

## Research



**Cite this article:** Paun LM, Colebank MJ, Olufsen MS, Hill NA, Husmeier D. 2020 Assessing model mismatch and model selection in a Bayesian uncertainty quantification analysis of a fluid-dynamics model of pulmonary blood circulation. *J. R. Soc. Interface* **17**: 20200886. <https://doi.org/10.1098/rsif.2020.0886>

Received: 2 November 2020

Accepted: 2 December 2020

### Subject Category:

Life Sciences—Mathematics interface

### Subject Areas:

biomathematics, biophysics, computational biology

### Keywords:

uncertainty quantification, model mismatch, model selection, MCMC, Gaussian processes, pulmonary circulation

### Author for correspondence:

L. Mihaela Paun

e-mail: l.paun.1@research.gla.ac.uk

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5238491>.

# Assessing model mismatch and model selection in a Bayesian uncertainty quantification analysis of a fluid-dynamics model of pulmonary blood circulation

L. Mihaela Paun<sup>1</sup>, Mitchel J. Colebank<sup>2</sup>, Mette S. Olufsen<sup>2</sup>, Nicholas A. Hill<sup>1</sup> and Dirk Husmeier<sup>1</sup>

<sup>1</sup>School of Mathematics and Statistics, University of Glasgow, Glasgow, G12 8QQ, UK

<sup>2</sup>Department of Mathematics, North Carolina State University, Raleigh, NC 27695, USA

**id** LMP, 0000-0002-8734-8135; MJC, 0000-0002-2294-9124; MSO, 0000-0003-2694-0044; DH, 0000-0003-1673-7413

This study uses Bayesian inference to quantify the uncertainty of model parameters and haemodynamic predictions in a one-dimensional pulmonary circulation model based on an integration of mouse haemodynamic and micro-computed tomography imaging data. We emphasize an often neglected, though important source of uncertainty: in the mathematical model form due to the discrepancy between the model and the reality, and in the measurements due to the wrong noise model (jointly called ‘model mismatch’). We demonstrate that minimizing the mean squared error between the measured and the predicted data (the conventional method) in the presence of model mismatch leads to biased and overly confident parameter estimates and haemodynamic predictions. We show that our proposed method allowing for model mismatch, which we represent with Gaussian processes, corrects the bias. Additionally, we compare a linear and a nonlinear wall model, as well as models with different vessel stiffness relations. We use formal model selection analysis based on the Watanabe Akaike information criterion to select the model that best predicts the pulmonary haemodynamics. Results show that the nonlinear pressure–area relationship with stiffness dependent on the unstressed radius predicts best the data measured in a control mouse.

## 1. Introduction

Computational haemodynamics models are emerging as powerful tools for analysing cardiovascular disease progression and the effects of treatments [1] by providing essential haemodynamic metrics which could not be obtained from *in vivo* experiments [2]. The ultimate goal is achieving personalized medicine, to allow patient-specific care and treatment. Before using the models for decision-making in the clinic, they must be calibrated and fitted to data, and their credibility rigorously tested by modelling all sources of uncertainty using statistical analysis.

The current study assesses the health of the pulmonary system by integrating imaging data (obtained with micro-computed tomography (CT)), blood pressure data (measured invasively via catheterization) and blood flow data (measured with ultrasound), using a one-dimensional (1D) fluid-dynamics model combined with statistical inference. Predictions of blood pressure, blood flow and vessel area are computed in an arterial network model constructed from micro-CT images from a control mouse, and the pressure predictions are compared to dynamic data in the main pulmonary artery (MPA).

We highlight the importance of determining the uncertainty when calibrating the model to the data. Our analysis includes the uncertainty in the model parameters (which are naturally variable), in the model form/structure (the discrepancy between the model and the reality), in the measurements (the noise model), and in the simulator output (e.g. the errors from numerically integrating the model equations).

Several previous studies [3–7] have developed 1D fluid-dynamics models predicting pulmonary blood flow and pressure. However, only a few [3,4] have aimed at devising subject-specific predictions by estimating model parameters. These studies minimize the least-squares error between the model output and the measurements. While these investigations provided valuable insight into the physiology, they ignored an essential source of uncertainty resulting from the inadequacy of the model form. Our work shows that the consequence of ignoring model discrepancy is biased haemodynamic predictions and parameters, and thus inability of reliably using these models in the clinic.

### 1.1. Model parameters and types

Similar to previous studies [3,4], the 1D models analysed here have two types of parameters: specifying the vessel network (radius, length and the connectivity of arteries) and the haemodynamics (pressure and flow). The current study focuses on inferring and analysing parameters intrinsic to the haemodynamic model, including the vessel stiffness and parameters specifying the micro-circulation (boundary conditions via three-element Windkessel models attached at the terminal vessels), in a fixed network. We examine several alternative models to select the model that best predicts the data, which is performed using statistical model selection criteria based on the Watanabe Akaike information criterion (WAIC) [8]. We compare two constitutive equations: a linear and a nonlinear wall model relating vessel pressure and area. We investigate if the vessel stiffness is (a) constant over the entire network (as suggested in [9]), (b) increases with decreasing radius (as suggested in [10]), or (c) should be estimated for each vessel separately. In addition, we estimate micro-circulation parameters by introducing global scaling factors for the boundary condition parameters at the terminal arteries [4,11].

### 1.2. Bayesian inference

Due to the limited data and the model complexity, these parameters may be highly correlated or unidentifiable, thus it may be unfeasible to estimate all parameters uniquely. To address these issues, we use a Bayesian approach, with the aim to obtain the posterior distribution of the parameters [11,12]. This is analytically intractable, so we use Markov chain Monte Carlo (MCMC) to sample parameters approximately from the posterior distribution (with an asymptotic convergence guarantee). The posterior parameter samples are then used to estimate the uncertainty of model predictions throughout the pulmonary arterial network.

### 1.3. Model mismatch

Several recent studies have incorporated Bayesian parameter estimation [3,4,13–15] and uncertainty quantification (UQ) [16,17], however, most of these studies ignore the model

mismatch. Our current study assumes that the model mismatch stems from two sources: (1) inadequate mathematical model (i.e. model discrepancy, since the mathematical model is not a perfect representation of the real system) [18] and (2) incorrect noise model (i.e. erroneously assuming independence when the errors are, in fact, correlated). A few studies have discussed the importance of allowing for model discrepancy. In an electrophysiology model, Lei *et al.* [19] incorporate the model discrepancy using GPs and autoregressive-moving-average (ARMA) models; the authors show using synthetic studies that ignoring the model form uncertainty leads to biased predictions and uncertainty underestimation. Additionally, Whittaker *et al.* [20] and Mirams *et al.* [21] discuss model discrepancy in a review of cardiac model calibration. Furthermore, a few studies have investigated the impact of making the wrong assumption for the measurement errors. For example, Konukoglu *et al.* [22] included an inhomogeneous variance, informed by the authors' experience with the data, in an electrophysiology model, finding that the noise model greatly influences the inference results. Despite these findings, most cardiovascular modelling studies do not account for model mismatch [23,24].

To investigate the importance of accounting for model mismatch, we employ a Bayesian approach quantifying the uncertainty in the mathematical and the noise model form based on data. The model mismatch is explicitly modelled using Gaussian processes (GPs) [25] following an approach by Kennedy *et al.* [18]. Our Bayesian inference framework uses MCMC to jointly sample the mathematical model parameters and the model mismatch (error model) parameters from their posterior distribution. Thus, uncertainties associated with parameters, model form and measurements are all accounted for in our analysis [26].

We use physiological and synthetic pressure data to examine the consequence of inferring parameters when suspected model mismatch is unaccounted for. Results show that ignoring model mismatch biases parameter estimates and underestimates uncertainty in parameter and output space, whereas our proposed method corrects this bias. In addition, we carry out a synthetic study displaying the effect of using data from multiple vessels on the parameter inference and UQ.

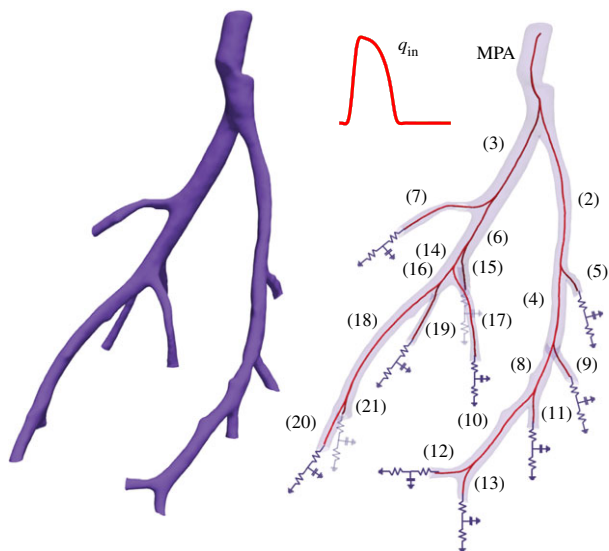
Finally, we perform model selection based on WAIC to discriminate between the two constitutive models, with a number of parameter constraints related to the vessel stiffness.

## 2. Data

### 2.1. Physiological data

This study compares model predictions to measured MPA blood pressure data from a control mouse lung (figure 1). The experimental protocols used to extract the haemodynamic and image data are summarized in our recent study [4], and a more detailed experimental protocol is found in [27,28]. We provide a brief overview of the data used in this study.

A 3D segmentation of the vessel geometry is obtained from micro-CT images of excised mice lungs as described in detail by Vanderpool *et al.* [28]. The image data are segmented, and using the Vascular Modeling ToolKit (VMTK,<sup>1</sup>



**Figure 1.** 3D smoothed segmented network from a micro-CT image of a healthy mouse lung (left) and the directional graph of the same network with vessel numbers attached (right). At the network inlet, we specified a flow waveform taken from measurements (see §2.1 for a description of the experimental data), and at the outlet of each terminal vessel, we attached a three-element Windkessel model with two resistors and a capacitor.

[29]) and custom algorithms discussed in [3] we obtain a 1D directional graph of the large vessels, perfusing each of the lung lobes. In this study, we assume that all vessels are straight, i.e. that they do not taper along their length. While many studies [10] examining flow in the systemic arteries account for vessel taper, the vessel radii in the large pulmonary arteries of mice show negligible inter-vessel tapering likely because the pulmonary tree is formed by rapidly branching vessels.

Dynamic pressure and flow waves were measured in the MPA. Pressure was measured using a 1.0-F pressure-tip catheter (Millar Instruments, Houston, TX) and recorded on a haemodynamic workstation (Cardiovascular Engineering, Norwood, MA) at 5 kHz. MPA flow velocity was simultaneously measured during catheterization on the same workstation via ultrasound (Visualsonics, Toronto, Ontario, CA) at a rate of 30 MHz [27]. The haemodynamic data analysed in this study include wave forms averaged over 20 cardiac cycles, using available ECG as a fiducial point.

## 2.2. Synthetic data

We also use synthetic data obtained from a forward simulation of the mathematical model. We generated synthetic, error-free data in all 21 vessels using the linear wall model and a radius-dependent exponential stiffness with parameter values consistent with the physiological data. To these data, we added non-stationary, additive Gaussian correlated errors generated using the same error parameters for all the vessels, assuming that the pressure transducer produces measurement noise that is independent of the measurement location (see §5.2 for numerical details). This error correlation induces a model mismatch if the standard assumption of independent and identically distributed (iid) measurement errors is made, i.e. if the wrong noise model is used. To make the synthetic data physiologically realistic, we ensured a signal-to-noise ratio of approximately 100 (see fig. 5 in [30])

and that the pressure monotonicity constraint is satisfied, i.e. that the pressure decreases as it approaches the periphery [31]. The mathematical model provides predictions satisfying the pressure monotonicity constraint. However, when adding noise, this is no longer guaranteed to hold. Therefore, to satisfy the monotonicity condition, the noise was constrained by using a rejection mechanism, i.e. any noise instantiation leading to a constraint violation was discarded. This mimics the experimental procedure by which data that appear physiologically unrealistic are disposed of.

## 3. Model

### 3.1. Fluid-dynamics model of the pulmonary circulation

This study uses our previously developed 1D fluid-dynamics model [4], predicting pressure, flow and cross-sectional area explicitly in the large pulmonary arteries (shown in figure 1). For each vessel, the 1D model is derived under the assumptions that blood is incompressible and that the flow is Newtonian, laminar and axisymmetric, and has no swirl. Under these assumptions, the Navier–Stokes equations describing conservation of mass and momentum are expressed by

$$\left. \begin{aligned} \frac{\partial A}{\partial t} + \frac{\partial q}{\partial x} &= 0, \\ \frac{\partial q}{\partial t} + \frac{\partial}{\partial x} \left( \frac{q^2}{A} \right) + \frac{A \partial p}{\rho \partial x} &= -\frac{2\pi\mu r}{\rho \delta} \frac{q}{A}, \end{aligned} \right\} \quad (3.1)$$

where  $x$  (cm) and  $t$  (s) denote the spatial and temporal coordinates, and  $p = \tilde{p}/\text{conv}$  (mmHg) denotes pressure ( $\tilde{p}$  ( $\text{g} \cdot \text{cm}^{-1} \text{s}^{-2}$ ) is in cgs units and  $\text{conv} = 1333.22$  (mmHg ( $\text{g} \cdot \text{cm}^{-1} \text{s}^{-2}$ ) $^{-1}$ ) is a conversion factor).  $\rho = 1.055 \text{ g ml}^{-1}$  is the blood density and  $\mu = 0.049 \text{ g cm}^{-1} \text{ s}^{-1}$  is the blood viscosity, assumed constant. Vessel radius and cross-sectional area are  $r(x, t)$  (cm) and  $A(x, t) = \pi r(x, t)^2$  ( $\text{cm}^2$ ), respectively. The volumetric flow rate  $q = A\bar{u}_x$  ( $\text{ml s}^{-1}$ ) is derived assuming a Stokes boundary layer velocity profile

$$u_x(r, x, t) = \begin{cases} \bar{u}_x, & 0 \leq r \leq \delta, \\ \bar{u}_x \frac{(R-r)}{\delta}, & \delta \leq r \leq R(x, t), \end{cases} \quad (3.2)$$

where  $\bar{u}_x$  is the average velocity,  $\delta = \sqrt{\mu T / 2\pi\rho}$  (cm) is the boundary layer thickness and  $T$  (s) is the length of the cardiac cycle.

