Surrogate modelling and uncertainty quantification for multiscale simulation

Ye, D.

Publication date
2022

Citation for published version (APA):
Chapter 3

UQ of a three-dimensional ISR model with surrogate modelling\(^1\)

3.1 Introduction

In-stent restenosis is a recurrence of coronary artery narrowing due to vascular injury caused by balloon dilation and stent placement. It may lead to the relapse of angina symptoms or to an acute coronary syndrome. An UQ of a model for in-stent restenosis with four uncertain parameters (endothelium regeneration time, the threshold strain for smooth muscle cell bond breaking, blood flow velocity and the percentage of fenestration in the internal elastic lamina) is presented. Two quantities of interest were studied, namely the average cross-sectional area and the maximum relative area loss in a vessel. Owing to the high computational cost required for UQ, a surrogate model, based on GP regression with proper orthogonal decomposition, was developed and subsequently used for model response evaluation in the UQ.

Generally GPs are designed for a scalar output and become cumbersome when multi-output is required as a result of the large kernel used for coregionalization. The complexity of multi-output Gaussian process (MOGP) is

Chapter 3

associated with the dimension of output and the number of training samples. The computational cost of MOGP can easily become prohibitively expensive if the desired output dimension is high. One alternative solution is to apply dimensionality reduction techniques, such as proper orthogonal decomposition (POD) [110], to the model response before regression. The regression prediction is hence no longer the model response but the projection coefficients of the response. Due to the limited amount of projection coefficients required for the reconstruction of the output space, several single-output GPs are sufficient in this case. This method has been widely applied for time-dependent problems [111, 112], computational fluid dynamics [113], etc.

Here, the uncertainty propagation due to four uncertain parameters of the ISR3D model is investigated. The QoIs are the average cross-sectional area of the lumen and the maximum relative area loss as a function time. We applied POD to reduce the dimension of the output and used GP regression as the surrogate model to map the uncertain inputs to the projection coefficients of the POD. With this computationally efficient surrogate model, uncertainty estimations and sensitivity analysis of the restenosis process are conducted and analysed.

The chapter is arranged as follows. The details of the ISR3D model are introduced in Section 3.2. The construction of the surrogate model with POD and GP is described in Section 3.3. The uncertain parameters and uncertainty estimations are presented in Section 3.4. The results of uncertainty estimates and sensitivity analysis are presented in Section 3.5 followed by a discussion in Section 3.6 and the conclusions in Section 3.7.

3.2 Three dimensional in-stent restenosis model

ISR3D is a multiscale computational model simulating the post-stenting neointima growth in a coronary artery [89, 90]. It mainly consists of three single-scale submodels: the Initial Condition (IC) model, the Smooth Muscle Cells (SMC) model (including details of the vascular wall, such as lamina and the endothelial cells) and the Blood Flow (BF) model. A schematic diagram of ISR3D is shown in Figure 3.1.
3.2. Three dimensional in-stent restenosis model

Figure 3.1: A schematic diagram of the ISR3D model. After the initial deployment of the stent with the IC model, the SMC and BF models run iteratively until the end of the simulation. At every time step the SMC model passes the current lumen geometry to the BF model, which then updates the blood flow and sends the wall shear stress back to the SMC model. The SMC model then computes neointimal growth for the next time step, based on the wall shear stress.

The SMC model has two parts: one deals with the biomechanics of the vessel wall post-stenting, while the other deals with the SMC biology, mainly in relation to proliferation and production of extracellular matrix. The mechanical part of the SMC model simulates the mechanical response of the vessel wall, based on cell-cell pairwise repulsive and attractive forces and calculating the cell displacements. Each SMC of the vessel wall is modelled as a spherical agent, and the interactions between them are provided by potential and bond forces. The effective radii of particles represent the radii of corresponding cells and changes during the growth governed by the biological solver [90].

The growth of individual SMCs is modelled by a cell cycle model, similar to the one described in [114]. Each cell can be in a state of growth (G1), synthesis/secondary growth/mitosis (S/G2/M), or a quiescent state (G0). Cells
evolve from one state to the next according to their internal clock, and stop or die under the influence of external factors such as contact inhibition (the mechanical stresses in between SMCs) or the concentration of nitric oxide. The biological model provides new radii, states of the cells as its output, and also the initial coordinates for newly formed cells. The processes that influence the cell life cycle take place in the 30 $\mu m$ neighbourhood around the cell; the time scale of one cycle is around 24-48 hours. Growth of the neointima takes several dozens of cell cycles and stops several weeks after the stenting procedure [90].

