Statistical metamodel of liner acoustic impedance based on neural network and probabilistic learning for small datasets

Amritesh Sinha, Christophe Desceliers, Christian Soize, Guilherme Cunha

Statistical metamodel of liner acoustic impedance based on neural network and probabilistic learning for small datasets

Amritesh Sinha[a], Christophe Desceliers[a,*], Christian Soize[a], Guilherme Cunha[b]

[a] *Université Gustave Eiffel, MSME UMR 8208, 5 bd Descartes, 77454 Marne-la-Vallée, France*
[b] *Airbus Operations SAS, 31300 Toulouse, France*

Abstract

The main novelty of this paper consists in presenting a statistical Artificial Neural Network (ANN) based model for a robust prediction of the frequency-dependent aeroacoustic liner impedance using an Aeroacoustic Computational Model (ACM) dataset of small size. The model, focusing on Percentage of Open Area (POA) and Sound Pressure Level (SPL) at a zero Mach number, takes into accounts uncertainties using a probabilistic formulation. The main difficulty in training an ANN-based model is the small size of the ACM dataset. The probabilistic learning carried out using the Probabilistic Learning on Manifold (PLoM) algorithm addresses this difficulty as it allows constructing a very large training dataset from learning the probabilistic model from a small dataset. A *prior* conditional probability model is presented for the PCA-based statistical reduced representation of the frequency-sampled vector of the log-resistance and reactance. It induces some statistical constraints that are not straightforwardly taken into account when training such an ANN-based model by classical optimizations methods under constraints. A second novelty of this paper consists in presenting an alternate solution that involves using conditional statistics estimated with learned realizations from PLoM. A numerical example is presented.

Keywords: Statistical metamodel, Neural Networks, Probabilistic learning on manifolds, Liner acoustic impedance, Small data.

*Corresponding author: C. Desceliers, christophe.desceliers@univ-eiffel.fr

Email addresses: amritesh.sinha@univ-eiffel.fr (Amritesh Sinha), christophe.desceliers@univ-eiffel.fr (Christophe Desceliers), christian.soize@univ-eiffel.fr (Christian Soize), guilherme.coelho-cunha@airbus.com (Guilherme Cunha)

Nomenclature

$[C_{\mathbf{H}\|\mathbf{W}}(\mathbf{w})]$	Conditional covariance matrix of $\mathbf{H}$ given $\mathbf{W} = \mathbf{w}$
$[C_{\mathbf{Q}}]$	Covariance matrix of $\mathbf{Q}$
$[C_{\mathbf{R}\|\mathbf{W}}(\mathbf{w})]$	Conditional covariance matrix of $\mathbf{R}$ given $\mathbf{W} = \mathbf{w}$
$[C_{\mathbf{V}\|\mathbf{W}}(\mathbf{w})]$	Conditional covariance matrix of $\mathbf{V}$ given $\mathbf{W} = \mathbf{w}$
$E\{\cdot\}$	Mathematical expectation operator
$J(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$	Cost function for ANN training
n_d	Number of values in $\mathcal{D}_{\mathbf{w}}$
$n_{\mathbf{w}}$	Number of control parameters
n_ω	Number of sampled frequencies
$p_{\mathbf{Q}\|\mathbf{W}}, p_{\mathbf{R}\|\mathbf{W}}, p_{\mathbf{V}\|\mathbf{W}}$	Conditional probability density functions
$p_{\mathbf{W}}$	Probability density function of $\mathbf{W}$
$r^{\,\mathrm{acm}}(\omega; \mathbf{w})$	Resistance from ACM
s	Silverman bandwidth
$v^{\,\mathrm{acm}}(\omega; \mathbf{w})$	Reactance from ACM
$\boldsymbol{\mu}^j$	j-th realization of conditional mean
$\mathbf{Q}, \mathbf{R}, \mathbf{V}$	Random vectors for impedance, resistance, and reactance
$\underline{\mathbf{q}}$	Empirical mean value of $\mathbf{Q}$
$\mathbf{q}^{\,\mathrm{acm},j}$	j-th frequency-sampled impedance from ACM
$\mathbf{q}^{j,k}$	k-th realization of frequency-sampled impedance given $\mathbf{W} = \mathbf{w}^j$
$\mathbf{q}_{\mathrm{ar}}^{\ell}$	ℓ-th additional realization of frequency-sampled impedance
$\mathbf{r}^{\,\mathrm{acm},j}$	j-th frequency-sampled resistance from ACM
$\mathbf{w} = (w_1, w_2)$	Control parameters (POA and SPL)
$\mathbf{w}^j$	j-th realization of control parameters
$\mathbf{w}^{j,k}$	Rewritting (with repitition) of $\mathbf{w}^{j,k} = \mathbf{w}^j$
$\mathbf{w}_{\mathrm{ar}}^{\ell}$	ℓ-th additional realization of control parameters
$\mathbf{W}$	Random vector of control parameters
$\mathbf{z}^{\,\mathrm{acm}}(\omega; \mathbf{w})$	Frequency-dependent acoustic impedance from ACM
$\boldsymbol{\zeta}^j$	j-th realization of covariance matrix parameters
$[\lambda]$	Diagonal matrix of eigenvalues
$[\phi]$	Matrix of eigenvectors
$\mathbf{H}$	Normalized random vector from PCA
$\boldsymbol{\mu}_{\mathbf{H}\|\mathbf{W}}(\mathbf{w};\, \boldsymbol{\theta}_1)$	ANN output for conditional mean of $\mathbf{H}$ given $\mathbf{W} = \mathbf{w}$
$\boldsymbol{\zeta}_{\mathbf{H}\|\mathbf{W}}(\mathbf{w};\, \boldsymbol{\theta}_2)$	ANN output for vectorized upper triangular elements of the matrix logarithm of $[C_{\mathbf{H}\|\mathbf{W}}(\mathbf{w})]$
$\mathcal{L}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$	Negative log-likelihood
ω	Frequency (rad/s)
$\boldsymbol{\theta}_1, \boldsymbol{\theta}_2$	Parameters of the ANN

ACM	Aeroacoustic Computational Model
ANN	Artificial Neural Network
BPF	Blade Passing Frequency
GKDE	Gaussian Kernel Density Estimation
MaxEnt	Maximum Entropy
PCA	Principal Component Analysis
PLoM	Probabilistic Learning on Manifold
POA	Percentage of Open Area
SPL	Sound Pressure Level
UHBR	Ultra-High Bypass Ratio

$\mathcal{D}_{\mathrm{acm}}$	ACM dataset
$\mathcal{D}^*_{\mathrm{acm}}$	Training dataset
$\mathcal{D}^*_{\mathrm{ar}}$	Learned dataset
$\mathcal{D}^*_{\mathbf{H}\|\mathbf{W}}$	GKDE-based estimates dataset
$\mathcal{D}_{\mathbf{w}}$	Set of control parameter values

1. Introduction

We are interested in the aircraft noise in the context of green aviation. This includes both external and internal noise relative to the aircraft for certification purposes. In modern turbofan engines, especially those with an ultra-high bypass ratio (UHBR), fan noise is a major contributor to overall noise. Fan noise comprises broadband and tonal noise. To attenuate the tonal noise component, particularly at blade passing frequency (BPF), targeted acoustic liners are used, while modifications to the liners geometry and intrinsic properties help dissipate the broadband noise component. For effectiveness, it is crucial to study liners under various operating conditions, including different flight scenarios. The design of liners is critically important and has been extensively studied (see [1, 2, 3, 4, 5, 6, 7]). Liner acoustic impedance can be calculated by using an Aeroacoustic Computational Model (ACM), which is computationally expensive and prohibitive high-fidelity computational model for optimizing aeroacoustic performances of the liners using a numerical optimization solver. Such ACM have been developed to predict liner performance, as detailed in [8, 9, 10, 11, 12, 13]. An uncertain ACM of the liner system, which allows for quantifying uncertainties in aeroacoustic models of liner performance, was also developed in [6].

The objective of this paper is to construct a statistical metamodel for which the outputs are the frequency-sampled impedance of the liner and the inputs are the control parameters that are the design parameters. The metamodel should have a low-computational cost to enable its usage in optimization of liners aeroacoustic performance *via* a numerical optimization solver. In addition, the gradients of the metamodel with respect to its inputs (the control parameters) should also have a low-computational cost.

In this paper, we construct the metamodel using a dataset, referred to as the ACM dataset, which includes samples of control parameters and the corresponding acoustic impedance, numerically simulated by ACM. Consequently, the dataset is small due to the prohibitive computational cost of ACM, which prevents the construction of a large dataset.

As only the components of control parameter are inputs to the metamodel, all other ACM parameters are unobserved (thus uncontrolled) and should be treated as random latent variables. Therefore, the liner acoustic impedance at any frequency should also be modeled as a random variable. A first novelty contribution of this paper is the methodology for constructing such a statistical metamodel that is driven by the physics contained in the small ACM dataset.

In this paper, the statistical metamodel is defined by the conditional probability distribution of frequency-sampled vector of the random acoustic impedance at sampled frequencies, given the control parameter. The MaxEnt principle, applied with the available information, is used to construct an informative *prior* model of this conditional probability distribution. It should be noted that the hyperparameters of the probabilistic model depend on the values of the control parameter and are modeled using fully connected feedforward neural networks yielding a statistical ANN-based metamodel. Such a statistical ANN-based metamodel is fitted on an *ad hoc* training dataset using the maximum likelihood principle.

