Stochastic mesoscale characterization of ablative materials for atmospheric entry

Florian Girault, Francisco Torres Herrador, Bernd Helber, Alessandro Turchi, Thierry Magin, Pietro Marco Congedo

Stochastic mesoscale characterization of ablative materials for atmospheric entry

F. Girault[a], F. Torres Herrador, B. Helber, A. Turchi, T. Magin[b], P.M. Congedo*[a,]

[a] *Inria, Center for Applied Mathematics, Ecole Polytechnique, Institut Polytechnique de Paris, Route de Saclay, 91120 Palaiseau, France*
[b] *von Karman Institute for Fluid Dynamics, Chaussée de Waterloo 72, 1640 Rhode-Saint-Genèse, Belgium*

Abstract

This work aims to estimate material properties at a mesoscopic level using the PuMA Software from an uncertainty quantification perspective. The stochastic behavior of PuMA, mainly related to the random distribution of the fibers, is an intrinsic source of uncertainty. The choice of some physical parameters, such as the fiber's thermal conductivity, is an additional source of uncertainties. The first contribution is a low-cost surrogate-based methodology with an unequal allocation scheme applied for the first time to the stochastic mesoscale characterization of ablative materials. A second contribution of this work is the uncertainty propagation and sensitivity analysis of the material properties, which also yields a systematic assessment of the choice of the voxel resolution for both the fibers and the domain. Precisely, the convergence of the quantities of interest can be surveyed, thus identifying the minimal reference elementary volume.

Keywords: uncertainty quantification, kriging, carbon/phenolic ablators, thermal protection systems, porous materials, microstructure, thermal conductivity.

Nomenclature

$\hat{V}_{s^2}()$ Variance estimate of the sampling variance

$\hat{V}_{SK}$	Mean predictor of the variance for Stochastic Kriging
$\hat{Y}(\mathbf{x}_0)$	Mean squared error optimal predictor
$\mathcal{M}(\mathbf{x})$	Realization of a mean zero random field
$\mathcal{Y}_j()$	Output of the j^{th} replication simulation
Σ_ε	Covariance matrix implied by the intrinsic noise
Σ_M	Correlation covariance matrix
$\varepsilon_j(\mathbf{x})$	*Intrinsic* noise
$\widehat{\mathbb{E}_G}$	Monte Carlo estimators of expectation
$\widehat{V}^\star$	Heteroscedastic predictor for the variance
$\{\mathbf{S}_i^b\}_{b=1,\ldots,B}$	B bootstrap samples
D_{vx}	Number of voxels of the domain size
$G(\mathbf{x})$	Random variable with values in $\mathbb{R}^n$
l_{vx}^{phy}	Voxel Length
n_i	Number of simulation replications
R_{phy}	Average radius of fiber
R_{vx}	Number of voxels of the radius
$\mathrm{S}^2()$	Estimated variance for Stochastic Kriging
$\mathbf{x}$	Set of design parameters
$\widehat{\mathbb{V}_G}$	Monte Carlo estimators of the variance
S_{ch}	Sobol index related to the Poisson process
S_{mix}	Sobol index related to both random inputs
S_{par}	Sobol index related to parameters

1. Introduction

Atmospheric entry is the passage of an object from outer space into the gases of the atmosphere of a planet. While traveling in the atmosphere, these bodies usually exceed sound speed at $7 - 8km.s^{-1}$. At these speeds, detached shock waves form around the object due to the intense compression of the particles over a very short distance. During this compression, a significant fraction of the vehicle's kinetic energy is transferred to the particles in the flow. This energy transfer results in a temperature rise of several thousand Kelvin and high chemical activity, which causes loss of mass (ablation) or even complete disintegration of the object. When dealing with manufactured spacecraft, it is essential to protect its integrity using a Thermal Protection System (TPS). The new generations of TPS ablative materials usually comprise an organic precursor such as Carbon Fiber (CF) felts and a polymeric isotropic matrix [1]. These materials can dissipate the impinging hyperthermal fluxes through a disintegration process by transforming the received thermal energy into chemical processes and degrading gradually. In contrast, the remaining virgin material can insulate the spacecraft. The design of the TPS still requires several improvements in physical models, experimental validation, and numerical predictions.

Characterization of material physical properties through numerical modeling is a vast and complex area of research. Thousands of contributions have been produced for many mechanical topics such as stress resistance, viscoelasticity, thermo-mechanical coupling, and fracture. This paper, however, focuses on porous materials such as those employed in TPS. Optimizing their properties requires a deep understanding of how they conduct, ensuring the correct quantity of material is used and correctly applied to the spacecraft. Some papers investigated the influence of uncertainties in porous media [2, 3]. Several papers already contributed to studying specifically carbon porous materials. Ref. [4] proposes a stochastic modeling method of carbon fiber reinforced composites with polynomial chaos. In [5, 6], they study properties of such material and the influence of design parameters thanks to Finite Element Analysis and tomog-

raphy. In [7], the impact of the very orientation of the fibers in high-fidelity simulations is considered. Finally, in [8, 1], heat capacity and pyrolysis of carbon composites at high temperature in the context of TPS design in the case of pyrolysis are studied.

The study presented in this paper considers uncertainties on the thermal conductivity within a charred material. All materials are manufactured with uncertain parameters, and controlling the impact of uncertainty on the variability of the Quantities of Interest (QOI) is necessary for reliability. Uncertainty Quantification (UQ) seeks to characterize and potentially reduce these uncertainties of models or physical systems. UQ has been applied to mechanical models, and some procedures have been described in [9]. Concerning TPS, some papers have already used UQ methodology: sensitivity analysis on pyrolysis gas composition [10]; the impact of several chemical reactions for laminar and turbulent aeroheating predictions for Mars entry [11]; uncertainty analysis for thermochemical ablation modeling for estimating effects of the uncertain estimates of a plasma wind-tunnel test conditions on the final results of the model [12]; effects of adaptations like dimension reduction and grid refinement on parameter influence over material temperature response and total recession [13].

In our study, we use the Porous Material Analysis (PuMA) Software developed by NASA [14] to obtain fiber intrinsic conductivities from effecting (macroscopic) thermal conductivity. This numerical tool can perform mechanical properties computations on digitized carbon preform domain. Its particularity is that it does not make an average computation on large domains like other existing software but rather on small domains of the typical size of a few hundred micrometers. The properties obtained in PuMA can then be used to inform macroscopic scale models. For now, some contributions have been made thanks to PuMA: [15] for weave generation capabilities; [16] for a preliminary analysis of the influence of microstructure on macroscopic properties. Note that properties can change when the size of the sample domain is decreased by introducing effects that were averaged out at bigger scales. This behavior is especially true

for highly porous materials. For this material, made of carbon fibers of the typical diameter of a few μm, porosity is of the order of 0.9, which means only 10 % of the entire volume is indeed conductive solid; the remaining part is schematically either gas or cohesive resin. Ultimately, averaging porosity and other parameters/properties may lead to underestimating the variations at smaller scales. The designing process should then identify the design parameters, their associated uncertainties, and the causes of non-parametric stochasticity.

The first contribution of this paper is to propose, for the first time in literature, a methodology for building a surrogate of the thermal conductivity of porous materials at a mesoscale level as a function of several geometrical and physical parameters. In particular, we propose a Stochastic Kriging approach with an unequal allocation methodology for replica, permitting the building of a surrogate at a low computational cost. We aim to build a surrogate for prediction purposes, uncertainty quantification, and sensitivity analysis. In particular, we consider uncertainties on some physical parameters and the intrinsic stochasticity of the PuMA solver. The second contribution is an uncertainty propagation analysis permitting the quantification of the variability of some material properties with respect to the sources of uncertainties considered. In particular, we assess the prediction of some quantities of interest when designing a given material regarding domain size and voxel resolution. This procedure could lead to the capability of preparing an experimental tomography campaign by computing the minimal size a sample should have to be representative and what variabilities should be expected when observing a material.

Section 2 is dedicated to briefly describing the models and capabilities of PuMA software. Then, in Section 3, we formalize the problem under a UQ perspective. In Section 4, we introduce the numerical tools needed for building the surrogate model and proceed to a sensitivity analysis. Finally, Section 5 illustrates the application of the proposed methodology to the real mechanical problem estimating the influence of each parameter over the output variability. Section 6 draws some conclusions and perspectives.

2. PuMA and geometrical resolution

In this section, we introduce some basics of the PuMA software and then focus on its parameters and geometrical resolution.

2.1. PuMA description

PuMA is a software developed by NASA to calculate material properties at the microscopic scale. From a 3D material sample, PuMA can compute macroscopic geometric properties such as tortuosity or porosity and physical properties such as thermal and electrical conductivities. Version 2.2 of PuMA is used for this study. PuMA divides the sample into core materials. The fiber-made porous material mainly comprises the fibers and the surrounding gas (see Fig. 1). Other features can be added, such as several types of fibers, sticky resin, gas concentrations, etc. PuMA permits the definition of internal subdomains to a certain extent.

