POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Executive Summary of the Thesis

Gaussian process emulators to accelerate sensitivity analysis and Bayesian parameter estimation: application to cardiovascular modeling

Laurea Magistrale in Mathematical Engineering - Ingegneria Matematica

Author: Alessandro Pirozzi

Advisor: Prof. Alfio Maria Quarteroni

Co-advisors: Stefano Pagani, Francesco Regazzoni

Academic year: 2021-2022

1. Introduction

Numerical models have been successfully applied to many complex engineering fields for in silico testing. However, some limitations may arise in uncertainty quantification routines because of the high computational costs required. In this work, we consider a method to significantly reduce this computational burden by replacing a high-fidelity mathematical model, obtained, e.g., with the finite element method or the finite volume method, with a computationally cheap surrogate model [8]. In particular, we employ a non-parametric regression method, namely the Gaussian process (GP), as a surrogate model, since it is completely data-driven and does not require an explicit knowledge about the functional relationship between input and output variables.
First, we highlight some of the main characteristics of Gaussian processes with the aim of better understanding how to manage them in more complex problems, such as the lumped parameter closed-loop model for the whole circulatory network. Then we take advantage of Gaussian processes prediction efficiency to accelerate the sensitivity analysis and the Bayesian parameter estimation by means of the Markov chain Monte Carlo (MCMC) method.

2. Gaussian processes

Let f denote an (unknown) function that maps the input $\mathbf{x} \in \mathcal{X}$ to the output $y \in \mathcal{Y}$, i.e. $f : \mathcal{X} \to \mathcal{Y}$. As reported in [4], a Gaussian process is a collection of random variables, any finite number of which has a joint Gaussian distribution. One of the main characteristics of Gaussian processes is that they are completely identified by a mean function and a covariance function. Indeed, given the mean function $\mu(\mathbf{x})$ and the covariance function $k(\mathbf{x}, \mathbf{x}')$, with $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$, the Gaussian process for a real process $f(\mathbf{x})$ can be written as

$$f(\mathbf{x}) \sim \mathcal{GP}(\mu(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')).$$

The mean function $\mu(\mathbf{x})$ reflects the average of all the functions in the distribution evaluated at a certain input $\mathbf{x}$. On the other hand, the covariance function $k(\mathbf{x}, \mathbf{x}')$ models the dependence between the function values at different input points $\mathbf{x}$ and $\mathbf{x}'$.

2.1. Covariance function and hyperparameters optimization

The covariance function represents a key element in Gaussian processes, since it determines the shape, smoothness and other important properties of the function that we want to model. For this reason, the choice of a suitable covariance function is crucial in order to obtain reliable predictions.
The covariance function used throughout this work is the squared exponential covariance function, also known as the exponentiated quadratic covariance function:

$$k(\mathbf{x}, \mathbf{x}'; \sigma, \lambda) = \sigma^2 \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\lambda^2}\right), \qquad (1)$$

where σ^2 is called amplitude, which determines the average distance of our function away from its mean, and λ is called length scale, which regulates the speed of decay of the correlation among the points.
These quantities are called hyperparameters and their values must be determined, even if it may not be easy in practical applications. This problem is treated by the training of the Gaussian process, which consists of the maximization of the log marginal likelihood with respect to the hyperparameters.

2.2. Predictive distribution

As explained in [4], the choice of the covariance function implies a distribution over functions

$$\mathbf{f}_* \sim \mathcal{N}(\mathbf{0}, K(X_*, X_*)),$$

where X_* represents the input values and the covariance matrix $K(X_*, X_*)$ is defined by using the squared exponential covariance function (1) elementwise.
As explained in [4], we are not interested in drawing random functions from the prior, but we actually want to exploit the information given by the training data about the function.
Given the training inputs and outputs $\{x_i, f_i | i = 1, \dots, n\}$ and the test inputs and outputs $\{x_{*,j}, f_{*,j} | j = 1, \dots, n_*\}$, the joint distribution in the noise free case of the training outputs $\mathbf{f}$ and of the test outputs $\mathbf{f}_*$ according to the prior is defined as

$$\begin{bmatrix} \mathbf{f} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K(X, X) & K(X, X_*) \\ K(X_*, X) & K(X_*, X_*) \end{bmatrix}\right).$$