A principal component analysis (PCA) is conducted on the outputs (frequency-sampled acoustic impedance) of the metamodel. This is not only for potential statistical reduction but also because decorrelation and centering of outputs enhance numerical conditioning, thereby facilitating the optimization process for fitting such a statistical ANN-based model to the training dataset. However, such a statistical decorrelation and centering of the outputs introduce constraints on the hyperparameters of the statistical metamodel and consequently, on the parameters of the statistical ANN-based metamodel. Therefore, given that these hyperparameters are modeled by fully connected feedforward networks, some deterministic constraints must be considered, yielding the development of constrained training algorithm for such networks. Such a constrained training algorithm is challenging and complex when dealing with mini-batches, and might rely on techniques such as Lagrange multipliers, penalization approaches, correction formulations, etc. A second novelty of this paper is the presentation of an unconstrained formulation for a statistical ANN-based metamodel that takes into account the statistical constraints arising from the PCA-based reduction. This is achieved by generating an *ad hoc* training dataset using Probabilistic Learning on Manifold [14, 15, 16, 17] and using the learned dataset for estimating conditional statistics in a nonparametric framework, based on Gaussian Kernel Density Estimation (GKDE) method. Finally, the statistical metamodel can be used to generate additional realizations of the frequency-sampled acoustic impedance vector, thereby mitigating missing data in the set of control parameters. Compared to the previous work [18], this ANN-based surrogate model is actually a complement. First, it does not require the training dataset for offline use, significantly reducing memory requirements. Second, it allows for more efficient computation of gradients with respect to input parameters, which is crucial for optimization problems.

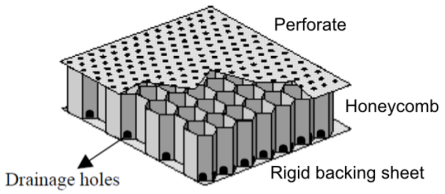


Figure 1: Schematic of the acoustic liner (adapted from [8]).

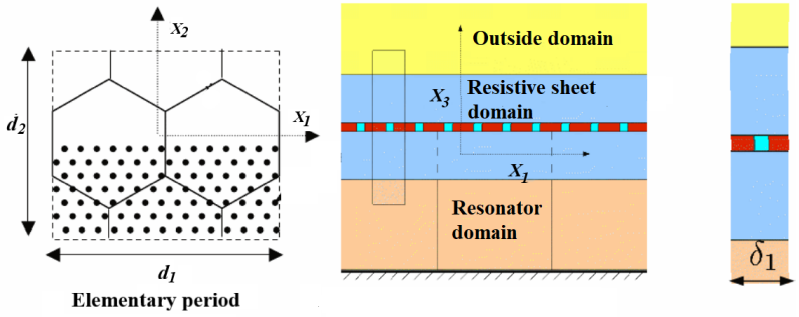


Figure 2: Reduced domain of the computational model (adapted from [8]).

Note that the objective of the paper is not to improve the ANN algorithms, but is to present a novel methodological application of ANNs to probabilistic ACM datasets for aeroacoustic liner impedance simulations. The novelty lies in how ANNs are used to model the hyperparameters of the probabilistic model (Sections 4.1 and 4.2) and in how to train this model using a learned GKDE-based estimates dataset by PLoM (Section 4.3). This approach allows us to effectively deal with small and probabilistic dataset, which is a significant challenge in this field and has not been widely addressed in the literature yet.

The paper is organized as follows: (1) Section 2 briefly defines the control parameters and the ACM for calculating the frequency-sampled vector of acoustic impedance, which is to be used for constructing the ACM dataset (small size). (2) Section 3 is dedicated to the parametric probabilistic modeling of the conditional probability distribution of the frequency-sampled vector of acoustic impedance given the control parameter. (3) Section 4 focuses on the statistical ANN-based metamodel given the control parameter. (4) Section 5 presents a numerical example, along with a discussion of the results.

2. Control parameters and ACM dataset

In this paper, the considered system is an acoustic liner consisting of a perforated plate, a honeycomb structure, and a rigid backing plate, as depicted in Fig. 1. The Mach number is assumed to be equal to zero, that is relevant for the ground configuration. Incorporating non-zero Mach numbers require additional analyses beyond the scope of this paper. A reduced domain, as described in [8], is used for the computational model (see Fig. 2). In this paper and for sake of simplicity, the liner system is parameterized by $n_{\mathbf{w}} = 2$ parameters (the control parameters) that are the Percentage of Open Area (POA) and the Sound Pressure Level (SPL), represented by $\mathbf{w} = (w_1, w_2)$, where w_1 corresponds to POA and w_2 to SPL.

For such a reduced domain, the frequency-dependent acoustic impedance is denoted as $\omega \mapsto \mathbf{z}^{\text{acm}}(\omega; \mathbf{w})$, where ω is the frequency (rad/s). Specifically, $\mathbf{z}^{\text{acm}}(\omega; \mathbf{w}) = r^{\text{acm}}(\omega; \mathbf{w}) + \imath\, v^{\text{acm}}(\omega; \mathbf{w})$, in which $r^{\text{acm}}(\omega; \mathbf{w})$ represents the resistance that is positive, $v^{\text{acm}}(\omega; \mathbf{w})$ the reactance that is real, $\imath = \sqrt{-1}$. Control parameter $\mathbf{w}$ belongs to an admissible set. The computational domain is centered around the resonator, within which the Navier-Stokes equations are solved. The computational model consists of 278 514 degrees of freedom. For given control parameter $\mathbf{w}$, and for each

sampled frequency $\omega_k = (k-1)\,\Delta\omega$ with $k = 1, \ldots, n_{\mathbf{w}}$ where $n_\omega = 7$, the ACM computes the resistance $r^{\mathrm{acm}}(\omega_k; \mathbf{w})$ and the reactance $v^{\mathrm{acm}}(\omega_k; \mathbf{w})$. The Navier-Stokes equations are solved using the numerical method presented in [8] for $\mathbf{w} = (w_1, w_2) \in [0.03, 0.1] \times [130, 145]$. As discussed in [18], ACM simulations are performed for $n_d = 48$ values $\mathbf{w}^1, \ldots, \mathbf{w}^{n_d}$ of $\mathbf{w}$, which constitute the set $\mathcal{D}_{\mathbf{w}}$. These points in $\mathcal{D}_{\mathbf{w}}$ are considered as realizations of a random vector $\mathbf{W}$ whose probability density function, $p_{\mathbf{W}}$, is unknown.

The ACM dataset $\mathcal{D}_{\mathrm{acm}}$ is then defined as the set of points $(\mathbf{w}^j, \mathbf{r}^{\mathrm{acm},j}, \mathbf{v}^{\mathrm{acm},j})$ in $\mathbb{R}^{n_{\mathbf{w}} \times n_\omega \times n_\omega}$ for $j = 1, \ldots, n_d$, in which $\mathbf{r}^{\mathrm{acm},j} = (r^{\mathrm{acm}}(\omega_1; \mathbf{w}^j), \ldots, r^{\mathrm{acm}}(\omega_{n_\omega}; \mathbf{w}^j))$ and $\mathbf{v}^{\mathrm{acm},j} = (v^{\mathrm{acm}}(\omega_1; \mathbf{w}^j), \ldots, v^{\mathrm{acm}}(\omega_{n_\omega}; \mathbf{w}^j))$. For each $j = 1, \ldots, n_d$, we introduce the $2\,n_\omega$-dimensional vector of frequency-sampled impedance $\mathbf{q}^{\mathrm{acm},j} = (\log \mathbf{r}^{\mathrm{acm},j}, \mathbf{v}^{\mathrm{acm},j})$ in which $\log \mathbf{r}^{\mathrm{acm},j} = (\log r^{\mathrm{acm}}(\omega_1; \mathbf{w}^j), \ldots, \log r^{\mathrm{acm}}(\omega_{n_\omega}; \mathbf{w}^j))$.

The model uncertainties are due to random latent parameters that, consequently, cannot be defined as control parameters. In order to take into account these model uncertainties, the random vectors $\mathbf{Q}, \mathbf{R}$, and $\mathbf{V}$ are introduced whose conditional probability density functions, given $\mathbf{W} = \mathbf{w}^j$, are denoted by $\mathbf{q} \mapsto p_{\mathbf{Q}|\mathbf{W}}(\mathbf{q}\,|\,\mathbf{w}^j)$, $p_{\mathbf{R}|\mathbf{W}}(\mathbf{r}\,|\,\mathbf{w}^j)$, and $p_{\mathbf{V}|\mathbf{W}}(\mathbf{v}\,|\,\mathbf{w}^j)$. These conditional probability density functions are constructed as explained in [18]. For instance, the conditional mean value of $\mathbf{Q}$ given $\mathbf{W} = \mathbf{w}^j$ is chosen as $\mathbf{q}^{\mathrm{acm},j}$ for $j = 1, \ldots, n_d$; the conditional dispersion coefficient given $\mathbf{W} = \mathbf{w}^j$ has been identified using experimental data. We then generate $m_d = 15$ statistically independent realizations $\mathbf{q}^{j,1}, \ldots, \mathbf{q}^{j,m_d}$ from $p_{\mathbf{Q}|\mathbf{W}}$ given $\mathbf{W} = \mathbf{w}^j$. Hence, training dataset $\mathcal{D}^*_{\mathrm{acm}}$ is made up of a total of $n_{\mathrm{acm}} = n_d \times m_d$ realizations $(\mathbf{w}^{j,k}, \mathbf{q}^{j,k})$ with $j = 1, \ldots, n_d$ and $k = 1, \ldots, m_d$, in which $\mathbf{w}^{j,k}$ is a rewriting of $\mathbf{w}^j$ that is independent of k (we introduce a repetition). For the sake of simplicity, all the realizations are rewritten as $(\mathbf{w}^j, \mathbf{q}^j)$ with $j = 1, \ldots, n_{\mathrm{acm}}$.

In [18], the PLoM (Probabilistic Learning on Manifold) is carried out to learn the joint probability density function $p_{\mathbf{Q},\mathbf{W}}$ of random vectors $\mathbf{Q}$ and $\mathbf{W}$ using $\mathcal{D}^*_{\mathrm{acm}}$ as a training dataset. PLoM also allows n_{ar} additional statistically independent realizations $\{(\mathbf{w}^\ell_{\mathrm{ar}}, \mathbf{q}^\ell_{\mathrm{ar}}),\ \ell = 1, \ldots, n_{\mathrm{ar}}\}$ to be generated, which constitute the learned dataset $\mathcal{D}^*_{\mathrm{ar}}$.

