



# Parameter subset selection techniques for problems in mathematical biology

Christian Haargaard Olsen<sup>1</sup> · Johnny T. Ottesen<sup>2</sup> · Ralph C. Smith<sup>1</sup> · Mette S. Olufsen<sup>1</sup>

Received: 8 February 2018 / Accepted: 6 October 2018 / Published online: 30 October 2018  
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

## Abstract

Patient-specific models for diagnostics and treatment planning require reliable parameter estimation and model predictions. Mathematical models of physiological systems are often formulated as systems of nonlinear ordinary differential equations with many parameters and few options for measuring all state variables. Consequently, it can be difficult to determine which parameters can reliably be estimated from available data. This investigation highlights pitfalls associated with practical parameter identifiability and subset selection. The latter refer to the process associated with selecting a subset of parameters that can be identified uniquely by parameter estimation protocols. The methods will be demonstrated using five examples of increasing complexity, as well as with patient-specific model predicting arterial blood pressure. This study demonstrates that methods based on local sensitivities are preferable in terms of computational cost and model fit when good initial parameter values are available, but that global methods should be considered when initial parameter value is not known or poorly understood. For global sensitivity analysis, Morris screening provides results in terms of parameter sensitivity ranking at a much lower computational cost.

**Keywords** Parameter identifiability · Parameter subset selection · Parameter estimation · Modeling

## 1 Introduction

The use of mathematical models allows for prediction of hidden dynamics, making it possible to test hypotheses in silico. It is desirable that models be as simple as possible, but not so simple that they fail to predict relevant dynamics. Ideally, models should not contain more parameters than required, and it should be possible to determine parameters uniquely from available data.

For an input–output relation  $y = f(\theta)$ , a set of inputs, e.g. parameters, initial or boundary conditions,  $\theta = [\theta_1, \dots, \theta_p]$ ,

are said to be identifiable at  $\theta^*$  if  $f(\theta) = f(\theta^*)$  implies that  $\theta = \theta^*$  for all admissible  $\theta$  [9,45]. Alternatively, if multiple parameter configurations yield the same output, the parameters are said to be *unidentifiable* [4,26].

Parameters that minimally affect outputs when varied across the admissible space are broadly termed *noninfluential*. Specifically, parameters  $\theta$  are defined as noninfluential on the space  $NI(\theta)$  if  $|f(\theta) - f(\theta^*)| < \varepsilon$  for all  $\{\theta, \theta^*\} \in NI(\theta)$ . The space of *influential* parameters  $I(\theta)$  is defined as the orthogonal complement of  $NI(\theta)$  [38]. The challenge of identifiability associated with noninfluential parameters may be enhanced when large parts of the system modeled is hidden for observation [4,5,31].

Parameters that do not uniquely influence the model output may stem from the model structure regardless of the parameter configuration and input [6,26], but may also arise due to restrictions in parameter values, model inputs, or outputs [16]. Following the definition from Bellman and Åström [6], parameters that are unidentifiable due to the model structure are referred to as *structurally unidentifiable*, whereas parameters that are unidentifiable due to practical restrictions, such as availability of data or physical restrictions, are referred to as *practically unidentifiable* [16].

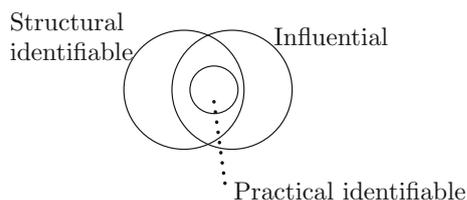
Communicated by Peter J. Thomas.

This article belongs to the Special Issue on *Control Theory in Biology and Medicine*. It derived from a workshop at the Mathematical Biosciences Institute, Ohio State University, Columbus, OH, USA.

✉ Mette S. Olufsen  
msolufse@ncsu.edu

<sup>1</sup> Department of Mathematics, NC State University, Raleigh, NC 27695, USA

<sup>2</sup> Department of Science and Environment, Roskilde University, 4000 Roskilde, Denmark



**Fig. 1** Relationship between influential, structural, and practical identifiable parameters. Note that an influential and structurally identifiable parameter may not necessarily be practically identifiable, whereas a practically identifiable parameter is both influential and structurally identifiable

Many complex models rely on numerical methods for finding solutions, which in itself brings about challenges associated with identifiability of the solution due to numerical precision, problem discretization, and how these may change with model parametrization. For such problems, it is interesting to study how identifiability challenges arises both from the structure of the model, but also from the limitations brought about by parametrization, data, model implementation, etc. This study focuses on methods for practical identifiability, which are based on numerical methods, rather than structural identifiability, which is often based on an algebraic approach. We refer readers to studies by others, e.g. Eisenberg et al. [11], Mahdi et al. [23], and Transtrum et al. [42] for works addressing structural identifiability from an algebraic perspective.

To illustrate challenges associated with identifiability, this investigation reviews a number of methods for determining which parameters have no effect on the model output and sets of parameters that may have a redundant effect on the model output. Methods assessing how variations (uncertainty) in parameters affect the model output are referred to as *sensitivity analysis* (SA), whereas methods for investigating interactions between parameters are referred to as *identifiability analysis* (IA). Figure 1 illustrates the relationship between the properties: Influential (sensitivity), structural, and practical identifiable. While it is rather obvious that not all influential parameters are structurally identifiable, it is less obvious why parameters can be practical but not structural identifiable. One example is highly—but not perfectly—correlated parameters. They can be impossible to estimate in practice with real data, while the algebraic structure of the model denotes the parameters identifiable. On the other hand practical identifiability requires that the parameters be both structurally identifiable and influential.

This study presents three methods for quantifying sensitivities: *derivative-based (local) sensitivities*, *Sobol indices* [39], and *Morris elementary effects* [27]. In addition, the *structural correlation method* (SCM) [9,28] and *orthogonal sensitivities method* (OSM) [22] are included to demonstrate practical identifiability and *subset selection*. These methods use statistical identifiability methods to perform subset selec-

tion and compare model fit for subsets recovered using each method. To investigate the efficiency of these methods, we apply them to a variety of examples of increasing complexity and with different identifiability challenges. The first example illustrates structurally unidentifiable parameters, the second discusses identifiability problems materializing due to sparse data, and the third, practical problems arising in a simple viscoelastic model with structurally identifiable parameters. Finally, the proposed methods are applied to analyze a more complex model predicting systemic arterial blood pressure in humans.

## 2 Methods

We present concepts and methods in the context of a system of ordinary differential equations defined by

$$\begin{aligned}\dot{x}(t) &= f(t, x(t), u(t), \theta) \\ y(t_i) &= h(x(t_i), u(t_i), \theta),\end{aligned}\quad (1)$$

where  $x$  is the model state vector,  $t$  is the independent variable,  $u(t)$  is the model input,  $\theta = \{\theta_1, \theta_2, \dots, \theta_p\}$  is the parameter vector, and  $h(x, u, \theta)$  is the observation function that relates the model states to the observable model output  $y(t_i)$  at  $t = t_i$ .

### 2.1 Sensitivity analysis

Local sensitivity analysis refers to analysis that quantifies how the model output changes with changes in parameter values. Global sensitivity analysis more broadly quantifies how variability or uncertainties in responses can be apportioned to variability or certainties in parameters considered throughout the admissible parameter space. We discuss three different methods. *Derivative-based (local) sensitivities* quantify the local influence of each parameter on the model output computed from the partial derivative of the model output with respect to each model parameter. This method is local in the sense that it is evaluated at a nominal set of parameter values. In addition, we present two global methods computing sensitivities over a parameter space: Sobol indices [39] and Morris elementary effects [27]. All methods presented here are chosen to study practical identifiability analysis and rely on numerical analysis of the model response. These are chosen as they reflect problems arising in analysis of real models composed of highly nonlinear systems, where numerical methods are required to obtain solutions. The response displayed by such models are greatly dependent on the parameter values. Therefore, practical methods for analyzing identifiability will also greatly depend on reasonable initial guess for parameter values and ranges. Given that models typically are built using physical principles we expect

that reasonable parameter range can be obtained from a priori knowledge of system behavior, such as steady-state or mean values.

### 2.1.1 Derivative-based (local) sensitivities

Local derivative-based sensitivities  $S_{\theta_i}^y(t)$  can be computed by differentiating the model output  $y(t)$  with respect to the parameters  $\theta_i$

$$S_{\theta_i}^y(t, \theta) = \frac{\partial y(t, \theta)}{\partial \theta_i}. \tag{2}$$

For some problems, these can be calculated analytically, but for more complex problems it is advantageous to numerically approximate local sensitivities. Local sensitivities are typically computed using *sensitivity equations, finite differences* or *automatic differentiation* [30,38].

The sensitivity equations method utilizes the fact that the sensitivities of the model states to parameter values  $S_{\theta_i}^x = \frac{\partial x}{\partial \theta_i}$  can be computed as solutions to the ODE

$$\begin{aligned} \frac{d}{dt} \left( \frac{\partial x}{\partial \theta_i} \right) &= \frac{df}{d\theta_i} + \frac{df}{dx} \frac{\partial x}{\partial \theta_i} \\ \frac{dS_{\theta_i}^x}{dt} &= J_{\theta_i} + J_x S_{\theta_i}^x, \end{aligned} \tag{3}$$

obtained by differentiating the model  $\frac{dx}{dt} = f(t, x, \theta)$  with respect to the parameters  $\theta_i$ . Here  $J_{\theta} = \frac{\partial f}{\partial \theta}$  denotes the *parameter Jacobian*, and  $J_x = \frac{\partial f}{\partial x}$  the *model Jacobian*, which can be calculated analytically. Subsequently, the sensitivities of the model output  $S_{\theta_i}^y$  can be calculated by application of the chain rule on the output function  $y = h(x, t, \theta)$ .

Alternatively, the derivative with respect to  $\theta_i$  can be approximated using the finite-difference scheme

$$\hat{S}_{\theta_i}^y(t) = \frac{y(t, \theta + h e_i) - y(t, \theta)}{h},$$

Here the step size  $h$  is chosen to reflect the precision of the model output, and  $e_i$  is the unit vector in the  $i$ 'th component direction. If the error in the model evaluation is on the order of  $\varepsilon$ , the step size should be  $h = \sqrt{\varepsilon}$  to obtain an error of same magnitude in the sensitivities [29]. Alternatively, one can avoid underflow due to subtraction by employing complex sums [25].