To close the system of equations, we add a constitutive pressure–area relation. We investigate two wall models: a linear elastic model [32]

$$p = \frac{4}{3} \chi \left( \sqrt{\frac{A}{A_0}} - 1 \right), \quad (3.3)$$

where  $\chi = Eh/r_0$  is the wall stiffness,  $E$  ( $\text{g cm}^{-1} \text{ s}^{-2}$ ) is Young's modulus in the circumferential direction,  $h$  (cm) is the wall thickness, and  $r_0$  (cm) is the unstressed vessel radius ( $A_0 = \pi r_0^2$ ); and an empirical nonlinear wall model [4] given by

$$p = \chi \tan \left[ \frac{\pi}{\gamma} \left( \frac{A}{A_0} - 1 \right) \right], \quad (3.4)$$

where  $\gamma > 0$  (dimensionless) is a scaling parameter specifying the maximal lumen area  $A_\infty$  for  $p \rightarrow \infty$ , and  $\chi > 0$  ( $\text{g cm}^{-1} \text{ s}^{-2}$ ) is a material parameter (defined similarly to  $Eh/r_0$ ) characterizing the half-max compliance pressure.

To investigate the claim of medial thickening and stiffening as arteries get smaller [33], we express the dependence of  $\chi$  on  $r_0$  via the equation

$$\chi(r_0, f_1, f_2, f_3) = f_1 \exp(f_2 r_0) + f_3, \quad (3.5)$$

$f_1$  ( $\text{g cm}^{-1} \text{s}^{-2}$ ),  $f_2$  ( $\text{cm}^{-1}$ ) and  $f_3$  ( $\text{g cm}^{-1} \text{s}^{-2}$ ) are constant parameters [10].

At the inlet to the network (shown in figure 1), we specify the flow (taken from measurements). Similar to previous studies [3,5,10], we assume flow conservation and pressure continuity

$$\text{and } \left. \begin{aligned} p_p(L, t) &= p_{d_1}(0, t) = p_{d_2}(0, t) \\ q_p(L, t) &= \sum_{i=1}^2 q_{d_i}(0, t), \end{aligned} \right\} \quad (3.6)$$

where  $p$  denotes the parent vessel,  $d_1$  and  $d_2$  are the daughter vessels, and  $L$  (cm) is the vessel length.

The micro-circulation is represented by a three-element Windkessel model (an RCR circuit) relating pressure and flow as

$$\left. \begin{aligned} \frac{dp(L, t)}{dt} - R_1 \frac{dq(L, t)}{dt} &= \\ q(L, t) \left( \frac{R_1 + R_2}{R_2 C} \right) - \frac{p(L, t)}{R_2 C} &= \end{aligned} \right\} \quad (3.7)$$

where  $R_1$ ,  $R_2$  ( $\text{g ml}^{-1} \text{cm}^{-1} \text{s}^{-1}$ ) are resistances, and  $C$  ( $\text{ml}^{-1} \text{cm}^{-1} \text{s}^{-2} \text{g}^{-1}$ ) is the capacitance.

### 3.2. Model parameters

The haemodynamic model has three types of parameters specifying blood characteristics, vessel tissue properties and micro-vasculature dynamics. Blood viscosity and density values are taken from the literature [34] and assumed constant. The stiffness in large vessels can be measured *ex vivo* via stress-strain testing. Finally, micro-vasculature parameters are prescribed in each terminal vessel's boundary condition.

In this study, we infer both wall model and boundary condition parameters. The linear wall model (3.3) has three parameters  $f_1$ ,  $f_2$ ,  $f_3$  characterizing the large vessel stiffness, while the nonlinear wall model (3.4) has four parameters  $\gamma$ ,  $f_1$ ,  $f_2$ ,  $f_3$ .

Three boundary condition parameters ( $R_1$ ,  $R_2$ ,  $C$ ) are specified for each terminal vessel. The network in figure 1 has 11 terminal vessels giving a total of 33 parameters. Since we only have data in the MPA, these parameters are not practically identifiable. To reduce parameter dimensionality, we introduce factors ( $\psi_1$ ,  $\psi_2$ ,  $c$ ) [4] scaling the nominal (initial) Windkessel parameters by

$$R_1^j = \psi_1 R_{01}^j, \quad R_2^j = \psi_2 R_{02}^j \quad \text{and} \quad C^j = c C_0^j, \quad (3.8)$$

where  $R_{01}^j$ ,  $R_{02}^j$ ,  $C_0^j$  are the nominal values for the  $j^{\text{th}}$  terminal vessel, computed using the junction conditions and Poiseuille's flow, as described in detail in our previous studies [3,4].  $R_1^j$ ,  $R_2^j$ ,  $C^j$  are the adjusted (estimated) values for the  $j^{\text{th}}$  terminal vessel.  $\psi_1$ ,  $\psi_2$ ,  $c$  are the scaling factors, common to all terminal vessels. The scaling factors are estimated from the available data.

Forward simulations were run for both the linear and nonlinear model to ensure physiologically plausible parameter bounds. Stiffness bounds ensure pressure sensitivity

to changes in  $\chi$ , while scaling factor bounds were constructed to ensure physiological pressures ( $12 \leq \max(p) \leq 35$  mmHg). These parameters are constrained in a univariate sense, but the parameters' behaviour in the joint space is unknown prior to carrying out the statistical analysis. Table 2 shows the univariate parameter ranges.

### 3.3. Overview of models: physiological hypotheses and model mismatch scenarios

Table 1 outlines the models considered in our work, which help explore several physiological hypotheses and model mismatch scenarios.

By analysing these models, for the physiological data described in §2.1, we test

- if the vessel wall model is:
  - (i) linear (equation (3.3)), and
  - (ii) nonlinear (equation (3.4))
- if the vessel stiffness:
  - (i) is constant, thus shared between the vessels ( $f_1$  in equation (3.5) is 0 and all vessels share one common  $f_3$  value),
  - (ii) is vessel-specific, thus independent of the vessel ( $f_1$  in equation (3.5) is 0 and every vessel has its own  $f_3$  value), and
  - (iii) is radius-dependent (expressed via equation (3.5))
- if the model mismatch, described in detail in §4.2, should be accounted for (no/yes).

## 4. Statistical methods

### 4.1. Data likelihood

We assume normally distributed errors (an assumption which we have checked by comparing the distribution of the residuals to a normal distribution, see electronic supplementary material, S4), and explore both iid errors and correlated errors. Under these assumptions, we express the likelihood function as:

- iid errors:  $\mathbf{y}(\mathbf{t}) \sim \mathcal{MVN}(\mathbf{m}(\boldsymbol{\theta}, \mathbf{t}), \sigma^2 \mathbf{I})$  (multivariate normal distribution), i.e.

$$p(\mathbf{y}|\boldsymbol{\theta}, \sigma^2) = \left( \frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp\left( -\frac{\sum_{i=1}^n (y(t_i) - m(\boldsymbol{\theta}, t_i))^2}{2\sigma^2} \right), \quad (4.1)$$

where

$$\sum_{i=1}^n (y(t_i) - m(\boldsymbol{\theta}, t_i))^2 = (\mathbf{y}(\mathbf{t}) - \mathbf{m}(\boldsymbol{\theta}, \mathbf{t}))^T (\mathbf{y}(\mathbf{t}) - \mathbf{m}(\boldsymbol{\theta}, \mathbf{t})) \quad (4.2)$$

is the Euclidean distance, and  $\mathbf{y}(\mathbf{t}) = (y(t_1), \dots, y(t_n))$  is the  $n$ -vector of temporal measurements,  $\mathbf{m}(\boldsymbol{\theta}, \mathbf{t})$  is the  $n$ -vector of temporal pressure predictions from the mathematical model evaluated with parameters  $\boldsymbol{\theta}$ , and  $\sigma^2$  is the measurement noise variance.

- Correlated errors:  $\mathbf{y}(\mathbf{t}) \sim \mathcal{MVN}(\mathbf{m}(\boldsymbol{\theta}, \mathbf{t}), \mathbf{C})$ , i.e.

$$p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{C}) = \det(2\pi\mathbf{C})^{-\frac{1}{2}} \exp\left( -\frac{1}{2} (\mathbf{y}(\mathbf{t}) - \mathbf{m}(\boldsymbol{\theta}, \mathbf{t}))^T \mathbf{C}^{-1} (\mathbf{y}(\mathbf{t}) - \mathbf{m}(\boldsymbol{\theta}, \mathbf{t})) \right), \quad (4.3)$$



**Table 1.** Models investigated: two constitutive models relating pressure–area (linear (equation (3.3)) and nonlinear (equation (3.4))) with several stiffness relations: constant ( $f_1$  in equation (3.5) is 0 and all vessels share one common  $f_3$  value), vessel-specific ( $f_1$  in equation (3.5) is 0 and every vessel has its own  $f_3$  value), or radius-dependent (expressed via equation (3.5)); and model and measurement error assumptions via including or ignoring model mismatch, described in detail in §4.2. For the nonlinear wall model, we do not consider the no model mismatch scenario based on conclusions drawn from the linear model, clearly supporting modelling the model mismatch. In addition, the vessel-specific stiffness scenario is not pursued due to the interaction between the parameters  $\chi$  and  $\gamma$  in equation (3.4), requiring vessel-specific ( $\chi, \gamma$ ). This would lead to a very large number of parameters being estimated, requiring extremely high computational efforts (simulations would most likely take months to complete).

vessel wall model	vessel stiffness type	allow for model mismatch	model abbreviation
linear	constant	no	A
	constant	yes	B/C
	radius-dependent	no	D
	radius-dependent	yes	E
	vessel-specific	no	F
	vessel-specific	yes	G
nonlinear	constant	yes	H
	radius-dependent	yes	I

where

$$(\mathbf{y}(\mathbf{t}) - \mathbf{m}(\boldsymbol{\theta}, \mathbf{t}))^T \mathbf{C}^{-1} (\mathbf{y}(\mathbf{t}) - \mathbf{m}(\boldsymbol{\theta}, \mathbf{t})) \quad (4.4)$$

is the Mahalanobis distance, and  $\mathbf{C}$  is the covariance matrix of the errors.

## 4.2. Model mismatch

The model mismatch function can be visualized by plotting the residuals, i.e. the difference between the partial differential equation (PDE) predictions and the measurements in time, see figure 8 (bottom right corner). A clear pattern is observed, indicating a correlation between the residuals. Such a plot should be used to decide on the appropriate error assumption.

When the model mismatch is neglected, the statistical model equation is equivalent to equation (4.1) and can be expressed as

$$\mathbf{y}(\mathbf{t}) = \mathbf{m}(\boldsymbol{\theta}, \mathbf{t}) + \mathbf{u}(\mathbf{t}) \quad \text{and} \quad \mathbf{u}(\mathbf{t}) \sim \mathcal{MVN}(\mathbf{0}, \sigma^2 \mathbf{I}), \quad (4.5)$$

where  $\sigma^2$  is the error parameter (known as measurement noise variance).

When incorporating the model mismatch with GPs [18] (details in electronic supplementary material, S2), the statistical model equation is equivalent to equation (4.3) and is given by,

$$\left. \begin{aligned} \mathbf{y}(\mathbf{t}) &= \mathbf{m}(\boldsymbol{\theta}, \mathbf{t}) + \boldsymbol{\Gamma}(\mathbf{t}) = \mathbf{m}(\boldsymbol{\theta}, \mathbf{t}) + \mathbf{f}(\mathbf{t}) + \mathbf{u}(\mathbf{t}) \\ \text{and} \quad \mathbf{f}(\mathbf{t}) &\sim \mathcal{GP}(\mathbf{0}, \mathbf{K}|\boldsymbol{\eta}), \quad \mathbf{u}(\mathbf{t}) \sim \mathcal{MVN}(\mathbf{0}, \sigma_n^2 \mathbf{I}), \end{aligned} \right\} \quad (4.6)$$

where  $\boldsymbol{\Gamma}(\mathbf{t})$  is the model mismatch function,  $\mathbf{f}(\mathbf{t})$  is a latent function,  $\mathcal{GP}(\mathbf{0}, \mathbf{K}|\boldsymbol{\eta})$  is a GP with zero mean and covariance matrix  $\mathbf{K}$ , which is a function of the error parameters  $\boldsymbol{\eta}$  (known as covariance function hyperparameters), and  $\sigma_n^2$  is the residual noise variance. In equation (4.3),  $\mathbf{C} = \mathbf{K} + \sigma_n^2 \mathbf{I}$ .

A neural network covariance function with hyperparameters  $w$  and  $b$  [25] is used to fit the GP to the residuals exhibiting non-stationarity (figure 8). Equation (4.5) is a limiting case of equation (4.6) by removing the GP contribution term,  $\mathbf{f}(\mathbf{t})$ , and setting  $\sigma_n^2 = \sigma^2$  (details in electronic supplementary material, S3).

The model mismatch  $\boldsymbol{\Gamma}(\mathbf{t})$  stems from two sources: (i) the model discrepancy between the real system and the mathematical model and (ii) the incorrect noise model (i.e. making the iid assumption for correlated measurement errors), thus

$$\boldsymbol{\Gamma}(\mathbf{t}) = \boldsymbol{\zeta}(\mathbf{t}) + \boldsymbol{\epsilon}(\mathbf{t}), \quad (4.7)$$

where  $\boldsymbol{\zeta}(\mathbf{t})$  is the model discrepancy function, and  $\boldsymbol{\epsilon}(\mathbf{t})$  is the noise model function. Note that in the present article, we distinguish between *model discrepancy* and *model mismatch* in the way described above, so these words are not used synonymously.

Since this is a retrospective data analysis study, it was intrinsically impossible to separate the contributions from the measurement error (noise model) and model error (model discrepancy), which are therefore modelled with one single GP, see §8 for a more thorough discussion on this.