3. Prior probabilistic model of the frequency-sampled impedance vector

In [18], the conditional probability density function $p_{\mathbf{Q}|\mathbf{W}}$ is estimated with the Gaussian Kernel Density Estimation (GKDE) using additional realizations (generated by PLoM) of $\mathbf{Q}$ and $\mathbf{W}$. In the context of the construction of a statistical metamodel based on a neural network, we need to introduce an algebraic representation of the conditional probability distribution of $\mathbf{Q}$ given $\mathbf{W}$, depending on hyperparameters. In this paper we have chosen a Gaussian model for which the hyperparameters are conditional mean value and the conditional covariance matrix of $\mathbf{Q}$ given $\mathbf{W}$. The neural network will be used for predicting these conditional hyperparameters. In this section, we then present the construction of the algebraic *prior* probabilistic model of $p_{\mathbf{Q}|\mathbf{W}}$. Nevertheless, since $\mathbf{Q}$ is in high dimension, we will introduce a statistical reduction $\mathbf{H}$ of $\mathbf{Q}$ using a PCA. Within Section 3.2, the conditional hyperparameters associated with such a *prior* probabilistic model are represented as functions of $\mathbf{w}$. The modeling of these functions is carried out using fully connected feedforward networks that are trained to map control parameter $\mathbf{w}$ onto a corresponding set of hyperparameters of the *prior* probabilistic model.

3.1. PCA-based statistical reduction $\boldsymbol{H}$ of $\boldsymbol{Q}$

A PCA is used to construct the statistical reduction of $\mathbf{Q}$, yielding a normalized random vector $\mathbf{H}$ (centered with identity covariance matrix). Random vector $\mathbf{H}$ is therefore written as $\mathbf{H} = [\lambda]^{-1/2}\,[\phi]^T\,(\mathbf{Q} - \underline{\mathbf{q}})$, in which $\underline{\mathbf{q}} = (1/n_{\mathrm{ar}}) \sum_{\ell=1}^{n_{\mathrm{ar}}} \mathbf{q}^\ell_{\mathrm{ar}}$ is the empirical mean value of random vector $\mathbf{Q}$, $[\lambda]$ is a $(m \times m)$ diagonal matrix, and $[\phi]$ is a $(n_q \times m)$ matrix whose columns are orthonormal vectors, and are such that $[C_{\mathbf{Q}}][\phi] = [\phi][\lambda]$. The estimate of the $(n_{\mathbf{q}} \times n_{\mathbf{q}})$ covariance matrix of $\mathbf{Q}$ is $[C_{\mathbf{Q}}] = (n_{\mathrm{ar}} - 1)^{-1} \sum_{\ell=1}^{n_{\mathrm{ar}}} (\mathbf{q}^\ell_{\mathrm{ar}} - \underline{\mathbf{q}})\,(\mathbf{q}^\ell_{\mathrm{ar}} - \underline{\mathbf{q}})^T$. The diagonal entries of $[\lambda]$ are the m-largest eigenvalues of $[C_{\mathbf{Q}}]$. By construction, the $\mathbb{R}^m$-valued random variable $\mathbf{H}$ is such

$$E\{\mathbf{H}\} = 0_m \quad , \quad E\{\mathbf{H} \otimes \mathbf{H}\} = [\mathrm{I}_m]\,. \tag{1}$$

3.2. Prior conditional probabilistic density function of $\boldsymbol{H}$ given $\boldsymbol{W}$

Let $\boldsymbol{\eta} \mapsto p_{\mathbf{H}|\mathbf{W}}(\boldsymbol{\eta}|\mathbf{w})$ be the conditional probability density function of $\mathbf{H}$ given $\mathbf{W}$. The *prior* conditional probability density function $p_{\mathbf{H}|\mathbf{W}}$ is constructed using the MaxEnt principle (see for instance [19]) with the following available information: (1) the support of $\boldsymbol{\eta} \mapsto p_{\mathbf{H}|\mathbf{W}}(\boldsymbol{\eta}|\mathbf{w})$ is $\mathbb{R}^m$, (2) the conditional mean value and the conditional covariance

matrix of $\mathbf{H}$ given $\mathbf{W} = \mathbf{w}$ are the vector $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w})$ and the matrix $[C_{\mathbf{H}|\mathbf{W}}(\mathbf{w})]$, which are estimated using the training dataset $\mathcal{D}^*_{\text{ar}}$ for each given value of $\mathbf{w}$. Therefore, $\mathbf{H}$ given $\mathbf{W} = \mathbf{w}$ is a multivariate Gaussian random variable with mean value $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w})$ and covariance matrix $[C_{\mathbf{H}|\mathbf{W}}(\mathbf{w})]$.

For any given $\mathbf{w}$ in its admissible set, the estimate of hyperparameters $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w})$ and $[C_{\mathbf{H}|\mathbf{W}}(\mathbf{w})]$, constructed using GKDE from nonparametric statistics and dataset $\mathcal{D}^*_{\text{ar}}$, are written as

$$\boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w}) = \frac{\sum_{\ell=1}^{n_{\text{ar}}} \boldsymbol{\eta}^{\ell}_{\text{ar}} \exp\left(-\frac{1}{2s^2}(\mathbf{w}-\mathbf{w}^{\ell}_{\text{ar}})^T\,[C_{\mathbf{W}}]^{-1}\,(\mathbf{w}-\mathbf{w}^{\ell}_{\text{ar}})\right)}{\sum_{\ell=1}^{n_{\text{ar}}} \exp\left(-\frac{1}{2s^2}(\mathbf{w}-\mathbf{w}^{\ell}_{\text{ar}})^T\,[C_{\mathbf{W}}]^{-1}\,(\mathbf{w}-\mathbf{w}^{\ell}_{\text{ar}})\right)}\,, \tag{2}$$

$$[C_{\mathbf{H}|\mathbf{W}}(\mathbf{w})] = \frac{\sum_{\ell=1}^{n_{\text{ar}}} \widetilde{\boldsymbol{\eta}}^{\ell}_{\text{ar}}(\mathbf{w})\,(\widetilde{\boldsymbol{\eta}}^{\ell}_{\text{ar}}(\mathbf{w}))^T \exp\left(-\frac{1}{2s^2}(\mathbf{w}-\mathbf{w}^{\ell}_{\text{ar}})^T\,[C_{\mathbf{W}}]^{-1}\,(\mathbf{w}-\mathbf{w}^{\ell}_{\text{ar}})\right)}{\sum_{\ell=1}^{n_{\text{ar}}} \exp\left(-\frac{1}{2s^2}(\mathbf{w}-\mathbf{w}^{\ell}_{\text{ar}})^T\,[C_{\mathbf{W}}]^{-1}\,(\mathbf{w}-\mathbf{w}^{\ell}_{\text{ar}})\right)}\,, \tag{3}$$

where (1) the $(n_{\mathbf{w}} \times n_{\mathbf{w}})$ matrix $[C_{\mathbf{W}}] = (n_{\text{ar}}-1)^{-1}\sum_{\ell=1}^{n_{\text{ar}}}(\mathbf{w}^{\ell}_{\text{ar}}-\underline{\mathbf{w}})\,(\mathbf{w}^{\ell}_{\text{ar}}-\underline{\mathbf{w}})^T$ is the estimate of the covariance matrix of $\mathbf{W}$ in which $\underline{\mathbf{w}} = (1/n_{\text{ar}})\sum_{\ell=1}^{n_{\text{ar}}}\mathbf{w}^{\ell}_{\text{ar}}$ is the estimate of the mean value of $\mathbf{W}$; (2) for all $\ell = 1,\ldots,n_{\text{ar}}$, we have $\boldsymbol{\eta}^{\ell}_{\text{ar}} = [\lambda]^{-1/2}[\phi]^T(\mathbf{q}^{\ell}_{\text{ar}}-\underline{\mathbf{q}})$ and $\widetilde{\boldsymbol{\eta}}^{\ell}_{\text{ar}}(\mathbf{w}) = \boldsymbol{\eta}^{\ell}_{\text{ar}} - \boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w})$; (3) s is the Silverman bandwidth given by

$$s = \left\{\frac{4}{n_{\text{ar}}(2+n)}\right\}^{1/(n+4)} \quad , \quad n = m + n_{\mathbf{w}} \tag{4}$$

Due to Eqn (1), $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w})$ and $[C_{\mathbf{H}|\mathbf{W}}(\mathbf{w})]$ have to satisfy the following equations,

$$E\{\mu_{\mathbf{H}|\mathbf{W}}(\mathbf{W})\} = 0_m \quad , \tag{5}$$

$$E\left\{[C_{\mathbf{H}|\mathbf{W}}(\mathbf{W})] + \boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{W})\boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{W})^T\right\} = [I_m] \quad . \tag{6}$$

With the proposed methodology, these two equations will automatically be satisfied.

