CSMA 2026
17ème Colloque National en Calcul des Structures
18-22 Mai 2026, Presqu'île de Giens (Var)

Efficient Estimation of A-basis and B-basis under Limited Data and Simplified Models

E. Donfack Siewe[1,2], J. Morio[2], S. Dubreuil[2], J.-P Navarro[1], C. Fagiano[3]

[1] *AIRBUS OPERATIONS SAS, Toulouse, France*
[2] *ONERA/DTIS, Université de Toulouse, Toulouse, France*
[3] *ONERA/DMAS, Université Paris Saclay, Châtillon, France*

Résumé — Estimation of A/B-basis values is essential for composite structure certification, traditionally grounded in extensive experimental datasets. However, this work extends this concept to numerical models, addressing the specific challenges of mixed aleatory and epistemic uncertainties inherent in virtual testing. We specifically target three sources of epistemic uncertainty : model identification, statistical estimation, and surrogate error. The proposed framework ensures an unbiased, efficient estimator and enables robust confidence bounds.
Mots clés — Conservative quantile estimation, Epistemic-based reliability analysis, Robust design.

1 General introduction

1.1 Objective

The context of this work is to ensure the reliability of complex engineering systems under uncertainty. In composite structures, materials and manufacturing processes significantly affect performance, making it essential to guarantee that requirements are met with a high degree of confidence even with limited data as required in the current regulation documentts (Cs-25) [1].

Historically, certification relies on conservative quantiles derived from extensive experimental testing, as standardized in the *Composite Materials Handbook* [2]. These are specifically defined as A-basis and B-basis values : the A-basis is the lower bound of the 95% confidence interval on the 1st percentile of the distribution (ensuring 99% reliability with 95% confidence), while the B-basis is the lower bound of the 95% confidence interval on the 10th percentile (ensuring 90% reliability).

This work targets the virtual estimation of these allowables within a numerical modeling framework. Unlike physical testing, virtual certification introduces specific challenges requiring a rigorous treatment of mixed uncertainties. Consequently, the proposed methodology explicitly integrates both aleatory and epistemic uncertainties, targeting key sources such as probabilistic model identification, statistical estimation uncertainty (due to a finite number of simulations), and surrogate model approximation error.

By providing robust confidence bounds for numerical predictions, this approach offers a transparent and theoretically justified way to derive A- and B-basis values from numerical simulations.

1.2 State of the art

Let us consider a computational model as a black-box function $\phi : \mathcal{X} \to \mathbb{R}, \quad \mathcal{X} \subseteq \mathbb{R}^d$ with input random vector $\mathbf{X} = (X_1, \dots, X_d)^\top$ of probability density function (pdf) $f_{\mathbf{X}}$. The model output is the random variable $Y = \phi(\mathbf{X})$, with unknown pdf f_Y and cumulative distribution function (cdf) F_Y. The α-quantile $\alpha \in [0,1])$ q_α of Y is defined as :

$$q_\alpha = \inf\{q \in \mathbb{R} \mid F_Y(q) \geq \alpha\}. \tag{1}$$

We denote by $\widehat{q}_\alpha$ a quantile estimator of q_α obtained by sampling method denoted by m (in practice m can be Monte Carlo Smpling, Importance Sampling...). In that case an asymptotic confidence interval at confidence level β can be generally derived and reads [3] :

$$\left[\widehat{q}_\alpha - z_{(1+\beta)/2} \cdot \sqrt{\mathbb{V}(\widehat{q}_\alpha)},\ \widehat{q}_\alpha + z_{(1+\beta)/2} \cdot \sqrt{\mathbb{V}(\widehat{q}_\alpha)}\right], \tag{2}$$

where $z_{(1+\beta)/2}$ denotes a $(1+\beta)/2$-quantile of the standard normal distribution. In composites structures certification, only the lower bound of this interval is relevant :

$$\widehat{q}_{\alpha,\beta,\inf} = \widehat{q}_\alpha - z_{(1+\beta)/2}\sqrt{\mathbb{V}(\widehat{q}_\alpha)}, \tag{3}$$

which defines the A- and B-basis values for $(\alpha,\beta) = (0.01, 0.95)$ and $(0.1, 0.95)$, respectively.

Several classical methods exist for quantile estimation, including the Monte Carlo method [4] and various variance reduction techniques such as Importance Sampling (IS) [5], Control Variates (CV) [6], and Stratified Sampling [7]. In industrial practice, however, the most widely adopted approach remains the Wilks estimator [8], which provides conservative quantile bounds without requiring explicit assumptions on the underlying distribution.