Nevertheless, in real-life situations, we only have access to noisy function values $y = f(\mathbf{x}) + \epsilon$, where ϵ is an independent identically distributed Gaussian noise. As a consequence, we get

$$\operatorname{cov}(\mathbf{y}) = K(X, X) + \sigma_n^2 I,$$

where σ_n^2 represents noise variance. Then the joint distribution of the noisy observed values and the function values at the test points under the prior can be written as

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K(X, X) + \sigma_n^2 I & K(X, X_*) \\ K(X_*, X) & K(X_*, X_*) \end{bmatrix}\right).$$

In conclusion, the predictive equations for Gaussian process regression in the noisy case are given by

$$\mathbf{f}_* \,|X_*, X, \mathbf{y} \sim \mathcal{N}(\bar{\mathbf{f}}_*, \operatorname{cov}(\mathbf{f}_*)),$$

where

$$\begin{aligned} \bar{\mathbf{f}}_* &= \mathbb{E}[\mathbf{f}_* \,|X_*, X, \mathbf{y}] \\ &= K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1}\mathbf{y}, \end{aligned}$$

$$\begin{aligned} \operatorname{cov}(\mathbf{f}_*) = K(X_*, X_*) - K(X_*, X)[K(X, X) + \\ \sigma_n^2 I]^{-1} K(X, X_*). \end{aligned}$$

2.3. Anisotropic covariance functions

The squared exponential covariance function (1) is by definition an isotropic covariance function. This means that it represents a suitable choice only if the output varies uniformly in all the directions [7]. Therefore, in order to obtain an anisotropic version of the squared exponential covariance function, [4] proposes to introduce the matrix $M = \operatorname{diag}(\boldsymbol{\lambda})^{-2}$ as follows:

$$\begin{aligned} k_y(x_i, x_j) = \sigma^2 \exp\Big(-\frac{1}{2}(x_i - \\ x_j)^T M (x_i - x_j)\Big) + \sigma_n^2 \delta_{ij}, \end{aligned}$$

where vector $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_D)^T$ contains the length scales with respect to each parameter, so that it is possible to detect, when present, the anisotropic behaviour.
In particular, Table 1 shows how the root mean squared error and the L^∞ error reduce as we pass from the isotropic covariance function to the anisotropic one in a simple benchmark problem, confirming that significant improvements are detectable if we modify the approach.

Error	Isotropic	Anisotropic
RMSE	0.1077	0.0671
L^∞ error	0.4339	0.2220

Table 1: Root mean squared error and L^∞ error in the isotropic and anisotropic cases.

2.4. High-fidelity circulation model

As explained in [1], since the numerical modeling of the cardiovascular system is a computationally expensive problem, some high-fidelity reduced models, such as the lumped parameter models, can be introduced in order to minimize the computational burden. These approaches are based on geometrical reduction, so that it is possible to significantly simplify the description of spatially distributed physical systems.
Since this kind of model is particularly popular to describe electric circuits, it is possible to establish an analogy between the circulatory network and an electric circuit. This means that a physical meaning for the electric elements we typically find in a circuit must be provided:

- the resistance R models the dissipation due to the fluid viscosity;
- the capacitance C models the vessel compliance due to the elasticity of the wall;
- the inductance L models the inertial properties of the fluid.

Therefore, the lumped parameter closed-loop circulation model proposed in [5] can be characterized as follows:

- the systemic and pulmonary circulations are modeled with resistance-inductance-capacitance (RLC) circuits, one for the arterial part and the other one for the venous part;
- the four chambers (atria and ventricles) are modeled by time-varying elastance elements;
- the cardiac valves are represented as non-ideal diodes.

