

Uncertainty Quantification and Sensitivity Analysis of Multi-parameter Models

Márcia R.S.S. Vagos

Simula Research Laboratory, Computational Physiology Department

University of Oslo, Department of Informatics

June 28, 2019

Contents

1	Introduction	3
1.1	Background and motivation	5
1.2	Sources of model uncertainty	6
1.3	Model building under uncertainty	7
2	Uncertainty quantification of multi-parameter models	9
2.1	Monte Carlo sampling	13
2.2	Latin-hypercube sampling	13
2.3	Generalized Polynomial Chaos Expansion	14
3	Sensitivity analysis of multi-parameter models	16
3.1	Local SA with Partial Derivatives	20
3.2	The Morris Method	21
3.3	Regression analysis	23
3.3.1	Multiple and Multivariate Linear Regression	24
3.3.2	Linear models with interaction terms	26
3.3.3	Least Squares Regression	26
3.3.4	Model adequacy and statistical significance	29
3.4	Correlation analysis	32
3.4.1	Non-linearities in the model	34
3.5	Sobol indices	36
3.6	Fourier Amplitude Sensitivity Test	40
4	Examples from cardiac modelling	44
5	Synthesis and conclusions	51

1 Introduction

Over the last decades, advancements in computer technology, including the exponential growth in computing power, have enabled the development of ever more complex computational tools and algorithms. Current supercomputers are able to store and analyse vast amounts of data, at much faster rates than before. At the same time, as scientific knowledge advances, the amount of data generated has become increasingly complex and hard to interpret. Thus, with these developments came the field of *computational modelling and simulation*, whereby physical processes and phenomena of interest are described in mathematical terms and implemented in a programming language to be processed by a computer.

Computational modelling has become an essential part of several different scientific disciplines, from physics to chemistry, biology, geosciences, climatology, material science, mechanical engineering, to name a few. Over the last decades, scientists and mathematicians have devoted a significant amount of effort to developing computational models that can emulate the structure and function of real physical systems, either at mechanistic or at a phenomenological level. A number of such models have been successfully implemented to generate new understanding of those systems, or to generate predictions of the behaviour of the systems. A few examples of these are models of fluid and solid mechanics, meteorology, biochemical reaction networks, hydrology, ecology, and systems biology. The latter has received particular attention due to the increasing availability and complexity of biological data, which required the development of specialized computational models and data analysis tools to integrate data from different spatial and temporal scales.

More recently, computational modelling, simulation and data analysis methods have been applied to various problems in biological sciences, giving rise to the field of *computational biology*. This is an interdisciplinary discipline that encompasses several other disciplines - such as biology, ecology, genetics, omics, biophysics, statistics, computer science - and that focus on developing/applying novel computational methods for knowledge discovery. Thus, models in computational biology can greatly differ in structure and application, spanning from DNA to single cell, to whole organ, to organism, and to population levels.

Computational modelling and simulation offers a fast and cheap way to test hypotheses

or new solutions, and thus it has become widespread in both scientific research and industrial practice. In healthcare, for instance, biomedical modelling has gained traction over the last few years, thanks to the improvements in bioimaging techniques, and has proven to be an asset in developing patient specific treatments. Similarly, the use of models and simulation in other fields has constantly shown to be of great value.

For computational models to be useful, they have to be able to reproduce certain phenomena or behaviours of interest, or to be able to make accurate predictions of how certain quantities of interest (QOI) evolve over time. This is termed predictive modelling, and one very well-known example is weather forecasting. In order for models to accurately and reliably reproduce or make predictions of the evolution of a system, they have to be populated with the 'correct' set of parameter values and initial conditions, which are in many cases uncertain or even completely unknown. Therefore, a critical part of the model development is to acknowledge model uncertainties. However, understanding how the choices of model parameters affect the result of a model, that is, how sensitive the model is to the values chosen can be quite challenging, in particular for complex models. Therefore, over the past decades, several techniques have been developed to deal with such uncertainties and variability in the data. Furthermore, quite often the experimental data available are highly variable, making it extremely difficult to determine the 'correct' values for the parameters. In this context, parameter identifiability becomes an issue in model development, and thus it is important to assess how the model is affected by each parameter. On the other hand, it is useful to quantify the uncertainty in the model predictions caused by the uncertainty in model parameters, so as to evaluate how confident we can be on the results.

The aim of the present essay is to provide an overview of methods for incorporating uncertainty and variability in model development and numerical simulations. In particular, the essay will focus on giving a brief overview of some of the most commonly used frameworks in uncertainty quantification, and sensitivity analysis. Additionally, I will provide some examples of context and applications of these methods whenever possible. Finally, I will provide a few examples of how some of the presented methods can be applied to the field of cardiac electromechanical modelling, which is an area of growing interest for the cardiology community.

1.1 Background and motivation

Despite the considerable differences among mathematical models, they are usually composed of a set of equations with several input parameters. Models typically produce output variables, or QOI, which depend on the choice of parameter values and initial conditions. Models vary considerably in size and complexity, and may be either linear or non-linear, monotonic or non-monotonic. Additionally, some models represent processes that are variable or stochastic in nature, and therefore model variables have to be defined as probabilistic quantities. Simpler models tend to have a smaller number of independent variables, whereas complex models can have up to thousands of variables which need to be defined and constrained.

A common approach to gain a deeper insight into the functioning of the model is to perform sensitivity analysis (SA) of the model to the input variables. SA enables the assessment of how much the outputs of the model, that is the physical quantities of interest, change when the input parameters are varied. This assessment is useful when the “real” values of the input parameters are either unknown, or are stochastic in nature. Model uncertainties due to ill-defined parameters can severely impact the accuracy of a single simulation, and abate confidence in the results.

SA has many applications in the model development process: (1) it allows to quantify the relationships between variables and outputs of a model, (2) to gain insight into the complexity of the model behaviour and the impact of model inputs in the response of the system, (3) to help identify the significant model parameters to be adjusted, and chose which parameters that can be frozen without significantly altering the results, and (4) to propose new hypothesis about mechanistic relationships between model parameters and outputs.

The concept of model sensitivity is closely related to uncertainty quantification (UQ). UQ consists in quantifying how much uncertainty in model parameters affects uncertainty the output of the model, in order to test hypotheses, or to make decisions on whether or not to trust the results of the model. imitations. Therefore, modellers (i.e., people who develop or use computational models) sometimes find it useful to quantify how much the uncertainty in input parameters can affect the outcome of a model, in order to test hypotheses, or to make decisions on whether or not to trust the results of the model. This method is, very

logically, termed uncertainty quantification (UQ).

The process of parametrization of the model can be extremely difficult when there is not sufficient knowledge of the system, and data to constrain in the parameter space. UQ and SA can be valuable tools in assisting model development and parametrization, since these frameworks provide additional knowledge of the functioning of the model and of the relationship between the multiple variables.

In this essay, I will focus on the differences between SA and UQ, and give some examples of common frameworks for performing SA and UQ in multi-variate models. This essay is not intended to be an exhaustive compilation of all the SA and UQ methods out there, as this list can be quite extensive, and I will also not go into the mathematical details of each of the presented methods. Instead, it is intended to give a general overview of the methods, and some common practical ways to implement them.

1.2 Sources of model uncertainty

Model uncertainty is often classified into two main categories: *epistemic* and *aleatory* uncertainty. *Epistemic uncertainty* is the uncertainty in model parameters associated with the lack of knowledge about the system and of the adequate value for the parameter(s). In contrast, *aleatory uncertainty* (often also termed aleatory variability) derives from the inherent randomness, due to the stochastic nature of the underlying process in the model. Epistemic uncertainty can be further differentiated into different categories depending of their sources:

- Experimental uncertainty (observation error): uncertainty due to uncontrollable measurement conditions, or measurement apparatus error. Can be reduced by repeating measurements and extracting statistical measures.
- Parametric uncertainty: unknown model parameters that cannot be measured or inferred, due to limited data or knowledge. These can be parameters with fixed but unknown value, or whose values vary in space or throughout the simulation.
- Uncertainty in model initial and boundary conditions: uncertainty in the initial state of the model, and on conditions that restrain the model output (based on known data

or model assumptions).

- Structural uncertainty: can also be referred to as model discrepancy, model bias, or model inadequacy. This uncertainty is a result of the lack of knowledge of the underlying physics of the system. It relates to how accurately a model describes the system, that is, how good of an approximation the model is to reality.
- Implementation or numerical uncertainty: this is the uncertainty due to numerical errors and approximations in the implementation of the mathematical model in a computational framework. Examples are finite element, or finite difference method for solving partial differential equations, numerical integration, or forward Euler method for solving ordinary differential equations.
- Interpolation uncertainty: uncertainty in simulated or experimental data points that are obtained from interpolation of known data points.

Forward uncertainty propagation allows to quantify statistical moments (mean, variance, etc.) or model outputs given uncertainty in the inputs. Typical applications for this type of UQ are to access model reliability and utility, and to perform SA for model parametrization and optimization.

Inverse uncertainty propagation allows to estimate uncertainty in input parameters given uncertainty in model outputs. This method can be used for bias correction (and thus quantification of model discrepancy) by estimating the bias function as the difference between the experimental and simulated observations given a distribution of the model parameters. This allows restraining probability distributions of model inputs with known experimental uncertainty of model outputs. Inverse uncertainty propagation can also be used for model calibration, by estimating the true value of the unknown parameters or their probability distributions.

1.3 Model building under uncertainty

Experimental data used to build and fit a model is often sparse, noisy and incomplete, and not measured with the specific purpose of building a computational model. There are also a

number of considerations regarding what simplifications can be made in the model in order to overcome incomplete knowledge or increase computational efficiency, without compromising the accuracy, and predictive power of the model. In fact, oversimplifications can result in the incapability of the model to reproduce the experimentally observed behaviours of interest. Therefore, building a computational model involves both the specification of the model structure and the parametrization of the model, which depend on the specific data set(s), and on the particular context in which the model will be used.

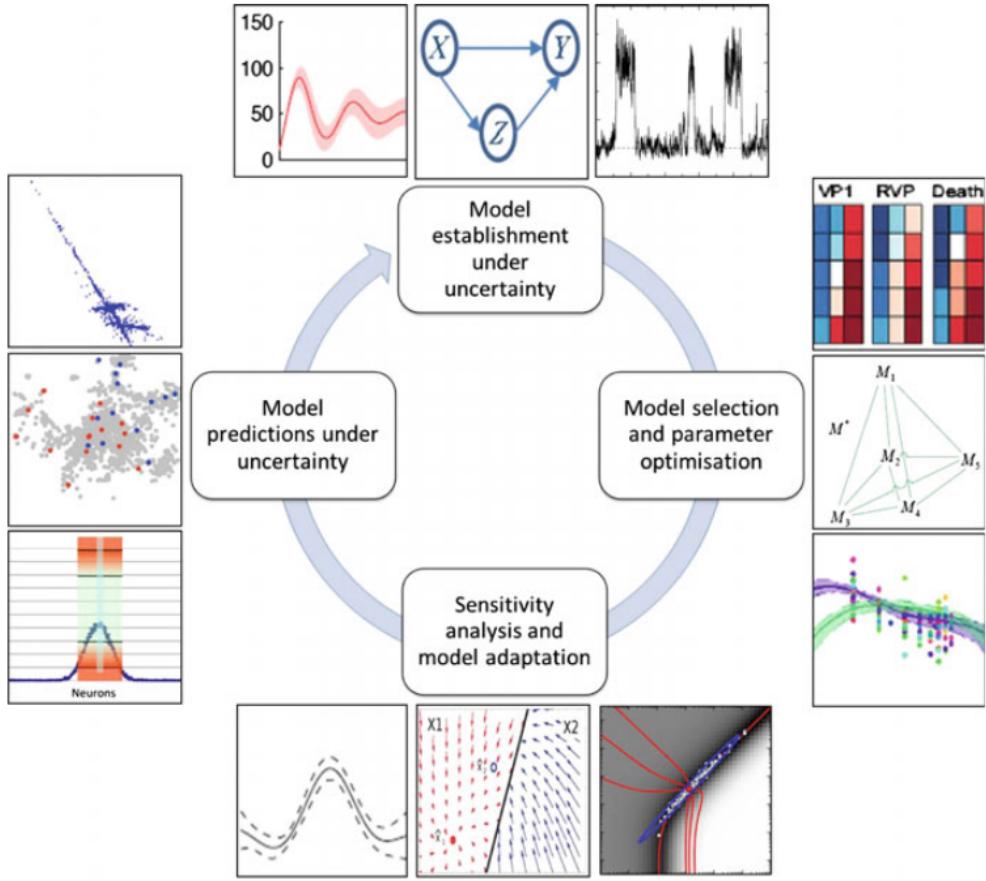


Figure 1: Schematic overview of the model development life cycle under uncertainty: (i) establishment of a model under uncertainty, (ii) model selection and parameter optimization, (iii) sensitivity analysis and model adaptation, (iv) model predictions under uncertainty. Reprinted from [1].

A crucial part in model development is parameter estimation and optimisation. Statistically-driven parameter optimisation techniques: simulated annealing, genetic algorithms, particle

swarm optimisation, etc. Global parameter search techniques include set inversion, constraint propagation techniques, and set-valued constraint propagation. Bayesian approaches can be used for subsequent model assessment.

Figure 1 showcases a life-cycle of typical model development under uncertainty. First, all uncertain parameters must be identified and, if relevant, structural uncertainties must also be included. Then, knowledge of the system can help select the most appropriate mathematical model for parameter optimization. This step can be assisted with UQ and SA techniques, by providing information on the most relevant parameter and which can be fixed to a nominal value before further analysis, as well as necessary changes in the model. Finally, several evaluations of the model are performed to produce model outputs incorporating model uncertainty.

2 Uncertainty quantification of multi-parameter models

The common way to assess model uncertainties is by treating the input factors as random variables, associated with a probability distribution, and performing several assessments of the model to estimate the uncertainty in the output responses.

Uncertainty quantification (UQ), or uncertainty analysis (UA), is a group of methods used to investigate the uncertainty in a model output that is generated from uncertainty in parameter inputs. The purpose of UQ is to quantify the degree of confidence in the values chosen for the model parameters, while it also allows to assess the reliability and sensitivity of complex models. More broadly, UQ methods can be used to gain insight into how uncertainties propagate through the model and thus can help guide model verification and model improvement. UQ has been successfully applied to computational models across several engineering problems.

UQ can be applied in two main problems: forward uncertainty propagation analysis, and parameter sensitivity analysis. **Forward uncertainty propagation** analyses how uncertainties in the input parameters propagate through the model and affect the output QOIs. In this framework, model parameters are treated as random variables with known probability distributions. This process is illustrated in Fig. 2. When *a priori* knowledge

to estimate the probability distributions is lacking probability density functions (PDF) can be used to defined model uncertainties. PDFs for the QOIs can then be constructed by performing multiple simulations with parameters values sampled from the PDF, using a suitable sampling method. In this framework, the uncertainty associated with the input variables can be propagated directly (forward) to the outputs via direct computation of the simulations.