Example. For $\beta = 0.95$ and $\alpha = 0.01$, Wilks' formula ensures that the smallest observation $Y_{(1)}$ in a sample of $N = 299$ exceeds the 1st percentile with 95% confidence : $\mathbb{P}(q_{0.01} \geq Y_{(1)}) \geq 0.95$. This conservative approach requires large samples and assumes perfect knowledge of input distributions, which is rarely the case in computationally expensive aerospace models.

These traditional estimators thus form the basis of conservative quantile estimation under idealized conditions. In practical engineering applications, however, limited data, restricted computational budgets, and surrogate model errors introduce significant *epistemic* uncertainties. The following section introduces a unified framework to address these challenges and provide statistically sound A- and B-basis estimates under mixed uncertainty.

2 Sources of epistemic uncertainty

Classical quantile estimators rely on idealized assumptions, such as perfect knowledge of input distributions, large independent and identically distributed samples, and negligible model error. In industrial applications, these assumptions rarely hold, leading to **epistemic uncertainty** due to limited knowledge that can, in principle, be reduced with additional data (See Figure 1).

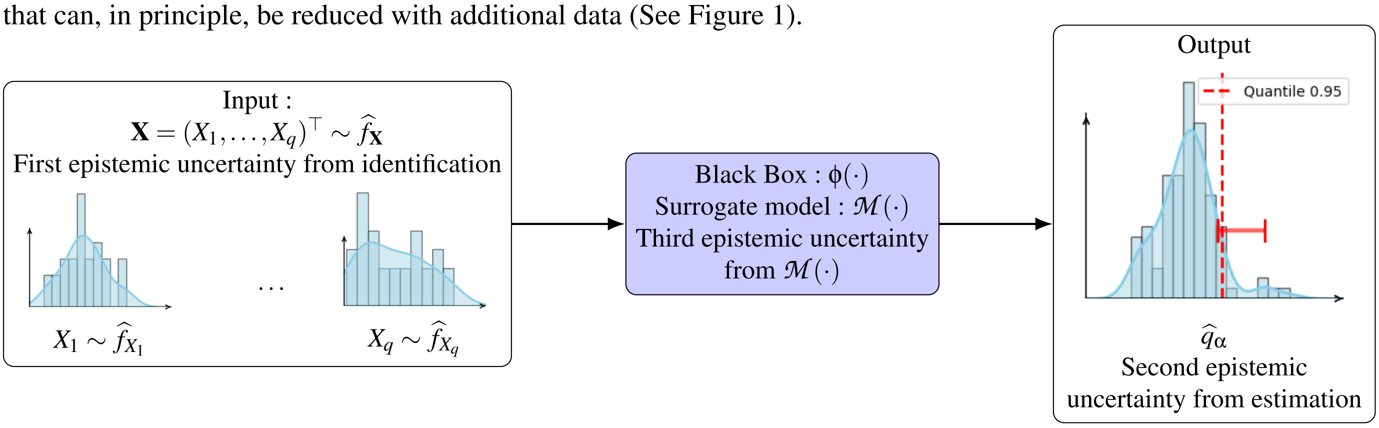

FIGURE 1 – General schematic of the uncertainty propagation for quantile estimation with true black box and surrogate model.

Three main sources of epistemic uncertainty affect quantile estimation :

1. **Input density estimation [9, 10] :** The true input distribution $f_{\mathbf{X}}$ is rarely fully known and may be inferred from limited experimental samples using parametric (e.g., Maximum Likelihood or Method of Moments) or non-parametric (e.g., Kernel Density Estimation) approaches. Small sample sizes can lead to poor model identification and an underestimation of variance. In the literature, this source of uncertainty is often addressed by applying bootstrap resampling [11] to the available experimental data.
2. **Finite simulation samples [9] :** Monte Carlo-based quantile estimators have variances that depend on the number of simulations. Increasing the number of simulations improves accuracy but is often computationally expensive for high-fidelity models. In the literature, this source of uncertainty is often addressed by applying bootstrap resampling [11] to the available simulation data.

3. **surrogate approximation [12] :** Surrogate models are used to reduce computational cost, but they introduce additional uncertainty due to limited training data or overfitting. This uncertainty must be quantified and propagated, for example through stochastic surrogate model realizations, bootstrap, or Bayesian methods, to ensure robust and credible quantile estimates.

