Parameter Estimation for Closed-Loop Lumped Parameter Models of the Systemic Circulation Using Synthetic Data

Nikolai L. Bjørdalsbakke[a], Jacob T. Sturdy[a], David R. Hose[b], Leif R. Hellevik[a]

[a] *Department of Structural Engineering, Norwegian University of Science and Technology (NTNU), Richard Birkelandsvei 1a, Trondheim, 7491, Norway*

[b] *Department of Infection, Immunity and Cardiovascular Disease, University of Sheffield, Beech Hill Road, Sheffield, S10 2RX, United Kingdom*

Abstract

Physics-based models can be applied to describe mechanisms in both health and disease, which has the potential to accelerate the development of personalized medicine. The aim of this study was to investigate the feasibility of personalizing a model of systemic hemodynamics by estimating model parameters.

We investigated the feasibility of estimating model parameters for a closed-loop lumped parameter model of the left heart and systemic circulation using the step-wise subset reduction method. This proceeded by first investigating the structural identifiability of the model parameters. Secondly we performed sensitivity analysis to determine which parameters were most influential on the most relevant model outputs. Finally, we constructed a sequence of progressively smaller subsets including parameters based on their ranking by model output influence. The model was then optimized to data for each set of parameters to evaluate how well the parameters could be estimated for each subset. The subsequent results allowed assessment of how different data sets, and noise affected the parameter estimates.

In the noiseless case, all parameters could be calibrated to less than 10^{-3}% error using time series data, while errors using clinical index data could reach over 100%. With 5% normally distributed noise the accuracy was limited to be within 10% error for the five most sensitive parameters, while the four least sensitive parameters were unreliably estimated for waveform data. The three least sensitive parameters were particularly challenging to estimate so these should be prioritised for measurement. Cost functions based on time series such as pressure waveforms, were found to give better parameter estimates than cost functions based on standard indices used in clinical assessment of the cardiovascular system, for example stroke volume (SV) and pulse pressure (PP). Averaged parameter estimate errors were reduced by several orders of magnitude by choosing waveforms for noiseless synthetic data. Also when measurement data were noisy, the parameter estimation procedure based on continuous waveforms was more accurate than that based on clinical indices. By application of the step-wise subset reduction method we demonstrated that by the addition of venous pressure to the cost function, or conversely fixing the systemic venous compliance parameter at an accurate value improved all parameter estimates, especially the diastolic filling parameters which have least influence on the aortic pressure.

1. Introduction

Mechanistic models have large potential within biomedical applications to possibly explain physiology and pathophysiology. Models which aim to describe physical systems are typically dependent upon model parameters to give an accurate description of reality. Parameters are primarily calibrated by using numerical optimization schemes and measured data. In biomedical applications it is desirable to predict a personalized response to a problem specific treatment or stimuli. Then using data measured for the present and previous states of an individual, predictions can be made about the future state of the personalized model, which has numerous applications in medicine. In the current literature there are many examples of cardiovascular models predicting different system states, medical intervention outcomes or disease [1, 2, 3, 4, 5]. Finding a model usable with minimal data from a typical clinical visit and low computational requirements is highly desirable to enable use of such models in the general health context. To investigate how well model parameters can be personalized in this context we approach the problem by using simple models.

In this manuscript we apply our analysis to a parsimonious model of the systemic circulation and the left ventricle to describe the cardiovascular state of an individual. The investigations are based on data types that should be available non-invasively and available outside extraordinary situations such as visits to the intensive care unit

(ICU). Since changes in individuals correspond to changes in model parameters it is important to find reliable methods to accurately and reliably estimate these parameters, such that the models can be used to monitor changes in response to for example therapy.

A large body of work has been done on model personalization within medical applications, but much of the research has focused on complex models that draw from a large set of measurements to personalize the model, and animal studies are often conducted to collect intensive and invasive measurement sets [6, 7]. Invasive data from human subjects has also been used as seen in work by Colunga et al. and Pant et al. [8, 9, 10]. For the individual providing such data it is less additional burden and risk if such measurements are non-invasive and limited to as few locations on the anatomy as possible. Model personalization of the cardiovascular system has been realised [8, 9, 10, 11, 12]. These studies often focus on the details of novel optimization approaches or the quality of outputs produced by the personalized models, while this manuscript focuses on the accuracy of the estimated parameters themselves. However, it remains a question just how much the available data set can be limited to successfully identify the necessary model parameters. The model must also be personalizable by the available data found within the context for model application, otherwise clinical application will likely be infeasible. The minimalistic model presented in this manuscript may provide a simple framework for monitoring changes in the systemic arterial and left heart hemodynamics.

Both model complexity and data modalities differ, while some studies use electronic data records from a large number of individuals, some efforts instead focus on intensive data from just a few individuals. Pant et al. [9, 10], fit the parameters of a lumped parameter model using an Unscented Kalman Filter (UKF) approach. The technique was demonstrated to be able to reliably estimate parameters for a three element Windkessel (WK) using a single cycle noisy synthetic waveform data. In a clinical environment Pant et al. used MRI measurements, waveform data from catheters, and doppler ultrasound to estimate personalized parameters in two individuals. Meiburg et al. apply a similar approach using a UKF to estimate parameters in a different model of the systemic circulation using synthetic waveform data which yields promising results [11]. Hann et al. have developed a method for personalization where parameters are continuously re-estimated based on the forward model solution, which has been applied to both synthetic and real waveform data in combination with clinical indices measured in the ICU [12]. Colunga et al., Marquis et al. and de Bournonville et al. take an approach to personalizing closed-loop lumped parameter models using invasive data sets to compute good nominal values before model optimization [6, 7, 8]. This approach usually relies on measurements from multiple vessel compartments or population data, and thorough parameter identifiability analysis. Patient specific modelling would likely benefit from more accurate measurement techniques, which would improve estimated parameters and possibly improve predictions based on these estimates. However, not only improved accuracy but also novel measurement techniques that would make data acquisition easier and less burdensome for both the individual and society would make patient specific modeling feasible where it would earlier be stopped by lack of available data.

The question of how well and which personalized model parameters can be identified is also model specific. Hann et al. take a structured approach to reducing the available data sets, while reducing the model complexity in the process [12]. We take a different approach where the problem complexity is reduced by sequentially fixing parameters, but mainly examine the question of how well estimated parameters correspond to parameters that are known to describe a given set of data, which has not been investigated in many contexts. We refer to these known parameters as the "true" parameters, throughout this study. A method for assessing the accuracy of the estimated parameters with respect to the true parameters is introduced in this manuscript. Further, we investigate how different available data sets and cost functions affect the accuracy and precision of parameter estimates, using standard optimization methods. Guided by sensitivity analysis, we assess the impact of reducing the estimated parameter subset by a method which is applicable to any deterministic computer model with continuous output ranges and constant parameters. Effectively, a practical identifiability analysis of the model is performed under different scenarios. This investigation is motivated by the hypothesis that determination of accurate personalized model parameters is essential for predicting future states of the system, and for using models as potential diagnosis support systems.

In order to be able to evaluate the accuracy of parameter estimates we focus on synthetic data generated from the model, as the *true* value of parameters is in general impossible to know for real data situations. For this investigation we use synthetic data with and without noise to investigate the best possible cases for estimation of parameters while still approximating real data. We emphasize that we wish to assess how different choices of parameter subsets affect estimation of parameter values rather than to identify the model configuration which best emulates the data. Attempting to estimate the parameters for real data belongs to future work, but will shed more light on how model discrepancy influences the resulting parameter estimates.

2. Method

2.1. Model

The motivation of this work was to determine if hemodynamic measurements could be applied to estimate personal mechanical parameters characterizing an individual's cardiovascular system. In general the parameters are defined in terms of a mechanistic model of the cardiovascular system. While various anatomically and physiologically detailed models have been developed to describe the human cardiovascular system [13, 14, 15], in this study we implemented a closed-loop lumped parameter model of the systemic circulation and left ventricle as a realistic candidate for routine clinical application. In contrast to some more detailed models the model applied in this work is suitable to real-time simulation. Perhaps more significantly, reducing the number of parameters required to specify the model is expected to reduce the measurement burden required to personalize the model. More detailed models often require direct imposition of population average or nominal values for parameters which may limit their potential for personalization.

The lumped parameter model in this work is similar to that previously applied by Segers et al. [16], however, instead of assuming a constant filling pressure for the left ventricle, we included a venous compartment to represent the left ventricular filling pressure (See Figure 1 for a graphical depiction of the model). This is a similar but simplified version of the approaches of Smith et al. [17] and Hann et al. [12]. Mathematically, the model consists of a system of differential equations that describe the time evolution of the state variables, which are the stressed volumes of the ventricle, arteries and veins. The pressures and flows are determined algebraically from the volumes.

$$\begin{aligned} \frac{dV_{sa}}{dt} &= C_{sa}\frac{dP_{sa}}{dt} = Q_{lvao} - Q_{sys} \\ \frac{dV_{sv}}{dt} &= C_{sv}\frac{dP_{sv}}{dt} = Q_{sys} - Q_{svlv} \\ \frac{dV_{lv}}{dt} &= Q_{svlv} - Q_{lvao} \end{aligned} \tag{1}$$

V_{sa}, V_{sv} and V_{lv} are the stressed blood volumes of the systemic arteries, systemic veins and left ventricle respectively. C_{sa} and C_{sv} are the volume compliance values of the systemic arteries and veins, while P_{sa} and P_{sv} are the corresponding pressures of these compartments. Q_{lvao} denotes the volume blood flow from the left ventricle to the systemic arteries, Q_{sys} is the flow between the systemic arteries and veins, and finally Q_{svlv} is the flow from the veins to the left ventricle. The left ventricular pressure is assumed to be a linear function of the volume of the left ventricle, $P_{lv} = E_{lv}(t)V_{lv}$, where $E_{lv}(t)$ is the elastance of the ventricle at time t. The elastance is modelled as a periodic function, which mimics the periodic contraction of the ventricle and the associated pressure gradient and ejection of blood into the arteries. The pressure in the arteries and veins are also modeled as linear functions of the corresponding volumes, $CP = V$, where C is the compliance of the respective compartment and describes aggregated stiffness of the arterial or venous walls. All equations describing the model are given in Appendix A.

The compartments and their connections are characterized by a set of mechanical parameters listed in Table 1 along with the symbols and selected reference values used in this study. Segers et al. report reference values for most of these parameters in both normotensive and hypertensive populations [18]. In this work we use the reference values for the normotensive population and manually adjusted the parameters not reported by Segers et al. The parameters originate from the mathematical description of each component. The lumped parameter approach represents the cardiovascular system as a set of compartments that contain a volume of blood at a particular pressure, and connections between these compartments which model the flow of blood between these compartments.

The flows between compartments are modeled as a linear function of the pressure difference between the compartments, $Q = \Delta P/R$, where R is a resistance and determines the mechanical energy required to sustain flow between compartments. The flows to and from the ventricle are slightly more complicated as the heart valves ensure these flows are always in the direction obeying the cardiac cycle. The valves are modeled as diodes such that the flow is linearly related to pressure for negative pressure gradients, and 0 otherwise. Note that flow from the the venous compartment, representing the systemic veins, directly enters the left ventricle across the mitral valve, which conceptually assumes the systemic venous pressure is identical to the pulmonary venous pressure. We therefore assume the effects of atrial dynamics are negligible, as previous works also have [12, 17], and that ventricular interaction effects are negligible.

The model presented in Figure 1 produces very similar arterial hemodynamics as the referenced model by Smith et al., the latter is a four element WK (4WK) as the former is a three element WK (3WK). Studies show that this version of the 4WK can be better optimized to data than the 3WK, but introduces another personalizable parameter and often yields similar parameter values for aortic characteristic impedance and resistance when fitted to real data [19]. Vachiéry et al. highlight that pulmonary arterial hypertension is challenging to diagnose and that even this condition may not be reflected in the left heart or systemic circulation [20]. In exercise, the pulmonary circulation and right heart are recognized to be influential also on arterial hemodynamics through limitations on cardiac output [21, 22]. The absence of the right heart and pulmonary circuit may then cause limited expressions of exercise hemodynamics and limit the applicability of

the model, but confirmation of this statement requires further investigation beyond the scope of this manuscript. By redirecting the venous return directly into the left ventricle, the cardiac preload is modeled by the venous pressure and flow rather than fixing the preload and venous properties as in an open-loop model. By circumventing the pulmonary circulation and right heart we also omit many potentially personalizable parameters, which may further complicate the process of personalizing the model with limited data. If the model is sufficiently simple so that it can be run and personalized in real time it may be used to monitor changes in resting hemodynamics. As discussed by Huberts et al., model accuracy lessens as the model is simplified [23], so an important question to be answered in a different setting is whether a simplified model is able resolve small changes in hemodynamics caused by a chosen form of stimulus.

For the remainder of this article, it is convenient to treat the model variables, e.g. P_{ao}, Q_{ao} or P_{sv}, as functions of the parameters $\boldsymbol{\theta}$ and time, $y(t, \boldsymbol{\theta})$, where y denotes a particular variable. Thus the values at a particular time point are denoted $y(t_k, \boldsymbol{\theta})$. $\boldsymbol{\theta}$ denotes the vector of all model parameters. These functions are approximated numerically using SciPy's implementation of the 4th-order Runge-Kutta (RK4) method to integrate the differential equations [24]. The resulting system of equations may be stiff, and therefore we also tested solving the system with the backward differentiation formula (BDF), but it was found to be much more computationally expensive than RK4 while results were found to be comparable.

