



**HAL**  
open science

## Calcul de probabilité de défaillance par échantillonnage d'importance sur méta-modèle de Krigeage pour un matériau viscoélastique

Aurélien Guesdon, Rivier Mickael, Elisa Bonigen, Marie Cottrel-Bussenault,  
Charles Elegbede

### ► To cite this version:

Aurélien Guesdon, Rivier Mickael, Elisa Bonigen, Marie Cottrel-Bussenault, Charles Elegbede. Calcul de probabilité de défaillance par échantillonnage d'importance sur méta-modèle de Krigeage pour un matériau viscoélastique. Congrès Lambda Mu 23 “ Innovations et maîtrise des risques pour un avenir durable ” - 23e Congrès de Maîtrise des Risques et de Sûreté de Fonctionnement, Institut pour la Maîtrise des Risques, Oct 2022, Paris Saclay, France. hal-03876600

**HAL Id: hal-03876600**

**<https://hal.science/hal-03876600v1>**

Submitted on 28 Nov 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Calcul de probabilité de défaillance par échantillonnage d'importance sur méta-modèle de Krigeage pour un matériau viscoélastique

## Calculation of failure probability using importance sampling on kriging surrogate model for a viscoelastic material

GUESDON Aurélien  
*ArianeGroup*

Les cinq chemins 33 187 Le Haillan  
aurelien.guesdon@ariane.group

RIVIER Mickael  
*ArianeGroup*

Les cinq chemins 33 187 Le Haillan  
mickael.rivier@ariane.group

BONIGEN Elisa  
*ArianeGroup*

Les cinq chemins 33 187 Le Haillan  
elisa.bonigen@ariane.group

COTTREL BUSSENAULT Marie  
*ArianeGroup*

Les cinq chemins 33 187 Le Haillan  
marie.cottrel@ariane.group

ELEGBEDE Charles  
*ArianeGroup*

Rue du Général Niox 33165 Saint  
Médard-en-Jalles  
charles.elegbede@ariane.group

**Résumé** — Cet article décrit la méthode de probabilisation de défaillance mécanique d'un corps de propulseur solide. En premier lieu la caractérisation des distributions statistiques des lois de comportement des matériaux, la réalisation d'un plan d'expérience calculs sur les paramètres influents, puis l'exploitation de ce plan par la construction d'un méta-modèle de Krigeage permettant de prédire les Coefficients de Sécurité (CS). Enfin, la réalisation des simulations de Monte-Carlo en utilisant la méthode d'échantillonnage d'importance sur ce méta-modèle pour déterminer la probabilité de défaillance selon le critère de rupture.

**Mots-clefs** — *Plan d'expérience, Krigeage, Propagation d'incertitudes, Echantillonnage d'importance, Viscoélastique*

**Abstract** — This article is describing the way to access to the probability of mechanical rupture of a solid propellant motor case. First, characterization of statistical distribution of mechanical behavior of a viscoelastic material (propellant and junctions), then the establishment of a Design Of Experiment (DOE) on important parameters based on mechanical calculations. Then, the exploitation of the DOE and the construction of a Kriging surrogate model allowing to predict safety coefficients. Finally, the realization of Monte-Carlo (MC) simulations on this surrogate model by using Importance-Sampling method and calculate the failure probability.

**Keywords** — *Design of Experiment, Kriging, Propagation of uncertainty, Importance sampling, Viscoelastic*

particulièrement complexe pour diverses raisons : le caractère viscoélastique du propegol et la diversité des cas de charges lors des phases de vie du propulseur (retrait thermique lors des phases de fabrication, gravité lors des phases de stockage, vibrations-chocs lors de phases de manutention et de vols, pression lors des phases d'allumage). Afin d'appréhender cette complexité de manière plus physique, une méthode de dimensionnement par mécanique de la rupture sur des profils de vie dimensionnant a été développée. La mise en place de cette méthode a été l'opportunité d'introduire une nouvelle méthode de probabilisation, elle aussi plus physique.