Finally, *automatic differentiation* (AD) is a third option that uses the chain rule while solving the differential equations by extending the basic numerical operations to include calculation of the derivative through recursive and automatic application of the chain rule [14]. AD routines have been developed for variety of programming languages, such as C/C++, Fortran, MATLAB and Python [38], and are now

included in optimization packages and modeling frameworks such as ACADO [17] and CasADi [3]. Accessibility of AD has improved tremendously in recent years and offers a tool for differentiation that ensures precision in the derivatives without the hassle of adjusting the step sizes or calculating the model Jacobian analytically [13].

### 2.1.2 Sobol sensitivity indices

This method was originally developed by Sobol [41], but several variations exist [33]. The description of the Sobol sensitivity indices and the algorithm for calculating first-order effects follows that of Smith [38], but with modifications to reflect the algorithm presented by Jansen [20].

Consider a scalar-valued nonlinear model

$$Y = f(\Theta) \tag{4}$$

with parameters represented by the vector of random variables  $\Theta = [\Theta_1, \dots, \Theta_p]$ . Here  $\Theta_i \in \Gamma_i \subset \mathbb{R}$ , where  $\Gamma_i$  is the domain of the random variable  $\Theta_i$ . The joint probability density function is defined as  $\rho_{\Theta}(\theta) = \prod_{k=1}^p \rho_{\Theta_k}(\theta_k)$ , where  $\rho_{\Theta_k}(\theta_k)$  denotes the marginal probability density function. For mutual independent variables, the joint distribution is simply the product of the marginal densities.

To avoid scaling issues, we transform parameters to  $\Gamma_k = [0, 1]$ . We assume that they are uniformly distributed to avoid inadvertent biases [10].

Consider the model decomposition

$$f(\theta) = \sum_{\mathbf{i} \subseteq \{1,2,\dots,p\}} f_{\mathbf{i}}(\theta_{\mathbf{i}}), \tag{5}$$

where  $f$  is the model output,  $\mathbf{i} = \{i_1, \dots, i_s\}$  is the parameter set with cardinality  $s$ ,  $\theta_{\mathbf{i}} = \{\theta_{i_1}, \dots, \theta_{i_s}\}$  and  $f_{\emptyset} = f_0$  [38, 40]. If

$$\int_0^1 f_{\mathbf{i}}(\theta_{\mathbf{i}}) \rho_k(\theta_k) d\theta_k = 0,$$

for any parameter  $\theta_k$  and  $\mathbf{i} = \{1, \dots, s\}$  that includes  $k$ , the decomposition is unique and the component functions are given by

$$f_{\mathbf{i}}(\theta_{\mathbf{i}}) = \int_{[0,1]^{p-s}} f(\theta) \rho_{\theta}(\theta_{\sim \mathbf{i}}) d\theta_{\sim \mathbf{i}} - \sum_{\substack{l' \subset \mathbf{i} \\ l' \neq \mathbf{i}}} f_{l'}(\theta_{l'}).$$

To determine the effect of the parameter interaction in the set  $\mathbf{i}$ , consider the conditional partial variance

$$D_{\mathbf{i}} = \int_{[0,1]^s} f_{\mathbf{i}}^2(\theta_{\mathbf{i}}) \rho_{\theta}(\theta_{\mathbf{i}}) d\theta_{\mathbf{i}} = \text{var}[\mathbb{E}(Y|\theta_{\mathbf{i}})] - \sum_{\substack{l' \subset \mathbf{i} \\ l' \neq \emptyset}} D_{l'},$$

and the total variance

$$D = \int_0^1 f^2(\theta) \rho_{\theta}(\theta) d\theta - f_0^2 = \sum_{\substack{i \subseteq \{1,2,\dots,p\} \\ i \neq \emptyset}} D_i.$$

The Sobol indices are defined as

$$S_i = \frac{D_i}{D} \text{ giving } \sum_{\substack{i \subseteq \{1,2,\dots,p\} \\ i \neq \emptyset}} S_i = 1. \tag{6}$$

The total sensitivity index

$$S_{T_i} \equiv \sum_{k, i \in k} S_k \tag{7}$$

describes the sensitivity of the model output variance to the parameter  $k$ , including the effects of interactions with other parameters.

A Monte Carlo approach for estimating the first-order indices  $S_i$  (6) and total indices  $S_{T_i}$  (7) requires  $M^2$  model evaluations for each parameter, where  $M$  is the number of points in the parameter space used to evaluate the conditional mean  $\mathbb{E}(Y|\theta_i)$ ; see Smith [38]. Saltelli et al. [33] proposed an algorithm that reduces the number of model evaluations required to  $M(p+2)$ , whereas Algorithm 1, used in our study, is due to Jansen [20] and requires  $2M(p+2)$  evaluations for a better trade-off between precision and least squares error. It should be noted that these algorithms truncate the variance decomposition after second order. If higher-order interactions are present in the model, these algorithms may yield incorrect estimates.

### 2.1.3 Morris elementary effects

Morris screening computes the sensitivity of the model output to each parameter  $\theta_i$  by sampling the *elementary effects* [27,38]

$$d_i(\Theta) = \frac{f(\Theta + \Delta e_i) - f(\Theta)}{\Delta},$$

from a grid with points  $\Gamma_\ell = \{0, \frac{1}{\ell-1}, \dots, 1\}$  in the parameter distribution range  $\Gamma = \{0, 1\}$ . The number of points (in each dimension) of the grid and their value  $\Delta$  is important as they determine the sampling of the parameter space. By choosing  $\ell$  to be even and  $\Delta = \frac{\ell}{2(\ell-1)}$ , it is possible to ensure an even probability sampling from  $\Gamma_\ell$  that guarantees equal probability for all points in the grid [38]. Similar to the Sobol indices  $f$  denote the scalar-valued nonlinear model given in (4).

### Algorithm 1 Sobol indices.

Using  $2M$  model evaluations for each of the  $p$  parameters:

1. Create two  $M \times p$  sample matrices

$$A = \begin{bmatrix} \theta_1^1 & \dots & \theta_i^1 & \dots & \theta_p^1 \\ \vdots & & & & \vdots \\ \theta_1^M & \dots & \theta_i^M & \dots & \theta_p^M \end{bmatrix},$$

$$B = \begin{bmatrix} \hat{\theta}_1^1 & \dots & \hat{\theta}_i^1 & \dots & \hat{\theta}_p^1 \\ \vdots & & & & \vdots \\ \hat{\theta}_1^M & \dots & \hat{\theta}_i^M & \dots & \hat{\theta}_p^M \end{bmatrix},$$

where  $\theta_i^j$  and  $\hat{\theta}_i^j$  are parameter samples drawn from the distribution corresponding to parameter  $i$ , for each sampled parameter configuration  $j = 1, \dots, M$ . In addition, create the matrix

$$C = [A \ B]^T,$$

where  $T$  denotes transpose.

2. Create two  $M \times p$  matrices

$$A_B^i = \begin{bmatrix} \theta_1^1 & \dots & \hat{\theta}_i^1 & \dots & \theta_p^1 \\ \vdots & & & & \vdots \\ \theta_1^M & \dots & \hat{\theta}_i^M & \dots & \theta_p^M \end{bmatrix},$$

$$B_A^i = \begin{bmatrix} \hat{\theta}_1^1 & \dots & \theta_i^1 & \dots & \hat{\theta}_p^1 \\ \vdots & & & & \vdots \\ \hat{\theta}_1^M & \dots & \theta_i^M & \dots & \hat{\theta}_p^M \end{bmatrix},$$

3. Compute the model output vectors

$$y_A = f(A), \quad y_B = f(B), \quad y_{A_B^i} = f(A_B^i), \quad y_{B_A^i} = f(B_A^i),$$

noting that

$$y_C = [y_A \ y_B]^T.$$

4. Estimates of the first-order sensitivity indices can be computed by

$$S_i^* = \frac{\frac{1}{M} \sum_{j=1}^M y_A^j y_{B_A^i}^j - y_A^j y_B^j}{\frac{1}{2M} \sum_{j=1}^M y_C^j y_C^j - \left(\frac{1}{M} \sum_{j=1}^M y_C^j\right)^2}$$

$$= \frac{\frac{1}{M} (y_A)^T (y_{B_A^i} - y_B)}{\frac{1}{2M} (y_C)^T y_C - \left(\frac{1}{M} \sum_{j=1}^M y_C^j\right)^2},$$

$$S_{T_i}^* = \frac{\frac{1}{M} \sum_{j=1}^M \left(y_A^j y_{A_B^i}^j\right)^2}{\frac{1}{2M} \sum_{j=1}^M y_C^j y_C^j - \left(\frac{1}{M} \sum_{j=1}^M y_C^j\right)^2}$$

$$= \frac{\frac{1}{M} (y_A - y_{A_B^i})^T (y_A - y_{A_B^i})}{\frac{1}{2M} (y_C)^T y_C - \left(\frac{1}{M} \sum_{j=1}^M y_C^j\right)^2}.$$

By estimating the mean and variance of the distribution  $|d_i(\theta)|$ , for which

$$u_i^* = \frac{1}{r} \sum_{j=1}^r \left| d_i^j(\theta) \right|,$$

$$\sigma_i^{2*} = \frac{1}{r-1} \sum_{j=1}^r \left( d_i^j(\theta) - \mu_i \right)^2,$$

$$\mu_i = \frac{1}{r} \sum_{j=1}^r d_i^j(\theta)$$

we obtain an estimate of the sensitivity noting large local changes attributed to parameter  $\theta_i$ .

The implementation presented in Algorithm 2 creates a trajectory through the grid in the parameter space, such that for each step in the trajectory, just one parameter is changed. Creating such a trajectory one *elementary effect* sampling can be obtained for each of the  $p$  parameters using  $p + 1$  function evaluations [27,38].

It should be noted that both the Morris- and Sobol- methods are based on the assumption of mutually independent parameters and can yield inaccurate results for correlated parameters. Whereas extensions to Sobol theory exist for correlated parameters, they require knowledge of the underlying distribution and can yield negative indices that are difficult to interpret.

### 2.2 Parameter interactions

We employ two methods for constructing parameter subsets with no parameter interactions using the structural correlation method (SCM) and the orthogonal sensitivities method (OSM). These methods are investigated. Both methods are based on the linearization of model output, which can be obtained without significant computational cost. Other methods, including the profile likelihood approach by Raue et al.[31], can provide more insight into parameter interactions, but require a significantly higher computational cost. Both the SCM and the OSM methods are based on analysis of the Fisher information matrix ( $F$ ) given by

$$F = S^T S, \tag{8}$$

where  $S$  is the sensitivity matrix defined in (2), evaluated at a nominal parameter value.