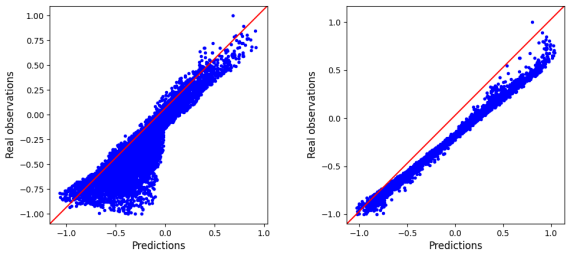
2.5. Numerical experiments

We analyze the obtained results by means of the root mean squared error and the L^∞ error between the high-fidelity circulation model introduced in Section 2.4 and its Gaussian process predictions. We test the prediction efficiency of Gaussian process regression in four different cases with 100, 500, 1000, 2000 training values and 9000 test values, whereas the considered output is the maximum pressure in the left ventricle $p_{\mathrm{LV}}^{\max}$. The idea is to let the parameters in Table 2 vary in their ranges used to train the Gaussian process, with the remaining ones which are fixed at their baseline value.

Case	Parameters
Case 1	$R_{\mathrm{AR}}^{\mathrm{SYS}}$, $C_{\mathrm{AR}}^{\mathrm{SYS}}$
Case 2	$R_{\mathrm{AR}}^{\mathrm{SYS}}$, $C_{\mathrm{AR}}^{\mathrm{SYS}}$, $E_{\mathrm{LV}}^{\mathrm{act}}$, $E_{\mathrm{LV}}^{\mathrm{pass}}$
Case 3	$R_{\mathrm{AR}}^{\mathrm{SYS}}$, $C_{\mathrm{AR}}^{\mathrm{SYS}}$, $E_{\mathrm{LV}}^{\mathrm{act}}$, $E_{\mathrm{LV}}^{\mathrm{pass}}$, $T_{\mathrm{LV}}^{\mathrm{contr}}$, $T_{\mathrm{LV}}^{\mathrm{rel}}$, $t_{\mathrm{LA}}^{\mathrm{del}}$, $t_{\mathrm{RA}}^{\mathrm{del}}$
Case 4	$R_{\mathrm{AR}}^{\mathrm{SYS}}$, $C_{\mathrm{AR}}^{\mathrm{SYS}}$, $E_{\mathrm{LV}}^{\mathrm{act}}$, $E_{\mathrm{LV}}^{\mathrm{pass}}$, $T_{\mathrm{LV}}^{\mathrm{contr}}$, $T_{\mathrm{LV}}^{\mathrm{rel}}$, $t_{\mathrm{LA}}^{\mathrm{del}}$, $t_{\mathrm{RA}}^{\mathrm{del}}$, $R_{\mathrm{VEN}}^{\mathrm{SYS}}$, $R_{\mathrm{VEN}}^{\mathrm{PUL}}$, $C_{\mathrm{VEN}}^{\mathrm{SYS}}$, $C_{\mathrm{VEN}}^{\mathrm{PUL}}$

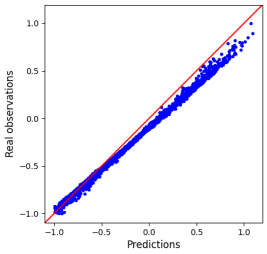
Table 2: Varying input parameters of the circulation model considered in the four cases.

Figure 1 shows that the prediction efficiency of the Gaussian process posterior predictive distribution improves as the number of training values increases.

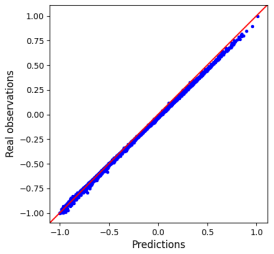


(a) 100 training values

(b) 500 training values



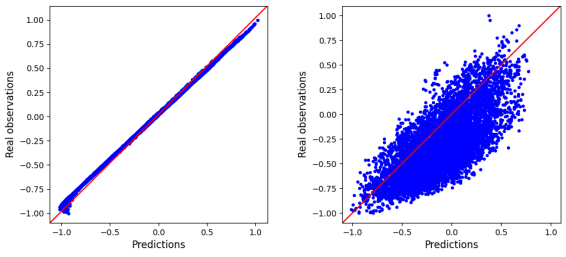
(c) 1000 training values



(d) 2000 training values

Figure 1: Behaviour of Gaussian process predictions with respect to the high-fidelity model observations with 100, 500, 1000 and 2000 training values as 4 parameters variate.