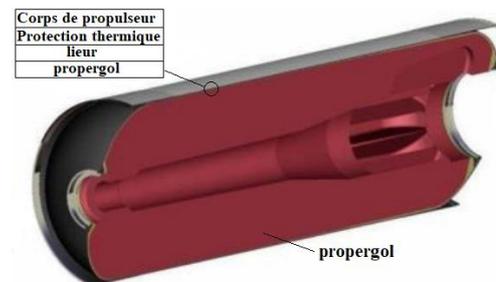


Fig. 1 Exemple de corps de propulseur à propegol solide

### I. INTRODUCTION

Le dimensionnement mécanique du propegol et de ses liaisons avec la structure du propulseur (Fig. 1) est une affaire

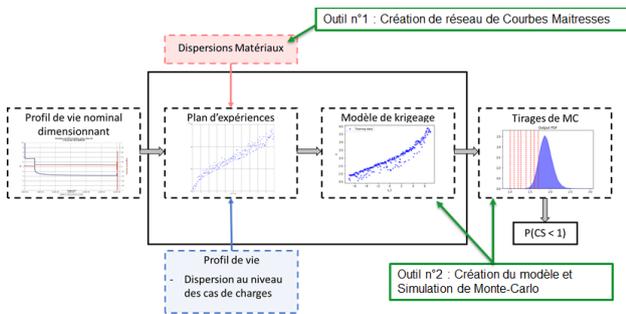


Fig. 2 Logigramme de la méthode de probabilisation

En effet, ces méthodes, basées sur des automatisations de boucles de calculs, permettent d'effectuer un grand nombre de calculs sans intervention humaine, laissant la possibilité de récupérer suffisamment de données pour calculer des probabilités de rupture fiable. La logique de cette méthode est divisée en plusieurs étapes (Fig. 2), la première consiste à identifier et caractériser la dispersion des paramètres influents ainsi que leurs corrélations. La seconde est de réaliser les plans d'expérience calculs suffisamment nombreux et bien distribués avant de les exploiter par la réalisation d'un méta-modèle de Krigeage. Enfin, des simulations de Monte-Carlo avec utilisation de l'échantillonnage d'importance sont réalisées sur ce méta-modèle dans le but d'obtenir les probabilités de ruptures souhaitées qui répondent aux exigences de sureté de fonctionnement du produit.

## II. CARACTERISATION DE LA DISPERSION DES PARAMETRES INFLUENTS

### A. Réseau de courbes maitresses

La majorité des paramètres influents sur le dimensionnement mécanique du corps de propulseur sont des paramètres viscoélastiques. Ces caractéristiques mécaniques de matériau viscoélastique dépendent du temps de sollicitation et de la température ; la combinaison de ces deux paramètres permet d'obtenir un temps équivalent temps-température. La dépendance des caractéristiques mécaniques viscoélastiques en fonction du temps équivalent suit usuellement une loi exponentielle. Cette loi peut comporter une ou plusieurs pentes (jusqu'à 2 pentes dans notre cas). On nomme cette loi courbe maitresse. C'est à partir de cette courbe maitresse qu'un calcul mécanique viscoélastique peut être mené. L'objectif final est de probabiliser la rupture mécanique, pour cela il faut intégrer une loi de distribution sur les paramètres influents. Pour effectuer ceci sur les paramètres

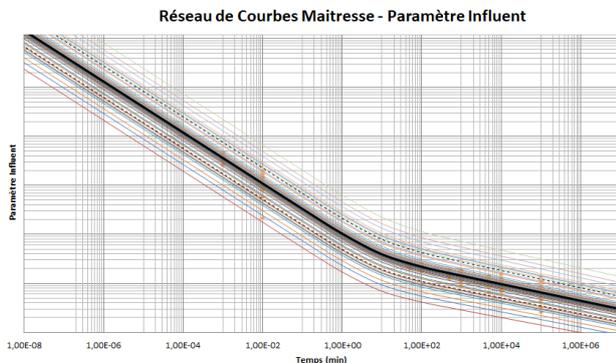


Fig. 3 Réseau de courbe maitresse d'un paramètre viscoélastique a 2 pentes

viscoélastiques, on utilise l'ensemble des résultats d'essais

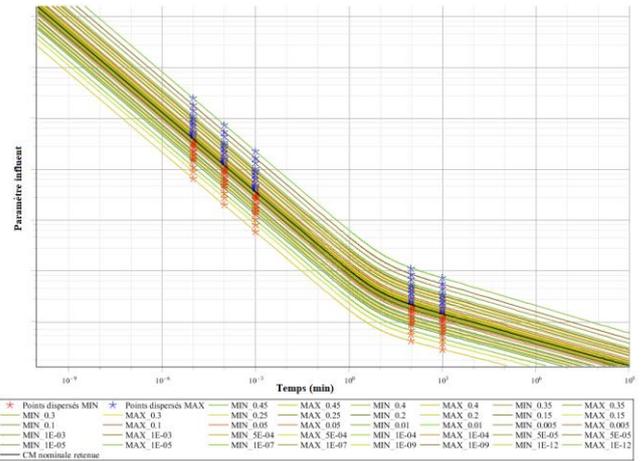


Fig. 4 Visualisation de l'outil de création de réseau de courbes maitresse (en échelle logarithmique)