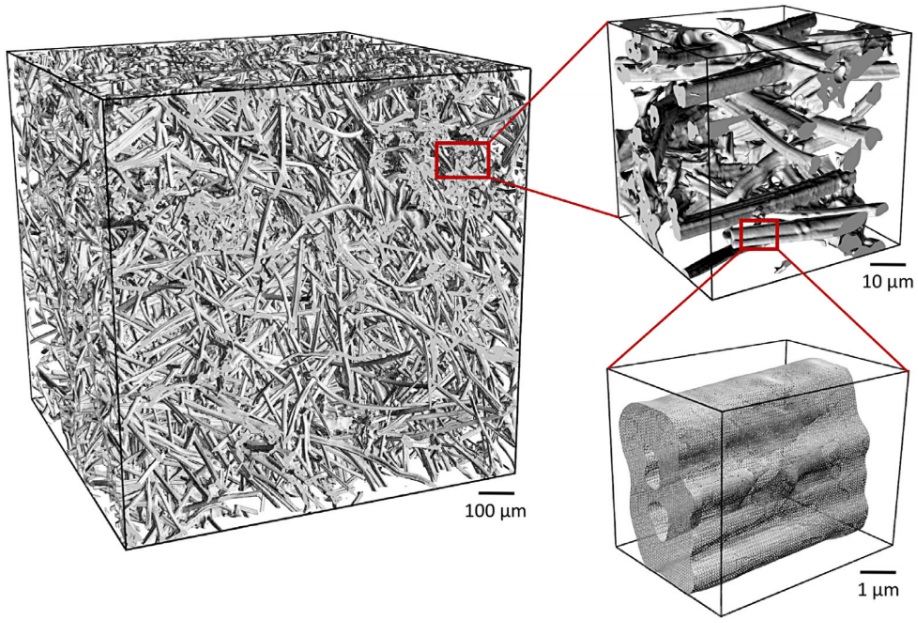


Figure 1: Typical microstructures dealt with by PuMA at several scales. Credits to [14].

The domain can be defined in two different ways. The first one is import-

ing a microtomography image into PuMA [6], such as a TIFF image: internal subdomains then correspond to an interval of greyscale. The second one is artificial generation. Given specific parameters and shapes, PuMA is capable of randomly generating domains. This second option is the one systematically used throughout this paper. In these domains, the basic length is the voxel, which is tantamount to pixel but in 3D. Once the domain is defined, PuMA can compute several QoI, for instance, the porosity p, the tortuosity, and thermal conductivity $\boldsymbol{\lambda}$, which depend on the material microstructure and its materials' conductivities.

This paper mainly focuses on the mean thermal conductivity of a sample, accounting for all the material constituents. The material is assumed to be locally isotropic. The so-called EJ-Heat method [17] is used, which relies on a finite-difference method to calculate the effective thermal conductivity of a material based on its constituent parts.

2.2. Geometrical resolution in PuMA

Concerning the geometrical resolution of PuMA, we systematically use the PuMA feature artificial generation [8], without considering the potential import of images from tomography. We use the subscript $_{phy}$ for physical quantities and $_{vx}$ for PuMA-related quantities. For example, R_{phy} is the average radius of fiber (a few micrometers), but R_{vx} is the number of voxels of the radius. The same notation holds for D, which is the domain size. The accuracy level in PuMA is then quantified through the number of voxels used for the fiber radius (R_{vx}). Note that in tomography, the apparatus's resolution defines a voxel's physical size. An illustration is reported in Fig. 2. It is straightforward to convert a length expressed in voxels to the physical space by specifying the length of a voxel. For example, in PuMA, the "Voxel Length" l_{vx}^{phy} is defined by $\frac{R_{phy}}{R_{vx}}$.

Simulating a specific material requires indicating the type of fiber and its radius (corresponding to R_{phy}). Once the type of fiber is chosen, the two tunable PuMA parameters are then R_{vx} and D_{vx}, where R_{vx} provides the level of resolution of the fiber and D_{vx} the size of the whole domain. Once R_{vx} is cho-

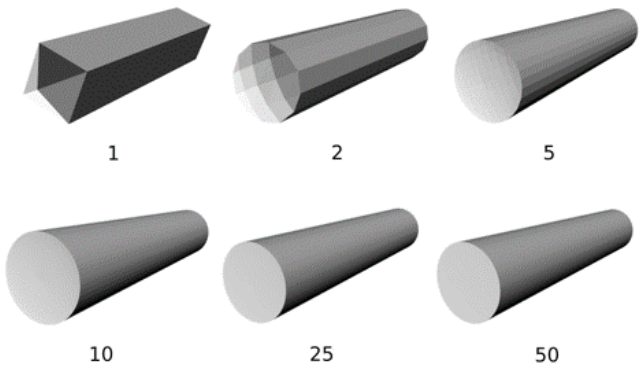


Figure 2: Fiber representation for different values of R_{vx}. The visualization is the output of a marching cube algorithm. Credits to [14].

sen, then we set l_{vx}^{phy} accordingly, following the previous formula for l_{vx}^{phy}. This procedure permits the investigation of the effect of the resolution for a specific physical size of the fibers. The numerical domain is a rectangular mesh of cubic voxels, where the reference length is D_{vx}.

Regarding the domain size, it is crucial to consider a domain large enough to be representative, i.e., permitting computing converged average properties of the material. Convergence is attained when the variation of the simulation output for an increasing domain size becomes lower than a certain threshold. Since the generation is stochastic, convergence should be assessed by looking at the conditional mean of the underlying distribution for each domain size. Note that the variance of the same distribution should converge to zero when the domain size tends to infinity. The additional problem here is that some physical parameters are known with a certain level of uncertainty (such as the fiber's thermal conductivity). Treating this uncertainty is also required to estimate the representative domain size reliably. At least, one should assess the sensitivity of the representative domain size to the input uncertainty.

Concerning the resolution, R_{vx} should be large enough to define the fiber correctly, and D_{vx} should not be too large for computational cost reasons (Note that a simulation with $D_{vx} = 100vx$ takes approximately three minutes when it

lasts several hours for $D_{vx} \sim 1000vx$, with potential convergence problems). In Section 5, we illustrate some results showing the implicit relation between the behavior of the quantity of interest concerning the choice of fiber and domain resolution.

3. Formulation of the stochastic problem

In this section, we address the issue of stochasticity more precisely. First, we perform an uncertainty characterization of PuMA and clarify the assumptions for building a predictive model for thermal conductivity. Last, we formalize the mathematical framework of this study, introducing notations that are useful for modeling the problem.

3.1. PuMA: Stochastic solver and parametric dependencies

Once the input parameters are specified, PuMA generates a geometry and estimates the intrinsically stochastic material properties due mainly to the fiber's orientation, radius, and length following some random distributions. Repetition of the PuMA simulations is then required to get averaged properties considering this intrinsic stochasticity. Note that, for a given porosity, an infinity of geometries is possible, and two different geometries do not lead to the same simulation outcome. Some other parameters, which are not well-known and are essentially the porosity p and the inherent properties of the fibers and the gas, are an additional source of uncertainty. Throughout this paper, PuMA has been treated as a stochastic solver with parametric dependencies.

We consider a material made of two phases: carbon fiber and gas. Both phases are homogenous (their properties do not depend on where they are measured). Conditions are uniform. All external parameters to the material (e.g., medium temperature) are constant in all the domains. Each phase has properties that do not vary with the location. Generally, thermal protection systems feature a geometrical plane with approximately distributed fibers. The material can then be assumed to be transverse isotropic. Later, what is related to the

plane, i.e., along the fiber, is called "In Plane"(IP), and what is associated with the normal direction, i.e., through the fiber, is called "Through Thickness"(TT). We refer to [18] for a detailed definition of these quantities.

Fibers are considered cylinders. This strong approximation (in real materials, they have a cylinder-like shape but are not straight and sometimes gather in clusters without a specific shape) is required to use the PuMA preform generation tool. Note that we allow the intersection of cylinders. The radius and length of fibers follow statistical distributions. Fiber orientation in space follows a transverse isotropic statistical distribution. That is to say, if we locate a fiber orientation on a 3D sphere of radius 1 in spherical coordinates (θ, Φ) (with $\mathbf{z}$ the normal vector), then Φ is identically distributed in $[0, 2\pi]$ and θ follows a law of expected value equal to $\frac{\pi}{2}$.

Concerning the stochasticity aspect of the model, the probability distribution of any QoI depends on the choice of geometric parameters laws and external parameters. Geometry parameters are fixed to represent as much real material as possible. Their values are taken from [19], where Calcarb® CBCF 18-2000 is studied. However, they play a role in stochasticity. Still, the model remains stochastic even with deterministic radius, length, and orientation.

3.2. Mathematical framework

For a given $\mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^d$ (where d is the number of input dimensions), $G(\mathbf{x})$ is a random variable with values in $\mathbb{R}^n$. We assume $G(\mathbf{x})$ is a second-order random variable with a Gaussian distribution. In this study, we propose a numerical strategy to build a surrogate model of $G(\mathbf{x})$ using the values estimated of G on a few training points. Once this surrogate model is built, we use it to compute the first and second moment of $G(\mathbf{x})$ with respect to $\mathbf{x}$. An additional question we address in this work is how to perform a sensitivity analysis of G with respect to the stochasticity and the parametric dependencies.