2.2. Data and measurements

A focus of this study was to determine how accurately and precisely specific model parameters could be estimated from various data sources. In order to quantify the performance of the estimation procedure, the true values of the parameters, $\boldsymbol{\theta}_{true}$, must be known. As this is not possible for real clinical data, we simulated realistic measurement data based on the numerical solution of the model for particular parameter values such that the true values of the parameters are known. The data generated in this way stimulates continuous waveform data for pressure and flow in the aorta and large systemic veins of the cardiovascular system (Using the notation proposed previously, $P_{ao}(t_k, \boldsymbol{\theta}_{true})$, $Q_{ao}(t_k, \boldsymbol{\theta}_{true})$, and $P_{sv}(t_k, \boldsymbol{\theta}_{true})$. In addition, measurement of common clinical indices was simulated by computing these from the waveform data, e.g. $P_{sys} = \max_k P_{ao}(t_k, \boldsymbol{\theta}_{true})$. The time varying measurements and clinical indices are depicted in Figure 2 and described in the following section. The model equations (Appendix A) were solved by the 4th order Runge-Kutta scheme. Model outputs were solved until they reached a steady periodic state, which was found to be reached by 10 heart cycles. We also tested the solution after more heart cycles but model outputs did not change and parameter estimates remained identical. In the following, we use the notation y_k^m to denote a measured value of the quantity y at time t_k if appropriate (y^m denotes a time independent measurement).

Symbol	Description	True values	Unit
C_{ao}	Systemic arterial compliance	1.13	$\frac{\text{mL}}{\text{mmHg}}$
C_{sv}	Systemic venous compliance	11.0	$\frac{\text{mL}}{\text{mmHg}}$
E_{max}	Maximal left ventricular elastance	1.5	$\frac{\text{mmHg}}{\text{mL}}$
E_{min}	Minimal left ventricular elastance	0.03	$\frac{\text{mmHg}}{\text{mL}}$
R_{mv}	Mitral valve resistance	0.006	$\frac{\text{mmHg s}}{\text{mL}}$
R_{sys}	Total systemic vascular resistance	1.11	$\frac{\text{mmHg s}}{\text{mL}}$
T	Heart period	0.85	s
t_{peak}	Time of peak ventricular elastance	0.3	s
V_{tot}	Total stressed blood volume	300	mL
Z_{ao}	Characteristic impedance of the aorta	0.033	$\frac{\text{mmHg s}}{\text{mL}}$

Table 1: The model parameters are listed with their corresponding symbols and reference values. For most parameters the reference values were reported by Segers et al. [18]. The remaining parameters have been manually tuned. We use these reference values as experimental estimates, and in our wording as the "true" parameters, to generate the data used for the parameter estimation procedure.

We expressed the time series simply as a list of continuous measurement points of pressure or flow. A series of 100 points per heart cycle were defined, corresponding in this example to a measurement frequency of approximately 117.6 Hz with the chosen heart rate. Pressure catheters or tonometry can be used to measure blood pressure waveforms [25], and doppler ultrasound is routinely applied to record flow velocity waveforms in the heart and aorta.

In clinical practice blood pressure is assessed by sphygmanometry of brachial systolic and diastolic blood pressure. These may be emulated as the maximum and minimum values of the simulated aortic pressure waveform, respectively. The model does not explicitly describe the brachial pressure, thus the simulated data may be thought of the best case where central aortic pressure is measured. Aortic pressure is typically comparable to brachial pressure (especially in healthy, young to middle aged individuals). The difference between systolic and diastolic pressures is the pulse pressure, which is also a common clinical measure often used to compute other hemodynamic quantities [26].

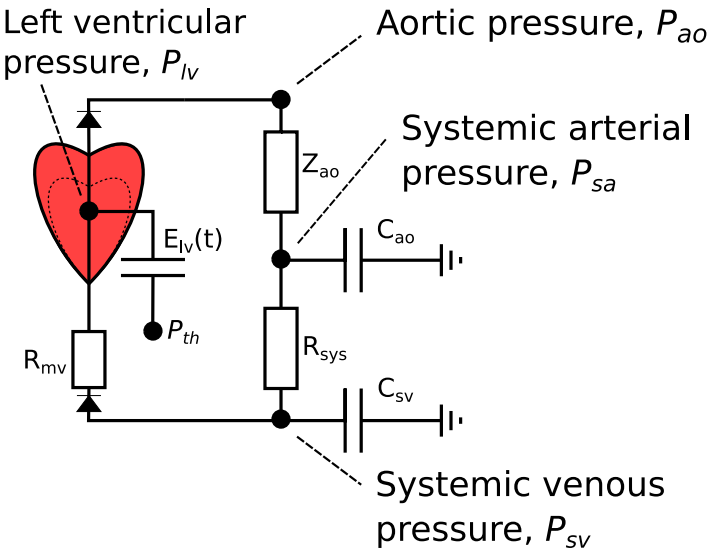

Figure 1: The closed-loop, lumped parameter model of the systemic circulation and the left ventricle. The circuit equivalent formulation of the model is depicted, with the pressures of each compartment, as well as most of the mechanical parameters. The model describes three compartments: the left ventricular, arterial and venous compartments. P_{th} is the intrathoracic pressure, P_{lv} is the left ventricular pressure and $E_{lv}(t)$ indicates the left ventricular elastance function. For an explanation of the remaining parameters, see 1.

The stroke volume and maximal aortic flow are two additional measures commonly used in clinical practice. These may both be derived from ultrasound measurements of the flow velocity and diameter of the left ventricular outflow tract [27]. The peak value corresponds to maximal aortic flow, while the integral of the volumetric flow over a single cardiac cycle gives the volume of blood ejected by the ventricle, the stroke volume. Since the three volume compartments of the model (Section 2.1) do not represent any anatomical features these measurements are emulated from the flow across the aortic valve.

The model generated data may be represented as

$$y_k^m = y(t_k, \boldsymbol{\theta}_{true}) \tag{2}$$

where the superscript m emphasizes this is the measured value and subscript k denotes the time point, where relevant. For the case where y_k^m is a scalar extracted from a time series output the k index is without meaning.

Realistic clinical measurements will have some noise. In this analysis, we investigate both the ideal case of perfect measurements and the more realistic case of noisy measurements. Noisy measurements are simulated by adding randomly sampled perturbations to the model outputs

$$y_k^{noisy} = y(t_k, \boldsymbol{\theta}_{true}) \times \xi_k \tag{3}$$

where ξ_k are independently and identically distributed normal random variables with mean 1 and standard deviation 0.05. In reality, the noise may be biased at least in parts of the measured signal, but this may depend on the measurement technique or equipment. To keep the case most general and for simplicity we use normal distributed noise even though more sophisticated noise distributions can be constructed in theory.

2.3. Structural Identifiability Analysis

Before investigating whether model parameters can be estimated in practice, a structural identifiability analysis will reveal if there is at all any possibility for estimating parameters from more realistic data. Structural identifiability only considers the model structure and data under perfect conditions.

Performing a global structural identifiability analysis means proving that a model formulation gives a unique model output for any given parameter vector $\boldsymbol{\theta}$. Alternatively, a model is globally structurally identifiable if for two distinct arbitrarily chosen vectors of parameters $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ from a parameter space Θ then $y(t_k, \boldsymbol{\theta}_1) = y(t_k, \boldsymbol{\theta}_2)$ holds if and only if $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$ for all k [28, 29]. A model is locally structurally identifiable if the condition $y(t_k, \boldsymbol{\theta}) \neq y(t_k, \boldsymbol{\theta}')$ is true for any $\boldsymbol{\theta}$ in an open interval around $\boldsymbol{\theta}'$ in the parameter space Θ. Therefore, for strictly local structural identifiability multiple but discretely different vectors of $\boldsymbol{\theta}$ can generate the same model outputs which causes additional concerns for practical parameter estimation, as there may be multiple parameter sets which recreates the output. However, it is not a given that the system has more than one solution within the domain for practical implementation. Practical identifiability analysis needs to be performed to assess this. Given that a model is locally identifiable, it can still be globally identifiable unless it is explicitly proven not to be.

Villaverde et al. have developed the STRIKE-GOLDD software package which determines the local structural identifiability of non-linear differential equation models by a method involving Lie derivatives [30]. We applied the MatLab implementation of STRIKE-GOLDD to our model with different data combinations to investigate identifiability. However, this software requires the right hand side of the system of ODEs to be infinitely differentiable functions with respect to both the state variables and parameters, and the valve models are not differentiable in this context. Therefore, we analyze the model in its ejection phase enforcing a closed mitral valve and open aortic valve using left ventricular pressure and aortic flow as the aortic pressure is equal to left ventricular pressure in the systolic ejection phase according to the model. For the diastolic phase systemic arterial pressure is given while the valves are open and closed oppositely to the ejection phase. The dynamic driver function for the model which is the left ventricular elastance function is dependent upon a time signal which is periodic with the length of the heart cycle. The time signal controlling the elastance cycle is the time

variable but defined as $u(t) = \tau = t \mod T$ and is treated as a known dynamic system input in this context. The full specification of the elastance function is given in Appendix A.

Pironet et al. perform a structural identifiability analysis by demonstrating that a unique solution for the parameter set can be found by using the information from perfectly observed model outputs [31]. The model examined by Pironet has a very similar formulation to our model, but includes more chambers, yet with a certain set of outputs the model was found to be globally structurally identifiable. Information from both the systolic and diastolic phases are used separately to identify parameters, as we have done using STRIKE-GOLDD. Parameters which are shown to be identifiable constant parameters are used to prove the identifiability of other parameters. Parameters identified during the model in the diastolic phase is used in identifying parameters in the ejection phase, and hence the reverse should also be valid. Some of the computations relevant to the identifiability analysis which required high amounts of memory were performed on the NTNU IDUN computing cluster [32].

2.4. Sensitivity analysis

Sensitivity analysis of the model quantities of interest can identify which parameters most influence the model's prediction of a quantity of interest. This method can subsequently be applied to identify which parameters may be estimated from given types of measurements, or conversely may be applied to identify which measurements are necessary to provide information about particular parameters [23, 33, 34]. We analyzed the sensitivity of the model outputs in terms of Sobol indices which are commonly used as a global measure of sensitivity [35]. Sobol indices quantify sensitivity as the proportion of variance of the model output attributable to variance of particular parameter values:

$$S_{M,i}(y) = \frac{\mathrm{Var}(\mathrm{E}(y|\theta_i))}{\mathrm{Var}(y)} \tag{4a}$$

$$S_{T,i}(y) = \frac{\mathrm{E}(\mathrm{Var}(y|\theta_{\sim i}))}{\mathrm{Var}(y)}. \tag{4b}$$

where $S_{M,i}$ and $S_{T,i}$ denote respectively the main and total sensitivity to parameter θ_i, while y denotes the function sensitive to the parameters $\boldsymbol{\theta}$, or for this analysis a model output.

To interpret these indices, we note that a quantity of interest with high values of either index for given parameter suggests that measurement of that quantity may provide substantial information about that parameter. However, if $S_{M,i}$ is low but $S_{T,i}$ is large then parameter θ_i impacts the quantity of interest primarily through interactions with other parameters and consequently may substantially affect the quantity of interest. Unfortunately due to its interactivity it may remain challenging to estimate due to its dependence on the values of other parameters.

In general, parameter estimation becomes more challenging as more parameters are estimated. In addition, strong interactions between parameters may impede efficient numerical optimization. Prior works have employed a number of subset selection methods to reduce the number of parameters varied while fitting the model to data [36, 37]. Most such methods are based on analysis of sensitivities and in general select subsets of parameters with high sensitivity to estimate, as the data will provide the most information about these. The complementary set of parameters with lower sensitivity is then fixed at nominal values, which is expected to have minimal impact on the fitting of the model to the data.

Sobol indices have a clear interpretation in terms of the behavior of individual variables, but the interpretation of the Sobol indices for many variables or for a time varying signal may not be as easily interpreted. For example simply averaging across variables or time will weight regions of low variance equally to those of high variance. To remedy this one may instead examine the variance weighted averages

$$TAS_{M,i}(y) = \frac{\sum_k S_{M,i}(y(t_k))\,\mathrm{Var}(y(t_k))}{\sum_k \mathrm{Var}(y(t_k))} \tag{5a}$$

$$TAS_{T,i}(y) = \frac{\sum_k S_{T,i}(y(t_k))\,\mathrm{Var}(y(t_k))}{\sum_k \mathrm{Var}(y(t_k))}. \tag{5b}$$

Here we abbreviate time-averaged as (TA). We computed the indices defined in (4) assuming all input parameters were independently and uniformly distributed over the range of 90% to 110% of their nominal values. The indices were estimated using a Monte-Carlo approach as proposed by Saltelli et al. with 2500 samples per parameter [35]. The weighted averages (5) were subsequently computed from these estimates of the Sobol indices [33, 38].

2.5. Synthetic data generation and parameter estimation

For the purposes of fitting a model to data a typical approach is to assume the data, y_k^m, are simply a perturbation of the values predicted by the model for the true parameter values:

$$y_k^m = y(t_k, \boldsymbol{\theta}_{true}) + E_k. \tag{6}$$

Given the data and the model, the objective is then to determine the value of $\boldsymbol{\theta}$ that produces model predictions best matching the data, i.e. with minimal E_k. To quantify how well a given value of parameters matches the data, a

cost function must be defined, for example

$$J(\boldsymbol{\theta}) = \sum_k (y(t_k, \boldsymbol{\theta}) - y_k^m)^2. \tag{7}$$

To estimate the parameters a minimization problem is solved

$$\hat{\boldsymbol{\theta}} = \arg\min J(\boldsymbol{\theta}) \tag{8}$$

which corresponds with ordinary least squares regression. The minimization problem is typically solved using a numerical optimization method, and the Trust Region Reflective algorithm (TRRA) was chosen for this study due to accepting parameter bounds, and is available through SciPy [24]. The constraint applied to the cost functions used in this investigation are given in higher detail in Appendix B.