As mentioned in the introduction, computational models are often rather complex and may contain several hundreds of parameters that need to be calculated every iteration, and thus each single simulation can become quite expensive computationally. Thus, it is desirable to be able to represent the overall behaviour of the model using a much more computationally tractable model. Therefore, in the UQ/SA framework it is often useful to analyse the workings of the model by replacing the model with a simpler “surrogate” model $f' \approx f$, also referred to as *metamodel*, which emulates the mapping of input to output variables. The surrogate model chosen is often a function with well-defined behaviour, and used as a *response surface* with high predictive capacity and low computational cost [2]. Polynomials, especially linear and quadratic functions, are common choices of metamodel, as well as splines, Gaussian processes, generalized linear models, among others. Quite often, these models are not able to emulate highly non-linear models with a sufficient level of predictive accuracy, and thus statistical learning models, such as support vector machines, and neural networks become preferable choices.

In order to define f' , some *a priori* knowledge of the system must be provided. When not enough *a priori* knowledge exists to define f' , it can be estimated by performing a “large enough” number of simulations with varying parameters to infer relationships between model inputs and outputs. The choice of a particular distribution will depend on *a priori* knowledge and assumptions of the nature of the data. When there is data characterizing the morphology of the uncertainty, this knowledge can be used to choose the most suitable probability distribution. However, when no prior knowledge is given about the nature of uncertainty, the best option is often to use the uniform distribution, since this makes the least number of assumptions on the likely of each particular value occurring. On the other

hand, when it is known that values around the nominal value are more likely to occur, for instance from experimental data, then it is often more reasonable to assume a normal distribution. Some authors also argue that it is in general preferable to obtain the probability distribution of the uncertainty in the input parameters by specifying selected quantiles of the CDF, rather than using ‘standard’ probability distribution [3]. However, doing this often requires expert knowledge, and sometimes experimental measurements, and thus might be in some cases impractical. Table 1 lists some of the commonly used probability distribution functions in UQ.

UQ results, i.e. the uncertainty in the output variables of a model, are usually presented are either statistical moments of the probability distribution, such as mean and variance, or as a reconstruction of the PDF, CDF, or even complementary CDF of the output uncertainty. Box plots can also be used to represent the output uncertainty, since they provide a compact and informative visualization of the probability distribution.

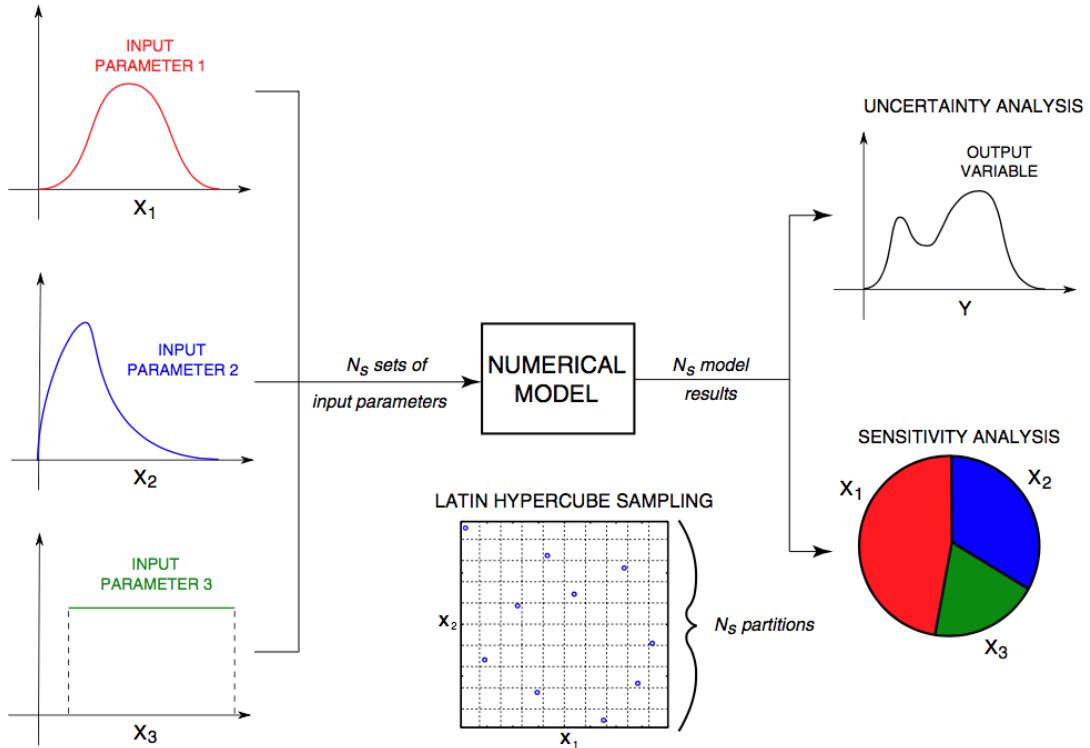
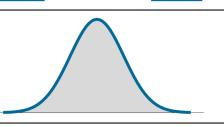
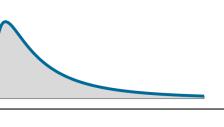
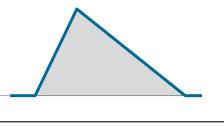
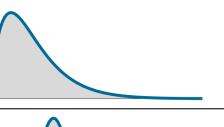
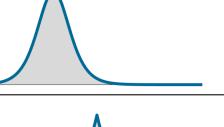
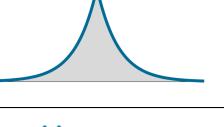


Figure 2: Reprinted from [4].

Table 1: Table of common probability distributions [5].

Name	Probability density function	Representation
Uniform	$f(x) = \frac{1}{b-a}, \quad a < x < b, -\infty < a < b < +\infty$	
Normal	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty$	
Log normal	$f(x) = \frac{1}{x\beta\sqrt{2\pi}} e^{-\frac{1}{2}(\ln(x/\alpha)/\beta)^2}, \quad x > 0$	
Triangular	$f(x) = \begin{cases} \frac{2(x-a)}{(b-a)(m-a)}, & a < x < m \\ \frac{2(b-x)}{(b-a)(b-m)}, & m \geq x < b \end{cases}$	
Gamma	$f(x) = \frac{x^{\beta-1}e^{-x/\alpha}}{\alpha^\beta \Gamma(\beta)}, \quad x > 0$	
Logistic	$f(x) = \frac{\lambda^\kappa \kappa e^{\kappa x}}{(1+(\lambda e^x)^\kappa)^2}, \quad -\infty < x < +\infty$	
Laplace	$f(x) = \begin{cases} (1/(\alpha_1 + \alpha_{12}))e^{x/\alpha_{11}}, & x < 0 \\ (1/(\alpha_1 + \alpha_{12}))e^{-x/\alpha_{11}}, & x \geq 0 \end{cases}$	
Poisson	$f(x) = \frac{\mu^x e^{-\mu}}{x!}, \quad x = 0, 1, 2, \dots$	

2.1 Monte Carlo sampling

Monte Carlo (MC) methods refer to a class of techniques of statistical sampling used to approximate solutions to problems without closed-form solution. An MC simulation consists of performing multiple model evaluations using random (or pseudo-random) sampling of probability distributions of model inputs. The MC method was initially formulated to estimate the integral of functions where it could not be determined analytically.

Each input parameter of the model is sampled from a specific of a probability distribution, depending on existing data and *a priori* information. If there are no available data, a natural choice is a uniform distribution (with a large enough range with minimum and maximum values for the parameters that covers the parameter space). If knowledge of the system exists that suggests a more frequent or expected value for a particular parameter, a common option would be to use a normal distribution with a large enough variance.

Typically, Monte Carlo (MC) methods are used as standard approach for UQ thanks to their accuracy and ease of implementation. On the down side, MC method require a high number of samples to converge. In fact, the accuracy of UQ results with the MC method is inversely related to the square root of the number of simulations (with the full model) performed. For complex models, they are very computationally expensive and cumbersome.

The MC method is a common approach to perform UQ due to its simplicity and good statistical results. However, this method is computationally expensive, and thus other methods have been developed that make UQ of large models with several uncertain parameters more computationally tractable. Improved sampling methods, such as the Latin-Hypercube sampling, and stochastic spectral methods such as the Polynomial Chaos Expansion, stochastic collocation, and stochastic reduced order models, are attractive alternatives.

2.2 Latin-hypercube sampling

The Latin-Hypercube sampling (LHS) method is a particular case of the MC method. This algorithm was developed to address the problem of under-sampling of values in the outer ranges of the pdfs. In this method, the input parameter space is partitioned into n intervals of equal probability that are randomly sampled. It is based on a technique called *stratified*

sampling without replacement. Stratification divides the cumulative probability distribution curve into equal intervals; one sample is then randomly taken from each interval without replacement so that the entire range for each parameter is explored and thus faithfully recreating the input pdf. Therefore, LHS allows for an unbiased estimate of the variation in the model output while requiring fewer samples than the standard random sampling with the MC method. Furthermore, this sampling scheme allows to uncover any hidden correlations between parameters.

The LHS algorithm can be applied to any type of continuous probability distributions, but typically the normal and uniform distributions are used. An example of how the LHS algorithm partitions the cumulative pdf is represented in Fig. 3.

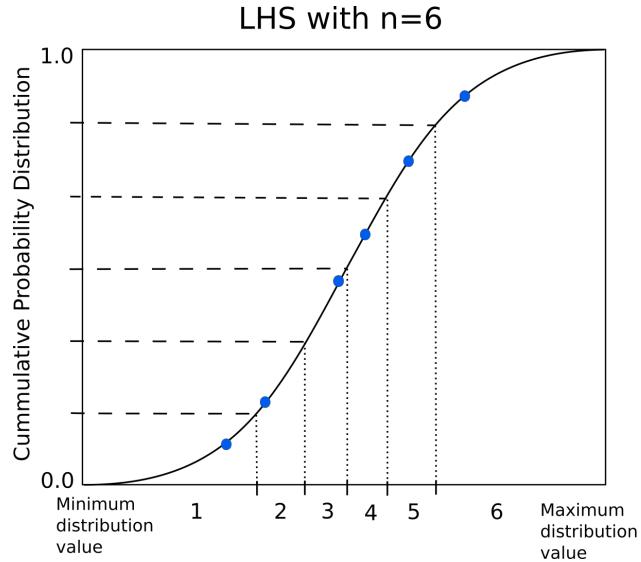


Figure 3: Schematic representation of the LHS method of the cumulative distribution function.

2.3 Generalized Polynomial Chaos Expansion

A common metamodeling approach is the so-called generalized polynomial chaos expansion (PCE). This method consists on building a response curve (*surrogate model*) based on polynomial functions to replace a computationally expensive model and perform dimensionality reduction. It is a common method for performing UQ and global SA.

In generalized PCE, a given variable is represented by means of second-order stochastic processes ($\hat{f}(w)$). The PCE is constructed with a truncated linear combination of polynomial basis functions. Each independent variable is replaced by a parametric equation in w . Thus, the expansion can be written as:

$$\hat{f}(w) = \sum_{k=0}^P C_k \Phi_k(\xi(w)) \quad (1)$$

Where $\Phi_k(\xi(w))$ are orthogonal polynomials of ξ of order P , and $\xi = (\xi_1, \dots, \xi_{1n})$ is a n -dimensional vector of Gaussian random variables with joint probability distribution $\rho(\xi)$, and $w \in \Omega$ is a random event. The polynomial basis $\{\Phi_k\}$ must satisfy the orthogonality criterion $\langle \Phi_i, \Phi_j \rangle = 0$. The number of terms P can be determined from the order of the expansion, and the number of random variables.

$$P + 1 = \frac{(n + p)!}{n!p!} \quad (2)$$

Given the basis functions satisfy the orthogonality property, the PCE coefficients C_k (or spectral modes of the expansion) can be written as:

$$C_k = \int_{\Omega} \hat{f}(\xi) \rho(\xi) \Phi_k(\xi) d\xi \quad (3)$$

The most challenging part in using the PCE method is to determine the coefficients of the expansion. There are two main classes of methods: intrusive, and non-intrusive. In intrusive methods, all dependent variables and random parameters in the governing equations are replaced with their polynomial chaos expansions; however, this is in general quite a demanding, time consuming task, requiring high computational cost. Therefore, intrusive methods are often not the preferred choice.

On the other hand, non-intrusive methods obtain approximations of the polynomial coefficients without making any modifications to the deterministic code, and thus do not require any additional modelling effort. There are several different approaches to non-intrusive methods: sampling based (for instance, Monte Carlo), point collocation based, and quadrature methods. Sampling-based often make use of random (eg., Monte Carlo) or LHS to estimate the above integrals. Given the complexity of these methods, here I will only focus on

the Point Collocation (PC) method. For a complete review on other methods to solve the polynomial expansions, refer to [3, 6].

The non-intrusive PC method consists of choosing an appropriate function basis ($\Phi_k(\xi)$ terms) to construct the response surface. Then, N_P random point vectors in the input space, and performing model evaluations for each of the random points. These will provide N_P estimations ($\hat{f}(\xi_n)$) of the response surface; typically, $N_P > (P + 1)$. These can then be replaced in eq. 4, yielding a system of linear equations, where the only unknowns are the spectral modes, C_P . If the system is over-determined (case in which there are more collocation points than polynomial terms), then the linear system can be solved with the Least Squares method.

$$\begin{bmatrix} \Phi_0(\xi_0) & \Phi_1(\xi_0) & \dots & \Phi_P(\xi_0) \\ \Phi_0(\xi_1) & \Phi_1(\xi_1) & \dots & \Phi_P(\xi_1) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_0(\xi_n) & \Phi_1(\xi_n) & \dots & \Phi_P(\xi_n) \end{bmatrix} \times \begin{bmatrix} C_0 \\ C_1 \\ \vdots \\ C_P \end{bmatrix} = \begin{bmatrix} \hat{f}(\xi_0) \\ \hat{f}(\xi_1) \\ \vdots \\ \hat{f}(\xi_n) \end{bmatrix} \quad (4)$$

Fig. 4 illustrates a response curve with the collocation points highlighted in black. The main advantages to using PCE is linked to its computational efficiency in performing multiple model evaluations and relatively cheap estimation of the expansion coefficients with methods like the Least Squares, making it therefore a very attractive method for UQ. Furthermore, the capability to tackle non-linearities in the model and to allow for different parameter distributions to be defined is another major advantage of PCE. On the down side, they require PC processes to be defined, which can be a non-trivial task.

3 Sensitivity analysis of multi-parameter models

The purpose of sensitivity analysis is to measure how much the variation in input factors contributes to the variation observed in the model responses, that is, how “influential” a certain parameter is to the output. These measures are called **sensitivity indices** (SI). Sensitivity analysis has many applications in modelling, as it allows:

1. to incorporate and understand the impact of uncertainty on model results

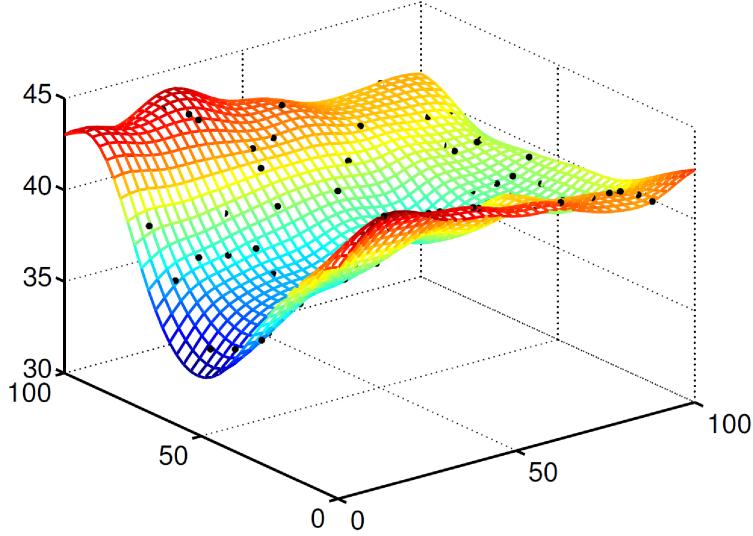


Figure 4: Representation of a search curve obtained with the Point Collocation method. Collocation points, shown in black, are used to perform evaluations of the model, which are then inserted into an over-determined system of linear equations to determine the PCE coefficients. Reprint from [7].

