Proceedings of the 4th International Conference on Statistics: Theory and Applications (ICSTA'22)
Prague, Czech Republic – July 2022
Paper No. 138
DOI: 10.11159/icsta22.138

Statistical Inference for Optimisation of Drug Delivery from Stents

L. Mihaela Paun[1], **André Fensterseifer Schmidt**[2], **Sean McGinty**[2], and **Dirk Husmeier**[1]
[1]School of Mathematics and Statistics, University of Glasgow, Glasgow G12 8SQ, UK
[2]Division of Biomedical Engineering, University of Glasgow, Glasgow G12 8QQ, UK
Emails: Mihaela.Paun@glasgow.ac.uk; a.fensterseifer-schmidt.1@research.gla.ac.uk; Sean.Mcginty@glasgow.ac.uk
Dirk.Husmeier@glasgow.ac.uk;

Abstract The current study employs state-of-the-art optimisation methods for estimation of unknown parameters in a mathematical model of highly non-linear partial differential equations describing drug delivery from a drug-eluding stent. A classical optimisation scheme entails enormous run times due to the need to numerically solve the computationally expensive equations a large number of times to obtain the objective (black-box) function. We address this issue by employing an efficient global optimisation scheme, i.e. Bayesian optimisation (BO). This scheme aims to find the optimum of the black-box function by using an emulator of the original objective function to select the next query point (while balancing exploration and exploitation), and sequentially refining the emulator. Additionally, the proposed optimisation scheme is adapted to scenarios where there are hidden constraints in parameter space by incorporating a classifier that learns the infeasible parameter domains. We demonstrate that given a fixed number of expensive mathematical model evaluations, the proposed BO scheme outperforms state-of-the-art classical optimisation methods in terms of accuracy.

Keywords: statistical inference, Bayesian optimisation, Gaussian processes, emulation, classification, drug delivery, stents

1. Introduction

Computer simulations based on mathematical models are powerful tools that enable understanding of disease mechanisms, treatment development or medical device implantation. For models to become clinically applicable, estimation of unknown model parameters with efficient statistical tools is pivotal. The present study focuses on applying state-of-the-art optimisation methods, namely BO in a drug-eluting stent application.

In modern interventional cardiology, stent implantation has become a common technique to treat obstructive artery disease. To minimise the acute healing response after the intervention, and thus improve the outcome, stents now incorporate antiproliferative drugs. These are designed to be locally delivered at the lesion site, and are typically embedded into the stent platform via a polymeric coating covering the stent struts. Critical to the therapy's success is that drug is released to the arterial wall at both an adequate magnitude and rate. Specifically, while it is desired that drug levels are maintained below a toxic level ('safety'), it is also essential that the receptors of the target cells are saturated with drug for a sufficiently long period for the treatment to be successful ('efficacy'). The challenge of balancing safety with efficacy is what motivates the present work.

From an engineering perspective, the physical properties of the polymeric coating and the physicochemical properties of the drug may in principle be tailored to achieve a desired drug release profile. For example, the composition and thickness of the polymer coating may be adjusted to modulate the drug release profile with the former influencing the effective diffusion coefficient of the drug. Despite the existence of a large number of clinically-available stent platforms with widely varying drug release properties, the optimal drug release profile has yet to be defined.

Computational modelling and simulation provides a useful platform for medical device development. Following a sufficient description of underlying physics and biology, one may translate the clinical problem into a virtual '*in silico*' model, enabling hypothesis testing and performance evaluation for virtually any given scenario. To the best of our knowledge, the only existing published work which presents a framework for optimisation of drug delivery from stents was provided by Bozsak et al. [1]. In their work, they devised a cost function which sought to ensure therapeutic but sub-toxic drug levels in the media layer of the arterial wall, while minimising drug concentration at the endothelial surface, with the later aiming to avoid delayed healing. Whilst their model incorporates complex physics within a 2D-axisymmetric multi-layer arterial wall, it does not account for specific binding of drug to target receptors (instead accounting 'total' drug binding) and is therefore unable to assess specific receptor saturation of drug, an indicator strongly linked to efficacy [2] [3]. Moreover, the optimization methodology is not designed to handle cases of hidden constraints in parameter space, as is the case in our work. In this

proof-of-concept study, we address each of the aforementioned limitations within a simplified computational representation of the problem of interest. Concisely, we model drug elution from a non-erodible polymer coating on a stent strut and subsequent transport, specific and non-specific binding within arterial tissue. All of the model parameters are extracted from the literature, with the exception of the design parameters of interest, namely the polymer coating thickness, drug diffusion coefficient in the polymer and initial drug loading: these are the parameters we seek to optimise. Moreover, we employ an improved optimization methodology, harnessing the latest advancements in statistical surrogate modelling.