Overall, accounting for all these sources of epistemic uncertainty is essential to obtain reliable and statistically sound estimates of conservative quantiles, especially in certification-relevant applications.

3 Proposed quantile estimation using CV and IS

The primary objective of this work is to account for the three different sources of epistemic uncertainty previously introduced and propagate them through the computational model. Our strategy combines two complementary ingredients : Importance Sampling (IS) and Control Variates (CV). IS allows independent treatment of two sources of epistemic uncertainty—probabilistic model identification and Monte Carlo sampling by generating simulation samples from an auxiliary distribution that differ from the distributions inferred from experimental data. CV leverages the surrogate model to reduce variance without introducing bias or compromising estimator accuracy.

In practice, IS and CV are jointly applied to estimate the cdf of the output, from which the quantile of interest, $\widehat{q}_\alpha$, is obtained along with its associated confidence interval.

3.1 Estimation of the cumulative distribution function by CV and IS

We assume a surrogate model $\mathcal{M}$ for which the quantile z_α is pre-estimated with high precision. The input distribution is denoted by $f_\mathbf{X}$, and we introduce a new auxiliary density g from which the Monte Carlo samples will be drawn.

We propose the following unbiased estimate of the cdf of Y, coupling CV and IS (CV-IS) :

$$\widehat{F_Y^{\text{CV-IS}}}(y) = \frac{1}{N_{\text{sim}}}\sum_{i=1}^{N_{\text{sim}}} \mathbf{1}_{\phi(\mathbf{X}_i)\leq y}\frac{f_\mathbf{X}(\mathbf{X}_i)}{g(\mathbf{X}_i)} - C(y)\left(\frac{1}{N_{\text{sim}}}\sum_{i=1}^{N_{\text{sim}}} \mathbf{1}_{\mathcal{M}(\mathbf{X}_i)\leq z_\alpha}\frac{f_\mathbf{X}(\mathbf{X}_i)}{g(\mathbf{X}_i)} - \mathbb{E}_{f_\mathbf{X}}\left[\mathbf{1}_{\mathcal{M}(\mathbf{X})\leq z_\alpha}\right]\right) \tag{4}$$

$$:= \widehat{F_Y^{\text{IS}}}(y) - C(y)(\widehat{h^{\text{IS}}} - \mathbb{E}_{f_\mathbf{X}}\left[\mathbf{1}_{\mathcal{M}(\mathbf{X})\leq z_\alpha}\right]), \tag{5}$$

with

$$\mathbf{X}_i \sim g, \quad \widehat{h^{\text{IS}}} = \frac{1}{N_{\text{sim}}}\sum_{i=1}^{N_{\text{sim}}} \mathbf{1}_{\mathcal{M}(\mathbf{X}_i)\leq z_\alpha}\frac{f_\mathbf{X}(\mathbf{X}_i)}{g(\mathbf{X}_i)}, \quad \mathbb{E}_{f_\mathbf{X}}\left[\mathbf{1}_{\mathcal{M}(\mathbf{X})\leq z_\alpha}\right] = \alpha \quad \text{and} \quad \lambda^{\text{IS}}(\mathbf{X}) = \frac{f_\mathbf{X}(\mathbf{X})}{g(\mathbf{X})}.$$

$\mathbb{E}$ is the mathematical expectation. The control variable, denoted here $\mathbf{1}_{\mathcal{M}(\cdot)\leq z_\alpha}\lambda^{\text{IS}}(\cdot)$, is based on the surrogate $\mathcal{M}$. Since the quantile z_α of the surrogate is pre-estimated with high precision, its theoretical mean is α. It is used to correct the importance sampling estimator ($\widehat{F_Y^{\text{IS}}}(y)$) by exploiting the correlation between the model ϕ and the surrogate $\mathcal{M}$.

