Gaussian process based surrogate modelling of acoustic systems

Thomas Kohlsche[1,*]**, Stephan Lippert**[1]**,** and **Otto von Estorff**[1]

[1] Institute of Modelling and Computation, Hamburg University of Technology

The numerical simulation of acoustic problems is, for itself, a quite difficult task since the underlying systems are usually highly complex with a broad frequency range and high sensitivity. Due to this complexity and the corresponding computational burden, tasks like optimization and uncertainty quantification (UQ) are seldom performed in acoustics. Especially when dealing with polymorphic uncertainties where combined techniques of UQ might be required, a direct use of the model is not viable. To allow such engineering tasks, the construction of a cheap surrogate or reduced model is common practice in order to allow a large number of model evaluations at low costs.
For acoustic systems, the construction of a reasonably accurate surrogate model can become a challenging task since many systems operate in the frequency domain where phenomena like resonance and interference can cause highly nonlinear responses. In this paper, a surrogate model based on the combination of a parametric model for capturing local nonlinearities and a random process regression for modelling the global trend is presented. The basic procedure based on previous works of the authors is extended to a predict the system response both for unobserved parameters and frequencies. The procedure is demonstrated for a representative example namely the acoustic simulation of a car interior, and further improvements in accuracy and efficiency by the usage of multilevel information are discussed.

1 Introduction

Within the last decades, acoustic properties have become an important design criterion for a large variety of structural design problems. On the one hand, law regulation become more strict concerning noise pollution. On the other hand, acoustic properties become more and more an aspect of comfort. The structural design nowadays is mostly done in a numerical way where the acoustic part is a challenging task since relevant systems are, from a geometrical point, of comparably large size and a broad frequency range has to be considered. The distinct frequency dependence can also lead to high sensitivities in the acoustic response of a structural system. A small variation in the geometry or a material parameter can lead to a shift in the eigenfrequencies which can cause resonances that were not foreseen by the underlying deterministic model. Such effects demonstrate the need to incorporate uncertainties in the numerical design of structures with respect to their acoustical properties.

The uncertainty quantification (UQ) in acoustic systems forms a relatively new field of application in the UQ community. Soize [1], among others, demonstrated a non-parametric approach by using randomized system matrizes for a vibro-acoustic model while Dammak et al. [2] used spectral stochastic methods to model parametric uncertainties in the acoustics of a vehicle cabin. Further applications can be found in the modelling of acoustic wave propagation in materials with spatially varying uncertain material properties [3] or acoustic scattering for obstacles with uncertain shapes [4]. These few examples of existing works are representative for most works on this field. They primarily try so solve one particular uncertain scenario and mostly use intrusive techniques, tailored for the situation but without providing a generally applicable technique.

A more general approach to handle parametric uncertainties is the construction of a surrogate model which, once generated, can be used for most common UQ technique. Unfortunately, most surrogate techniques are based on the assumption of a certain smoothness in the response. This assumption can be completely wrong when frequency depending responses are modelled. It has been shown in multiple works that well established approaches like the polynomial chaos expansion show poor convergence for frequency responses close to eigenfrequencies, i.e. when resonances occur [5].

The authors proposed in [6] a methodology based on universal kriging that uses local parametric models to represent local nonlinear responses combined with a Gaussian process (GP) regression to model the global trends. The procedure showed promising results for the examined test problem where the response depending on some uncertain parameter has been modelled for each frequency of interest individually.

The construction of the surrogate for fixed frequencies has the drawback that possibly correlated effects of the parameters and the frequencies in the response can not be captured. As an example, a variation of a parameter may lead to a shift of the eigenfrequencies. When the surrogate is constructed in parameter space for a fixed set of frequencies, these shifts in frequency direction are not learned by the surrogate. Therefore, a high sample resolution in frequency direction is required to ensure that all relevant aspects in the response are captured. In the following, the approach from [6] is extended to learn also the response in frequency direction to reduce the number of required sampling points and to provide response predictions for unobserved frequencies.