The challenge, from an optimization point of view, is that even a simplistic representation of the computational model entails enormous computational times when a classical iterative parameter optimisation scheme is employed. This is because a large number of optimisation iterations is needed, and at every iteration, the highly non-linear model partial differential equations need numerically solving several times to compute the objective (black-box) function and possibly, its derivative with respect to the parameters.

Recent methodological advancements in statistical surrogate modelling can be used to reduce computational costs by training a computationally cheap statistical emulator to replace the original objective function, as e.g. discussed in [4]. Our approach to building an emulator is based on Gaussian Processes (GPs) [5]. BO [6] can then be employed, with the aim to minimise the original objective function with the fewest number of expensive model evaluations. The optimisation is shifted to that of a computationally cheap acquisition function, constructed based on the emulated objective function, ensuring an exploration-exploitation trade-off [6].

A further complication in this work is that there are hidden constraints on the parameters, i.e. we require parameter values which produce a model output (maximum drug concentration value) below some set toxic threshold. To deal with this, our approach is to build a GP classifier that learns the infeasible parameter domain by assigning a 'probability of success' to parameter sets, and the acquisition function is accordingly modified.

We demonstrate that, given a fixed number of expensive model evaluations, the proposed Bayesian optimisation scheme outperforms state-of-the-art classical optimisation methods in terms of accuracy.

2. Applications

2.1. Toy problem

We first illustrate the performance of the BO method on a toy problem. We aim to minimise the following function:

$$f(x) = x^2 - 100, \text{where } x \in [-10, 5], \tag{1}$$

subject to the constraint that $x < -1$. The minimum of $f(x)$ is 0, but given the constraint, the optimum found should be -1.

2.2. Stent model description

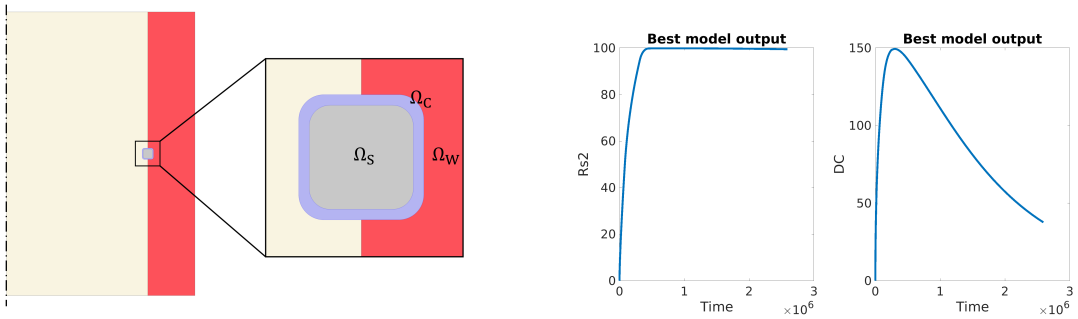
The model presented in this work represents drug delivery from one strut of a drug-eluting stent half-embedded into a 2D-axisymmetric single-layer arterial wall (panel (a) of Figure 1).

The model domains are: (i) strut (Ω_S), (ii) polymeric coating (Ω_C), and (iii) arterial wall (Ω_W). The strut domain is assumed to be impenetrable to drug and is therefore excluded from drug transport physics. Drug transport within the coating of thickness h_{coat} is described via a linear diffusion equation:

$$\frac{\partial c_{\text{coat}}}{\partial t} = D_{\text{coat}} \nabla^2 c_{\text{coat}}, \quad \text{in} \quad \Omega_C \tag{2}$$

where c_{coat} $[mol/m^3]$ is the concentration of drug in the coating, assumed to take the constant value c_0 initially and D_{coat} $[m^2/s]$ is the diffusion coefficient of the drug in the coating. Continuity of concentration and flux is imposed across the coating and interfaces. Thus, drug released from the coating is transferred to the arterial wall, where it is transported as a result of diffusion and advection and undergoes non-linear reversible binding reactions. The equation governing the concentration of free drug in the arterial wall, c_{wall}, is:

$$\frac{\partial c_{\text{wall}}}{\partial t} + \mathbf{u} \cdot \nabla c_{\text{wall}} = \nabla \cdot (D_{\text{wall}} \nabla c_{\text{wall}}) + R_1 + R_2, \quad \text{in} \quad \Omega_W \tag{3}$$