The optimal (in terms of variance reduction) parameter $C(y)$ is defined by [13] :

$$C(y) = \frac{\text{Cov}_g[\widehat{F_Y^{\text{IS}}}(y), \widehat{h^{\text{IS}}}]}{\mathbb{V}_g[\widehat{h^{\text{IS}}}]},$$

where Cov and $\mathbb{V}$ denote the covariance and variance under the simulation distribution. Using this optimal value, the variance of the cdf estimator in Eq. (5) becomes :

$$\mathbb{V}_g[\widehat{F_Y^{\text{CV-IS}}}(y)] = \mathbb{V}_g[\widehat{F_Y^{\text{IS}}}(y)]\,[1-\rho^2(y)] = \frac{\mathbb{E}_g[\mathbf{1}_{\phi(\mathbf{X})\leq y}\lambda^{\text{IS}}(\mathbf{X})^2] - F_Y(y)^2}{N_{\text{sim}}}[1-\rho^2(y)],$$

where $\rho(y)$ is the correlation coefficient between $\mathbf{1}_{\phi\leq y}\lambda^{\text{IS}}$ and $\mathbf{1}_{\mathcal{M}\leq z_\alpha}\lambda^{\text{IS}}$. For **strong correlation** ($\rho(y) \approx \pm 1$), the variance is greatly reduced and the control variate is highly effective. For **weak correlation** ($\rho(y) \approx 0$), there is little to no variance reduction : the control variate provides no significant benefit, but it does not degrade the estimator either. In practice, this means that even if the surrogate model is inaccurate, the base estimator remains unbiased and unaffected. This robustness is one of the main advantages of using control variates.

3.2 Quantile estimation by CV-IS

Using the cdf CV-IS estimator given by Eq. (5), the α-quantile estimator is :

$$\widehat{q_\alpha^{\text{CV-IS}}} = \inf\left\{ q \in \mathbb{R} \ \middle| \ \widehat{F_Y^{\text{CV-IS}}}(q) \geq \alpha \right\}. \tag{6}$$

This estimator is asymptotically normal with the reduced variance :

$$\mathbb{V}_g[\widehat{q_\alpha^{\text{CV-IS}}}] = \frac{\mathbb{E}_g[\mathbf{1}_{\phi(\mathbf{X})\leq q_\alpha}\lambda^{\text{IS}}(\mathbf{X})^2] - \alpha^2}{N_{\text{sim}} f_Y^2(q_\alpha)}[1-\rho^2(q_\alpha)], \tag{7}$$

with

$$\rho(q_\alpha) = \text{Cor}_g[\mathbf{1}_{\phi\leq q_\alpha}\lambda^{\text{IS}}, \mathbf{1}_{\mathcal{M}\leq z_\alpha}\lambda^{\text{IS}}].$$

As for the cdf, for **good correlation :** $\rho(q_\alpha) \approx \pm 1$ the variance is greatly reduced and the control variate is highly effective and for **poor correlation :** $\rho(q_\alpha) \approx 0$ there is little to no variance reduction : the control variate provides no significant benefit, but it does not degrade the estimator either.

4 Quantile estimation under three sources of epistemic uncertainty

We will consider a given scenario where the following elements are available :

— a probabilistic input model has been identified from a given experimental dataset of size N_{test}, denoted $\widetilde{\mathbf{x}}_{\text{obs}}$. From this dataset, a conditional input distribution $\widehat{f}_{\mathbf{X}|\widetilde{\mathbf{X}}_{\text{obs}}}(\cdot \mid \widetilde{\mathbf{x}}_{\text{obs}})$ has been inferred.

— a Monte Carlo (MC) simulation sample of size N_{sim}, denoted $\widetilde{\mathbf{x}} = (\mathbf{x}_1, \ldots, \mathbf{x}_{N_{\text{sim}}})$, is available, with independent and identically distributed components following a known auxiliary distribution g.

— a stochastic surrogate model $\mathcal{M}$ (such as a Gaussian Process) trained on N_{DoE} samples, for which a realization (or trajectory) $M(\cdot)$ is available and the quantile z_α is known (can be easily estimated with high precision).

The quantile estimator proposed in Eq. (6) can be interpreted as a black-box that takes the three previous parameters as inputs and returns a quantile :

$$\widehat{q_\alpha^{\text{CV-IS}}}(M, \widetilde{\mathbf{x}}_{\text{obs}}, \widetilde{\mathbf{x}}) = \inf\left\{ q \in \mathbb{R} \ \middle| \ \widehat{F_Y^{\text{CV-IS}}}(q, M, \widetilde{\mathbf{x}}_{\text{obs}}, \widetilde{\mathbf{x}}) \geq \alpha \right\}, \tag{8}$$