For least-squares optimization problems, the goal is to determine a parameter vector  $\theta^*$  that minimizes the objective function

$$SS = \frac{1}{2} \sum_{i=0}^{n-1} r_i^2, \quad r_i = v_i - y_i, \tag{9}$$

where  $r_i$  are the residuals, i.e.  $v_i$  the measured data and  $y_i = f(t_i)$  the  $i$ th output of the function. For this optimum, the

---

### Algorithm 2 Morris Indices

---

For each  $j = 1, \dots, r$ :

1. Sample an initial parameter vector  $\theta^*$  from the uniform distribution  $\mathcal{U} [0, 1]^p$ .
2. Create a  $(p + 1) \times p$  permutation matrix  $B^*$  where each row represents a parameter configuration. A completely deterministic permutation matrix is given by

$$B_j^* = J_{p+1,p} \theta^* + \Delta B,$$

where  $J_{p+1,p}$  is a  $(p + 1) \times p$  matrix of ones,  $\Delta$  is the chosen grid step size, and  $B$  is a lower triangular matrix of ones. Each row in  $B^*$  differs in exactly one column, corresponding to the parameter for which the step is made. A random permutation matrix can be obtained by applying the orientation matrix

$$B_j^* = \left( J_{p+1,p} \theta^* + \frac{\Delta}{2} [(2B - J_{p+1,p}) D^* + J_{p+1,p}] \right) P^*,$$

where the  $p \times p$  matrix  $D^*$  is diagonal with elements randomly chosen from the set  $\{-1, 1\}$ , and the  $p \times p$  matrix  $P^*$  is created by perturbing the columns of a  $p \times p$  identity matrix.

3. Evaluate the model output for each of these  $p + 1$  configurations, in the  $(p + 1) \times 1$  vector

$$y_j = f(B_j^*).$$

4. Through analysis of  $B^*$  determine which two consecutive rows  $k$  and  $k + 1$  that corresponds to each parameter  $i$ , and calculate the corresponding elementary effect,

$$d_i^j = \frac{y_{j,k+1} - y_{j,k}}{\Delta}.$$

Finally, the sampling mean  $\mu_i^*$  and variance  $\sigma_i^{2*}$  for each parameter  $i$  can be estimated by

$$u_i^* = \frac{1}{r} \sum_{j=1}^r \left| d_i^j(\theta) \right|,$$

$$\sigma_i^{2*} = \frac{1}{r-1} \sum_{j=1}^r \left( d_i^j(\theta) - \mu_i \right)^2,$$

$$\mu_i = \frac{1}{r} \sum_{j=1}^r d_i^j(\theta).$$


---

gradient vector is zero, i.e.  $g(\theta^*) = 0$ . The second-order Taylor expansion of the objective function at the optimum point  $\theta^*$  in the parameter space is given by

$$SS(\theta) = SS(\theta^*) + \Delta^T g + \Delta^T H \Delta + \mathcal{O}(\Delta^3), \tag{10}$$

where  $\Delta = (\theta - \theta^*)$ ,  $g$  is the gradient

$$g_j = \frac{\partial SS}{\partial \theta_j} = \sum_{i=0}^{n-1} r_i \frac{\partial r_i}{\partial \theta_j}$$

and  $H$  is the Hessian—a symmetric matrix with second-order partial derivatives—of the objective function with entries

$$H_{jk} = \frac{\partial g_j}{\partial \theta_k} = \sum_{i=0}^{n-1} \frac{\partial r_i}{\partial \theta_k} \frac{\partial r_i}{\partial \theta_j} + r_i \frac{\partial^2 r_i}{\partial \theta_j \partial \theta_k}.$$

Ignoring higher-order terms of the Taylor expansion in (10), and re-arranging terms yields

$$SS(\theta) = SS(\theta^*) + \Delta^T H \Delta$$

since  $g(\theta^*) \approx 0$ . If  $H$  is singular, there exists some  $\Delta \neq 0$ , such that  $H\Delta = \mathbf{0}$ , which in turn means that  $SS(\theta) = SS(\theta^*)$ , and thereby, that the minimizer of the objective function is not unique and some parameters are unidentifiable.

When  $r$  is small or when  $r$  varies linearly near a minimum of (9), a reasonable approximation to the Hessian is obtained by neglecting the second-order term [47]

$$H_{jk} \approx \sum_{i=0}^{n-1} \frac{\partial r_i}{\partial \theta_k} \frac{\partial r_i}{\partial \theta_j}.$$

Since  $r_i = v_i - y_i$  the derivative of the residual is given by

$$\frac{\partial r_i}{\partial \theta_j} = -\frac{\partial y_i}{\partial \theta_j} = -S_{ij}$$

and the Hessian approximation can be formulated as matrix product

$$H \approx S^T S = F$$

and the FIM can be used in place of the Hessian for local identifiability analysis. Whereas the previous presentation is formulated using the absolute sensitivities, similar arguments can be made for relative sensitivities

$$\tilde{S}_{ij} = \frac{\partial \ln y_i}{\partial \ln \theta_j}.$$

Defining  $\tilde{y}_i = \ln y_i$  and  $\tilde{\theta}_j = \ln \theta_j$ , the Hessian of the relative objective function is then formulated as

$$\tilde{H} = \tilde{S}^T \tilde{S},$$

which can be approximated by the FIM using the relative sensitivities.

As noted previously, parameter identifiability requires that the Hessian  $H \approx S^T S$  is non-singular. An immediate consequence of this is that the columns of  $S$  need to be nonzero and cannot be linearly dependent.

## 2.2.1 Structural correlation method

Correlations between parameters can be used to describe how the parameter values depends on each other when fitting experimental data from the same system. With a series of estimates for parameters  $a$  and  $b$ , the covariance and correlation are given by

$$\begin{aligned} \text{Cov}[a, b] &= E[(a - E[a])(b - E[b])], \\ \text{cor}[a, b] &= \frac{\text{Cov}[a, b]}{\sqrt{\text{Cov}[a, a]\text{Cov}[b, b]}}. \end{aligned}$$

While this makes it clear how to understand what correlations are, the computational cost associated with this method is high.

For problems with constant measurement variance  $\sigma$ , the parameter covariance is the inverse of the FIM [given in (8)], i.e.

$$C = F^{-1},$$

which enables the calculation of correlations by

$$c_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}.$$

In the structural correlation method, a pair of parameters with a large correlation—their uncertainty is strongly coupled—cannot both be estimated [19]. For practical applications, a correlation threshold  $\gamma < 1$  is set to determine pairs of parameters that are correlated. For each pairwise correlated parameters, the least sensitive is identified as described in Algorithm 3.

## 2.2.2 Orthogonal sensitivities

The orthogonal sensitivities method developed by Li et al. [22] combines two of the methods presented by Miao et al. [26]: the principal component analysis (PCA) and the orthogonal method. Whereas the structural correlation method is based on removing pairwise correlation, the orthogonal sensitivities method constructs a subset by adding one parameter at the time. For each parameter, an importance index  $e$  is calculated using PCA and each time the subset of already selected parameters is updated, we compute the orthogonality index  $d$  for the remaining parameters. At each iteration the parameter with the highest product  $I = ed$  is added to the subset [22].

**Algorithm 3** Structural Correlation Method

1. For each parameter, compute the norm of the sensitivity of the model output to this parameter. We choose the 2-norm

$$\tilde{s}_j = \sqrt{\sum_{i=0}^{N-1} S_j^2(t_i)}.$$

2. Fix insensitive parameters for which  $\tilde{s}_j < \sqrt{\varepsilon}$ , where  $\varepsilon$  is the precision used when calculating the sensitivities.
3. Repeat the following steps.
  - (a) Calculate the covariance matrix by inverting the FIM,  $C = (F)^{-1} = (S^T S)^{-1}$ , and determine correlations as described in Sect. 2.2.1. If the largest correlation is smaller than some value  $\gamma$ , the subset reduction procedure is complete. If the matrix is invertible, some parameters are perfectly correlated, and analysis can only be continued when one of these fixed.
  - (b) For the correlation with the largest absolute value, fix the parameter with the lowest overall sensitivity, determined by the norm of the sensitivities  $\tilde{s}_j$ .
  - (c) The parameter to be fixed is noted, and the column of  $S$  corresponding to this parameter are removed from  $S$ , before correlations are calculated again, and the process is repeated until the stopping criteria in (a) is satisfied.

**3 Results**

In this section, we illustrate challenges associated with the various identifiability concepts using the algorithms summarized in Sect. 2. This will be accomplished using four examples illustrating different causes of unidentifiability: problem structure, availability of data, and a mixture of both. Note that for the global sensitivity methods, the parameter values are mapped from [0, 1] to relevant intervals when the model is evaluated.

**3.1 Structural unidentifiability**

Consider a distensible blood vessel with a volume given by  $x$ . Blood flow into the vessel  $q_{in} = bu(t)$  is proportional to  $u(t)$  with proportionality  $b$ , while the flow out is proportional to the volume  $q_{out} = -ax$ . Here  $a$  is the constant of proportionality. Assume data  $y$  is proportional to the blood volume,  $y = cx$ , and treat the initial blood volume as a parameter  $x_0$ . For this system, the change in blood volume  $x$  can be quantified by the differential equation

$$\begin{aligned} \dot{x}(x, t) &= -ax + bu(t), \quad x(0) = x_0 \\ y(x) &= cx, \end{aligned} \tag{13}$$

where  $a, b, c$  and  $x_0$  are model parameters. The model output is taken to be

$$y(t) = cx(t) = cx_0 e^{-at} + cb \int_0^t e^{-a(t-s)} u(s) ds. \tag{14}$$

Note that the parameter  $c$  only appears in combination either with the initial condition parameter  $x_0$  or with the parameter

**Algorithm 4** Orthogonal Sensitivities Method

1. For each parameter  $j$  calculate the Importance Index

$$e_j = \frac{\sum_{i=1}^m \lambda_i Q_{ji}}{\sum_{i=1}^m \lambda_i}, \tag{11}$$

where  $\lambda_i$  is the  $i$ 'th eigenvalue of the matrix  $F = \tilde{S}^T \tilde{S}$  with the corresponding eigenvector being the  $i$ 'th column of the matrix  $Q$ .

2. Select the parameter with the highest value of  $e$  as the first parameter,  $\theta_{k_1}$ .
3. Repeat the following steps until no more parameters can be added to the subset.