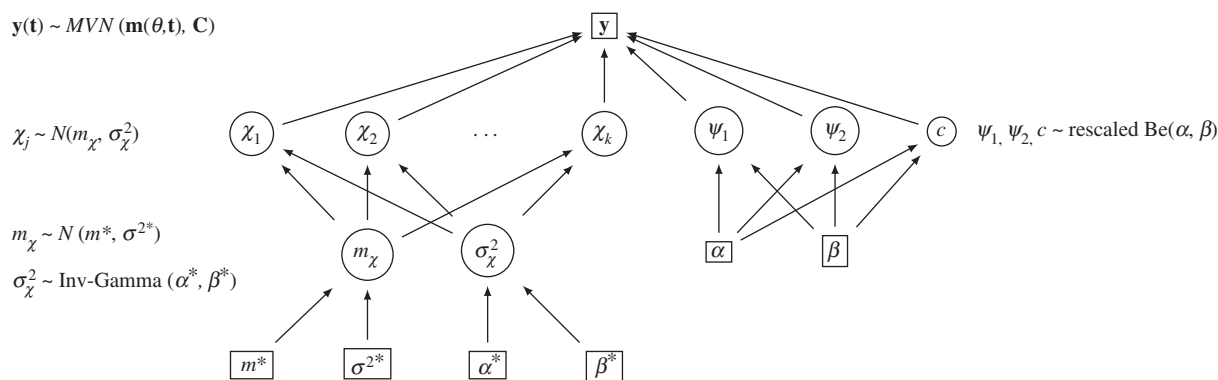
Possible causes for the measurement error correlation are: the temporal nature of the data (measurements at the current time point depend on measurements at previous time points), and smoothing and averaging of the data. Possible causes for the model discrepancy are: numerical errors (e.g. numerical integration of the PDEs), model assumptions (e.g. purely elastic vessel walls, or the 1D model simplification), uncertainty of the network geometry (kept fixed), and inconsistency between network geometry and haemodynamic data, which come from different mice.

## 4.3. Prior distributions

### 4.3.1. Biophysical parameters

*Constant or radius-dependent stiffness models.* For all models with constant or radius-dependent stiffness (table 1), we used a rescaled beta distribution for the biophysical parameters to ensure positive support within physiologically realistic ranges ( $[l_i, u_i]$ ) [4,9],  $\theta_i \sim \text{rescaled beta}(1, 1)$ ,  $l_i \leq \theta_i \leq u_i$ , where  $i = 1, \dots, k$ , with  $k$  being the parameter dimensionality.

*Vessel-specific stiffness in a Bayesian hierarchical model.* Different pulmonary arteries may have different vessel wall stiffness values, but since all vessels have similar tissue composition,



**Figure 2.** Bayesian hierarchical model used for the vessel-specific stiffness analysis. The data, denoted by  $\mathbf{y}(\mathbf{t})$ , are assumed to follow a multivariate normal distribution with mean  $\mathbf{m}(\boldsymbol{\theta}, \mathbf{t})$  and covariance matrix  $\mathbf{C}$ . If iid errors are assumed (i.e. the model mismatch is ignored),  $\mathbf{C}$  is a diagonal matrix,  $\mathbf{C} = \sigma^2 \mathbf{I}$  (where  $\sigma^2$ : noise variance and  $\mathbf{I}$ : identity matrix), and if correlated errors are assumed (i.e. the model mismatch is incorporated),  $\mathbf{C}$  is a full matrix. The biophysical parameters,  $\boldsymbol{\theta} = (\chi_1, \dots, \chi_k, \psi_1, \psi_2, c)$  (described in §3.2), and the hyperparameters,  $m_\chi, \sigma_\chi^2$ , are *a priori* drawn from the distributions indicated in the graphical model. The circle represents variable quantities, which are inferred using MCMC, and the rectangle stands for fixed quantities. Inference in this model is analytically intractable, and we resort to a Gibbs sampling scheme, discussed in electronic supplementary material, equations (24)–(26). Note that a modification of this model, where an additional edge is introduced from  $\sigma_\chi^2$  to  $m_\chi$  allows these two parameters to be integrated out in closed form, potentially leading to a more efficient sampling scheme; however, these equations are less intuitive, so the details have been relegated to electronic supplementary material, S5.

the parameters are related. A Bayesian hierarchical model [35] is needed to incorporate our prior notion that the vessel stiffnesses are similar, and this model provides a mechanism of information sharing among the vessel stiffness parameters. The dependence of the stiffness parameters can be captured by using a common ‘population’ prior distribution, from which vessel-specific stiffness parameters are sampled. Next, to allow the stiffness parameters to influence each other, we introduce a layer of priors for the hyperparameters of the population distribution. This construct enables the hyperparameters to be variable, ensuring a dependency between the stiffness parameters, and the hyperparameters’ uncertainty is naturally modelled. The result is a Bayesian hierarchical model (shown in figure 2), which tends to avoid overfitting the existing data by allowing information sharing between the stiffness parameters. This model subsumes two simpler models as limiting cases: the model where all vessels have the same stiffness (when the prior distribution of the vessel-specific stiffness parameter collapses to a delta spike), and the model of independent vessel-specific stiffness parameters without information sharing (when the prior distribution of the stiffness parameters is the uniform distribution). Electronic supplementary material, S5 offers more details of the Bayesian hierarchical model.

Under this model (figure 2), the three Windkessel parameters  $\psi_1, \psi_2, c$  are assumed common to all the vessels, and to ensure positive support (within physiological ranges) for them, we used a rescaled beta distribution.

### 4.3.2. Error parameters

For the analysis neglecting the model mismatch (i.e. assuming iid errors, see equations (4.1) and (4.5)), we place a conjugate weakly informative inverse-gamma prior on the error parameter,  $\eta = \{\sigma^2\}$ :  $\sigma^2 \sim \mathcal{IG}(a, b)$ , with  $a = 0.001$  and  $b = 0.001$ , leading to an  $\mathcal{IG}$  posterior distribution.

For the model mismatch analysis (i.e. assuming correlated errors, see equations (4.3) and (4.6)), the hyperparameters of the GP neural network covariance function  $\boldsymbol{\eta} = \{w, b\}$ , are

given a log uniform distribution with the range chosen based on maximizing the profile log likelihood (see electronic supplementary material, S8.4 for details).

## 4.4. Posterior inference with Bayesian methods

The posterior distribution is computed as

$$p(\boldsymbol{\theta}, \boldsymbol{\eta} | \mathbf{y}) \propto p(\mathbf{y} | \boldsymbol{\theta}, \boldsymbol{\eta}) p(\boldsymbol{\theta}, \boldsymbol{\eta}), \quad (4.8)$$

where  $\boldsymbol{\theta}$  are the biophysical parameters and  $\boldsymbol{\eta}$  are the error parameters. In this study, we pursue Bayesian inference based on sampling the biophysical and error parameters from their posterior distribution (equation (4.8)), see electronic supplementary material, S6 for details.

## 4.5. Bayesian model selection: WAIC

WAIC [8] is used for model selection as its computation is straightforward from the MCMC posterior samples (see electronic supplementary material, S1.1 for the mathematical details). Out of a number of candidate models, the model which registers the lowest WAIC score is best supported by the data.

# 5. Simulations

## 5.1. Code

Our statistical methods were implemented in Matlab (Mathworks, Natick, MA) and simulations were run on a RedHat Enterprise Linux 6 machine with Intel(R) Xeon(R) CPU E5-2680 v. 2 2.80GHz and 32GB RAM. The simulated pressure waveforms were obtained by numerically solving the PDEs in equations (3.1)–(3.7) using a two-step Lax–Wendroff scheme [36] implemented in  $C^{++}$  by Olufsen *et al.* [4,10], which is second-order accurate in space and time. A spatial and temporal discretization of  $\Delta x = 0.025$  (mm) and  $\Delta t = 1.34 \times 10^{-5}$  (s) were used after testing different discretizations. This satisfied the Courant–Fredrich–Levy (CFL) condition and ensured minimal numerical error. We implemented the GP models

using the GPstuff toolbox [37] and the MCMC convergence diagnostics using the MCMC toolbox [38].

## 5.2. Set-up

We split the simulations into two categories. The simulations in the first category use the measured MPA pressure data for all the models summarized in §3.3 and tables 1 and 2. Simulations in the second category use synthetic data (20 data instantiations) generated in 1, 3 or 21 vessels from the linear wall model with radius-dependent stiffness and correlated errors (model E in table 2) created using a GP with a neural network kernel.

## 5.3. Computational efficiency

For some of the simulations in the first category, models A, B and H, summarized in tables 1 and 2, we focus on computational efficiency, thus we implement MCMC with a GP surrogate (emulator) for the posterior distribution [12]. This approach is motivated by the high computational complexity of repeated numerical integrations of the PDEs in the Bayesian analysis. The method, described in detail in electronic supplementary material, S7, significantly speeds up computationally expensive simulations, which is essential if the analysis performed here is to be translated to human medicine and practical clinical decision support. Constructing emulators for models with large numbers of parameters poses computational challenges due to the large dimension of the input space that has to be covered, which is beyond the remit of the present paper.

## 6. Results

### 6.1. Importance of correcting for model mismatch

We compare inference results based on MCMC between the conventional method ignoring model mismatch and our proposed approach, which explicitly incorporates the model mismatch, defined in equation (4.7), with GPs. The results are shown for synthetic and physiological data. Convergence of MCMC methods was tested using the Geweke test [39] (the  $p$ -values from the  $Z$  test were greater than 0.05) and the Brooks multivariate potential scale reduction factor (MPSRF) [40] (ensuring that MPSRF  $\leq 1.1$ ).

#### 6.1.1. Synthetic data

We generated synthetic data using model E (table 2) with additive, correlated Gaussian errors in the MPA (20 different data instantiations), which mimics the physiological data (as described in §2.2). We then ran two MCMC simulations: one which incorporates the model mismatch, and another simulation which does not. Parameter estimates obtained from these two simulations are compared to the ground truth parameter values in table 3 using the relative sum of squared errors (SSE),

$$\sum_{i=1}^k \left( \frac{\theta_i - \hat{\theta}_i}{\theta_i} \right)^2, \quad (6.1)$$

which is the relative deviation in Euclidean space of the estimated values from the true parameter values. We also show the median marginal and joint posterior density value of the true parameter vector,  $\theta = (f_1, f_2, f_3, \psi_1, \psi_2, c)$ , under the assumed model, as found from 20 synthetic datasets. To

**Table 2.** Models analysed for the measured data: the constitutive models (linear and nonlinear) with model parameters ( $f_1, f_2, f_3, \gamma, \psi_1, \psi_2, c$ ) prior ranges. We indicate whether the model mismatch is incorporated (by yes or no), and if it is, the hyperparameters  $w$  and  $b$  for the GP model mismatch are given. In the right-most column, the symbol '+' indicates that the emulation approach was used to accelerate the MCMC simulations, while '-' indicates that the standard MCMC was used. The stiffness relation used is given in equation (3.5).

model abbreviation	model	no. parameters	$f_1$ ( $\times 10^3$ )	$f_2$	$f_3$ ( $\times 10^4$ )	$\gamma$	$\psi_1$	$\psi_2$	$c$	model mismatch	$w$ ( $\times 10^4$ )	$b$	emulator
A	linear	4	0	not used	(2, 10)	—	(0.05, 2.50)	(0.05, 2.50)	(0.05, 2.50)	no	—	—	+
B	linear	6	0	not used	(2, 10)	—	(0.05, 2.50)	(0.05, 2.50)	(0.05, 2.50)	yes	(1, 9)	(1, 500)	+
C	linear	6	0	not used	(2, 10)	—	(0.05, 2.50)	(0.05, 2.50)	(0.05, 2.50)	yes	(1, 9)	(1, 500)	—
D	linear	6	(1, $10^4$ )	(-300, -50)	(3, 6)	—	(0.05, 2.50)	(0.05, 2.50)	(0.05, 2.50)	no	—	—	—
E	linear	8	(1, $10^4$ )	(-300, -50)	(3, 6)	—	(0.05, 2.50)	(0.05, 2.50)	(0.05, 2.50)	yes	(1, 9)	(1, 500)	—
F	linear	24	0	not used	(2, $10^*$ )	—	(0.05, 2.50)	(0.05, 2.50)	(0.05, 2.50)	no	—	—	—
G	linear	26	0	not used	(2, $10^*$ )	—	(0.05, 2.50)	(0.05, 2.50)	(0.05, 2.50)	yes	—	—	—
H	nonlinear	7	0	not used	(3, 50)	(1, $2\pi$ )	(0.05, 2.50)	(0.05, 2.50)	(0.05, 2.50)	yes	(1, 9)	(1, 500)	+
I	nonlinear	9	(5, $10^2$ )	(-200, 0)	(1, 50)	(1, $2\pi$ )	(0.05, 2.50)	(0.05, 2.50)	(0.05, 2.50)	yes	(1, 9)	(1, 500)	—

Legend: \* in column  $f_3$  indicates that 90% prior probability has been placed on these bounds as part of a Bayesian hierarchical scheme (figure 2) to infer 21 individual vessel stiffness parameters.