*3.3. Statistically independent realizations of **R** and **V** given **W***

For given $\mathbf{w}$, let $\mathbf{R}(\mathbf{w})$ and $\mathbf{V}(\mathbf{w})$ be the $\mathbb{R}^{n_\omega}$-valued random variables defined in Section 2. For any given $\mathbf{w}$ in its admissible set, N statistically independent realizations $\boldsymbol{\eta}^1(\mathbf{w}),\ldots,\boldsymbol{\eta}^N(\mathbf{w})$ of $\mathbf{H}$ given $\mathbf{W} = \mathbf{w}$ are generated using the multivariate Gaussian random distribution whose mean value is $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w})$ and covariance matrix is $[C_{\mathbf{H}|\mathbf{W}}(\mathbf{w})]$. We then deduce N statistically independent realizations $\mathbf{q}^1(\mathbf{w}),\ldots,\mathbf{q}^N(\mathbf{w})$ of random vector $\mathbf{Q}$ given $\mathbf{W} = \mathbf{w}$ such that, for $j = 1,\ldots,N$, $\mathbf{q}^j(\mathbf{w}) = \underline{\mathbf{q}} + [\Phi]\,[\lambda]^{1/2}\,\boldsymbol{\eta}^j(\mathbf{w})$. For $j = 1,\ldots,N$, the block decomposition of vector $\mathbf{q}^j(\mathbf{w})$ is written as $(\mathbf{q}^j_{\mathbf{R}}(\mathbf{w}),\,\mathbf{v}^j(\mathbf{w}))$ with $\mathbf{q}^j_{\mathbf{R}}$ and $\mathbf{v}^j(\mathbf{w})$ being two n_ω dimensional vectors. Obviously, $\mathbf{v}^1(\mathbf{w}),\ldots,\mathbf{v}^N(\mathbf{w})$ are statistically independent realizations of $\mathbf{V}(\mathbf{w})$ and the statistically independent realizations $\mathbf{r}^1(\mathbf{w}),\ldots,\mathbf{r}^N(\mathbf{w})$ of $\mathbf{R}(\mathbf{w})$ are such that $\mathbf{q}^j_{\mathbf{R}}(\mathbf{w}) = \log(\mathbf{r}^j(\mathbf{w}))$. Note that, for given $\mathbf{w}$, conditional mean vectors $\boldsymbol{\mu}_{\mathbf{R}|\mathbf{W}}(\mathbf{w})$ and $\boldsymbol{\mu}_{\mathbf{V}|\mathbf{W}}(\mathbf{w})$, and the conditional covariance matrices $[C_{\mathbf{R}|\mathbf{W}}(\mathbf{w})]$ and $[C_{\mathbf{V}|\mathbf{W}}(\mathbf{w})]$ are estimated using the statistically independent realizations $\mathbf{r}^j(\mathbf{w}),\ldots,\mathbf{r}^N(\mathbf{w})$ and $\mathbf{v}^j(\mathbf{w}),\ldots,\mathbf{v}^N(\mathbf{w})$. Note that vector $\boldsymbol{\mu}_{\mathbf{R}|\mathbf{W}}(\mathbf{w})$ and matrix $[C_{\mathbf{R}|\mathbf{W}}(\mathbf{w})]$ can also be written as

$$\boldsymbol{\mu}_{\mathbf{R}|\mathbf{W}}(\mathbf{w}) = \exp\left(\boldsymbol{\mu}_{\mathbf{Q_R}|\mathbf{W}}(\mathbf{w}) + \frac{1}{2}\text{diag}\,[\mathrm{C}_{\mathbf{Q_R}|\mathbf{W}}(\mathbf{w})]\right) \quad , \tag{7}$$

$$[C_{\mathbf{R}|\mathbf{W}}(\mathbf{w})] = \exp\left([C_{\mathbf{Q_R}|\mathbf{W}}(\mathbf{w})] - 1\right) \odot \left(\boldsymbol{\mu}_{\mathbf{R}|\mathbf{W}}(\mathbf{w})\,\boldsymbol{\mu}_{\mathbf{R}|\mathbf{W}}(\mathbf{w})^T\right)\,, \tag{8}$$

where $[A] \odot [B]$ stands for Hadamard product of matrices $[A]$ and $[B]$; $\exp([A])$ is the element-wise exponential and diag [A] is the vector made up of the diagonal entries of a given matrix $[A]$. The block decomposition of $\boldsymbol{\mu}_{\mathbf{Q}|\mathbf{W}}(\mathbf{w}) = \underline{\mathbf{q}} + [\Phi]\,[\lambda]^{1/2}\,\boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w})$ into n_ω dimensional vectors $\boldsymbol{\mu}_{\mathbf{Q_R}|\mathbf{W}}(\mathbf{w})$ and $\boldsymbol{\mu}_{\mathbf{V}|\mathbf{W}}(\mathbf{w})$, is written as

$$\boldsymbol{\mu}_{\mathbf{Q}|\mathbf{W}}(\mathbf{w}) = (\boldsymbol{\mu}_{\mathbf{Q_R}|\mathbf{W}}(\mathbf{w}),\,\boldsymbol{\mu}_{\mathbf{V}|\mathbf{W}}(\mathbf{w}))\,. \tag{9}$$

The block decomposition of matrix $[C_{\mathbf{Q}|\mathbf{W}}(\mathbf{w})] = [\Phi]\,[\lambda]^{1/2}\,[C_{\mathbf{H}|\mathbf{W}}(\mathbf{w})]\,[\lambda]^{1/2}\,[\Phi]^T$ into $(n_\omega \times n_\omega)$ matrices $[C_{\mathbf{Q_R}|\mathbf{W}}(\mathbf{w})]$, $[C(\mathbf{w})]$, and $[C_{\mathbf{V}|\mathbf{W}}(\mathbf{w})]$, is written as

$$[C_{\mathbf{Q}|\mathbf{W}}(\mathbf{w})] = \begin{pmatrix} [C_{\mathbf{Q_R}|\mathbf{W}}(\mathbf{w})] & [C(\mathbf{w})] \\ [C(\mathbf{w})] & [C_{\mathbf{V}|\mathbf{W}}(\mathbf{w})] \end{pmatrix}\,. \tag{10}$$

Concerning the random frequency-sampled vector $\mathbf{V}(\mathbf{w})$, vector $\boldsymbol{\mu}_{\mathbf{V}|\mathbf{W}}(\mathbf{w})$ and matrix $[C_{\mathbf{V}|\mathbf{W}}(\mathbf{w})]$ are directly obtained from block decomposition of vector $\boldsymbol{\mu}_{\mathbf{Q}|\mathbf{W}}(\mathbf{w})$ and matrix $[C_{\mathbf{Q}|\mathbf{W}}(\mathbf{w})]$ given by Eqs. (9) and (10).

4. Statistical ANN-based metamodel

4.1. Fully connected feedforward neural network

Deterministic mappings $\mathbf{w} \mapsto \mu_{\mathbf{H}|\mathbf{W}}(\mathbf{w})$ and $\mathbf{w} \mapsto [C_{\mathbf{H}|\mathbf{W}}(\mathbf{w})]$ may have a complex behavior, not only because their supports are multidimensional, but also because the underlying physical process is complex. In such a case, fully connected feedforward neural network is well adapted to represent such deterministic mappings. We then consider a fully connected feedforward neural network $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w}\ \boldsymbol{\theta}_1)$ with parameter $\boldsymbol{\theta}_1$, which is constructed in order to model deterministic mapping $\mathbf{w} \mapsto \mu_{\mathbf{H}|\mathbf{W}}(\mathbf{w})$. However, a representation of $\mathbf{w} \mapsto [C_{\mathbf{H}|\mathbf{W}}(\mathbf{w})]$ by a fully connected feedforward network is not straightforward because, for each given $\mathbf{w}$, the output has to be a positive-definite matrix. To circumvent this apparent difficulty, the matrix logarithm of the $(m \times m)$ symmetric covariance matrix $[C_{\mathbf{H}|\mathbf{W}}(\mathbf{w})]$ is calculated for each given $\mathbf{w}$, which yields a $(m \times m)$ symmetric matrix $[\log C_{\mathbf{H}|\mathbf{W}}(\mathbf{w})]$. If all the $m(m+1)/2$ entries of the upper triangular block of matrix $[\log C_{\mathbf{H}|\mathbf{W}}(\mathbf{w})]$ are collected into the $m(m+1)/2$ dimensional vector $\boldsymbol{\zeta}_{\mathbf{H}|\mathbf{W}}(\mathbf{w})$ then, a fully connected feedforward network $\boldsymbol{\zeta}_{\mathbf{H}|\mathbf{W}}(\mathbf{w}\ \boldsymbol{\theta}_2)$ with parameter $\boldsymbol{\theta}_2$ is constructed for modeling the deterministic mapping $\mathbf{w} \mapsto \boldsymbol{\zeta}_{\mathbf{H}|\mathbf{W}}(\mathbf{w})$. Then, for a given $\mathbf{w}$, the output of $\boldsymbol{\zeta}_{\mathbf{H}|\mathbf{W}}(\mathbf{w};\ \boldsymbol{\theta}_2)$ with parameter $\boldsymbol{\theta}_2$ is used in order to assemble matrix $[\log C_{\mathbf{H}|\mathbf{W}}(\mathbf{w};\ \boldsymbol{\theta}_2)]$. Then, matrix $[C_{\mathbf{H}|\mathbf{W}}(\mathbf{w};\ \boldsymbol{\theta}_2)]$ is calculated as the matrix exponential of $[\log C_{\mathbf{H}|\mathbf{W}}(\mathbf{w};\ \boldsymbol{\theta}_2)]$. We then defined as the statistical ANN-based metamodel, the probabilistic model in Section 3 where $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w};\ \boldsymbol{\theta}_1)$ and $[C_{\mathbf{H}|\mathbf{W}}(\mathbf{w};\ \boldsymbol{\theta}_2)]$ are used for modelling $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w})$ and $[C_{\mathbf{H}|\mathbf{W}}(\mathbf{w})]$. Consequently, the statistical ANN-based model is defined as the probability density function $\boldsymbol{\eta} \mapsto p_{\mathbf{H}|\mathbf{W}}(\boldsymbol{\eta}|\mathbf{w};\ \boldsymbol{\theta}_1,\ \boldsymbol{\theta}_2)$ that is a multivariate Gaussian probability density function with mean value $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w};\ \boldsymbol{\theta}_1)$ and covariance matrix $[C_{\mathbf{H}|\mathbf{W}}(\mathbf{w};\ \boldsymbol{\theta}_2)]$.

This ANN-based approach is actually a complement to the model presented in [18]. The previous model relies on conditional statistics based on the PLoM generation of the learned dataset requiring the training dataset. The proposed ANN-based model, once trained, can make predictions without referencing the learned dataset, at the price of introducing a probabilistic simplification concerning the probability distributions. This feature significantly reduces the computational resources required for deployment. In addition, the ANN-based model can be viewed as another type of representation of the surrogate model in [18].