Four cost functions were considered in this work, corresponding to three sets of clinical measurements. The first corresponds with fitting the model to standard clinical measurements

$$\begin{aligned} J(\boldsymbol{\theta}) = &\left(\frac{P_{sys}(\boldsymbol{\theta}) - P_{sys}^m}{K_{psys}}\right)^2 \\ &+ \left(\frac{P_{dia}(\boldsymbol{\theta}) - P_{dia}^m}{K_{pdia}}\right)^2 \\ &+ \left(\frac{PP(\boldsymbol{\theta}) - PP^m}{K_{pp}}\right)^2 \\ &+ \left(\frac{SV(\boldsymbol{\theta}) - SV^m}{K_{SV}}\right)^2 \\ &+ \left(\frac{Q_{max}(\boldsymbol{\theta}) - Q_{max}^m}{K_{qmax}}\right)^2. \end{aligned} \tag{9}$$

The K_j are scaling factors, where subscript j indicates a type of measurement the parameter normalizes. These were used to ensure that all quantities contributing to the cost function were weighted approximately equally. See Appendix B for the values used in our computations.

The second cost function corresponds to making use of continuous pressure and flow waveform data which is not typically measured in the clinic:

$$\begin{aligned} J(\boldsymbol{\theta}) = &\sum_k^N \left(\frac{P_{ao,k}(\boldsymbol{\theta}) - P_{ao,k}^m}{K_p}\right)^2 \\ &+ \sum_k^N \left(\frac{Q_{ao,k}(\boldsymbol{\theta}) - Q_{ao,k}^m}{K_q}\right)^2 \end{aligned} \tag{10}$$

where k indexes the sample times of the measurements for a total N points per measurement type.

The third cost function requires measurement of the venous systolic and diastolic pressure values:

$$\begin{aligned} J(\boldsymbol{\theta}) = &\left(\frac{P_{sys}(\boldsymbol{\theta}) - P_{sys}^m}{K_{psys}}\right)^2 \\ &+ \left(\frac{P_{dia}(\boldsymbol{\theta}) - P_{dia}^m}{K_{pdia}}\right)^2 \\ &+ \left(\frac{PP(\boldsymbol{\theta}) - PP^m}{K_{pp}}\right)^2 \\ &+ \left(\frac{SV(\boldsymbol{\theta}) - SV^m}{K_{SV}}\right)^2 \\ &+ \left(\frac{Q_{max}(\boldsymbol{\theta}) - Q_{max}^m}{K_{qmax}}\right)^2 \\ &+ \left(\frac{P_{sv,sys}(\boldsymbol{\theta}) - P_{sv,sys}^m}{K_{psvsys}}\right)^2 \\ &+ \left(\frac{P_{sv,dia}(\boldsymbol{\theta}) - P_{sv,dia}^m}{K_{psvdia}}\right)^2. \end{aligned} \tag{11}$$

The fourth cost function requires measurement of the venous pressure wave form in addition to the aortic pressure and flow:

$$\begin{aligned} J(\boldsymbol{\theta}) = &\sum_k^N \left(\frac{P_{ao,k}(\boldsymbol{\theta}) - P_{ao,k}^m}{K_p}\right)^2 \\ &+ \sum_k^N \left(\frac{Q_{ao,k}(\boldsymbol{\theta}) - Q_{ao,k}^m}{K_q}\right)^2 \\ &+ \sum_k^N \left(\frac{P_{sv,k}(\boldsymbol{\theta}) - P_{sv,k}^m}{K_{psv}}\right)^2. \end{aligned} \tag{12}$$

As many optimization methods such as the Quasi-Newton method used in this analysis may be attracted to local minima, we attempted to mitigate this by performing the numerical optimization with several initial parameter values randomly sampled according to the formula

$$\theta_{sampled,i} = \theta_{ref,i}(1 + \delta_i), \tag{13}$$

where δ_i are stochastic values drawn independently from a normal distribution with a zero mean and standard deviation of 0.3. $\boldsymbol{\theta}_{ref}$ was a set of reference parameters used to sample initial guesses arbitrarily chosen within a physiologically realistic combination of parameters. They are not equal to $\boldsymbol{\theta}_{true}$, as to not center the distribution of the sampled parameters at the desired cost function minimum. See Appendix B for the list of reference values.

2.6. Quantities of interest

The main goal of fitting the model to personal data is to accurately estimate $\boldsymbol{\theta}_{true}$, thus the measure of error for this is $\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_{true}$, where $\hat{\boldsymbol{\theta}}$ is the estimated parameter vector. $\boldsymbol{\theta}_{true}$ are the synthetic reference parameters, listed

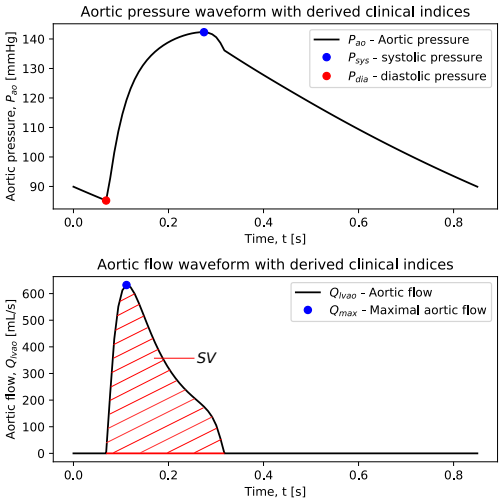

Figure 2: Most of the data types and measurements are depicted, as well as their relations.

in Table 1. We evaluated how well local optimization algorithms could calibrate the lumped parameter model by recovering the "true" parameters when we used no prior knowledge about the parameters aside from what was considered approximately physiologically realistic ranges for the parameters. We generated the data sets used in this study with the "true" parameter values, $\boldsymbol{\theta}_{true}$, set to the values reported by Segers et al. for normotensive individuals [18]. (The remaining parameters of the parameter vector which are not reported by Segers et al. were chosen by manual tuning and their values are reported in Table 1.) Thus, the true parameters, $\boldsymbol{\theta}_{true}$, are mostly chosen to be values used in the literature.

Specifically, to recover these parameters we optimized the model outputs according to the measurements in Section 2.2. Table 2 lists the relevant quantities, and Figure 2 depicts them.

Model output/ measurements	Derived clinical index
Aortic pressure waveform, P_{ao}	- Aortic systolic and diastolic pressures, P_{sys} and P_{dia} - Pulse Pressure, PP
Aortic flow waveform, Q_{ao}	- Maximal aortic flow, $Q_{ao,max}$ - Stroke Volume, SV
Systemic venous pressure waveform, P_{sv}	- Venous systolic and diastolic avgeraged pressures, $P_{sv,sys}$ and $P_{sv,dia}$

Table 2: The model waveform outputs are listed along with any derived quantities that are or serve as approximations to clinical indices. Abbreviations are also specified.

2.7. The stepwise subset reduction method

Based on the total Sobol indices in (4), and (5) we ranked parameters with respect to aortic pressure, and devised an approach to investigate how well adding additional less sensitive parameters to the fitting procedure affects parameter estimation. We named the method the stepwise subset reduction method (SSRM).

The purpose of parameter estimation based on the model presented here is to develop a method for identifying personal characteristics from clinically measured data. In this context, the personal characteristics are the parameter values $\boldsymbol{\theta}_{true}$; however, simultaneous estimation of all parameter values may be challenging, particularly when measurements are noisy. We thus sought to investigate how this approach performed for various subsets of parameters in order to determine if any less influential model parameters should be fixed in order to improve the robustness of this approach. For similar models, calibration of stable parameters with low variability have been demonstrated [6, 12].

For each cost function and data set (to be described later), we performed the numerical optimization procedure outlined in the previous section for each subset of parameters (described in the following section) for $N_s = 50$ sets of $\boldsymbol{\theta}_{sampled}$ as initial guesses for the parameters. For each set of initial guesses, the cost function at the termination of the optimization procedure was recorded. Subsequently the minimum cost function observed in the subset was identified, and only the cases where the cost function was less than 125% of the minimum observed cost function were retained for evaluation of the subset.

The subsets of parameters were defined as follows. At maximum all nine parameters (see Table 1) were estimated at once as described in the previous paragraph. After fitting all nine model parameters we repeated this step eight times, while fixing one additional parameter at its true value according to $\boldsymbol{\theta}_{true}$ in ascending order of sensitivity per repetition. In practice this means that the subset selection was reduced by one additional parameter at each iteration. See Table 3 for the full sequence of subsets in the SSRM. All nine parameters are estimated N_s times, then the eight most sensitive parameters are estimated N_s times, and so on until only the most sensitive parameter is estimated. From here on this method is referred to as the stepwise subset reduction method (SSRM).

The SSRM was then performed for ten different estimation scenarios. First, for standard clinical indices as specified in equation (9) with (III) and without noise (I). Subsequently, for aortic flow and pressure waveforms (10) with (IV) and without noise (II). Another four scenarios (V - VIII) were all based on noisy waveform data (3). Two approaches (V and VI) included venous information as per equations (11) and (12), respectively. Two additional scenarios focused on the impact of fixing either t_{peak} (VII) or

C_{sv} (VIII) at their true values by applying the SSRM for estimating the remaining parameters based on noisy aortic data (11).

The motivation for the investigations with fixed t_{peak} (VII) or C_{sv} (VIII) was the observation that by inclusion or removal of certain parameters in the estimated subset, the estimation error of the remaining parameters changed by an order of magnitude. The addition of more data from the venous compartment was motivated by the observation that many of the least sensitive and hardest to estimate parameters relate to ventricular filling, which in turn is largely determined by venous pressure.

As stated, the SSRM was conducted by fixing the model parameters not in the subset at their true values. To investigate the effect on parameter estimates by fixing the parameters at wrong values, we repeated the SSRM with the cost function in equation (10) but fixed the parameters at 30% above their true value (X), and the effect on parameter estimates was observed.

In the previous estimation cases (I) - (VIII) and (X), the same sampled values of noise, ξ_k, were used in all optimization runs. To assess the general impact of noise, we estimated parameters from N_s cases of normally distributed noise (3) applied to the data in equation (10) (IX). For each sample of ξ_k, we estimated parameters from only one initial parameter guess. The sampling values in Table B.6 from Appendix B were used as initial guesses in this exercise, which were confirmed to estimate all parameters to negligible error without noise.

The performance of the parameter estimation procedure was evaluated for each subset and cost function by calculating the mean absolute percentage error (MAPE) between estimated and true parameter values

$$MAPE_i = \frac{1}{N_s}\sum_{j=1}^{N_s}\left|\frac{\hat{\theta}_{i,j} - \theta_{true,i}}{\theta_{true,i}}\right| \cdot 100\%. \qquad (14)$$

i indicates the i-th parameter of the vector $\boldsymbol{\theta}$, while j indicates estimate number out of N_s estimates. The mean percentage error (MPE) is defined nearly identically, but without the modulus operation. The standard deviation of the MAPE measurements was computed as

$$\begin{aligned} APE_i\ STD &= \mathrm{Std}\left(\left|\frac{\hat{\theta}_i - \theta_{true,i}}{\theta_{true,i}}\right| \cdot 100\%\right) \\ &= \sqrt{\frac{1}{N_s}\sum_{j=1}^{N_s}(APE_{i,j} - MAPE_i)^2} \end{aligned} \qquad (15)$$

The workflow for calculating parameter estimation error for a selected cost function and number of fixed parameters is illustrated in Figure 3. For each such configuration

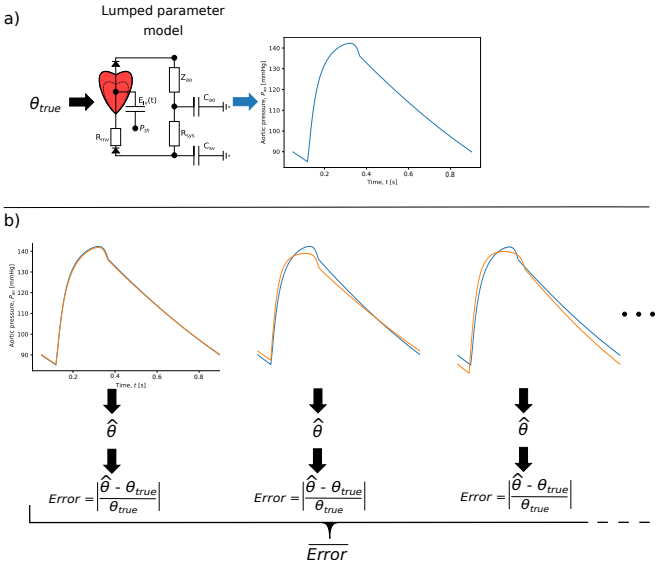

Figure 3: An illustration of SSRM for one chosen cost function is described. (a) A "true" parameter set ($\boldsymbol{\theta}_{true}$) is defined and used to generate corresponding model output, which is illustrated by an aortic pressure curve in this example. In (b) multiple sampled $\boldsymbol{\theta}_{sampled}$ initial parameter values are used to estimate N_s parameter sets $\hat{\boldsymbol{\theta}}$. The pipeline is repeated nine or eight times as the parameter subset selection is reduced by one parameter for each repetition. For this manuscript the entire SSRM is repeated in ten different estimation scenarios.

the non-linear optimization was repeated $N_s = 50$ times with a new initial parameter value $\boldsymbol{\theta}_{sampled}$ from equation (13) for each run. The features of all the applied cost functions are organized in subsection 2.8.