Sensitivity analysis can be used to assess the respective impacts of parametric and non-parametric uncertainty on the variability of G. Sensitivity analysis has been employed in several contexts. When collecting data, it can help iden-

tify main variables and uncertainties for optimizing additional collections [20]. Furthermore, sensitivity analysis can play an important role in model verification and validation when developing and refining models (*e.g.* [21, 22]). It can also provide insight into the robustness of model results when making decisions (*e.g.* [23, 24]).

Ref. [25] has proposed a novel framework dealing with stochasticity as an independent source of variability for Sobol decomposition. Ref. [26] suggests a general method for global sensitivity analysis in the case of stochastic simulation with uncertain parameters. Inspired from [25], we introduce first the classical Sobol-Hoeffding (SH) decomposition [27] from which we derive an expression for the sensitivity indices, which provide a measure for the impact of all different sources on the global variability. Assuming $\mathbf{x}$ and $\boldsymbol{\xi}$ are independent variables, the SH decomposition separates G into a sum of independent and orthogonal elements:

$$G(\mathbf{x}, \boldsymbol{\xi}) = \bar{G} + G_{par}(\mathbf{x}) + G_{ch}(\boldsymbol{\xi}) + G_{mix}(\mathbf{x}, \boldsymbol{\xi}) \tag{1}$$

where $\bar{G}$ is a constant, $G_{par}(\mathbf{x})$ is a random functional that depends only on the parameters, $G_{ch}(\boldsymbol{\xi})$ depends only on the Poisson processes, and $G_{mix}(\mathbf{x}, \boldsymbol{\xi})$ depends on both random inputs. Orthogonality gives uniqueness of the decomposition:

$$\begin{aligned} G_{par}(\mathbf{x}) &= \mathbb{E}\left[G|\mathbf{x}\right] - \mathbb{E}\left[G\right] \\ G_{ch}(\boldsymbol{\xi}) &= \mathbb{E}\left[G|\boldsymbol{\xi}\right] - \mathbb{E}\left[G\right] \\ G_{mix}(\mathbf{x}, \boldsymbol{\xi}) &= G(\mathbf{x}, \boldsymbol{\xi}) + \mathbb{E}\left[G\right] - \mathbb{E}\left[G|\mathbf{x}\right] - \mathbb{E}\left[G|\boldsymbol{\xi}\right] \end{aligned} \tag{2}$$

Orthogonality also gives the following decomposition:

$$\mathbb{V}\left[G\right] = V_{par} + V_{ch} + V_{mix} \tag{3}$$

Finally, the Sobol indices are:

$$S_{par} = \frac{V_{par}}{\mathbb{V}\left[G\right]}, \quad S_{ch} = \frac{V_{ch}}{\mathbb{V}\left[G\right]}, \quad S_{mix} = \frac{V_{mix}}{\mathbb{V}\left[G\right]} \tag{4}$$

Here, we have exposed the decomposition for only two parameters (one uncertain and one stochastic). However, the method can apply to an undefined number of parameters if they are mutually independent [26].

4. Surrogate-based numerical methods

Evaluating statistics of material properties estimated with PuMA requires computational costs to be too high if Monte Carlo-based sampling is used. Many surrogate modeling strategies for Uncertainty Propagation in stochastic solvers have been proposed in literature (see, for example, [28] and [29]). Polynomial chaos expansions (PCE) [30] and low-rank tensor approximations (LRA) have been used in other methods such as Generalized Additive Models (GAM) [31, 32], and to perform sensitivity analysis [28]. Other methods [33, 34] focus on predicting statistics such as mean, variance, and quantiles. The surrogating techniques may differ. For example, in [35], GAM models were used to predict the mean and the variance, while in [36], the mean function was predicted by assuming that the output is a mixture of normal distributions. Kriging methods have proved to be able to perform forward uncertainty propagation [37, 34], notably in UQ and sensitivity analysis context. Here, we work in the framework of the stochastic kriging methodology for estimating an unknown response surface with uncertainty [34]. In particular, we propose a novel algorithm and explain how to use the surrogate model for sensitivity analysis with a Monte Carlo algorithm.

4.1. Stochastic Kriging

We aim at modeling an unknown function $\mathrm{y}(\mathbf{x})$ where $\mathbf{x} = (x_1, ..., x_d)^T$ is a set of design parameters. For example, $\mathbf{x}$ might account for material porosity, gas thermal conductivity, fibers' orientation etc. and $\mathrm{y}(\mathbf{x})$ the associated global thermal conductivity. Let $\mathbf{X} = (\mathbf{x}_i, n_i)_{i=1,...,N}$ be the *design of experiment* (*DoE*) of our model, where n_i is the number of simulation replications taken at

design setting $\mathbf{x}_i$. Let's then note:

$$\bar{\mathcal{Y}}(\mathbf{x}_i) = \frac{1}{n_i}\sum_{j=1}^{n_i}\mathcal{Y}_j(\mathbf{x}_i), \tag{5}$$

where $\mathcal{Y}_j(\mathbf{x}_i)$ is the output of the j^{th} replication simulation at design point $\mathbf{x}_i$ so that $\mathbf{S}_i = \{\mathcal{Y}_j(\mathbf{x}_i), i = 1, ..., n_i\}$ is the set of simulation outputs at design point $\mathbf{x}_i$, and let $\bar{\mathcal{Y}} = (\bar{\mathcal{Y}}(\mathbf{x}_1)...\bar{\mathcal{Y}}(\mathbf{x}_N))^T$. Ref. [38] shows that for linear predictor, the mean squared error optimal predictor and the associated mean squared error (MSE) are given by

$$\hat{Y}(\mathbf{x}_0) = \boldsymbol{\beta}^T\mathbf{f}(\mathbf{x}_0) + \Sigma_M(\mathbf{x}_0, .)^T(\Sigma_M + \Sigma_\varepsilon)^{-1}(\bar{\mathcal{Y}} - F\boldsymbol{\beta}), \tag{6}$$

$$MSE(\mathbf{x}_0) = \Sigma_M(\mathbf{x}_0, \mathbf{x}_0) - \Sigma_M(\mathbf{x}_0, .)^T(\Sigma_M + \Sigma_\varepsilon)^{-1}\Sigma_M(\mathbf{x}_0, .), \tag{7}$$

with $\{f_j; j = 1, ..., P\}$ a vector of known functions of $\mathbf{x}$, $\boldsymbol{\beta}$ a vector of unknown parameters, $\mathcal{M}(\mathbf{x})$ a realization of a mean zero random field, Σ_M the $N \times N$ correlation covariance matrix between design points with (p, q) element equal to $Cov[\mathcal{M}(\mathbf{x}_p), \mathcal{M}(\mathbf{x}_q)]$, $\Sigma_M(\mathbf{x}_0, .) = (Cov[\mathcal{M}(\mathbf{x}_0), \mathcal{M}(\mathbf{x}_1)], ..., Cov[\mathcal{M}(\mathbf{x}_0, .), \mathcal{M}(\mathbf{x}_N)])^T$, which is the $N \times 1$ vector of correlation covariance between the point of interest $\mathbf{x}_0$ and the design points. The term $\varepsilon_j(\mathbf{x})$, called *intrinsic* noise because it is inherent to the sampling simulation variability, is the realization of a Gaussian random variable, with Σ_ε the $N \times N$ covariance matrix implied by the intrinsic noise over design points with (p, q) element equal to $Cov[\bar{\varepsilon}(\mathbf{x}_p), (\bar{\varepsilon}(\mathbf{x}_q)]$. Finally, F is the $N \times P$ matrix with elements $F_{pq} = f_q(\mathbf{x}_p)$.

Stochastic Kriging (SK) methodology proposed by [34] is a surrogate tool for approximating a mean response surface associated with a stochastic simulation. It accounts for sampling uncertainty (Σ_ε, also called *intrinsic uncertainty*) inherent in a stochastic simulation in addition to the response-surface uncertainty (Σ_M, also called *extrinsic uncertainty*). In [34], a kriging surrogate is built for the variance ($\hat{V}_{SK}$ is denoted hereafter as the associated mean predictor of the variance for SK). Since the variance is not observable, even at design points, it is estimated as follows:

$$\mathcal{S}^2(\mathbf{x}_i) = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}(\mathcal{Y}_j(\mathbf{x}_i) - \bar{\mathcal{Y}}(\mathbf{x}_i))^2. \tag{8}$$

In [34], $\hat{V}_{SK}$ is an interpolator, meaning that $\hat{V}_{SK}(\mathbf{x}_i) = \mathcal{S}^2(\mathbf{x}_i)$ at design points. [34] shows that using Eq. 8 does not introduce prediction bias. The confidence of the prediction can be questionable, but Ref. [34] argues that for n_i not too small (typically superior to 10), the MSE only slightly increases.