2. to do model optimization by prioritising the parameters are relevant for analysis, and fixing the others to their nominal values
3. to analyse model behaviour in comparison to other models
4. to do model validation based on observations.

Thus, SA provides a framework for parameter identifiability, dimensionality reduction, and parameter constraint and optimization. SA is a natural expansion of the UQ methods presented above, since it requires a large number of model simulations to be performed to obtain the distributions of the model responses, or quantities of interest (QOI). From these, relationships between model inputs and outputs, and among inputs, can be estimated to infer the role of the input parameters in the model.

The choice of sensitivity analysis method depends on the nature of the input variables and of the responses. Several different approaches have been developed to extract *importance measures* to estimate **sensitivity indices**. In a general way, the various methods have been classified as **screening methods**, **sampling-based**, **variance-based**, and **moment-**

independent SA. SA encompasses a wide range of methods, so the purpose of this essay is not to give a comprehensive overview of all of the SA, but rather to give a brief insight into some of the main methods used. For a complete review of SA, refer to [8].

Additionally to the classification above, SA methods can also be divided into two categories: Local SA, and Global SA:

Local sensitivity analysis (LSA) → Explores the effect of small changes of model parameters close to their nominal value; this is the preferred method when the input factors such as parameters or initial conditions are known with little uncertainty. The main approach is to examine the partial derivative of the response with respect to the input factors.

Global sensitivity analysis (GSA) → parameters are varied simultaneously within a wide range, and variance in outputs correspond to the sum of variance in each input parameter individually plus covariance between input variables.

Local methods for SA or one-at-a-time (OAT) sensitivity analysis are the most commonly used methods due to their simplicity and ease of implementation, but they assume that model response to different inputs is independent. This assumption is quite often not true, and thus SA results can be biased. Global SA methods overcome this limitation by varying the input parameters simultaneously, thus quantifying the sensitivity of to that input by averaging the output over the other input parameters. The drawback is that GSA are more computationally costly and can become infeasible for expensive models. However, GSA that use a statistical emulator in place of the model (for instance, the Sobol method and FAST) can greatly reduce the computational cost of SA by requiring fewer model runs.

Additionally, SA methods can be subdivided into three broad classes: *sampling-based*, *variance-based*, and *screening* methods. The sampling-based methods rely on the construction of a dataset of input and output pairs of values, in order to quantify relationships between the two groups. Examples of sampling-based SA methods include multi-parameter sensitivity analysis, regression models, and correlation analysis. In contrast, variance-based methods infer the sensitivities of the model based on decomposition of the output variance in components attributable to each input variables. Methods in this category include the Sobol Indices, and the Fourier Amplitude Sensitivity Test. The screening methods lie in

between local and global SA since they scan the whole parameter space (so they are global in a sense), but using local derivatives to estimate the average effect of one variable at a time.

The particular effect of the variable to be quantified also varies with the SA method. *Partial* sensitivity indices are based on importance measures that quantify only the marginal effect of an input variable on the outcome, that is the effect of varying that input alone. In contrast, *total* sensitivity indices measure the combined marginal effect and interactions with other inputs on the variance of the output.

SA techniques will perform differently for different types of computational problems. The choice of particular SA method mostly depends on the nature of the underlying model (non-linearities, monotonicity, and interactions among input variables), the structure of the data (continuous, sparse, binominal), and the desired sensitivity effect (partial versus total sensitivities).

Another aspect of SA is the analysis of the evolution of sensitivity indices in time. This is called **dynamical**, or **time-varying SA**. I will not cover this in the present essay, but the basic idea behind dynamical SA is that in models with a time dependent component, the importance of the various parameters to the model output may vary over the time scale. When dealing with narrow time spans, sensitivities may be invariant in time and so checking SI in various time points might be unnecessary. However, if time-varying sensitivities have cannot be disregarded, sensitivities should be taken at different time points. If no *a priori* knowledge exists on which are the critical time points to do the analysis, SA can be done in an exploratory way by performing multiple analyses at regular time intervals and checking for any time dependencies.

The next subsections will give an overview of four SA methods commonly used in multi-parameter models. It is not meant to give an exhaustive description of all SA methods available. For further reading on sensitivity analysis methods and reviews, see [8, 9, 3, 10, 11].

3.1 Local SA with Partial Derivatives

The Local Partial Derivatives method is a method of local SA that quantifies fluctuations in the model output due to the effect small perturbations in the input parameters [12]. Typically, model parameters are varied one-at-a-time in a small degree, eg., 0.1 %, while all the other input factors are fixed at their nominal values. Local sensitivity indices are quantified by first-order partial derivatives of the model output are taken with respect to the individual input variables.

Partial derivatives measure the change in the slope of the tangent to the point $(\mathbf{x}, f(\mathbf{x}))$ in the direction of x_i^0 , fixing the remaining inputs, and thus do not take into account the effects of interactions among input factors. Therefore, global methods have been introduced to overcome this limitation. Additionally, since the sensitivities provided by the local method consider only the effects of variability in input parameters in a small range around their nominal values, this method is not suitable for performing SA in models with non-linearities and largely varying input parameters. For a model of N input variables:

$$y = f(x_1, x_2, \dots, x_N) = f(\mathbf{x}), \quad (5)$$

where $\{x_1, x_2, \dots, x_N\}$ is a vector of N independent parameters with nominal values $\tilde{\mathbf{x}} = \{\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_N\}$, the sensitivity index of is given by the partial derivatives:

$$S_i = \frac{\partial f(\tilde{\mathbf{x}})}{\partial x_i}, \quad (6)$$

where S_i is the sensitivity index of the i th input parameter, and measures how much changing x_i influences the model output.

Since different input parameters can be expressed in different units and differ by several orders of magnitude, it is common to normalize the sensitivity indices. The normalization can be done with respect to either the mean value, the standard deviation, or as a percentage of parameter change:

$$\text{Mean value: } S_i = \frac{\tilde{x}_i}{f(\tilde{\mathbf{x}})} \frac{\partial f(\tilde{\mathbf{x}})}{\partial x_i} \quad (7)$$

$$\text{Standard deviation: } S_i = \frac{\sigma(x_i)}{\sigma(f(\tilde{x}))} \frac{\partial f(\tilde{x})}{\partial x_i} \quad (8)$$

$$\text{Percent change: } S_i = \frac{f(x_1, \dots, x_i + p\% \cdot x_i, \dots, x_N)}{p\% \cdot x_i} - \frac{f(x_1, \dots, x_i, \dots, x_N)}{p\% \cdot x_i} \quad (9)$$

The main advantage of this method in comparison to global SA methods is its ease of implementation and low computational cost. Local SA can be used in the first stage of model development and analysis to obtain an overview of the model behaviour and pinpoint model inconsistencies.

3.2 The Morris Method

The Morris method, or Elementary Effects method, is the most used screening method in sensitivity analysis [13]. This method is used to identify non-influential variables in a model with a large number of variables. It uses a one-at-a-time (OAT) approach, whereby each input parameter is varied at once, thus allowing to quantify the main effects of varying each input individually. Like other screening methods, the Morris method provides quantitative measures of the sensitivities of the model, allowing to rank the inputs in order of importance. However, unlike other methods, this method does not allow to quantify the relative importance of the input variables.

Consider the model $Y = f(X_1, X_2, \dots, X_k)$. In this method, the input space is discretized into a k -dimensional grid with n levels. In each OAT analysis, a random trajectory in the input space is defined by randomly sampling the input space in a random direction of variation. Each trajectory has $(k + 1)$ points where inputs are varied in steps $\Delta = \{0, 1/(p-1), 1/(p-1), \dots, 1\}$, as illustrated in Fig. 5. Each trajectory provides an estimate of the *elementary effect* of each variable X_j on Y , E_j .

$$E_j^{(i)} = \frac{f(X_1, \dots, X_{i-1}, X_i + \Delta, X_{i+1}, \dots, X_k) - f(\mathbf{X})}{\Delta} \quad (10)$$

A number r of OAT analyses is performed for each input to estimate two importance measures, μ_j and σ_j :

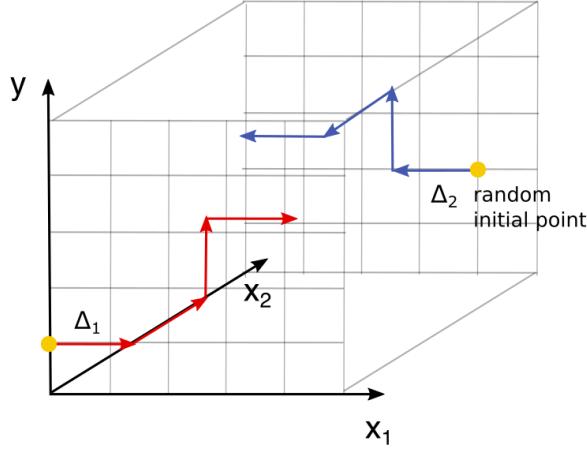


Figure 5: Schematic representation of the Morris method, showing two different trajectories (red and blue) starting from different initial random points (x_1, x_2) .

$$\mu_j^* = \frac{1}{r} \sum_{i=1}^r |E_j^{(i)}| \quad (11)$$

$$\sigma_j = \sqrt{\frac{1}{r} \sum_{i=1}^r \left(E_j^{(i)} - \frac{1}{r} \sum_{i=1}^r E_j^{(i)} \right)^2} \quad (12)$$

μ_j^* measures the overall importance of the input X_j on the model output Y , and σ_j measures the non-linear and/or interaction effects of the inputs. The use of the mean of the absolute values of the elementary effects, proposed by Campolongo et al. [14], resolved the problem of compensatory effects with opposite signs in non-monotonic models. The relationship between the values of μ_j , and σ_j allow to group the variables into three distinct groups: inputs with negligible effects (small μ_j , and large σ_j); inputs with large effects without interactions (large μ_j , and small σ_j); and inputs with non-linear and/or interactions effects (large μ_j , and large σ_j).

The Morris method is useful method for performing SA given its simplicity. It allows screening model with a large number of parameters, without requiring certain properties such as monotonicity or additivity. Furthermore, it provides a measure of the main effects of a variable, including to some extent the effects arising from interactions with other variables [14]. The main limitation is that the design proposed by Morris [] consists on varying

parameters one-at-a-time, and therefore similarly to the partial derivatives method, it does not provide a measure of the output sensitivity to an input taking averaged in relation to the effects of the remaining inputs.

3.3 Regression analysis

Global SA can be performed with classical regression methods. This is a sampling-based form of SA, in that both inputs and outputs have to be sampled in order to provide the raw data to fit a regression model.

Regression provides an algebraic representation of the relationship between multiple input variables X_i , and an output variable y . Thus, it is a form of predictive modelling which investigates the relationship between one or more independent variables (predictors), and a dependent variable (response).

Regression methods essentially find the ‘best-fit’ of a set of data points to some sort of mathematical relation, whether that is linear, polynomial, logarithmic, exponential, or any other function. More typically, linear regression is used for fitting a straight line through a set of data points (observations) using some goodness-of-fit criterion. The mathematical expression is defined by a set of regression coefficients which provide information on:

- the significance of relationships between dependent variable and independent variable;
- the strength of impact of multiple independent variables on a dependent variable.

Given this, regression coefficients can be used to provide a qualitative measure of sensitivity indices of the model to the input parameters, and thus rank the inputs in order of their importance. There are several different regression models (linear, polynomial, logarithmic, exponential, logistic etc.), and the particular choice of model depends on three main aspects: (i) the number of independent variables, (ii) the type of dependent variables, and (iii) the trends in the data. In Multiple Linear Regression (MLR), a straight line is fit to the data as a function of more than one independent variables. If several output variables are considered, it is called Multivariate Linear Regression (MVLR), or general Linear Model. I will be addressing only the cases of MLR and MVLR in the following subsection, as these have a broad range of applications.

3.3.1 Multiple and Multivariate Linear Regression

In most computational models a given response is influenced by more than one predictor. Therefore, it is often more useful to model the responses as a function of several predictors. This is called **Multiple Linear Regression (MLR)**. In MLR, each observation of the dependent variable y_j can be expressed as a linear combination of the observed independent variables x_{ij} and the regression coefficients β_{ij} , plus an intercept β_{0j} and a residual term ϵ , much like in the simple linear regression model.

$$y_j = \beta_0 + \beta_1 x_{j1} + \beta_2 x_{j2} + \cdots + \beta_p x_{jp} = \beta_0 + \sum_{i=1}^p \beta_i x_{ji} + \epsilon \quad (13)$$

Where p is the number of predictors of the linear model (or input parameters), and y_{ij} is the j^{th} response. The regression coefficients correspond to the slope of the line on the simple regression model. Typically, the number of independent observations of predictors and responses is much larger than the number of predictors. This ensures that the system is overdetermined and improves the prediction capacity of the model. In matrix notation, the model can be written as:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \times \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_p \end{bmatrix}$$

$$\mathbf{y} = \mathbf{X} \cdot \boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (14)$$

where p is the number of varied input parameters, and ϵ is the residual term of the model, which is assumed to have identical a normal distributions and be uncorrelated. Typically $n \gg p$ and $n \gg k$. In matrix representation, each independent observation of the predictors is a row in the \mathbf{X} matrix, and each response (ie, the output of each individual simulation with the computational model) is an entry in the \mathbf{y} vector.

In the case of the MVLR with multiple responses, the equations above can be written as:

$$y_{ij} = \beta_{0j} + \sum_{j=1}^p \beta_{ji} x_{ij} + \epsilon_{ij} \quad (15)$$

Similarly to MLR, here the outputs from each independent simulation constitute a single row in the \mathbf{Y} matrix, and each column in the \mathbf{B} matrix constrains the regression coefficients of the linear model. The system can be written as:

$$\begin{bmatrix} y_{11} & y_{12} & \dots & y_{1k} \\ y_{21} & y_{22} & \dots & y_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \dots & y_{nk} \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ 1 & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \times \begin{bmatrix} \beta_{01} & \dots & \beta_{0k} \\ \beta_{11} & \dots & \beta_{1k} \\ \vdots & \ddots & \vdots \\ \beta_{p1} & \dots & \beta_{pk} \end{bmatrix} + \begin{bmatrix} \epsilon_{11} & \dots & \epsilon_{1k} \\ \epsilon_{21} & \dots & \epsilon_{2k} \\ \vdots & \ddots & \vdots \\ \epsilon_{n1} & \dots & \epsilon_{nk} \end{bmatrix}$$

$$\mathbf{Y} = \mathbf{X} \cdot \mathbf{B} + \mathbf{E} \quad (16)$$

where n is the number of samples, p is the number of varied input parameters, and k is the number of outputs from the model. Again, the residuals random and are normally distributed with zero mean $(\epsilon_{i1}, \dots, \epsilon_{ik}) \sim N(0, \Sigma)$.