The sequence of subsets in SSRM	Subset parameters
#1	V_{tot}, E_{max}, C_{ao}, R_{sys}, t_{peak}, C_{sv}, E_{min}, Z_{ao}, and R_{mv}
#2	V_{tot}, E_{max}, C_{ao}, R_{sys}, t_{peak}, C_{sv}, E_{min}, and Z_{ao}
#3	V_{tot}, E_{max}, C_{ao}, R_{sys}, t_{peak}, C_{sv}, and E_{min}
#4	V_{tot}, E_{max}, C_{ao}, R_{sys}, t_{peak}, and C_{sv}
#5	V_{tot}, E_{max}, C_{ao}, R_{sys}, and t_{peak}
#6	V_{tot}, E_{max}, C_{ao}, and R_{sys}
#7	V_{tot}, E_{max}, and C_{ao}
#8	V_{tot}, and C_{ao}
#9	V_{tot}

Table 3: A table representing the sequence of model parameter subsets as estimated by the SSRM.

2.8. Summary

Parameter sensitivity was quantified by estimating variance weighted averages and total Sobol indices as per equations (5b) and (4), respectively. Based on the sensitivities to the timeseries output for the aortic pressure, we ranked the model parameters. To assess the impact of adding parameters of varying sensitivity to the estimation procedure we developed the SSRM (stepwise subset reduction method). The SSRM was applied by ranking model parameters by sensitivity, then estimating all model parameters in the first "step", and then reducing the esti-

mated subset by the least sensitive parameter for the second "step" and making new parameter estimates. The subset reduction continues stepwise until only the most sensitive parameter is estimated. The SSRM was repeated in 10 estimation scenarios:

I For a cost function using clinical indices
II For a cost function using waveform data
III For a cost function using clinical indices with noise
IV For a cost function using waveform data with noise
V For a cost function using clinical indices with noise including systolic and diastolic venous pressure
VI For a cost function using waveform data with noise including the venous pressure waveform.
VII For a cost function using waveforms with noise where the t_{peak} is always fixed at its true value.
VIII For a cost function using waveforms with noise where the C_{sv} is always fixed at its true value.
IX For a cost function using waveforms where noise is varied N_s times for each step in the SSRM.
X For a cost function using waveforms where parameters are fixed at 30% above their true value as they are left out of the estimated parameter subset.

After these steps were completed we computed and examined the MAPE and MPE for the estimated parameters compared to the true parameters to assess accuracy, precision and bias.

3. Results

3.1. Structural Identifiability Analysis

Local structural identifiability was analyzed using the STRIKE-GOLDD software. The measured aortic pressure is defined as $P_{ao} = \max P_{lv}, P_{sa}$ in our model, which means that during ejection the aortic pressure is equal to the left ventricular pressure, but is equal to the systemic arterial pressure in diastole. When given aortic flow and left ventricular pressure during ejection all model parameters were found to be locally structurally identifiable except R_{mv} as it was eliminated from the model description by closing the mitral valve. The filling phase model was analyzed using only systemic arterial pressure as aortic flow is zero at this point in the heart cycle. The parameters R_{sys}, C_{ao}, Z_{ao}, and C_{sv} were taken as parameters known to be structurally identifiable upon the second analysis and the rest of the parameters were also found to be locally structurally identifiable in this analysis given waveform data.

3.2. Sensitivity analysis

The sensitivity analysis yielded a parameter ranking based on the sensitivity of the parameters to model outputs. Table 5 shows the ranking according to Sobol indices (4b) for systolic and diastolic pressure and variance weighted sensitivity (5b) to dynamic output signals. Table 4 displays the parameter sensitivities, which for dynamic signals are the total variance weighted average from equation (5b) over a heart cycle and for the clinical indices are the total Sobol indices from equation (4b).

Model output	$TAS_{T,i}(P_{ao})$	$TAS_{T,i}(V_{lv})$	$TAS_{T,i}(Q_{lvao})$	P_{sys}	P_{dia}
V_{tot}	$6.8 \cdot 10^{-1}$	$6.3 \cdot 10^{-1}$	$1.2 \cdot 10^{-1}$	$7.1 \cdot 10^{-1}$	$6.6 \cdot 10^{-1}$
E_{max}	$1.1 \cdot 10^{-1}$	$6.4 \cdot 10^{-2}$	$1.0 \cdot 10^{-1}$	$9.9 \cdot 10^{-2}$	$1.1 \cdot 10^{-1}$
C_{ao}	$8.3 \cdot 10^{-2}$	$1.0 \cdot 10^{-1}$	$4.7 \cdot 10^{-2}$	$1.5 \cdot 10^{-1}$	$3.7 \cdot 10^{-3}$
R_{sys}	$6.5 \cdot 10^{-2}$	$1.7 \cdot 10^{-2}$	$1.8 \cdot 10^{-1}$	$1.5 \cdot 10^{-2}$	$2.1 \cdot 10^{-1}$
t_{peak}	$3.6 \cdot 10^{-2}$	$1.6 \cdot 10^{-1}$	$5.7 \cdot 10^{-1}$	$8.3 \cdot 10^{-3}$	$3.8 \cdot 10^{-3}$
C_{sv}	$2.4 \cdot 10^{-2}$	$2.1 \cdot 10^{-2}$	$4.0 \cdot 10^{-3}$	$2.4 \cdot 10^{-2}$	$2.3 \cdot 10^{-2}$
E_{min}	$1.3 \cdot 10^{-2}$	$1.6 \cdot 10^{-2}$	$4.3 \cdot 10^{-3}$	$1.6 \cdot 10^{-2}$	$8.1 \cdot 10^{-3}$
Z_{ao}	$3.0 \cdot 10^{-4}$	$4.0 \cdot 10^{-5}$	$1.0 \cdot 10^{-2}$	$2.0 \cdot 10^{-4}$	$3.9 \cdot 10^{-5}$
R_{mv}	$2.0 \cdot 10^{-4}$	$1.7 \cdot 10^{-3}$	$2.0 \cdot 10^{-4}$	$3.0 \cdot 10^{-4}$	$3.8 \cdot 10^{-5}$

Table 4: Sensitivity values for the parameters to different model outputs. The values given for the time averaged outputs are the total variance weighted averages, see (5), and the derived clinical indices are given as the total Sobol indices (4).

Model output	$TAS_{T,i}(P_{ao})$	$TAS_{T,i}(V_{lv})$	$TAS_{T,i}(Q_{lvao})$	P_{sys}	P_{dia}
1st	V_{tot}	V_{tot}	t_{peak}	V_{tot}	V_{tot}
2nd	E_{max}	t_{peak}	R_{sys}	C_{ao}	R_{sys}
3rd	C_{ao}	C_{ao}	V_{tot}	E_{max}	E_{max}
4th	R_{sys}	E_{max}	E_{max}	C_{sv}	C_{sv}
5th	t_{peak}	C_{sv}	C_{ao}	E_{min}	E_{min}
6th	C_{sv}	R_{sys}	Z_{ao}	R_{sys}	t_{peak}
7th	E_{min}	E_{min}	E_{min}	t_{peak}	C_{ao}
8th	Z_{ao}	R_{mv}	C_{sv}	R_{mv}	Z_{ao}
9th	R_{mv}	Z_{ao}	R_{mv}	Z_{ao}	R_{mv}

Table 5: Based on the sensitivity values calculated and presented in Table 4, the following rankings of how sensitive the chosen outputs are to each respective parameter have been compiled. The parameters at the 1st position are the most influential parameters to the model output they are listed under.

3.3. Parameter estimation

Computation of the MAPE for all $N_s = 50$ model evaluations yielded substantial errors in some estimated parameters for most cost functions. The exception was the noiseless case with the cost function in equation (10) (II) where all parameters were estimated to an error of order of magnitude 10^{-3} percent or less, see Figure 4. The bar graphs presented from figures 4 - 12 present the parameter MAPE or MPE with standard deviations for all parameter subsets in the SSRM, cases I - X. The bars corresponding to n parameters are the errors of the estimates for the n most sensitive parameters while the rest of the less sensitive parameters are fixed at their true values. Consequently, the figures allow the reader to assess how increasing or decreasing the number of estimated parameters affect the parameter estimation error.

For the cost functions using clinical indices (I), all errors were considerably larger when compared to the waveform based error estimates (II). For the subsequent plots

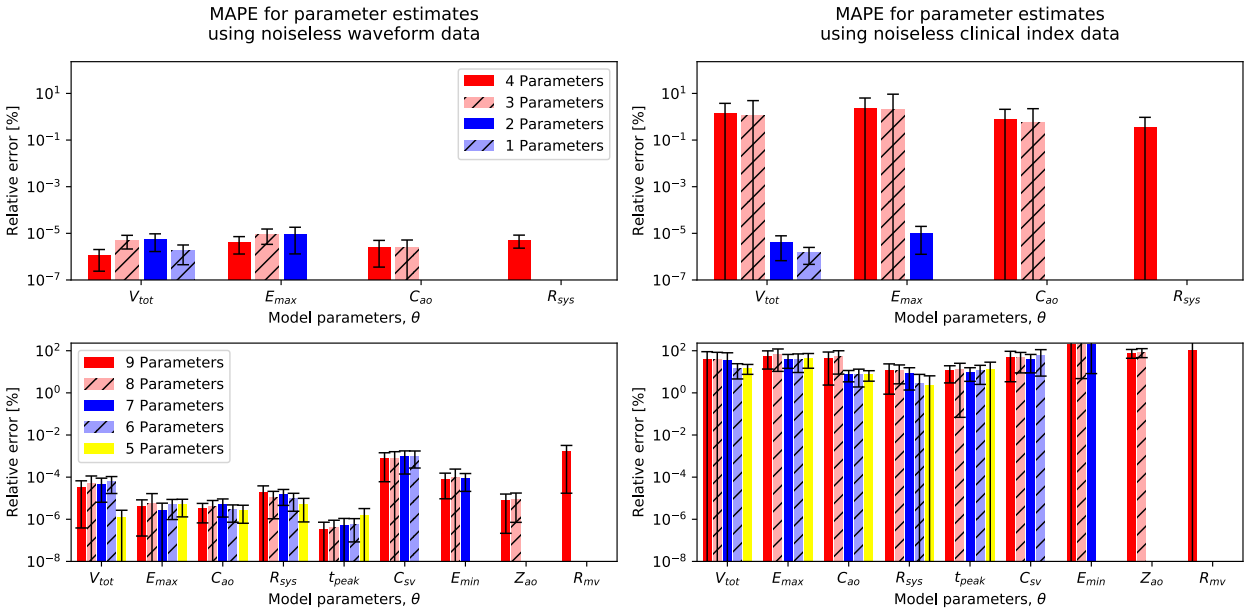

Figure 4: The plots display a comparison of the MAPE for the parameter estimates using cost functions (10) (left) and (9) (right). These results correspond to scenarios (II) and (I). All N_s estimates are presented in this graph. The bottom row of figures illustrate how the estimates change when you estimate all nine parameters at once and gradually fix one by one parameter at their "true" or correct values. The top row illustrates the same procedure when the five least sensitive parameters are always fixed. The number of parameters indicates how many parameters are estimated while the rest are fixed.

only the parameter estimates with the smallest cost functions, which were here defined to be less than 125% of the smallest achieved cost function value, are included. Limiting the selection of estimates to the best cost functions is common practice when performing repeated estimates. This filtering improved the clinical index based estimates for the subsets containing up to four of the most sensitive parameters when all other parameters were fixed at their true value, and made these comparable to the waveform estimates, see Figure 5. Small improvements were seen for most subsets on the left hand side of the figure, but for subsets with more than five parameters, there was no clear trend that the parameter estimates improved.

All parameters of our model were estimated with good accuracy when noise was omitted for the cost function in equation 10. For clinical indices only the estimates of the four most sensitive parameters were reasonably accurate. The maximal MAPE for the waveform parameter estimates was on the order of 10^{-3}%, while for clinical indices this was order 10^{2}%, as shown in Figure 5.

Figure 6 shows little noticeable bias for parameter estimates based on waveform data, though the estimates of C_{sv} and R_{mv} have minor negative and positive bias, respectively. The estimates based on clinical indices are all biased to a noticeable degree with especially large positive biases for $E_{max}, C_{ao}, C_{sv}, E_{min}$ and Z_{ao}.

For the cost function in equation (9), when more than four parameters were estimated, the accuracy was reduced and the worst cases yielded an average error of over 100% for some of the least sensitive parameters. Estimates for the four least sensitive parameters in the model: C_{sv}, E_{min}, C_{sv}, and Z_{ao}, generally displayed larger errors, emphasizing that these parameters were challenging to estimate.

Figure 7 shows the results of the SSRM for the same cost functions as in Figure 5, but with noise (III and IV). The noise reduced the accuracy of estimates, especially for the waveform cost function.

Including either venous indices or waveform data as in cost functions (11) (V) and (12) (VI) improved estimates for at least the seven most sensitive parameters as seen in Figure 8. The estimates of E_{min} and C_{sv}, which largely determine ventricular filling in the model, improved noticeably for both cost functions. However, the cost function based on clinical indices still resulted in estimates with substantially larger errors in the majority of cases.