---

* Corresponding author: e-mail mub@tuhh.de, phone +49 40 428 78 30 70, fax +49 40 428 78 43 53

PAMM · Proc. Appl. Math. Mech. 2019;**19**:e201900471.   www.gamm-proceedings.com   **1 of 4**

https://doi.org/10.1002/pamm.201900471   © 2019 The Authors *Proceedings in Applied Mathematics & Mechanics* published by Wiley-VCH Verlag GmbH & Co. KGaA Weinheim

2 Gaussian process regression

General GP formulation

The usage of random processes for regression is most prominent in geostatistics where it is well known under the name kriging. Within the last decades, random processes, and especially GPs as a special case, have experienced a significant success as computer learning techniques for multiple fields of application like regression or classification. Rasmussen and Williams in their well-known work [7] provide an introduction of GP for machine learning whereas Forrester et al. [8] provide a practical overview of GPs for surrogate modelling in engineering applications.

GP regression is a non-parametric, data-driven approach to learn a functional mapping $\boldsymbol{y} = \mathcal{M}(\boldsymbol{x})$ between an input space $\boldsymbol{x} \in \mathbb{R}^d$ and the output space $\boldsymbol{y}$ from a given set of known data points $\boldsymbol{X} = [X_1 \dots X_N]$ and $\boldsymbol{Y} = \mathcal{M}(\boldsymbol{X}) = [Y_1 \dots Y_N]$. For the problem discussed in this paper, the mapping $\mathcal{M}$ can be any numerical model, $\boldsymbol{x}$ is a set of input parameters and $\boldsymbol{y}$ is the model output. The data is usually gathered by running the model for a number of sampling points, chosen by a design of experience. Given the data set, the GP regression assumes the observed points to be realizations of a random process with Gaussian distribution,

$$\mathcal{M}(\boldsymbol{X}) \approx \mathcal{G}\left(\mu(\boldsymbol{x}), \sigma^2 k(\boldsymbol{x}, \boldsymbol{x}^*)\right), \tag{1}$$

where $\mu(\boldsymbol{x})$ is the underlying mean function, σ^2 is the standard deviation and $k(\boldsymbol{x}, \boldsymbol{x}^*)$ depicts the correlation function, prescribing the correlation between two points $(\boldsymbol{x}, \boldsymbol{x}^*)$. In the general case, a Gaussian kernel defining the correlation and a parametric model modelling the mean are used,

$$k(x, x^*) = \exp\left(-\frac{1}{2}\left(\frac{x - x^*}{\theta}\right)^2\right) \qquad \text{and} \qquad \mu(\boldsymbol{x}) = \sum_i^m f_i(\boldsymbol{x})\beta_i = \boldsymbol{f}(\boldsymbol{x})\boldsymbol{\beta}. \tag{2}$$

Since the GP does not use any parametric functions that are fitted to the data, it is referred to as a non-parametric approach. It nevertheless contains a set of so-called hyper parameters σ^2, θ which defines the correlation length, and the weights $\boldsymbol{\beta}$ of a set of m basis functions $\boldsymbol{f}(\boldsymbol{x})$. For a d-dimensional input, θ becomes a d-by-d matrix. This leads to the total set of hyper parameters $[\boldsymbol{\beta}, \sigma^2, \boldsymbol{\theta}]$ that have to be determined such that the resulting GP is most likely to be responsible for the observed realizations. Taking the known data as a finite set of realizations of a random process results in a multivariate Gaussian distribution,

$$\boldsymbol{Y} \quad \sim \mathcal{N}(\boldsymbol{\mu}, \sigma^2 \boldsymbol{\Sigma}), \tag{3}$$

$$\text{with} \quad \mu_i = \mu(X_i), \qquad i = 1 \dots N \qquad \text{and} \qquad \Sigma_{i,j} = k(X_i, X_j), \qquad i, j = 1 \dots N. \tag{4}$$