On the other hand, it is possible to detect a worsening of the predictions as the number of parameters gets larger (e.g., the RMSE increases from 0.0463 with 100 training values and 2 parameters to 0.2508 with 100 training values and 12 parameters), as shown in Figure 2. However, even if a more complicated model implies a worse prediction ability of the Gaussian process, an improvement is always noticeable by increasing the number of training values (e.g., the L^∞ error decreases from 0.7601 with 100 training values and 4 parameters to 0.0690 with 2000 training values and 4 parameters).



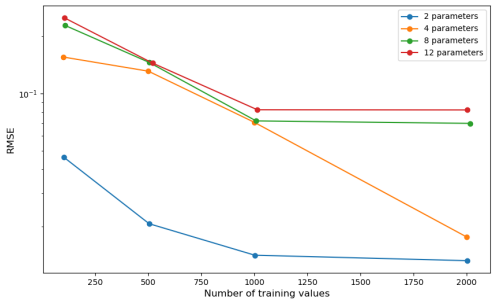
(a) 2 parameters (b) 12 parameters

Figure 2: Behaviour of Gaussian process predictions with respect to the high-fidelity model observations with 100 training values for the cases with 2 parameters in panel (a) and 12 parameters in panel (b) respectively.

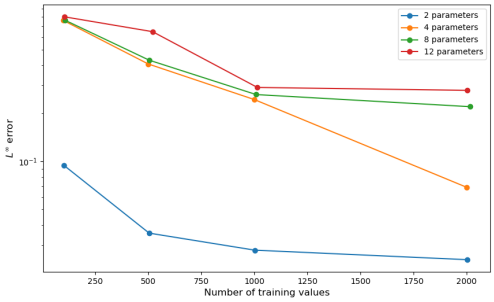
Indeed, Figure 3 shows that the errors increase as the number of parameters gets larger, since more varying input parameters imply a more complex behaviour to be predicted, but they also tend to decrease when the number of training values increases.
We can conclude that a similar trend is recognizable for all the cases, but the higher the variability of the output, the more demanding the training of the Gaussian process, since an increase in the number of training values implies a significant increase in the required computational time, as shown in Table 3.

N° of training values			
100	**500**	**1000**	**2000**
$\sim$ 31 s	$\sim$ 446 s	$\sim$ 3051 s	$\sim$ 13561 s

Table 3: Computational time required by Gaussian process training depending on the number of training values.



(a) RMSE



(b) L^∞ error

Figure 3: Root mean squared error and L^∞ error for each number of parameters as the size of the training sample varies.

3. Sensitivity analysis

Sensitivity analysis quantifies how sensitive model outputs are with respect to changes in model inputs. It can be used to determine a criterion to rank the most influential input parameters and, conversely, to identify which of them do not have a strong effect on a specific output.
As reported in [3], in general, sensitivity analysis approaches can be divided in two groups: local and global methods. Basically, local methods are employed to analyze the impact of input parameters at a specific point in the parameter space, whereas global methods examine the sensitivity with regard to the entire parameter distribution.
In this work, we consider a variance-based global sensitivity method that describes the amount of output variance generated from the variation of each parameter [6]. This analysis is performed by means of the Sobol indices, which provide direct information about the contribution of a specific parameter on the total variance of the output.