4.2. The proposed approach based on joint surrogates

In this section, we address the issue of building a surrogate for both the mean prediction and the variance. One approach is building a polynomial approximation of the mean and variance separately [39], based on repeated calculations with the same set of controllable input variables. This dual modeling approach has been successfully applied to robust conception problems. However, our purpose here is to fit both mean and variance accurately. Within this context, it has been shown that the dual model performs worse than the joint model [40]. For example, [41] applies a specific joint metamodel developed in [42] in an accidental scenario in a pressurized water reactor. This section describes the method used in this paper, which is an extension of SK.

4.2.1. Surrogates estimation

Inspired from [43], the variance surrogate is built by adding an extrinsic noise. It is estimated with *Bootstrap Sample Variances* because the comparison with other methods proves it is adapted to variable replication allocation schemes (which is discussed further), with performances comparable to the best estimator. We draw B bootstrap samples each of size n_i from $\mathbf{S}_i$, resulting in B bootstrap samples $\{\mathbf{S}_i^b\}_{b=1,\dots,B}$ where $\mathbf{S}_i^b = \{\mathcal{Y}_j^b(\mathbf{x}_i)\}_{j=1,\dots,n_i}$. Then the bth sample variance $s_{i,b}^2$ at design point $\mathbf{x}_i$ is computed with $\mathbf{S}_i^b$. Finally, the variance estimate of the sampling variance is:

$$\hat{V}_{s^2}(\mathbf{x}_i) = \frac{1}{B-1}\sum_{b=1}^{B}(s_{i,b}^2 - \overline{s_{i,.}^2})^2 \qquad \text{with} \quad \overline{s_{i,.}^2} = \frac{1}{B}\sum_{b=1}^{B} s_{i,b}^2 \tag{9}$$

From this equation, we can build a heteroscedastic predictor $\widehat{V}^\star$ for the variance [43], with $\mathcal{S}^2$ (resp. $\hat{V}_{s^2}$) an estimator of the variance (resp. noise) at the design points.

The predictor for the mean is given by:

$$\widehat{Y}^{\star}(\mathbf{x}_0) = \boldsymbol{\beta}^T \mathbf{f}(\mathbf{x}_0) + \Sigma_M(\mathbf{x}_0, .)^T (\Sigma_M + \widehat{\Sigma}_{\varepsilon}^{\star})^{-1} (\bar{\mathcal{Y}} - F\boldsymbol{\beta}) \tag{10}$$

where $\widehat{\Sigma}_{\varepsilon}^{\star} = \text{Diag}\{\frac{\widehat{V}^{\star}(\mathbf{x}_1)}{n_1}, ..., \frac{\widehat{V}^{\star}(\mathbf{x}_N)}{n_N}\}$. Note that the variance and mean surrogates hyperparameters optimization processes are separated but not uncorrelated.

4.2.2. Allocation methodology

The number of replications massively influences the overall computation cost. In this paper, we assume that every simulation costs the same budget, regardless of the set of input parameters, and that the budget is equal to the maximum number of simulations, indicated here as C. Since the DoE is defined as $\mathbf{X} = (\mathbf{x}_i, n_i)_{i=1,...,N}$, the following relation holds $\sum_{i=1}^{N} n_i \leq C$. Note that increasing the number of replications at a design point improves the accuracy of the two surrogates. It can be shown that it decreases the noise level when estimating the mean and the variance.

The most basic strategy is distributing the budget equally among all design points. This strategy is not optimal since it does not take advantage of prior knowledge of the objective function. Nonetheless, it does not need extra information and is decided before the simulation runs. In this paper, we propose an *Unequal IMSE minimizing allocation scheme*, permitting the distribution of the budget in an optimal way to each design point. The IMSE (Integral Min Square Root Error) strategy was first proposed in [34], and we discuss how this has been implemented in our framework.

Let $\mathcal{X}$ be the d-dimensional space of interest from which the design points $\mathbf{x}_i$

have been chosen. Let $\mathbf{n} = (n_1, ..., n_N)$. The problem can be written as follows:

$$\begin{aligned}
\text{minimize} \quad & IMSE(\mathbf{n}) = \int_{\mathcal{X}_0} MSE(\mathbf{x}_0; \mathbf{n}) d\mathbf{x}_0 \\
\text{s.t} \quad & \mathbf{n} \in \mathbb{N}^N, \mathbf{n}^T \mathbf{1}_N \leq C \\
\text{where} \quad & MSE(\mathbf{x}_0; \mathbf{n}) = \Sigma_M(\mathbf{x}_0, \mathbf{x}_0) - \Sigma_M(\mathbf{x}_0, .)^T (\Sigma_M + \Sigma_\varepsilon(\mathbf{n}))^{-1} \Sigma_M(\mathbf{x}_0, .) \\
\text{with} \quad & \Sigma_M \text{ and } \Sigma_\varepsilon \text{ the optimized covariance} \\
& \text{associated to the joint surrogate}
\end{aligned}$$

with $\Sigma(\mathbf{n}) = \Sigma_M + \Sigma_\varepsilon(\mathbf{n})$. Ref. [34] shows the solution with the integrity constraint relaxed satisfies $n_i^\star \propto \sqrt{V(\mathbf{x}_i) C_i(\mathbf{n}^\star)}$ where $C_i(\mathbf{n}) = \left[\Sigma(\mathbf{n})^{-1} \mathbf{W} \Sigma(\mathbf{n})^{-1}\right]_{ii}$ and $\mathbf{W}$ is a $N \times N$ matrix $W_{ij} = \int_{\mathcal{X}_0} k(\mathbf{x}_i, \mathbf{x}_0) k(\mathbf{x}_j, \mathbf{x}_0) d\mathbf{x}_0$

Under the assumption that N is large enough and training designs are equally spatially distributed, it can be shown that:

$$n_i = C \frac{\sqrt{V_i}}{\Sigma_{j=1}^N \sqrt{V_j}} \tag{11}$$

It can also be shown that it is a valid allocation choice for the surrogate on the variance. Exploiting the computational budget optimally requires building the two surrogates adaptively. The whole algorithm describing the approach proposed in this paper is reported in the following. A Latin Hypercube Sampling (LHS) is systematically used for the initial Experiment design. More efficient and adaptive strategies could be implemented.

Algorithm 1: Joint metamodel algorithm

Step 0: Choose number of training points N, total budget C, number of allocation steps r and bootstrap number B.

Step 1: Generate training set $(\mathbf{x}_i)_{i=1,...,N}$.

Step 2: Allocate and evaluate initial number of repetitions n_i for each training point $\mathbf{x}_i$ ($n_i \geq 5$ suggested) : $(\mathcal{Y}_j(\mathbf{x}_i)_{j=1,...,n_i})$.

Step 3: Evaluate variance at each training point

$\mathcal{S}^2(\mathbf{x}_i) = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (\mathcal{Y}_j(\mathbf{x}_i) - \bar{\mathcal{Y}}(\mathbf{x}_i))^2$. Eq.8

Step 4: **for** *k between 1 and r* **do**

Define $C_k = \frac{kC}{r}$ as temporary total budget.

Update budget allocation : $n_i = C_k \frac{\sqrt{\mathcal{S}^2(\mathbf{x}_i)}}{\sum_{j=1}^{N} \sqrt{\mathcal{S}^2(\mathbf{x}_j)}}$. Eq.11

Perform new evaluations $\mathcal{Y}_j(\mathbf{x}_i)$.

Update $\mathcal{S}^2(\mathbf{x}_i)$ for all training points.

end

Step 5: **for** *i between 1 and N* **do**

Draw B bootstrap samples each of size n_i from $(\mathcal{Y}_j(\mathbf{x}_i)_{j=1,...,n_i})$ and compute their variance $s^2_{i,b}$.

Compute variance noise $\hat{V}_{s^2}(\mathbf{x}_i) = \frac{1}{B-1} \sum_{b=1}^{B} (s^2_{i,b} - \overline{s^2_{i,.}})^2$. Eq.9

end

Step 5: Build joint variance heteroscedastic predictor $\hat{V}^\star$ based on $(\mathcal{S}^2(\mathbf{x}_i))_{i=1,...,N}$ and $\widehat{\Sigma}^{V\star}_{\varepsilon} = (\hat{V}_{s^2}(\mathbf{x}_i))_{i=1,...,N})$.

Step 6: Build joint mean heteroscedastic predictor $\hat{Y}^\star$ based on $(\bar{\mathcal{Y}}(\mathbf{x}_i))_{i=1,...,N}$ and $\widehat{\Sigma}^{Y\star}_{\varepsilon} = \mathrm{Diag}\{\frac{\hat{V}^\star(\mathbf{x}_1)}{n_1}, ..., \frac{\hat{V}^\star(\mathbf{x}_N)}{n_N}\}$.

4.3. Numerical test-case

The overall methodology is assessed on the 3D-Hartmann function (see [44] for more details) defined as follows:

$$m(\mathbf{x}) = -\sum_{i=1}^{4} \alpha_i \exp\left(-\sum_{j=1}^{3} A_{ij}(x_j - P_{ij}^2)\right) \tag{12}$$

where values for α, A and P can be found in [44].