**Table 3.** Results obtained when allowing for or ignoring model mismatch, defined in equation (4.7), on synthetic data generated from model E in table 2 with correlated errors. First row: standard approach ignoring model mismatch; second row: the proposed new method, where a GP mismatch model has been introduced. The relative sum of squared errors (SSE), as well as the median posterior distribution of the true parameter vector,  $\theta = (f_1, f_2, f_3, \psi_1, \psi_2, c)$ , under the assumed model are presented (median calculated from 20 datasets). Marginal and joint posteriors were obtained from the MCMC samples with kernel density estimation (first entry in last column) and Chib's method [41] (second entry in last column). Parameters were scaled to the same order of magnitude.

allow for model mismatch	SSE	median $p(\theta_i y)$	median $p(\theta y)$
no	0.02	$(9.8 \times 10^{-08}, 0.004,$ 0, 0, 0, 0)	0, 0
yes	0.03	$(9.7 \times 10^{-08}, 0.004,$ 0.0001, 4.36, 5.25, 1.30)	1037, 648

obtain the marginal posterior density of the true parameter vector, we used the kernel smoothing function estimate for univariate data with the optimal bandwidth for normal densities [42]. To check for consistency of the results, the joint posterior density was obtained in two ways: using the multivariate kernel density estimation with the bandwidth estimated with Silverman's rule [43], and using Chib's method (see §2.1 in [41]). The parameters were scaled to the same order of magnitude, as both methods were affected by having parameters with different orders of magnitude. Figure 3 displaying the marginal posterior density values of the parameters for three of the 20 datasets shows that with the standard method neglecting model mismatch, the ground truth parameter value lies in the tail of the posterior distribution for most cases investigated, and the posterior uncertainty is underestimated. By contrast, with our proposed method allowing for model mismatch, the posterior distribution contains the true parameter and the posterior uncertainty is wider. For the complete set of results, we refer to table 3, which shows that neglecting model mismatch leads to a lower (better) relative SSE for the parameter estimates. However, as seen from figure 3, a small SSE does not rule out the possibility of seriously underestimating the uncertainty. A better measure, to capture both estimation accuracy and UQ, is the marginal posterior density of the true parameters. Here, we see that the median marginal posterior density value of the true parameter with the standard method is substantially lower (worse) for the identifiable parameters  $f_3$ ,  $\psi_1$ ,  $\psi_2$  and  $c$ .

### 6.1.2. Physiological data

For the physiological data, table 4 shows comparative results from the MCMC analysis for the nine models explored. We present the median posterior density value (50th posterior quantile) for each of the models' parameters, and the associated 95% posterior credible interval obtained from the 2.5% and 97.5% quantiles of the MCMC posterior samples. WAIC scores, calculated from 1000 random

MCMC samples, are compared to the Euclidean distance (equation (4.2)) obtained with the median posterior parameter value. A lower WAIC score indicates a better model.

Models incorporating the model mismatch, defined in equation (4.7), record a lower (better) WAIC and a higher (worse) Euclidean distance in output space compared to models which ignore it, implying that the former are better supported by the data, and that minimizing the Euclidean distance, which is equivalent to minimizing the mean squared error (MSE), is a sub-optimal inference procedure. The reason is that the MSE does not take the error correlation into account and does not penalize models for poor UQ. This is illustrated in figure 4 showing that the posterior uncertainty in parameter space is much wider when allowing for model mismatch, which aligns with findings from the synthetic study. Moreover, parameters  $f_3$  and  $\psi_1$  have different posteriors depending on whether the model mismatch is incorporated (figure 4). Additionally, figure 5 displays the posterior uncertainty in output space for pressures in several vessels using the linear model with constant stiffness ignoring or correcting for the model mismatch (models A and B in table 2). Similar to the posterior uncertainty in parameter space, the posterior uncertainty in output space is much wider when correcting for the model mismatch, which is also shown in electronic supplementary material, table S1; there we provide the time-averaged 95% explanatory and predictive credible interval width for the pressure data from every model.

### 6.2. Parameter posteriors

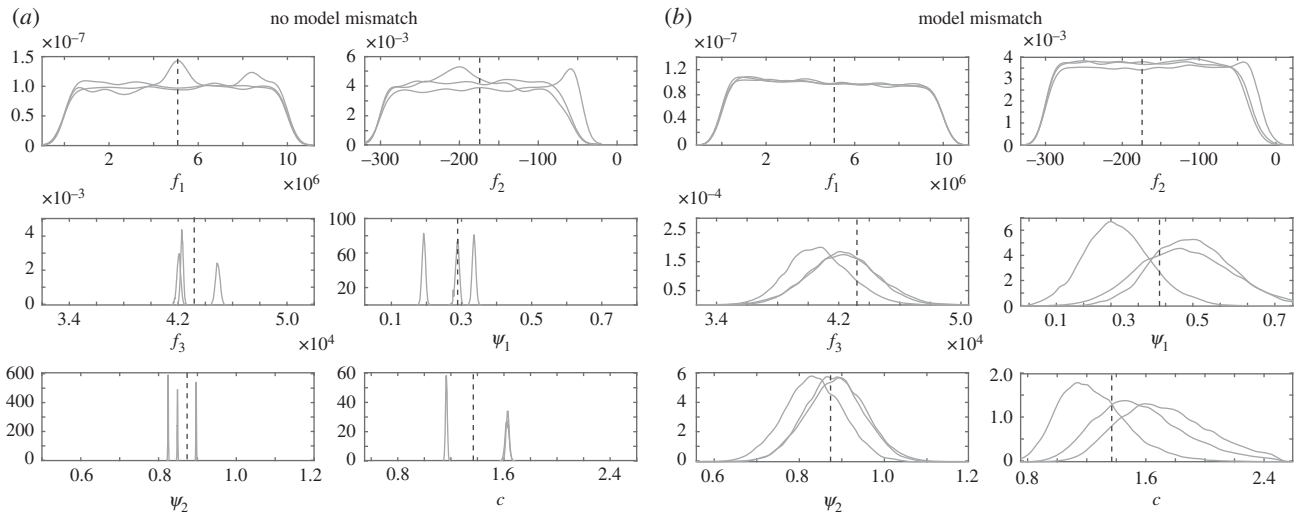
Figure 4 shows the posterior correlations and the marginal posterior densities for the linear models A, and B and nonlinear model I. The marginal posterior distributions have one clear mode and correlations between the parameters are negligible for the linear models. When the nonlinear model is used, we plot  $\chi(f_1, f_2, f_3)$  in equation (3.5) rather than the individual  $f_1, f_2, f_3$  parameters, since the term  $f_2 r_0$  is close to 0 (compare prior range for  $f_2$  in table 2 to posterior uncertainty interval in table 4), leading to unidentifiability of  $f_1$  and  $f_3$ .

### 6.3. Vessel wall model

Table 4 indicates that the nonlinear wall model outperforms the linear wall model, since the former registers a lower (better) WAIC score than the latter (linear model G: -4522, nonlinear model I: -4530). More specifically, results suggest that the nonlinear model with radius-dependent stiffness (model I) is preferred, as it registers the lowest (best) WAIC score computed based on the MPA-measured pressure data.

Based on the WAIC scores in table 4, we conclude that out of all the models investigated, the best linear wall model is that with vessel-specific stiffness (model G), and the best nonlinear model is that with an exponential radius-dependent stiffness (model I), incorporating the model mismatch in both models. The nonlinear model with vessel-specific stiffness is not considered due to the interaction between the parameters  $\chi$  and  $\gamma$  in equation (3.4), requiring vessel-specific ( $\chi, \gamma$ ). This would lead to a large number of parameters being estimated, requiring demanding computational efforts (simulations would most likely take months to complete.<sup>2</sup>)





**Figure 3.** Inference results for three synthetic datasets generated from model E in table 2 with correlated errors using (a) the standard method, which ignores model mismatch, defined in equation (4.7), and (b) a GP introduced to allow for model mismatch. We show marginal posterior densities for the parameters of the exponential radius-dependent stiffness linear model ( $f_1$ ,  $f_2$ ,  $f_3$ ,  $\psi_1$ ,  $\psi_2$ ,  $c$ ), where the stiffness is given by equation (3.5). The different densities per parameter correspond to three random datasets out of 20 (for complete results, see table 3). The black dashed vertical line marks the ground truth parameter values which generated these data.

#### 6.4. Vessel wall stiffness

Table 4 shows a lower WAIC score for the linear wall model with vessel-specific stiffness (model G) relative to the other linear models, which assume constant or radius-dependent stiffness (B, E). The exponential radius-dependent stiffness model (E) has the same WAIC score as the constant stiffness model (B), suggesting that the additional model complexity is not beneficial, as the  $f_1$  and  $f_2$  parameters are non-influential (their marginal posterior distributions are uniform on the prior range), see the left panel of figure 6. In addition, the 21 stiffness model reveals that the median posterior stiffness values are nearly all similar (right panel of figure 6). However, the stiffness becomes increasingly variable for small-radius vessels, which is also evident from the 95% credible interval width presented in electronic supplementary material, S9. This suggests that the Bayesian hierarchical model should allow for vessel-specific variance, thus the common variance  $\sigma_x^2$  in figure 2 should be replaced by a variance-covariance matrix. This method extension will lead to a substantial increase in the computational complexity due to higher parameter space dimension, and thus, it is subject to future work. Almost all the stiffness values are of the order  $10^4$ , which is the regime where the MPA systolic, diastolic and pulse pressures are sensitive to changes in stiffness (figure 7). The plot is produced with a set of scaling parameters consistent with the physiological data ( $\psi_1 = 0.30$ ,  $\psi_2 = 0.97$ ,  $c = 1.23$ ).

In addition, table 4 shows that the stiffness in the nonlinear model has only a weak dependence on the radius (model I), as expressed in equation (3.5), since the term  $f_2 r_0$  is close to 0 (compare prior range for  $f_2$  in table 2 to posterior uncertainty interval in table 4).

#### 6.5. Model fits

Figures 8 and 9 show the model fits for all the models analysed. The median pressure predictions obtained using the MCMC-simulated posterior parameter values (figure 8) are qualitatively similar for all nine models investigated. Pressure

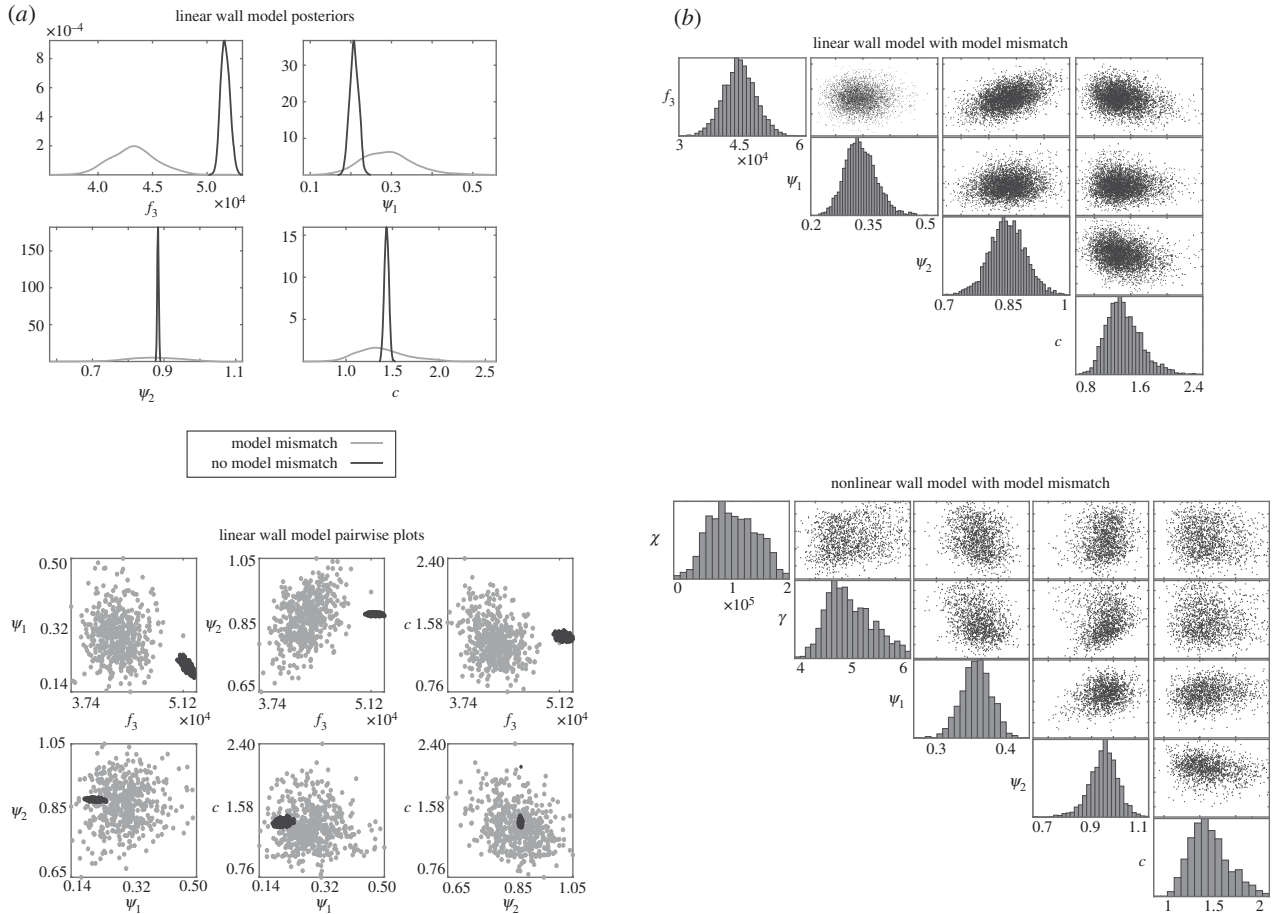
predictions in the MPA are compared between all models, all producing a waveform similar to the measured data. The best linear model (model G) fits the measured data better in the diastolic phase, but gives a peak shift in the systolic phase (figure 8). On the other hand, the best nonlinear model (model I) provides a better fit in the systolic phase, but has a slight discrepancy in diastole. Generally, the pressure increases more steeply in the systolic phase for the nonlinear model compared to the linear model.

Figure 8 also shows that the pressure predicted with the nonlinear model is slightly higher than that predicted with the linear model for all 21 vessels. In addition, the predictions obtained with the linear model with 21 individual stiffnesses while ignoring model mismatch (model F) provides fits similar to the other models in the proximal arteries, but predicts downstream pressure waves with different shape and increased oscillatory behaviour. This suggests that assuming a vessel-specific stiffness and ignoring the model mismatch provides poor extrapolation performance. When analysing the median flow predictions obtained from the parameter posteriors (figure 9), we note that all are again very similar in shape, except for the outlier model F. We observe an unequal flow distribution between the right and left side of the tree for all the models except model F.