disponibles et valides effectués (points oranges sur la Fig. 3) sur nos matériaux et on applique la formule de l'intervalle de prévision d'une régression (1) issue de [1] pour prédire sur l'ensemble de domaine de temps équivalent souhaité ce qu'on appelle le réseau de courbes maitresses (Fig. 3). Ce réseau, donne pour une probabilité donnée, la courbe maitresse correspondante.

$$Y_0 \pm t_{1-\frac{\alpha}{2}}(ddl)\sigma \sqrt{1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum_{i=1}^n (X_i - \bar{X})^2}} \quad (1)$$

Ce processus est réalisé en 3 étapes (exemple pour une courbe maitresse à 2 pentes) :

- Etablir la régression logarithmique entre la caractéristique étudiée et le temps équivalent
  - Pour les temps dit « courts »
  - Pour les temps dit « longs »
- Déterminer les points dit « dispersés » aux probabilités recherchées (voir les croix bleus et rouges sur la Fig. 4)
- Etablir les régressions entre les points dispersés iso-probables selon le standard de courbe souhaité (formule)

Ce processus a été partiellement automatisé grâce au développement d'un outil interne ArianeGroup utilisant la base de données des essais valides et guidant l'utilisateur dans les étapes précédemment décrites.

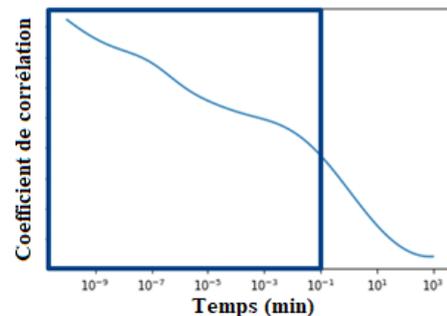


Fig. 5 Evaluation de la corrélation entre 2 paramètres viscoélastiques en fonction du temps équivalent

## B. Corrélation entre paramètres influents

Le propergol est un matériau qui, suite à sa fabrication, montre de fortes variations de modules. Cette variation permet de mettre en évidence des corrélations avec certaines de ses caractéristiques à rupture comme la contrainte. Cette corrélation peut affecter de façon non négligeable les probabilités de ruptures et doit donc être caractérisée. En effet, lorsque le module augmente, les contraintes dans le matériau augmentent. Mais la capacité à rupture évolue elle aussi dans le même sens, ce qui fait que le Coefficient de sécurité reste relativement stable.

Différentes corrélations ont été établies entre module propergol et capacité à rupture. La difficulté réside dans l'estimation d'un coefficient de corrélation moyen ou conservatif (en termes de tenue mécanique) sur l'ensemble du domaine de temps équivalent utile et étudié (Fig. 5). Cette estimation est réalisée pour chacune des conditions d'essais, pour les essais de traction comme de fluage.

### III. REALISATION DES PLANS D'EXPERIENCE

Des études préalables de sensibilité aux paramètres d'entrée des coefficients de sécurité du corps de propulseur ont permis d'établir les paramètres influents et ceux à exclure de la probabilisation. Ces études ont consisté à réaliser des plans d'expérience simple et peu coûteux dans l'objectif de hiérarchiser les paramètres influents et surtout les ordres de grandeurs de ces influences. Les valeurs des paramètres non retenus ont été choisies pour assurer le conservatisme de la méthode.

Pour chacun des 15 profils de vie dimensionnants retenus, un plan d'expérience factoriel complet a été réalisé. Les paramètres influents ont été adimensionnés notamment pour assurer l'exploitation des résultats.

TABLE I. LISTE DE PARAMETRES INFLUENTS ET DOMAINE DE VALIDITE DU PLAN D'EXPERIENCE

Variabes	Valeur nominale	Borne min	Borne max
Raideur propergol	0	$-7\sigma$	$+7\sigma$
Coefficient de dilatation propergol	0	$-7\sigma$	$+7\sigma$
Capacité à rupture propergol	0	$-7\sigma$	$+7\sigma$
Capacité à rupture liaisons propergol	0	$-7\sigma$	$+7\sigma$

Les paramètres sont centrés sur leurs valeur nominale et s'étendent jusqu'à  $\pm 7\sigma$ . Le domaine de validité a été choisi volontairement très large afin d'être en capacité d'évaluer des probabilités très faibles (mais non nulles si possible) même lorsque les coefficients de sécurité sont très élevés.