(a) (b)

Fig. 1: Panel (a): Schematic of the 2D-axisymmetric geometry employed, highlighting the stent strut Ω_S , drug-containing coating Ω_C and arterial wall Ω_w . The axis of symmetry is represented by the dotted line. The lumen, although simplified as an infinite sink of drug in this model, is illustrated in light yellow. Panel (b): Forward model output: receptor saturation, $Rs2$ defined in eq (6) and drug content, DC defined in eq (7) produced based on optimum parameter values found with Bayesian optimisation.

where D_{wall} $[m^2/s]$ is the diffusivity tensor for free drug in the wall, $\mathbf{u}$ $[m/s]$ is the transmural fluid velocity as a result of the pressure gradient across the wall (steady-state solution, provided by Darcy's Law), and R_1 and R_2 refer to drug lost from the free phase as a result of non-specific and specific binding. The non-linear reversible binding reactions are governed by:

$$R_1 = -\frac{\partial b_1}{\partial t} = -k_{1\text{ on}}\, c_{\text{wall}}\, (B_1 - b_1) + k_{1\text{ off}}\, b_1, \quad \text{in} \quad \Omega_W \tag{4}$$

$$R_2 = -\frac{\partial b_2}{\partial t} = -k_{2\text{ on}}\, c_{\text{wall}}\, (B_2 - b_2) + k_{2\text{ off}}\, b_2, \quad \text{in} \quad \Omega_W, \tag{5}$$

where b_1 and b_2 $[mol/m^3]$ refer to the concentrations of non-specific and specific bound drug, respectively. The parameters $k_{1\text{ on}}$, $k_{2\text{ on}}$ $[s^{-1}mol^{-1}m^3]$, $k_{1\text{ off}}$ and $k_{2\text{ off}}$ $[s^{-1}]$ are the respective binding rates while B_1 and B_2 $[mol/m^3]$ are the respective binding site densities. Drug transport within the lumen is not explicitly modelled, since rapid clearance occurs due to fast blood flow. Instead, the lumen-wall interface is split into 2 parts where drug may be lost to the lumen: the denuded endothelium (infinite sink for drug, adjacent to the strut) and healthy endothelium (semi-permeable membrane governed by the Kedem-Katchalsky equations). At the perivascular side of the arterial wall, an infinite sink condition is imposed.

The total drug content (DC) in the wall, expressed as μg drug/g tissue, is a common measure used to infer safety. Moreover, at least for the drug sirolimus, the number of specific receptors that are bound to drug ($Rs2$, expressed as a % of the total) has been linked to efficacy. These two time-dependent quantities are defined by:

$$Rs2 = \frac{1}{B_2 V_{\text{wall}}} \int_{\Omega_W} b_2 \, d\Omega_W, \tag{6}$$

$$DC = \frac{M_W}{\rho V_{\text{wall}}} \int_{\Omega_W} (c_{\text{wall}} + b_1 + b_2) \, d\Omega_W, \tag{7}$$

where M_W $[kg/mol]$is the molecular weight of the drug, while V_{wall} $[m^3]$ and ρ $[kg/m^3]$ are the volume and density of the tissue, respectively. The initial drug loading in the coating is given by:

$$m_{\text{drug}} = c_0\, M_W\, V_{\text{coat}}. \tag{8}$$

where V_{coat} $[m^3]$ is the volume of the coating domain. The current model is built to allow the iterative change of three design parameters within prescribed ranges, as indicated in Table 1. Altering these parameters directly affects the drug release dynamics of the model, and, therefore, the indicators of safety and efficacy. The goal in this work is to infer the design

Parameter name	Notation	Range
Initial drug loading	m_{drug}	$(5-25)\ \mu g$
Coating diffusion coefficient	D_{coat}	$(10^{-18}-10^{-13})\ m^2/s$
Coating thickness	h_{coat}	$(1-10)\ \mu m$

Table 1: Stent design model parameters to be estimated with Bayesian optimisation.

parameters that minimise the quantity

$$T - \int_0^T Rs2\ dt \tag{9}$$

with $Rs2$ defined in eq (6) and $T = 30$ days taken as the typical healing time, subject to the constraint that the maximum value of DC (defined in eq (7)) remains below some safety threshold (assumed to be 150 μg drug/g tissue here).