Figure 9 shows predicted pressure–area relations using the posterior median parameter values. The best nonlinear model (model I) consistently predicts larger areas than the best linear model (model G) for the proximal vessels (see predictions in MPA and vessels 2 and 3), and the opposite trend is observed for most terminal vessels (see predictions in electronic supplementary material, S9). Furthermore, we observe that the nonlinear model with constant stiffness (model H) gives systematically larger area values than the nonlinear model with radius-dependent stiffness (model I) except in vessels 1 and 3, which aligns with the former model having a smaller stiffness than the latter (table 4). The linear model with 21 individual stiffnesses without model mismatch (model F) gives drastically different results than the other models for some of the vessels (e.g. vessels 4, 8, 12 in

**Table 4.** Summary of the MCMC simulation results on measured data for the constitutive models considered (linear and nonlinear) with model parameters ( $f_1, f_2, f_3, \gamma, \psi_1, \psi_2, c$ ). We indicate whether model mismatch, defined in equation (4.7), is incorporated (by yes or no), and if it is, the parameters  $w$  and  $b$  for the GP model mismatch are given. In the right-most column, the symbol ‘+’ indicates that the emulation approach was used to accelerate the MCMC simulations, while ‘-’ indicates that the standard MCMC was used. We have run 5000 MCMC iterations for the models which use the emulator (models A, B and H); 300 000 for the vessel-specific stiffness models which do not use the emulator (models F and G); and 150 000 MCMC iterations for the rest of the models which also do not use the emulator (models C, D, E, I) (the emulator approach needs much fewer iterations—see the electronic supplementary material, S7 for an explanation). We indicate the median posterior distribution value, as well as the 95% credible interval based on the posterior distribution. We display the WAIC score calculated from 1000 MCMC samples and the Euclidean distance obtained based on the posterior median parameter values. The exponential stiffness relation used is given in equation (3.5). If 21 individual stiffness parameters are inferred (figure 2), marked by \* in the table, we list the stiffness values in electronic supplementary material, S9.

model abbreviation	model	$f_1$ ( $\times 10^5$ )	$f_2$	$f_3$ ( $\times 10^4$ )	$\gamma$	$\psi_1$	$\psi_2$	$c$	model mismatch	$w$ ( $\times 10^4$ )	$b$	WAIC	Euclidean distance	emulator
A	linear	0	not used	5.17 (5.09, 5.25)	—	0.21 (0.19, 0.23)	0.88 (0.88, 0.89)	1.44 (1.39, 1.49)	no	—	—	408	66	+
B	linear	0	not used	4.31 (3.91, 4.72)	—	0.28 (0.18, 0.42)	0.87 (0.73, 1.00)	1.35 (0.98, 1.96)	yes	5.36 (4.32, 7.09)	137 (102, 197)	-4515	551	+
C	linear	0	not used	4.31 (3.91, 4.71)	—	0.29 (0.18, 0.41)	0.87 (0.73, 1.00)	1.34 (0.97, 1.94)	yes	5.41 (4.27, 7.09)	138 (101, 198)	-4515	571	—
D	linear	52.6 (2.59, 97.8)	-162 (-293, -51.1)	5.17 (5.10, 5.25)	—	0.21 (0.19, 0.23)	0.89 (0.88, 0.89)	1.45 (1.40, 1.50)	no	—	—	397	64	—
E	linear	51.1 (2.80, 97.4)	-171 (-293, -56.6)	4.32 (3.91, 4.72)	—	0.29 (0.18, 0.42)	0.87 (0.74, 1.01)	1.34 (0.96, 1.93)	yes	5.39 (4.28, 7.08)	138 (102, 200)	-4515	552	—
F	linear	0	not used	*	—	2.45 (2.33, 2.50)	0.21 (0.17, 0.26)	2.08 (0.11, 2.48)	no	—	—	-107	26	—
G	linear	0	not used	*	—	0.46 (0.19, 1.05)	0.86 (0.72, 1.10)	1.12 (0.45, 1.86)	yes	4.65 (3.72, 6.25)	161 (115, 262)	-4522	242	—
H	nonlinear	0	not used	9.17 (6.66, 12.1)	5.18 (4.38, 6.07)	0.37 (0.27, 0.47)	0.94 (0.80, 1.06)	1.60 (1.15, 2.27)	yes	4.91 (3.97, 6.44)	178 (126, 286)	-4522	384	+
I	nonlinear	0.58 (0.20, 0.97)	-6.00 (-18.0, -2.13)	2.00 (1.09, 3.40)	5.09 (4.43, 6.13)	0.34 (0.24, 0.43)	0.97 (0.82, 1.09)	1.58 (1.16, 2.27)	yes	4.45 (3.62, 5.76)	196 (130, 320)	-4530	417	—



**Figure 4.** (a) Marginal posterior densities (top) and pairwise scatterplots (bottom) of the posterior sample (obtained with MCMC) for the constant stiffness linear wall model with the standard method ignoring the model mismatch (equation (4.7)) (black) versus our proposed GP mismatch model (grey), i.e. models A and B in table 2. (b) Pairwise scatterplots between the MCMC posterior parameter samples of the linear model with constant stiffness and model mismatch (top), i.e. model B in table 2, and nonlinear model with radius-dependent stiffness and model mismatch (bottom), i.e. model I in table 2. For the nonlinear model, we express  $\chi(f_1, f_2, f_3)$  in equation (3.5) instead of individual parameters  $f_1, f_2, f_3$  due to parameter identifiability issues, see §6.2 for a discussion on this. Here, we show the distribution of  $\chi(f_1, f_2, f_3)$  in equation (3.5) for radius  $r_0$  corresponding to vessel 1, the MPA, but the pattern of the distribution is similar for the other vessels.

electronic supplementary material, S9), which further indicates that this model can lead to drastic changes in downstream predictions.

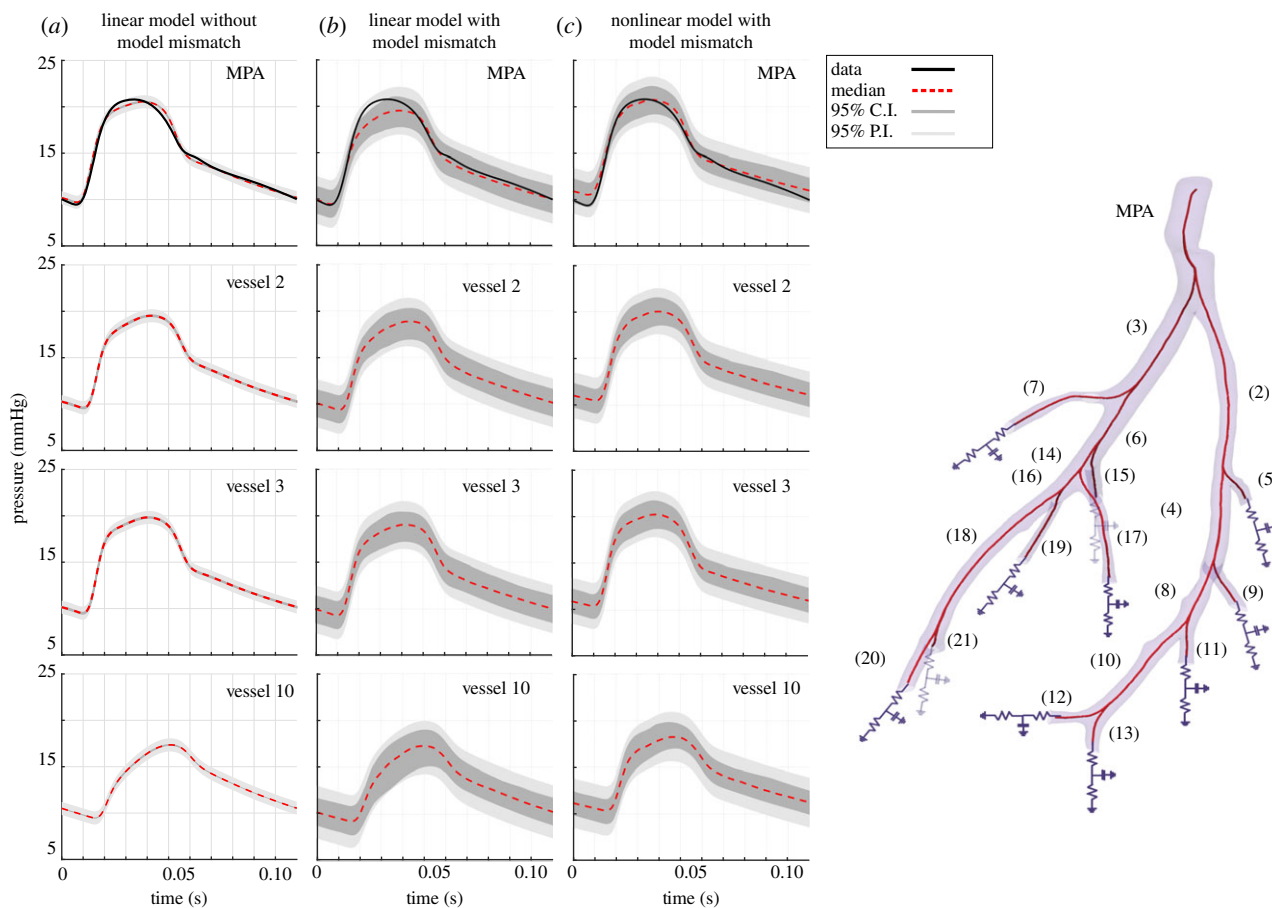
## 6.6. Future experimental design

We test by a synthetic study if parameters  $f_1$  and  $f_2$  in model E, i.e. the linear model with exponential radius-dependent stiffness (equation (3.5)) become influential as complementary data from downstream vessels are added. We generated synthetic data from this model and, as described in §2.2, we added additive correlated Gaussian errors to them. We created 20 synthetic datasets and applied MCMC to infer the data-generating parameter values. In figure 10, we show the agglomerated MCMC posterior distributions from all 20 data instantiations, and superimpose the true parameter values (the peak of the agglomerated distributions should coincide with the true parameter values). We stress that this is purely for visualization purposes, since agglomerated results over different datasets is a non-conventional Bayesian approach. For a fully Bayesian approach, we calculate the marginal posterior distribution, as well as the joint posterior distribution of the true parameter for each of the datasets, and find the median over the datasets, shown in table 5. For the joint posterior distribution, we used the multivariate kernel density estimation. Figure 10 shows that the

peak of the agglomerated distributions aligns with the ground truth parameter values for the influential parameters, validating our inference procedure. Even with data from more than one vessel (3 or 21 vessels),  $f_1$  and  $f_2$  parameters remain non-influential (close to uniform marginal posterior density). The uncertainty for all the other parameters ( $f_3$ ,  $\psi_1$ ,  $\psi_2$  and  $c$ ) is reduced and the distributions become increasingly focused around the true parameter values with increasing complementary data. Additionally, in table 5, we quantify how the marginal and joint posterior densities values of the true parameters increase with the amount of vessel data.

## 6.7. Accuracy of emulator

Table 4 and figure 11 show that the parameter inference results obtained with the emulation approach (model B) are comparable with those obtained with the conventional method (model C), i.e. the marginal posterior densities overlap. This suggests that no bias is introduced by the emulator. This finding is further confirmed by similar results in output space between the two approaches, summarized in electronic supplementary material, table S1 and figure S11, i.e. the median pressure signals and 95% posterior predictive intervals are very similar.



**Figure 5.** Ninety-five per cent credible intervals (C.I.) and prediction intervals (P.I.) for the pressure prediction from the linear model with one stiffness and ignoring model mismatch in equation (4.7) (left) (model A in table 2), the linear model with one stiffness and model mismatch correction (centre) (model B in table 2), and the nonlinear model with radius-dependent stiffness and model mismatch correction (right) (model I in table 2), obtained from MCMC posterior samples. We superimpose the measured pressure data in the MPA and the median prediction, and show plots in three other vessels (for all the other vessels, see electronic supplementary material, S9). This figure illustrates the predictive performance of these models (A, B, I) and helps visualize pressure profiles in parts of the vessel network for which no experimental data are available. The figure is not used to justify our model selection results, for which table 4 should be used.

## 7. Discussion

In this study, we have explored several mathematical models of the pulmonary circulation (a linear and a nonlinear wall model with different vessel wall stiffness assumptions) and two error models capturing or ignoring the model mismatch, as defined in equation (4.7). We have used Bayesian inference to find the model that can best predict the measured MPA blood pressure, while providing UQ associated with that pressure prediction. We have also tested the validity of our parameter inference procedure by a synthetic study.