The set of hyper parameters that fits best to the observed data can be determined by a bayesian approach or by maximum likelihood estimation (MLE) [9]. For the more common MLE, the so called concentrated log-likelihood has to be maximized, e.g., by a genetic optimization algorithm. Once the hyper parameters are determined, the mean and covariance of the posterior distribution of the response $\boldsymbol{Y}_*$ for a set of unobserved points $\boldsymbol{X}_*$ can be calculated by

$$\boldsymbol{\mu}^* = \mu(\boldsymbol{X}_*) + \boldsymbol{\Sigma}_*^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{Y} - \mu(\boldsymbol{X})) \qquad \text{and} \qquad \boldsymbol{\Sigma}^* = \boldsymbol{\Sigma}_{**} - \boldsymbol{\Sigma}_*^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_*, \tag{5}$$

with $\boldsymbol{\Sigma}_* = k(\boldsymbol{X}, \boldsymbol{X}_*)$ and $\boldsymbol{\Sigma}_{**} = k(\boldsymbol{X}_*, \boldsymbol{X}_*)$. The mean value can then be used as a prediction for the unobserved points while the covariance can be used to determine confidence intervals.

Semi-parametric model for systems with nonlinear responses

The characteristics of the GP regression depend mainly on the choice of the kernel function. If a smooth kernel is chosen, the resulting model will be smooth while a non smooth kernel will lead to a non smooth model. For the systems described in the introduction, a global smooth behaviour with local nonlinear and possibly peak-like, non smooth responses is possible. For such a scenario, neither a smooth nor a non smooth kernel will give good results. To handle this problem, the semi-parametric properties of a GP with a parametric mean function can be used. The basis functions for the mean $\boldsymbol{f}(x)$ can be chosen such that they are well suited to model the few local nonlinearities adequately while the globally smooth response is captured by the GP with a smooth kernel function. The challenge is then to find suitable basis functions. In [6], the authors proposed to use models of multiple fidelity levels to perform a grid search in order to locate the critical areas. The different levels can be, e. g., generated by coarsening the spatial discretization of the model. The low level information is cheap to obtain which allows a grid search in the parameter space with a high resolution. The results may not be accurate but give a rough estimate of the critical areas. By successively refining the model discretization and simultaneously isolating the critical areas, only a comparatively small number of evaluations around possible peaks have to be evaluated at the highest level of fidelity. The data can then be used to generate local models as basis functions for the GP regression.

3 Acoustic application case for a surrogate model

Test case model settings

As a realistic test scenario, the interior acoustics of a simplified car cabin model is simulated using the boundary element method. The system is excited by a Neumann boundary condition simulating a loudspeaker and the sound pressure level (SPL) at the drivers position is evaluated as the quantity of interest over a frequency range from 50 to 550 Hz. The wall impedance Z of the sidewalls is taken as the variable parameter for which the surrogate should be constructed. A low fidelity model with roughly 5% of the full models number of degrees of freedom is used for the localization of critical areas in the response. Four negative peaks in the response can be found in the examined parameter-frequency-range due to negative interference in the closed cavity of the car. For each peak, a basis function of the shape

$$f_i(F, Z) = -\log\left(\left[a_0(F - \bar{F}_i)^2 + a_1(Z - \bar{Z}_i)^2\right]^{\frac{1}{a_2}}\right) \tag{6}$$

is fitted to the data with $\bar{F}_i$ and $\bar{Z}_i$ being the approximate frequency and parameter value at which the i-th peak occurs and $[a_0, a_1, a_2]$ being free parameters. A reference solution is generated to compute the errors of the surrogates. For the given modelling scenario, a parameter-only model (P-model) is generated first at fixed frequencies in steps of 5 Hz. At each modelled frequency, 12 equidistant sampling points are taken in parameter direction to create the surrogate. At each frequency, a new GP regression with a Gaussian kernel function is performed to predict the response in parameter direction. The P-model is compared to a parameter-frequency-surrogate (PF-model).

Constructing the surrogate in both parameter and frequency direction at the same time has some significant advantages and drawbacks at the same time. A PF-model requires less sampling points since the resolution in frequency direction can be reduced. Furthermore, predictions for unobserved frequencies can be made which also makes it possible to consider variations in the frequencies of the excitations. As a downside, using one global model with only one set of hyper parameters for the whole domain is less flexible to react to locally varying response characteristics. Furthermore, the size of the data for the GP increases drastically which can lead to ill-posed covariance matrices. The PF-model is generated from a structured grid with the same 12 parameter sampling points for each frequency, but evaluated at less frequencies than the P-model which leads to a reduction of the computational effort. A Gaussian kernel with two independent correlation lengths $[\theta_Z, \theta_F]$ is used.