with

$$\widehat{F_Y^{\text{CV-IS}}}(y, M, \widetilde{\mathbf{x}}_{\text{obs}}, \widetilde{\mathbf{x}}) = \frac{1}{N_{\text{sim}}}\sum_{i=1}^{N_{\text{sim}}} \mathbf{1}_{\phi(\mathbf{x}_i)\leq y} \frac{\widehat{f}_{\mathbf{X}|\widetilde{\mathbf{X}}_{\text{obs}}}(\mathbf{x}_i \mid \widetilde{\mathbf{x}}_{\text{obs}})}{g(\mathbf{x}_i)} - \widehat{C}(y)\left(\frac{1}{N_{\text{sim}}}\sum_{i=1}^{N_{\text{sim}}} \mathbf{1}_{M(\mathbf{x}_i)\leq z_\alpha} \frac{\widehat{f}_{\mathbf{X}|\widetilde{\mathbf{X}}_{\text{obs}}}(\mathbf{x}_i \mid \widetilde{\mathbf{x}}_{\text{obs}})}{g(\mathbf{x}_i)} - \alpha \right).$$

By generating N bootstrap samples of $\widetilde{\mathbf{x}}_{\text{obs}}$ and $\widetilde{\mathbf{x}}$, and N trajectories of $\mathcal{M}$, we obtain the quantiles sample :

$$\left(\widehat{q_\alpha^{\text{CV-IS}}}(\mathcal{M}_i, \widetilde{\mathbf{X}}_{\text{obs}^i}, \widetilde{\mathbf{X}}_i) \right)_{1\leq i\leq N},$$

which represent the samples of the distribution $\widehat{q}_\alpha^{\text{CV-IS}}(\mathcal{M}, \widetilde{\mathbf{X}}_{\text{obs}}, \widetilde{\mathbf{X}})$ in the augmented space corresponding to the three sources of uncertainty where $(\mathcal{M}, \widetilde{\mathbf{X}}_{\text{obs}}, \widetilde{\mathbf{X}})$ are the random variables. From this distribution, one can obtain its β-level confidence interval by :

$$\left[\widehat{q_{\alpha,\text{inf}}^{\text{CV-IS}}}(\beta),\ \widehat{q_{\alpha,\text{sup}}^{\text{CV-IS}}}(\beta) \right], \tag{9}$$

where $\widehat{q_{\alpha,\text{inf}}^{\text{CV-IS}}}(\beta)$, and $\widehat{q_{\alpha,\text{sup}}^{\text{CV-IS}}}(\beta)$ are the empirical quantiles at levels $(1-\beta)/2$ and $(1+\beta)/2$ of the sample $(\widehat{q_\alpha^{\text{CV-IS}}}(\mathcal{M}_i, \widetilde{\mathbf{X}}_{\text{obs}^i}, \widetilde{\mathbf{X}}_i))_{1\leq i\leq N}$. The A-basis and B-basis values correspond directly to the lower confidence bound quantile estimator for specific α and β :

$$\text{A-basis} = \widehat{q_{0.01,\text{inf}}^{\text{CV-IS}}}(0.95) \quad \text{and,} \quad \text{B-basis} = \widehat{q_{0.1,\text{inf}}^{\text{CV-IS}}}(0.95).$$

The estimates obtained using this approach are highly consistent and explicitly account for all epistemic uncertainties.

5 Numerical applications

The feasibility and relevance of the proposed approach are illustrated through numerical examples. In the considered case studies, it is important to note that the probabilistic model is assumed to be known solely for the purpose of generating the reference test sample $\widetilde{\mathbf{x}}_{\text{obs}}$, composed of N_{test} experiments (See example in Figure 3).

The auxiliary density used in this case for simulation sample is :

$$\mathbf{x} \in \mathbb{R}^d, \quad g(\mathbf{x}) = \frac{1}{N}\sum_{\ell=1}^{N} \widehat{f}_{\mathbf{X}|\widetilde{\mathbf{X}}_{\text{obs}}}(\mathbf{x} \mid \widetilde{\mathbf{X}}_{obs^{\ell}}).$$

This density is used for the purpose of generating the reference simulation sample $\widetilde{\mathbf{x}}$, composed of N_{sim} samples.

The surrogate models employed are Gaussian processes, for which the trajectories can be efficiently generated using the Karhunen–Loève (KL) decomposition [15].

Furthermore, the bootstrap (BS) sample size is fixed at $N = 10^3$ and the objective is to estimate both A-basis and B-basis under the presence of three sources of epistemic uncertainty.