3. Statistical inference methods

3.1. Problem formulation

We formulate the statistical problem as follows: minimize $f(\boldsymbol{\theta})$ w.r.t. $\boldsymbol{\theta}$ subject to the constraint $g(\boldsymbol{\theta}) < T$, where in our case, $f(.)$ is given by the expression in eq (9), which is a functional depending on the parameters $\boldsymbol{\theta} = (m_{\text{drug}}, D_{\text{coat}}, h_{\text{coat}})$ of Table 1 that need estimating. The function $g(.)$ is given by eq (7), and T is a threshold set as explained in Section 2.2.

3.2. Classical optimisation

Classical constrained optimisation methods are reviewed in [7]. This includes sequential quadratic programming (SQP) [8], which is a local, gradient-based algorithm that uses derivatives of the black-box objective function to guide the search towards the local minima (where the gradient becomes small). Since it is a local optimisation algorithm, it can be run from several initial values (e.g. taken from a space filling design such as Sobol sequence [9]), and the lowest value of the objective function can be taken as the global minimum.

3.3. Bayesian optimisation

Bayesian optimisation [6] is a global, derivative-free method suitable for computationally expensive objective functions. The method efficiently searches for the optimum value of the black-box function by using and sequentially refining a surrogate model of the black-box function, based on which a computationally cheap function (acquisition function, AF) is constructed. The optimisation of the expensive objective function is now shifted to that of the computationally cheap AF, which gives the next query point following an exploitation-exploration principle, as explained in Section 3.3.2.

3.3.1. Gaussian Processes

In this work we build the surrogate model for the black-box objective function using GPs. A GP prior is placed on the distribution of the unknown black-box function to reflect uncertainty in the functional form. By allowing for noisy black-box evaluations $\mathbf{y}$ [1]:

$$\mathbf{y}|\mathbf{f}, \sigma^2 \sim \mathcal{N}(\mathbf{f}, \sigma^2\mathbf{I}), \quad \mathbf{f}(\boldsymbol{\Theta})|\boldsymbol{\gamma} \sim \mathcal{GP}(\mathbf{m}(\boldsymbol{\Theta}), \mathbf{K}|\boldsymbol{\gamma}), \tag{10}$$

where $\mathbf{f}$ is a vector of scalar-valued functions evaluated at the set of design points $\boldsymbol{\Theta}$, $\mathbf{m}(\Theta) = (m(\boldsymbol{\theta}_1), \ldots, m(\boldsymbol{\theta}_n))$ is the mean vector and $\mathbf{K} = [k(\boldsymbol{\theta}_i, \boldsymbol{\theta}_j)]_{i,j=1}^n$ is the covariance matrix of $\mathbf{f}$. $\boldsymbol{\gamma}$ contains the covariance function (kernel) hyperparameters, and σ^2 is the observational noise variance. The covariance function, $k(\boldsymbol{\theta}, \boldsymbol{\theta}'|\boldsymbol{\gamma})$, gives the smoothness and variability of the latent functions. Here, $D = \{\boldsymbol{\Theta}, \mathbf{y}\}$ is the set of training points, which sequentially expands during BO. The posterior predictive

[1] To allow for, e.g., errors from the numerical integration of the PDEs, or the fact that we add a small jitter term to the diagonal of the covariance matrix $\mathbf{K}$ for numerical stability during inversion.

distribution for the scalar-valued function $\tilde{f}$ at a new point $\tilde{\boldsymbol{\theta}}$ can be derived using Bayes rule:

$$\tilde{f}|D,\boldsymbol{\gamma},\sigma^2 \sim \mathcal{N}(m_{\mathrm{p}}(\tilde{\boldsymbol{\theta}}),k_{\mathrm{p}}(\tilde{\boldsymbol{\theta}},\tilde{\boldsymbol{\theta}}'|\boldsymbol{\gamma})), \tag{11}$$

$$m_{\mathrm{p}}(\tilde{\boldsymbol{\theta}}) = \boldsymbol{k}(\tilde{\boldsymbol{\theta}},\boldsymbol{\Theta}|\boldsymbol{\gamma})(\mathbf{K}+\sigma^2\mathbf{I})^{-1}\mathbf{y}, \tag{12}$$

$$k_{\mathrm{p}}(\tilde{\boldsymbol{\theta}},\tilde{\boldsymbol{\theta}}'|\boldsymbol{\gamma}) = k(\tilde{\boldsymbol{\theta}},\tilde{\boldsymbol{\theta}}'|\boldsymbol{\gamma}) - \boldsymbol{k}(\tilde{\boldsymbol{\theta}},\boldsymbol{\Theta}|\boldsymbol{\gamma})(\mathbf{K}+\sigma^2\mathbf{I})^{-1}\boldsymbol{k}(\boldsymbol{\Theta},\tilde{\boldsymbol{\theta}}'|\boldsymbol{\gamma}), \tag{13}$$