Au final, c'est un peu plus de 3000 résultats disponibles pour chacun des profils de vie qui permettront de créer les méta-modèles. De ces résultats sont extraits les Coefficients de Sécurité de chacune des zones d'intérêts du corps de propulseur (une dizaine de zones).

### IV. EXPLOITATION DES PLANS D'EXPERIENCE

Les bases de données obtenues sur les plans d'expériences présentés ci-dessus sont exploitées pour construire des modèles de substitution des différents Coefficients de Sécurité (CS). Cela permet d'obtenir un modèle prédictif de chaque coefficient à coût de calcul négligeable, qui pourra être utilisé

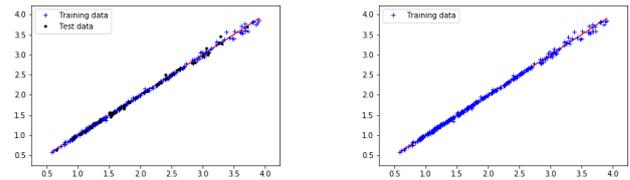


Fig 6. Exemples de Goodness-of-fit de validation (à gauche) et complet (à droite).

pour l'estimation empirique des probabilités de défaillance et des quantiles associés.

Il est ici proposé de construire des modèles par processus Gaussiens, aussi appelés modèles de krigeage. Ces derniers sont en effet généralement très performants lorsque le nombre de dimensions d'entrée est réduit et que la base de données ne dépasse pas quelques milliers d'échantillons, ce qui est le cas ici.

En pratique, par soucis de parcimonie, des plans d'expériences plus légers seront extraits des 3000 points mentionnés plus haut. L'impact de ce choix pourra être quantifié dans la métrique de qualité du modèle, et dans l'erreur induite par le métamodèle sur les statistiques d'intérêt.

#### A. Construction du métamodèle de krigeage

Les métamodèles de krigeage sont construits sur les données des plans d'expériences grâce à une librairie python interne reposant sur le paquet open-source GPy [2], développé par l'Université de Sheffield.

L'une des forces principales de ces modèles est qu'en plus de la prédiction, appelée moyenne prédictive, le krigeage retourne une métrique d'incertitude associée à cette prédiction, appelée variance prédictive. Dans le cas d'un krigeage sans tendance, ces deux grandeurs sont données en (2) et (3).

$$\mu(x) = k_*^T(x)(K + \sigma_n^2 I)^{-1}y, \quad (2)$$

$$\sigma^2(x) = k_{**}(x) - k_*^T(x)(K + \sigma_n^2 I)^{-1}k_*(x). \quad (3)$$

Une tendance constante est obtenue implicitement par normalisation et centrage des valeurs de CS. Les dimensions d'entrée sont également normalisées afin de faciliter au maximum l'étape d'optimisation des hyper-paramètres. Cela est particulièrement important car nous utilisons ici la technique d'*Automatic Relevance Determination* (ARD), associant un hyper-paramètre différent à chaque dimension d'entrée, et permettant la construction d'un modèle de krigeage anisotrope. L'optimisation de la vraisemblance est réalisée par l'algorithme L-BFGS-B avec 10 restarts pour limiter le risque de tomber dans un minimum local. Enfin, pour un maximum de généralité, nous utilisons l'un des noyaux universels les plus communs dans la littérature : le noyau Matern5/2.

#### B. Validation du métamodèle

En pratique, nous proposons de construire deux métamodèles de krigeage. Le premier sur 80% des données disponibles, afin d'obtenir des scores de validation et de vérifier l'absence de sur-apprentissage. Le second sur la totalité de la base de donnée, en vérifiant la cohérence avec les scores du premier sur la base d'entraînement, afin d'utiliser

toute l'information disponible pour l'estimation des probabilités de défaillance à venir.