As explained in [2], given a real process $f : \mathbb{R}^d \to \mathbb{R}$, it is possible to write

$$Y = f(\mathbf{X}),$$

where the input $\mathbf{X} = (X_1, \dots, X_d)^T$ consists of d statistically independent random variables with known distributions, since the exact value of the input parameters is unknown. Similarly, this makes the model output Y a random variable as well. The first order Sobol indices are defined as

$$S_i = \frac{\mathbb{V}_i}{\mathbb{V}(Y)},$$

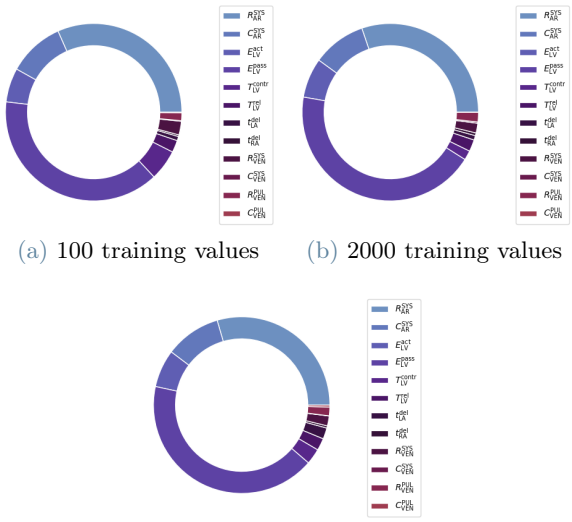
where

$$\mathbb{V}_i = \mathbb{V}(f_i(X_i)) = \mathbb{V}_{X_i}(\mathbb{E}_{\mathbf{X}_{\sim i}}[Y|X_i]).$$

The notation $\mathbf{X}_{\sim i}$ is used to indicate the set of all the input factors excluding X_i.

3.1. Numerical tests

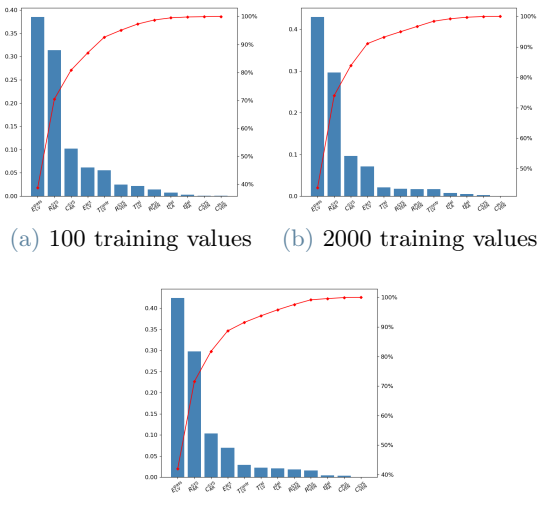
We perform the sensitivity analysis by means of the first order Sobol indices in the four cases reported in Table 2, whereas the considered output is always the maximum pressure in the left ventricle $p_{\text{LV}}^{\max}$. The goal is to prove that it is possible to obtain a reliable ranking of the input parameters with respect to Gaussian process predictions by comparing these results with those obtained in the high-fidelity model.
These results are in line with what we observed in Section 2.5; indeed, first order Sobol indices with respect to Gaussian process predictions are approximately equivalent to those obtained with the high-fidelity model in the first two cases, where a better posterior predictive distribution is obtained. On the other hand, when many varying input parameters are involved, the estimation becomes more challenging. Indeed, Figures 4–5 show that the values of the first order Sobol indices significantly change for some parameters from the lowest number of training values to the greatest one.
We can observe that the reliability of the sensitivity analysis performed on Gaussian process predictions is rather low for a small number of training values. For instance, the order of parameters with lower impact on the output considerably changes from the case with minimum number of training values to the case with maximum number of them. However, these charts suggest that it is always possible to obtain a significant improvement by increasing the size of the training sample, even if, as we reported in Table 3, this implies a trade-off in terms of computational cost.



(a) 100 training values (b) 2000 training values

(c) High-fidelity model

Figure 4: Comparison between pie charts with respect to Gaussian process predictions for the minimum/maximum number of training values and the one obtained from the high-fidelity circulation model in the case with 12 parameters.