The j^{th} simulation output on design point $\mathbf{x}_i$ is generated as a stationary Gaussian random field $\mathcal{Y}_j(\mathbf{x}_i) = m(\mathbf{x}_i) + \varepsilon_j(\mathbf{x}_i)$, with $\varepsilon_j(\mathbf{x}_i)$ i.i.d $\mathcal{N}(0, V(\mathbf{x}_i))$ and $V(\mathbf{x}_i) = \delta|m(\mathbf{x}_i)|$ where δ is a magnitude parameter. We define an empirical root mean squared error (ERMSE) over a set of R points noted $(\mathbf{p}_i)_{i=1,\dots,R}$ to compare models. For a given experiment subscripted e (*i.e.* a surrogate built on a training set), the error of the estimation surrogate $\hat{G}_e$ with respect to the true function g is:

$$ERMSE_e = \sqrt{\frac{1}{R}\sum_{i=1}^{R}\left(\hat{G}_e(\mathbf{p}_i) - g(\mathbf{p}_i)\right)^2} \tag{13}$$

with $(\mathbf{p}_i)_{i=1,\dots,R}$ fixed for all experiments and generated by a LHS in $[0,1]^3$. Then, we can make statistics over several experiments and compare ERMSEs.

We report in Table 1 the results obtained with the same budget $C = 2e3$ and $\delta = 1$. Note that values are written under the format: mean over 100 experiments of ERMSE on mean function - mean over 100 experiments of ERMSE on variance function (if applicable). We compare an equal allocation scheme with a kriging model assuming there is no noise, the homoscedastic model (HO, assuming a uniform unknown noise on data, *i.e.* the basics kriging model for stochastic simulation), and SK with an unequal allocation scheme for the model proposed in this paper, based on joint surrogates.

For each surrogate, prediction quality first decreases and then increases when the number of training points increases. Initially, there is a lack of information with a few training points. After some iterations, the lack of replications at training points leads to under-fitting because of significant noise. Using the unequal allocation scheme improves the quality of mean predictions. Because the local variance is estimated as the model progresses, significant errors in variance estimations with a few samples can lead to poor allocation schemes. Our joint surrogate performs better than all the other models. Learning the variance as new samples come in, the joint surrogate shows similar results for mean prediction than SK (which does not estimate variance). When it does not perform better for mean prediction, a slight quality loss on mean leads to a gain

for variance prediction (for $N = 40$, a mean loss of 3% leads to an improved variance by 35%).

N	IN	HO	SK	Joint metamodel
250	m : 1.184	m : 0.256	m : 0.193, V : 0.497	m : 0.197, V : 0.468
200	m : 0.731	m : 0.262	m : 0.186, V : 0.561	m : 0.175, V : 0.426
100	m : 0.395	m : 0.278	m : 0.193, V : 0.518	m : 0.167, V : 0.324
40	m : 0.337	m : 0.330	m : 0.294, V : 0.615	m : 0.284, V : 0.374

Table 1: ERMSE obtained with an equal allocation scheme for IN (Interpolator), HO (Homoscedastic) and SK, and with an unequal allocation scheme for Joint surrogate

4.4. Sensitivity analysis

The approach proposed in the previous section permits the construction of a surrogate capable of predicting the inherent variance of the output on the domain. We can recall the concepts illustrated in Section 3 to present the SH decomposition. Let $(\hat{G}^\star, \hat{V}^\star)$ be the mean and variance surrogates associated to G. We assume each component of $\mathbf{x}$ is defined as a nominal value and an uncertainty law (typically $x_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$). $\boldsymbol{\xi}$ also has its own known law.

Let $\mathbb{X}$ and $\boldsymbol{\mathcal{E}}$ be two sample sets of S independent realizations of the parameters and stochastic parameters. As discussed before, $\mathbf{X}_i$ is a sample of mutually independent parameters, whereas $\boldsymbol{\xi}_i$ is an initialized seed. From these samples, Monte Carlo estimators of expectation and variance are:

$$\widehat{\mathbb{E}_G} = \frac{1}{S}\sum_{i=1}^{S} \hat{G}(\mathbf{X}_i, \boldsymbol{\xi}_i) \tag{14}$$

$$\widehat{\mathbb{V}_G} = \frac{1}{S}\sum_{i=1}^{S} \hat{G}(\mathbf{X}_i, \boldsymbol{\xi}_i)^2 - \widehat{\mathbb{E}_G}^2 \tag{15}$$

Now, let $\widetilde{\mathbb{X}}$ and $\widetilde{\boldsymbol{\mathcal{E}}}$ be independent replicas of $\mathbb{X}$ and $\boldsymbol{\mathcal{E}}$. Monte Carlo method

gives:

$$\widehat{V_{par}} = \frac{1}{S-1}\sum_{i=1}^{S} \hat{G}(\mathbf{X}_i, \boldsymbol{\xi}_i)\hat{G}(\widetilde{\mathbf{X}}_i, \boldsymbol{\xi}_i) - \widehat{\mathbb{E}_G}^2 \tag{16}$$

$$\widehat{V_{ch}} = \frac{1}{S-1}\sum_{i=1}^{S} \hat{G}(\mathbf{X}_i, \boldsymbol{\xi}_i)\hat{G}(\mathbf{X}_i, \widetilde{\boldsymbol{\xi}}_i) - \widehat{\mathbb{E}_G}^2 \tag{17}$$

where $\widehat{V_{par}}$ and $\widehat{V_{ch}}$ correspond to the parts of the variance justified by the parameters and the Poisson processes, respectively.

Associated sensitivity indices (see Section 3.2) are then finally estimated as:

$$\widehat{S_{par}} = \frac{\widehat{V_{par}}}{\widehat{\mathbb{V}_G}}, \quad \widehat{S_{ch}} = \frac{\widehat{V_{ch}}}{\widehat{\mathbb{V}_G}}, \quad \widehat{S_{mix}} = 1 - \widehat{S_{par}} - \widehat{S_{ch}} \tag{18}$$

5. Results

This section illustrates the main results obtained by applying the UQ methodology presented in the previous section to compute statistical moments of ablative material properties. In particular, all the analyses rely on synthetic data generated using PuMA for some specific conditions.

Our study is applied to TC2 fibers pre-treated at $2500K$ as described in [45]. These conditions are close to the ones of Calcarb® CBCF 18-2000 pre-treated at $2000K$, but TC2 is the same type (rayon), and $2000K$ is relatively close enough to consider their respective performances to be equivalent. We consider the following assumptions and conditions:

- Fiber (resp. gas) conductivity, λ_f (resp. λ_g) is only a function of temperature T. Note that the pressure is not a parameter, except in the case of a rarefied medium;
- The considered temperatures are $T = 855, 1380, 1875K$;
- The considered gases are helium, argon, and nitrogen;
- Once the temperature is chosen, the gas conductivity is supposed to be known without uncertainty. Its value is computed at ambient pressure with Mutation^{++} [46].

For a given temperature and gas, we consider fiber conductivity as a function of p and λ_f, which are affected by the following uncertainties (For both parameters, we consider a Gaussian distribution). Nominal porosity p is the mean porosity of Calcarb® CBCF 18-2000, *i.e.* $p = 0.89$ with an uncertainty of $2\sigma_p = 0.01$. Ref. [45] contains the nominal values and uncertainty of fiber conductivity. We use the simple formula, *i.e.* $\lambda_f = \rho c_p \kappa$, where ρ is the local density, c_P is the local heat capacity, and κ is the local thermal diffusivity. The additional source of uncertainty comes from the intrinsic stochasticity of PuMA outcomes.

The application of the UQ methodology is straightforward. We perform an independent UQ study for each gas and temperature by building a Design of Experiment with the two uncertain parameters, *i.e.* p and λ_f.

In the following, we discuss first the notion of convergence in the case of a stochastic solver such as PuMA. This aspect is particularly relevant to compute the needed resolution for the fibers and the domain to identify the reference elementary volume robustly. The remaining part of the section illustrates several analyses of the influence of several physical parameters of interest, showing the interest of the UQ methodology in this problem and permitting some relevant physical interpretations of the observed behaviors.

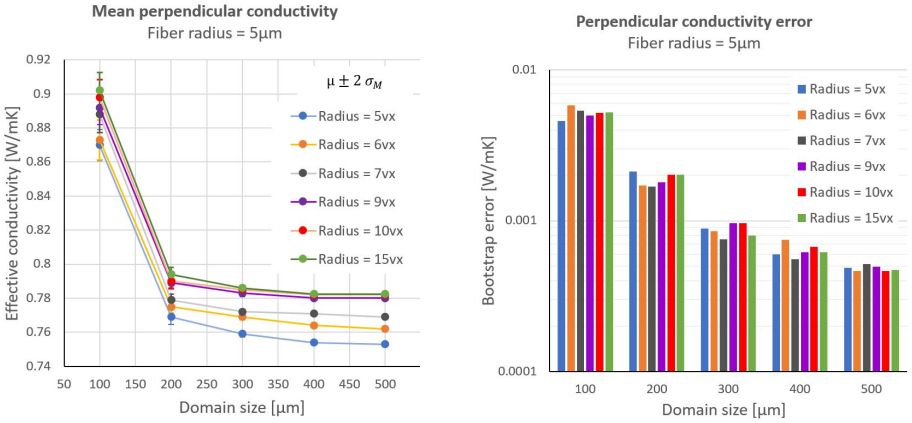
5.1. Convergence of PuMA solver

We focus here on the influence of mesh resolution and other numerical parameters to attain converged prediction of the quantities of interest. In particular, we discuss the choice of domain size and the number of voxels to represent the fibers and the domain. Statistical quantities are computed over 100 repetitions, each at nominal parameters $p = 0.85$, $\lambda_f = 12.0 W/m.K$, $\lambda_g = 0.04 W/m.K$. There is no uncertainty on R_{vx} but $l_{vx} = 80 R_{vx} \pm 50\%$.