The BF model is a pressure-driven fluid dynamics model, which provides relevant ranges of shear stresses on the vessel walls. The solver receives the lumen geometry every timestep from the SMC solver, simulates the steady-state blood flow and returns the wall shear stress information to the SMC model. The blood is assumed to be incompressible and Newtonian, and is modelled by the lattice Boltzmann method (LBM) [115] in a three dimensional rectangular mesh (D3Q19). The inlet boundary condition for velocity is set to a parabolic profile and its maximum velocity is defined as one of the uncertain parameters. A Dirichlet pressure boundary condition is assigned at the outlet and the vessel wall is defined as a non-slip condition. The simulation is implemented with Palabos [116].

The initial stent deployment is performed by IC model. The stent is expanded radially with a capsule-shaped balloon until it reaches a predefined deployment depth. As there is no uncertain input of the UQ experiment related to the IC model and all the simulations start from exactly the same post-deployment state, we exclude the IC model from the execution of the UQ. For further details about the ISR3D, see [89, 90]. A public version of the ISR3D model, which is studied in this work, can be found on Github$^2$. In the UQ experiments described here, the scenario of stenting a small porcine coronary vessel with 2 $mm$ diameter is simulated. The simulation domain is limited on the outside by the middle layer of the artery wall, the tunica media. The entire length of the vessel is set to be 18 $mm$ with a tunica media thickness of 0.35 $mm$ and 1 $mm$ lumen radius. The entire vessel is

$^2$https://github.com/ISR3D/ISR3D
3.2. Three dimensional in-stent restenosis model

Figure 3.2: The simulation outcomes of ISR3D model (left) and corresponding QoIs measured over time (right). On the left-hand side, the blue part denotes the vessel wall, the beige part is the stent and the red part denotes the neointima. On the right-hand side, the average cross-sectional area and maximum relative area loss of the vessel at each day are measured and computed.

assumed to be slightly curved to obtain a more realistic blood flow pattern in the vessel. The stent applied in the simulations is made of intersecting spiral elements (shown in Figure 3.2 left). It can be viewed as a simplified version of the NIR stent [117] and the deployment depth is set to be 0.25 mm. The model is set to simulate the restenosis process up to 30 days after stenting.

The computational cost of ISR3D with a vessel and a stent of this size is rather expensive. A single run of the ISR3D simulation takes 500 to 600 core hours on a supercomputer node (a node with $2 \times 12$-core 2.6 GHz Intel Xeon E5-2690 v3 CPUs), depending on the total amount of neointima growth. For non-intrusive UQ methods, a large number of evaluations of the model are required for the statistical analysis and this becomes impractical for such a computationally expensive model. Therefore, to perform the UQ efficiently, a data-driven surrogate model based on GPs and POD is developed to learn the latent function between the uncertain inputs and the QoIs, and applied to evaluate the model response in the UQ.

Two QoIs are measured in the UQ experiment: the average cross-sectional area of the vessel lumen, and the maximum relative area loss. The lumen cross-sectional areas along the centerline of the vessel are obtained using an
open-source toolkit VMTK\(^3\). The average value of this area over the considered vessel model at each time step are used to evaluate how the uncertain parameters influence the total amount of neointima growth over time (shown in Figure 3.2 right). The relative area loss of the vessel shows the relative amount of neointima growth compared with the initial post-stenting cross-sectional area. Clinically, the restenosis is defined as the renarrowing of the lumen to more than 50\% occlusion [81]. The maximum value of relative area loss of a vessel offers us a criterion to judge whether the restenosis happens or not. Note that both QoIs are evaluated as a function of time. The values at consecutive timesteps are highly correlated with each other.

### 3.3 Surrogate modelling

#### 3.3.1 Proper orthogonal decomposition on model response

Assume the response of the model is a series of responses (here, average cross-sectional areas of the lumen) over time \( \mathbf{y} \in \mathbb{R}^{N_t} \), where \( N_t \) is the dimension of the output vector. The POD method can be applied to approximate the model responses by projecting the response to a low-rank space. The POD can be realized in three schemes, Karhunen-Loeve decomposition, principal component analysis, and the singular value decomposition (SVD). In this work, the SVD method is applied for the decomposition [118].