4.2. Statistical ANN-based metamodel for regression with the learned dataset $\mathcal{D}^*_{\text{ar}}$

As it is usually the case for most regression problems, parameters $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are adjusted by fitting the statistical ANN-based metamodel to a suitable training dataset. In this paper, such fitting is carried out in minimizing with respect to $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ the negative-log-likelihood $\mathcal{L}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ of the statistical ANN-based model. Obviously, $\mathcal{D}^*_{\text{acm}}$ and $\mathcal{D}_{\text{acm}}$ are not suitable as training datasets for such fitting process since their sizes are small (less than one hundred elements). It is well known that, in the framework of statistical ANN-based modeling (as well for deterministic ANN-based modeling), small training datasets yields *overfitting* models that are unable to predict their targets well enough for unseen values of their inputs. On the other hand, since the size n_{ar} of learned dataset $\mathcal{D}^*_{\text{ar}}$ is as large as needed, such dataset is completely suitable as a training dataset. Using learned dataset $\mathcal{D}^*_{\text{ar}}$, the negative-log-likelihood to be minimized is written as

$$\mathcal{L}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = -\sum_{\ell=1}^{n_{\text{ar}}} \log\ p_{\mathbf{H},\mathbf{W}}(\boldsymbol{\eta}^\ell_{\text{ar}},\ \mathbf{w}^\ell_{\text{ar}};\ \boldsymbol{\theta}_1,\ \boldsymbol{\theta}_2) \quad , \tag{11}$$

$$= -\sum_{\ell=1}^{n_{\text{ar}}} \log\ \Big(p_{\mathbf{H}|\mathbf{W}}(\boldsymbol{\eta}^\ell_{\text{ar}}|\mathbf{w}^\ell_{\text{ar}};\ \boldsymbol{\theta}_1,\ \boldsymbol{\theta}_2)\, p_{\mathbf{W}}(\mathbf{w}^\ell_{\text{ar}}) \Big) \quad , \tag{12}$$

$$= -\sum_{\ell=1}^{n_{\text{ar}}} \log\ p_{\mathbf{H}|\mathbf{W}}(\boldsymbol{\eta}^\ell_{\text{ar}}|\mathbf{w}^\ell_{\text{ar}};\ \boldsymbol{\theta}_1,\ \boldsymbol{\theta}_2) - \sum_{\ell=1}^{n_{\text{ar}}} \log p_{\mathbf{W}}(\mathbf{w}^\ell_{\text{ar}}) \quad . \tag{13}$$

Minimizing $\mathcal{L}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ with respect to $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ is equivalent to minimizing the cost function $J(\boldsymbol{\theta}_1,\ \boldsymbol{\theta}_2)$ defined by

$$J(\boldsymbol{\theta}_1,\ \boldsymbol{\theta}_2) = \frac{1}{2}\sum_{\ell=1}^{n_{\text{ar}}} \log\ \Big(\det\left[C_{\mathbf{H}|\mathbf{W}}(\boldsymbol{\eta}^\ell_{\text{ar}};\ \boldsymbol{\theta}_2)\right] \Big) + \frac{1}{2}\sum_{\ell=1}^{n_{\text{ar}}} \big(\boldsymbol{\eta}^\ell_{\text{ar}} - \boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w}^\ell_{\text{ar}};\ \boldsymbol{\theta}_1)\big)^T [C_{\mathbf{H}|\mathbf{W}}(\mathbf{w}^\ell_{\text{ar}};\ \boldsymbol{\theta}_2)]^{-1} \big(\boldsymbol{\eta}^\ell_{\text{ar}} - \boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w}^\ell_{\text{ar}};\ \boldsymbol{\theta}_1)\big) . \tag{14}$$

Note that such minimization should be performed under the constraints defined by Eqs. (5) and (6). Classically, such constraints would be taken into account introducing Lagrange multipliers [20, 21, 22], augmented Lagrangian [21, 22], penalty methods [20, 21, 22], barrier methods [22], projected gradient methods [23] or Sequential Quadratic Programming [22]. It is not straightforward to implement such methods in the framework of fully connected feedforward neural network training when mini-batch are required due to constraints on RAM availability due to CPU or GPU limitations. In this paper and as explained in Section 4.3, we take advantage of the PLoM in order to fit the probabilistic model by adjusting $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ such that constraints defined by Eqs. (5) and (6) are automatically satisfied.

4.3. Statistical ANN-based metamodel for regression with a learned GKDE-based estimates dataset

It should be noted that minimizing $J(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ with respect to $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ under constraints defined by Eqs. (5) and (6) is equivalent to construct the likelihood-based statistical estimators of conditional mean and conditional covariance matrix of $\mathbf{H}$ given $\mathbf{W} = \mathbf{w}$. Such statistical estimators are different from the GKDE-based statistical estimators defined by Eqs. (2) and (3). Therefore, in the context of constructing $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w}; \boldsymbol{\theta}_1)$ and $\boldsymbol{\zeta}_{\mathbf{H}|\mathbf{W}}(\mathbf{w}; \boldsymbol{\theta}_2)$, an alternative strategy to the typical approach of minimizing the negative-log-likelihood $\mathcal{L}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ in regression problems, as presented in Section 4.2, is proposed. This alternative entails generating a GKDE-based estimates dataset $\mathcal{D}^*_{\mathbf{H}|\mathbf{W}}$ made up of pre-calculated estimates of conditional mean and conditional covariance of $\mathbf{H}$ given $\mathbf{W}$, using GKDE-based estimators as defined by Eqs. (2) and (3). Then, parameters $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are adjusted such that statistical ANN-based metamodel defined by $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w}; \boldsymbol{\theta}_1)$ and $\boldsymbol{\zeta}_{\mathbf{H}|\mathbf{W}}(\mathbf{w}; \boldsymbol{\theta}_2)$ fits dataset $\mathcal{D}^*_{\mathbf{H}|\mathbf{W}}$. Such a strategy requires a very large dataset $\mathcal{D}^*_{\text{ar}}$ in order to construct a large enough dataset $\mathcal{D}^*_{\mathbf{H}|\mathbf{W}}$ using GKDE-based estimators. Such an adapted very large dataset $\mathcal{D}^*_{\text{ar}}$ can easily be constructed by PLoM. Therefore, dataset $\mathcal{D}^*_{\mathbf{H}|\mathbf{W}}$ consists of $\nu < n_{\text{ar}}$ elements $(\boldsymbol{\mu}^1, \boldsymbol{\zeta}^1), \ldots, (\boldsymbol{\mu}^\nu, \boldsymbol{\zeta}^\nu)$ defined as follows. For all $j = 1, \ldots, \nu$,

$$\boldsymbol{\mu}^j = \frac{\sum_{\ell=\nu+1}^{n_{\text{ar}}} \boldsymbol{\eta}^\ell_{\text{ar}} \exp\left(-\frac{1}{2s^2}(\mathbf{w}^j_{\text{ar}} - \mathbf{w}^\ell_{\text{ar}})^T [C_{\mathbf{W}}]^{-1} (\mathbf{w}^j_{\text{ar}} - \mathbf{w}^\ell_{\text{ar}})\right)}{\sum_{\ell=\nu+1}^{n_{\text{ar}}} \exp\left(-\frac{1}{2s^2}(\mathbf{w}^j_{\text{ar}} - \mathbf{w}^\ell_{\text{ar}})^T [C_{\mathbf{W}}]^{-1} (\mathbf{w}^j_{\text{ar}} - \mathbf{w}^\ell_{\text{ar}})\right)} \quad , \tag{15}$$

and the $m(m+1)/2$ dimensional vector $\boldsymbol{\zeta}^j$ collects all the entries of the upper triangular part of the matrix logarithm $[\log C^j]$ of the matrix $[C^j]$, defined as

$$[C^j] = \frac{\sum_{\ell=\nu+1}^{n_{\text{ar}}} \widetilde{\boldsymbol{\eta}}^\ell_{\text{ar}}(\mathbf{w}^j_{\text{ar}}) (\widetilde{\boldsymbol{\eta}}^\ell_{\text{ar}}(\mathbf{w}^j_{\text{ar}}))^T \exp\left(-\frac{1}{2s^2}(\mathbf{w}^j_{\text{ar}} - \mathbf{w}^\ell_{\text{ar}})^T [C_{\mathbf{W}}]^{-1} (\mathbf{w}^j_{\text{ar}} - \mathbf{w}^\ell_{\text{ar}})\right)}{\sum_{\ell=\nu+1}^{n_{\text{ar}}} \exp\left(-\frac{1}{2s^2}(\mathbf{w}^j_{\text{ar}} - \mathbf{w}^\ell_{\text{ar}})^T [C_{\mathbf{W}}]^{-1} (\mathbf{w}^j_{\text{ar}} - \mathbf{w}^\ell_{\text{ar}})\right)} \quad . \tag{16}$$

Note that $n_{\text{ar}} - \nu$ is the number of realizations used for the GKDE-based estimates of conditional mean and conditional covariance of $\mathbf{H}$ given $\mathbf{W} = \mathbf{w}^j_{\text{ar}}$. Hence, the least-square estimation of parameters $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ is obtained as the parameters that minimize the cost function $\mathcal{J}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ that is written as

$$\mathcal{J}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \sum_{j=1}^{\nu} \|\boldsymbol{\mu}^j - \boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w}^j_{\text{ar}}; \boldsymbol{\theta}_1)\|^2 + \sum_{j=1}^{\nu} \|\boldsymbol{\zeta}^j - \boldsymbol{\zeta}_{\mathbf{H}|\mathbf{W}}(\mathbf{w}^j_{\text{ar}}; \boldsymbol{\theta}_2)\|^2 \quad , \tag{17}$$

This optimization problem is solved using ADAM (ADAptive Moment estimation) algorithm [24] with a learning rate γ scheduler that adjusts the learning rate over the course of training as

$$\gamma(k) = \max(\gamma_o \, \alpha^{(k-1)/\Delta}, \gamma_{\min}) \, , \tag{18}$$

where k is the epoch, γ_o is the initial learning rate (default value is $\gamma_o = 0.001$), $\alpha = 0.95$ is the decay factor, $\Delta = 5$ is the decay period, and $\gamma_{\min} = 1 \times 10^{-7}$ is the minimum leaning rate. For the ADAM algorithm, the parameters are fixed as $\beta_1 = 0.9, \beta_2 = 0.99$, and $\epsilon = 1 \times 10^{-8}$.