- (a) Calculate the projection of the sensitivity vector  $\tilde{S}_{\theta_{k+1}}$  for each parameter  $j$  that has not been selected, onto the space spanned by the sensitivities of the  $k$  already selected parameters  $\theta_{\ell_1}, \dots, \theta_{\ell_k}$ ,  $\tilde{S}_s = \text{Span}(\tilde{S}_{\theta_{\ell_1}}, \dots, \tilde{S}_{\theta_{\ell_k}})$ ,

$$\tilde{s}_j = \sum_{i=\theta_{\ell_1}}^{\theta_{\ell_k}} \alpha_i \tilde{S}_i.$$

This is accomplished by finding the coefficients  $\alpha_i$  that solves

$$\min_{\alpha_i} \frac{1}{2} (\tilde{S}_j - \tilde{s}_j)^T (\tilde{S}_j - \tilde{s}_j),$$

which is equivalent to solving the linear system

$$\begin{bmatrix} \tilde{S}_{\theta_{\ell_1}}^T \tilde{S}_{\theta_{\ell_1}} & \dots & \tilde{S}_{\theta_{\ell_k}}^T \tilde{S}_{\theta_{\ell_1}} \\ \vdots & \ddots & \vdots \\ \tilde{S}_{\theta_{\ell_1}}^T \tilde{S}_{\theta_{\ell_k}} & \dots & \tilde{S}_{\theta_{\ell_k}}^T \tilde{S}_{\theta_{\ell_k}} \end{bmatrix} \alpha = \begin{bmatrix} \tilde{S}_j^T \tilde{S}_{\theta_{\ell_1}} \\ \vdots \\ \tilde{S}_j^T \tilde{S}_{\theta_{\ell_k}} \end{bmatrix}, \tag{12}$$

as long as the columns corresponding to the already selected parameters are linearly independent.

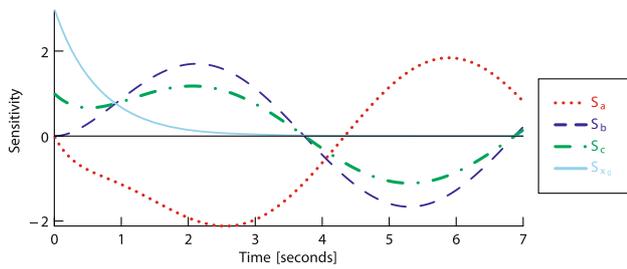
- (b) The linear independence of each parameter  $j$  is determined by considering the sine of the angle between the projection  $\tilde{s}$  and  $\tilde{S}_j$ . This yields the *orthogonality index*,

$$d_j = \sin \left[ \arccos \left( \frac{\tilde{S}_j^T \tilde{s}}{\|\tilde{S}_j\| \|\tilde{s}\|} \right) \right].$$

- (c) Select the parameter  $j$  that has the highest *Identifiability Index*,

$$I = e_j d_j,$$

Note: Relative sensitivities were used in [22] but absolute sensitivities may be used as well.



**Fig. 2** Sensitivities expressed in (15) for the linear input example given in (13, 14) (color figure online)

*b*. Hence, *c*, *b* and *x*<sub>0</sub> are not mutually identifiable since they are not uniquely determined by the data.

The local sensitivities can be calculated analytically by way of the ODE system

$$\begin{aligned} \frac{\partial y}{\partial a}(t, \theta) &= -ctx_0e^{-at} - cb \int_0^t (t-s)e^{-a(t-s)}u(s)ds, \\ \frac{\partial y}{\partial b}(t, \theta) &= c \int_0^t e^{-a(t-s)}u(s)ds, \\ \frac{\partial y}{\partial c}(t, \theta) &= x_0e^{-at} + b \int_0^t e^{-a(t-s)}u(s)ds, \\ \frac{\partial y}{\partial x_0}(t, \theta) &= ce^{-at}. \end{aligned} \tag{15}$$

Using the input function  $u(x) = \sin(x)$ , parameters  $a = 1.5, b = 2, c = 3$  and  $x_0 = 1$  for  $t_i \in [0, 7]$  yield the sensitivities shown in Fig. 2.

When all parameters are included, or when only *a* is excluded, the matrix  $S^T S$  is singular due to linear dependence between the sensitivities to the parameters *b*, *c* and *x*<sub>0</sub>. This is indicated by condition numbers on the order of  $10^{16}$ , whereas fixing of *b*, *c* or *x*<sub>0</sub> reduces the condition number to  $10^1$ . We note that this matches well with the seemingly linear dependence of the sensitivities shown in Fig. 2.

Using the orthogonal sensitivities (OSM) in Algorithm 4, we obtained the ranking compiled in Table 1. We note that:

1. *a* is ranked as most influential. The sensitivity plot in Fig. 2 confirms this, yet *b* and *c* are almost as influential.
2. *b* is ranked 2<sup>nd</sup>. The orthogonality of *b* is only 0.45 which is lower than that of *x*<sub>0</sub>, but *b* is chosen since the importance index is higher.
3. *x*<sub>0</sub> is ranked 3<sup>rd</sup>. The importance index of *x*<sub>0</sub> is lower (0.18) than that of *c* (0.39), which is understandable considering the magnitude of the sensitivities in Fig. 2. However, since the sensitivity to *x*<sub>0</sub> has a larger projection (0.97) to those already selected, *x*<sub>0</sub> is chosen over *c*.

**Table 1** (a) Ranking obtained using the orthogonal sensitivities method (OSM)

Rank	Parameter	<i>e</i>	<i>d</i>	<i>I</i>
(a) OSM				
1	<i>a</i>	0.66		
2	<i>b</i>	0.45	0.53	0.24
3	<i>x</i> <sub>0</sub>	0.18	0.97	0.17
4	<i>c</i>	0.39	0.00	0.00
Parameter	<i>S</i> <sub><i>i</i></sub> <sup>*</sup>	<i>S</i> <sub><i>T</i><sub><i>i</i></sub></sub> <sup>*</sup>	$\mu_i^*$	$\sigma_i^*$
(b) Global				
<i>a</i>	0.28	0.29	297.81	162.62
<i>b</i>	0.01	0.01	76.35	67.97
<i>c</i>	0.31	0.42	374.46	150.44
<i>x</i> <sub>0</sub>	0.24	0.32	322.08	120.32

Columns are importance (PCA) score (*e*), orthogonality (*d*) and total identifiability ( $I = ed$ ). (b) Sobol first order (*S*<sub>*i*</sub><sup>\*</sup>) and total effect (*S*<sub>*T*<sub>*i*</sub></sub><sup>\*</sup>), and Morris elementary effects mean and variance ( $\mu_i^*$  and  $\sigma_i^*$ )

4. *c* is ranked last. When *a*, *b*, *x*<sub>0</sub> have already been chosen, the orthogonality index of *c* is 0.00. This confirms the observation that the sensitivities are linearly dependent.

Since there is a linear dependence in *S*, the FIM (8) is not invertible. Hence, it is impossible to calculate statistical correlations of the parameters using the Cramer–Rao bound [26]. As suggested by the structural correlation method (SCM) one have to remove the least sensitive parameter if the sensitivity matrix is singular and so on until a non-singular sensitivity matrix is obtained. Removing one of the parameters *b*, *c* or *x*<sub>0</sub> the correlation matrix can be calculated. This results in off-diagonal elements with an absolute value significantly lower than 1, indicate that all remaining parameters are identifiable when *b*, *c* or *x*<sub>0</sub> is fixed. The problem with linear dependence could also be remedied by non-dimensionalization, here employed by scaling the model with respect to the initial condition, e.g. letting  $\tilde{x} = x/x_0, \tilde{x}_0 = 1$ , which in turn would give the output  $\tilde{y}y(t) = c\tilde{x}(t) = ce^{-at} + c\tilde{b} \int_0^t e^{-a(t-s)}u(s)ds$ , where  $\tilde{b} = b/x_0$ .

One important final remark is that fixing parameters reduces the degrees of freedom of the model and changes the model. For this example, the interaction between *b*, *c* and *x*<sub>0</sub> makes it impossible to determine which two parameters are identifiable. Since fixing one parameter effectively changes the model, it does not make sense to compare parameter values for *x*<sub>0</sub> in an optimization where *a* is fixed with another value obtained with *b* fixed. These two estimates of *x*<sub>0</sub> should not be expected to be similar as they are de facto parameters of different models. In other cases, it may be that noninfluential parameters are fixed. In contrast to cases with

parameter interactions, this does not introduce a bias, as the noninfluential parameters have no impact on model behavior, and therefore also no impact on parameter estimates.

For the global sensitivity analysis methods the parameters are sampled from a uniform distribution ranging from  $\frac{3}{4}$  to  $\frac{5}{4}$  times the nominal value used for the local sensitivities. The model is evaluated for the times  $t_i = i/100, i \in [0, 700]$ , and the output considered is the sum of squares,  $\sum_{i=0}^{700} y^2(t_i)$ . The total sensitivity index for the Sobol method and the elementary effects of the Morris method yields the same ordering of the parameters:  $c, x_0, a, b$ . These are different from the results obtained with the local methods, and also from what one would expect observing the local sensitivities in Fig. 2.

### 3.2 Practical unidentifiability

For practical unidentifiability, we consider two cases. The first case illustrates how parameters can be unidentifiable due to practical restrictions in model output measurements. This example shows how such restrictions can render a parameter noninfluential or fail to highlight differences in model effects. The second case is a model of a viscous material using a Voigt body model. While it is structurally identifiable, like the example in the first case, it is not clear whether the parameters are practically identifiable.

#### 3.2.1 Case A: Aliasing

Consider the function

$$f(t, \theta) = \sum_{k=1}^p \theta_k^2 f_k(t), \quad f_k(t) = \sin(2\pi kt). \quad (16)$$

If the function is restricted to a certain rate, one may experience *aliasing*—appearing when an oscillating signal is recorded at a rate that make the signal appear to be oscillating at a lower frequency or not at all. This example is a nonlinear version of the aliasing example presented by Smith [38].

Assume that the function  $f$  can be evaluated at  $n$  evenly spaced values  $t_i = i\Delta t, i = 0, \dots, n-1, \Delta t = \frac{1}{n-1}$ , and that the parameters are restricted to the interval  $(0, 1)$ . For example, for  $n-1 = 4$  the function can be observed/evaluated at the points  $t = 0, 1/4, 1/2, 3/4, 1$ . For  $p = 4$  and  $n-1 = \{4, 8, 16, 256\}$  the results are shown in Fig. 3a–d. Panel (a) shows that for  $n-1 = 4$ , there are only two nonzero basis functions and hence it is impossible to estimate all four parameters. Furthermore, the basis functions,  $f_1$  and  $f_3$  are linear dependent, implying that all possible model outputs can be represented by either of these functions.