Pour chaque modèle, les scores  $R^2$  et de *Root Mean Square Error* (RMSE) relative sont retournés sur les données d'entraînement. De plus un graphe de *Goodness-of-fit* est tracé, représentant les valeurs prédites en ordonnée, en fonction des valeurs réelles de CS. Un modèle parfait aura donc l'ensemble des points sur la diagonale  $y = x$ . De plus, pour le modèle de validation, entraîné sur 80% des données uniquement, le score  $Q^2$  ( $R^2$  sur les données de test) et la RMSE relative sur les points de test sont aussi calculés.

Un exemple de construction de modèle est donné en Fig. 6, avec les graphes de *Goodness-of-fit* associés au modèle de validation et au modèle complet. Dans ces deux graphes, les points d'entraînement apparaissent en croix bleues et les points de test en rond noir. Les scores  $Q^2$  et  $R^2$  de ces deux modèles sont de 0.997 et 0.998 respectivement. Les comportements des modèles sont ici extrêmement similaires, et ils possèdent tous deux de très bons scores. Les différentes analyses statistiques ci-dessous peuvent donc être réalisées sur ces modèles sans crainte de biais important dans les résultats.

## V. ESTIMATION DES STATISTIQUES DE DEFAILLANCE

Les métamodèles construits ci-dessus peuvent être exploités pour obtenir des statistiques d'intérêt. En effet, les paramètres d'entrée de ces modèles (module matériau, capacité à rupture, ...) sont supposés aléatoires, et leurs distributions et corrélations ont été déterminées. L'objectif de cette section est donc de propager ces incertitudes d'entrée à travers les métamodèles, d'estimer des statistiques d'intérêt et de quantifier les différentes sources d'erreur.

Deux statistiques principales sont estimées dans cette étude. La première correspond directement à la probabilité du CS de descendre en dessous d'un certain seuil. A l'inverse, la seconde correspond au quantile de CS correspondant à une probabilité de défaillance fixée.

En pratique, ces statistiques sont ici estimées empiriquement, grâce à des techniques d'échantillonnage de type Monte Carlo (MC). Une méthode par échantillonnage d'importance est notamment proposée pour améliorer la qualité d'estimation des différentes statistiques.

### A. Estimations par échantillonnage d'importance

L'estimation Monte Carlo empirique des statistiques d'intérêt est par définition sujette à une variance d'estimation, dépendant directement du nombre d'échantillons exploités. Dans le contexte d'estimation de très faibles probabilités, ou de quantiles extrêmes, l'ordre de grandeur de l'estimation peut devenir inférieur à celui de l'écart-type d'estimation, ce qui représente une limite ferme à l'approche Monte Carlo. Cela vient intuitivement du fait que les échantillons tirés pour l'estimation empirique des statistiques sont tirés depuis la distribution d'entrée. Ainsi, dans le cas de très faibles probabilités, l'immense majorité des échantillons sera tirée dans la zone acceptable, et très peu amèneront à un cas de défaillance.

La méthode d'échantillonnage d'importance (*Importance Sampling*, IS) est une technique de réduction de variance d'estimation qui permet de grandement améliorer la précision des statistiques empiriques dans ce contexte. Cette approche permet d'estimer ces statistiques extrêmes, avec les distributions d'entrée données, à partir d'échantillons

provenant d'une distribution différente. L'objectif est donc de tirer beaucoup plus d'échantillons dans la zone de défaillance, puis de pondérer chaque tirage par la probabilité réelle associée. En pratique, les estimations MC et IS de probabilité de défaillance sont réalisées avec les formules (4) et (5).

$$p_{MC} = \sum_i \mathbb{1}(x_i \leq s), \quad (4)$$

$$p_{IS} = \sum_i \left( \mathbb{1}(y_i \leq s) \times \frac{f(y_i)}{g(y_i)} \right), \quad (5)$$

avec  $f$  la densité de probabilité (*Probability Density Function*, PDF) réelle des paramètres d'entrée et  $g$  la PDF d'échantillonnage IS. Les échantillons  $x_i$  sont tirés depuis  $f$  et les  $y_i$  depuis  $g$ . L'opérateur  $\mathbb{1}$  est la fonction indicatrice, permettant par la somme de compter le nombre d'occurrences d'un événement.