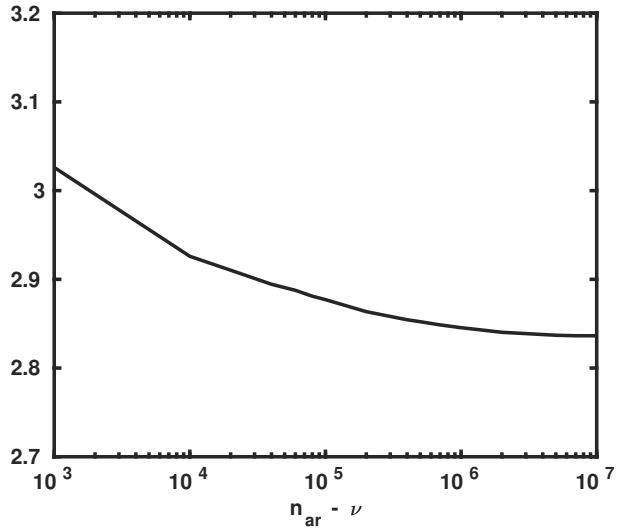


Figure 3: Statistical convergence analysis of the GKDE estimator for the conditional covariance matrix $[C_{\mathbf{H}|\mathbf{W}}(\mathbf{w})]$ of $\mathbf{H}$ given $\mathbf{W} = \mathbf{w}$. Graph of $n_{\text{ar}} \mapsto (n_{\text{ar}} - \nu + 1)^{-1} \sum_{j=\nu+1}^{n_{\text{ar}}} \| [C_{\mathbf{H}|\mathbf{W}}(\mathbf{w}_{\text{ar}}^j)] \|^2$. Horizontal axis: $n_{\text{ar}} - \nu$.

5. Numerical applications

5.1. Architecture of the statistical ANN-based metamodel

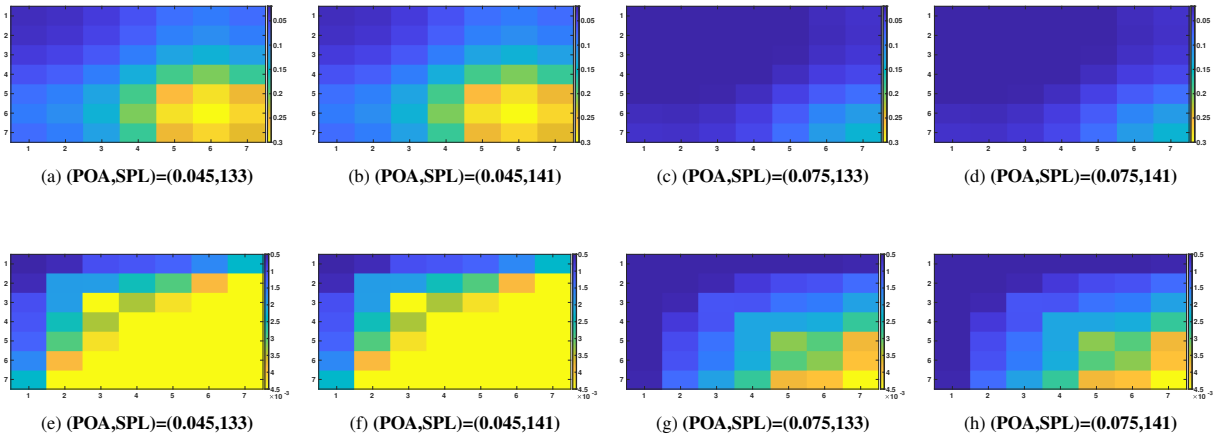
Concerning the architecture of the statistical ANN-based metamodel, rather than constructing two multi-outputs fully connected feedforward networks (one for $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w}; \boldsymbol{\theta}_1)$ and one for $\boldsymbol{\zeta}_{\mathbf{H}|\mathbf{W}}(\mathbf{w}; \boldsymbol{\theta}_2)$), we choose to construct m single output fully connected feedforward networks $\{\boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w}; \boldsymbol{\theta}_{1,1})\}_1, \ldots, \{\boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w}; \boldsymbol{\theta}_{1,m})\}_m$ and $d = m(m+1)/2$ single output fully connected feedforward networks $\{\boldsymbol{\zeta}_{\mathbf{H}|\mathbf{W}}(\mathbf{w}; \boldsymbol{\theta}_{2,1})\}_1, \ldots, \{\boldsymbol{\zeta}_{\mathbf{H}|\mathbf{W}}(\mathbf{w}; \boldsymbol{\theta}_{2,d})\}_d$, where $\boldsymbol{\theta}_1$ is rewritten as $\boldsymbol{\theta}_1 = (\boldsymbol{\theta}_{1,1}, \ldots, \boldsymbol{\theta}_{1,m})$ and $\boldsymbol{\theta}_2$ is rewritten as $\boldsymbol{\theta}_2 = (\boldsymbol{\theta}_{2,1}, \ldots, \boldsymbol{\theta}_{2,d})$. The training is carried out in parallel on a cluster of 3 *Tesla V100* GPUs. For each fully connected feedforward network, there are four hidden layers, the number of units is 20, 250, 75 and 25 respectively. This architecture implies a total number of 26 314 parameters (biases and weights). For the first and the fourth layer, Glorot [25] initialization is used and for the second and third layer, He [26] initialization is used. Rectified linear unit (ReLU) activation functions is used for each of the four hidden layers.

5.2. Statistical convergence analysis for the learned GKDE-based estimates dataset

A statistical convergence analysis is carried out with respect to number ν (see section 4.3). Figure 3 shows the graph of $n_{\text{ar}} \mapsto (n_{\text{ar}} - \nu + 1)^{-1} \sum_{j=\nu+1}^{n_{\text{ar}}} \| [C_{\mathbf{H}|\mathbf{W}}(\mathbf{w}_{\text{ar}}^j)] \|^2$ in which $\| \cdot \|$ is the Frobenius norm. The horizontal axis of Fig. 3 is the number $n_{\text{ar}} - \nu$ statistically independent realizations used in order to construct the GKDE-based estimates of conditional covariance of $\mathbf{H}$ given $\mathbf{W} = \mathbf{w}$ with $\nu = 60\,000$. It can be shown that convergence is reached for $n_{\text{ar}} - \nu = 200\,000$, that is to say $n_{\text{ar}} = 260\,000$.

*5.3. Conditional covariance matrices of **R** and **V** given **W***

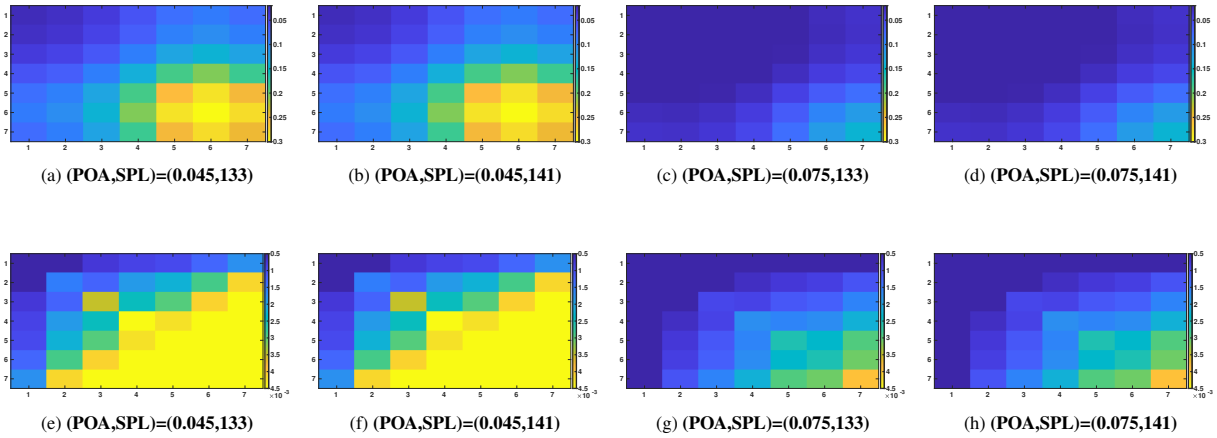
Using the database $\mathcal{D}_{\text{ar}}^*$ and conditional statistics based on GKDE estimation, the $(m \times m)$ conditional covariance matrix $[C_{\mathbf{Q}|\mathbf{W}}(\mathbf{w})]$ can be calculated. The expression is not included in this paper but is similar to Eq. (3) using $\mathbf{q}_{\text{ar}}^j$ instead of $\boldsymbol{\eta}_{\text{ar}}^j$ and with a Silverman bandwidth $n = n_{\mathbf{q}} + n_{\mathbf{w}}$ in Eq. (4). Using Eqs. (7) to (10) allow for calculating conditional covariance matrices of resistance $[C_{\mathbf{R}|\mathbf{W}}(\mathbf{w})]$ and reactance $[C_{\mathbf{V}|\mathbf{W}}(\mathbf{w})]$ for given $\mathbf{W} = \mathbf{w}$. Furthermore, as explained in Section 3.3 and given models $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w}; \boldsymbol{\theta}_1)$ and $\boldsymbol{\zeta}_{\mathbf{H}|\mathbf{W}}(\mathbf{w}; \boldsymbol{\theta}_2)$, conditional covariance matrix $[C_{\mathbf{R}|\mathbf{W}}(\mathbf{w}; \boldsymbol{\theta}_1, \boldsymbol{\theta}_2)]$ (resp. $[C_{\mathbf{V}|\mathbf{W}}(\mathbf{w}, \boldsymbol{\theta}_2)]$) of $\mathbf{R}(\mathbf{w})$ (resp. $\mathbf{V}(\mathbf{w})$) can be constructed for given $\mathbf{W} = \mathbf{w}$. The entries of matrices $[C_{\mathbf{R}|\mathbf{W}}(\mathbf{w})]$ and $[C_{\mathbf{V}|\mathbf{W}}(\mathbf{w})]$ (calculated by GKDE) are displayed in Fig. 4. The entries of matrices $[C_{\mathbf{R}|\mathbf{W}}(\mathbf{w}; \boldsymbol{\theta}_1, \boldsymbol{\theta}_2)]$ and $[C_{\mathbf{V}|\mathbf{W}}(\mathbf{w}; \boldsymbol{\theta}_2)]$ (calculated using the statistical ANN-based metamodel) are displayed in Fig. 5. The frequency-sampled resistance is displayed in Figs 4a to 4d and in 5a to 5d. The frequency-sampled reactance is