The evolution of the effective thermal conductivity as a function of the domain size for a fixed radius resolution is reported in Fig. 3(a). Note that several levels of resolutions are taken into account. Here, the mean and the error interval are computed considering 100 repetitions of the stochastic solver. At $300\mu m$, the thermal conductivity value is nearly converged (the final value is

within 95%). Secondly, when R_{vx} increases, the converged value increases until a value around $0.782W/m.K$ is reached.

In Fig. 3(b), the estimated error on the mean (computed by bootstrap) is reported again as a function of the domain size. As expected, it features two trends. First, when increasing the domain size, the error decreases (Fig. 3(b)): samples are enough to all be representative, and therefore, the error on the mean becomes small. On the contrary, for small domains, the error on the mean is significant, and even the mean value seems very far from the convergence value, which is supposed to be the most accurate (See Fig. 3(a)). Only the results obtained in in-plane (IP) are plotted since the ones in TT have the same trend.



(a) (b)

Figure 3: Convergence of QoI as a function of R_{vx} and D_{phy}. (a) Mean of repetitions and (b) error on the mean.

This analysis illustrates reliable estimations for choosing an optimal resolution of the fiber and the whole domain. Specifically, we adopt the following choices, i.e., $R_{vx} = 8vx$ and $D_{phy} = 300\mu m$. This last value can be interpreted as a reliable estimation of the minimal reference elementary volume.

Furthermore, we choose 30° as the angle variation parameter for sampling fibers' orientation. Finally, the physical values for radius and length are $R_{phy} = 5 \pm 0.625\mu m$ and $L_{phy} = 800 \pm 500\mu m$. We account for both the great stochas-

ticity of fiber length and the possibility of fiber clusters.

5.2. UQ-based analysis

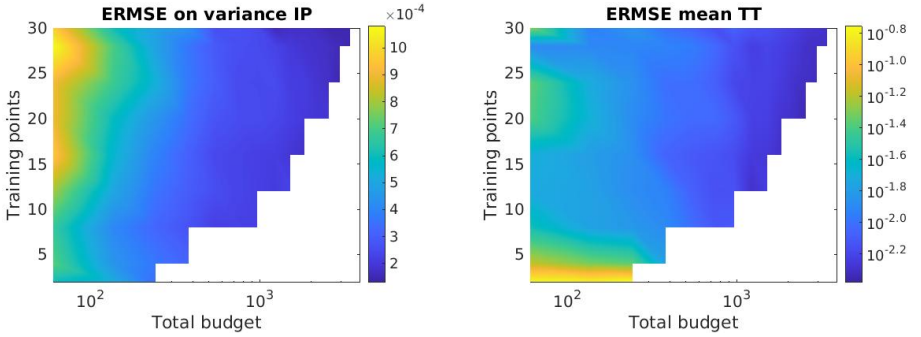
We focus here on building a surrogate of thermal conductivity and evaluating the impact of uncertainty and stochasticity on the variability of the output. The results are concentrated on two axes. First, we analyze helium as a medium gas for several temperatures and see the impact of the temperature on the Sobol indices (fibers and gas conductivities should vary). Then, we assess the influence of the choice of gas (nitrogen and argon) by fixing the temperature, thus changing the gas's conductivity.

The *DoE* is generated for each numerical experiment based on nominal values and uncertainties of p and λ_f. Several training points equal to $N_{max} = 30$ are generated with LHS. We have applied the unequal allocation scheme with four allocation steps. It has resulted in a total budget of $C_{max} = 3800$, distributed among the 30 training points. Note that we use UQLab project [47] on MATLAB for coding and solving. We also checked that distributions of λ (TT and IP) converge to Gaussian distributions.

5.2.1. Convergence of metamodel

Assessing the convergence of the surrogate model requires the computation of a metric with respect to the validation set. In this case, we compute an Empirical Root Mean Squared Error, denoted hereafter as ERMSE (defined precisely in Eq. 13, representing the error of the estimation surrogate with respect to the true function. In the following, we compute the ERMSE on mean and variance to evaluate the errors of the method when increasing the number of training points taken into account and increasing the allocated budget. However, ERMSE demands reference values to compare predictions. Empirical mean and variance in a set of 30 training points used as reference are compared to the prediction given by the surrogate models.

Figure 4 shows the result of convergence. As expected, increasing C and N decreases the error. However, increasing N with fixed C does not necessarily lead to better performance. A priori estimations of QoI are less accurate because the same budget is allocated to more points. This behavior is especially true for the variance at low budgets (Fig. 4(a)). This effect disappears at a higher budget because it is compensated by the hyperparameters optimization process. Furthermore, convergence is relatively fast when increasing N when C is larger than 200: the function varies quite slowly, and stochasticity is more responsible for inaccuracy than a lack of training points. Overall, hyperparameters do not show irregular behaviors, and MSEs of both mean and variance surrogates are consistent.



(a) (b)

Figure 4: Convergence of ERMSEs when increasing C and N. (a) In Plane variance ERMSE in $[W/m.K]^2$; b) Mean Through Thickness in $[W/m.K]$. The reference is the empirical QoI at all training points.

5.2.2. Mean prediction

First, we evaluate the behavior of the conductivity with respect to the temperature in the case of helium. IP and TT conductivities are plotted in Fig. 5.

For helium, increasing temperature leads to a rise in conductivity. This trend is the same for most gases and fiber types. Ref. [45] shows that fiber conductivity is an increasing function of temperature for this temperature in-

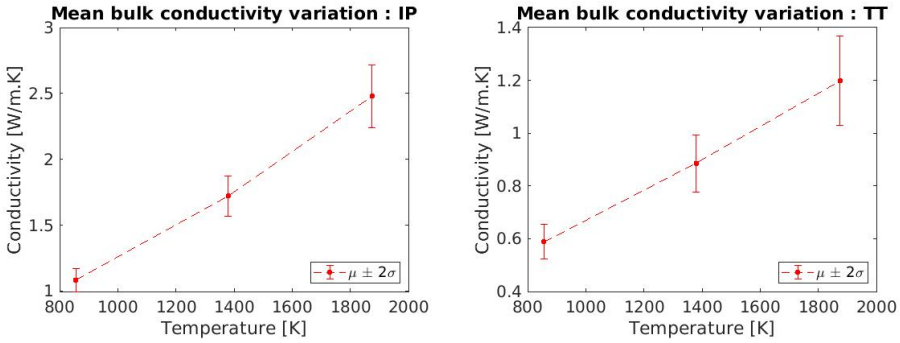


Figure 5: Influence of temperature on total sample conductivity in case of helium. Means and total variability error bars are plotted.

terval. This behavior is confirmed also by Mutation^{++} outcome simulation for gas conductivity.

Finally, IP conductivity is more significant than TT conductivity due to the geometry and fibers' orientation preferentially in the plane. Figure 6 reports the influence of the choice of the gas on the total conductivity for a fixed intermediate temperature of $T = 1380K$. For this temperature, $\lambda_{helium} > \lambda_{nitrogen} > \lambda_{argon}$ holds. Since interactions between gas and fiber are neglected, this behavior is consistent with the results.

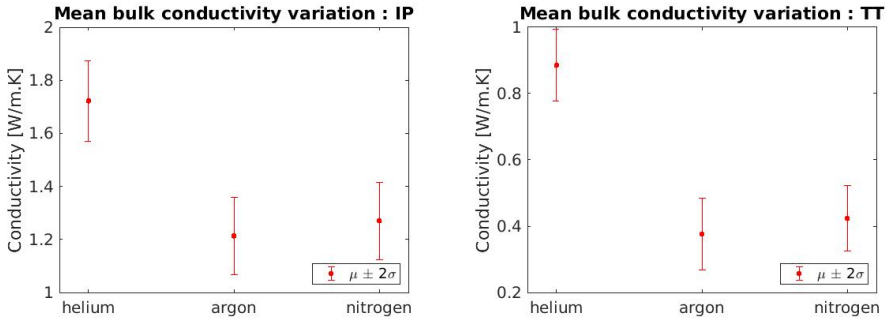


Figure 6: Influence of gas on total sample conductivity for $T = 1380K$ (Means and total variability error bars are plotted).

5.2.3. Sensitivity analysis

This section focuses on computing the output's variability and its decomposition after variable components. We recall uncertainty on porosity and fiber conductivity and model intrinsic stochasticity as sources of variability. For mean analysis, we first study variability for several temperatures and then for several gases.