Figure 6 schematically illustrates the typical dimensions of the data in MVLR of computational models with a large number of observations.

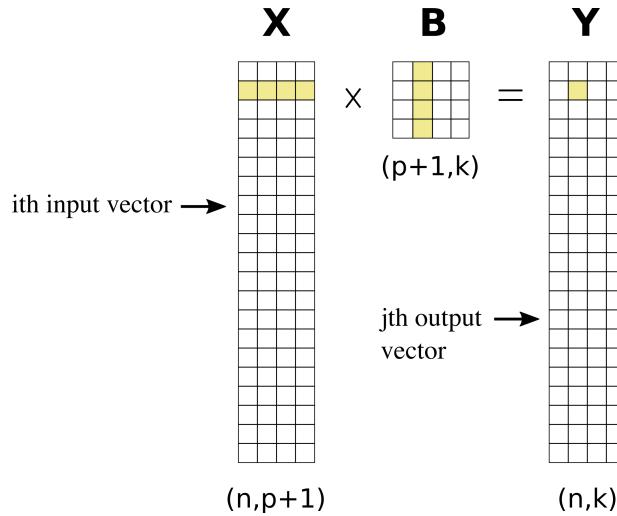


Figure 6: Representation of the input and output data structure and in MVLR models.

In order to eliminate effects of different units among variables, regression coefficients are often expressed as normalized by the variances:

$$SRC_j = \hat{\beta}_j \sqrt{\frac{Var(X_j)}{Var(Y)}} \quad (17)$$

This is called the Standard Regression Coefficient (SRC). SRC_j they measure the variable importance based on their effect on y relative to $\text{Var}(y)$ of perturbing X_i from $\mathbb{E}(X_i)$ by a fixed fraction of $\text{Var}(X_i)$. The larger the magnitude of β_i , the more important is the input X_i to the output y . A negative coefficient means that the independent and the dependent variables vary in opposite directions Importantly, the coefficients of regression in linear regression can only provide a *quantitative* measure of sensitivities if the data is perfectly linearly related, that is, if the coefficient of determination of the regression (introduced in the next subsection) is approximately 1. Also, note that the SRCs provide a useful measure of variable importance only if all the X_i are independent, and they do not provide reliable indications otherwise. .

3.3.2 Linear models with interaction terms

The MVLR model can be expanded to include interaction terms between one or more predictor variables of order >1 :

$$y = \beta_0 + \sum_{j=1}^p \beta_j f_k(x_{1,x_2}, \dots, x_p) + \epsilon \quad (18)$$

$$\hat{y} = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p + \beta_{p+1} x_1^2 + \beta_{p+2} x_1 x_2 + \dots \quad (19)$$

Although a such a model is a polynomial model on x of order 2, the model also can be seen as a linear combination of the β -parameters, with the x -terms being seen as the weights. Although the MVLR allows for a high degree of flexibility for the incorporation of interactions among variables, it is a highly complex process as it requires *a priori* knowledge of the covariances of the independent variables. There are techniques that can assist in uncovering hidden relationships between input variables, but in many cases these interactions terms can be neglected with significantly compromising the accuracy of the predictions.

3.3.3 Least Squares Regression

The method of least squares is the most common approach in regression analysis to approximate the solution of overdetermined systems. This method is also often referred to as Least

Squares regression. It allows to determine unbiased estimates of the β , often termed as least squares estimators ($\hat{\beta}$):

$$\hat{\mathbf{y}} = \mathbf{X} \cdot \hat{\beta} \quad (20)$$

When used with a linear regression model, it is called Ordinary Least Squares (OLS). The OLS fits a hyperplane into a $(p + 1)$ -dimensional space that minimizes the sum of squared residuals. "Least squares" means that the overall solution minimizes the sum of the squares of the residuals of the analytical and the predicted solutions.

The residuals are defined as:

$$\epsilon = \mathbf{y} - \mathbf{X}\beta = \mathbf{y} - \hat{\mathbf{y}} \quad (21)$$

Thus, the sum of squared residuals is given by:

$$\sum_{i=1}^n \epsilon_i^2 = \epsilon' \epsilon = (\mathbf{y} - \mathbf{X}\hat{\beta})'(\mathbf{y} - \mathbf{X}\hat{\beta}) \quad (22)$$

Thus, the problem can be formulated as:

$$\min_{\beta \in \mathbb{R}^{p+1}} \|\mathbf{y} - \mathbf{X}\beta\|^2 = \min_{\beta \in \mathbb{R}^{p+1}} \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 \quad (23)$$

By the Orthogonal Decomposition Theorem, the residuals are minimal when ϵ_i are orthogonal to the columns of \mathbf{X} , resulting in a closed-form expression for calculating the $\hat{\beta}$:

$$\mathbf{X}'(\mathbf{y} - \mathbf{X}\hat{\beta}) = 0 \Leftrightarrow \hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad (24)$$

Which has a unique solution if the inverse of $(\mathbf{X}'\mathbf{X})$ exists, that is, if the columns of \mathbf{X} are linearly independent (the predictors are all uncorrelated). In the case of MVR, the OLS problem has the same solution, where $\hat{\beta}$ and \mathbf{y} are replaced by the multivariate matrices:

$$\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \Rightarrow \hat{\beta}_j = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}_j \quad (25)$$

It is common practice to centre and normalize \mathbf{X} and \mathbf{Y} prior to computing the $\hat{\beta}_{ij}$. This allow to compare the $\hat{\beta}_{ij}$ against each other, and makes it so that the intercept term β_{0j} can

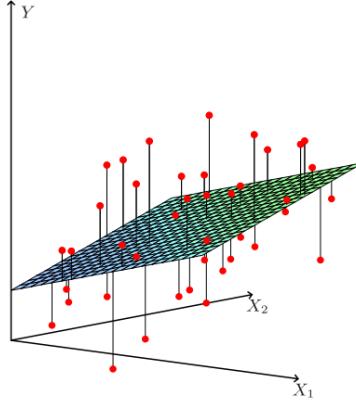


Figure 7: Least squares fitting. Reprint from [15].

be interpreted as the expected value of y_j . Transformation of the data can be done with z -scores $z_i = \frac{x_i - \bar{x}}{\sigma}$.

A property of the Least Squares estimators $\hat{\beta}$ is that they are unbiased, and the residuals are normally distributed with zero-mean and are uncorrelated. The residual variance is defined by the sum of squares of the residuals:

$$\text{MLR: } \hat{\sigma}^2 = \frac{\sum_{i=1}^n \epsilon_i^2}{n - p - 1} \quad (26)$$

$$\text{MVR: } \hat{\Sigma}^2 = \frac{\sum_{i=1}^n \epsilon_i \epsilon_i'}{n - p - 1} = \frac{\sum_{i=1}^n (\mathbf{y}_i - \hat{\mathbf{y}}_i)(\mathbf{y}_i - \hat{\mathbf{y}}_i)'}{n - p - 1} \quad (27)$$

Additionally, the residual variance is model dependent, meaning that it depends on which predictors are used to build the linear model. Therefore, a natural question that arises is whether the particular choice of input parameters is the most adequate for the model. In general, the choice of predictors should be such to minimize residual variance. Model adequacy can be tested through several metrics. The F -test statistic

By replacing eq. 24 into eq. 20, we get:

$$\hat{\mathbf{y}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{H}\mathbf{y} \quad (28)$$

\mathbf{H} is called the Hat-matrix, and it is useful since it maps the vector of observed values \mathbf{y} onto the vector of predictions $\hat{\mathbf{y}}$, thus allowing to account for approximation errors when using linear regression models.

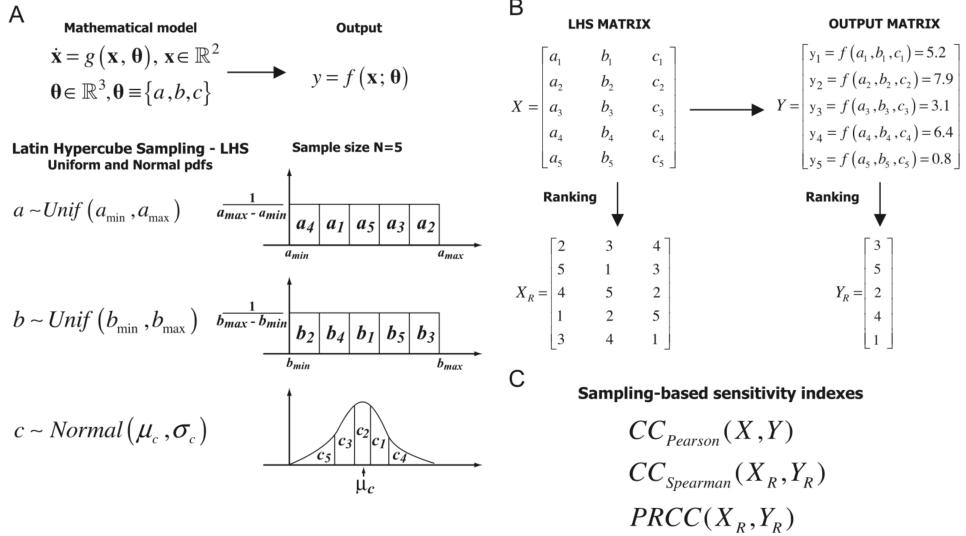


Figure 8: Schematic overview of UQ and SA performed with the the LHS sampling and PRCC methods. A: the mathematical model is defined consisting of a system of ordinary differential equations. LHS method is used to randomly sample the input parameter space, where input parameter values are specified with probability distributions functions; $N_S = 5$ samples are drawn from each distribution. B: the randomly picked samples are structured into an input matrix X , and N_S simulations are run to produce the output matrix Y . X and Y are rank-transformed into X_R and Y_R , respectively. C: Pearson correlation coefficients ($CC_{Pearson}$) are determined from X and Y ; Spearman Rank CC ($CC_{Spearman}$), and Partial Rank CC (PRCC) are determined from X_R and Y_R . Reprint from [16].

3.3.4 Model adequacy and statistical significance

As mentioned, the sensitivity measures introduced above are based on the assumption of a linear relationship between predictive variables and model responses. The validity of the linear hypothesis can be confirmed by statistical testing.

There are several ways to evaluate model adequacy, or the “goodness of fitness”, for regression problems. In linear regression, one common measure is the Coefficient of Determination (R^2). The R^2 represents the proportion of variation in Y that is explained by the predictors, that is, how strong is the linear dependency of the variables.

This is estimated from three different sums-of-squares:

$$\text{Sum-of-squares Error: } SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (29)$$

$$\text{Sum-of-squares Regression: } SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \quad (30)$$

$$\text{Sum-of-squares Total: } SST = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (31)$$

It can be verified that $SST = SSR + SSE$. In MVR models, these quantities become crossproducts:

$$\text{Sum-of-squares Error: } SSE_{CP} = \sum_{i=1}^n (\mathbf{y}_i - \hat{\mathbf{y}}_i)(\mathbf{y}_i - \hat{\mathbf{y}}_i)' \quad (32)$$

$$\text{Sum-of-squares Regression: } SSR_{CP} = \sum_{i=1}^n (\hat{\mathbf{y}}_i - \bar{\mathbf{y}})(\hat{\mathbf{y}}_i - \bar{\mathbf{y}})' \quad (33)$$

$$\text{Sum-of-squares Total: } SST_{CP} = \sum_{i=1}^n (\mathbf{y}_i - \bar{\mathbf{y}})(\mathbf{y}_i - \bar{\mathbf{y}})' \quad (34)$$

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}, \quad 0 \leq R^2 \leq 1 \quad (35)$$

A similar metric to the R^2 is the **F -test** statistic. The F -test for regression tests the hypothesis whether any of the p predictor variables have a relationship with the response, that is the ratio of explained to unexplained variance in the response:

$$H_0 : \beta_1 = \cdots = \beta_{p-1} = 0$$

$$H_1 : \beta_j \neq 0, \text{ for at least one value of } j$$

The F -test compares the F^* statistic to the F distribution. The F -table [17] provides an estimative of the p -value corresponding to the F -value given n and p ($p^* = P(F_{p,n-p-1 > F^*})$).

$$F^* = \frac{SSR/p}{SSE/(n-p-1)} \sim F_{p,n-p-1} \quad (36)$$

Similarly, the significance of each of the $\hat{\beta}$'s can be tested with a ***t*-test**. The hypotheses to be tested are where a $\hat{\beta}_j$ is zero or not:

$$H_0 : \beta_j = 0, \quad H_1 : \beta_j \neq 0$$

The *t* statistic is determined as:

$$t = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)}, \quad se(\hat{\beta}_j) = \sqrt{\frac{SS_E}{(n - 2)}} \quad (37)$$

where $se(\hat{\beta}_j)$ is the standard error of β_j . The *t* statistic can be compared with the the *Student's t distribution* to determine the *p*-value.

A simple way to test if a group of predictors can be removed from the model, is to subdivide the predictors and respective regressors into two subgroups, with one group containing $q < p$ predictors to be tested, and the other group the remaining $p - q$ predictors. Then, a new hypothesis can be formulated:

$$H_0 : \beta_{q+1} = \beta_{q+2} = \dots = \beta_p = 0$$

$$H_1 : \text{at least one } \beta_k \neq 0$$

We can then build a full model and a reduced model excluding the q predictors, and compare the F^* -statistic to infer on their significance.

$$\text{Full model: } y_i = \hat{\beta}_0 + \sum_{i=1}^p \hat{\beta}_i x_{ij} + \epsilon_i$$

$$\text{Reduced model: } y_i = \hat{\beta}_0 + \sum_{i=1}^q \hat{\beta}_i x_{ij} + \epsilon_i$$

$$F^* = \frac{(SSR_{Full} - SSR_{Red})/(p - q)}{SSE_{Full}/(n - p - 1)} \sim F_{p-q, n-p-1} \quad (38)$$

All of these statistical tests can also be performed for the MVLR problem as well. For a more in-depth read on Linear Regression and the Least Squares method, see [15].

3.4 Correlation analysis

Much like linear regression, correlation analysis quantifies the extent of dependence or association between two random variables, regardless of the causality relationship between them. Correlation allows to determine trends in the data, and therefore to make predictions of one variable given the other variable is known. For linear models, correlation is determined by the Pearson correlation coefficient (PCC):

$$\rho_{X_j Y} = \frac{\text{Cov}(X_j, Y)}{\sqrt{\text{Var}(X_j)\text{Var}(Y)}} = \frac{\sum_{i=1}^N (X_j^{(i)} - \mathbb{E}(X_j))(Y_i - \mathbb{E}(Y))}{\sqrt{\sum_{i=1}^N (X_j^{(i)} - \mathbb{E}(X_j))^2} \sqrt{\sum_{i=1}^N (Y_i - \mathbb{E}(Y))^2}} \quad (39)$$

$$-1 \leq \rho_{x_i y} \leq 1$$

Where $\rho_{X_j Y}$ is the CC between an input X_j and the output Y , N is the sample size, and k is the number of independent variables. $\text{Cov}(X_j, Y)$ is the covariance between X_j and Y , and $\text{Var}(X_j)$, $\text{Var}(Y)$ are the variances of X_j and Y , respectively, and \bar{x} and \bar{Y} are the sample means. A $\rho_{X_j Y}$ close to zero indicates the absence of a linear relationship, as illustrated in Fig. ???. Additionally, a negative coefficient means that the independent and dependent variables vary in opposite directions, just like in regression analysis. Note that a zero-valued $\rho_{X_j Y}$ does not indicate the existence of a well-defined nonlinear association between X_j and Y .