In (b) three basis function are nonzero, but no basis functions are linearly dependent. Hence, the parameters determining the weight of the nonzero basis functions are

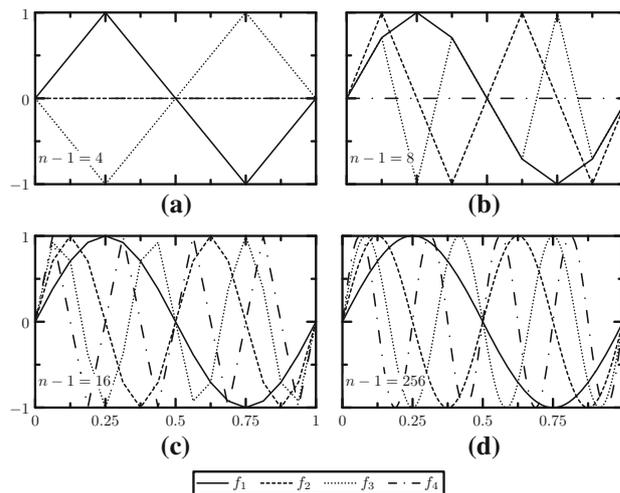


Fig. 3 Basis functions for the aliasing Example (16) evaluated at  $t = 0, 1/n, 2/n, \dots, n-1/n, 1$  (plotted with linear interpolation between the points) for the three measurement resolutions and full time series

identifiable. For (c, d) all basis functions are nonzero and linearly independent, implying that all parameters are identifiable. The configurations corresponding to panel (a–c) will be analyzed using the methods in Sect. 2.2.1.

The local sensitivities can be calculated analytically. They are given by

$$\frac{\partial f}{\partial \theta_k}(t, \theta) = 2\theta_k f_k(t). \quad (17)$$

Since it has already been shown that the basis functions are linearly dependent if evaluated at  $n-1 = 4, 8$ , the sensitivities will also be linearly dependent. Therefore, it is known in advance that it is not possible to estimate all parameters uniquely for these configurations. Since the basis functions are linearly independent for  $n-1 = 16$  so are the sensitivities, and all parameters are identifiable.

**Structural correlation method** Since there is a linear dependence between the four basis functions (two of them being zero), the FIM  $F = S^T S$  is not invertible, and therefore a linear combination of columns is trivially zero. Calculating the condition numbers, the  $F = S^T S$ -matrix for different values of  $n$  reveals that there is linear dependence in the sensitivities for  $n-1 = 4, 8$ , but not for  $n-1 = 16$ . As the FIM is invertible for  $n-1 = 16$ , correlations can be calculated. This results in the correlation matrix with all zero entries indicating that no parameters are correlated.

**Orthogonal sensitivities method** Table 2 shows the rankings obtained using the orthogonal sensitivities method for  $n-1 = 4, 8, 16$ . For  $n-1 = 4$  only one parameter is found

to be identifiable, whereas three are found identifiable for  $n - 1 = 8$ , and all parameters are identifiable for  $n - 1 = 16$ .

For the global sensitivity analysis methods we assume that parameters are independent and uniformly distributed,

$$\theta_i \sim \mathcal{U}(0, 1).$$

From the local analysis and arguments determining the dimension of the spanned space, it is expected that parameters  $\theta_2$  and  $\theta_4$  are noninfluential for  $n - 1 = 4$ , that  $\theta_4$  is noninfluential for  $n - 1 = 8$ , and that all parameters are influential for  $n - 1 = 16$ . In addition, it is expected that the effect of all influential parameters on the model output is similar in magnitude.

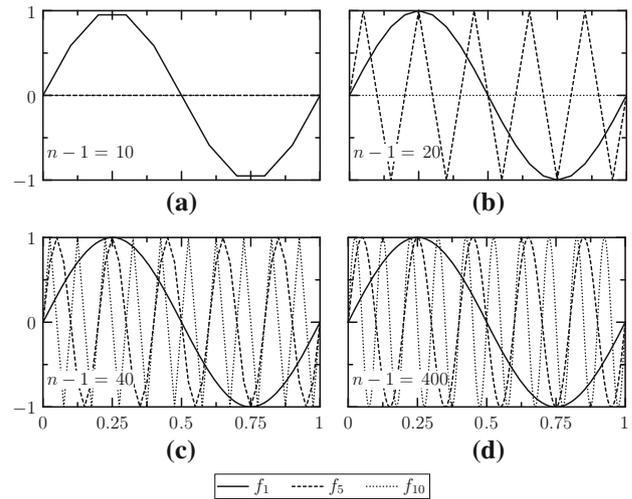
**Sobol indices** The estimated first-order Sobol indices  $S_i^*$  and total indices  $S_{T_i}^*$  are listed in Table 2. As expected, for  $n - 1 = 4$  only  $\theta_1$  and  $\theta_3$  have an effect on the model output, while  $\theta_2$  and  $\theta_4$  have no effect. Also, note that the total effect estimates  $S_{T_i}^*$  are similar for  $\theta_1$  and  $\theta_3$ .

Likewise, for  $n - 1 = 8$  and  $n - 1 = 16$  the estimated indices match the expected behavior from the analysis of local sensitivities and the basis functions of the model.

**Morris indices** All parameters are expected to have equal sensitivities, due to equal weights and same mean value. Parameter sets reducing to the same basis functions ( $n - 1 = 4$ ) are expected to show some second- or higher-order effects. The resulting indices are shown in Table 2. The ordering predicted by the two global methods is consistent.

Next, we expand this aliasing example to  $p = n - 1 = 10$  with results shown in Table 3. Both Sobol indices and Morris elementary effects suggests that all parameters are equally influential except for  $\theta_5$  and  $\theta_{10}$ , which have negligible impact on the model output. This is confirmed by Fig. 4 where the graphs of the corresponding basis functions are plotted for the given resolution, and three higher resolutions. When  $n - 1 = 10$ , all samples of basis functions corresponding to  $\theta_5$  and  $\theta_{10}$  are zero. This result is consistent with the rankings obtained using OSM, also shown in Table 3. As indicated by these tables, there are two parameters  $\theta_5$  and  $\theta_{10}$  that do not contribute to the model output. For these subsets, the FIM is singular, and correlations cannot be calculated by inverting the FIM. Here consideration of which parameters' fixation reduces the FIM condition number the most is required.

For the problem with  $p = 4$  and  $n - 1 = 8$ , due to the resolution the parameter  $\theta_4$  is noninfluential, while  $\theta_1, \theta_2, \theta_3$  are linearly independent, and are therefore all identifiable. Adding to the sum an interaction term  $\theta_1\theta_4 \sin(2\pi t)$ , gives the response function



**Fig. 4** Basis functions for the expanded aliasing Example (16) evaluated at  $t = 0, 1/n, 2/n, \dots, n-1/n, 1$  (plotted with linear interpolation between the points) for  $n - 1 = 10$  and two other measurement resolutions and full resolution time series

$$f(t, \theta) = \theta_1\theta_4 \sin(2\pi t) + \sum_{k=1}^4 \theta_k^2 \sin(2\pi kt). \tag{18}$$

One would expect that  $\theta_4$  now has influence on the model output, but that a large part of the influence is through interaction. Table 4 shows the ordering from OSM using parameter values  $\theta_i = 1$ , as well as Sobol indices and Morris elementary effects. From OSM, we observe that the parameter  $\theta_4$  has an impact (the parameter  $e = 0.18$ ), but it does not contribute any dynamics not already represented by the other parameters, namely  $\theta_1$  that it is linearly dependent on. Again, since there is a linear dependence in the sensitivities, correlations cannot be calculated from FIM. For the global methods, we note that all parameters are found to be influential, confirming that these methods measure sensitivity and not identifiability.

### 3.2.2 Case B: Voigt body model

The previous aliasing example illustrated a case of practical unidentifiability. Unfortunately, for most models identifiability issues are more complex, as illustrated in this example.

Consider the two-element Voigt body model, shown in Fig. 5, which models the viscoelastic strain of the baroreceptor neurons in large arteries [2,8,24]. The deformation of the components can be described by the differential equations

$$\begin{aligned} \frac{d\varepsilon_1}{dt} &= -(\alpha_1 + \alpha_2 + \beta_1)\varepsilon_1 + (\beta_1 - \beta_2)\varepsilon_2 + (\alpha_1 + \alpha_2)\varepsilon_w \\ \frac{d\varepsilon_2}{dt} &= -\alpha_2\varepsilon_1 - \beta_2\varepsilon_2 + \alpha_2\varepsilon_w, \end{aligned} \tag{19}$$

**Table 2** (a) Ranking obtained using the orthogonal sensitivities method for  $n - 1 = \{4, 8, 16\}$

Rank	$n - 1 = 4$				$n - 1 = 8$				$n - 1 = 16$			
	Par	$e$	$d$	$I$	Par	$e$	$d$	$I$	Par	$e$	$d$	$I$
(a) OSM												
1	$\theta_3$	0.71			$\theta_1$	0.55			$\theta_3$	0.38		
2	$\theta_1$	0.71	0.00	0.00	$\theta_3$	0.51	1.00	0.51	$\theta_2$	0.33	1.00	0.33
3	$\theta_2$	0.00	0.00	0.00	$\theta_2$	0.48	1.00	0.48	$\theta_1$	0.33	1.00	0.33
4	$\theta_4$	0.00	0.00	0.00	$\theta_4$	0.00	0.00	0.00	$\theta_4$	0.25	1.00	0.25
Par	$n - 1 = 4$				$n - 1 = 8$				$n - 1 = 16$			
	$S_i^*$	$S_{Ti}^*$	$\mu_i^*$	$\sigma_i^*$	$S_i^*$	$S_{Ti}^*$	$\mu_i^*$	$\sigma_i^*$	$S_i^*$	$S_{Ti}^*$	$\mu_i^*$	$\sigma_i^*$
(b) Global												
$\theta_1$	0.05	0.85	1.10	1.24	0.29	0.32	1.15	1.33	0.20	0.29	1.31	1.56
$\theta_2$	0.00	0.00	0.00	0.00	0.41	0.36	1.12	1.31	0.18	0.25	1.36	1.62
$\theta_3$	0.15	0.90	1.10	1.25	0.39	0.37	1.17	1.34	0.23	0.29	1.33	1.61
$\theta_4$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.16	0.25	1.33	1.61

Columns denote the importance (PCA) score ( $e$ ), orthogonality ( $d$ ), and total identifiability ( $I = ed$ ). (b) Sobol and Morris indices for the parameters in Case A: Aliasing (Sect. 3.2.1) with  $m = 5$  and  $n - 1 = 4, 8, 16$  parameters. Here  $S_i^*$  and  $S_{Ti}^*$  denote the first-order sensitivity indices, while  $\mu_i^*$  and  $\sigma_i^*$  are the Morris indices

**Table 3** Sensitivity and identifiability analysis results for the aliasing example (Sect. 3.2.1) with  $p = n - 1 = 10$