Figures 9 displays the results found when t_{peak} was fixed at its correct value and the subset selection method was performed only on the remaining parameters (VII). A comparison of Figure 9 and the left panel of Figure 7 reveals that the errors of the parameter estimates were not substantially affected by fixing t_{peak}. The same analysis was conducted for C_{sv} (VIII), but in this case all estimates were improved when estimating five or more parameters

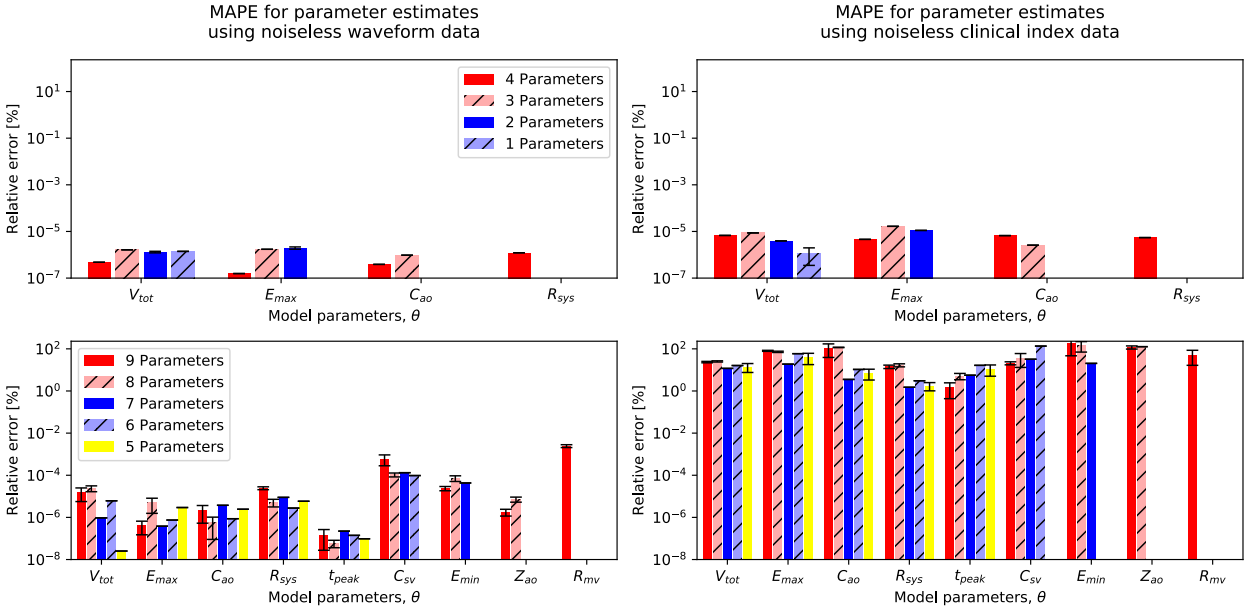

Figure 5: The plots display a comparison of the MAPE for the parameter estimates using cost functions (10) (left) and (9) (right). These results correspond to scenarios (II) and (I). Only the estimates with a cost function value below 125% of the smallest cost function found for each parameter subset were included. The bottom row of figures illustrate how the estimates change when you estimate all nine parameters at once and gradually fix one by one parameter at their "true" or correct value. The top row illustrates the same procedure when the five least sensitive parameters are always fixed. The number of parameters indicates how many parameters are estimated while the rest are fixed.

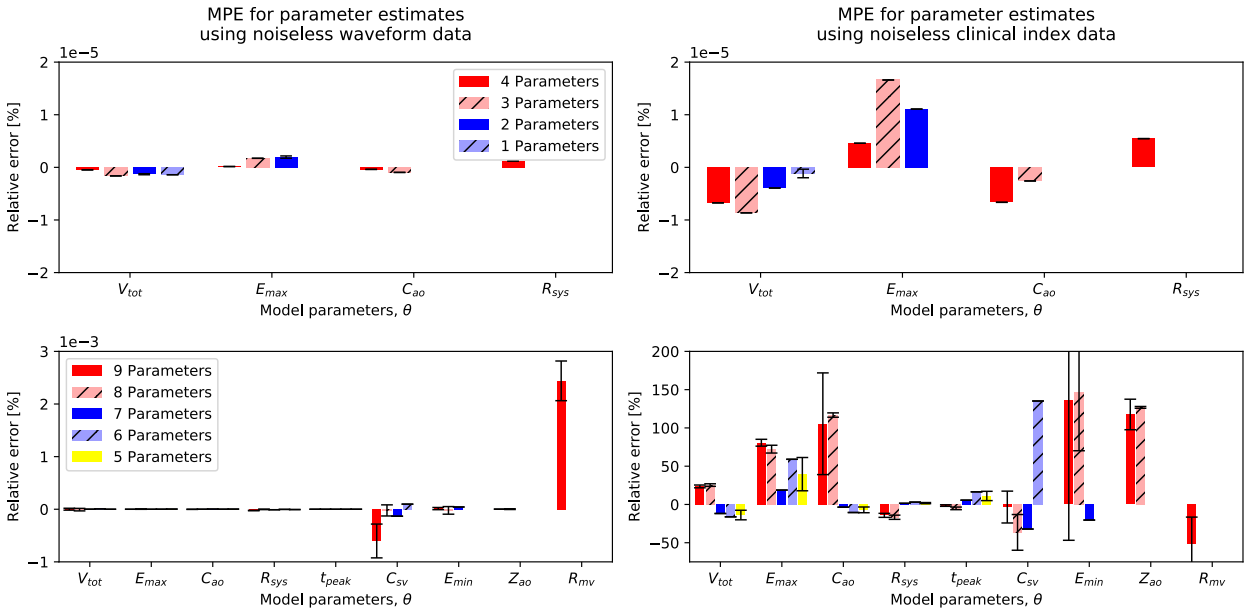

Figure 6: The plots display a comparison of the MPE for the parameter estimates using cost functions (10) (left) and (9) (right). These results correspond to scenarios (II) and (I). Only the estimates with a cost function value below 125% of the smallest cost function found for each parameter subset were included. The MPE metric gives an indication of whether estimates are biased toward being higher or lower than the true value. The bottom row of figures illustrate how the estimates change when you estimate all nine parameters at once and gradually fix one by one parameter at their "true" or correct value. The top row illustrates the same procedure when the five least sensitive parameters are always fixed. The number of parameters indicates how many parameters are estimated while the rest are fixed.

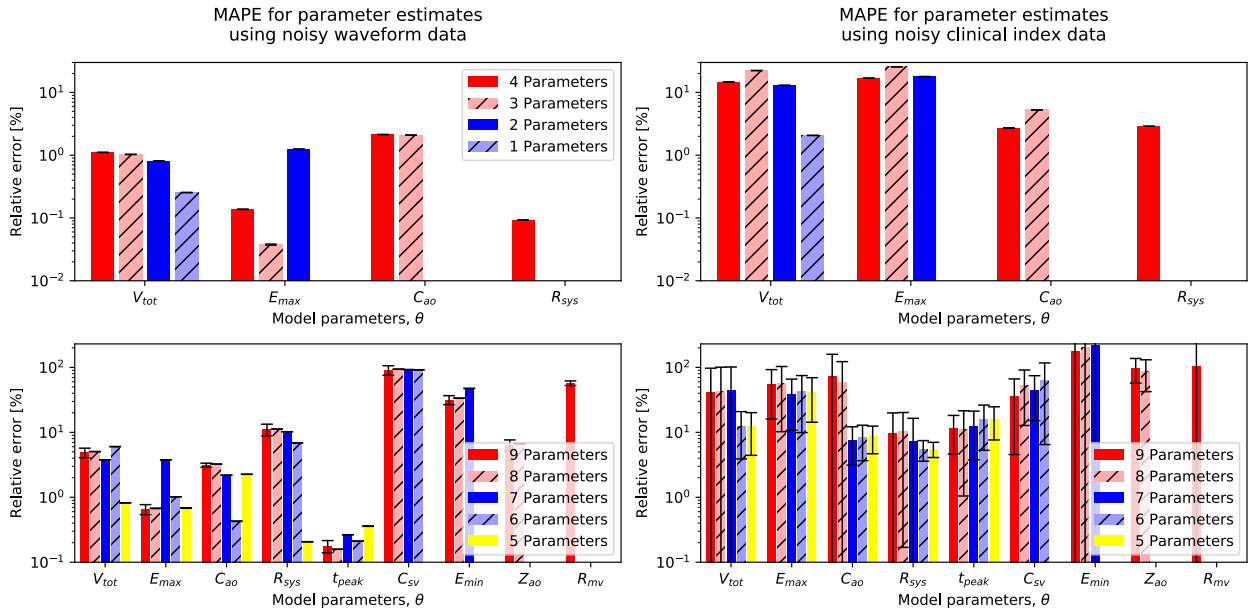

Figure 7: The plots display a comparison of the MAPE for the parameter estimates using cost functions (10) (left) and (9) (right) where 5% normally distributed noise is added to the signals to which the model is fitted. These results correspond to scenarios (IV) and (III). Only the estimates with a cost function value below 125% of the smallest cost function found for each parameter subset were included. The bottom row of figures illustrate how the estimates change when you estimate all nine parameters at once and gradually fix one by one parameter at their "true" or correct value. The top row illustrates the same procedure when the five least sensitive parameters are always fixed. The number of parameters indicates how many parameters are estimated while the rest are fixed.

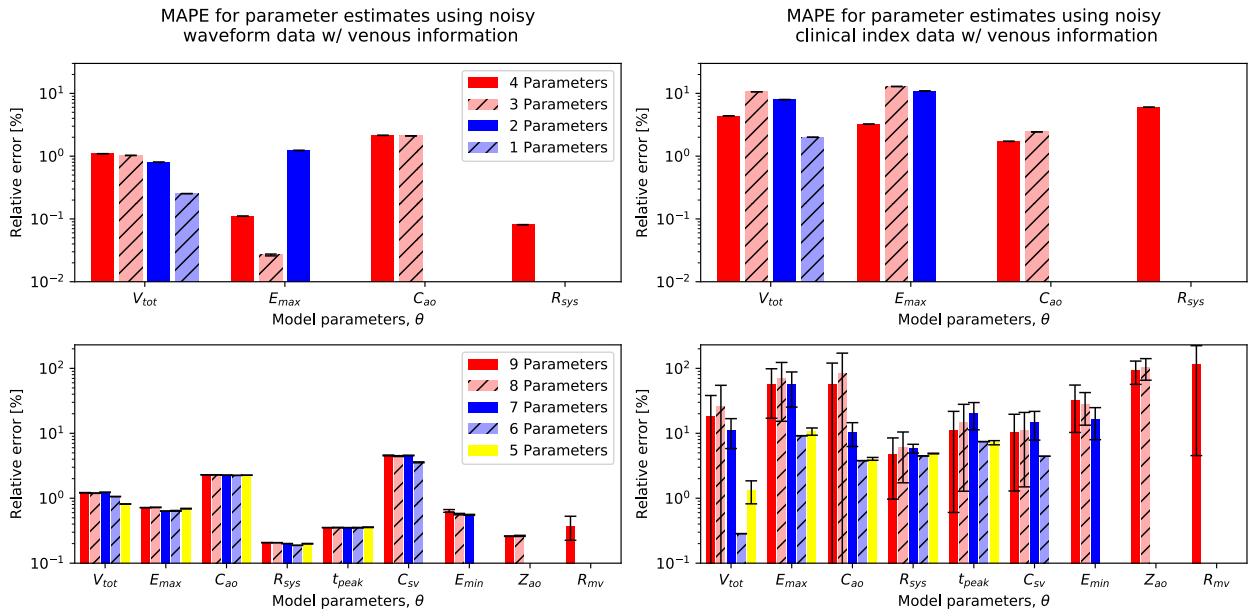

Figure 8: The plots display the results of parameter the MAPE for the parameter estimates using cost function (10) (left) and and (9) (right) for noisy data with added terms containing information from the systemic venous pressure waveform. These results correspond to scenarios (VI) and (V). Only the estimates with a cost function value below 125% of the smallest cost function found for each parameter subset were included. The bottom row of figures illustrate how the estimates change when you estimate all nine parameters at once and gradually fix one by one parameter at their "true" or correct value. The top row illustrates the same procedure when the five least sensitive parameters are always fixed. The number of parameters indicates how many parameters are estimated while the rest are fixed.

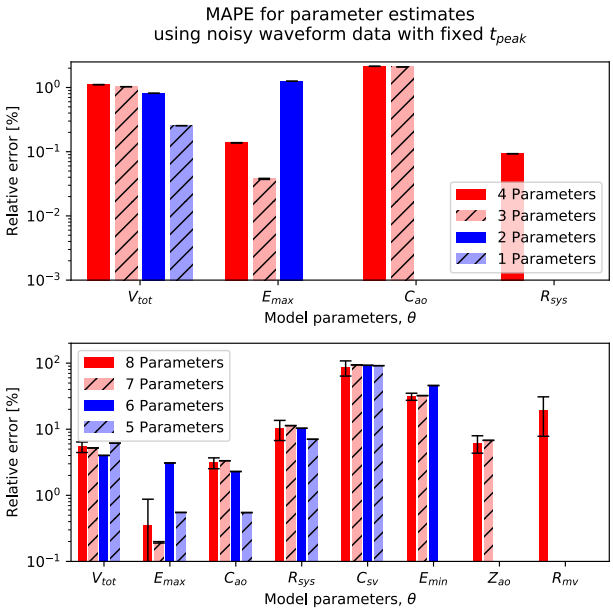

Figure 9: The plots display the MAPE for the parameter estimates using cost function (10) with noise, while always keeping parameter t_{peak} fixed at its true value. These results correspond to scenario (VII). Only the estimates with a cost function value below 125% of the smallest cost function found for each parameter subset were included. The bottom row of figures illustrate how the estimates change when you estimate all nine parameters at once and gradually fix one by one parameter at their "true" or correct value. The top row illustrates the same procedure when the five least sensitive parameters are always fixed. The number of parameters indicates how many parameters are estimated while the rest are fixed.