5.1 Description of the model

Consider a cantilever beam test case where the mean deflection at the free end is induced by the applied load represented in Figure 2. The function ϕ writes in the following analytical form with d = 5 uncertain inputs :

$$\phi(F, L, E_{YM}, b, h) = \frac{4FL^3}{E_{YM}bh^3} \tag{10}$$

where F is the transverse load applied at the free end, L is the beam length, E_{YM} is the Young's modulus and bh is the cross-section. The actual distribution of each input variable are taken from [14] (see Table 1) and all the variables are assumed independent. This probabilistic model is assumed to be unknown and only enables the generation of the N_{test}-experiment sample on which a KDE (kernel density estimation) is performed (See example in Figure 3).

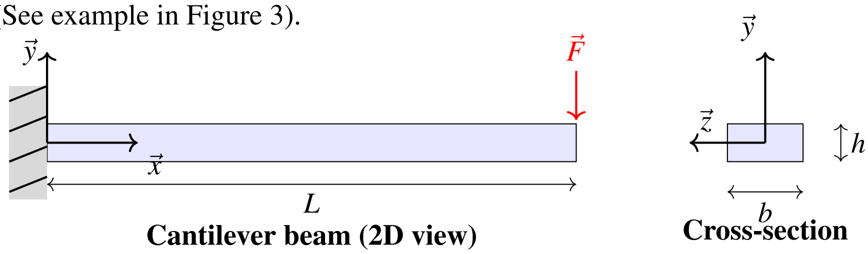

FIGURE 2 – Illustration of a cantilever beam of length L, Young's modulus E_{YM} and cross-section $b \times h$ with an applied transverse load F.

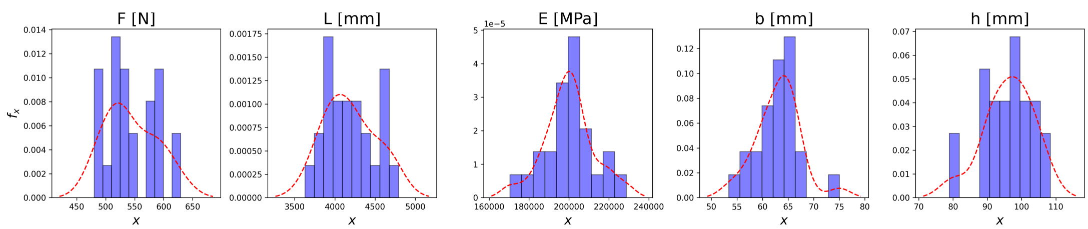

FIGURE 3 – Illustration of database used for test sample.

TABLE 1 – Input random variable distributions for the cantilever beam test case. CV denotes the coefficient of variation.

Variable	**Symbol, Units**	**Distribution**	**Mean**	**CV**
Transverse load	F, N	Lognormal	556.8	0.08
Length	L, mm	Normal	4290	0.10
Young's modulus	E_{YM}, MPa	Lognormal	2.105	0.06
Width	b, mm	Normal	62	0.10
Height	h, mm	Normal	98.7	0.10

5.2 Results

The results are obtained for different given values of $(N_{\mathrm{DoE}}, N_{\mathrm{test}}, N_{\mathrm{sim}})$. The description of these values are in Table 2.

TABLE 2 – Definition of sample sizes corresponding to the three sources of epistemic uncertainty considered in this framework.

Notation	**Description**
N_{test}	Number of physical experimental data points available to identify the probability distributions of the input parameters.
N_{DoE}	Number of high-fidelity model evaluations (Design of Experiments) used to build and train the surrogate model.
N_{sim}	Number of numerical statistical samples used to compute the quantile estimator.

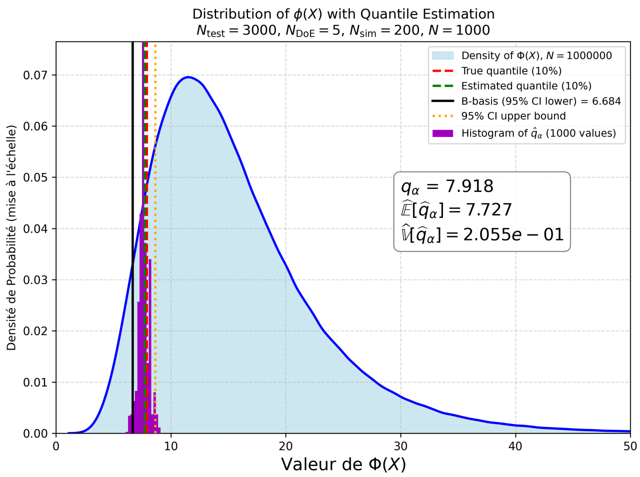

FIGURE 4 – Variance of the estimated quantity arising solely from sampling in simple Monte Carlo simulations. The figure displays the histogram of the estimated quantile and the corresponding B-basis.