Surrogate results

To compare the quality of the surrogates, the root mean squared error (RMSE) and the maximum error, both absolute and relative, are calculated. The RMSE provides an indicator for the global approximation while the maximum deviation indicates a local mismatch, especially near the peaks. In table 1, the results for the P-model compared to the PF-model with different frequency step sizes are shown. Four different scenarios are shown: taking the same number of sampling points as the P-model and taking $\frac{1}{2}$ / $\frac{1}{3}$ / $\frac{1}{5}$ of the points. The number of sampling points is directly proportional to the primary computational costs resulting from the evaluations of the full model. There are three primary observations that need to be discussed. First, some PF-models seem to outperform the theoretically more flexible and robust P-model, even with far less data used for the model generation. The explanation can be found in fig. 1 and 2. For the given sampling, the P-model has trouble to capture the smooth regions for low impedances. The PF-model can also use the information from nearby points in frequency direction which allows a better representation of the smooth regions with, in total, less data points. Nevertheless, the PF-models perform worse in regions of rapid changes for high impedances which demonstrates the downsides.

As a second point, a look at the maximum deviations and the peaks is important. Due to the basis functions, all models independent of the sampling size capture the nonlinear responses but show maximum deviations of more than 7 % in any model. In the context of a good surrogate, errors of this magnitude are usually unacceptable but for practical applications of the surrogate like UQ or optimization, the modeller is usually not interested in the exact values near such nonlinear responses. These points represent potentially unstable regions which are very important to be qualitatively included in the model to warn the modeller against large uncertainties and non robust solutions but the exact values do rarely matter.

Last, the effect of different sampling sizes have to be discussed. It shows that building a PF-model with the same sampling density as the P-model results in a rather poor result. If the size of the training data is too large, the GP can run into trouble finding an optimal set of hyper parameters. This effect is comparable to overfitting phenomena for classical interpolation techniques. Taking too few sampling meanwhile leads to the GP missing important information in regions of rapid changes. All in all, the results show that a GP surrogate is basically suitable to predict both in parameter and frequency space but such a

Table 1: RMSE and maximum errors for the different surrogate models.

	P-model	PF-model			
		$\frac{1}{1}$ samples	$\frac{1}{2}$ samples	$\frac{1}{3}$ samples	$\frac{1}{5}$ samples
RMSE (abs. in dB / rel. in %)	1.07 / 1.01	0.89 / 0.87	0.67 / 0.62	0.89 / 0.76	1.77 / 1.50
max. error (abs. in dB / rel. in %)	7.98 / 7.52	12.47 / 15.64	5.81 / 7.29	8.21 / 8.43	11.17 / 11.47

PF-model is far more prone to misinterpretations of the data. Without a reference solution available as in the shown test case, forecasting whether the predictions are good and the sampling is chosen well is a tough task.

One possible way to reduce this problems is to use multi-fidelity information to improve the model.
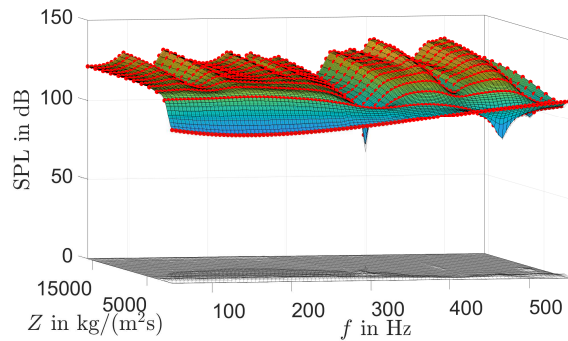

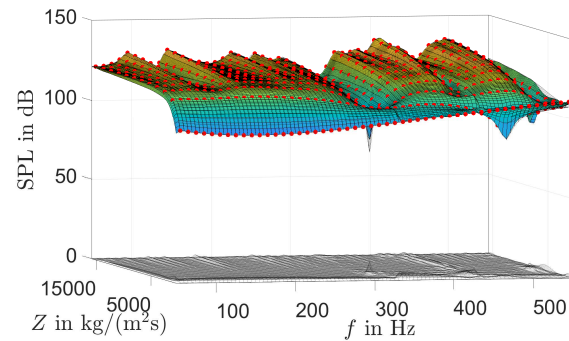**Fig. 1:** P-model response with sampling points (red) and local error (grey).