Consider a snapshot matrix \( \mathbf{S} \in \mathbb{R}^{N_t \times N_s} \) consisting of \( N_s \) number of the model responses \( \{ \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{N_s} \} \):

\[
\mathbf{S} = [\mathbf{y}_1 | \mathbf{y}_2 | \ldots | \mathbf{y}_{N_s}], \quad \text{where } N_t \gg N_s.
\]  

The snapshot matrix can be decomposed into three matrices using singular value decomposition:

\[
\mathbf{S} = \mathbf{U} \Sigma \mathbf{V}^T,
\]

where \( \mathbf{U} \) and \( \mathbf{V} \) denote the left and right orthonormal matrices, respectively. \( \Sigma \) denotes a diagonal matrix with singular values \( \sigma_i \), where \( i = 1, \ldots, N_s \) and

\(^3\)http://www.vmtk.org/
3.3. Surrogate modelling

\[ \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_{N_s} \geq 0. \]

The objective of POD is to find out a set of orthogonal bases \( \tilde{\Phi} = \{ \phi_1, \phi_2, \ldots, \phi_k \} \) from the space \( \mathcal{L} = \{ \Phi \in \mathbb{R}^{N_t \times k} : \Phi^T \Phi = I \} \) containing all possible orthogonal bases, such that the error introduced by the projection to low dimensional space could be minimized,

\[
\min_{\Phi \in \mathcal{L}} \sum_{i=1}^{N_s} \| y_i - \Phi \Phi^T y_i \|_{L^2}^2,
\]

where \( \| \cdot \|_{L^2} \) denotes the \( L^2 \) norm. By the Eckart-Young theorem [119], the orthogonal basis with the basis vectors \( \{ u_i \}_{i=1}^k \) taken from the \( i \)th column of \( U \) is the solution to such an optimization problem. The relative energy captured by the projection to such low dimensional space consisting of the first \( k \) columns of \( U \) can be evaluated by [120]:

\[
R_{en} = 1 - \frac{\sum_{i=1}^{N_s} \| y_i - \tilde{\Phi} \tilde{\Phi}^T y_i \|_{L^2}^2}{\sum_{i=1}^{N_s} \| y_i \|_{L^2}^2} = \frac{\sum_{i=1}^{k} \sigma_i^2}{\sum_{i=1}^{N_s} \sigma_i^2}.
\] (3.4)

We assume that, if the relative energy \( R_{en} \) is higher than 99.9\%, the approximation reconstructed by the first \( k \) bases performs well enough. Since the values of \( \sigma_i \) decay rapidly, a small \( k \) would be sufficient to achieve the relative energy threshold. Once the basis vectors are obtained, any model response can be approximated by:

\[ y \approx \hat{y} = \sum_{i=1}^{k} \alpha_i \phi_i, \]

where \( \alpha_i \) are the projection coefficients.

3.3.2 Gaussian process regression

Assume that a model response \( y \in \mathbb{R} \) is generated by the function \( y = f(x) + \epsilon \) with a corresponding input \( x \in \mathbb{R}^d \), and \( \epsilon \) denotes the noise of the measurement or stochasticity of the model and assumes that a normal distribution is followed: \( N(0, \sigma_n^2) \). A GP can be defined as a collection of random variables and any finite number of the random variables follows joint Gaussian distribution [98],

\[
f(x) \sim GP(m(x), k(x, x')),
\] (3.5)
where $\mathcal{GP}$ denotes a GP prior over the space of functions specified by its mean function $m(x)$ and covariance functions $k(x, x')$. Generally the mean function is set to be zero without loss of generality. The kernel functions specify how the random variables are correlated with each other and also imply the smoothness of the functions. One of the common choices is the radial basis function kernel with automatic relevant determination (ARD) [98],

$$k(x, x') = \sigma_f^2 \exp \left( -\frac{1}{2} \sum_{i=1}^{d} \frac{(x_i - x_i')^2}{\ell_i^2} \right),$$  

(3.6)

where $\sigma_f^2$ is the signal variance and $\ell_i$ denotes the lengthscale for each input dimension. For a regression problem, an independent Gaussian kernel with variance $\sigma_n^2$ is used to specify the noise in the function. These hyperparameters in the kernel will be determined via the optimization of likelihood function with observed data collection $(X, y) = \{(x_i, y_i)\}_{i=1}^{N}$,

$$\arg \max_{\theta} \log p(y|X, \theta) = \arg \max_{\theta} \left[ -\frac{1}{2} y^\top (K + \sigma_n^2 I)^{-1} y - \frac{1}{2} \log |K + \sigma_n^2 I| - \frac{n}{2} \log 2\pi \right],$$  

(3.7)

where $\theta = \{\sigma_f, \sigma_n, \ell_1, \ldots, \ell_d\}$ and $K = k(X, X)$. To predict the model response at an unevaluated location $x^*$, the GP prior can be rewritten into:

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

(3.8)

Conditioning on the observed data, the predictive distribution of the new point $x^*$ also follows a normal distribution,

$$f(x^*)|X, y, x^* \sim \mathcal{N}(\bar{y}^*, \text{Var}(y^*)), $$  

(3.9)

where

$$\bar{y}^* = k(x^*, X)[k(X, X) + \sigma_n^2 I]^{-1} y,$$

$$\text{Var}(y^*) = k(x^*, x^*) - k(x^*, X)[k(X, X) + \sigma_n^2 I]^{-1} k(X, x^*).$$
The $\bar{y}^*$ stands for the mean prediction of the response and $\text{Var}(y^*)$ is the predictive variance indicating the uncertainty of the prediction.