Performing variance decomposition as introduced in Section 4 requires that model stochasticity is independent of parameters. When using PuMA, stochasticity is related to a seed at the beginning of the generation. Unless there is apriori knowledge about the link between parameters and seed choice when generating random samples (see [29]), there is no dependency between stochasticity and parametric uncertainty, permitting the use of SH decomposition.

First, we verify the convergence of the QoI in addition to the convergence of ERMSEs. Figure 7 shows the sensitivity analysis results from the ERMSE convergence analysis.

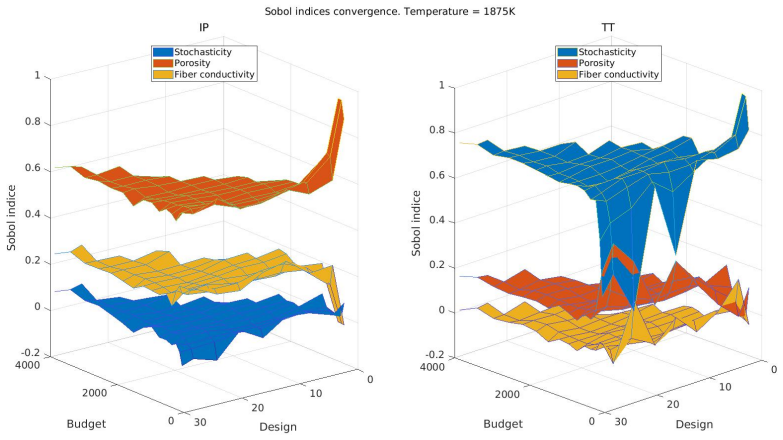


Figure 7: Convergence of Sobol indices when increasing C and N.

Convergence is fast for ERMSEs as well (in both budget and number of training points terms). This fact means our surrogate method is accurate and does not need as many simulations to reach convergence. Figure 8 shows the

results from the sensitivity analysis for helium for the three temperatures.

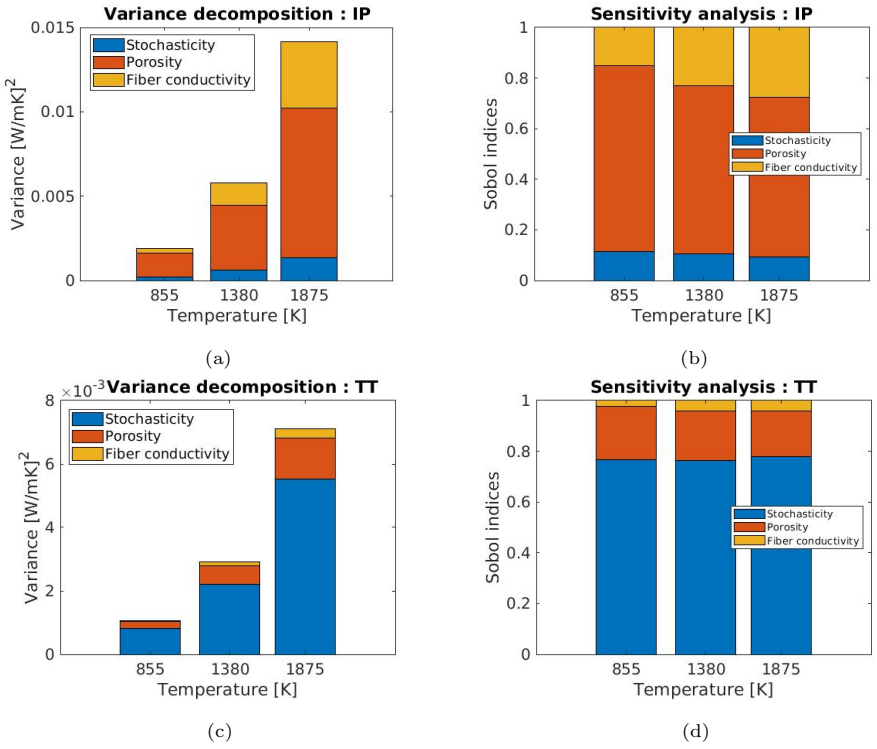


Figure 8: (a) (resp. (c)) shows variance decomposition of the total variability In Plane (resp. Through Thickness) for three temperatures. (b) and (d) show associated Sobol decomposition. Gas is helium.

Note that stacked variance for a given temperature directly equals the square root of the associated total error bar divided by four of figure 5. Regarding Sobol indices (Fig. 8(b) and 8(d)), there is a clear hierarchy between components, different in IP and TT cases. Concerning variance magnitude (Fig. 8(a) and 8(c)), total variability increases with temperature, as well as all its components. At least two reasons can be given. First, fiber conductivity uncertainty rises with temperature, so it seems normal that its associated component increases. Notably, the influence of fiber conductivity increases with temperature. Secondly, mean fiber conductivity also increases with temperature: a change in geometry, therefore, has a more significant impact on the output. The remaining variance

is due to more complex mutual relations between parameters and stochasticity, as well as the relative increase of λ_f with respect to λ_g.

Note that the stochasticity component is much higher in the through-plane case. This trend is explained as follows. The value of λ_{TT} is smaller than λ_{IP}, and λ_{TT} is the mean of two quantities that are supposed to follow approximately the same type of law as λ_{IP}: therefore variance is lower in the plane. Figure 9 shows the results from the sensitivity analysis for the three gases for the intermediate temperature.

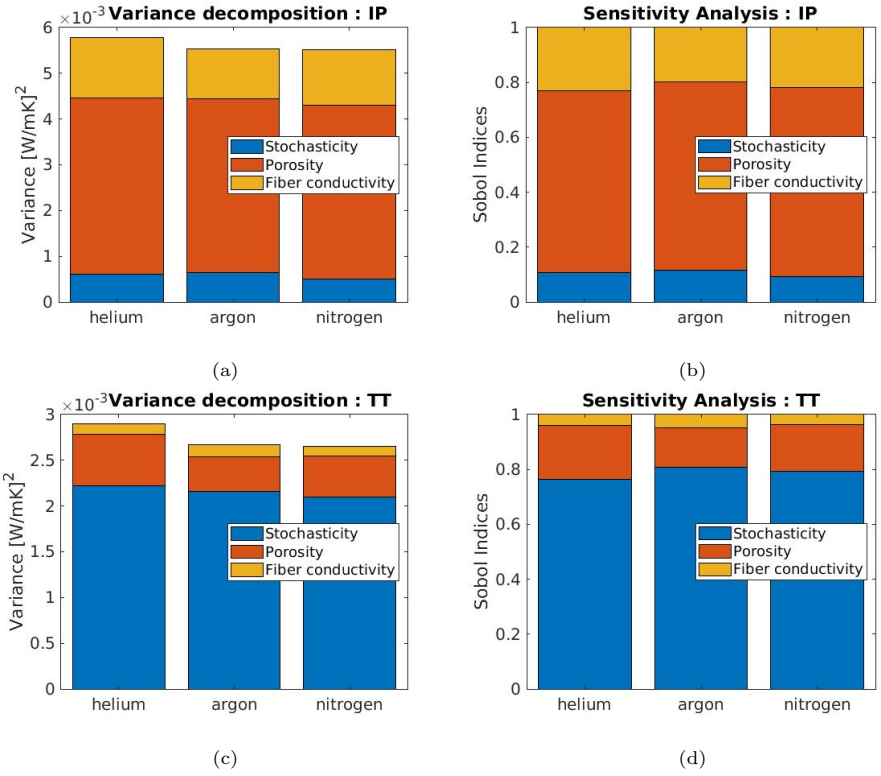


Figure 9: (a) (resp. (c)) shows variance decomposition of the total variability In Plane (resp. Through Thickness) for three gases. (b) and (d) show associated Sobol decomposition. Temperature is $T = 1380K$.

The main difference between each gas is, thus, the nominal gas conductivity. As observed in the previous section, it does change the mean conductivity of the

sample. However, the impact on variance is much slighter. Figures 9 (a) and (c) seem to show that it decreases total variability, but figures 9 (b) and (d) do not show significant differences between gases regarding repartition of indices. Furthermore, estimating the error made by the Monte Carlo estimation method of Sobol indices and drawing a particular trend is problematic.

6. Conclusions and perspectives

This paper illustrates a UQ methodology applied to characterize the microscopic properties of composite materials. First, we introduced a novel methodology to propagate uncertainties through the PuMA software, which provides a stochastic response under a specific choice of deterministic inputs. The proposed technique can estimate the mean and variance of any quantity of interest at a moderate computational cost, considering the solver's stochasticity and the considered uncertainties. Additionally, we showed how a sensitivity analysis can be performed by splitting the contribution to the variance of the quantity of interest as a summation of the stochastic part and the one associated with the uncertain parameters.

We illustrated an uncertainty propagation study considering an artificial material generated through PuMA, considering uncertainty on porosity and the thermal conductivity of the fiber. Given the current knowledge of fiber and gas properties, we showed that the thermal conductivity variability is mainly explained by the uncertainty on the porosity for in-plane conductivity and inherent stochasticity for through-thickness conductivity. We observed that the choice of the gas may influence mean conductivity as expected but much less the total variance. In practice, this behavior indicates that when external conditions other than temperature change, the choice of the gas does not impact the overall variability. As expected, when temperature increases, thermal conductivity increases. However, this is also true for its variance. Furthermore, the higher the temperature, the higher the Sobol index of fiber conductivity.