The configuration shown in Figure 4 corresponds to $(N_{\mathrm{sim}}, N_{\mathrm{test}}, N_{\mathrm{DoE}}) = (200, 3000, 3)$. In this scenario, the input distributions are accurately identified (N_{test} is large), but the surrogate model is poorly trained (N_{DoE} is small). This figure illustrates the quantile estimator ($\widehat{q}_\alpha$), its estimated variance ($\widehat{\mathbb{V}}[\widehat{q}_\alpha]$), and the resulting B-basis value under the effect of epistemic uncertainty.

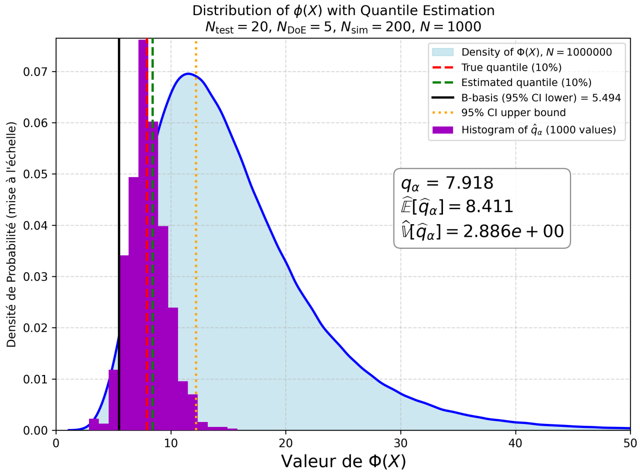

FIGURE 5 – Variance of the estimated quantity resulting from both Monte Carlo sampling and limited input data. The figure displays the histogram of the estimated quantile and the corresponding B-basis.

The configuration shown in Figure 5 corresponds to $(N_{\text{sim}}, N_{\text{test}}, N_{\text{DoE}}) = (200, 20, 3)$. In this scenario, the input data are limited ($N_{\text{test}} = 20$), leading to high variability, and the surrogate model is poorly trained ($N_{\text{DoE}} = 3$). The figure clearly shows a significantly increased estimator variance, which is directly caused by the limited input data. This demonstrates that even with a poorly trained surrogate model, the estimator remains unbiased and robust, thanks to the use of control variates.

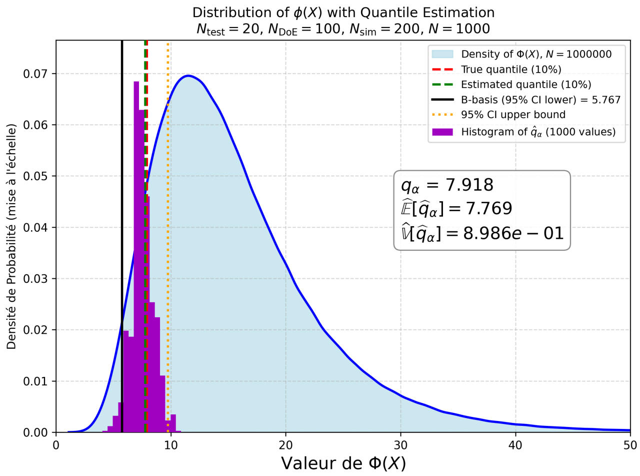

FIGURE 6 – Variance resulting from simple Monte Carlo simulations and limited input data using an accurate surrogate model. The figure displays the histogram of the estimated quantile and the corresponding B-basis.

The configuration shown in Figure 6 corresponds to $(N_{\text{sim}}, N_{\text{test}}, N_{\text{DoE}}) = (200, 20, 100)$. In this scenario, the input data are limited ($N_{\text{test}} = 20$), but the surrogate model is well trained ($N_{\text{DoE}} = 100$). This configuration leads to a significant reduction in the variance of the estimated quantile. The low variance confirms that the surrogate model effectively accelerates convergence without compromising the estimator's accuracy, a benefit attributed to the Control Variates method.

A comparative summary of the numerical results for these three configurations is provided in Table 3, illustrating the specific impact of input identification, simulation size, and surrogate quality on the quantile estimator.