Generally, the GP regression is applied as a surrogate model to infer the latent function between uncertain inputs and QoIs. However after the decomposition of the model response by SVD, both evaluated and unevaluated model responses can be represented by the projection coefficients on the chosen orthogonal bases, therefore the GP is now used to learn the mapping between uncertain inputs and projection coefficients of POD and predicts the new coefficients for unevaluated points. The details of the procedure are shown in Algorithm 3.1.

\begin{algorithm}
\caption{Constructing a surrogate model for ISR3D with GP and POD}

\textbf{Training}
1. Evaluate $N_{\text{train}}$ number of samples using ISR3D and collect the training data $\{(x_i, y_i)\}_{i=1}^{N_{\text{train}}}$.
2. Construct the snapshot matrix $[y_1 | y_2 | \cdots | y_{N_s}]$ and perform SVD to obtain $k$ orthogonal bases $\tilde{\Phi}$ based on the relative energy threshold.
3. Project the output of the training data to each basis and compute the projection coefficients: $\{(x_i, y_i)\}_{i=1}^{N_{\text{train}}} \xrightarrow{\text{POD}} \{(x_i, \alpha_j(x_i))\}_{i=1}^{N_{\text{train}}}$, where $j = 1, \cdots, k$.
4. Train $j$-th single-output GP with uncertain inputs and projection coefficients $\{(x_i, \alpha_j(x_i))\}$.

\textbf{Prediction}
1. For an unevaluated point $x^*$, use GPs to predict its projection coefficients $\alpha_j(x^*)$, where $j = 1, \cdots, k$.
2. Reconstruct the corresponding model response $y^* = \sum_{j=1}^{k} \alpha_j(x^*)\phi_j$.
\end{algorithm}

3.4 Uncertainty quantification

3.4.1 Uncertain parameters

The four epistemic uncertainties considered in the forward uncertainty propagation of ISR3D include endothelium regeneration time, blood flow velocity, the threshold strain for SMC bond breaking, and the percentage of fenestration
by area in the IEL. Note that all the uncertain parameters except the blood flow velocity are parameters of the SMC submodel.

**Endothelium regeneration time**: The endothelium regeneration starts right after the denudation caused by the balloon dilation and stent deployment. With sufficiently high wall shear stress from blood flow, the endothelium releases nitric oxide, which behaves as the inhibitor of the proliferation of SMCs. Therefore the rate of endothelial regrowth significantly influences the growth of neointima. In the ISR3D, the regeneration of endothelium cells is modelled to increase linearly up to a coverage of 59% after 3 days, followed by a full recovery to 100% after a certain number of days given by the uncertain input [89]. This setting is based on experimental results from Nakazawa et al. [121]. However, the exact time for re-endothelialization may vary with many factors, such as the severity of vessel injury, the types of stenting and the degrees of inflammatory response [122]. In order to study this uncertain parameter, we consider an average endothelium regeneration time of 15 days based on the experimental data from Nakazawa et al. [121], and vary it from 10 to 20 days in the UQ.

**Threshold relative strain**: The threshold strain is the maximum strain that can be obtained before the bonds between SMCs break. Generally, during the stenting process, the vessel wall is overstretched in the circumferential direction, and therefore the connections between the SMCs (e.g. collagen fibers) are possibly broken and cause microfractures in the tissue. These microfractures may cause inflammation and contribute to the proliferation of SMCs after stenting.

Our choice on the uncertainty of the breaking strain is inferred from stretching experiments [123, 124] in which the mechanical responses of the coronary arteries under the stretch condition were gauged. The result demonstrated that the first intimal rupture occurred at around 110% strain, and the strain-stress curve became non-smooth when strain reached approximately 120% . Therefore, we consider the threshold strain around the experimental rupture value 1.1 with an uncertainty of +/-20% in our UQ experiment. Note that the measurements in [123, 124] started from an unstrained sample, while in our model the vessel
is pre-strained by 30% owing to it being pressurized by the flowing blood inside it.

**Blood flow velocity**: Blood flow, as one of the mechanical factors, also plays an important role in the growth of neointima [125, 126]. High enough wall shear stress in the vessel accelerates the production of nitric oxide in endothelial cells, which acts as an inhibitor of SMC proliferation.