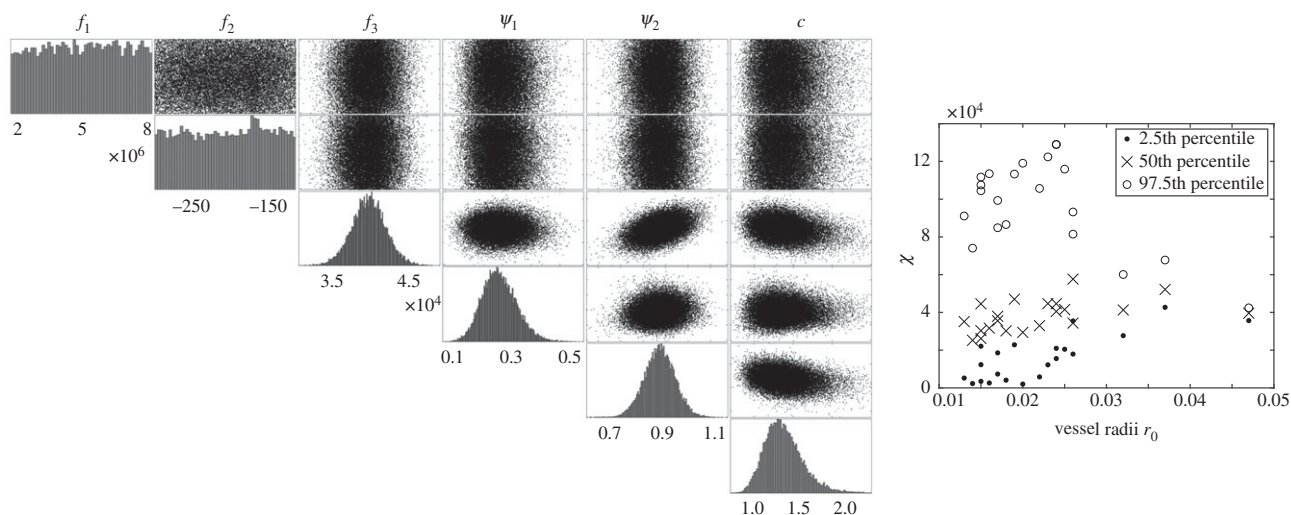
### 7.1. Importance of correcting for model mismatch

Neglecting the model mismatch by obtaining point estimates based on MSE minimization biases parameter estimates and underestimates uncertainty in parameter and output space. This finding is based on synthetic data with known parameters, and tallies with results from the physiological data. The model mismatch is a consequence of wrong measurement error assumptions (i.e. iid for correlated measurement errors) and ignoring model discrepancy between the real system and the mathematical model. In this study, we proposed a method based on GPs to allow for the model mismatch, which circumvents the limitations outlined above. Figure 3 clearly illustrates that the posterior uncertainty in parameter space is under-dispersed when the

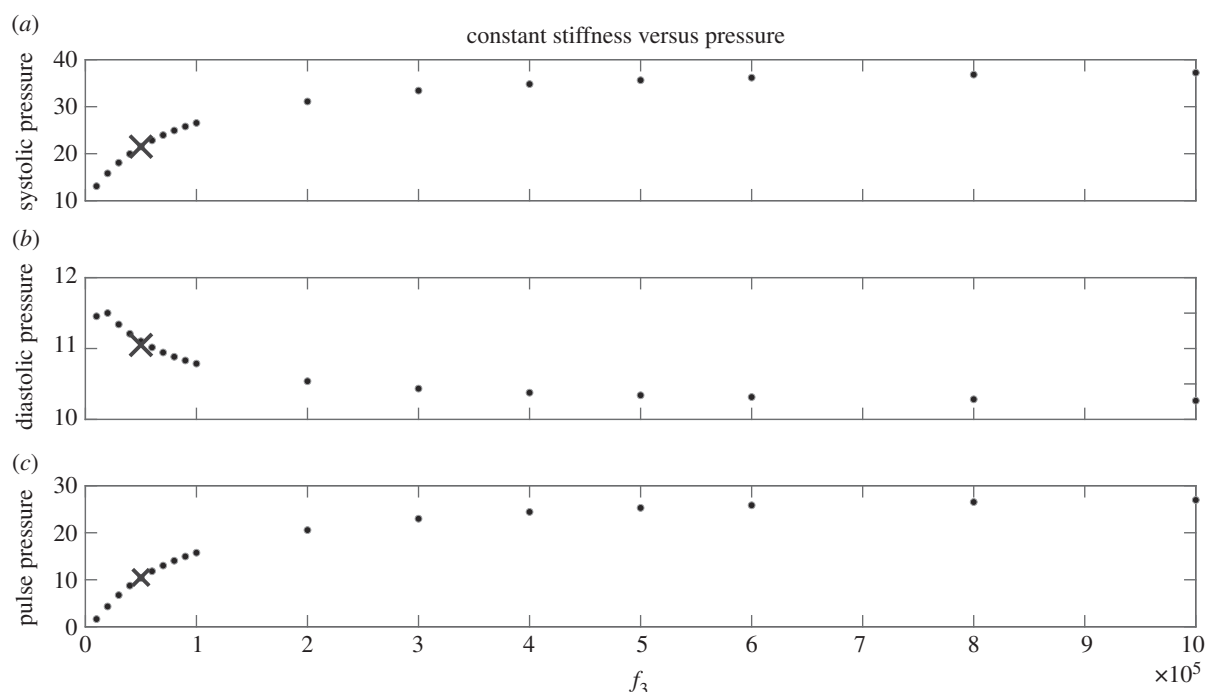
model mismatch is neglected, and the true (data-generating) parameter values lie in the tail of the posterior distribution for most datasets; however, this is not the case for our proposed method of model mismatch. This is in line with results from the measured data, as evident in figure 4a, that also shows very narrow uncertainty bounds in parameter space and in output space (figure 5). Moreover, the model selection criteria (WAIC in table 4), clearly and consistently favours the models with model mismatch, further strengthening our statement.

We emphasize that most studies in the literature rely on minimizing the Euclidean distance (i.e. MSE) in equation (4.2), which implicitly ignores the model mismatch. This approach is equivalent to maximizing the likelihood in equation (4.1) under the assumption of additive Gaussian iid errors. However, in the presence of model mismatch, the estimates that minimize the MSE are different from the estimates which maximize the likelihood in equation (4.3). Our work demonstrates that ignoring model mismatch leads to biased point estimates, and thus incorrect predictions and uncertainty underestimation. Wider uncertainty bounds in output space, as seen in figure 5, reflect more adequately variations in pulmonary pressure due to the natural inter-subject factors (e.g. effects of the respiratory cycle). These are well known [6] and should be contained within the uncertainty bounds of the model.





**Figure 6.** Left panel: MCMC results (marginal posterior densities and scatterplots) based on the linear wall model with exponential stiffness (equation (3.5)) and correcting for model mismatch (defined in equation (4.7)), for measured data. Right panel: MCMC results for 21 individual stiffness values corresponding to every vessel radius  $r_0$  for the linear wall model and correcting for model mismatch. Here, we show the posterior median value for each vessel stiffness parameter, and the 2.5th and 97.5th quantiles.



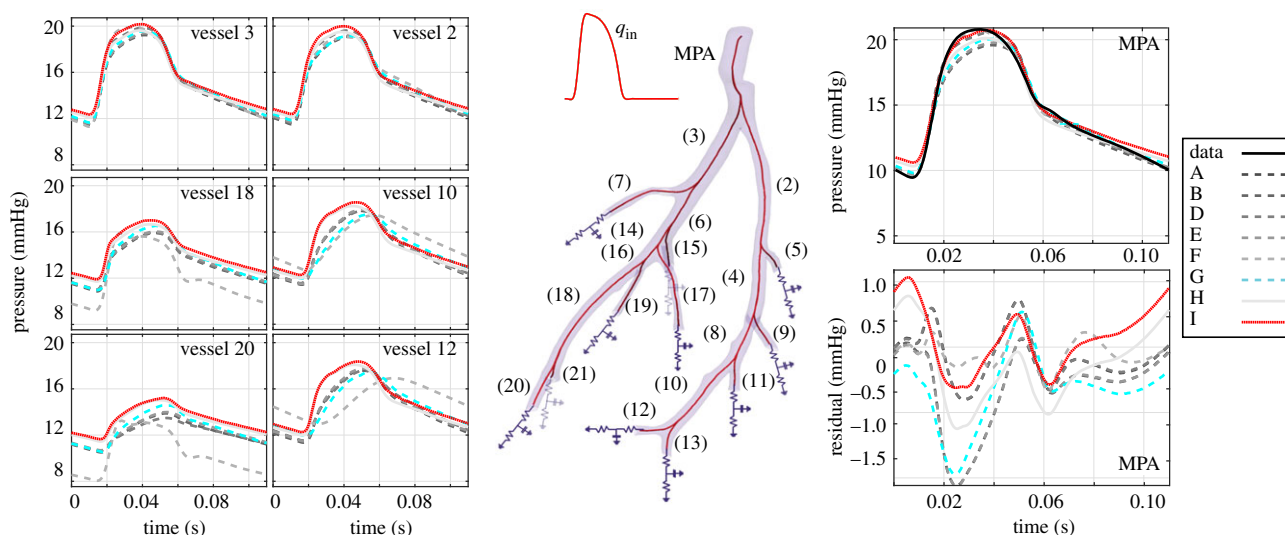
**Figure 7.** Relation between the radius-constant stiffness and the systolic, diastolic and pulse pressure for the linear wall model. Here, we vary  $f_3$  within the range  $[10^4, 10^6]$ ,  $f_1$  is set to 0,  $f_2$  can take any value in the exponential radius-dependent expression in equation (3.5), and the Windkessel parameters  $\psi_1$ ,  $\psi_2$ ,  $c$  are kept fixed to 0.30, 0.97, 1.23, which are plausible values for the measured data. A similar trend is observed when the Windkessel parameters are fixed to other values, or when the nonlinear wall model is used. We mark by a cross point the  $f_3$  stiffness value estimated from the measured data using the constant stiffness linear model.

To our knowledge, our study is one of the first to focus on parameter estimation in cardiovascular modelling while incorporating model mismatch. A notable exception is Lei *et al.* [19], who explore model discrepancy in cardiac electrophysiology, and the authors also show through synthetic studies that ignoring the model-form uncertainty produces biased predictions and uncertainty underestimation, which agrees with our findings.

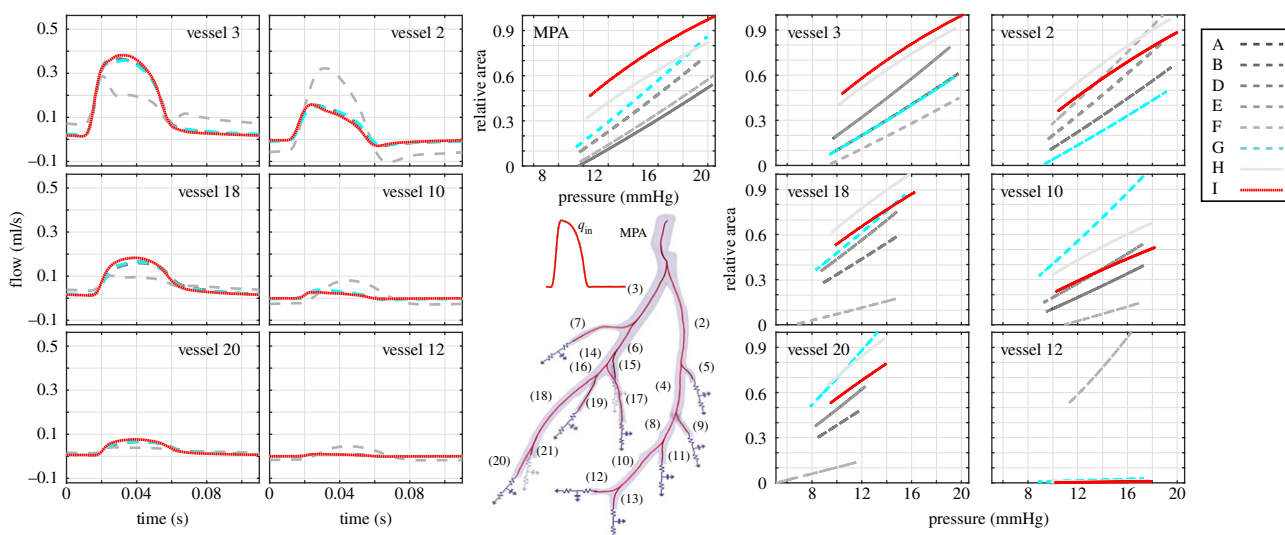
## 7.2. Vessel wall model

The WAIC scores in table 4 suggest that the nonlinear wall model (model I) is better supported by the MPA-pressure

data than the linear wall model (model G), since the former model registers a lower WAIC score. Our analysis indicates that, out of all the models investigated, the model that is most likely under the data is the nonlinear model with a slight dependence on the vessel radius, i.e. model I, recording the lowest WAIC score. This finding agrees with other studies—e.g. the study by Valdez *et al.* [44] on pressure area dynamics in systemic arteries of control sheep and the study by Pillhwa *et al.* [7] analysing distensibility of pulmonary arteries in control mice. The study in [45] provided experimental stress–strain relations in control and hypoxic pulmonary arteries, illustrating a predominant viscoelastic effect and further suggesting that a nonlinear elastic wall is



**Figure 8.** Pressure predictions obtained using the MCMC posterior samples for the parameters from all the models considered—see §3.3 and tables 2 and 4 for a summary of the models, which are denoted by A–I in the figure legend. We show the median pressure signal for seven of the 21 blood vessels in time (see electronic supplementary material, S9 for all the other vessels). We superimpose the measured pressure data in the MPA (top right). Examples of pressure residuals, that is, the difference between the predicted and measured blood pressure, are shown in the bottom right panel.



**Figure 9.** Flow (left side) and pressure–area (right side) predictions obtained using the MCMC posterior sample for the parameters from all the models considered, see §3.3 and tables 2 and 4 for a summary of the models, which are denoted by A–I in the figure legend. We show the median flow predictions and pressure versus standardized cross-sectional area predictions for seven of the 21 blood vessels (see electronic supplementary material, S9 for all the other vessels). The area,  $A_i$  is standardized per vessel  $i$  to lie between  $[0,1]$  using the expression:  $(A_i - l_i)/(u_i - l_i)$ , where  $l_i, u_i$  are the maximum and minimum area value for vessel  $i$ , listed in electronic supplementary material, S9.