	$\theta_1$	$\theta_2$	$\theta_3$	$\theta_4$	$\theta_5$	$\theta_6$	$\theta_7$	$\theta_8$	$\theta_9$	$\theta_{10}$
(a) Global										
$S_i^*$	0.05	0.05	0.06	0.04	0.00	0.05	0.05	0.06	0.05	0.00
$S_{Ti}^*$	0.19	0.18	0.18	0.19	-0.02	0.17	0.18	0.17	0.20	-0.02
$\mu_i^*$	9.78	8.83	9.25	8.90	0.00	8.82	9.96	8.88	9.22	0.00
$\sigma_i^*$	112.00	105.67	116.31	93.10	0.00	103.31	134.31	102.61	117.50	0.00
Rank	Parameter	$e$	$d$	$I$						
(b) OSM										
1	$\theta_2$		0.36							
2	$\theta_8$		0.36	0.98	0.35					
3	$\theta_4$		0.30	1.00	0.30					
4	$\theta_6$		0.30	0.98	0.30					
5	$\theta_1$		0.25	0.97	0.24					
6	$\theta_9$		0.25	0.98	0.24					
7	$\theta_7$		0.23	0.92	0.21					
8	$\theta_3$		0.23	0.98	0.22					
9	$\theta_{10}$		0.00	0.00	0.00					
10	$\theta_5$		0.00	0.00	0.00					

(a) Sobol indices, computed using Jansen’s implementation, and Morris elementary effects. Here  $S_i^*$  and  $S_{Ti}^*$  denote the first-order sensitivity indices, while  $\mu_i^*$  and  $\sigma_i^*$  are the Morris indices. (b) Ranking obtained using OSM. Here ( $e$ ) denote the importance (PCA) score ( $e$ ), ( $d$ ) Orthogonality, and ( $I = ed$ ) total identifiability

where  $\alpha_i = K_N/B_i$  and  $\beta_i = K_i/B_i$ . Let  $\varepsilon_1$  be the model output, the Heaviside function be the model input  $\varepsilon_w = H$ , and nominal parameter values be  $\alpha_1 = 2, \alpha_2 = 1, \beta_1 = 2, \beta_2 = 0.4$ . For the global sensitivity analysis methods, we sample the parameter values with uniform probability on the intervals of 75–125% of the nominal value, e.g.  $\alpha_1 \in 2 \times (0.75, 1.25)$ . We take the output variable to be  $\int_0^5 \varepsilon_1^2(t)dt$ . The model is solved for  $t \in (0, 5)$ , with the

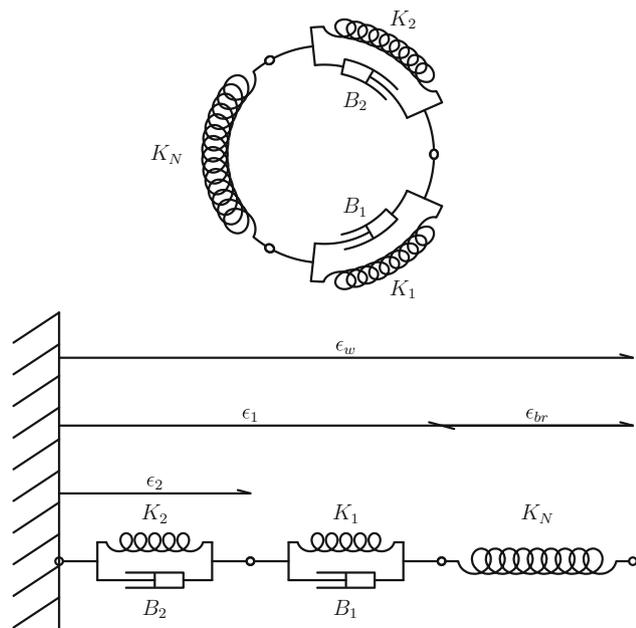
Heaviside function changing at  $t = 0.5$ . Parameters  $\alpha_1, \alpha_2$  determine how the strain is distributed between the two Voigt bodies, whereas  $\beta_1, \beta_2$  determine the relaxation time for each Voigt body.

The local sensitivities are calculated using sensitivity equations and are plotted in Fig. 6, whereas analysis results are presented in Table 5.

**Table 4** Results for linear aliasing with interaction term with  $p = 4, n - 1 = 8$

	$\theta_1$	$\theta_2$	$\theta_3$	$\theta_4$
Global				
$S_i^*$	0.61	0.09	0.09	0.08
$S_{Ti}^*$	0.73	0.13	0.13	0.20
$\mu_i^*$	19.95	11.85	11.06	4.00
$\sigma_i^*$	298.03	87.54	93.07	18.81
Rank	Par	$e$	$d$	$I$
OSM				
1	$\theta_1$	0.53		
2	$\theta_3$	0.22	1.00	0.22
3	$\theta_2$	0.22	1.00	0.22
4	$\theta_4$	0.18	0.00	0.00

Left: global measures. Right: OSM ranking



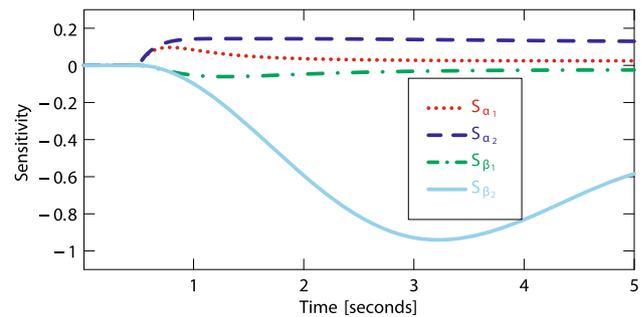
**Fig. 5** The Voigt body model used in the baroreceptor reflex model to describe viscoelastic strain of the baroreceptor neurons

Mahdi et al. [23] found that the model is structurally identifiable, when regarding wall strain  $\epsilon_w$  as the model input and  $\epsilon_1$  as the model output. Our analysis suggests that some parameters may be difficult to determine in practice. OSM yields that  $\alpha_1$  and  $\beta_1$  have low identifiability indices based on low importance indices, while SCM yields large correlations between parameters in the pairs  $(\alpha_1, \alpha_2)$ ,  $(\alpha_1, \beta_2)$  and  $(\alpha_2, \beta_2)$ . The global sensitivity analysis methods yield similar influence for all parameters.

**Table 5** Sensitivity and identifiability analysis results for the Voigt body model

	$\alpha_1$	$\alpha_2$	$\beta_1$	$\beta_2$
(a) Global				
$S_i^*$	0.19	0.50	0.14	0.11
$S_{Ti}^*$	0.15	0.53	0.16	0.17
$\mu$	0.05	0.09	0.05	0.05
$\sigma_i$	0.01	0.01	0.01	0.00
Rank	Par	$e$	$d$	$I$
(b) OSM ranking				
1.00	$\beta_2$	0.97		
2.00	$\alpha_2$	0.18	0.42	0.08
3.00	$\alpha_1$	0.04	0.23	0.01
4.00	$\beta_1$	0.05	0.17	0.01
	$\alpha_1$	$\alpha_2$	$\beta_1$	$\beta_2$
(c) SCM correlations				
$\alpha_1$	–	–0.90	–0.56	–0.91
$\alpha_2$	–0.90	–	0.85	0.98
$\beta_1$	–0.56	0.85	–	0.75
$\beta_2$	–0.91	0.98	0.75	–

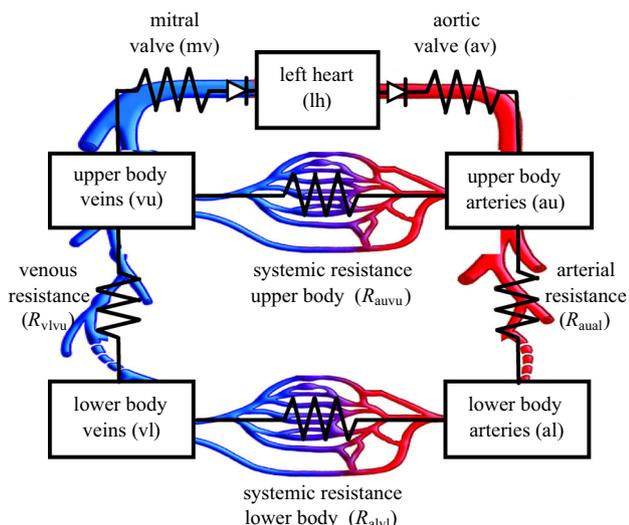
(a) Sobol indices ( $M = 50,000$ ) and Morris elementary effects ( $r = 40$ ) from Global analysis. (b) OSM ranking. (c) Initial correlations from SCM



**Fig. 6** Sensitivities for the Voigt body model (color figure online)

### 3.3 Application to a physiological model

For the final application in mathematical biology, we consider a model of the cardiovascular system that describes the blood pressure in the systemic circulation of the human body. The model is derived from an electrical circuit analogy with a forcing function that models the pumping heart. It is described by a system of linear differential equations with an explicit forcing function. While the model is fairly simple, the relationship between, and identifiability of, the model parameters is not obvious due to the nonlinear dependence of model output on the parameters. We will analyze



**Fig. 7** Overview of the cardiovascular model. All compartments represent a given volume and pressure, compartments are compliant modeled via a capacitor linked to ground, and flow between compartments is resistive. The heart compliance is time-varying representing pumping of the heart. Valves on both sides of the heart are modeled as diodes (color figure online)

the model, build subsets and test the ability of the subsets to fit experimental data.

The model lumps the circulation into 5 compartments; Upper body arteries (au), lower body arteries (al), lower body veins (vl), upper body veins (vu) and the heart (h). For a more thorough description of the model see [46]. Figure 7 shows a schematic of the model, where components are depicted with the symbol of their electrical equivalent.

The model dynamics can be described from equations relating pressure (analogous to voltage)  $p_i$ , flow (analogous to current)  $q_i$  and volume (analogous to charge)  $V_i$ . As for most compartment models, the change in volume  $V$  is given by

$$\frac{dV}{dt} = q_{in} - q_{out},$$

and the flow and pressure related as described by Ohm’s law, i.e.

$$q = \frac{p_{in} - p_{out}}{R},$$

where  $R$  is the resistance, and finally pressure and volume are related as

$$p - p_{ext} = C (V - V_{un}). \tag{20}$$

Here  $C$  is the compartment compliance (analogous to capacitance),  $V_{un}$  the compartment unstressed volume, and  $p_{ext} = 0$  is the tissue pressure.