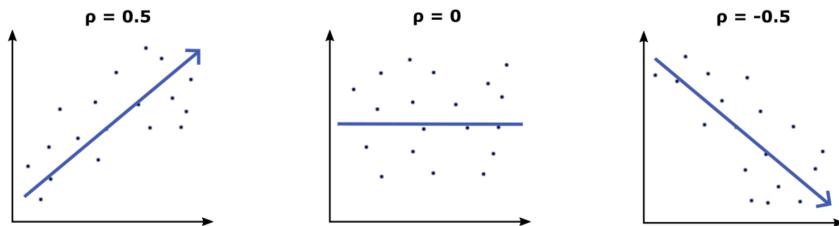


Figure 9: Schematic representation of correlation coefficients for three different types of data.

As previously mentioned, the Pearson correlation coefficient provides a measure of strength of linear association between X_j and Y , and therefore can be used as sensitivity indices to

rank variable importance. Unlike regression coefficients, which are influenced by the distributions assigned to X_j , correlation tends to exclude the effects of other independent variables, as well as their assumed distributions. If two random variables are statistically independent, the value assigned to one variable does not affect the probability distribution of the other. However, if correlations between the independent variables exist, this is no longer the case, and analyses based on the Pearson CC can be inaccurate. Note that, if X_j are all statistically independent, the PCCs and SRCs give the same variable importance rankings. Like in regression analysis, CC are more of a qualitative SA than quantitative.

To improve the accuracy of the PCC when one variable is numerically associated with X_i , partial correlation coefficients can be used instead. **Partial correlation** characterizes the linear relationship between X_j and Y after the linear effects of the remaining inputs are discounted. This is done by building two regression models:

$$\widehat{X}_{-j} = c_0 + \sum_{\substack{p=1 \\ p \neq j}}^n c_p X_p \quad (40)$$

$$\widehat{Y}_{-j} = b_0 + \sum_{\substack{p=1 \\ p \neq j}}^n b_p X_p \quad (41)$$

where \widehat{X}_{-j} denotes the prediction of the linear model expressing X_j in relation to the other inputs, and \widehat{Y}_{-j} is the prediction of the linear model without X_j . Thus, two new variables can be defined as $(X_j - \widehat{X}_{-j})$, and $(Y - \widehat{Y}_{-j})$, with the partial correlation coefficients being the CC between the residuals $(X_j - \widehat{X}_{-j})$ and $(Y - \widehat{Y}_{-j})$:

$$\hat{\rho}_{X_j Y} = \rho(X_j - \widehat{X}_{-j}, Y - \widehat{Y}_{-j}) = \frac{\text{Cov}(X_j - \widehat{X}_{-j}, Y - \widehat{Y}_{-j})}{\sqrt{\text{Var}(X_j - \widehat{X}_{-j}) \text{Var}(Y - \widehat{Y}_{-j})}} \quad (42)$$

The **T-test** statistic allows to test the significance of a CC (if it significantly differs from zero), and if two CCs are significantly different from each other:

$$T = r \sqrt{\frac{(N - 2 - p)}{1 - r^2}} \sim t_{N-2-p} \quad (43)$$

where T follows a Student's t distribution with $(N - 2 - p)$ degrees of freedom, and $p = 0$ for the Pearson CCs. Again, N is the sample size, and p is the number of inputs whose effects are discounted when calculating the CCs. The significance can be inferred from the corresponding p -value of the t distribution.

In order to test if two CCs (r_1, r_2) are significantly different from each other, the **Z -test** can be performed on log transformed CCs:

$$\dot{r} = \frac{1}{2} \ln \left| \frac{1+r}{1-r} \right| : N(\mu, \frac{1}{\sqrt{N-3-p}}) \quad (44)$$

where again p is the number of inputs whose effects are discounted when calculating the PRCCs, and $p = 0$ for the Pearson CCs. The log-transformation is used here to conform the data to the normal distribution. The significance can be inferred from the corresponding p -value of the normal distribution.

$$z = \frac{\dot{r}_1 - \dot{r}_2}{\sqrt{\frac{1}{N_1-3-p_1} + \frac{1}{N_2-3-p_2}}} : N(0, 1) \quad (45)$$

3.4.1 Non-linearities in the model

The above mentioned statistical test allow to test the validity of the linearity hypothesis. If the linearity between inputs and outputs cannot be assumed, the above mentioned sensitivity measures can still be used by applying a **rank transformation** to the data, providing that outputs vary monotonically with the inputs. The data (X^n, Y^n) is rank transformed to (R_X^n, R_Y^n) by replacing the values in each column by their respective ranks.

The equivalent of the Standard Regression Coefficient for rank transformed data is the Standard Rank Regression Coefficient (SRCC). Similarly, the Pearson CC becomes the Spearman Correlation Coefficients, or the Spearman Rho, and the PCC becomes the Partial Rank Correlation Coefficients (PRCC). PRCC is a robust sensitivity measure for non-linear but monotonic relationships between X_j and Y as long as little or no correlation exists between the inputs. Figure 10 shows an example of the different types of CCs of the same variables, showing that these can uncover substantially different trends in the data.

Other methods that can be used for nonlinear models are nonlinear regression, where the X_i variables are expressed as functions of nonlinear operators, such as exponentials and sinusoids. Another commonly used technique is nonparametric regression. These methods will not be covered in the present essay, but more detailed information can be found in [3, 8]

Graphical techniques can also be quite useful in SA for detecting nonlinear or other major relationships between the inputs and outputs. Therefore, inspection of scatterplots constitutes a natural starting point in the analysis of complex systems, and can provide insight into the most appropriate SA strategy. For instance, scatter plots allow to visually inspect the nature of the trends in the data, as exemplified in Fig. 10. Although scatter plots allow to infer the nature of the dependencies in the model, it does not provide information on the interaction effects among input variables. Therefore, it usually requires some level of knowledge of the most significant interactions in order to select the variables to display on the same 2D or 3D plot. Other types of plots, such as coweb plots [18], can help visualize common trends in the data by plot “simulation trajectories” of the input parameters, as shown in Fig. 11.

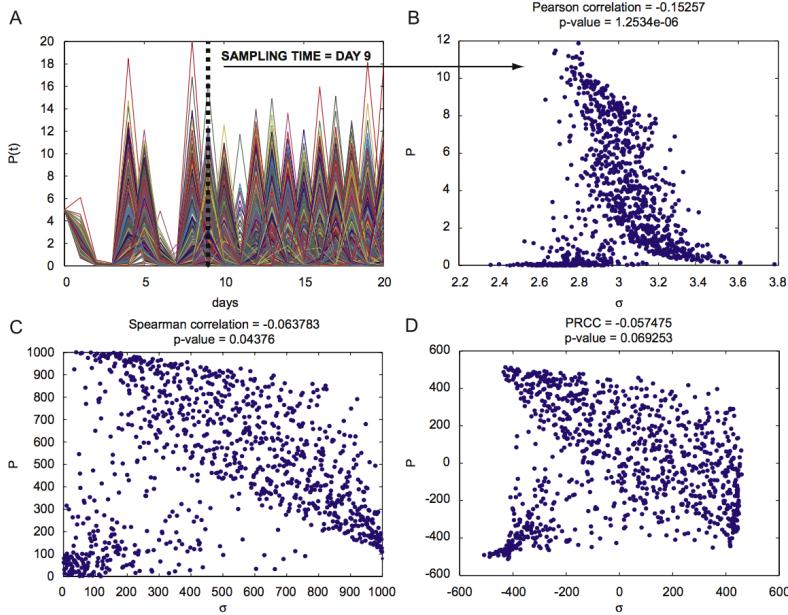


Figure 10: Plot of the output of the Lotka-Volterra model (A), in respect to one of its input variables, and scatter plots showing the Pearson CC (B), the Spearman CC (C), and the PRCC. Reprint from [16].

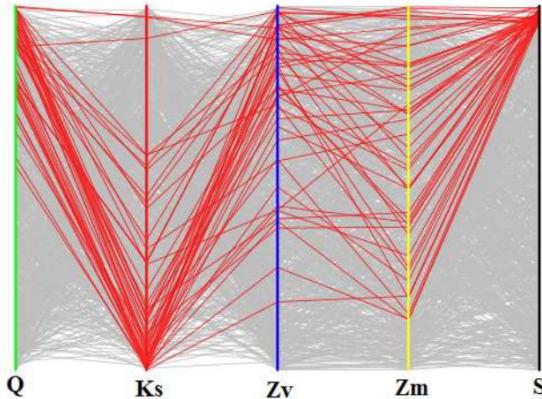


Figure 11: Example of a coweb plot showing a set of trajectories of the simulations corresponding to the 5% largest output values. Reprinted from [9].

3.5 Sobol indices

When models are non-linear and non-monotonic, the preferred choice pf SA method is the so-called **variance-based SA**. This method is based on *functional decomposition of variance*, whereby variance in the model output is decomposed into fractions of variance attributable to inputs of groups of inputs. This method (as well as the eFAST method presented in the next section) is closely related to analysis of variance (ANOVA), and so it falls into the general framework of **ANOVA-like decomposition** of factorial analysis [19]. Importantly, in contrast to regression and correlation analysis, variance decomposition methods can provide actual quantitative measures of sensitivity, and they are model independent in the sense they they do not require the model to be linear or monotonic to be effective.

There are several ways in which to variance can be decomposed, but once the variance fractions are known, it is possible to estimate sensitivity indices as ratios of the fractional variance of inputs to the total variance in the output. This framework was introduced by Ilia Sobol, and thus the sensitivity indices are known as Sobol Indices [20, 21].

Consider the mathematical model described by the function $Y = f(\mathbf{X})$, where $\mathbf{X} = (X_1, \dots, X_n)$ is a vector of input parameters defined in a unit n -dimensional unit cube denoted by I^n , such that each $0 \leq X_i \leq 1$. If $f(\mathbf{X})$ is integrable and defined in I_n , $f(\mathbf{X})$ can be decomposed in an expansion of orthogonal summands of different dimensions (i.e.,

dependencies on input parameters of different orders):

$$f = f_0 + \sum_i f_i(X_i) + \sum_{i < j} f_{ij}(X_i, X_j) + \cdots + f_{12\dots n}(X_1, \dots, X_n), \quad (46)$$

With total number of summands equal to 2^n , f_0 is constant and integrals of the $f_{i_1\dots i_s}$ are all zero (orthogonality). This expansion is unique if:

$$\int_0^1 f_{i_1\dots i_s}(x_{i_1}, \dots, x_{i_s}) dx_{i_k} = 0, \quad 1 \leq k \leq s, \{i_1, \dots, i_s\} \in \{1, \dots, n\} \quad (47)$$

$$1 \leq i_1 < \dots < i_s \leq n$$

It follows that f_0 is a constant:

$$\int f(\mathbf{X}) d\mathbf{X} = f_0. \quad (48)$$

where $f_0 = \mathbb{E}[Y]$ is a constant representing the expectation, or mean value, of the output. By assuming that all integrals of $f(\mathbf{X})$ are square integrable in I^n , we can find the relationship:

$$\int f^2 dX - f_0^2 = \sum_{s=1}^n \sum_{i_1 < \dots < i_s} \int f_{i_1\dots i_s}^2 dx_{i_1} \cdots dx_{i_s}. \quad (49)$$

The above relationship can be represented in terms of the variances:

$$\begin{aligned} \text{Var}(Y) &= \sum_{s=1}^n \sum_{i_1 < \dots < i_s} D_{i_1\dots i_s}(Y) = \\ &= \sum_{i=1}^n D_i(Y) + \sum_{i < j}^d D_{ij}(Y) + \cdots + D_{1,\dots,n}(Y) \end{aligned} \quad (50)$$

where $\text{Var}(Y)$ denotes the variance of Y , the term $D_i(Y)$ denotes the variance associated with variable X_i , and $D_{ij}(Y)$ denotes the variance associated with the interaction between variables X_i and X_j . The relationship above represents the partitioning of the total variance in the output ($D(Y)$) into components of increasing dimensionality ($D_{i_1\dots i_s}$). The terms on the first summation on the right-hand side of equation 50 represent the contribution from each X_i to the variance of Y ; the second summation represents the variance attributed to the interaction between different variables.

The Sobol method decomposes the variance observed in the model output \mathbf{Y} into the contributions associated to each of the input variables X_i , that is, their 'importance' to the variance in the output Y . The single impact of x_i is given by the conditional variance of Y ($D(Y)$) by fixing only X_i to a single value:

$$D_i(Y) = \text{Var}[\mathbb{E}(Y|X_i)] \quad (51)$$

One of the strengths of the Sobol method method is that it allows to take into consideration the variance associated with an input. The impact of the interaction between variables x_i and x_j is given by:

$$D_{ij}(Y) = \text{Var}[\mathbb{E}(Y|X_i, X_j)] - D_i(Y) - D_j(Y) \quad (52)$$

The variance of the distribution of the outputs can then be used to calculate sensitivity indices of Y to the inputs. These are called Sobol indices (SI). The SI are naturally derived by normalizing the decomposition of variances presented above to the variance in the output:

$$S_i = \frac{D_i(Y)}{D(Y)}, \quad S_{i_1 \dots i_s} = \frac{D_{i_1 \dots i_s}(Y)}{D(Y)} \quad (53)$$

These ratios are global sensitivity indices since they capture the variance attributed to the variables directly and via interaction with other variables. All $S_{i_1 \dots i_s}$ lie in the range $[0 : 1]$ and their sum is equal to 1:

$$\sum_{s=1}^n \sum_{i_1 < \dots < i_s} S_{i_1 \dots i_s} = \sum_i^n S_i + \sum_{i < j} S_{i \dots j} = 1 \quad (54)$$

The Sobol indices S_i are called "first-order sensitivity indices", or "main effect indices". They quantify how much the total variance of the model is reduced by *fixing* the i th parameter. Indices $S_{i \dots j}$ represent the higher order SI, quantifying the mixed effects of fixing x_i in the variance of the other variables; these can be used to assess the degree of interdependency of the variables.

The total-order sensitivity indices S_i^T correspond to the total part of the output variance attributed to each X_i , including the marginal effect and the interactions effects with other parameters:

$$S_i^T = S_i + \sum_{i < j} + \sum_{j \neq i, k \neq i, j < k} S_{ijk} + \dots = \sum_{l \in \#i} S_l \quad (55)$$

where $\#i$ are all subsets of $\{1, \dots, d\}$ that include i . For instance, in a model with 3 independent variables, the total SI of the parameters 1 would be $S_1^T = S_1 + S_{12} + S_{13} + S_{123}$. S_i^T also takes into account the variance attributable to interactions with other variables. The larger the S_i and the S_i^T , the more sensitive the model is to that variable. SI are analogous to the R^2 in regression, which represent the proportion of variance explained by parameter x_i .

The total-order indices encode the contributions of parameters through non-linear of interaction effects. If these interaction effects are negligible (all parameters are uncorrelated), then

$$\sum_{i=1}^n S_i \simeq 1. \quad (56)$$

In this case, the model is said to be *additive*. If the sum in eq. 56 is > 1 , then the interaction effects have to be considered.