(a) **(POA,SPL)=(0.045,133)** (b) **(POA,SPL)=(0.045,141)** (c) **(POA,SPL)=(0.075,133)** (d) **(POA,SPL)=(0.075,141)**

(e) **(POA,SPL)=(0.045,133)** (f) **(POA,SPL)=(0.045,141)** (g) **(POA,SPL)=(0.075,133)** (h) **(POA,SPL)=(0.075,141)**

Figure 4: Conditional covariance matrices with GKDE-based estimation of the conditional covariance of resistance (Figs.4a to 4d) and reactance (Figs. 4e to 4h) given four different values of $\mathbf{w}$



(a) **(POA,SPL)=(0.045,133)** (b) **(POA,SPL)=(0.045,141)** (c) **(POA,SPL)=(0.075,133)** (d) **(POA,SPL)=(0.075,141)**

(e) **(POA,SPL)=(0.045,133)** (f) **(POA,SPL)=(0.045,141)** (g) **(POA,SPL)=(0.075,133)** (h) **(POA,SPL)=(0.075,141)**

Figure 5: Conditional covariance matrices with statistical ANN-based metamodel estimation of the conditional covariance of resistance (Figs.5a to 5d) and reactance (Figs. 5e to 5h) given four different values of $\mathbf{w}$

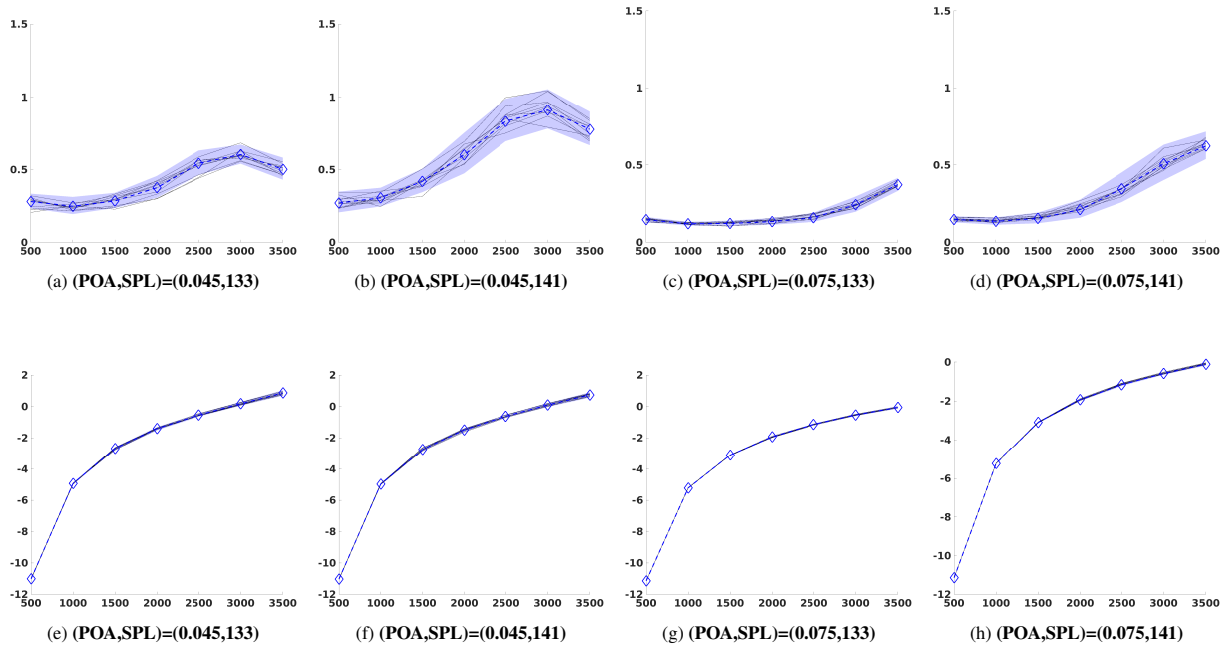
displayed in Figs 4e to 4h and 5e to 5h. It should be noted that the considered values of the control parameters are indicated in each sub-figure. It can be seen that the matrices are not diagonal and that a correlation exits between the different frequency points of the frequency-sampled impedance. It can also be seen that GKDE-based estimate in Fig. 4 and statistical ANN-based metamodel estimate in Fig. 5 of the covariance matrices are quantitatively the same.

5.4. Frequency-sampled impedance using the statistical ANN-based metamodel

For any value of $\mathbf{w}$ in its admissible set, the conditional statistics (mean values and confidence regions) of $\mathbf{R}(\mathbf{w})$ and $\mathbf{V}(\mathbf{w})$ are estimated using the statistically independent realizations $\mathbf{r}^1(\mathbf{w};\, \boldsymbol{\theta}_1,\, \boldsymbol{\theta}_2)$, $\ldots$, $\mathbf{r}^N(\mathbf{w};\, \boldsymbol{\theta}_1,\, \boldsymbol{\theta}_2)$ and $\mathbf{v}^1(\mathbf{w};\, \boldsymbol{\theta}_1,\, \boldsymbol{\theta}_2), \ldots, \mathbf{v}^N(\mathbf{w};\, \boldsymbol{\theta}_1,\, \boldsymbol{\theta}_2)$ as presented in Section 3.3 and obtained using $\boldsymbol{\mu}_{\mathbf{H}|\mathbf{W}}(\mathbf{w};\, \boldsymbol{\theta}_1)$ and $\boldsymbol{\zeta}_{\mathbf{H}|\mathbf{W}}(\mathbf{w};\, \boldsymbol{\theta}_2)$ as explained in Section 4.1.

Figures 6 shows 10 realizations of frequency-sampled resistances (Figs 6a to 6d) and 10 corresponding realizations of frequency-sampled reactance (the values of the control parameters are indicated in each sub-figure). The dashed blue line (resp. the blue domain) is the conditional mean value (resp. the conditional confidence region) of the frequency-sampled resistance (Figs. 6a to 6d) and reactance (Figs. 6a to 6d).

The conditional mean value and the conditional confidence region with probability level $P_c = 98\%$, of the frequency-sampled impedance given $\mathbf{W} = \mathbf{w}$, corresponding to the statistical ANN-based metamodel, are presented



(a) **(POA,SPL)=(0.045,133)** (b) **(POA,SPL)=(0.045,141)** (c) **(POA,SPL)=(0.075,133)** (d) **(POA,SPL)=(0.075,141)**

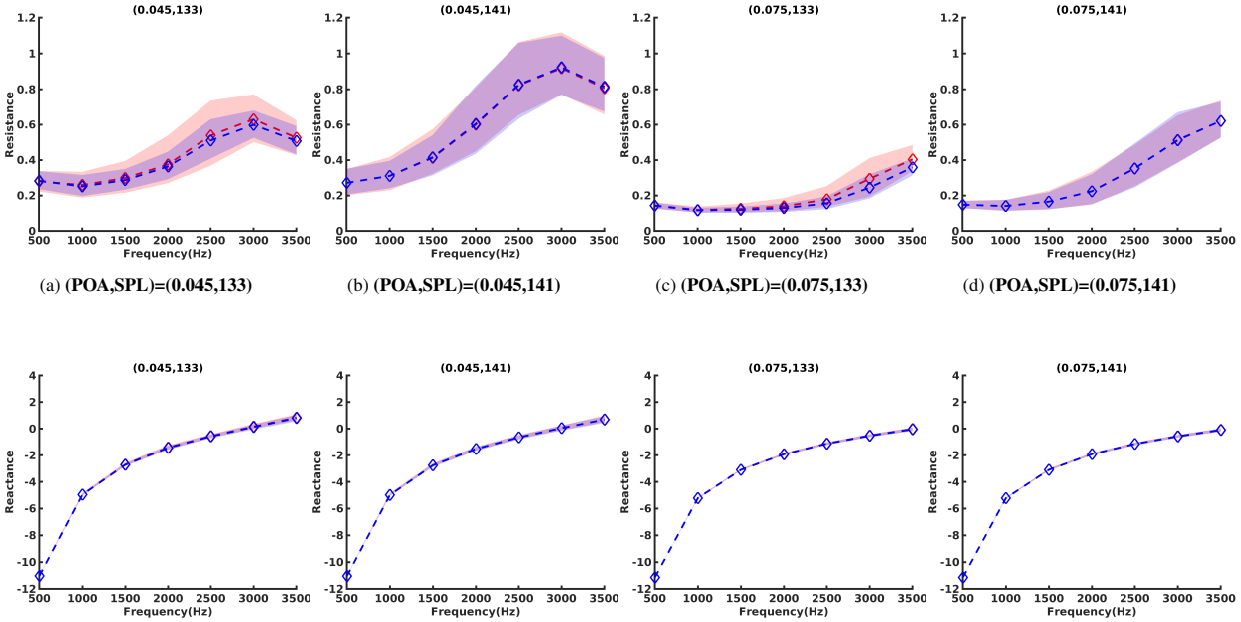
(e) **(POA,SPL)=(0.045,133)** (f) **(POA,SPL)=(0.045,141)** (g) **(POA,SPL)=(0.075,133)** (h) **(POA,SPL)=(0.075,141)**

Figure 6: Random generations of frequency-sampled impedance using the statistical ANN-based metamodel. Conditional mean value (dashed blue line) and conditional confidence region (blue region) with a probability level $P_c = 98\%$ for the resistance ((a) to (d)) and reactance ((e) to (h)). The 10 thin black lines represent realizations of the frequency-sampled impedance generated by the statistical ANN-based metamodel for given $\mathbf{w}$. Horizontal axis is the frequency in Hz

in Figs. 7 and are compared with results presented in [18] (GKDE based estimates). Figures 7a to 7d correspond to the resistance and Figures 7e to 7h correspond to the reactance (the values of the control parameters are indicated in each sub-figure). The red dashed line (resp. the red domain) shows the conditional mean value (resp. the confidence region) from [18]. The blue dashed line (resp. the blue domain) shows the conditional mean value (resp. confidence region) from the statistical ANN-based metamodel. It can be seen that, concerning the conditional mean values of the frequency-sampled impedance, there is a good match between the statistical ANN-based metamodel presented in this paper and the previous results presented in [18]. The statistical ANN-based metamodel, which has been fitted on the learned GKDE-based estimates dataset $\mathcal{D}^*_{\mathbf{H}|\mathbf{W}}$ (with 60 000 realizations), shows conditional confidence region that are quantitatively the same as those calculated by GKDE-based estimate using the learned ACM dataset $\mathcal{D}^*_{\text{ar}}$ (260 000 realizations) with manually adjusted Silverman bandwidth s to the value 0.1186.