The left ventricular pressure is predicted similar to (20), except it is formulated using a time-varying elastance function  $E_{lv}(t)$  (inverse of compliance,  $E = 1/C$ ) given by

$$p_{lv} = E_{lv}(t)(V - V_{un}),$$

$$E_{lh}(t) = \begin{cases} \frac{E_M - E_m}{2} \left( 1 - \cos \left( \pi \frac{t}{T_M} \right) \right) + E_m & 0 \leq t \leq T_M \\ \frac{E_M - E_m}{2} \left( 1 + \cos \left( \pi \frac{t - T_M}{T_R} \right) \right) + E_m & T_M \leq t \leq T_M + T_R \\ E_m & T_M + T_R < t. \end{cases}$$

Figure 8 shows experimental blood pressure data and model output from the upper body arteries, computed using nominal parameter values. Since the blood pressure model does not have any description of the fluid dynamics of the blood, it is not able to predict the second peak in each cycle, as it originates from a reflective wave in the arteries. In attempt to encapsulate the modeled dynamics, we have selected two points in each cycle where the arterial blood pressure is not affected by the reflected wave for optimization during systole and diastole. Subsequently, for SA we selected five points centered around each of these, as illustrated in Fig. 8, resulting in the output times  $t = t_i, i \in [1, 5 \times 2 \times N]$  for  $N$  cycles. Inspired by Williams et al. [46], we added total blood volume and cardiac output (CO)—the total blood flow—to the model output, but only evaluate these at systole; that is,  $t = t_{8+10(j-1)}, j \in [1, N]$ ,

$$y^m = [p_m(t_1), \dots, p_m(t_{10N}),$$

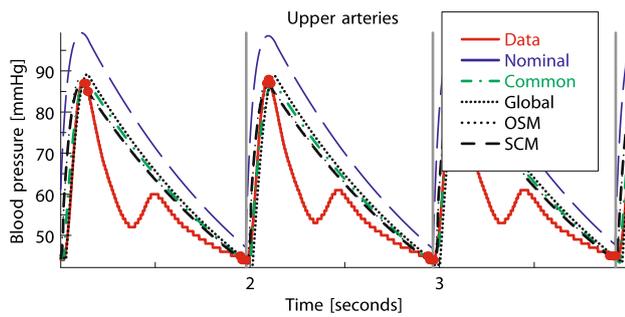
$$CO_m(t_8), \dots, CO_m(t_{8+10(N-1)}),$$

$$V_m(t_8), \dots, V_m(t_{8+10(N-1)})].$$

To balance the importance between matching pressure, blood volume, and cardiac output, blood volume and cardiac output are scaled by a constant factor  $\alpha = 5$ . The resulting residual is given by

$$r = \frac{1}{10N + \alpha 2N} \left[ \frac{y_1^m - y_1^d}{y_1^d}, \dots, \frac{y_{10N}^m - y_{10N}^d}{y_{10N}^d}, \right. \\ \left. \alpha \frac{y_{10N+1}^m - y_{10N+1}^d}{y_{10N+1}^d}, \dots, \alpha \frac{y_{12N}^m - y_{12N}^d}{y_{12N}^d} \right].$$

Initial values for parameter values and state variables (given in Table 8) are calculated from steady-state data, as described by Williams et al. [46]. These values are used to calculate local sensitivities via sensitivity equations. For the global sensitivity analysis methods, to restrict parameters to positive values, the parameters are assumed to be log-normally distributed. Parameter values are sampled using a uniform distribution for the Morris and Sobol methods and



**Fig. 8** Zoom of experimental blood pressure data (red solid), and model output for nominal (blue long dashed) and optimized parameter values (green dash-dotted and black dotted, short dashed and dashed). The red dots show the time points used for SA and optimization and the vertical lines (gray) the start of each contraction cycle of the heart. The optimized curves are colored green for the combined method, and black for the individual SA methods. The structural correlation method (SCM) and orthogonal sensitivities method (OSM) solutions are virtually indistinguishable for large parts of the plot (color figure online)

translated to corresponding percentile on log-normal distribution for the parameters. As the Morris methods sample from the extreme points on  $\ell$ -grid, the sampled points have been scaled to fall within the 1 and 99 percentiles, since the 0 percentile corresponds to 0, and the 100 percentile does not have a well-defined value on the log-normal distribution. The global methods consider the least squares error  $J = r^T r$  for the analysis.

Model data include upper body arterial pressure and heart rate measured over 32 cycles for a subject resting in supine position. Heart rate is used as a model input to set the length of each cardiac cycle, and blood pressure is to be compared to the model output. The fact that data are only available from one compartment means that the model is likely over-parameterized. As common from electrical circuits, it can be observed that the two paths from the arterial to the venous side becomes almost equivalent. For this reason, interactions are expected between corresponding parameters in the upper and lower part of the circulation.

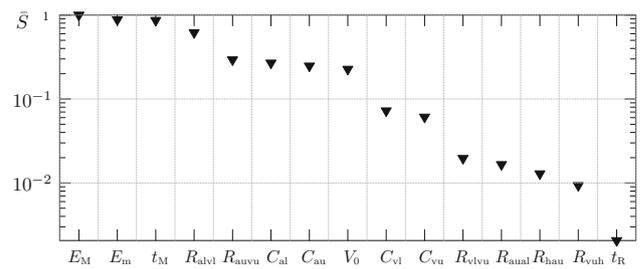
Figure 9 shows the ordering when parameters are ranked by the two-norm of sensitivities. Fixing the three least influential parameters, and then iteratively fixing the least influential parameter for the largest correlation pair until no correlation exceed 0.9 yields the subset

$$\theta_{SCM} = \{E_M, E_m, t_M, R_{alvl}, C_{al}, C_{vl}, R_{vlvu}, R_{aual}\},$$

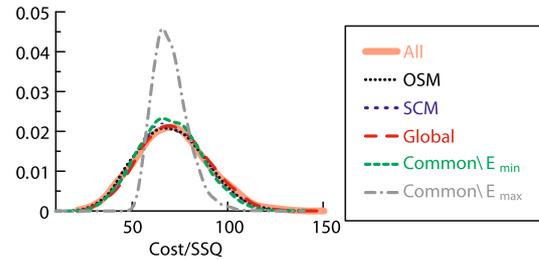
using the SCM method.

Considering parameters with an  $I$ -value lower than  $5 \times 10^{-3}$  as unidentifiable, the OSM method yields the parameters

$$\theta_{OSM} = \{E_M, E_m, t_M, R_{alvl}, C_{al}, C_{vu}, C_{au}, V_0\}$$



**Fig. 9** Logarithm of two-norms of local relative sensitivities, estimated using sensitivity equations and scaled by largest two-norm



**Fig. 10** Least squares error distribution with random sampling for different parameter sets. Each curve represents the distribution of 5000 sampled configurations smoothed using kernel density estimation (KDE) (color figure online)

as shown in Table 6.

Table 6 also shows the sensitivity indices obtained for the Sobol and Morris methods. The two methods agree remarkably well, yielding just two parameters that differ in order. Choosing eight parameters, as with the local methods, the global sensitivity analysis methods yields the subset

$$\theta_{Sobol} = \theta_{Morris} = \{R_{alvl}, R_{aavu}, C_{al}, C_{au}, E_m, E_M, C_{vu}, C_{vl}\}.$$

The three subsets have four parameters in common,

$$\theta_{common} = \{E_M, E_m, R_{alvl}, C_{al}\},$$

which we will consider for optimization.

We used the Levenberg–Marquardt optimization method [21] to optimize the different subsets minimizing the least squares error, fixing all parameters not included in the subsets at their nominal values. Table 7 summarizes the optimization results: listing the number of estimated parameters, the least squares error, and the number of steps required by the optimization routine. Figure 8 shows the upper arterial pressure, simulated using the optimized parameter values for each parameter set, and Table 8 shows the estimated parameter values. The optimization was run over the same time points used in the SA, and the plot shows a zoom over one cardiac cycle. From Fig. 8 it is clear that more than one parameter set can be used to obtain a reasonable fit.

**Table 6** (a) Results and ordering due to the OSM method for the blood pressure circulation model

(i) OSM				(ii) Sobol			(iii) Morris		
Par	<i>e</i>	<i>d</i>	<i>I</i>	Par	$\mu$	$\sigma^2$	Par	$\mu$	$\sigma^2$
$E_M$	0.49			$R_{alvl}$	0.64	1.21	$R_{alvl}$	0.47	0.48
$E_m$	0.42	0.99	0.42	$R_{auvu}$	0.16	0.31	$R_{auvu}$	0.23	0.25
$t_M$	0.46	0.63	0.28	$C_{al}$	0.12	0.18	$C_{al}$	0.180.19	
$R_{alvl}$	0.29	0.47	0.14	$C_{au}$	0.066	0.14	$C_{au}$	0.15	0.16
$C_{al}$	0.10	0.20	0.021	$E_m$	0.044	0.086	$E_m$	0.12	0.13
$C_{vu}$	0.017	0.72	0.013	$E_M$	0.003	0.015	$E_M$	0.058	0.057
$C_{au}$	0.089	0.050	0.0044	$C_{vu}$	0.002	0.0012	$C_{vu}$	0.015	0.019
$V_0$	0.12	0.014	0.0016	$C_{vl}$	0.00028	0.00087	$C_{vl}$	0.012	0.015
$R_{av}$	0.0067	0.11	0.00076	$t_M$	0.00097	0.00079	$t_M$	0.010	0.011
$R_{vlvu}$	0.0094	0.042	0.00039	$R_{vlvu}$	0.00002	0.00020	$V_0$	0.0054	0.00581
$t_R$	0.00079	0.14	0.00011	$V_0$	0.00052	0.00015	$R_{vlvu}$	0.0046	0.0062
$R_{aual}$	0.0056	0.015	0.000086	$R_{aual}$	0.00008	0.00003	$R_{aual}$	0.0020	0.0025
$R_{auvu}$	0.14	0.00019	0.000027	$R_{mv}$	0.00000	0.00001	$R_{mv}$	0.0011	0.0013
$C_{vl}$	0.0077	0.00063	0.000005	$t_R$	0.00000	0.00000	$t_R$	0.00057	0.00061
$R_{mv}$	0.0045	0.0011	0.000005	$R_{av}$	0.00000	0.00000	$R_{av}$	0.00016	0.00032

The parameters below  $V_0$  are considered unidentifiable due to having  $I < 10^{-4}$ . (b, c) Ordering by the Sobol total effects (5000 model evaluations) and Morris elementary effects (480 evaluations) for the blood pressure circulation model

**Table 7** Optimization results for the different subsets

Subset	# parameters	LSE	Steps
Unoptimized*	15	67.8	
Common	4	1.17	14
SCM	8	0.950	11
OSM	8	0.901	16
Global	8	1.14	27

\*Unoptimized parameters are fixed at their nominal values

\*\*LSE least squares error

As expected, the estimated parameter values vary reflecting the bias associated with fixing certain parameters at specific values.