where $k(.)$ denotes a scalar-valued function and $\boldsymbol{k}(.)$ is a vector-valued function. Additionally, we train a GP classification model to model the probability of the hidden constraint being satisfied:

$$\lambda|g \sim \text{Bernoulli}(p(\lambda=1|g)), \quad \mathbf{g}(\boldsymbol{\Theta})|\boldsymbol{\gamma}^* \sim \mathcal{GP}(\mathbf{m}^*(\boldsymbol{\Theta}),\mathbf{K}^*|\boldsymbol{\gamma}^*), \tag{14}$$

where $\boldsymbol{\lambda}$ are the success labels {-1,1}, $\mathbf{m}^*(\Theta) = (m^*(\boldsymbol{\theta}_1),\ldots,m^*(\boldsymbol{\theta}_n))$ is the mean vector of $\mathbf{g}$, $\mathbf{K}^*$ is the covariance matrix, and $\boldsymbol{\gamma}^*$ are the kernel hyperparameters.

3.3.2. Acquisition function

Several acquisition functions (AFs) have been proposed in the literature [6, 10]. In this work we chose a widely used AF, the upper confidence bound (UCB), which in its original formulation has the following form [10]:

$$u(\tilde{\boldsymbol{\theta}}) = m_p(\tilde{\boldsymbol{\theta}}|D_l) + \beta_{l+1}^{\frac{1}{2}}\sqrt{k_{\mathrm{p}}((\tilde{\boldsymbol{\theta}},\tilde{\boldsymbol{\theta}})|D_l)} \tag{15}$$

at iteration l. Here D_l is the training set at iteration l; $m_p(\tilde{\boldsymbol{\theta}}|D_l)$ and $k_{\mathrm{p}}((\tilde{\boldsymbol{\theta}},\tilde{\boldsymbol{\theta}})|D_l)$ are the GP posterior predictive mean and variance, calculated using eqns (12) and (13); $\beta_{l+1} = 2\log\left(\frac{(l+1)^{\frac{d}{2}+2}\pi^2}{3\delta}\right)$, with d being the dimension of the new point $\tilde{\boldsymbol{\theta}}$; and $\delta = 0.1$ [10]. The UCB function takes higher values in regions of high objective function values, i.e. for large values of the GP posterior predictive mean $m_p(\tilde{\boldsymbol{\theta}}|D_l)$ (exploitation) and in regions of high uncertainty, i.e. for large values of the GP posterior predictive standard deviation $\sqrt{k_{\mathrm{p}}((\tilde{\boldsymbol{\theta}},\tilde{\boldsymbol{\theta}})|D_l)}$ (exploration).

3.3.3. Hidden constraints weighted acquisition function

To account for the hidden constraints in the parameter domain, we scale the UCB in eq (15) by the GP classifier-predicted probability that the proposed parameter vector returns a maximum DC value below the threshold, as follows:

$$u_{\mathrm{HCW}}(\tilde{\boldsymbol{\theta}}) = u(\tilde{\boldsymbol{\theta}})p(g(\tilde{\boldsymbol{\theta}}) < T). \tag{16}$$

3.3.4. Method outline

Below we give a brief method outline:

- Draw parameter values from a space filling design in parameter space (e.g. a Latin hypercube [11]), and perform an expensive model evaluation for each parameter vector to obtain $f(\boldsymbol{\theta})$ and $g(\boldsymbol{\theta})$ as indicated in Section 3.1. All training points are used to create a GP classifier, and only those points for which $g(\boldsymbol{\theta}) < T$ are used to create a GP emulator.
- Run the following for a fixed budget (number of black-box evaluations):
  - Find a new parameter vector $\tilde{\boldsymbol{\theta}}$ that maximises the acquisition function $u_{HCW}(.)$ in eq (16).
  - Perform an expensive model evaluation to obtain $f(\tilde{\boldsymbol{\theta}})$ and $g(\tilde{\boldsymbol{\theta}})$.
  - Add $(\tilde{\boldsymbol{\theta}}, \mathbb{1}(g(\tilde{\boldsymbol{\theta}}) < T))$ (where $\mathbb{1}$ is the indicator function) to the list of training points for the GP classifier.
  - If $g(\tilde{\boldsymbol{\theta}}) < T$, add $(\tilde{\boldsymbol{\theta}}, f(\tilde{\boldsymbol{\theta}}))$ to the list of training points for the GP emulator.
  - Refine GP emulator and classifier by re-optimising the kernel hyperparameters and observation noise variance.

