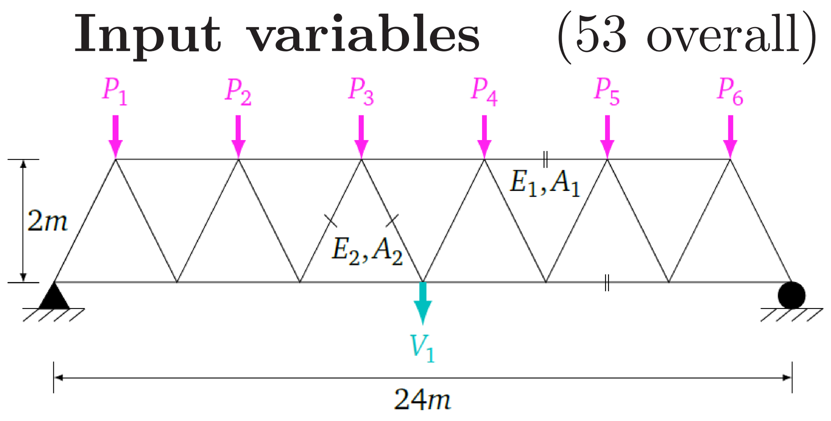
Variance-based learning criterion
- EFF function [1] → select new point
- Variance-based (Vb) criterion → $V_G \leq V_X$

Variance-based stopping criterion
- Total variance → $V_G + V_X \leq \bar{V} = \overline{coV}P_f^2$
- Bootstrap → verify that V_{GX} can be neglected
- Used for sampling and overall procedure

Numerical Results

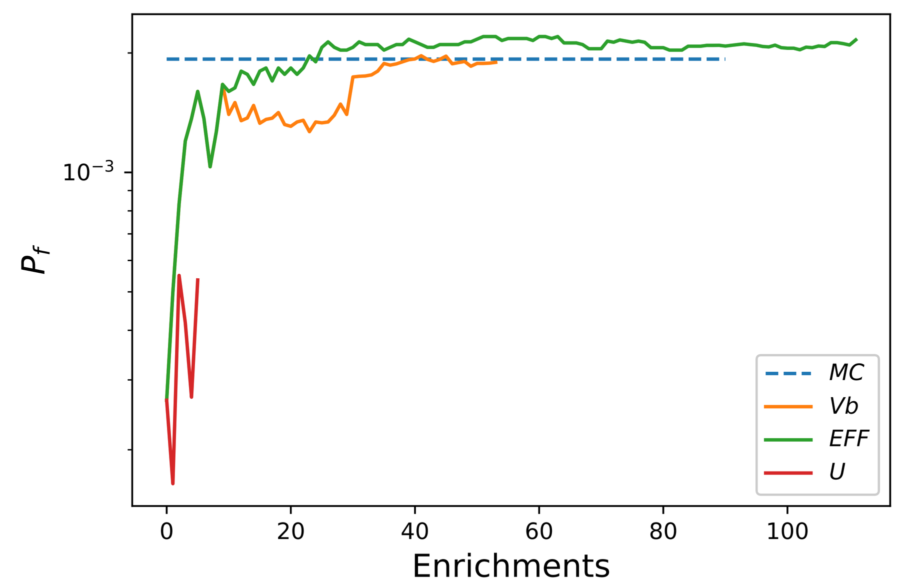
Comparison with U (from AK-MCS [3]) and EFF approach [1].