TABLE 3 – Summary of results for the three configurations. It highlights the impact of input identification (N_{test}), simulation sample (N_{sim}), and surrogate quality (N_{DoE}) on the estimators $\widehat{\mathbb{E}}[\widehat{q}_\alpha]$ and $\widehat{\mathbb{V}}[\widehat{q}_\alpha]$.

Case	$\mathbf{N_{sim}}$	$\mathbf{N_{test}}$	$\mathbf{N_{DoE}}$	$\widehat{\mathbb{E}}[\widehat{\mathbf{q}}_\alpha]$	$\widehat{\mathbb{V}}[\widehat{\mathbf{q}}_\alpha]$	Observation
Fig. 4	200	3000	5	7.727	2.06×10^{-1}	Low variance due to accurate identification.
Fig. 5	200	20	5	8.411	2.89×10^{0}	High variance caused by limited data (N_{test}).
Fig. 6	200	20	100	7.769	8.99×10^{-1}	Significant variance reduction via Control Variates.

6 Conclusion

This work presents a unified and rigorous framework for quantile estimation based on numerical simulation under combined aleatory and epistemic uncertainties, explicitly including probabilistic model identification, statistical estimation, and surrogate model approximation errors. The methodology proposes a robust and statistically justified way to derive A- and B-basis values from numerical simulations, especially in a context with limited data and fixed computational budgets.

A key feature is the use of variance reduction via Control Variates , where the surrogate model acts as a control tool rather than a surrogate replacement, ensuring unbiased and consistent estimators.

The approach is currently being evaluated in a multiscale framework for stiffened composite panels within an industrial Research and Technology (R&T) context.

Références

[1] E. U. A. S. Agency, *Certification Specifications for Large Aeroplanes : CS-25*, Technical Report, EASA, 2019.

[2] CMH-17, *Composite Materials Handbook, Volume 17 : Polymer Matrix Composites*, ASTM International, 2012.

[3] R. J. Serfling, *Approximation Theorems of Mathematical Statistics*, John Wiley & Sons, New York, 1980.

[4] R. A. S. Cardoso, et al., *A-Basis and B-Basis Buckling Allowables for Composite Semi-Wing – Application of Monte Carlo Simulations and Sensitivity Analysis*, in Proceedings of the 34th Congress of the International Council of the Aeronautical Sciences (ICAS), 2024.

[5] S. Tokdar and R. Kass, *Importance Sampling : A Review*, Technical Report TR-09-01, Department of Statistical Science, Duke University, 2009.

[6] F. J. Hickernell, C. Lemieux, and A. B. Owen, *Control Variates for Quasi-Monte Carlo*, Technical Report, Hong Kong Baptist University, University of Calgary, and Stanford University, this version : April 2003, original : August 2001.

[7] H. M. Samawi, O. A. Al-Sagheer, and A. I. Al-Omari, *Stratified Sampling and Its Applications in Statistics*, *Communications in Statistics - Simulation and Computation*, vol. 48, no. 3, pp. 664–676, 2019.

[8] S. S. Wilks, *Determination of Sample Sizes for Setting Tolerance Limits*, *Annals of Mathematical Statistics*, vol. 12, no. 1, pp. 91–96, 1941.

[9] C. Surget, S. Dubreuil, J. Morio, C. Mattrand, J. Bourinet, and N. Gayton, “A sensitivity analysis based trade-off between probabilistic model identification and statistical estimation,”

[10] G. Sarazin, “Analyse de sensibilité fiabiliste en présence d'incertitudes épistémiques introduites par les données d'apprentissage,” Ph.D. thesis, Institut Supérieur de l'Aéronautique et de l'Espace, Toulouse, 2021.

[11] B. Efron, *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia : SIAM, 1982.

[12] M. Menz, “Réduction du coût numérique des méthodes d'apprentissage actif pour l'analyse de fiabilité,” Ph.D. thesis, Université de Toulouse, 2020.

[13] C. Cannamela, J. Garnier, and B. Iooss, “Controlled Stratification for Quantile Estimation,” *Annals of Applied Statistics*, vol. 2, no. 4, pp. 1554–1580, 2008.

[14] B. Li, L. Zhang, X. Zhu, X. Yu, and X. Ma, “Reliability analysis based on a novel density estimation method for structures with correlations,” *Chinese Journal of Aeronautics*, vol. 30, no. 3, pp. 1021–1030, 2017.

[15] J. Villemonteix, E. Vazquez, and E. Walter. An informational approach to the global optimization of expensive-to-evaluate functions. *Journal of Global Optimization*, 44 :509–534, 2009.