Les estimations de quantiles suivent des formules légèrement plus complexes, notamment tirées de [4]. Pour une probabilité  $p$  donnée, ces estimations passent toutes deux par la définition de  $F$ , la fonction de densité cumulée empirique (*Empirical Cumulative Density Function*, ECDF), avec ensuite  $q = F^{-1}(p)$ . Les ECDF par Monte Carlo et échantillonnage d'importance sont données en (6) et (7).

$$F_{MC}(z) = \frac{1}{N} \sum_i \mathbb{1}(x_i \leq z), \quad (6)$$

$$F_{IS}(z) = \frac{1}{N} \sum_i \left( \mathbb{1}(y_i \leq z) \times \frac{f(y_i)}{g(y_i)} \right). \quad (7)$$

Il est à noter que cette formule s'applique pour les quantiles associés à de probabilités très proches de 0. Les différentes formules sont explicitées et quantitativement comparées dans [4].

Le point central de l'utilisation de cette méthode IS est donc le choix de cette nouvelle distribution d'échantillonnage, permettant de mettre plus de poids dans les zones de défaillance. Dans notre cas, nous proposons d'utiliser une distribution uniforme sur l'intégralité du domaine d'applicabilité du plan d'expériences, contenant les supports des distributions des paramètres d'entrée.

### B. Intervalles d'erreur des estimations

Les estimations présentées ci-dessus sont affectées de deux types d'erreur :

- une erreur due à l'estimation empirique, avec un échantillonnage donné,
- une erreur due à l'utilisation d'une prédiction des CS par métamodèle.

#### 1) Erreurs dues à l'échantillonnage

Comme évoqué précédemment, l'estimation empirique des statistiques d'intérêt possède une certaine variance, bien que grandement réduite par l'utilisation de l'échantillonnage d'importance. Cette variance serait notamment visible en réalisant l'estimation de statistique un grand nombre de fois. Cela induit cependant un coût informatique non-négligeable qui peut être aisément accéléré par l'utilisation d'une méthode de bootstrap. Cette technique consiste aussi à calculer les statistiques empiriques sur un grand nombre de jeux de données, mais dont les valeurs sont tirées avec remise depuis le jeu de données de la première estimation.

Cette approche possède le premier avantage de ne pas nécessiter d'évaluations supplémentaires du métamodèle, ce

qui accélère sensiblement l'analyse, mais surtout un deuxième avantage critique : la possibilité d'obtenir analytiquement les dispersions associées aux estimations de quantile. Ces dernières étant significativement plus coûteuses que les estimations de probabilité, cela permet un important gain en temps de calcul.

Les erreurs d'estimation liées à l'échantillonnage seront donc estimées par bootstrap empirique (tirages avec remise) pour les probabilités de défaillance, et par bootstrap analytique, avec les formules présentées ci-après, pour les quantiles. Dans les deux cas, cela permet d'obtenir une distribution correspondant aux valeurs possibles de la statistique estimée.

Dans le cas de l'estimation du quantile de probabilité  $p$  par simple Monte Carlo, la distribution bootstrap correspond, pour chaque échantillon  $i$  du tirage préalablement trié, à la probabilité que le  $p \times N$ -ième élément d'un tirage bootstrap tombe sur cet échantillon. La formule (8) permet d'évaluer la probabilité (par bootstrap) que le quantile tombe sur l'échantillon  $x_i$ .

$$P_{bs}[q = x_i] = P_{S_i}[S_i \geq pN] - P_{S_{i-1}}[S_{i-1} \geq pN], \quad (8)$$

où  $S_i$  est le nombre de succès d'une loi binomiale de probabilité  $\frac{i}{N}$  avec  $N$  répétitions.

Dans le cas où l'estimation de quantile est réalisée avec échantillonnage d'importance, la formule (7) utilisant une loi binomiale n'est plus applicable. Il faut en effet prendre en compte le fait que chaque échantillon est associé à un certain ratio de pondération  $\frac{f(y_i)}{g(y_i)}$ . Nous proposons donc la formule (9), qui se base sur le théorème central limite pour approcher une somme de variables aléatoires i.i.d. par une distribution Gaussienne.

$$P_{bs}[q = x_i] \approx P_{Z_i}[Z_i \geq p] - P_{Z_{i-1}}[Z_{i-1} \geq p], \quad (9)$$

avec  $Z_i$  une Gaussienne de moyenne  $\sum_{k=1}^i \frac{\rho_k}{N}$  et d'écart-type

$$\sqrt{\frac{\sum_{k=1}^i \rho_k \mathbb{1}(x_k \leq x_{i-1})}{N}}.$$

Un exemple de distribution d'erreur par bootstrap utilisant les formules (6) et (7) est donné en Fig. 7 dans le cas d'une estimation de quantile  $10^{-6}$  avec des échantillonnages MC et IS de  $10^5$  points. On peut remarquer que la distribution associée au Monte Carlo n'est pas du tout convergée tandis que celle par échantillonnage d'importance est bien resserrée autour de la valeur réelle -4.753.