Input variables (53 overall)

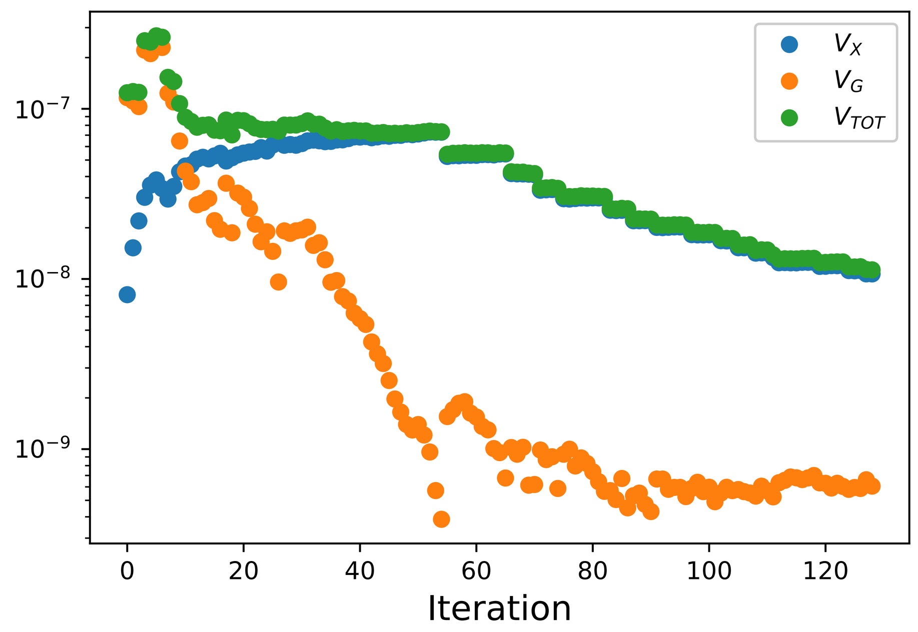


Limit state function

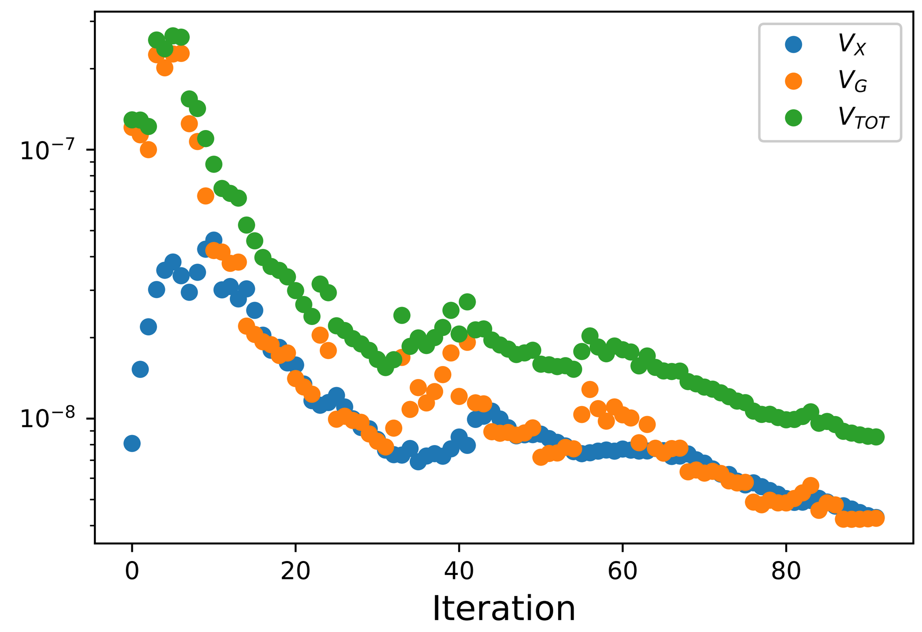
$$g(V_{all}, X_S) = V_{all} - V_1(X_S)$$



Evolution of P_f as function of DoE updates



Variance decomposition for EFF



Variance decomposition for Vb

Conclusion

We achieved the main objective of conceiving a ML-based approach for reliability applications:

- scalable, thus able to extend Kriging exploration features to high-dimensional problems
- efficiently updating the training set with the most relevant inputs
- reducing computational efforts with respect to other referenced methods
- able to distinguish the different sources of error, in terms of variance, of the P_f estimation
- providing a global measure of variability of the failure probability estimate
- guaranteeing a good overall accuracy by balancing the sources of uncertainty.