The overall approach proposed in this paper permits a systematic assessment

of the choice of the voxel resolution for both the fibers and the domain. Precisely, the convergence of the quantities of interest can be surveyed, thus identifying the minimal reference elementary volume. We consider this result very interesting for future studies preparing experimental tomography campaigns. Another potential research could target the direct use of tomography images under an uncertainty quantification perspective.

References

[1] F. Torres-Herrador, et al, Decomposition of carbon/phenolic composites for aerospace heatshields: Detailed speciation of phenolic resin pyrolysis products, Aerosp. Sci. Technol. 119 (2021) 107079.

[2] A. R. Valdez, et al, Uncertainty quantification and sensitivity analysis for relative permeability models of two-phase flow in porous media, J. Pet. Eng. 192 (2020) 107297.

[3] V. Ciriello, F. P. J. de Barros, Characterizing the influence of multiple uncertainties on predictions of contaminant discharge in groundwater within a lagrangian stochastic formulation, Water Resour. Res. 56 (10).

[4] M. Thapa, S. Mulani, R. Walters, Stochastic multi-scale modeling of carbon fiber reinforced composites with polynomial chaos, Composite Structures.

[5] L. N. Collins, S. A. Roberts, Mesoscale simulation of woven composite design decisions, CoRR abs/2104.13554.

[6] F. Panerai, et al, Micro-tomography based analysis of thermal conductivity, diffusivity and oxidation behavior of rigid and flexible fibrous insulators, INT J HEAT MASS TRAN 108 (2017) 801–811.

[7] F. Semeraro, et al, Anisotropic analysis of fibrous and woven materials part 1: Estimation of local orientation, Comput. Mater. Sci. 178 (2020) 109631.

[8] F. Torres-Herrador, et al, Determination of heat capacity of carbon composites with application to carbon/phenolic ablators up to high temperatures, Aerosp. Sci. Technol. 108 (2021) 106375.

[9] S. Mahadevan, B. Liang, Error and uncertainty quantification and sensitivity analysis in mechanics computational models, Int. J. Uncertain. Quantif. 1 (2011) 147–161.

[10] M. Rivier, et al, Ablative thermal protection system under uncertainties including pyrolysis gas composition, Aerosp. Sci. Technol. 84 (2019) 1059–1069.

[11] W. Xiaoyong, et al, Uncertainty analysis of laminar and turbulent aeroheating predictions for mars entry, INT J HEAT MASS TRAN 112 (2017) 533–543.

[12] A. Turchi, et al, Thermochemical ablation modeling forward uncertainty analysis—part ii: Application to plasma wind-tunnel testing, Int. J. Therm. Sci. 118 (2017) 510–517.

[13] P. Rostkowski, et al, Effects of problem complexity reduction on parameter sensitivity and classification in charring ablator scenarios, Aerosp. Sci. Technol. 124 (2022) 107522.

[14] J. Ferguson, et al, Puma: the porous microstructure analysis software, SoftwareX 7 (2018) 81–87.

[15] S. Visser, et al, Microscale artificial weave generation capabilities for thermal protection system material modelling, in: 11th Ablation Workshop, University of Minnesota, 2019.

[16] J. R. Orero, Computational study on the effect of the microstrucutre on macroscopic properties for carbon fiber felts, Master's thesis (2020).

[17] A. Wiegmann, A. Zemitis, Ej-heat: A fast explicit jump harmonic averaging solver for the effective heat conductivity of composite materials, Tech. Rep. 94, Fraunhofer (ITWM) (2006).

[18] J. Ferguson, et al, Update 3.0 to "puma: The porous microstructure analysis software", SoftwareX 15 (2021) 100775. doi:https://doi.org/10.1016/j.softx.2021.100775.

[19] G. Vignoles, et al, Ablative and catalytic behavior of carbon-based porous thermal protection materials in nitrogen plasmas, Carbon 134 (2018) 376–390.

[20] A. Cullen, H. Frey, Probabilistic Techniques in Exposure Assessment. A Handbook for Dealing with Variability and Uncertainty in Models and Inputs, Plenum Press, New York and London, 1999.

[21] J. Kleijnen, Statistical validation of simulation models, European Journal of Operational Research 87 (1) (1995) 21–34.

[22] J. Kleijnen, R. Sargent, A methodology for fitting and validating metamodels in simulation, European Journal of Operational Research 120 (1) (2000) 14–29.

[23] A. Phillips, D. Janies, W. Wheeler, Multiple sequence alignment in phylogenetic analysis, Molecular phylogenetics and evolution 16 (2000) 317–30.

[24] A. Saltelli, K. Chan, E. M. Scott, Sensitivity Analysis, John Wiley & Sons, Ltd.: West Sussex, England., 2000.

[25] O. L. Maitre, O. Knio, Pc analysis of stochastic differential equations driven by wiener noise, Reliab. Eng. Syst. Saf. 135 (2015) 107–124.

[26] M. N. Jimenez, O. L. Maitre, O. Knio, Global Sensitivity Analysis in Stochastic Simulators of Uncertain Reaction Networks, J. Chem. Phys. 145.

[27] I. Sobol, On sensitivity estimation for nonlinear mathematical models, Matematicheskoe Modelirovanie 2 (1) (1990) 112–118.

[28] B. Sudret, S. Marelli, J. Wiart, Surrogate models for uncertainty quantification: An overview, in: 2017 11th European Conference on Antennas and Propagation (EUCAP), 2017, pp. 793–797.

[29] S. Azzi, Surrogate modeling of stochastic simulators, Ph.D. thesis (2020). URL http://www.theses.fr/2020IPPAT009

[30] N. Fajraoui, S. Marelli, B. Sudret, Sequential design of experiment for sparse polynomial chaos expansions, SIAM-ASA J. Uncertain. Quantif. 5 (2017) 1061–1085.

[31] R. Higdon, Generalized Additive Models, Springer New York, New York, NY, 2013. doi:10.1007/978-1-4419-9863-7_1197.

[32] X. Zhu, B. Sudret, Global sensitivity analysis for stochastic simulators based on generalized lambda surrogate models, Reliab. Eng. Syst. Saf. 214 (2021) 107815.

[33] M. Plumlee, R. Tuo, Building accurate emulators for stochastic simulations via quantile kriging, Technometrics 56 (4) (2014) 466–473.

[34] B. Ankenman, B. Nelson, J. Staum, Stochastic Kriging for Simulation Metamodeling, Operations Research 58 (2) (2010) 371–382.

[35] R. Rigby, D. Stasinopoulos, Construction of reference centiles using mean and dispersion additive models, J R Stat Soc Ser D 49 (2000) 41 – 50.

[36] B. J. Reich, E. Kalendra, C. B. Storlie, H. D. Bondell, M. Fuentes, Variable selection for high dimensional bayesian density estimation: application to human exposure simulation, J R Stat Soc Ser C 61 (1) (2012) 47–66.

[37] N. Razaaly, P. Congedo, Novel algorithm using active metamodel learning and importance sampling: Application to multiple failure regions of low probability, J. Comput. Phys. 368.

[38] C. Rasmussen, C. Williams, Gaussian Processes for Machine Learning, the MIT Press, 2006.

[39] G. Vining, R. Myers, Combining taguchi and response surface philosophies: A dual response approach, J QUAL TECHNOL 22 (1) (1990) 38–45.

[40] I. Zabalza, et al, Prediction and density estimation of a horizontal well productivity index using generalized linear, European Association of Geoscientists & Engineers, 1998.

[41] A. Marrel, B. Iooss, Advanced methodology for uncertainty propagation in computer experiments with large number of inputs : application to accidental scenario in a pressurized water reactor, in: European Safety and Reliability Conference (ESREL - 2018), Trondheim, Norway, 2018.

[42] A. Marrel, et al, Global Sensitivity Analysis of Stochastic Computer Models with joint metamodels, Stat Comput 22 (2012) 833–847.

[43] W. Wang, X. Chen, The effects of estimation of heteroscedasticity on stochastic kriging (2016). doi:10.1109/WSC.2016.7822100.

[44] F. Girault, P. M. Congedo, Uncertainty Characterization of ablative materials for atmospheric reentry of spacecrafts using PuMA, Research Report RR-9417, Inria Saclay -Île de France (Sep. 2021).
URL https://inria.hal.science/hal-03384925

[45] C. Pradère, et al, Thermal properties of carbon fibers at very high temperature, Carbon 47 (3) (2009) 737–743.

[46] J. Scoggins, et al, Mutation++: Multicomponent thermodynamic and transport properties for ionized gases in c++, SoftwareX 12 (2020) 100575.

[47] S. Marelli, B.Sudret, Uqlab: A framework for uncertainty quantification in matlab, The 2nd International Conference on Vulnerability and Risk Analysis and Management (2014) 2554–2563doi:10.1061/9780784413609.257.