### 2) Erreurs dues au métamodèle de krigeage

Nous proposons d'approximer les potentielles erreurs d'estimation dues à l'utilisation du métamodèle de krigeage de manière empirique. Un modèle de krigeage est en effet un modèle statistique, dont des réalisations fonctionnelles

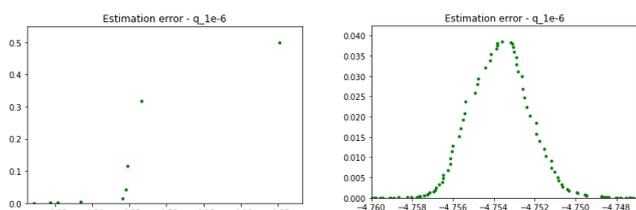


Fig. 7 Distribution bootstrap MC (à gauche) et IS (à droite) d'un quantile  $10^{-6}$ .

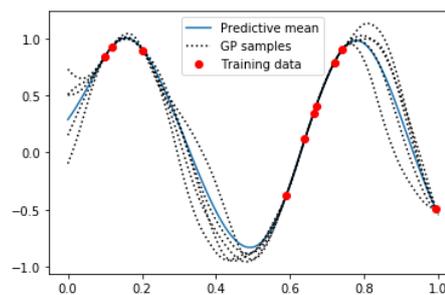


Fig. 8 Exemples de réalisations fonctionnelles.

peuvent être tirées, comme illustré en Fig. 8. La supposition sous-jacente est donc que la fonction réelle soit issue d'un processus stochastique Gaussien, dont les paramètres correspondent bien aux hyper-paramètres optimisés lors de construction du modèle. En pratique, cette supposition est généralement satisfaite, notamment quand les modèles ont de bons scores.

La méthode proposée consiste alors à tirer un grand nombre de réalisations fonctionnelles et d'estimer la statistique d'intérêt sur chacune d'elles, afin d'obtenir des échantillons de valeurs possibles pour cette statistique. Une méthode de reconstruction de densité à noyau (*Kernel Density Estimation*, KDE) est ensuite utilisée pour obtenir une distribution de probabilité continue.

La mise en application de la méthodologie proposée ici nécessite le tirage simultané d'un grand nombre de valeurs pour chaque estimation de statistique, ce qui implique un important coût de calcul, que nous évitons grâce à l'approximation par la méthode des points d'induction proposée dans [5].

Un exemple de reconstruction KDE de l'erreur due au modèle pour l'estimation d'un quantile  $10^{-9}$  est donné en Fig. 9 avec la distribution orange. Cette dispersion est comparée à l'erreur due à l'échantillonnage, présentée dans la section précédente et représentée sous la forme de la distribution discrète verte.

### 3) Couplage des erreurs

Les deux sources d'erreur décrites ci-dessus sont couplées pour obtenir un intervalle de confiance final autour de la statistique estimée. Pour cela, les deux distributions sont sommées sous supposition d'indépendance. Dans l'optique d'un taux de confiance à 90%, les quantiles 10% et 90% de cette distribution couplée seront finalement retournés.

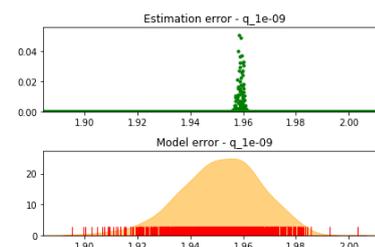


Fig. 9 Exemples d'erreurs dues à l'échantillonnage et au modèle.

## VI. RESULTATS

Dans le but d'automatiser l'obtention des résultats, un outil interne ArianeGroup a été développé. A partir de la base de données des résultats de calculs, cet outil permet dans un premier temps de générer les méta-modèles puis dans un second temps d'effectuer les simulations de Monte-Carlo sur celui-ci. Cet automatiser était indispensable au vu du nombre de méta-modèles et de calculs de probabilités qui devaient être effectués. Plus de 100 méta-modèles ont été construits et plusieurs calculs de probabilités de défaillances (variations dans les hypothèses) ont été réalisés pour chacun d'eux afin de couvrir l'ensemble des configurations possibles et de déterminer des probabilités de défaillances conservatives.