6. Conclusions and perspectives

In this paper, a statistical ANN-based metamodel of the frequency-sampled liner acoustic impedance has been presented, for which only a small ACM data is available to fit its parameters. The control parameters are the POA and SPL with Mach number being kept at zero. The latent (unobserved) parameters introduce uncertainties that are taken into account by a probabilistic model introduced in [18]. The construction of the statistical metamodel uses a PCA to construct a statistical reduced representation of the frequency-sampled vector of the random log-resistance and the random reactance. For fitting the statistical ANN-based metamodel, a big training dataset is constructed using probabilistic learning on manifolds (PLoM). A *prior* conditional probability distribution of the reduced representation given the control parameters is then constructed and assumed to be Gaussian, yielding a multivariate log-normal distribution for the resistance and a multivariate Gaussian distribution for the reactance. The metamodels of the hyperparameters of such conditional probabilistic model is presented as fully-connected feedforward neural networks. As the reduced representation is modeled by a centered and normalized random vector, some constraints have to be taken into account in the minimization of the negative-log likelihood when fitting the parameters (biases and weights) of the neural network. The constrained problem is transformed in an unconstrained one, requiring the construction of



(a) **(POA,SPL)=(0.045,133)** (b) **(POA,SPL)=(0.045,141)** (c) **(POA,SPL)=(0.075,133)** (d) **(POA,SPL)=(0.075,141)**

(e) **(POA,SPL)=(0.045,133)** (f) **(POA,SPL)=(0.045,141)** (g) **(POA,SPL)=(0.075,133)** (h) **(POA,SPL)=(0.075,141)**

Figure 7: Conditional mean values and conditional confidence region with a probability level $P_c = 98\%$ of the frequency-sampled impedance, for the resistance ((a) to (d)) and reactance ((e) to (h)). The red dashed curve and the red zone represent the conditional mean values and conditional confidence intervals estimated from [18] and the Silverman bandwidth fixed at $s = 0.1186$ for all the figures. The blue dashed curve and the blue zone represent the conditional expectation and conditional confidence intervals estimated using the statistical ANN-based metamodel of liner impedance. Horizontal axis is frequency in Hz.

a second training dataset to estimate the conditional mean vectors and conditional covariance matrices for which the learned realizations are generated using PLoM. The novelty of this paper lies in the methodology used to construct a statistical ANN-based metamodel of liner acoustic impedance, which can be used as a low-computational cost metamodel to predict the confidence interval and the mean value of the impedance given any value of the control parameter. The gradients of the mean values and confidence regions can also easily be derived using classical backpropagation algorithms for very cheap computational cost. The statistical ANN-based metamodel presented in this paper is a complement to the model proposed in [18]. Firstly, it allows for offline use, as it does not require access to the learned dataset for making predictions. Secondly, the gradient computations with respect to input parameters are expected to be computationally cheaper, facilitating its use in optimization problems. These improvements, combined with the ANN ability to capture complex, multidimensional relationships, make this new model a flexible and efficient tool for optimizing liner acoustic impedance across the parameter space. Such work can be extended to non-zero Mach in the ACM dataset.

7. Acknowledgments

This work was supported by DGAC (Direction Générale de l'Aviation Civile), by PNRR (Plan National de Relance et de Résilience Français) and by NextGeneration EU via the project MAMBO (Méthodes Avancées pour la Modélisation du Bruit moteur et aviOn).

References

[1] B. van Den Nieuwenhof, Y. Detandt, G. Lielens, E. Rosseel, C. Soize, V. Dangla, M. Kassem, A. Mosson, Optimal design of the acoustic treatments damping the noise radiated by a turbo-fan engine, in: 23rd AIAA/CEAS Aeroacoustics Conference, 2017, p. 4035. doi:10.2514/6.2017-4035.

[2] D. M. Nark, M. G. Jones, Design of an advanced inlet liner for the quiet technology demonstrator 3, in: 25th AIAA/CEAS Aeroacoustics Conference, 2019, p. 2764. doi:10.2514/6.2019-2764.

[3] D. L. Sutliff, D. M. Nark, M. G. Jones, N. H. Schiller, Design and acoustic efficacy of a broadband liner for the inlet of the dgen aero-propulsion research turbofan, in: 25th AIAA/CEAS Aeroacoustics Conference, 2019, p. 2582. doi:10.2514/6.2019-2582.
[4] A. T. Chambers, J. M. Manimala, M. G. Jones, Design and optimization of 3d folded-core acoustic liners for enhanced low-frequency performance, AIAA Journal 58 (1) (2020) 206–218. doi:10.2514/1.J058017.
[5] E. Özkaya, N. R. Gauger, J. A. Hay, F. Thiele, Efficient design optimization of acoustic liners for engine noise reduction, AIAA Journal 58 (3) (2020) 1140–1156. doi:10.2514/1.J057776.
[6] V. Dangla, C. Soize, G. Cunha, A. Mosson, M. Kassem, B. Van den Nieuwenhof, Robust three-dimensional acoustic performance probabilistic model for nacelle liners, AIAA Journal 59 (10) (2021) 4195–4211. doi:10.2514/1.J060299.
[7] A. M. Spillere, D. S. Braga, L. A. Seki, L. A. Bonomo, J. A. Cordioli, B. M. Rocamora Jr, P. C. Greco Jr, D. C. dos Reis, E. L. Coelho, Design of a single degree of freedom acoustic liner for a fan noise test rig, International Journal of Aeroacoustics 20 (5-7) (2021) 708–736. doi:10.1177/1475472X211023831.
[8] M. Lavieille, T. Abboud, A. Bennani, N. Balin, Numerical simulations of perforate liners: Part i - model description and impedance validation, in: 19th AIAA/CEAS Aeroacoustics Conference, 2013. doi:10.2514/6.2013-2269.
[9] B. Van Antwerpen, Y. Detandt, D. Copiello, E. Rosseel, E. Gaudry, Performance improvements and new solution strategies of Actran/TM for nacelle simulations, in: 20th AIAA/CEAS Aeroacoustics Conference, 2014, p. 2315. doi:10.2514/6.2014-2315.
[10] L. Pascal, E. Piot, G. Casalis, A new implementation of the extended helmholtz resonator acoustic liner impedance model in time domain caa, Journal of Computational Acoustics 24 (01) (2016) 1663–1674. doi:10.1142/S0218396X15500150.
[11] L. Casadei, H. Deniau, E. Piot, T. Node-Langlois, Time-domain impedance boundary condition implementation in a cfd solver and validation against experimental data of acoustical liners, eForum Acusticum (2020) 359–366doi:10.48465/fa.2020.0088.
[12] V. Dangla, C. Soize, G. Cunha, A. Mosson, M. Kassem, B. Van Den Nieuwenhof, Stochastic computational model of 3d acoustic noise predictions for nacelle liners, in: AIAA Aviation 2020 Forum, 2020, p. 2545. doi:10.2514/6.2020-2545.
[13] J. Winkler, J. M. Mendoza, C. A. Reimann, K. Homma, J. S. Alonso, High fidelity modeling tools for engine liner design and screening of advanced concepts, International Journal of Aeroacoustics 20 (5-7) (2021) 530–560. doi:10.1177/1475472X211023884.
[14] C. Soize, R. Ghanem, Data-driven probability concentration and sampling on manifold, Journal of Computational Physics 321 (2016) 242–258. doi:10.1016/j.jcp.2016.05.044.
[15] C. Soize, R. Ghanem, Probabilistic learning on manifolds, Foundations of Data Science 2 (3) (2020) 279–307. doi:10.3934/fods.2020013.
[16] C. Soize, R. Ghanem, Probabilistic learning on manifolds (PLoM) with partition, International Journal for Numerical Methods in Engineering 123 (1) (2022) 268–290. doi:10.1002/nme.6856.
[17] C. Soize, Software_PLoM_with_partition_2021_06_24 (Jun. 2021).
URL https://hal-upec-upem.archives-ouvertes.fr/hal-03275052
[18] A. Sinha, C. Soize, C. Desceliers, G. Cunha, Aeroacoustic liner impedance metamodel from simulation and experimental data using probabilistic learning, AIAA Journal 61 (11) (2023) 4926–4934. doi:10.2514/1.J062991.
[19] C. Soize, Uncertainty quantification, Springer, 2017. doi:10.1007/978-3-319-54339-0.
[20] D. P. Bertsekas, Constrained Optimization and Lagrange Multiplier Methods, Academic Press, 1982.
[21] D. Luenberger, Optimization by Vector Space Methods, John Wiley & Sons, 1997.
[22] J. Nocedal, S. Wright, Numerical Optimization, Springer, 2006. doi:10.1007/978-0-387-40065-5.
[23] P. H. Calamai, J. J. Moré, Projected gradient methods for linearly constrained problems, Mathematical programming 39 (1) (1987) 93–116. doi:10.1007/BF02592073.
[24] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization (2017). arXiv:1412.6980.
[25] X. Glorot, Y. Bengio, Understanding the difficulty of training deep feedforward neural networks, in: Y. W. Teh, M. Titterington (Eds.), Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Vol. 9 of Proceedings of Machine Learning Research, PMLR, Chia Laguna Resort, Sardinia, Italy, 2010, pp. 249–256.
[26] K. He, X. Zhang, S. Ren, J. Sun, Delving deep into rectifiers: surpassing human-level performance on imagenet classification, in: 2015 IEEE International Conference on Computer Vision (ICCV), 2015, pp. 1026–1034. doi:10.1109/ICCV.2015.123.