(a) 100 training values (b) 2000 training values

(c) High-fidelity model

Figure 5: Comparison between Pareto charts with respect to Gaussian process predictions for the minimum/maximum number of training values and the one obtained from the high-fidelity circulation model in the case with 12 parameters.

4. Bayesian parameter estimation

As reported in [6], the patient-specific personalization of the circulation model requires the estimation of several parameters starting from clinical measurements. However, not all the required scalar quantities are usually available for this purpose and, when present, they are affected by noise.
For this reason, we exploit the Markov chain Monte Carlo method, which represents a suitable Bayesian method that allows us to solve the inverse problem (i.e. estimating parameters from outputs) taking into account the impact of the noise that affects the measurement of the quantities of interest and that reflects in uncertainty on parameters.
This method is non-intrusive, since it analyzes a large number of model evaluations for different parameter values with the aim of providing a suitable approximation of the posterior distribution. Moreover, since it only requires the model outputs, we can use Gaussian processes instead of the high-fidelity circulation model in order to reduce the computational time.

4.1. Numerical tests

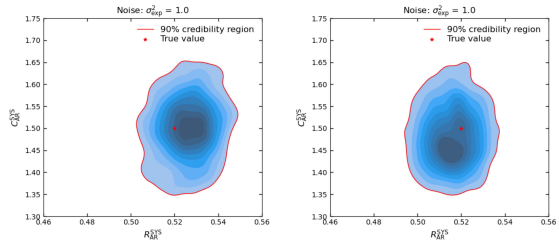
To prove the capability of Gaussian processes to accelerate the estimation of parameters, we proceed as follows:

- we perform a simulation with the high-fidelity circulation model from which we derive a set of observations $\mathbf{y}$ to which we add a synthetic measurement noise;
- we employ the Gaussian process instead of the high-fidelity circulation model in the Bayesian estimation of parameters performed by means of the MCMC method;
- we validate the obtained results with respect to the values used to generate the observations.

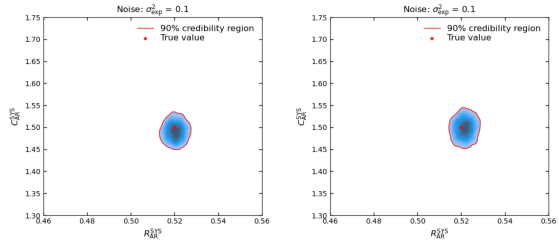
We consider a couple of parameters, namely the systemic arterial resistance $R_{\mathrm{AR}}^{\mathrm{SYS}}$ and the systemic arterial capacitance $C_{\mathrm{AR}}^{\mathrm{SYS}}$, whereas the remaining ones are fixed at their baseline value. On the other hand, the quantities of interest considered are the minimum and maximum arterial pressures. For parameters under investigation we prescribe a value that is slightly different from the baseline one, such as $R_{\mathrm{AR}}^{\mathrm{SYS}} = 0.52$ $\mathrm{mmHg\,s\,mL^{-1}}$ and $C_{\mathrm{AR}}^{\mathrm{SYS}} = 1.5$ $\mathrm{mL\,mmHg^{-1}}$. Then we demonstrate that it is possible to reduce the computational burden associated to the MCMC method for parameter estimation by means of Gaussian process emulators.
The measurement errors are added as artificial noises with zero mean and variance $\sigma_{\mathrm{exp}}^2 = \{1.0, 0.1, 0.01\}$ $\mathrm{mmHg^2}$ over the output values given by the high-fidelity model. On the other hand, the error corresponding to the usage of Gaussian process predictions for outputs estimation in the MCMC method is given by the observations noise variances obtained from the optimization of GP hyperparameters.
Figure 6 shows the comparison between the output of the Bayesian parameter estimation for the parameters pair $(R_{\mathrm{AR}}^{\mathrm{SYS}}, C_{\mathrm{AR}}^{\mathrm{SYS}})$ obtained by using the high-fidelity circulation model and the one obtained by exploiting Gaussian process predictions. In particular, we can observe the 90% credibility region and the prescribed exact value of parameters. Notice that for each value of the noise, the credibility region contains the exact value of the parameters and, as expected, for larger values of the noise, the size of the credibility region increases, since the estimate is more uncertain.
Moreover, it is possible to observe that the size of the credibility region in the high-fidelity case is slightly smaller for each value of the measurement noise. However, this does not represent an advantageous trade-off if compared with the Gaussian process based approach, since it results in a computationally expensive method (see Table 4) with no substantial improvements, especially as the measurement noise gets larger.