There are several advantages of using variance-based methods for GSA:

- They allow to compute the sensitivities of the model across the whole input space;
- They can deal with non-linearities in the model;
- They take into account the interactions between input variables in non-additive systems.

In order to compute the Sobol indices of a process, each component of the variance has to be computed. The The Sobol Indices can be computed with Monte Carlo sampling [20, 22], but this method can be very computational costly for complex and non-linear models. Therefore, other schemes such as quasi-Monte Carlo, and FAST [23, 24] have been used to reduce the computational cost. With MC, the Sobol Indices can be directly computed from $f(x)$ by estimating the expected values of the quantities from multiple model evaluations. The

Sobol Indices can also be directly estimated from a surrogate model f' , with analytical solution of the integral components. For instance, the variance decomposition of the Polynomial Chaos Expansion presented above have an exact form, and therefore the SI can be computed analytically.

3.6 Fourier Amplitude Sensitivity Test

An alternative method to determine sensitivity indices based on decomposition of variance was proposed by Cukier et al.: the Fourier Amplitude Sensitivity Test (FAST) [25, 23]. This method was originally developed in the context of the analysis of coupled chemical reaction systems for the computation of the first order sensitivity indices. Later, Saltelli et al. [24] developed the extended FAST (eFAST) method based on the original FAST algorithm which allowed also to calculate total-order sensitivity indices. Similar to the Sobol method, FAST partitions the variance in the output among the contributions of each input variable. The key difference is that the FAST method uses a sampling scheme based on sinusoidal functions, and multiple Fourier series decomposition for the partitioning of the variance. One main advantage of the FAST and eFAST method over other variance-based methods is its lower computational cost due to the efficient sampling scheme it uses.

Consider again the model $Y = f(\mathbf{X})$, with $\mathbf{X} = \{X_1, X_2, \dots, X_k\}$. The method works by varying each model parameter at a different set of integer frequencies, which serves as a unique identifier of that parameter. In the first step, each input parameter is transformed into a new independent variable s , and sampled with a one-dimensional sinusoidal function called *search curve*. The search curve explores the entire n -dimensional input space with multiple parameter combinations.

$$x_l(s) = G_l(\sin w_l s), \forall l = 1, 2, \dots, n, -\infty < s < +\infty \quad (57)$$

where s is a scalar parameter, G_i is a transformation function, and $\{w_i\}$ are a set of different frequencies associated to each input factor. $f(\mathbf{X})$ can then be expressed as $f(s) = f(X_1(s), \dots, X_n(s))$. The exact form of $x_i(s)$ depends on the sampling distribution desired. Saltelli et al. [24] proposed the following as transformation function, although other functions

have also been proposed [25, 23]:

$$x_l(s) = \frac{1}{2} + \frac{1}{\pi} \arcsin(\sin(w_l s + \phi_l)), \forall l \in [1, n] \quad (58)$$

where N_s is the number of samples per parameter, and it is equal to $2Mw_{max} + 1$, where M is an interference factor (typically 4), and w_{max} is the maximum frequency assigned to the parameter. The original model $f(\mathbf{x})$ is therefore expressed as a function of a single independent variable: $y = f(s) = f(x_1(s), x_2(s), \dots, x_N(s))$.

A set of rules dictate which values of frequencies are allowed such that the search curve forms a closed path covering the entire input space. According to the Nyquist-Shannon sampling theorem, the frequencies must less than $1/2$ of the sampling frequency to avoid aliasing effects. Additionally, to ensure that the whole parameter space is explored, the set of frequencies w_l chosen for each x_i must not be a linear combination of the remaining frequencies, i.e., have no harmonics in common (incommensurate). If this is the case, the search curve covers the entire input space as s varies from 0 to $+\infty$, and according to the ergodic theorem, it is possible to express the variance of $f(\mathbf{X})$ (which is a multi-dimensional integral of (X_1, X_2, \dots, X_N)) as a one-dimensional integral of s :

$$\sigma^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f^2(s) ds - \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) ds \right]^2. \quad (59)$$

Additionally, if w_i are positive integers, then $f(s)$ is periodic in $(-\pi; \pi)$ [25]. The frequencies encoded in each parameter is then propagated through the model by performing several model evaluations with the parametric search curves. The frequencies can then be detected in the output by computing the Fourier coefficients of the model results. The Fourier coefficients can be computed as one-dimensional integrals, assuming the fundamental frequencies follow the properties explained above. The Fourier series of $f(s)$ and its coefficients are defined as:

$$f(s) = \sum_{p=-\infty}^{+\infty} \{A_{pw_l} \cos(pw_l s) + B_{pw_l} \sin(pw_l s)\} \quad (60)$$

$$A_{pw_l} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) \cos(pw_l s) ds \quad (61)$$

$$B_{pw_l} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) \sin(pw_ls) ds \quad (62)$$

$$p \in \mathbb{Z} = \{-\infty, \dots, -1, 0, 1, \dots, \infty\},$$

where A_{pw_l} and B_{pw_l} are the real and imaginary parts of the Fourier coefficients of the l th parameter. Note the infinite integral in the equation above as a consequence of the transformed function $f(s)$ being periodic in 2π . To evaluate the above integrals, the input space is sampled by drawing N values from $s_k = \frac{2k\pi}{N}, k = 1, \dots, N$.

According to Parseval's Theorem, the marginal and total variances, σ_i^2 and σ^2 , can then be estimated from the Fourier coefficients for each fundamental w_l and its harmonics, pw_l , as:

$$\sigma_l^2 = 2 \sum_{p=-\infty}^{+\infty} (A_{pw_l}^2 + B_{pw_l}^2) \quad (63)$$

$$\sigma^2 = 2 \sum_{j=-\infty}^{+\infty} (A_j^2 + B_j^2) \quad (64)$$

where $p \neq 0$ and $j \neq 0$, and $(A_j^2 + B_j^2)$ are the sum of squares of the Fourier coefficients of all integer frequencies (total variance). In practice, integer frequencies are only approximately incommensurate to an order of M , and thus A_{pw_l}, B_{pw_l} can only be approximated with a given error, $A_{pw_l}^*, B_{pw_l}^*$.

The partial variance of parameter X_l (the fraction of variance of the output due to the uncertainty in the l th parameter when the output is averaged over the uncertainties of all the other parameters $X_i, i \neq l$) is given by:

$$S_{w_l}^* = \frac{\sum_{p=-(N/2-1)}^{N/2} (A_{pw_l}^* |^2 + B_{pw_l}^* |^2)}{\sum_{j=-(N/2-1)}^{N/2} (A_j^* |^2 + B_j^* |^2)} = \frac{(\sigma_i^*)^2}{(\sigma^*)^2} \quad (65)$$

This is the main effects of the parameter X_l . The parameters with the largest influence on the output correspond to the highest frequency components in the frequency spectrum.

The **eFAST** method [24] was introduced as a variation of the well-established FAST to allow for the calculation of total-order effects of any input variable. In eFAST, each X_i is varied at a high frequency, while varying the remaining parameters at low and non-unique

frequencies. This provides an estimate of the contributions of the complementary set of i (that is, the variance attributable to all all parameters expect i , thus allowing to calculate the total-order sensitivity indices, S_i^T . The remainder portion of variance that that is not accounted for by the first-order or total-order indices corresponds to non-linear or interaction effects. This, S_i^T can also be used to infer the degree of additivity of the model.

The error made in the estimations of the variance components can be reduced by performing N_R resampling with different parameter search curves, each specified by a random phase shift, and then by taking the mean of the calculated variances. the eFAST algorithm is then repeated $k \times N_S \times N_R$ times to perform a full SA.

Marino et al. [16] proposed a method for determining the statistical significance of the sensitivity indices obtained with the eFAST method. They proposed to include of a dummy parameter in the analysis (in this case, a parameter non-existent in the model), which will result in a non-zero sensitivity index due to aliasing or interaction effects. Since in the eFAST method the input space is randomly sampled N_R times, it yields multiple estimates of the variance associated with each parameter. Then, a two-sample t -test can be used to compare the distributions of the estimated sensitivity measures S_i and S_i^T of each input parameter with those of the dummy parameter. The null hypothesis is that parameters with total-order sensitivity index less than or equal to that of the dummy parameter should not be considered significant. For further details see supplementary material of [16].

To illustrate how this method works, Fig. 12 shows the results of performing eFAST on the Lotka-Volterra of predator-prey dynamics from [16]. Input parameters σ and β are sampled with two integer approximately incommensurate frequencies (A), corresponding to normal distributions. The model is then solved for each input parameter combination (B). Obtain the Fourier coefficients of the output in (B) shows different spectral components (C), with the highest peaks corresponding to the frequencies of the varies input parameters. (D) To calculate the first-order sensitivity index of parameter σ , the values in the spectral analysis in (C) corresponding to frequency 31 are summed and normalized by the total variance. The sensitivity of the complementary set of parameters is calculated by summing the remaining components. The total-order sensitivity index is the sum of the first-order and higher-order

effects.

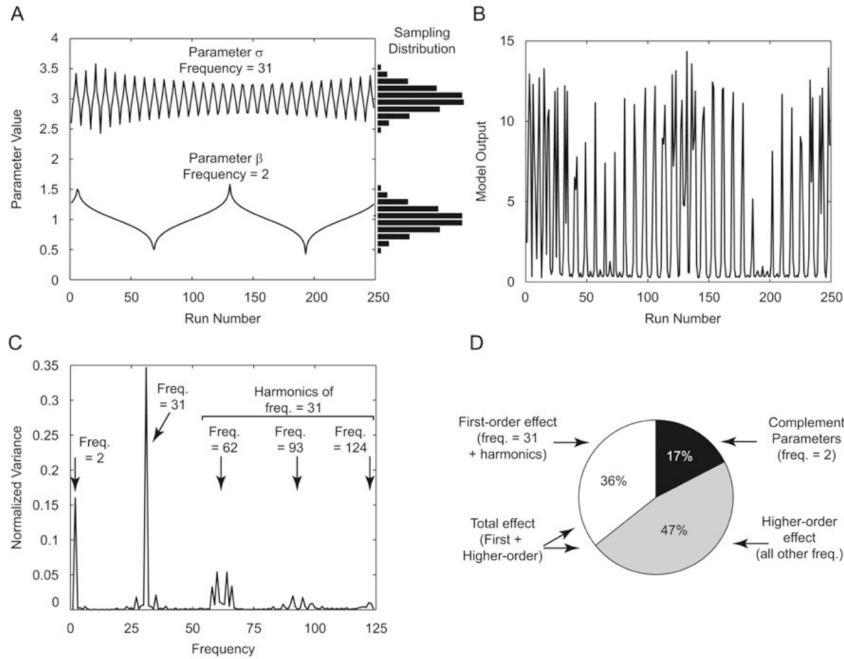


Figure 12: eFAST analysis of the Lotka-Volterra model output. Reprinted from [16].

4 Examples from cardiac modelling

SA frameworks have shown to be useful in helping researchers better understand the inner structure of the models for knowledge discovery and predictive modelling, with extensive applications in numerous mathematical models in different areas of study. In particular, the cardiac modelling field has seen extensive advancement over recent years thanks to the availability of experimental data, and the development of tools for solving differential equations in models of heart mechanics and electrophysiology (EP). Models of cardiac function have become multi-scale and highly complex in that they integrate models spanning different organizational levels, ranging from molecular (ion channels, and contractile machinery), to cellular, tissue and whole organ levels. Models of single cell function, for instance, can include dozens of differential equations coupled to several constitutive equations, and include over a hundred model parameters. These models have been widely used to advance our understanding of fundamental mechanisms of cardiac function and the pathophysiology of

heart diseases, to develop novel treatment strategies, and surgical guidance.

Physiological variability in cardiac EP is linked to several factors, such as, the stochastic nature of ion channel gating, natural cell variability (variable cell size, ion channel expression, etc.), and the non-linear behaviour of AP restitution. These manifest in variable ion channel function, calcium signalling, and electrical propagation in the tissue. Therefore, deterministic cardiac models that do not account for uncertainty in the parameters are prone to provide unreliable results and lead to misleading conclusions. Inherent uncertainties and variability should therefore be properly addressed to improve reliability of cardiac simulations and enable predictive modelling in clinical practice. In that respect, interest in UQ and SA has grown among the cardiac modelling community over recent years, with the importance of handling uncertainties and variability in cardiac modelling being addressed by some authors [26, 27].

In this section I will present a few examples of how UQ and SA has been applied in cardiac modelling to help answers question such as “*what are the most influential parameters in determining the observed mechanical and electrical behaviour of cardiac models?*”, “*what are the physiological ranges of parameters that result in a particular behaviour of interest?*”, and “*how does the choice of parameters affect the predictions of the model?*”.

In the work of Hurtado et al. [28], the authors used PCE and the elementary effects method to perform forward uncertainty propagation and sensitivity analysis on two coupled models of excitation-contraction of a 3D model stripe of a cardiac tissue. The authors looked at how variability in maximal conductances of major ionic currents affect selected QOIs, such as the action potential duration (APD90), calcium transient amplitude (CaT), cardiac stretch, and mechanical stress. Their results suggested a linear relationship between maximal conductances and the quantities of interest for a variability in parameters up to 25%. Therefore, the authors used linear response surfaces to compute the empirical probability density functions of model responses. They found that a larger variability of the model parameters resulted in a larger variability of the model outputs, as expected (Fig. 13). Furthermore, APD90 and CaT were mostly affected by the maximal conductances of L-type calcium current (ICaL), and the slow delayed rectifier potassium current (IKs), and

that stretch and stress were mostly influenced by ICaL alone, as shown in Fig. 14. This suggested a poor mechanoelectric feedback in models of cardiac electromechanics.

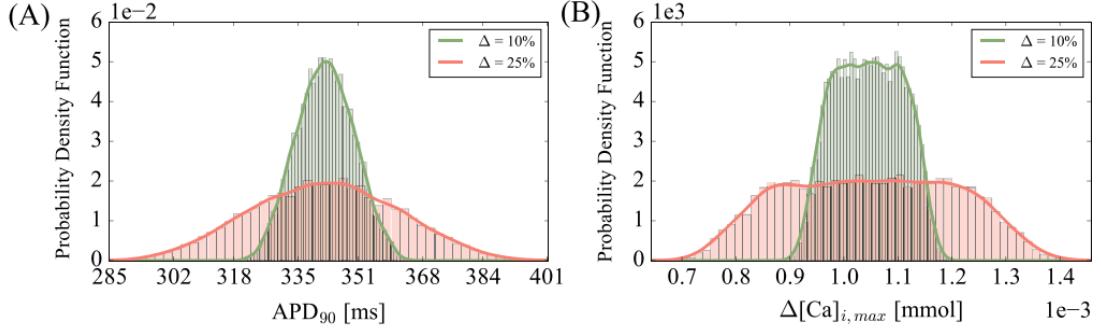


Figure 13: Empirical probability density functions of the electromechanical model corresponding to 10% and 25% parameter variability for APD90 (A), and CaT ($\Delta[Ca]_{i,max}$) (B). Reprinted from [28].