As mentioned before, the blood flow in the simulation is modelled as a steady flow with a constant parabolic inlet boundary condition, since we are mainly interested in the time-averaged values of the wall shear stress. The velocity data from [127] were applied to compute time-averaged velocity and converted to the parabolic profiles, the maximum velocity of which is 0.266 m/s. Owing to the measurement error and potential variety of velocity for individual vessels, we presume a large uncertainty in the data and vary it by 50% based on the average values 0.266 m/s.

**Fenestration percentage**: The IEL is modelled in ISR3D as a layer of agents on the inner surface of the vessel wall [89]. The fenestrations on IEL significantly affect the initial growth of SMCs as they allow SMCs to migrate into the blood vessel and start proliferating there. However, the SMCs in ISR3D are not able to change shape to migrate through the fenestrations, unlike real SMCs. Therefore, a certain percentage of IEL agents are switched to SMCs in ISR3D, to obtain a smaller amount of very large fenestrations, with the same total surface area as in the experiment. The fenestration percentage is calculated as the total area of all fenestrations divided by the total area of the IEL. The uncertainty ranges for this parameter are obtained from [128], where the percentage of fenestration in the hypercholesterolaemic group is approximately 7.5% and in the control group is approximately 3.5%. To include and study the scenarios for both cases, we consider the parameters to vary from 2% to 10%.

The ranges of all the uncertain parameters mentioned above are given in Table 3.1 and the distributions of the uncertainties are all assumed to be uniform.
Table 3.1: Ranges, units and coefficients of variation (CVs) of uncertain parameters of the ISR3D model. Note that the relative threshold strain is calculated with 30% pre-strained. The measurements in [123, 124] started from an unstrained sample while in our model the vessel is pre-strained by 30%. Therefore, strain listed by Holzapfel et al, \( \sigma_{\text{abs}} \), is scaled to obtain the relative deformation of our pre-strained tissue by \( \sigma_{\text{rel}} = \frac{\sigma_{\text{abs}} + 1}{1.3} - 1 \).

\[\text{3.4.2 Uncertainty estimations and sensitivity analysis}\]

For the UQ we applied the quasi-Monte Carlo (qMC) sampling method with Sobol sequence [129]. The method allows the sample to be more evenly distributed in the domain, which leads to a better convergence rate than the standard random sampling.

To investigate the uncertainty propagation of the uncertain inputs through the model, mean, variance, probability density function (PDF), and coefficient of variation are estimated. In addition, global sensitivity analysis has been conducted to study how much each uncertain input has contributed to the uncertainty of QoIs. The variance-based method (Sobol method) [31] is applied, which assumes that the latent functions \( f(x) \) can be decomposed into a combination of functions of individual uncertain inputs and their higher order interactions, which also leads to the following decomposition of the variance [31]:

\[
\text{Var} f(x) = \sum_i V_i + \sum_i \sum_{j>i} V_{ij} + \cdots + V_{12...d}, \tag{3.10}
\]

where \( V_i, V_{ij}, V_{12...d} \) stand for the partial variance contributed by the \( i \)-th uncertain input, by the interactions between the \( i \)-th and \( j \)-th uncertain inputs and by the higher order interactions. The first order Sobol indices indicate the independent contributions from the partial variance of each single uncertain input.
input:

\[ S_i = \frac{\text{Var}_{x_i}}{\text{Var} f(x)} = \frac{\text{Var}_{x_i} (E_{x_{\sim i}} (f(x) \mid x_i))}{\text{Var} f(x)}, \quad (3.11) \]

where \( x_{\sim i} \) denotes a vector of all uncertain parameters in \( x \) except \( x_i \). The total sensitivity indices take all the relevant contributions of an uncertain input into account:

\[ S_{T_i} = \frac{\text{Var}_{x_i}^{\text{total}}}{\text{Var} f(x)} = 1 - \frac{\text{Var}_{x_{\sim i}} (E_{x_i} (f(x) \mid x_{\sim i}))}{\text{Var} f(x)}. \quad (3.12) \]

All the sensitivity indices mentioned above are computed by Saltelli’s method [31].