Gaussian process based approach	**High-fidelity model**
$\sim$ 1153 s	$\sim$ 13728 s

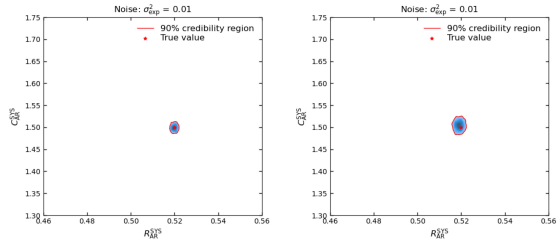
Table 4: Computational time required for Bayesian parameter estimation by exploiting Gaussian processes and the high-fidelity circulation model.



(a) $\sigma^2_{\text{exp}} = 1.0$ mmHg^2



(b) $\sigma^2_{\text{exp}} = 0.1$ mmHg^2



(c) $\sigma^2_{\text{exp}} = 0.01$ mmHg^2

Figure 6: Posterior distributions on the parameters pair $(R^{\text{SYS}}_{\text{AR}}, C^{\text{SYS}}_{\text{AR}})$ computed by means of the MCMC method with the high-fidelity model (on the left) and the Gaussian process for $\sigma^2_{\text{exp}} = \{1.0, 0.1, 0.01\}$ mmHg^2.

5. Conclusions

In this work, we observed that Gaussian processes are a powerful tool in making predictions also thanks to the fact that they are easily interpretable and not particularly computationally demanding. Indeed, we proved that Gaussian processes provide accurate estimations in sensitivity analysis and in Bayesian parameter estimation.

Some issues may arise when the number of varying input parameters is large, even if this problem could be successfully overcome by increasing the number of training values. However, this is not always possible in real-life problems; indeed, it may happen that a sufficiently large amount of data is not available for practical difficulties linked to data collection.

References

[1] M. Corti. Effects of Atrial Fibrillation on Left Atrium Haemodynamics: A Patient Specific Computational Fluid Dynamics Study. Master's thesis, School of Industrial and Information Engineering, 2021.

[2] H. Mohammadi, P. Challenor, and C. Prieur. Variance-based global sensitivity analysis of numerical models using R, 2022.

[3] S. Pagani. *Reduced-Order Models for Inverse Problems and Uncertainty Quantification in Cardiac Electrophysiology*. PhD thesis, School of Industrial and Information Engineering, 2017.

[4] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[5] F. Regazzoni, M. Salvador, P. C. Africa, M. Fedele, L. Dedè, and A. Quarteroni. A cardiac electromechanical model coupled with a lumped-parameter model for closed-loop blood circulation. Part I: model derivation. *Mox Reports*, pages 1–23, 2020.

[6] F. Regazzoni, M. Salvador, L. Dedè, and A. Quarteroni. A machine learning method for real-time numerical simulations of cardiac electromechanics. *Computer methods in applied mechanics and engineering*, pages 1–26, 2022.

[7] F. Vivarelli and C. K. I. Williams. Discovering hidden features with Gaussian processes regression. *Advances in Neural Information Processing Systems*, pages 613–619, 1999.

[8] Z. Zhou, Y. S. Ong, P. B. Nair, A. J. Keane, and K. Y. Lum. Combining Global and Local Surrogate Models to Accelerate Evolutionary Optimization. *IEEE Transactions on Systems, Man, and Cybernetics*, pages 66–76, 2007.