(compare Figure 10 and the left panel of Figure 7). Most errors were reduced by an order of magnitude, with the exceptions of E_{min} and R_{mv} for which the improvements were between 14.9% and 88.4%.

One estimation case was performed where the noise added to the data was varied for each estimation rather than the initial parameter guesses (IX), see Figures 11 and 13. Some negative biases were observed for R_{sys} and C_{sv}, while the least sensitive parameter R_{mv} exhibited a large positive bias.

Figure 12 shows the results of applying the SSRM to cost function (10), but where fixed parameters were set at 30% above their true value (X). Estimated errors were shown to increase as more parameters were left out of the estimated subsets.

4. Discussion

We sought to evaluate the potential for estimation of personal cardiovascular parameters for a lumped parameter model that may serve as the basis for predicting long-term changes of the resting systemic hemodynamics in an

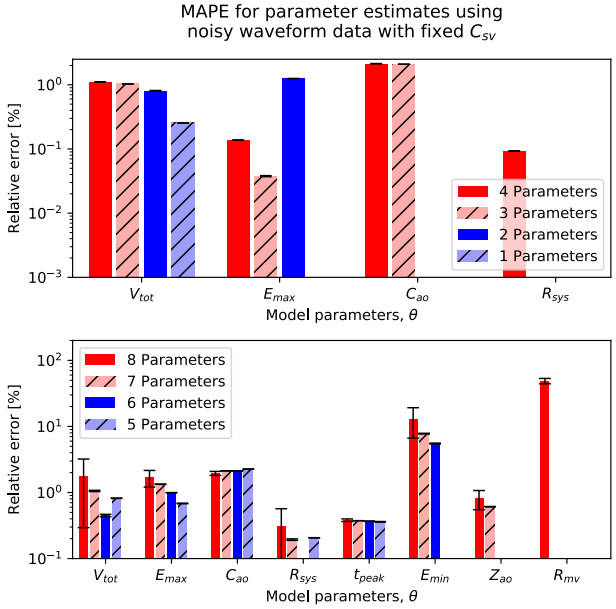

Figure 10: The plots display the MAPE for parameter estimates using cost function (10) with noise, while always keeping parameter C_{sv} fixed at its true value. These results correspond to scenario (VIII). Only the estimates with a cost function value below 125% of the smallest cost function found for each parameter subset were included. The bottom row of figures illustrate how the estimates change when you estimate all nine parameters at once and gradually fix one by one parameter at their "true" or correct value. The top row illustrates the same procedure when the five least sensitive parameters are always fixed. The number of parameters indicates how many parameters are estimated while the rest are fixed.

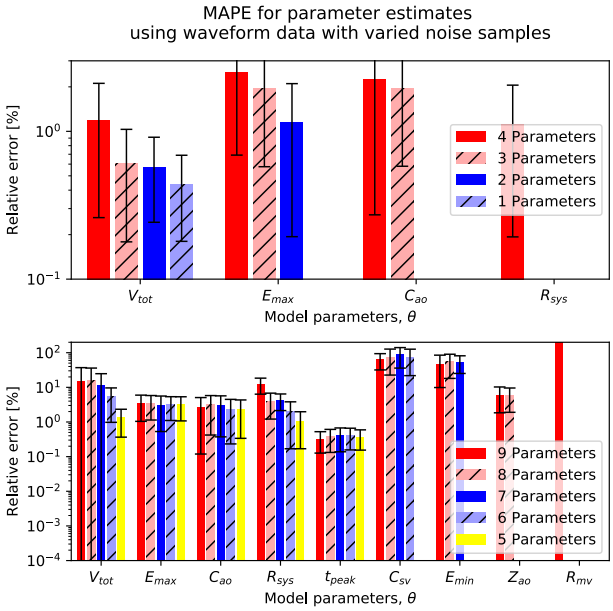

Figure 11: The plots display the MAPE for parameter estimates using cost function (10) with $N_s = 50$ different noise cases, but the same initial parameter guess for each case. These results correspond to scenario (IX). Only the estimates with a cost function value below 125% of the smallest cost function found for each parameter subset were included. The bottom row of figures illustrate how the estimates change when you estimate all nine parameters at once and gradually fix one by one parameter at their "true" or correct value. The top row illustrates the same procedure when the five least sensitive parameters are always fixed. The number of parameters indicates how many parameters are estimated while the rest are fixed.

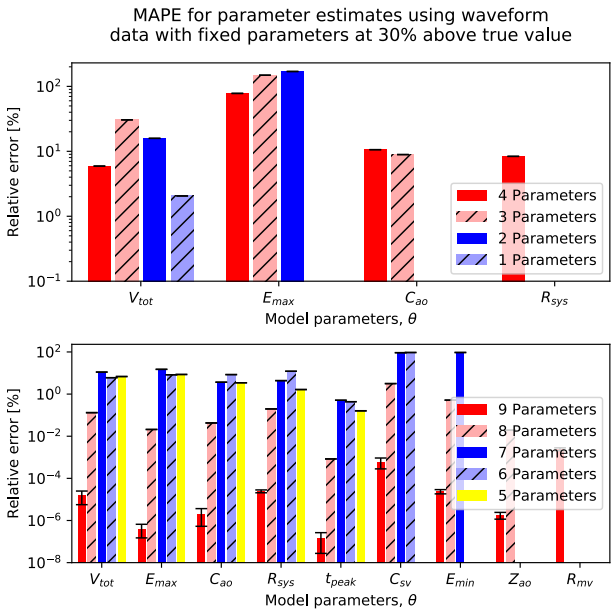

Figure 12: The plots display model fits using cost function (10), but where fixed parameters are fixed at 30% above their true value. These results correspond to scenario (X). Only the estimates with a cost function value below 125% of the smallest cost function found for each parameter subset were included. The bottom row of figures illustrate how the estimates change when you estimate all nine parameters at once and gradually fix one by one parameter at their "true" or correct value. The top row illustrates the same procedure when the five least sensitive parameters are always fixed. The number of parameters indicates how many parameters are estimated while the rest are fixed.

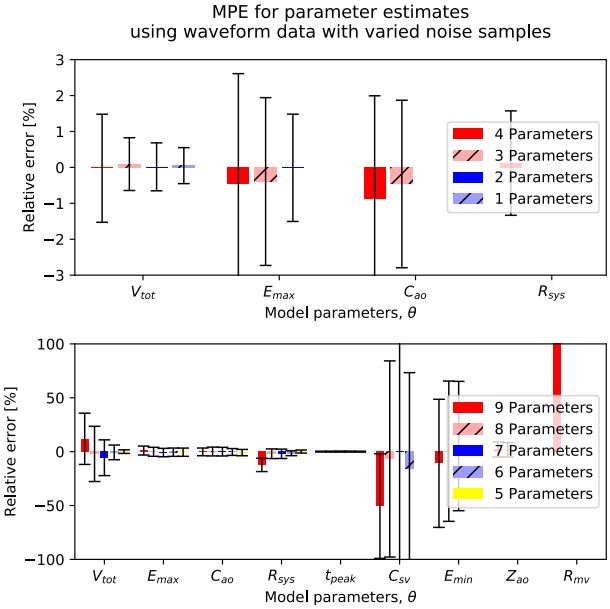

Figure 13: The plots display the MPE for parameter estimates using cost function (10) with N_s different noise cases, but the same initial parameter guess for each case. Only the estimates with a cost function value below 125% of the smallest cost function found for each parameter subset were included. The MPE metric gives an indication of whether estimates are biased toward being higher or lower than the true value. The bottom row of figures illustrate how the estimates change when you estimate all nine parameters at once and gradually fix one by one parameter at their "true" or correct value. The top row illustrates the same procedure when the five least sensitive parameters are always fixed. The number of parameters indicates how many parameters are estimated while the rest are fixed.

individual. In order to know that it is theoretically possible to recover the model parameters at all we first performed a structural identifiability analysis. The model was found to be locally structurally identifiable, hence demonstrating that it is possible to find at least one identifiable parameter vector for a given model output. As the possibility of estimating the model parameters depends on the data available, we compared an approach using only commonly measured clinical quantities to more data rich approaches using waveform data describing pressures and flows. Both types of data sets were synthetically generated from simulations of a computer model. In addition, we expected that attempting to estimate all parameters in the model simultaneously might inhibit the quality of the parameter estimates, so we applied the SSRM to investigate the benefits of fixing some of the parameters to generic values, while estimating the remaining parameters. This performance of the parameter estimation was evaluated by the SSRM for ten distinct combinations of measurement data and constraints.

The results indicate that it is possible to estimate the most influential parameters to the aortic pressure output signal within an accuracy of 10^{-5}% using both cost functions (9) and (10), if the rest of the parameters are set at their correct value. Introducing noise on the other hand, worsens the accuracy for all parameters as seen in Figure 7, yet the time series results are still most reliable despite the added noise. These results are also only achieved in one model realization, so how well this result generalizes to other models is yet uncertain.

Gill et al. analyzed the appropriate step sizes for numerical differentiation of functions with numerical errors and showed that the step size should be larger than the square root of the expected relative error [39]. As the model is solved by a Runge-Kutta numerical integration procedure with relative error tolerance set to 10^{-9}, the minimum step-size for numerical differentiation of the model output is $10^{-4.5}$. Consequently, it should not be expected to estimate parameters to greater accuracy than this when using a numerical gradient based optimization method. The results show the parameters are estimated to the expected accuracy, well below any physically significant error level.

The addition of noise reduced accuracy in the parameter estimates based on waveform data, see the left hand side of Figure 7. The four previously well estimated parameters became biased when the model was fitted to clinical indices, although still with high precision when a subset of four or fewer parameters were estimated, as seen on the right hand side of Figure 7. When five or more parameters were estimated, the noise did result in poorer estimates compared to the noiseless case shown on the right hand side of Figure 5. By further inspection, it emerged that the waveform data estimates were more accurate than for

clinical indices when using noisy data in terms of average error, especially when estimating only the most sensitive parameters.

Inclusion of venous information in both cost function classes proved beneficial, especially for estimates concerning parameters important pertaining to the filling phase of the heart, in this case C_{sv} and E_{min}. This was not unexpected as more data in a compartment which is the source of inflow for the heart should provide more information to these parameters.

Figure 7 shows that the average errors decreased substantially after C_{sv} was removed from the estimated subset. C_{sv} was the first parameter to appear with errors on the scale of 10^2 when added to the estimated subset. The results of fixing C_{sv} for all estimated subsets, can be seen in Figure 10. A substantial improvement in most estimation errors was observed. Similarly to adding venous information to the cost functions, estimation of the ventricular filling parameters improved, but also estimation of some more sensitive parameters, indicating that this parameter has a larger influence on the estimates of other parameters. Therefore, making accurate estimates for the central systemic venous compliance C_{sv} can be an alternative to inconvenient pressure measurements in the central veins.

An interesting observation can be made from Figure 8 for the waveform data parameter estimates: As soon as 5 or more parameters are selected in the estimated subset, the estimation error for all parameters are practically constant and unperturbed by estimation of additional parameters. This signifies that the deviation from true parameters is sufficiently small, which according to the figure is 10% or less, such that the estimation error in the least sensitive parameters does not influence the other estimates in any way. It appears that the same minimum is found in the cost function regardless of the additional lesser sensitive parameters, and this must then mean the noise level in the data limits more accurate optimization with respect to parameters. For the large majority of parameters the error is well below 5% which was the specified noise amplitude. Consequentially, if the venous pressure waveform is available, all parameters should be obtainable to within 10% accuracy based on the noise sample used. Compared to Figure 12 where parameters are fixed at 30% above their true value, but without noise and venous data we find errors of approximately 10% for most estimates when less than 7 parameters are estimated, but some errors approach 100% as progressively fewer parameters are estimated. Therefore, the estimation procedure seems more sensitive to poor estimates for fixed parameters than to noise in the data. This is even the case when the three least sensitive parameters are fixed at their wrong values.

In Figure 13 we observe that for most parameters, the estimation error is averaged out and shows little or no bias for most parameters over a larger number of noisy samples as one might expect for Gaussian noise. The four least sensitive parameters do show some bias which emphasizes that estimates of these can not be trusted in general. In Figure 11 presenting MAPE, the errors are maximally 15.4% for the five most sensitive parameters, and up to 88.1% for the next three most sensitive parameters, and even over 10^3% for R_{mv}. Therefore, noise can have a large impact on a single cycle of data. However the estimate error variance is for the most part quite low, so based on this it appears that it is not the details of the different noise samples which mainly dictate the estimation error, but rather the presence of normally distributed noise of the given magnitude.