### 3.5 Results

To train the surrogate models, 512 samples were generated by the qMC method and evaluated by the ISR3D model. Before the surrogate model was deployed to the UQ experiment, the surrogates were validated with a four-fold cross-validation. We measured the approximation error of both POD and GP regression with the relative \( L^2 \) norm:

\[
e_{\text{POD}} = \frac{1}{N_{cv}} \sum_{i=1}^{N_{cv}} \frac{\|y_i - \tilde{\Phi} \tilde{\Phi}^T y_i\|_{L^2}}{\|y_i\|_{L^2}}, \quad e_{\text{GP}} = \frac{1}{N_{cv}} \sum_{i=1}^{N_{cv}} \frac{\|y_i - \tilde{\Phi} \alpha(x_i)\|_{L^2}}{\|y_i\|_{L^2}}.
\quad (3.13)
\]

where \( N_{cv} \) denotes the number of samples used for cross-validation. In the cross-validation of POD, a certain number of snapshots were randomly taken from the training dataset and used to construct the snapshot matrix for SVD. The validation dataset was used to measure the approximation error. The relative \( L^2 \) error of the POD approximation with a different number of snapshots of both QoIs are shown in Figure 3.3. The average relative \( L^2 \) error gradually decreases to around 0.07% and 0.3% respectively with the number of snapshots reaching 100. The tendency of the curve shows that the error has almost converged to a limit; a further increase in the number of snapshots will not greatly improve performance. The low standard deviation of the error shows that there is no significant influence on the choice snapshots. Therefore, we randomly chose 100 snapshots from the output of the training data for the
Figure 3.3: Relative $L^2$ error of the POD with different numbers of snapshots for two QoIs, average cross-sectional area (a) and maximum relative area loss (b) using cross-validation. The error bar stands for the standard deviation computed from 100 replications.

Figure 3.4: A comparison between expected QoIs and predicted QoIs over all time steps from a cross-validation, where the diagonal lines denote the precise predictions of expected values. (a) ACSA, average cross-sectional area; (b) MRAL, maximum relative area loss.

POD in the construction of the surrogate model.

To test the performance of the GP regression, another four-fold cross-validation was performed with 100 repetitions. The predicted projection coefficients were first used to reconstruct their original model responses and subsequently compared with the expected output from the validation dataset. Comparisons of the predicted QoIs versus expected QoIs over all time steps...
are demonstrated in Figure 3.4. The resulting points are clustered around the diagonal line indicating that the GP has inferred the underlying functions well. The average relative $L^2$ error is 0.52% for the average cross-sectional area and 2.52% for the relative maximum area loss.

After the validations of surrogate models, the UQ experiments for both QoIs were performed. We applied the qMC method to draw $10^5$ samples from the uncertain input domain and fed them to the surrogate models. The mean and 50%, 75% and 95% percentile estimations of average cross-sectional area over time are shown in Figure 3.5. The corresponding histogram and PDFs of day 5, 10, 15, 20, and 30 are also shown in the same figure. The initial average cross-sectional area after stenting was around 3.17 mm$^2$. With the evolution of time, the cross-sectional area gradually reduced owing to the neointimal growth. The mean estimation of the average cross-sectional area shows that the neointimal growth was slow at the beginning but started to accelerate after day 1. An almost linear growth between day 1 and day 10 was observed followed by a descending growth rate until all the growth stopped at around day 22. The upper boundary of the 95% percentile shows that some samples stopped growing shortly after day 10 owing to the short re-endothelialization time, while a few other cases did not stop before 22 days.

The PDFs and histograms in Figure 3.5 show the details of the distributions of day 5, 10, 15, 20, and 30. On day 5, most of the samples cluster around 2.8 mm$^2$ and a small number of the samples have a lower average cross-sectional area up to 2.63 mm$^2$. A certain number of samples already stopped growing between day 10 and 20. The early stop usually means a small amount of neointima and contributes to the right tail of the distributions (around 2.4 mm$^2$ to 2.6 mm$^2$), while the rest of the samples still shifted towards the left owing to the growth. The difference between day 20 and day 30 is minor, indicating that the growth in most of the samples had stopped before day 20.

Similar patterns can be observed for the maximum relative area loss in Figure 3.6. The distribution at day 30 shows that most of the simulations ended up with 30% to 60% area loss. Assuming that the restenosis happened when the area loss reached 50%, about 5%, 16% and 18% of the simulations
Figure 3.5: Mean, 50%, 75% and 95% percentiles of the average cross-sectional area of the lumen over time with qMC sampling and corresponding histograms and PDFs at day 5, 10, 15, 20 and 30.

had reached the restenosis threshold at day 15, 20 and 30, respectively.

Table 3.2 provides detailed information on the mean, standard deviation (SD), CVs at day 5, 10, 15, 20 and 30 of both QoIs computed by 100 replications of the UQ experiment. Around 11.3% and 16.6% of uncertainty are observed from the average cross-sectional area and maximum relative area loss
3.5. Results

Figure 3.6: Mean, 50%, 75% and 95% percentiles of the maximum relative area loss over time with qMC sampling and corresponding histograms and PDFs at day 5, 10, 15, 20 and 30.