4. Simulation study

The software used to run the optimisation routine is `Matlab` (for the toy problem) and `Livelink for Matlab` (for the stent application), where the latter interfaces Matlab and COMSOL Multiphysics (the stents model equations in Section 2.2

are solved numerically in `COMSOL Multiphysics 5.6`). The stationary part of the simulation employed a Multifrontal Massively Parallel Sparse (MUMPS) direct solver, with relative tolerance of 10^{-3}. The partial differential equations of the time-dependent part were solved with a Parallel Direct Sparse Solver (PARDISO), with relative tolerance of $5 \cdot 10^{-3}$. Transient steps were discretised with the Backward Differentiation Formula (BDF) method, with variable order of accuracy between 1 and 5. Meshing resolution in the coating domain was linked to the value of h_{coat}, guaranteeing sufficient mesh quality and convergence for all parameter values within the studied range. Values for the fixed model parameters (based on the drug sirolimus) and for the free stent design parameters were derived from the literature [2, 3, 12]. The `GPstuff` toolbox [13] was used to construct the GP models and run BO. Classical optimisation was run using the `fmincon` function in Matlab (with the SQP algorithm). Simulations were run on an Ubuntu machine with Intel(R) Xeon(R) Platinum 8276L CPU @ 2.20GHz with HT and 1TB RAM. One typical run of the model takes roughly one hour and a half.

The simulation set up for the optimisation is as follows: the classical optimisation was run from 5 different initial parameter values taken from a Sobol sequence, while for BO, we considered 5 initial designs, each with a different number of points: 20, 30, 40, 50 and 60 drawn from a Sobol sequence. We allowed a fixed budget of 90 black-box evaluations for the two optimisation schemes. After constructing the initial design with the parameter ranges given in Table 1, we found that with the limited number of initial points, there were no runs for which the maximum DC value was below the threshold. Based on this we refined the parameter ranges as follows: $m_{\text{drug}} \in [5, 10]$, $D_{\text{coat}} \in [10^{(-18)}, 10^{(-16)}]$, $h_{\text{coat}} \in [7, 10]$.

For BO, we used the SQP algorithm to minimise the UCB acquisition function with 72 initialisations. We added the global minimum, and the local minima (a maximum of 10 values) as training points to the GP emulator and classifier. Between iterations, the GP emulator and classifier were refined, i. e. kernel hyperparameters and likelihood noise variance optimised. For the GP emulator we used a squared exponential kernel and for the classifier a Matèrn 3/2 kernel, with the conditional posterior distribution being approximated with Laplace approximation.

5. Results and discussion

5.1. Toy problem

We first illustrate the performance of BO on the toy problem described in Section 2 .1. Using BO we minimise the AF plotted in the middle panel of Figure 2, and the optima points at every iteration are superimposed in the GP regression graph (left side panel) – the accumulation of points around the boundary (threshold). The latter graph shows that the optima points added as training points for the emulator are in the feasible domain, i.e. where $x < -1$. Thus, in this region, the uncertainty is low, and the emulator predicts the true function well (the GP mean is close to the true function). An accumulation of points is seen around the boundary (-1), and the global optimum is found very close to this boundary (as green dashed and red dashed lines are close together). This global optimum gives the minimum AF value in the middle panel. The right side panel displaying the probability of success for the variable x shows that the classifier learns the true constraint function, and the exploration around the boundary indicates that the BO method finds the global minimum of the constraint problem.

5.2. Stents application

Figure 3 shows the results obtained from carrying out both a classical optimisation (with the SQP algorithm) from 5 different initial parameter values, and our BO scheme with 5 different initial designs. Firstly, we note that after the 90 black-box evaluations (the allowed budget), the classical optimisation returns parameter values that are far from convergence. Generally, the objective function decreases very slowly, and the final parameter values highly depend on the initialisation. We note also that the classical optimisation algorithm spends a substantial number of black-box evaluations on approximating derivatives by numerically solving the mathematical equations. This explains the regular plateaus in the trajectory of the objective function.

In contrast, the trajectory of the incumbent minimum objective function value (i.e. the minimum value so far), which partially explains the plateau seen [2], shows a much more quickly decreasing pattern, and by 90 black-box evaluations, all 5 graphs reach a function value between 0.046 and 0.048. The graph shown as red line reaches the minimum value out of all 5, and we regard that as our best optimum. It appears that more iterations/evaluations are needed to improve convergence of

[2] A plateau can also be due to parameter values being proposed in the infeasible domain, which is more unlikely to happen as the accuracy of the classifier increases.