From our investigation it was also demonstrated that fixing parameters at their wrong values will substantially affect estimates. We observed that the estimation error fell in a range of 0.1% - 100% for all parameters as seen in Figure 7. Adding 30% to the fixed parameter values was then determined to be a suitable level for comparison on average. Figure 12 shows that the estimation errors then fell in the range of 10^{-6}% - 50% for most parameters in the model, regardless of which estimated subset was considered. When all 9 parameters are estimated, none are fixed at erroneous values. As more parameters were fixed, the estimation error also grew rapidly. This was true even for a noiseless case, implying that parameters set at wrong values may strongly influence the other parameter estimates. To be able to personalize, or calibrate this model means that the three or four least sensitive parameters should be estimated or measured to high accuracy, before the chosen least squares optimization methods can estimate the rest to good accuracy, unless venous pressure data are available. Some of three least sensitive parameters can be estimated from pressure flow and/or volume measurements, but most often require measurements in or around the left ventricle [40, 41], which is costly and burdensome for an individual. In the case of E_{min} one may perhaps estimate this from the end-diastolic pressure-volume relationship but there are not any known examples of this being demonstrated.

The SSRM was performed in the sequence prescribed by the parameter sensitivity ranking according to the aortic pressure time series in Table 5, because this is among the most easily observed outputs to predict and is therefore also used to constrain the cost function either as a time series or as extrema values. The optimization approach presented relies solely on the local Quasi-Newton optimization method TRRA, as opposed to more computationally expensive global optimization methods such as for instance Markov Chain Monte Carlo methods, or genetic algorithms [12]. Partial motivation for this work is to develop methods where computational time is another parameter to be minimized and is the reason why computa-

tionally cheaper algorithms are preferred in this investigation. Other notions of sensitivity could have been chosen to rank the parameters, but the total variance weighted averaged Sobol indices were chosen in this instance. For the notions of sensitivity tested the least sensitive three parameters were often the same. On the other hand there were some variations among the more sensitive parameters, but there were no dominating patterns.

Only for a limited subset of the most sensitive parameters is the local optimization by TRRA robust for parameter estimation given that all other parameters are fixed at their true values. Robust means here that the results are practically independent of the initial parameter values, and that the estimates are made to low error. The four least sensitive parameters C_{sv}, E_{min}, Z_{ao} and R_{mv} could not be reliably estimated in all cases as they were estimated with large errors. The conclusion must be that the inclusion of these parameters makes the cost function very flat around the true minimum in parameter space and causes difficulties for numerical optimization method. Conclusions based on the estimates of these parameters can therefore not be trusted in a majority of cases. Counter intuitively, the second least influential parameter Z_{ao} could for cost functions using waveforms often be better estimated compared to the other three least influential parameters. This may be explained by limited resolution in the sensitivity analysis. For parameters of sensitivities below 0.1 it is challenging to make an accurate ranking, and hence Z_{ao} might actually be more influential than the other three parameters since they were all estimated to similar low values.

The data used in this analysis is purely synthetic, which in contrast to real data exhibit no model discrepancy, and has no structured noise or hidden bias. Still, in this perfect case it is not always possible to fit the model to its own output data using the applied methods given randomly sampled initial parameter guesses. This lends credence to the methods presented by Colunga et al., and de Bournonville et al. which emphasize that using good nominal values used as initial values for the optimization problem is necessary for accurate parameter estimation [8, 7]. To make the estimation scenario more realistic we applied Gaussian noise to the output signals from the model. The standard clinical method of measuring blood pressure is to measure the systolic and diastolic blood pressure by cuff plethysmography and the measurement uncertainty is widely regarded to be ± 5 mmHg. Using 100 mmHg as a reference scale for blood pressure and adding noise from a normal distribution with 5% standard deviation the error should be comparable and even introduce larger deviations at some points, making this a realistic measurement error. During measurement of blood flow velocity a slight misalignment of only 2 deg between measurement probe and direction of blood flow may introduce large errors, so as a general rule of thumb for Doppler velocity uncertainty is $\pm 10\%$. As we have observed when comparing the results from using noiseless to noisy data the results are majorly affected by adding 5% normally distributed noise, and we hypothesize that knowing personalized parameters accurately is essential to make personalized predictions about an individual. However, it would also be beneficial if measurements could be made more readily available and less burdensome to the patient such that more data can be collected and make it easier to perform analysis for an individual at all, rather than be limited by not having any data to attempt this.

This study can be described as a practical identifiability study where we attempted to personalize a model to a given set of data under both ideal and noiseless conditions. The structural identifiability was demonstrated and supported by the demonstration of practically recovering the true parameters in the noiseless case using waveforms. The other scenarios investigated in this paper showed that the model was not practically identifiable by the given set of measurements with noise. By application of the STRIKE-GOLDD software tool we showed that the model was locally structurally identifiable using aortic pressure and flow waveforms. Local structural identifiability implies that there may be an infinite set of solutions to the problem, yet we seem to find the correct cost function minimum in the noiseless case using waveforms which may support that there is only one solution in the chosen domain of valid parameters.

The estimation results were chosen by picking the once with the best cost function values, and therefore the limit was set at a cost function value better than the 125% of the best cost function value. Controlling the cost function by picking the best results are common practice in optimization problems. The result however is that in some cases depending on the variation in the results all results are included if all estimates yield an equally good cost function value and the operation means nothing is done. In other cases the limit may be too strict and only a few of the results are used in the final average. This also means that the standard deviation bars are calculated based on little data and may be uncertain.

In most of the results we observed uneven improvement in estimates as the subset size of estimated parameters decreased. Some of the uneven variation may be explained by the addition of noise and finite points of the waveform data. The minimum which the optimization algorithm finds for noisy data may not correspond exactly to the solution defined by $\theta_{\mathbf{true}}$. Therefore, as different parameter subsets are optimized for the perturbed optimal solution the estimates may unevenly approach more accurate estimates as the metric used can compare a cost function minimum which differs slightly from the minimum defined by the noiseless waveforms.

The results of our study are somewhat in contrast to prior works on parameter estimation and subset selection for cardiovascular models [36, 42] which propose that fixing insensitive and some strongly correlated parameters may improve parameter estimates and model fits. Some of these methodologies have been tested primarily on fitting models to experimental or clinical data, and have not been evaluated on synthetic data sets where the true values of model parameters can be directly known. Many such studies argue that parameter estimates show good convergence to the estimated values (e.g. Pant et al. [10]) and some have employed statistical approaches to quantify uncertainty and variability of the parameter estimates such as Colunga et al. and Marquis et al. [8, 6]. These studies have estimated lower uncertainties in parameter estimates than have been found in our synthetic approach, but use more sophisticated optimization methods. While the models, measurements, and parameter estimation methods differ in numerous ways, the patterns in these results may well be a general feature of lumped parameter circulatory models.

5. Conclusion

We have applied the step-wise subset reduction method (SSRM) to a closed-loop lumped parameter model in order to investigate how well this model can be calibrated using limited, in this case synthetic, data. The envisioned application of the presented model is to predict long term changes in resting systemic hemodynamics. Consequentially we have demonstrated a framework for assessing the accuracy and precision of parameter estimation for different subsets of model parameters. We first performed a structural identifiability analysis which revealed the model to be structurally identifiable with waveform data from the aorta. The local optimization methods applied in this analysis were sensitive to initial parameter guesses, and we investigated the effect on the precision and accuracy of estimated parameters. We demonstrated, that using waveform data as opposed to scalar clinical indices improves the accuracy of parameter estimates, and the waveform data cost functions are far more robust in terms of providing the best estimates despite the introduction of normally distributed noise to the indices. Standard local optimization methods can be used for model calibration of the five most sensitive parameters in the presented simple-lumped parameter model of the systemic circulation and left ventricle. The most sensitive parameters are generally recovered with errors less than 10% given that the other parameters are fixed at correct values, and noise is normally distributed at 5%. We also showed that when parameters were fixed at 30% above their true value, estimation error can be comparable to the error introduced by noise even if only the three least sensitive parameters are fixed at erroneous values. The four least sensitive parameters could not be generally reliably estimated by this method without more information, and no conclusions should be based on estimates of these parameters. The three least sensitive parameters especially should be considered for measurements. The addition of venous information to the cost function further improved parameter estimates, and after analysis by the SSRM the central venous compliance of the systemic circulation was found to be a positive influence on the less sensitive parameter estimates if it could be measured and fixed at an accurate value.

6. Acknowledgements

Some of the computations were performed on resources provided by the NTNU IDUN/EPIC computing cluster [32].

References

[1] P. D. Morris, D. Ryan, A. C. Morton, R. Lycett, P. V. Lawford, D. R. Hose, and J. P. Gunn, "Virtual fractional flow reserve from coronary angiography: modeling the significance of coronary lesions: results from the virtu-1 (virtual fractional flow reserve from coronary angiography) study.," JACC. Cardiovascular interventions, vol. 6, pp. 149–57, Feb 2013.

[2] E. Kung, G. Pennati, F. Migliavacca, T.-Y. Hsia, R. Figliola, A. Marsden, and A. Giardini, "A simulation protocol for exercise physiology in fontan patients using a closed loop lumped-parameter model.," Journal of biomechanical engineering, vol. 136, pp. 0810071–08100714, Aug 2014.

[3] L. Fresiello, B. Meyns, A. Di Molfetta, and G. Ferrari, "A model of the cardiorespiratory response to aerobic exercise in healthy and heart failure conditions," Frontiers in Physiology, vol. 7, 2016.

[4] A. Di Molfetta, A. Amodeo, M. G. Gagliardi, M. G. Trivella, L. Fresiello, S. Filippelli, A. Toscano, and G. Ferrari, "Hemodynamic effects of ventricular assist device implantation on norwood, glenn, and fontan circulation: A simulation study.," Artificial organs, vol. 40, pp. 34–42, Jan 2016.

[5] H. Ho, H. B. Yu, A. Bartlett, and P. Hunter, "An in silico pipeline for subject-specific hemodynamics analysis in liver surgery planning," Computer Methods in Biomechanics and Biomedical Engineering, vol. 23, no. 4, pp. 138–142, 2020. PMID: 31928213.

[6] A. D. Marquis, A. Arnold, C. Dean-Bernhoft, B. E. Carlson, and M. S. Olufsen, "Practical identifiability and uncertainty quantification of a pulsatile cardiovascular model," Mathematical Biosciences, vol. 304, pp. 9 – 24, 2018.

[7] S. de Bournonville, A. Pironet, C. Pretty, J. G. Chase, and T. Desaive, "Parameter estimation in a minimal model of cardio-pulmonary interactions," Mathematical Biosciences, vol. 313, pp. 81 – 94, 2019.

[8] A. L. Colunga, K. G. Kim, N. P. Woodall, T. F. Dardas, J. H. Gennari, M. S. Olufsen, and B. E. Carlson, "Deep phenotyping of cardiac function in heart transplant patients using cardiovascular system models," The Journal of Physiology, vol. 598, no. 15, pp. 3203–3222, 2020.

[9] S. Pant, B. Fabrèges, J. Gerbeau, and I. E. Vignon-Clementel, "A methodological paradigm for patient-specific multi-scale cfd simulations: from clinical measurements to parameter estimates for individual analysis.," International Journal for Numerical Methods in Biomedical Engineering, vol. 30, no. 12, pp. 1614 – 1648, 2014.

[10] S. Pant, C. Corsini, C. Baker, T.-Y. Hsia, G. Pennati, and I. E. Vignon-Clementel, "Data assimilation and modelling of patient-specific single-ventricle physiology with and without valve regurgitation," Journal of Biomechanics, vol. 49, no. 11, pp. 2162

– 2173, 2016. Selected Articles from the International Conference on CFD in Medicine and Biology (Albufeira, Portugal – August 30th - September 4th, 2015).

[11] R. Meiburg, W. Huberts, M. C. M. Rutten, and F. N. van de Vosse, "Uncertainty in model-based treatment decision support: Applied to aortic valve stenosis," International Journal for Numerical Methods in Biomedical Engineering, vol. 36, no. 10, p. e3388, 2020.

[12] C. Hann, J. Chase, T. Desaive, C. Froissart, J. Revie, D. Stevenson, B. Lambermont, A. Ghuysen, P. Kolh, and G. Shaw, "Unique parameter identification for cardiac diagnosis in critical care using minimal data sets," Computer Methods and Programs in Biomedicine, vol. 99, no. 1, pp. 75 – 87, 2010.

[13] N. Stergiopulos, J. J. Meister, and N. Westerhof, "Determinants of stroke volume and systolic and diastolic aortic pressure.," The American journal of physiology, vol. 270, pp. H2050–9, Jun 1996.

[14] Y. Shi, P. Lawford, and R. Hose, "Review of zero-d and 1-d models of blood flow in the cardiovascular system," BioMedical Engineering OnLine, vol. 10, p. 33, Apr. 2011.

[15] J. P. Mynard and J. J. Smolich, "One-dimensional haemodynamic modeling and wave dynamics in the entire adult circulation," Annals of Biomedical Engineering, vol. 43, pp. 1443–1460, June 2015.

[16] P. Segers, N. Stergiopulos, P. Verdonck, and R. Verhoeven, "Assessment of distributed arterial network models," Medical and Biological Engineering and Computing, vol. 35, pp. 729–736, Nov. 1997.

[17] B. W. Smith, J. Chase, R. I. Nokes, G. M. Shaw, and G. Wake, "Minimal haemodynamic system model including ventricular interaction and valve dynamics," Medical Engineering & Physics, vol. 26, no. 2, pp. 131 – 139, 2004.

[18] P. Segers, N. Stergiopulos, and N. Westerhof, "Quantification of the contribution of cardiac and arterial remodeling to hypertension," Hypertension, vol. 36, no. 5, pp. 760–765, 2000.