Apart from the uncertainty estimations, sensitivity analysis has also been performed. The sensitivity analysis was performed with $5 \times 10^5$ samples using the Sobol sequence and was repeated 100 times to compute the confidence interval. The first order indices of the four uncertain inputs over time for both
Table 3.2: Mean, standard deviation (SD), CV (in percentage) and percentage of restenosis at day 5, 10, 15, 20 and 30 for both QoIs computed from 100 repetitions of the UQ estimation. Owing to the large number of samples used, the confidence interval of the evaluations for each estimate is extremely small and is not shown here.

QoIs are shown in Figure 3.7. The confidence interval can hardly be seen in the figure, indicating extremely small uncertainty in our sensitivity estimations. The first order indices quantify the direct effect of each uncertain input on the total variance of the QoIs. For both QoIs, a dominant influence of the fenestration percentage on the variance can be observed at the initial stage of the process and keeps decreasing over time. It has almost no impact after 10 days. The blood flow velocity is a critical factor on the growth throughout the entire process and shows significant influences on variance in between day 5 and day 10, and gradually falls to around 0.2, while the endothelium regeneration times show an increasing effect and play the most important role after 13 days. The sensitivity to threshold strain is relatively small compared to the
other uncertain inputs. The total order indices measure both the independent and higher order interaction effect of an uncertain input on the total variance of QoIs. The total order indices of both QoIs are very similar to their first order results meaning that there is little higher order interaction between the uncertain inputs.

To further investigate the relations between uncertain inputs and restenosis, scatter distributions and histograms of the samples which reached the restenosis threshold at day 15, 20 and 30 are shown in Figure 3.8. Note that the threshold strain is not shown in the Figure since the sensitivity analysis result suggested that it is relatively not important in the process.

In the left column of Figure 3.8, scatter distributions of samples in terms of fenestration percentage and re-endothelialization time are shown. The range of re-endothelialization time falls between days 14 to 20, meanwhile a clear degression tendency can be observed from the corresponding histogram. The range of fenestration percentage shows that the restenosis can happen even with the lowest fenestration percentage, but the probability decreases slightly as the percentage drops. The middle column demonstrates the scatter distributions of re-endothelialization time and blood flow velocity. At day 15, only the cases with rather low blood flow velocity (under 0.27 \( m/s \)) reached the restenosis threshold. However at the end of the simulations, the upper bound rose to 0.38 \( m/s \). Unlike the left and right columns, a clear separation can be found between restenotic samples and the rest of the domain. The right column is based on fenestration percentage and blood flow velocity. Same patterns could be observed. The influence of fenestration percentage is rather minor while the value of blood flow velocity significantly affects the possibility of restenosis.

The speedup of the entire UQ experiment using the surrogate model has also been estimated. Table 3.3 shows the details of the computational cost, including the average core hour for model evaluation with ISR3D and surrogate model \( T_{ISR} \), training data generation \( T_{sample} \), and surrogate training \( T_{train} \). Both training and prediction of a surrogate model were extremely fast. The most computationally expensive part was the generation of training data with ISR3D. Each evaluation took around 585 core hour on average. Since in this case
Figure 3.8: Restenosis sample scatter distributions and corresponding histograms at day 15, 20 and 30.

\[ N_{\text{UQ}} T_{\text{ISR}} + T_{\text{train}} \text{ is negligible compared with } T_{\text{train}}, \text{ we find that the speedup equals } N_{\text{UQ}} / 512 \approx 976.6. \]

### 3.6 Discussion

The result of surrogate modelling shows that the combination of POD and GP regression performs well. The decomposition and reconstruction of the model response with POD save the computational effort for regression and provide a convenient and consistent way to cover the entire model response over time. In this work, the snapshot matrix was constructed by 100 randomly- cho-
3.6. Discussion

Table 3.3: The computation cost of the ISR3D model and the surrogate model. $T_{\text{ISR}}$ denotes the core hour to finish one single run of the simulation. Each simulation was performed exclusively on a node with $2 \times 12$-core 2.6 GHz Intel Xeon E5-2690 v3 (Haswell) CPUs of Dutch supercomputer *Cartesius*. $T_{\text{train}}$ and $T_{\text{sample}}$ stand for the sample generation time and training time for the surrogate model, respectively. $N_{\text{UQ}}$ denotes the number of samples used in a UQ experiment.