Models of single cell EP have also been subjected to SA, mainly to help researchers dissect the role of the several model parameters in the action potential (AP) and calcium transient morphologies under a varied set of conditions. Furthermore, it has been observed that these models are degenerate in that different sets of model parameters can produce similar AP and CaT shapes (Fig. 15), and thus SA can help guide the best choice of parameter sets. Examples in this area include the work of Sobie and Sarkar et al. [29, 30], Britton et al. [31], and Sanchez et al. [32]. In this approach, the authors used LHS to sample the parameter space, assuming a uniform distribution of the parameters, producing a large number of samples (up to 10.000). These samples are then used to perform an equal number of model evaluations to produce what is called *population of models* (PoM). Each model in the PoM outputs a number of QOIs, typically being APD90, CaT, resting membrane potential, upstroke velocity, and CaT duration. Because all these quantities can be measured experimentally, the populations are subsequently *experimentally calibrated* to eliminate unphysiological models from the subsequent analyses. Most commonly varied input parameters that are related to the ionic currents, and in particular, the maximal conductances since these are known to great affect the behaviour of EP models. The results of the model simulations are then linearly regressed onto the varied model parameters to

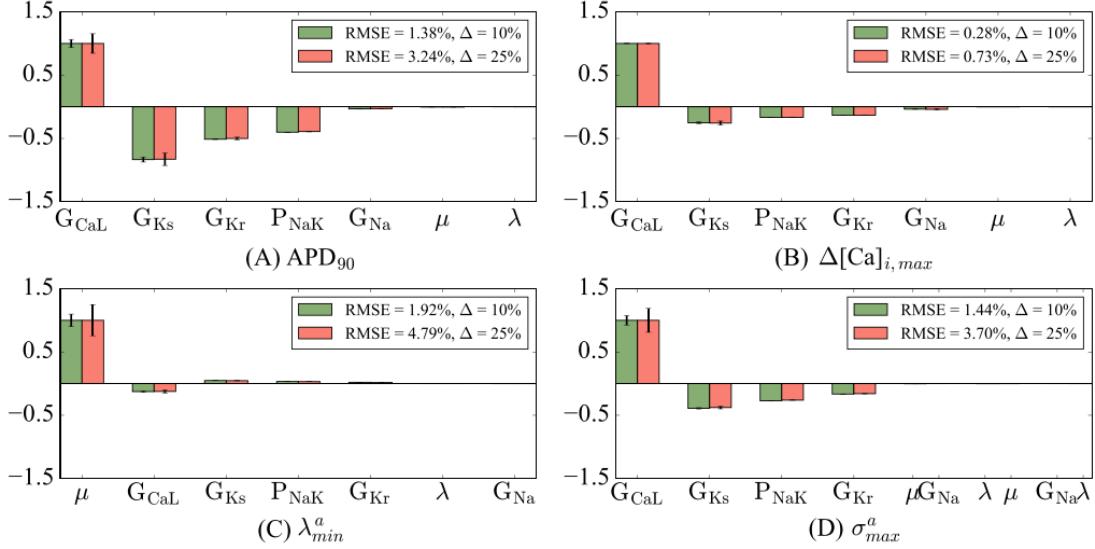


Figure 14: Parameter sensitivity measures μ^* (bar) and σ^* (error bars) from elementary-effects analysis of the electromechanical model: APD90 (A), CaT ($\Delta[Ca]_{i,max}$) (B), stretch (λ_a) (C), and stress ($\sigma_{a,max}$) (D). Reprinted from [28].

obtain regression coefficients that can be used to rank the parameters in order of their importance, as shown in Fig. 16.

Furthermore, output QOI can be used to segment calibrated PoM in accordance to distinguishable model behaviours, for instance APD90 above or below a certain threshold. This approach allows to test hypotheses related to the existence of significant differences in parameter distributions of both sub-populations. This analysis can provide useful insight into potential driving mechanisms of either behaviour. For instance, in [32] the authors constructed two PoMs of atrial cell models for three different cell models - the Maleckar, Courtemanche, and Grandi models; one PoM consisting of normal cells, and the other consisting of a diseased phenotype (CAF), and then compared the parameter distributions underlying each of the experimentally calibrated PoMs, revealing some significant differences among the two groups of parameters (Fig. 17), which can indicate the potential role of the underlying mechanisms in the development of the diseased phenotype .

In [33], Romero et al. performed local SA in a model of the human ventricular cellular electrophysiology to investigate how natural variability or drug-induced alterations in ionic

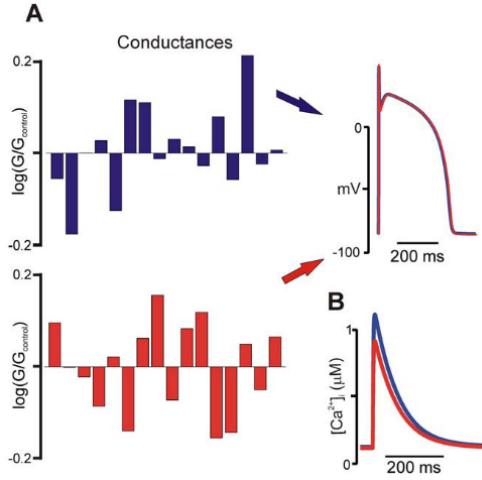


Figure 15: Non-uniqueness of EP model simulations, showing two different combinations of parameters of ionic conductances resulting in nearly identical action potential morphology (A). However, intracellular calcium transients produced by the same parameter combinations show some differences (B), suggesting that such information could be used to distinguish between the two parameter sets. Reprinted from [30].

current properties are related to arrhythmic risk. The authors varied the maximal conductances and time constants of the main transmembrane ionic currents involved in AP repolarization, one parameter at a time by -30% and -15%, +15% and +30%. Although performed as a local method, the analysis in this study provided some insights into the relative importance of the varied parameters to action potential properties, calcium and sodium dynamics, as well as their rate dependence.

In another example, Hu et al. [34] used the generalised PCE framework to perform forward UQ on multi-scale models of cardiac tissue with parametric uncertainty at the ion channel and cell levels. They looked at the effect of inserting uncertainties in several parameters related to two potassium channel on several multi-scale model outputs, such as steady state activation and inactivation of ion channels, AP duration, conduction velocity and spiral wave dynamics. Additionally, they performed variance-based SA with MC sampling to uncover the relevant parameters for these model responses in order to select which parameters to consider for uncertainty propagation. They also compared the accuracy and computational burden of the PCE method as compared to performing MC simulations, concluding

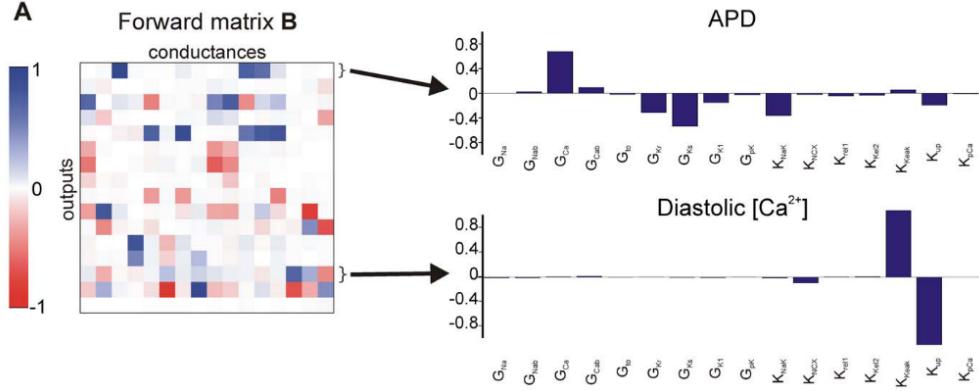


Figure 16: Parameter sensitivities obtained from regression analyses shown as “heat maps”, with white representing values near zeros, and blue and red indicating positive and negative values, respectively. Each row of the forward matrix \mathbf{B} represents the contributions of each of the conductances to a particular output, with APD and CaT shown in bar graphs to the right. Reprinted from [30].

that PCE was an much more effective method for computing parameters of the uncertainty distributions of the outputs (mean and variance), requiring only a fraction of the computational time required by using MC with 10.000 samples, while achieving comparable accuracy. As an example, Fig. 18 shows the uncertainty in steady state activation and inactivation of IK_{to} with 5% and 10% parameter variability achieved with both the PCE and MC methods.

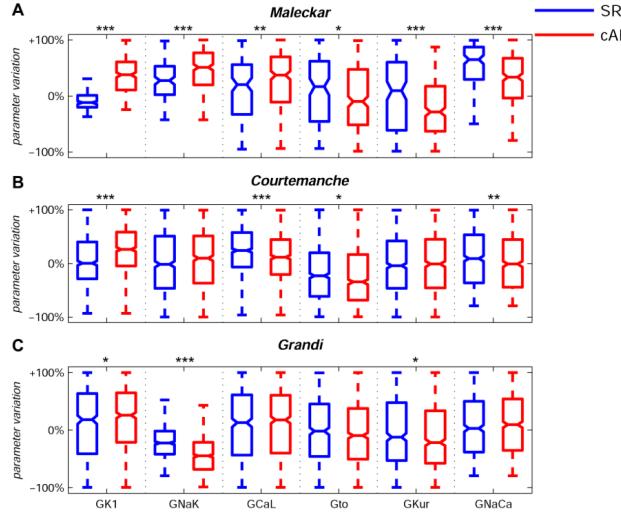


Figure 17: Boxplots showing the ranges of variability of ionic conductances of two distinct populations of models obtained with three different human atrial AP models: Malekar (A), Courtemanche (B) and Grandi (C) human atria AP models. The distinct populations correspond to in normal cells (blue) and diseased cells (red). Each boxplot represents the range covered by the ionic conductances. Reprinted from [32].

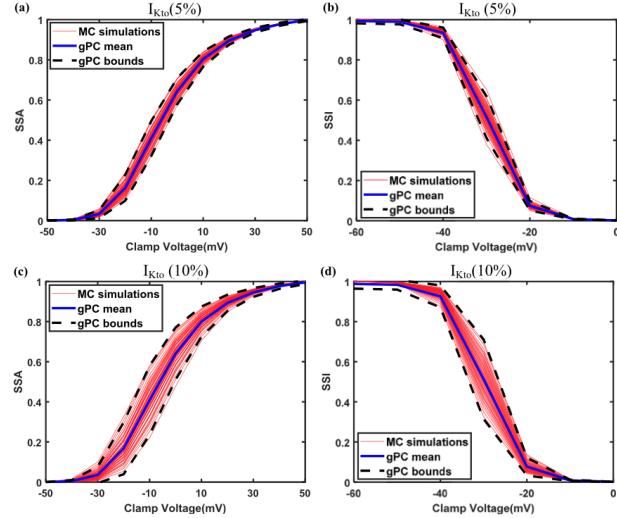


Figure 18: Comparison between the PCE method and MC simulation. (a) and (c) show the upper and lower bounds of the steady state activation curve of I_{Kto} , and (b) and (c) show the upper and lower bounds of the steady state inactivation curve of I_{Kto} . Reprinted from [34].

5 Synthesis and conclusions

The present essay aimed at providing an overview of some of the most commonly used uncertainty quantification and sensitivity analysis techniques. The extensive existing literature on UQ and SA, both regarding the development of methods and application to numerous real problems, is more than a convincing indicator of usefulness and applicability of these techniques. Therefore, an exhaustive review of the literature was out of the scope of this work. However, the methods presented here were selected so as to provide a wide enough range of different approaches and applicabilities.

The SA methods presented in this essay, although similar in some aspects, do not provide exactly the same information. Therefore, in certain instances it may be advantageous to perform more than one analysis, for instance, correlation and variance decomposition, in order to gain more insight into the sensitivities of the model. Furthermore, the availability of SA methods is vast and the choice of best method can become difficult. Some authors have reviewed the different SA methods, with the assumptions and applications implicit to each one of them, in an attempt to provide some guidance in choosing the SA method for the particular model at hand. A few examples of such synthesis are presented in Fig. 20 and 21 below.

Figure 19: Table comparing different SA methods based on several comparison criteria. Reprint from [9].

Criteria for comparison	Commonly used global sensitivity analysis methods				
	Weighted average of local sensitivity analysis (WALS)	Partial rank correlation coefficient (PRCC)	Multi-parametric sensitivity analysis (MPSA)	Fourier amplitude sensitivity analysis (FAST)	Sobol
Discrete inputs	Yes	Yes	Yes	Yes	Yes
Model independence	No	No	No	Yes	Yes
Non-linear, input-output relationship	Yes	Yes	Yes	Yes	Yes
Non-monotonic input-output relationship	Yes	No	Yes	Yes	Yes
Robustness	Yes	Yes	Yes	Yes	Yes
Reproducibility	Yes	Yes	Yes	Yes	Yes
Ability to apportion the output variance	No	No	No	Yes	Yes
Higher order interaction of parameters	No	No	No	Yes	Yes
Quantitative measure for ranking	Yes	Yes	Yes	Yes	Yes
Computational efficiency	Yes	Yes	Yes	No	No

Sensitivity analysis is an growing field of study, and its importance is becoming ever more

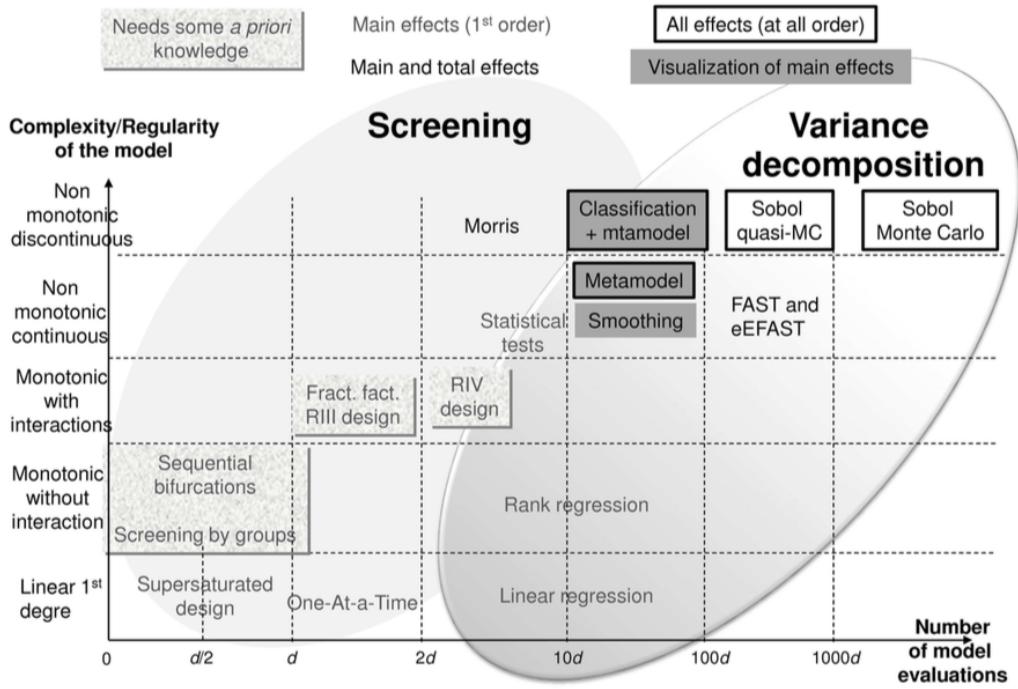


Figure 20: A graphical representation of different SA methods comparing their performance in terms of number of model evaluations required for different types of models. Reprint from [9].

evident with the increasing use of computational modelling in several different areas. In the field of cardiac modelling in particular, modellers are increasingly recognizing the importance of acknowledging the uncertainty and variability in model, and thus more and more UQ and SA are being included in model development routines. It is important to for modellers to have a level of understanding of how variability and uncertainty affect the model predictions and confidence in the results. Although it is not yet standard routine, the development of UQ and SA tools will potentially facilitate the adoption of these techniques in a wider setting of model developers, and with that contribute to more accurate and reliable computational models. The present essay aimed at giving a small step towards disseminating the knowledge of UQ and SA among non-experts.