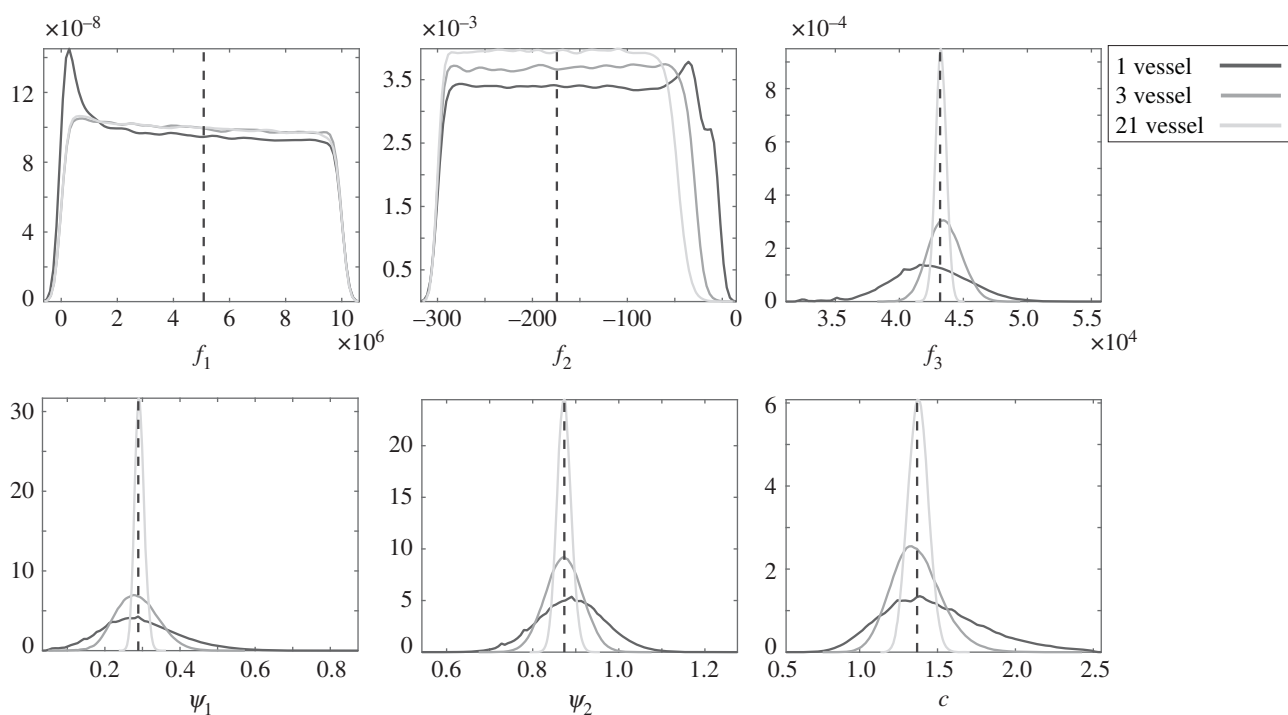
more appropriate for modelling pulmonary haemodynamics. A visual comparison of the measured data with the predictions obtained using the linear or nonlinear model (figure 5) cannot be used as an objective metric to choose between models since the model which gives the closest prediction in Euclidean space to the measured data is not necessarily the most consistent with the data (e.g. for misspecified mechanistic models). This emphasizes the need to carry out a formal model selection analysis.

### 7.3. Vessel wall stiffness

Out of all linear models, model G, i.e. the model with vessel-specific stiffness (with model mismatch) describes best the pressure data (table 4). As expected physiologically, an increased wall stiffness leads to increased systolic and pulse

pressures (figure 7), with more dynamic changes in these values occurring at a lower stiffness range. Table 4 shows that the estimated stiffness in the linear models are within this range, suggesting accurate depiction of healthy haemodynamics, regardless of stiffness model. Additionally, given the data, an exponential radius dependence of the stiffness (model E) leads to non-influential parameters with nearly flat posteriors (figures 6 and 10).

Regarding the nonlinear models, our model selection results support a slight radius dependent stiffness—model I. Previous investigations [46] have shown that both wall thickness ( $h$ ) and tissue properties ( $E$ ) are drastically different in pulmonary arteries in pulmonary hypertension. This encourages future investigations into whether our model selection results are consistent in specimens with pulmonary hypertension.



**Figure 10.** Synthetic data results obtained by agglomeration of MCMC posterior samples over 20 data instantiations. We show marginal posterior densities for the parameters  $f_1$ ,  $f_2$ ,  $f_3$ ,  $\psi_1$ ,  $\psi_2$ ,  $c$  of the linear wall model with exponential stiffness,  $\chi(f_1, f_2, f_3)$  given in equation (3.5). We superimpose results for three simulations: one which uses synthetic data from one vessel (MPA) for inference—dark grey line, a second one which uses data from three vessels (MPA and its two daughter vessels)—medium grey line and a third one which uses data from all 21 vessels—light grey line. The black dashed vertical line marks the ground truth parameter values which generated these data.

**Table 5.** Inference results obtained using synthetic data, to which we added additive, correlated Gaussian error, from one vessel (MPA), three vessels (MPA and its two daughter vessels) and all 21 vessels. The model mismatch was included in the analysis, and the data were generated using the linear wall model with exponential stiffness,  $\chi(f_1, f_2, f_3)$ , given in equation (3.5). The median marginal and joint posterior density, of the true parameter vector,  $\theta = (f_1, f_2, f_3, \psi_1, \psi_2, c)$  are presented for each of the three scenarios (median calculated from 20 datasets). Joint and marginal posteriors were computed using the MCMC samples with kernel density estimation. Parameters were scaled to the same order of magnitude.

data	median $p(\theta y)$	median $p(\theta y)$
one vessel	$(9.8 \times 10^{-08}, 0.004, 0.0001, 4.30, 5.22, 1.30)$	$2.8 \times 10^{+03}$
three vessels	$(9.9 \times 10^{-08}, 0.004, 0.0003, 6.78, 9.61, 2.61)$	$3.4 \times 10^{+04}$
21 vessels	$(9.9 \times 10^{-08}, 0.004, 0.001, 38.4, 27.0, 6.39)$	$3.9 \times 10^{+06}$

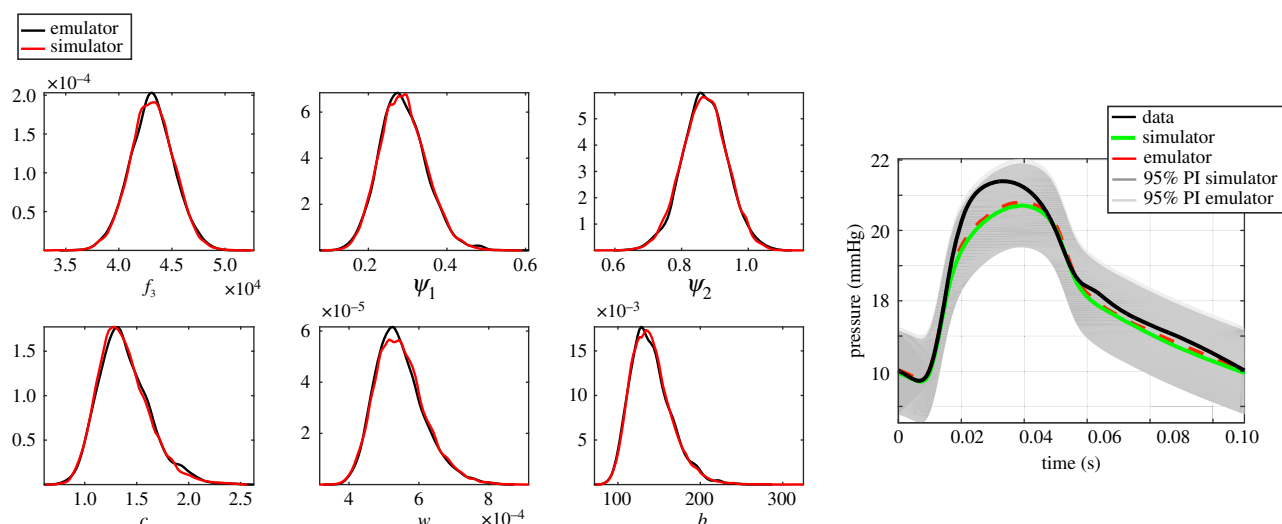
#### 7.4. Model fits

The pressure predictions shown in figure 8 deviate from model to model in arteries distal to the left and right pulmonary artery (vessels 2 and 3, respectively). While predictions look qualitatively similar, it is clear that the model used can lead to significant changes in downstream predictions. An understanding of how model type affects predictions down the pulmonary arteries is critical for future use of

mathematical models in disease prognostication. For instance, pulmonary diseases like pulmonary hypertension remodel smaller arteriolar segments initially, making vessel stiffness a critical parameter in the development of disease [47–49]. The flow and pressure–area graphs show a more dramatic change between model types, which is expected as distal flow and dynamic area data are not available. This variability is important when considering lesions, i.e. pulmonary emboli, that can lead to obstructions in the pulmonary arteries, limiting perfusion to the alveoli for blood re-oxygenation. The pressure–area relations in figure 9 show that the inferred parameters for the nonlinear wall-model provide a nearly linear pressure–area curve, contrary to the findings in [45]. We suspect that inference using both pressure and dynamic area data will better illustrate dissimilarities between the two wall models, and hypothesize that additional pressure–flow data in distal arteries would allow for inference of additional Windkessel scaling factors in the network.

#### 7.5. Parameter unidentifiability

Our analysis of the linear model with exponential radius-dependent stiffness shows that using complementary data from downstream vessels does not resolve the unidentifiability of  $f_1$  and  $f_2$  in equation (3.5) (figure 10). Thus, additional pressure data do not carry information about the non-influential parameters. If the model has structural unidentifiabilities, subsequent predictions are unreliable, and can lead to spurious diagnoses or sub-optimal treatments [50,51]. For this reason, it is imperative that in our exponential radius-dependent models, the entire expression  $\chi(f_1, f_2, f_3)$  in equation (3.5) is interpreted, and not the individual parameters,  $f_1, f_2, f_3$ .



**Figure 11.** Uncertainty quantification in input (left panel) and output space (right panel) obtained with the emulation MCMC method (model B in table 2) and the standard MCMC method (model C).

## 7.6. Future experimental design

Our analysis reveals that when complementary data are used, the parameters  $f_3$ ,  $\psi_1$ ,  $\psi_2$  and  $c$  are more accurately estimated (but not  $f_1$  and  $f_2$ ; see §7.5 for a justification), and our study allows the reduction in the estimation uncertainty to be quantified (table 5). This may be used in future experimental design, when deciding whether to record measurements in vessels beyond MPA. Furthermore, results in figure 10 show that the true parameter values are accurately inferred, validating our inference procedure.

## 7.7. Real-time treatment planning

A long-term goal of our project is real-time, personalized treatment planning. Therefore, once the model selection procedure finds the ‘best’ model, predictions from that model should be computationally efficient. We show that this can be accomplished using efficient surrogate models in place of the computationally expensive PDE model (see table 2; electronic supplementary material, S9). In principle, the emulation approach could be performed for the vessel-specific nonlinear model, which was not explored in this study due to very computationally costly simulations. However, emulation would most likely require high efforts due to a large number of model parameters (i.e. 42 vessel-stiffness parameters, three Windkessel factors and two error parameters). These parameters might be unidentifiable, leading to a high-dimensional parameter domain needing coverage by the emulator, requiring a long time for training, as well as the implementation of more sophisticated GPs (e.g. sparse GPs [52]). This is certainly a highly topical problem with formidable methodological challenges for cutting edge machine learning research. However, pursuing this is far beyond the remit of the present work.

## 8. Limitations and future directions

The physiological conclusions from the model selection analysis according to which the nonlinear model is preferred over the linear model is based on the study of just one mouse. Moreover, it would be interesting to compare the performance of the asymptotically based WAIC approach for model

selection to an approach which does not rely on asymptotics, e.g. marginal likelihood [53], whose calculation, however, comes at a significantly higher computational cost.

Given that this is a retrospective data analysis study, with limited data (a single MPA pressure signal), unspecified machine precision for data measurement, unknown data smoothing and averaging technique applied to the raw data, and lack of prior knowledge of the model discrepancy function, it is intrinsically impossible to distinguish and separately incorporate the model discrepancy and the measurement errors (noise) model. In principle, a strongly informative prior on the model discrepancy function could help separate the contributions from the model and measurement errors [54]. Multi-scale vessel wall models that include fluid–structure interactions at individual cell level, or 3D computational fluid-dynamics models, may be too complex for inference, but could refine prior knowledge. Running forward simulations with both high and medium fidelity models for a space filling design in parameter space and then fitting a GP to the differences in output space could give a more realistic prior for future inference applications. However, these higher fidelity fluid-dynamics models come with their own modelling assumptions and non-measurable parameters, hence a prior built on these models may be inherently biased.

This study analyses both boundary condition and vessel stiffness parameters. We only examined one boundary condition model and estimated scaling factors that adjust nominal Windkessel parameters in each terminal vessel. This boundary condition ignores the fractal structure of the downstream vascular, which is accounted for by the structured tree model [5,10,48]. However, these methods can be used with any boundary condition, and will be pursued in the future.

The uncertainty in the vessel network, which is kept fixed, will also be modelled in the future [3,11]. Minor losses at junctions, unaccounted for here, are also subject to uncertainty and can further introduce a model mismatch. This is systematically addressed by the proposed inference scheme, which provides a mechanism to prevent such model mismatch from causing any bias in the parameter estimation and haemodynamic predictions.



Additionally, the MPA inflow boundary condition could be replaced by a coupling of the MPA with a right ventricle model [55].

## 9. Conclusion

Our study uses Bayesian analysis techniques to approximately infer the posterior distribution with the aim to quantify the uncertainty of model parameters and haemodynamic predictions in a 1D fluid-dynamics model of the pulmonary circulation.

Our main contribution is to draw attention to an often neglected source of uncertainty: in the mathematical model form, caused by the discrepancy between the real system and the model, and in the measurements due to the wrong noise model (jointly called ‘model mismatch’).

Additionally, we explored several mathematical models (a linear and a nonlinear wall model with different vessel wall stiffness assumptions: constant, vessel-specific or radius-dependent stiffness), and error models (via the inclusion of a model mismatch). We implemented Bayesian model selection based on WAIC to find the model that can most accurately predict the MPA pressure and provide adequate uncertainty quantification in the pressure predictions.

Our study clearly demonstrates that the widely used least-squares fit method ignores model mismatch, biasing parameter estimates and model predictions, and underestimating uncertainty in parameter and output space. We circumvent these issues by incorporating the model mismatch using GPs.

Additionally, we found that the MPA-measured pressure data best supports the nonlinear wall model with a weak exponential radius-dependent stiffness.

Lastly, our synthetic study validates our inference procedure, identifies those parameters that benefit from complementary data distal to the MPA, and quantifies the reduction in their intrinsic estimation uncertainty, which may help better design future experiments.

**Data accessibility.** This article has no additional data.

**Competing interests.** We declare we have no competing interests.

**Funding.** This work is part of the research programme of the Centre for Multiscale Soft Tissue Mechanics with Application to Heart & Cancer (SofTMech), funded by the Engineering and Physical Sciences Research Council (EPSRC) of the UK, grant reference no. EP/N014642/1. D.H. is supported by a grant from the Royal Society of Edinburgh, award no. 62335. M.S.O. is supported by the National Science Foundation (NSF) grant reference no. NSF-DMS 1615820, and by NSF-DMS 1246991, and M.J.C. is supported by the American Heart Association (AHA) grant no. 19PRE34380459.