[19] P. Segers, E. R. Rietzschel, M. L. De Buyzere, N. Stergiopulos, N. Westerhof, L. M. Van Bortel, T. Gillebert, and P. R. Verdonck, "Three- and four-element windkessel models: assessment of their fitting performance in a large cohort of healthy middle-aged individuals.," Proceedings of the Institution of Mechanical Engineers. Part H, Journal of engineering in medicine, vol. 222, pp. 417–28, May 2008.

[20] J.-L. Vachiéry and S. Gaine, "Challenges in the diagnosis and treatment of pulmonary arterial hypertension," European Respiratory Review, vol. 21, no. 126, pp. 313–320, 2012.

[21] A. La Gerche, D. J. Rakhit, and G. Claessen, "Exercise and the right ventricle: a potential Achilles' heel," Cardiovascular Research, vol. 113, pp. 1499–1508, 08 2017.

[22] T. D. Miller and N. S. Anavekar, "Right ventricle and exercise capacity," Circulation: Cardiovascular Imaging, vol. 9, no. 11, p. e005703, 2016.

[23] W. Huberts, S. G. Heinen, N. Zonnebeld, D. A. van Den Heuvel, J.-P. P. de Vries, J. H. Tordoir, D. R. Hose, T. Delhaas, and F. N. van de Vosse, "What is needed to make cardiovascular models suitable for clinical decision support? a viewpoint paper," Journal of Computational Science, vol. 24, pp. 68–84, 2018.

[24] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. Jarrod Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. Carey, l. Polat, Y. Feng, E. W. Moore, J. Vand erPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and S. . . Contributors, "Scipy 1.0: Fundamental algorithms for scientific computing in python," Nature Methods, 2020.

[25] H. Miyashita, "Clinical assessment of central blood pressure.," Current hypertension reviews, vol. 8, pp. 80–90, May 2012.

[26] J. Lee, J. Sohn, J. Park, S. Yang, S. Lee, and H. C. Kim, "Novel blood pressure and pulse pressure estimation based on pulse transit time and stroke volume approximation.," Biomedical engineering online, vol. 17, p. 81, Jun 2018.

[27] C. W. L. Chin, H. J. Khaw, E. Luo, S. Tan, A. C. White, D. E. Newby, and M. R. Dweck, "Echocardiography underestimates stroke volume and aortic valve area: implications for patients with small-area low-gradient aortic stenosis.," The Canadian journal of cardiology, vol. 30, pp. 1064–72, Sep 2014.

[28] H. Miao, X. Xia, A. S. Perelson, and H. Wu, "On identifiability of nonlinear ode models and applications in viral dynamics.," SIAM review. Society for Industrial and Applied Mathematics, vol. 53, pp. 3–39, Jan 2011.

[29] L. Ljung and T. Glad, "On global identifiability for arbitrary model parametrizations," Automatica, vol. 30, no. 2, pp. 265–276, 1994.

[30] A. F. Villaverde, A. Barreiro, and A. Papachristodoulou, "Structural identifiability of dynamic systems biology models," PLOS Computational Biology, vol. 12, pp. 1–22, 10 2016.

[31] A. Pironet, P. C. Dauby, J. G. Chase, P. D. Docherty, J. A. Revie, and T. Desaive, "Structural identifiability analysis of a cardiovascular system model," Medical Engineering & Physics, vol. 38, no. 5, pp. 433–441, 2016.

[32] M. Själander, M. Jahre, G. Tufte, and N. Reissmann, "EPIC: An energy-efficient, high-performance GPGPU computing research infrastructure," 2019.

[33] V. G. Eck, W. P. Donders, J. Sturdy, J. Feinberg, T. Delhaas, L. R. Hellevik, and W. Huberts, "A guide to uncertainty quantification and sensitivity analysis for cardiovascular applications," International Journal for Numerical Methods in Biomedical Engineering, vol. 32, no. 8, pp. n/a–n/a, 2016.

[34] W. P. Donders, W. Huberts, F. N. Vosse, and T. Delhaas, "Personalization of models with many model parameters: an efficient sensitivity analysis approach.," International Journal for Numerical Methods in Biomedical Engineering, vol. 31, no. 10, pp. n/a – N.PAG, 2015.

[35] A. Saltelli and P. Annoni, "How to avoid a perfunctory sensitivity analysis," Environmental Modelling & Software, vol. 25, no. 12, pp. 1508 – 1517, 2010.

[36] M. S. Olufsen and J. T. Ottesen, "A practical approach to parameter estimation applied to model predicting heart rate regulation," Journal of Mathematical Biology, vol. 67, pp. 39–68, July 2013.

[37] C. H. Olsen, J. T. Ottesen, R. C. Smith, and M. S. Olufsen, "Parameter subset selection techniques for problems in mathematical biology," Biological Cybernetics, vol. 113, pp. 121–138, Apr. 2019.

[38] A. Alexanderian, P. A. Gremaud, and R. C. Smith, "Variance-based sensitivity analysis for time-dependent processes," Reliability Engineering & System Safety, vol. 196, p. 106722, 2020.

[39] P. E. Gill, W. Murray, M. A. Saunders, and M. H. Wright, "Computing Forward-Difference Intervals for Numerical Optimization," SIAM Journal on Scientific and Statistical Computing, vol. 4, pp. 310–321, June 1983.

[40] H. S. Roshdy, A. M. Meshrif, and I. I. El-Dosouky, "Value of the mitral valve resistance in evaluation of symptomatic patients with mild and moderate mitral stenosis–a dobutamine stress echocardiographic study.," Echocardiography (Mount Kisco, N.Y.), vol. 31, pp. 347–52, Mar 2014.

[41] R. Kelly and D. Fitchett, "Noninvasive determination of aortic input impedance and external left ventricular power output: A validation and repeatability study of a new technique," Journal of the American College of Cardiology, vol. 20, no. 4, pp. 952 – 963, 1992.

[42] A. Cintrón-Arias, H. T. Banks, A. Capaldi, and A. L. Lloyd, "A sensitivity matrix based methodology for inverse problem formulation," Journal of Inverse and Ill-posed Problems, vol. 17, Jan. 2009.

Appendices

Appendix A. Model Equations

The model is comprised as a system of nonlinear ODEs describing the volume state variable of the three model compartments. The dynamic elastance function of the heart contributes a source term to the system of ODEs.

$$\begin{aligned} \frac{dV_{sa}}{dt} &= C_{sa}\frac{dP_{sa}}{dt} = Q_{lvao} - Q_{sys} \\ \frac{dV_{sv}}{dt} &= C_{sv}\frac{dP_{sv}}{dt} = Q_{sys} - Q_{svlv} \\ \frac{dV_{lv}}{dt} &= Q_{svlv} - Q_{lvao} \end{aligned} \tag{A.1}$$

V_{sa}, V_{sv} and V_{lv} are the stressed blood volumes of the systemic arteries, systemic veins and left ventricle respectively. C_{sa} and C_{sv} are the compliance values of the systemic arteries and veins, while P_{sa} and P_{sv} are the corresponding pressures of these compartments. Q_{lvao} denotes the volume blood flow from the left ventricle to the systemic arteries, Q_{sys} is the flow between the systemic arteries and veins, and finally Q_{svlv} is the flow from the veins to the left ventricle. The two remaining state variables, pressure and flow, are mainly modelled by linear relationships as expressed below

$$\begin{aligned} V_{sa} &= C_{sa}P_{sa} \\ V_{sv} &= C_{sv}P_{sv} \\ P_{lv} &= E_{lv}(t)V_{lv} + P_{th}(t) \\ E_{lv}(t) &= (E_{\max} - E_{\min})e(t) + E_{\min} \\ P_{ao} &= \max\left[P_{sa}, P_{lv}\right] \\ Q_{lvao} &= I(P_{lv} > P_{sa})\frac{P_{lv} - P_{sa}}{Z_{ao}} \\ Q_{svlv} &= I(P_{sv} > P_{lv})\frac{P_{sv} - P_{lv}}{R_{mv}} \\ Q_{sys} &= \frac{P_{sa} - P_{sv}}{R_{sys}}. \end{aligned} \tag{A.2}$$

P_{ao} indicates aortic pressure, while P_{lv} denotes left ventricular pressure. Z_{ao} is the characteristic aortic impedance while R_{sys} and R_{mv} are total systemic vascular and mitral valve resistances. The indicator function $I(x)$ has the value 1 when the argument x is true and 0 when x is false. The activation function $e(\tau)$ is defined as

$$e(\tau) = \alpha \times \frac{(\tau/a_1)^{n_1}}{1 + (\tau/a_1)^{n_1}} \times \frac{1}{1 + (\tau/a_2)^{n_2}} \tag{A.3}$$

where τ is position in the cardiac cycle between the end of the last diastolic period and the end of the next diastolic period $\tau = 1$. The parameters a_1 and n_1 determine the shape of contraction and a_2 and n_2 determine the shape of relaxation of the elastance curve and the timing of peak elastance. The choice of values for these parameters are identical to those of Stergiopulos et al. [13]. We wrote the parameter values for a_1 and a_2 in terms of the ratio of $\frac{t_{peak}}{T}$, and set $\alpha = 1.672$, to ensure normalization of the curve. t_{peak} describes the time of peak ventricular elastance, and therefore determines when the left ventricular elastance $E_{lv}(t)$ reaches E_{max}. The intrathoracic pressure function P_{th} describes the external pressure effects on the ventricular muscle aside from pressure gradients inside the blood vessels and is here modeled as a constant of $P_{th}(t) = -4$ mmHg. Otherwise the parameters are as defined by Table 1.

The V_{tot} parameter describes total stressed volume and is enforced by setting intial compartment volumes and pressures such that the total stressed volume equals the parameter value. The model was demonstrated to conserve the volume and hence the total blood volume will not change. The inital volumes and pressures are set according to the equations

$$\begin{aligned} V_{ao,0} &= C_{ao}P_{ao,0}, \text{and} \\ P_{sv,0} &= \frac{V_{tot} - V_{ao,0} - V_{lv,0}}{C_{sv}} \end{aligned} \tag{A.4}$$

where the initial aortic pressure is set to $P_{ao,0} = 100$ mmHg, and initial left ventricular volume is set to $V_{lv,0} = 100$ mL. The initial venous pressure is denoted by $P_{sv,0}$.

Appendix B. Algorithm specification

The algorithm applied for computations within this study was chosen to be the Trust Reflective Region Algorithm as implemented in SciPy version 1.4.1 [24].

The function scipy.optimize.least_squares() function is applied given the arguments listed below and a initial parameter guess. The reference parameters for the initial parameter guesses are listed in 1 and are sampled as vectors with normally distributed noise with a standard deviation of 30% as per equation (13). Each parameter is assigned its own random perturbation, except T which is always fixed. The parameters are also assigned upper and lower bounds, which are set to the values given in Table B.6. The initial parameter guesses are limited to fall within these bounds.

Scaling factors used to balance the different components in the cost functions as specified in section 2 are listed in Table B.7.

Otherwise some function specific parameters for the accuracy of the method are set as $xtol = 2.3 \cdot 10^{-16}$, $ftol =$

Parameter	Upper bounds	Lower bounds	Sampling value	Units
C_{ao}	10.0	0.5	1.0	$\frac{\text{mL}}{\text{mmHg}}$
C_{sv}	30.0	0.5	10.0	$\frac{\text{mL}}{\text{mmHg}}$
E_{max}	5.0	0.9	2.0	$\frac{\text{mmHg}}{\text{mL}}$
E_{min}	1.0	0.0	0.06	$\frac{\text{mmHg}}{\text{mL}}$
R_{mv}	0.1	0.0	0.003	$\frac{\text{mmHg s}}{\text{mL}}$
R_{sys}	3.0	0.5	1.0	$\frac{\text{mmHg s}}{\text{mL}}$
T	0.85	0.85	0.85	s
t_{peak}	0.75	0.05	0.32	-
V_{tot}	2000.	50.	250.	mL
Z_{ao}	1.0	0.0	0.1	$\frac{\text{mmHg s}}{\text{mL}}$

Table B.6: All model parameters that are assigned to be personalizable are listed along with their upper and lower bounds as chosen for this study. The sampling value is the mean value of the normal distribution from which initial parameter guesses are sampled.

Symbol	Value	Unit
K_p	100.0	mmHg
K_q	500.0	$\frac{\text{mL}}{\text{s}}$
K_{PP}	40.0	mmHg
K_{psys}	120.0	mmHg
K_{pdia}	80.0	mmHg
K_{psvsys}	20.0	mmHg
K_{psvdia}	20.0	mmHg
K_{psv}	20.0	mmHg
K_{SV}	100.0	mL
K_{qmax}	500.0	mL

Table B.7: The scaling factors K which are used to balance and approximately normalize the terms in the specified cost functions. Subscripts: p - Aortic pressue waveform, q - Aortic flow, PP - pulse pressure, psys - Systolic aortic pressure, pdia - Diastolic aortic pressure, psvsys - Systolic averaged venous pressure, psvdia - Diastolic averaged venous pressure, psv - Averaged venous pressure waveform, SV - stroke volume, and qmax - maximal aortic flow.

$2.3 \cdot 10^{-16}, gtol = 2.3 \cdot 10^{-16}$ and $diff\_step = 1. \cdot 10^{-3}$.

For sampling noise applied to time series signal according to formula (3) the numpy seed, np.random.seed() function was initialized at the value 87654321. The seed was set at 112233, for sampling initial parameter guesses, $\theta_{sampled}$, as prescribed by formula (13).