Fig. 2: PF-model ($\frac{1}{2}$) response with sampling points (red) and local error (grey).

Integration of multi-fidelity information

Using low fidelity information to learn the high fidelity response of a complex system has gained a lot attention within the last decade. Le Gratiet [9] introduced multiple ways to include information from different levels in a GP while Bieler et al. [10] showed the technique for complex UQ scenarios. The idea is to build the GP $\mathcal{G}^{(i)}$ at fidelity level $i = 1 \ldots s$, where s is the highest level, from a GP constructed for the level below, $\mathcal{G}^{(i-1)}$, plus an additional GP $\delta^{(i)}$ modelling the discrepancy between the levels,

$$\mathcal{G}^{(i)}(x) = \rho^{(i-1)} \cdot \mathcal{G}^{(i-1)}(x) + \delta^{(i)}(x), \qquad i = 2 \ldots s. \tag{7}$$

The scaling factor $\rho^{(i-1)}$ is usually assumed to be a constant and has to be learned as an additional hyper parameter. The modelling scenario described in this paper is somehow predestined to include multi-fidelity information since they are already available from the localization procedure of the critical areas. Building a multi-fidelity model in a straight forward way does indeed improve the global quality of the surrogate but unfortunately it corrupts the approximation of the local peaks. The critical regions move slightly with different levels of fidelity while the global trends are basically the same. This leads to the GP on highest level receiving misleading information from the lower levels. Due to the fact that the low level information still gives a very good basis almost everywhere, it is weighted much higher than any basis function approximating the peaks, which are only valid close to the critical regions but nowhere else. In a certain way, the low level information suppress the basis functions used before to capture the peaks. The multi-fidelity approach is still promising but a complex, locally changing scaling $\rho^{(i-1)}(x)$ is required that is able to suppress the low level information near the critical areas.

4   Conclusion

In this paper, the general applicability of GP regressions for the construction of surrogates for acoustic responses has been demonstrated and their ability to capture the distinctive frequency response characteristic has been shown. To improve the utility and efficiency of the approach, an inclusion of the frequency in the surrogate approximation has been introduced and demonstrated. Ongoing research currently tries to include multi fidelity information to construct efficient and reliable surrogates for frequency depending responses.

References

[1] C. Soize, Uncertainty Quantication - An Accelerated Course with Advanced Applications in Computational Engineering, Interdisciplinary Applied Mathematics, Vol. 47 (Springer, 2017).
[2] K. Dammak, S. Koubaa, A. E. Hami, L. Walha, and M. Haddar, Applied Acoustics **144**, 113–123 (2019).
[3] S. Mishra, C. Schwab, and J. Sukys, Multi-Level Monte Carlo Finite Volume methods for uncertaintyquantification of acoustic wave propagation in randomheterogeneous layered medium, Tech. rep., ETH Zurich, 2014.
[4] R. Hiptmair, L. Scarabosio, C. Schillings, and C. Schwab, Advances in Computational Mathematics **44**(5), 1475–1518 (2018).
[5] E. Jacquelin, S. Adhikari, J. J. Sinou, and M. I. Friswell, Journal of Engineering Mechanics **141**(4), 04014145 (2014).
[6] D. Pivovarov, K. Willner, P. Steinmann, S. Brumme, M. Müller, T. Srisupattarawanit, G. P. Ostermeyer, C. Henning, T. Ricken, S. Kastian et al., GAMM-Mitteilungen p. e201900010 (2019).
[7] C. E. Rasmussen and C. K. I. Williams, Gaussian processes for machine learning (The MIT press, 2004).
[8] A. I. J. Forrester, A. Sóbester, and A. J. Keane, Engineering Design via Surrogate Modelling - A Practical Guide (Wiley, 2008).
[9] L. Le Gratiet, Multi-fidelity GP regression for computer experiments, PhD thesis, Université Paris-Diderot, Paris VII, France, 2013.
[10] J. Biehler, M. Mäck, J. Nitzler, M. Hanss, P. S. Koutsourelakis, and W. A. Wall, GAMM-Mitteilungen p. e201900008.