**Acknowledgements.** The authors thank Umar Qureshi for his important contribution in the early stages of this work. The data were made available by N. Chesler, Department of Biomedical Engineering, University of Wisconsin, Madison.

## Endnotes

<sup>1</sup><http://vmtk.org>.

<sup>2</sup>For a comparison, the elapsed time required to complete simulations with the linear model G was roughly six weeks.

## References

- Su Z, Hunter K, Shandas R. 2011 Impact of pulmonary vascular stiffness and vasodilator treatment in pediatric pulmonary hypertension: 21 patient-specific fluid–structure interaction studies. *Comput. Methods Programs Biomed.* **108**, 617–628. (doi:10.1016/j.cmpb.2011.09.002)
- Kheifets V *et al.* 2015 Patient-specific computational modeling of blood flow in the pulmonary arterial circulation. *Comput. Methods Programs Biomed.* **120**, 88–101. (doi:10.1016/j.cmpb.2015.04.005)
- Colebank M, Paun L, Qureshi M, Chesler N, Husmeier D, Olufsen M, Ellwein L. 2019 Influence of image segmentation on one-dimensional fluid dynamics predictions in the mouse pulmonary arteries. *J. R. Soc. Interface* **16**, 20190284. (doi:10.1098/rsif.2019.0284)
- Qureshi M, Colebank M, Paun L, Chesler N, Haider M, Hill N, Husmeier D, Olufsen M. 2018 A computational study of pulmonary hemodynamics in healthy and hypoxic mice. *Biomech. Model. Mechanobiol.* **18**, 219–243. (doi:10.1007/s10237-018-1078-8)
- Umar Qureshi M, Vaughan G, Sainsbury C, Johnson M, Peskin C, Olufsen M, Hill N. 2014 Numerical simulation of blood flow and pressure drop in the pulmonary arterial and venous circulation. *Biomech. Model. Mechanobiol.* **13**, 1137–1154. (doi:10.1007/s10237-014-0563-y)
- Clipp R, Steele B. 2012 An evaluation of dynamic outlet boundary conditions in a 1D fluid dynamics model. *Math. Biosci. Eng.* **9**, 61–74. (doi:10.3934/mbe.2012.9.61)
- Lee P, Carlson BE, Chesler N, Olufsen MS, Qureshi MU, Smith NP, Sochi T, Beard DA. 2017 Heterogeneous mechanics of the mouse pulmonary arterial network. *Biomech. Model. Mechanobiol.* **15**, 1245–1261. (doi:10.1007/s10237-015-0757-y)
- Watanabe S. 2010 Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *J. Mach. Learn. Res.* **11**, 3571–3594.
- Krenz G, Dawson A. 2003 Flow and pressure distributions in vascular networks consisting of distensible vessels. *Am. J. Physiol.-Heart Circ. Physiol.* **284**, H2192–H2203. (doi:10.1152/ajpheart.00762.2002)
- Olufsen M, Peskin CS, Kim W, Pedersen EM, Nadim A, Larsen J. 2000 Numerical simulation and experimental validation of blood flow in arteries with structured-tree outflow conditions. *Ann. Biomed. Eng.* **28**, 1281–1299. (doi:10.1114/1.1326031)
- Colebank MJ, Umar Qureshi M, Olufsen MS. 2019 Sensitivity analysis and uncertainty quantification of 1-D models of pulmonary hemodynamics in mice under control and hypertensive conditions. *Int. J. Numer. Methods Biomed. Eng.* e3242.
- Mihaela Paun L, Colebank M, Umar Qureshi M, Olufsen M, Hill N, Husmeier D. 2019 MCMC with delayed acceptance using a surrogate model with an application to cardiovascular fluid dynamics. In *Proc. of the Int. Conf. on Statistics: Theory and Applications (ICSTA'19)*.
- Paun LM, Qureshi MU, Colebank M, Hill NA, Olufsen MS, Haider MA, Husmeier D. 2018 MCMC methods for inference in a mathematical model of pulmonary circulation. *Stat. Neerlandica* **72**, 306–338. (doi:10.1111/stan.12132)
- Schiavazzi D, Baretta A, Pennati G, Hsia T, Marsden A. 2017 Patient-specific parameter estimation in single-ventricle lumped circulation models under uncertainty. *Int. J. Numer. Method Biomed. Eng.* **33**, e02799. (doi:10.1002/cnm.2799)
- Schiavazzi D, Arbia G, Baker C, Hlavacek A, Hsia T, Marsden A, Vignon-Clementel I. 2016 Uncertainty quantification in virtual surgery hemodynamics predictions for single ventricle palliation. *Int. J. Numer. Methods Biomed. Eng.* **32**, e02737. (doi:10.1002/cnm.2737)

16. Melis A, Clayton R, Marzo A. 2017 Bayesian sensitivity analysis of a 1D vascular model with Gaussian process emulators. *Int. J. Numer. Methods Biomed. Eng.* **33**, e2882. (doi:10.1002/cnm.2882)
17. Eck VG, Donders WP, Sturdy J, Feinberg J, Delhaas T, Hellevik LR, Huberts W. 2016 A guide to uncertainty quantification and sensitivity analysis for cardiovascular applications. *Int. J. Numer. Methods Biomed. Eng.* **32**, e02755. (doi:10.1002/cnm.2755)
18. Kennedy MC, O'Hagan A. 2001 Bayesian calibration of computer models. *J. R. Stat. Soc. B (Stat. Methodol.)* **63**, 425–464. (doi:10.1111/1467-9868.00294)
19. Lei CL *et al.* 2020 Considering discrepancy when calibrating a mechanistic electrophysiology model. *Phil. Trans. R. Soc. A* **378**, 20190349. (doi:10.1098/rsta.2019.0349)
20. Whittaker DG, Clerx M, Lei CL, Christini DJ, Mirams GR. 2020 Calibration of ionic and cellular cardiac electrophysiology models. *WIREs Syst. Biol. Med.* **12**, e1482. (doi:10.1002/wsbm.1482)
21. Mirams GR, Pathmanathan P, Gray RA, Challenor P, Clayton RH. 2016 Uncertainty and variability in computational and mathematical models of cardiac physiology. *J. Physiol.* **594**, 6833–6847. (doi:10.1113/JP271671)
22. Konukoglu E *et al.* 2011 Efficient probabilistic model personalization integrating uncertainty on data and parameters: application to eikonal-diffusion models in cardiac electrophysiology. *Prog. Biophys. Mol. Biol.* **107**, 134–46. (doi:10.1016/j.pbiomolbio.2011.07.002)
23. Johnstone RH, Chang ET, Bardenet R, de Boer TP, Gavaghan DJ, Pathmanathan P, Clayton RH, Mirams GR. 2016 Uncertainty and variability in models of the cardiac action potential: can we build trustworthy models? *J. Mol. Cell. Cardiol.* **96**, 49–62. special Issue: Computational Modelling of the Heart. (doi:10.1016/j.yjmcc.2015.11.018)
24. Coveney S, Clayton RH. 2018 Fitting two human atrial cell models to experimental data using Bayesian history matching. *Prog. Biophys. Mol. Biol.* **139**, 43–58. (doi:10.1016/j.pbiomolbio.2018.08.001)
25. Rasmussen C, Williams C. 2005 *Gaussian processes for machine learning (adaptive computation and machine learning)*. Cambridge, MA: The MIT Press.
26. Xu T, Valocchi AJ, Ye M, Liang F. 2017 Quantifying model structural error: efficient Bayesian calibration of a regional groundwater flow model using surrogates and a data-driven error model. *Water Resour. Res.* **53**, 4084–4105. (doi:10.1002/2016WR019831)
27. Tabima D, Roldan-Alzate A, Wang Z, Hacker T, Molther R, Chesler N. 2012 Persistent vascular collagen accumulation alters hemodynamic recovery from chronic hypoxia. *J. Biomech.* **45**, 799–804. (doi:10.1016/j.jbiomech.2011.11.020)
28. Vanderpool R, Kim A, Molthen R, Chesler N. 2011 Effects of acute Rho kinase inhibition on chronic hypoxia-induced changes in proximal and distal pulmonary arterial structure and function. *J. Appl. Physiol.* **110**, 188–198. (doi:10.1152/jappphysiol.00533.2010)
29. Antiga L, Piccinelli M, Botti L, Ene-lordache B, Remuzzi A, Steinman DA. 2008 An image-based modeling framework for patient-specific computational hemodynamics. *Med. Biol. Eng. Comput.* **46**, 1097–1112. (doi:10.1007/s11517-008-0420-1)
30. Triantafyllou C, Polimeni J, Wald L. 2011 Physiologically noise and signal-to-noise ratio in fMRI with multi-channel array coils. *NeuroImage* **55**, 597–606. (doi:10.1016/j.neuroimage.2010.11.084)
31. Fung Y-C. 2013 *Biomechanics: circulation*. Boca Raton, FL: Springer Science & Business Media.
32. Safaei S *et al.* 2016 Roadmap for cardiovascular circulation model. *J. Physiol.* **594**, 6909–6928. (doi:10.1113/JP272660)
33. Townsley MI. 2012 Structure and composition of pulmonary arteries, capillaries, and veins. *Compr. Physiol.* **2**, 675–709.
34. Windberger U, Bartholovitsch A, Plasenzetti R, Korak KJ, Heinze G. 2003 Whole blood viscosity, plasma viscosity and erythrocyte aggregation in nine mammalian species: reference values and comparison of data. *Exp. Physiol.* **88**, 431–440. (doi:10.1113/eph8802496)
35. Gelman A, Carlin J, Stern H, Dunson D, Vehtari A, Rubin D 2013 *Bayesian data analysis*, 3rd edn. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis.
36. Lax P, Wendroff B. 1960 Systems of conservation laws. *Commun. Pure Appl. Math.* **13**, 217–237. (doi:10.1002/cpa.3160130205)
37. Vanhatalo J, Riihimäki J, Hartikainen J, Jylänki P, Tolvanen V, Vehtari A. Bayesian modeling with Gaussian processes using the GPstuff Toolbox. (<http://arxiv.org/abs/1206.5754>).
38. Laine M. 2007 MCMC Toolbox for Matlab. <http://helios.fmi.fi/ainema/dram/>.
39. Geweke J. 1992 Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In *Bayesian statistics* (eds JM Bernardo, JO Berger, AP Dawid, AFM Smith), vol. 4, pp. 169–193. Oxford, UK: Clarendon Press.
40. Brooks S, Gelman A. 1998 General methods for monitoring convergence of iterative simulations. *J. Comput. Graph. Stat.* **7**, 434–455.
41. Chib S, Jeliazkov I. 2001 Marginal likelihood from the Metropolis–Hastings output. *J. Am. Stat. Assoc.* **96**, 270–281. (doi:10.1198/016214501750332848)
42. Bowman AW, Azzalini A. 1997 *Applied smoothing techniques for data analysis*. New York, NY: Oxford University Press Inc.
43. Silverman B. 1986 *Density estimation for statistics and data analysis*. London, UK: Chapman & Hall/CRC, Chapman and Hall.
44. Valdez-Jasso D. 2010 Modeling and identification of vascular biomechanical properties in large arteries. Ph.D. thesis, North Carolina State University, Raleigh, NC.
45. Wang Z, Lakes RS, Golob M, Eickhoff JC, Chesler NC. 2013 Changes in large pulmonary arterial viscoelasticity in chronic pulmonary hypertension. *PLoS ONE* **8**, e78569. (doi:10.1371/journal.pone.0078569)
46. Tuder R. 2017 Pulmonary vascular remodeling in pulmonary hypertension. *Cell Tissue Res.* **367**, 643–649. (doi:10.1007/s00441-016-2539-y)
47. Kheifets VO, O'Dell W, Smith T, Reilly JJ, Finol EA. 2013 Considerations for numerical modeling of the pulmonary circulation—a review with a focus on pulmonary hypertension. *J. Biomech. Eng.* **135**, 061011. (doi:10.1115/1.4024141)
48. Yang W, Dong M, Rabinovitch M, Chan FP, Marsden AL, Feinstein JA. 2019 Evolution of hemodynamic forces in the pulmonary tree with progressively worsening pulmonary arterial hypertension in pediatric patients. *Biomech. Model. Mechanobiol.* **18**, 779–796. (doi:10.1007/s10237-018-01114-0)
49. Lungu A, Wild J, Capener D, Kiely D, Swift A, Hose D. 2014 MRI model-based non-invasive differential diagnosis in pulmonary hypertension. *J. Biomech.* **47**, 2941–2947. (doi:10.1016/j.jbiomech.2014.07.024)
50. Saccomani M, Thomasset K. 2018 The union between structural and practical identifiability makes strength in reducing oncological model complexity: a case study. *Complexity* **2018**, 2380650. (doi:10.1155/2018/2380650)
51. Ryser MD, Gulati R, Eisenberg MC, Shen Y, Hwang ES, Etzioni RB. 2019 Identification of the fraction of indolent tumors and associated overdiagnosis in breast cancer screening trials. *Am. J. Epidemiol.* **188**, 197–205. (doi:10.1093/aje/kwy214)
52. Snelson E, Ghahramani Z. 2006 Sparse Gaussian processes using pseudo-inputs. In *Advances in neural information processing systems 18* (eds Y Weiss, B Schölkopf, JC Platt), pp. 1257–1264. Cambridge, MA: MIT Press.
53. Friel N, Pettitt AN. 2008 Marginal likelihood estimation via power posteriors. *J. R. Stat. Soc. B (Stat. Methodol.)* **70**, 589–607. (doi:10.1111/j.1467-9868.2007.00650.x)
54. Brynjarsdóttir J, O'Hagan A. 2014 Learning about physical parameters: the importance of model discrepancy. *Inverse Prob.* **30**, 114007. (doi:10.1088/0266-5611/30/11/114007)
55. Mynard JP, Davidson MR, Penny DJ, Smolich JJ. 2012 A simple, versatile valve model for use in lumped parameter and one-dimensional cardiovascular models. *Int. J. Numer. Methods Biomed. Eng.* **28**, 626–641. (doi:10.1002/cnm.1466)