<table>
<thead>
<tr>
<th>UQ method</th>
<th>$T_{\text{ISR}}$ (core hour)</th>
<th>$T_{\text{train}}$ (core hour)</th>
<th>$T_{\text{sample}}$ (core hour)</th>
<th>$N_{\text{UQ}}$</th>
<th>Speedup of UQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>qMC</td>
<td>585.1</td>
<td>/</td>
<td>/</td>
<td>585.1 × 512</td>
<td>5 × 10^5</td>
</tr>
<tr>
<td>qMC (surrogate)</td>
<td>$6.1 \times 10^{-8}$</td>
<td>$3.1 \times 10^{-2}$</td>
<td>/</td>
<td>/</td>
<td>976.6</td>
</tr>
</tbody>
</table>
of cells at the lumen surface could already have their growth inhibited by nitric oxide. After day 10, the influence of the blood flow velocity drops gradually and is replaced by re-endothelialization. Figure 3.8 shows that, at the end of the simulations, the influences of fenestration percentage are relatively minor compared with the effect of blood flow velocity and endothelium regeneration time. This suggests that the scenarios with a high fenestration percentage, such as hypercholesterolaemia, might not have a high impact on restenosis probability if other parameters such as endothelium regeneration time can be strictly controlled.

In the design of the UQ experiment, the choices of uncertain ranges and distributions are described in Section 3.4.1. For all uncertain parameters, the average values and some of the uncertainty ranges are directly obtained or calculated from experimental results. However, owing to the lack of data on real population distributions of these parameters, the uncertainty ranges and their corresponding distributions are mainly defined by our approximations. They are simplified, and are likely to be different from the actual distribution. Changing the input distribution naturally will lead to a different output distribution in the UQ. Our choice of uniform distributions and approximation of the ranges can illustrate the role of the input parameters and their effect on the restenosis process to a certain extent, but further UQ experiments based on actual distribution data of these parameters are still required and the outcome of which can also be applied to verify against clinical restenosis data. If such a realistic input distribution can be obtained, e.g. from a synthetic population, the trained surrogate can be used to perform the UQ quickly, as long as the population data lies in the same range used for generating the training sample.

In this work, we studied four biological uncertain parameters. We quantified their uncertainty propagation and sensitivity for two QoIs adapted for in-silico models from clinically recognized metrics. This helps us to better understand the underlying contribution of these parameters to restenosis. In addition to the investigated biological factors, other factors and scenarios can also be studied via ISR3D, for example, variability in the stenting procedure, such as deployment depth or malapposition of the stent. Through the UQ analysis,
the potential effect of such factors can be quantified and studied. Additionally, different scenarios, such as small/large vessel diameters and the tortuosity of the stented vessel, can also significantly influence the outcome of a simulation. We leave the study of these factors, which all affect the initial shape of the stented vessel, to our future work.

The ISR3D model itself has several limitations. First, it does not account for the inflammation processes, which are important during the early stages of post-stenting. Second, the geometry used in the UQ experiment is not based on any particular vessel, and instead is a piece of a perfectly cylindrical tube, and the stent fits the curvature of the vessel perfectly and is radially expanded in a uniform way. Additionally, a steady flow approximation was used for our BF model. The inlet was set to a parabolic profile, assuming the vessel’s shape upstream of the inlet is unknown. It should be noted that the curvature affects the flow pattern to a great extent, leading to helical flows and affecting the wall shear stress also inside the stented region. The effects of the curvature are studied to some extent in our earlier publications; see for example, figure 9 in [90]. Similarly, finite element computational studies confirm the same effect of the curvature in [131]. These studies demonstrate that it is possible to capture the helical flow with a parabolic inlet condition in the regions of developed flow, but the flow pattern may be unrealistic close to the inlet.

A combination of the limitations listed above may result in the underestimation of restenotic growth. For example, Morton et al. [132] reported an area loss of 62% for porcine vessels of a similar diameter and deployment depth with an NIR stent, which is very close to the upper bound of the distribution predicted by the model. Nevertheless, the experimental values lie within the distribution, further confirming that the ranges selected for UQ reasonably overlap with the physiological ranges. There are also other limitations in the model we use, which are discussed in detail in [89, 90].

3.7 Conclusion

The UQ and sensitivity analysis of a multiscale model ISR3D was performed. The uncertainty propagation from four parameters, endothelium regeneration
time, threshold strain, percentage of fenestration and blood flow velocity to two QoIs, the average cross-sectional area and the maximum relative area loss; are investigated. Owing to the high computational cost of ISR3D, surrogate modelling techniques were applied. The QoIs over time were decomposed by POD and the resulting projection coefficients were learned by a GP regression model. Cross-validations are applied to validate the performance of the surrogate model. The surrogate model was subsequently deployed in the UQ experiment to replace the original model. The UQ and sensitivity analysis results showed that the blood flow velocity and endothelium regeneration time have a significant influence on the neointima growth and result in restenosis, while the impact of the fenestration percentage is limited and the threshold strain barely has any influence on the process.