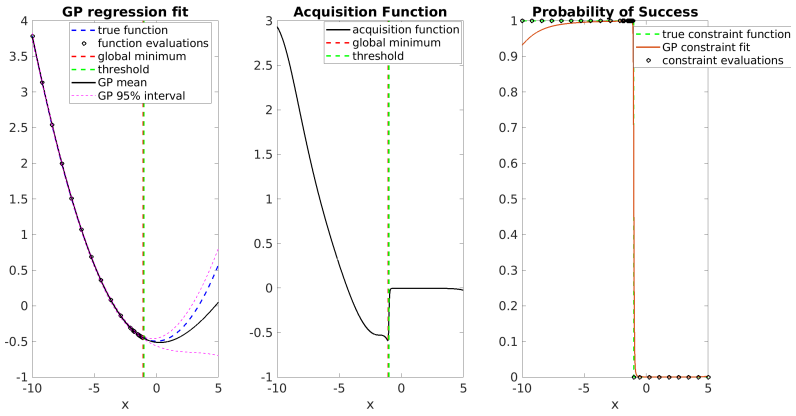


Fig. 2: Bayesian optimisation (BO) results for the toy problem. We show the GP regression fit (left side panel), and the acquisition function, which BO minimises (middle panel). The global minimum is also superimposed as a vertical dashed red line, found very close to the threshold (imposed by the constraint). The probability of success predicted by the classifier (right side panel) gives the feasible and infeasible domains for x. Note: the true and acquisition function values in the first two panels are scaled.

the BO scheme. Nevertheless, we often only have a set budget, and running the algorithm with that fixed budget helps modify future problem designs.

With the current fixed budget, the optimum parameter values $[m_{\text{drug}}, D_{\text{coat}}, h_{\text{coat}}] = [10, 4.5 \times 10^{(-17)}, 10]$ are on the boundary for two of the parameters. This aligns with an exploratory analysis that revealed monotonicity of the objective function (not shown here due to space constraints). Using the optimum parameter values, we ran the forward model and plotted the output, $Rs2$ and DC in Figure 1b. The optimal design parameters obtained are consistent with those of a slow release drug coating formulation and a moderate drug dose, consistent with the literature on sirolimus-eluting stents [2] [3], where it is expected that slow and sustained release is beneficial.

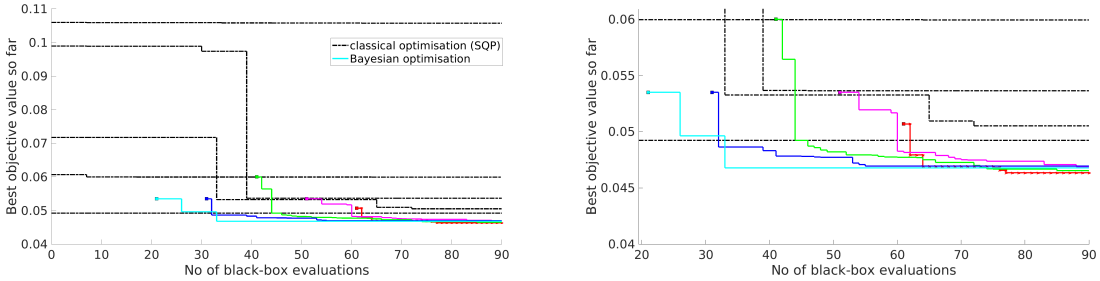


Fig. 3: Optimisation results for the stents application. A budget of 90 forward evaluations is allowed. We have run the classical optimisation from 5 different initial values and the trajectories are shown as black dashed lines in the left side panel. For the Bayesian optimisation (BO) we have taken 5 initial designs, each design with a different number of training points: 20 (corresponding to cyan colour), 30 (blue), 40 (green), 50 (magenta) and 60 (red), for which the forward model is evaluated. The remaining number of model evaluations (out of 90) is used as the number of BO iterations for the 5 BO simulations run, i.e. 70 (cyan), 60 (blue), 50 (green), 40 (magenta) and 30 iterations (red). The BO trajectories show the incumbent minimum, i.e. the best objective value so far. For clearer visibility, the panel on the right-hand side shows the incumbent minimum objective function trajectory zoomed in.

6. Conclusions

In this study we have used Bayesian optimisation to perform efficient st atistical in ference of un known pa rameters in a mathematical model describing drug delivery from drug-eluting stents. We have adapted the Bayesian optimisation scheme to handle hidden constraints in parameter space by incorporating a classifier that learns the infeasible parameter domains. Our results indicate that with a fixed budget (number of expensive forward model evaluations), the proposed Bayesian optimisation scheme is more efficient than state-of-the-art classical optimisation methods.