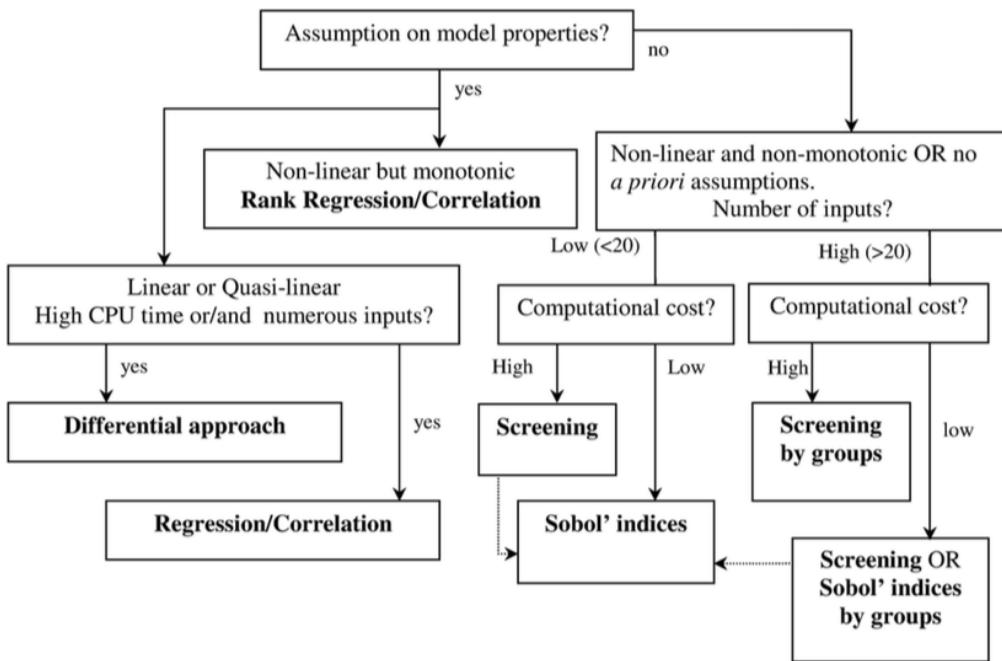


Figure 21: Decision diagram for choosing the most appropriate SA method based on type of model, and computational cost. Reprint from [35].

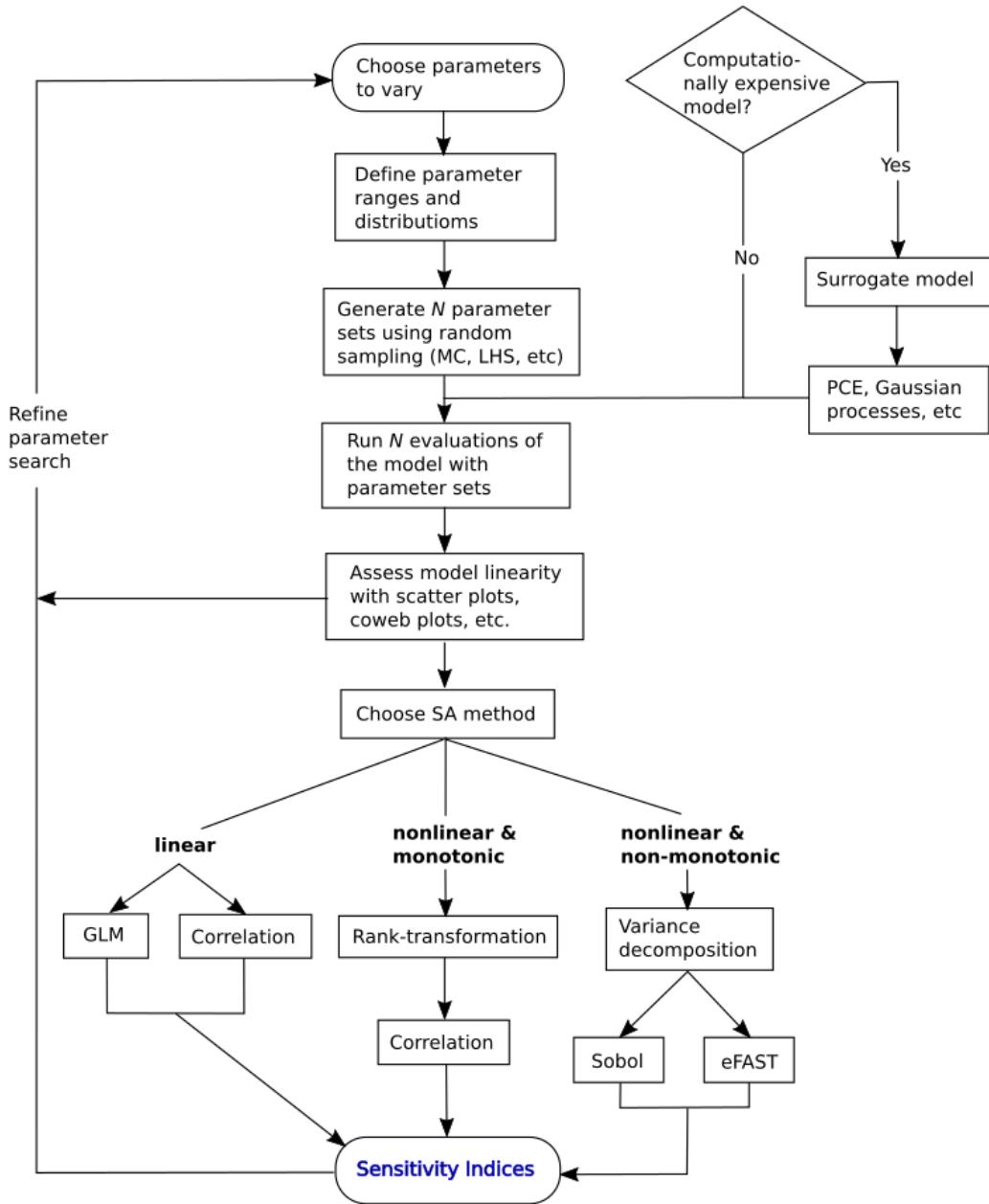


Figure 22: A flow diagram of the general SA framework.

References

- [1] D. G.-C. Lisbeth Geris, *Uncertainty in Biology, A Computational Modeling Approach.* [Online]. Available: <https://www.springer.com/gp/book/9783319212951>
- [2] B. Iooss, F. V. Dorpe, and N. Devictor, “Response surfaces and sensitivity analyses for an environmental model of dose calculations,” *Reliability Engineering & System Safety*, vol. 91, no. 10, pp. 1241 – 1251, 2006, the Fourth International Conference on Sensitivity Analysis of Model Output (SAMO 2004). [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0951832005002334>
- [3] J. Helton, J. Johnson, C. Sallaberry, and C. Storlie, “Survey of sampling-based methods for uncertainty and sensitivity analysis,” *Reliability Engineering & System Safety*, vol. 91, no. 10, pp. 1175 – 1209, 2006, the Fourth International Conference on Sensitivity Analysis of Model Output (SAMO 2004). [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0951832005002292>
- [4] M. de' Michieli Vitturi, S. Engwell, A. Neri, and S. Barsotti, “Uncertainty quantification and sensitivity analysis of volcanic columns models: Results from the integral model plume-mom,” *Journal of Volcanology and Geothermal Research*, vol. 326, pp. 77 – 91, 2016, numerical models of volcanic eruption plumes: inter-comparison and sensitivity. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0377027316300269>
- [5] L. Leemis. Univariate distributions relationships. [Online]. Available: <http://www.math.wm.edu/~leemis/chart/UDR/UDR.html>
- [6] N. Wiener, “The homogeneous chaos,” *American Journal of Mathematics*, vol. 60, no. 4, pp. 897–936, 1938. [Online]. Available: <http://www.jstor.org/stable/2371268>
- [7] Polynomial-chaos-kriging. [Online]. Available: <https://www.sudret.ibk.ethz.ch/research/past-projects/polynomial-chaos-kriging.html>
- [8] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola, *Global Sensitivity Analysis: The Primer.*

- [9] B. Iooss and P. Lemaître, “A review on global sensitivity analysis methods,” in *Uncertainty management in Simulation-Optimization of Complex Systems: Algorithms and Applications*, C. Meloni and G. Dellino, Eds. Springer, 2015. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-00975701>
- [10] A. Saltelli, S. Tarantola, and F. Campolongo, “Sensitivity analysis as an ingredient of modeling,” *Statistical Science*, vol. 15, no. 4, pp. 377–395, 2000. [Online]. Available: <http://www.jstor.org/stable/2676831>
- [11] A. Saltelli, S. Tarantola, F. Campolongo, and M. Ratto, *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*.
- [12] T. Turányi, “Sensitivity analysis of complex kinetic systems. tools and applications,” *Journal of Mathematical Chemistry*, vol. 5, no. 3, pp. 203–248, Sep 1990. [Online]. Available: <https://doi.org/10.1007/BF01166355>
- [13] M. D. Morris, “Factorial sampling plans for preliminary computational experiments,” *Technometrics*, vol. 33, no. 2, pp. 161–174, 1991. [Online]. Available: <https://www.tandfonline.com/doi/abs/10.1080/00401706.1991.10484804>
- [14] F. Campolongo, J. Cariboni, and A. Saltelli, “An effective screening design for sensitivity analysis of large models,” *Environmental Modelling & Software*, vol. 22, no. 10, pp. 1509 – 1518, 2007, modelling, computer-assisted simulations, and mapping of dangerous phenomena for hazard assessment. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1364815206002805>
- [15] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, ser. Springer Series in Statistics. New York, NY, USA: Springer New York Inc., 2001. [Online]. Available: <https://www.springer.com/gp/book/9780387848570>
- [16] S. Marino, I. B. Hogue, C. J. Ray, and D. E. Kirschner, “A methodology for performing global uncertainty and sensitivity analysis in systems biology,” *Journal of Theoretical Biology*, vol. 254, no. 1, pp. 178 – 196, 2008. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0022519308001896>

- [17] J. Troisi. Statistical tables.
- [18] D. Kurowicka and R. Cooke, *Uncertainty Analysis with High Dimensional Dependence Modelling*.
- [19] G. E. B. Archer, A. Saltelli, and I. M. Sobol, “Sensitivity measures,anova-like techniques and the use of bootstrap,” *Journal of Statistical Computation and Simulation*, vol. 58, no. 2, pp. 99–120, 1997. [Online]. Available: <https://doi.org/10.1080/00949659708811825>
- [20] I. M. Sobol, “Sensitivity estimates for nonlinear mathematical models,” *Matematicheskoe Modelirovanie*, vol. 2, 01 1990.
- [21] I. Sobolâš, “Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates,” *Mathematics and Computers in Simulation*, vol. 55, no. 1, pp. 271 – 280, 2001, the Second IMACS Seminar on Monte Carlo Methods. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378475400002706>
- [22] A. Saltelli, “Making best use of model evaluations to compute sensitivity indices,” *Computer Physics Communications*, vol. 145, no. 2, pp. 280 – 297, 2002. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0010465502002801>
- [23] R. Cukier, H. Levine, and K. Shuler, “Nonlinear sensitivity analysis of multiparameter model systems,” *Journal of Computational Physics*, vol. 26, no. 1, pp. 1 – 42, 1978. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/0021999178900979>
- [24] A. Saltelli, S. Tarantola, and K. P.-S. Chan, “A quantitative model-independent method for global sensitivity analysis of model output,” *Technometrics*, vol. 41, no. 1, pp. 39–56, 1999. [Online]. Available: <https://doi.org/10.1080/00401706.1999.10485594>
- [25] R. I. Cukier, C. M. Fortuin, K. E. Shuler, A. G. Petschek, and J. H. Schaibly, “Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. i

- theory,” *The Journal of Chemical Physics*, vol. 59, no. 8, pp. 3873–3878, 1973. [Online]. Available: <https://doi.org/10.1063/1.1680571>
- [26] W. Huberts, S. G. Heinen, N. Zonnebeld, D. A. van den Heuvel, J.-P. P. de Vries, J. H. Tordoir, D. R. Hose, T. Delhaas, and F. N. van de Vosse, “What is needed to make cardiovascular models suitable for clinical decision support? a viewpoint paper,” *Journal of Computational Science*, vol. 24, pp. 68 – 84, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1877750317307901>
- [27] G. R. Mirams, P. Pathmanathan, R. A. Gray, P. Challenor, and R. H. Clayton, “Uncertainty and variability in computational and mathematical models of cardiac physiology,” *The Journal of Physiology*, vol. 594, no. 23, pp. 6833–6847. [Online]. Available: <https://physoc.onlinelibrary.wiley.com/doi/abs/10.1113/JP271671>
- [28] D. E. Hurtado, S. Castro, and P. Madrid, “Uncertainty quantification of 2 models of cardiac electromechanics,” *International Journal for Numerical Methods in Biomedical Engineering*, vol. 33, no. 12, p. e2894, e2894 cnm.2894. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.2894>
- [29] E. A. Sobie, “Parameter sensitivity analysis in electrophysiological models using multivariable regression,” *Biophysical Journal*, vol. 96, no. 4, pp. 1264 – 1274, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0006349508032256>
- [30] A. X. Sarkar and E. A. Sobie, “Regression analysis for constraining free parameters in electrophysiological models of cardiac cells,” *PLOS Computational Biology*, vol. 6, no. 9, pp. 1–11, 09 2010. [Online]. Available: <https://doi.org/10.1371/journal.pcbi.1000914>
- [31] O. J. Britton, A. Bueno-Orovio, K. Van Ammel, H. R. Lu, R. Towart, D. J. Gallacher, and B. Rodriguez, “Experimentally calibrated population of models predicts and explains intersubject variability in cardiac cellular electrophysiology,” *Proceedings of the National Academy of Sciences*, vol. 110, no. 23, pp. E2098–E2105, 2013. [Online]. Available: <https://www.pnas.org/content/110/23/E2098>

- [32] C. S'anchez, A. Bueno-Orovio, E. Wettwer, S. Loose, J. Simon, U. Ravens, E. Pueyo, and B. Rodriguez, “Inter-subject variability in human atrial action potential in sinus rhythm versus chronic atrial fibrillation,” *PLOS ONE*, vol. 9, no. 8, pp. 1–14, 08 2014. [Online]. Available: <https://doi.org/10.1371/journal.pone.0105897>
- [33] L. Romero, E. Pueyo, M. Fink, and B. Rodríguez, “Impact of ionic current variability on human ventricular cellular electrophysiology,” *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 297, no. 4, pp. H1436–H1445, 2009, pMID: 19648254. [Online]. Available: <https://doi.org/10.1152/ajpheart.00263.2009>
- [34] Z. Hu, D. Du, and Y. Du, “Generalized polynomial chaos-based uncertainty quantification and propagation in multi-scale modeling of cardiac electrophysiology,” *Computers in Biology and Medicine*, vol. 102, pp. 57 – 74, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0010482518302622>
- [35] E. de Rocquigny, N. Devictor, and S. Tarantola, *Uncertainty in Industrial Practice: A guide to Quantitative Uncertainty Management*.