### A. Génération des méta-modèles

Dans le but de concentrer les modèles sur les zones de défaillances et de limiter les temps de calculs, le nombre de données pour construire la méta-modèle a été limité (<1000 expériences). Sur l'ensemble des méta-modèles construits, tous possèdent de très bon score (coefficient de corrélation de Pearson > 99%). Une visualisation de l'onglet pour générer les méta-modèles de l'outil industrialisé est visible en Fig. 10. On retrouve les estimateurs  $Q^2$  et  $R^2$ , ainsi que la visualisation des *goodness-of-fit* de validation et complet.

### B. Simulations de Monte-Carlo

Pour chaque méta-modèle, des simulations de Monte-Carlo comportant différentes configurations de lois de distributions (gaussienne, gaussienne tronquée, uniformes) et différentes valeurs de corrélations ont été réalisées en fonction des caractéristiques propres au cas étudié. Une visualisation de l'onglet pour calculer la probabilité de rupture de l'outil industrialisé est visible en Fig. 11, on y rentre les données d'entrée de la simulation de MC : lois de distributions, coefficient de corrélation, critères de rupture (pour



Fig. 10 Visualisation de l'outil interne ArianeGroup pour la création du méta-modèle

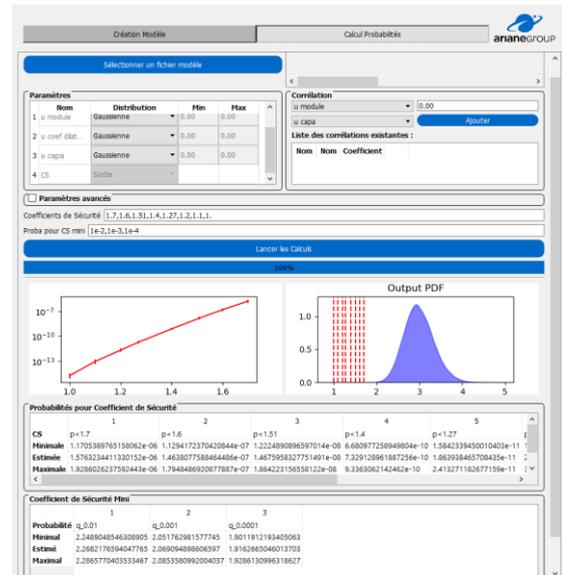


Fig. 11 Visualisation de l'outil interne ArianeGroup pour le calcul de la probabilité de défaillance

l'estimation de la probabilité de rupture) et la probabilité recherchée (pour l'estimation des quantiles). On obtient les probabilités et quantiles calculées avec estimation d'erreur pour chaque valeur de critère et de probabilité, ainsi qu'une distribution du CS étudié.

## VII. CONCLUSIONS

Les résultats obtenus ont permis de répondre aux différentes exigences de sureté de fonctionnement du produit et ainsi d'apporter une justification mécanique fine et robuste du corps de propulseur.

Les gains principaux apportés par cette méthode sont d'une part, le fait qu'elle soit basée sur des simulations ce qui limite les approximations de linéarité par exemple. D'autre part, elle apporte un gain de précision et de temps par rapport à une approche MC classique.

Une des améliorations envisagées est l'introduction de construction adaptative de plans d'expérience dont l'objectif serait de limiter le plan d'expérience aux seuls calculs nécessaires à l'évaluation de probabilité en ciblant des expériences proches de la zone de défaillance. Ce gain de temps pourrait permettre d'appliquer cette méthode sur des calculs plus complexes et chronophages (calculs 3D, thermostructuraux, ...).

## REFERENCES

- [1] G. W. Snedecor, G. W. Cochran, "Méthodes Statistiques 6e édition," Association de coordination technique agricole, pp. 148–189, 1971.
- [2] GPy, "GPy: A Gaussian process framework in python", <http://github.com/SheffieldML/GPy>, since 2012.
- [3] C. K. Williams and C. E. Rasmussen, "Gaussian processes for machine learning", Cambridge, MA: MIT press, 2006.
- [4] P. W. Glynn, "Importance sampling for Monte Carlo estimation of quantiles", Mathematical Methods in Stochastic Simulation and Experimental Design: Proceedings of the 2nd St. Petersburg Workshop on Simulation, Publishing House of St. Petersburg University, 1996.
- [5] M. Rivier, "Low-cost methods for constrained multi-objective optimization under uncertainty", Doctoral dissertation, Institut Polytechnique de Paris, 2020.