7. Limitations and future work

Future work includes an exploration of other acquisition functions for Bayesian optimisation (e.g. Expected improvement [6]). While this study has focused on optimisation, i.e. finding a point estimate, future investigations will focus on Bayesian uncertainty quantification. Given the high computational complexity of numerically solving the mathematical equations, an approach based on drawing samples from the approximate posterior distribution defined by a surrogate model for the mathematical model [14] could be employed. Further and substantial development of the stent model can provide a more realistic representation of the arterial environment, improving the relevance of drug transport and retention behaviours. Alternative objective functions and quantities of interest may be explored, as the understanding of indicators of safety and efficacy is deepened.

Acknowledgments

This work has been funded by EPSRC, grant reference number EP/T017899/1 (research hub for statistical inference in complex cardiovascular and cardiomechanic systems) and EP/S030875/1 (research centre for multiscale soft-tissue mechanics). We thank Agnieszka Borowska for her early contribution to this work.

References

[1] F. Bozsak, D. Gonzalez-Rodriguez, Z. Sternberger, P. Belitz, T. Bewley, J.-M. Chomaz, and A. I. Barakat, "Optimization of drug delivery by drug-eluting stents," *PloS one*, vol. 10, no. 6, p. e0130182, 2015.

[2] A. R. Tzafriri, A. Groothuis, G. S. Price, and E. R. Edelman, "Stent elution rate determines drug deposition and receptor-mediated effects," *Journal of Controlled Release*, vol. 161, no. 3, pp. 918–926, 2012.

[3] C. M. McKittrick, S. McKee, S. Kennedy, K. Oldroyd, M. Wheel, G. Pontrelli, S. Dixon, S. McGinty, and C. McCormick, "Combining mathematical modelling with in vitro experiments to predict in vivo drug-eluting stent performance," *Journal of Controlled Release*, vol. 303, pp. 151–161, 2019.

[4] L. M. Paun and D. Husmeier, "Markov chain Monte Carlo with Gaussian processes for fast parameter estimation and uncertainty quantification in a 1D fluid-dynamics model of the pulmonary circulation," *International Journal for Numerical Methods in Biomedical Engineering*, p. e3421, 2020.

[5] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005.

[6] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas, "Taking the Human Out of the Loop: A Review of Bayesian Optimization," *Proceedings of the IEEE*, vol. 104, pp. 148–175, 2016.

[7] M. Kochenderfer and T. Wheeler, *Algorithms for Optimization*. MIT Press, 2019.

[8] P. Boggs and J. Tolle, "Sequential quadratic programming for large-scale nonlinear optimization," *Journal of Computational and Applied Mathematics*, vol. 124, no. 1–2, pp. 123 – 137, 2000. Numerical Analysis 2000. Vol. IV: Optimization and Nonlinear Equations.

[9] P. Bratley and B. Fox, "Algorithm 659: Implementing Sobol's Quasirandom Sequence Generator," *ACM Trans. Math. Softw.*, vol. 14, pp. 88–100, Mar. 1988.

[10] Z. Wang, S. Mohamed, and N. de Freitas, "Adaptive Hamiltonian and Riemann Manifold Monte Carlo Samplers," in *Proceedings of the 30th International Conference on International Conference on Machine Learning - Volume 28*, ICML'13, pp. III–1462–III–1470, JMLR.org, 2013.

[11] M. D. McKay, R. J. Beckman, and W. J. Conover, "Comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 21, no. 2, pp. 239–245, 1979.

[12] N. Ding, S. D. Pacetti, F.-W. TANG, M. Gada, and W. Roorda, "Xience v™ stent design and rationale," *Journal of Interventional Cardiology*, vol. 22, pp. S18–S27, 2009.

[13] J. Vanhatalo, J. Riihimäki, J. Hartikainen, P. Jylänki, V. Tolvanen, and A. Vehtari, "GPstuff: Bayesian Modeling with Gaussian Processes," *J. Mach. Learn. Res.*, vol. 14, pp. 1175–1179, Apr. 2013.

[14] L. M. Paun, M. Colebank, M. Umar Qureshi, M. Olufsen, N. Hill, and D. Husmeier, "MCMC with Delayed Acceptance using a Surrogate Model with an Application to Cardiovascular Fluid Dynamics," in *Proceedings of the International Conference on Statistics: Theory and Applications (ICSTA'19)*, 08 2019.