To determine the variability of the model output, we sampled 5000 points from each subset parameter space and calculated the least squares error at each point. Figure 10 shows the least squares error distribution for each subset using Gaussian kernel density estimation [37], as well as for the full subset with all parameters included. It is evident that the OSM, SCM and global subsets align well with the nominal distribution. As the common subset aligned well, but slightly worse than the others, we investigated the distribution when reducing the subset even more. As the figure shows, the distribution is fairly similar when using the common subset and removing  $E_m$ , however when removing  $E_M$  the distribution deviates.

**Table 8** Nominal and optimized parameter values for the cardiovascular model

Parameter	Nominal	Optimized			
		Common	SCM	OSM	Global
$C_{au}$	0.951	–	–	0.784	0.290
$C_{al}$	1.04	0.015	0.849	0.492	0.043
$C_{vl}$	42.2	–	29.7	–	0.746
$C_{vu}$	32.8	–	–	28.5	89.1
$R_{aual}$	0.009	–	1.4e-13	–	–
$R_{auvu}$	1.66	–	–	–	16.5
$R_{alvl}$	0.799	3.78	1.05	1.99	4.10
$R_{vlvu}$	0.003	–	1.3e-9	–	–
$R_{vuh}$	0.001	–	–	–	–
$R_{hau}$	0.001	–	–	–	–
$V_0$	10.0	–	–	0.014	–
$E_m$	0.023	0.034	0.023	0.027	0.045
$E_M$	2.31	1.16	1.13	1.01	1.17
$t_M$	0.200	–	0.176	0.165	–
$t_R$	0.150	–	–	–	–

Nominal values are calculated as described in the study by Williams et al. [46]

### 4 Discussion

In this investigation, we presented methods and algorithms for quantifying model sensitivity to parameters and consistently identifiable parameter subsets. We used partial derivatives of the model output with respect to param-

ters [26,28,32,34] to locally estimate the parameter influence, and Sobol Indices [38,41] and Morris Measures [27,38] for global sensitivity analysis.

Whereas the latter allowed us to determine which parameters were noninfluential, they do not specify if any parameters are functionally related or the model over-parameterized. To remedy this challenge we considered two identifiability analysis methods: The structural correlation method (SCM) [28] and the orthogonal sensitivities method (OSM) [22]. The SCM constructs subsets by initially fixing noninfluential parameters, and then subsequently fixing parameters one by one until no linear pairwise correlations are left [28].

The OSM ranks the parameters by first selecting the most influential parameters based on a principal component analysis (PCA) of the local sensitivities, subsequently adding one parameter at the time to the ranking. The parameter added is the most identifiable found from analysis of its influence and the linear independence of its sensitivities to the space spanned by the sensitivities of the already ranked parameters [22]. Other methods aiming at using linear independence in sensitivities include the *orthogonal method* by Yao et al. [26,48], SVD-QR [12,29].

For the examples in Sects. 3.1 and 3.2, the global sensitivity analysis methods agreed, they both correctly specified parameters designated as noninfluential. While both OSM and SCM were able to determine structural and practical unidentifiability in these examples, the FIM is not invertible. If the FIM is not invertible, it is necessary to first remove parameters for which local sensitivities are linearly dependent on others as described in the SCM algorithm. Moreover, in Sect. 3.1 we showed how non-dimensionalization may be used to reformulate the model and reduce the parameter space for structural unidentifiability. This method is not specific to the example presented here but can often help reduce complexity of any given model. The downside is that they may be more difficult to interpret directly as they need to be re-dimensionalized to have physical units.

A more complex and practically applied example is shown in Sect. 3.2.2—a Voigt body is used to describe the viscoelastic properties of arterial wall. For this example, the SA methods showed that all parameters were influential, agreeing with the previous study by Mahdi, Meshkat and Sullivant [23], which showed that they are structural identifiable, yet both the OSM and the SCM suggested that given available input and output data, the model parameters are practically unidentifiable.

Finally, the analysis of the cardiovascular model in Sect. 3.3 showed that global sensitivity to parameters may be different from the local sensitivity, which is expected for a nonlinear model [32,35]. Since OSM and SCM relies heavily on the two-norm of the local sensitivities, it is likely that subsets built using local sensitivities may be different for other model configurations. This may warrant an additional anal-

ysis after optimization to ensure that a better subset does not exist, and underlines the importance of good initial parameter configuration for optimization. Moreover, we noticed that the local subsets predicted by SCM and OSM appeared to be more efficient than the global set, with SCM reaching an optimum in fewer optimization steps, while OSM reaches a slightly better fit. In this regard, it should be noted that Levenberg–Marquardt, the optimization routine used, is a gradient based method [21], and as such relies on linear independence of the sensitivities, and thereby favors methods that seek to optimize linear independence of the linearized system. The parameter sets with eight parameters did slightly better than the set of the four common parameters. In general none of the subsets generated by the individual methods did much better than the smaller common set. One may use model comparison methods such as the Akaike information criterion [1] to quantify the (dis)advantage of including additional parameters for a slightly better fit. In terms of maintained variability of model output, the suggested methods for subset selection all produced similar results. But we also saw that the parameter set could be reduced even more without losing the original variability. Moreover, as observed in Table 8 the estimated parameters vary dependent on the chosen subset, this is expected as here we fix noninfluential parameters biasing the prediction. Therefore, care must be taken if influential parameters are fixed and results should be interpreted relative to the physiological problem in question.

These tasks are by no means special to physiological modeling, but may be seen in many other areas of mathematical modeling. Other obvious applications involves Chemical reaction models, or any other area where ODE or PDE models are constructed to match experimental data.

It is important to be aware that this is not the only nor the typical application of SA, and that other aspects of the analysis might have greater importance in other applications. A good general introduction to many aspects of SA is given by Saltelli et al. [34]. We have used SA as a tool to reduce the number of parameters in what Saltelli et al. [35] calls *factors' fixing in hyper-parameterized* models. Another use is *Factors' prioritization* where SA is applied to pinpoint which parameter to measure to reduce the model output uncertainty the most [35], and is perhaps the most frequent application of SA.

Depending on the problem under consideration, different strategies for subset selection may be viable. Based on the optimization of the cardiovascular model considered here, the local methods appear to superior in terms of computational effort and model fit. This is, however, based on having a reasonable initial parameter configuration and using a gradient based optimization routine. For models where a good initial configuration is not available, or where the exploration of potential behavior is prioritized, the global methods might be a better choice. If computational efficiency is essential,

Morris elementary effects yields similar results to Sobol Indices at a fraction of the computational cost. One may also consider methods such as FAST [36] for more efficient calculation of Sobol Indices, or moment independent SA methods [7,44] that consider the model output probability density function rather than the variance. Iooss and Lemaître [18] provide a nice overview of global SA methods. In addition Markov Chain Monte Carlo (MCMC) methods such as DRAM [15] or DREAM [43] could be considered for estimating parameter uncertainties and correlations. However, these methods are typically applied after identifiable parameters have been isolated, and as such are beyond the scope of this study.

In addition to considering which method(s) to apply, it is important to remember that fixing parameters and reducing parameter subsets may introduce a bias in the model in question. In practice this implies that a parameter value estimated using one subset, cannot be compared (or expected to be similar) to values obtained using other subsets, as the parameter has different possibilities to influence the model output through interactions with other parameters.

**Acknowledgements** Olsen and Olufsen was supported in part by the National Science Foundation (Grant NSF-DMS 1022688 and NSF-DMS 1557761) as well as the Virtual Physiological Rat Project funded by the National Institute of General Medical Sciences (NIH-NIGMS P50 GM094503).

## References

1. Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Autom Control* AC-19(6):716–723
2. Alfrey K (1997) Model of the aortic baroreceptor in rat. PhD thesis, MS thesis. Rice University, Houston
3. Andersson J, Åkesson J, Diehl M (2012) Casadi: a symbolic package for automatic differentiation and optimal control. In: Forth S, Hovland P, Phipps E, Utke J, Walther A (eds) Recent advances in algorithmic differentiation. Springer, Berlin, pp 297–307
4. Ashyraliyev M, Fomekong-Nanfack Y, Kaandorp JA, Blom JG (2009) Systems biology: parameter estimation for biochemical models. *Febs J* 276(4):886–902
5. Ashyraliyev M, Jaeger J, Blom JG (2008) Parameter estimation and determinability analysis applied to drosophila gap gene circuits. *BMC Syst Biol* 2(1):83 (19 pages)
6. Bellman R, Åström KJ (1970) On structural identifiability. *Math Biosci* 7(3):329–339
7. Borgonovo E, Castaings W, Tarantola S (2012) Model emulation and moment-independent sensitivity analysis: an application to environmental modelling. *Environ Model Softw* 34:105–115. <https://doi.org/10.1016/j.envsoft.2011.06.006>. <http://www.sciencedirect.com/science/article/pii/S1364815211001617>
8. Bugenhagen S, Cowley A, Beard D (2010) Identifying physiological origins of baroreflex dysfunction in salt-sensitive hypertension in the Dahl SS rat. *Physiol Genom* 42:23–41
9. Daun S, Rubin J, Vodovotz Y, Roy A, Parker R, Clermont G (2008) An ensemble of models of the acute inflammatory response to bacterial lipopolysaccharide in rats: results from parameter space reduction. *J Theor Biol* 253(4):843–853
10. Devroye L (1986) Sample-based non-uniform random variate generation. In: Proceedings of the 18th conference on winter simulation. ACM, pp 260–265
11. Eisenberg M, Harsh J (2017) A confidence building exercise in data and identifiability: modeling cancer chemotherapy as a case study. *J Theor Biol* 431:63–78
12. Ellwein L, Pope S, Xie A, Batzel J, Kelley C, Olufsen M (2013) Patient-specific modeling of cardiovascular and respiratory dynamics during hypercapnia. *Math Biosci* 241(1):56–74
13. Fournier DA, Skaug HJ, Ancheta J, Ianelli J, Magnusson A, Maunder MN, Nielsen A, Sibert J (2012) Ad model builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optim Methods Softw* 27(2):233–249. <https://doi.org/10.1080/10556788.2011.597854>
14. Griewank A (1989) On automatic differentiation. *Math Program Recent Dev Appl* 6:83–107
15. Haario H, Laine M, Mira A, Saksman E (2006) Dram: efficient adaptive MCMC. *Stat Comput* 16:339–354
16. Holmberg A (1982) On the practical identifiability of microbial growth models incorporating Michaelis–Menten type nonlinearities. *Math Biosci* 62(1):23–43
17. Houska B, Ferreau HJ, Diehl M (2011) Acado toolkit: an open-source framework for automatic control and dynamic optimization. *Opt Control Appl Methods* 32(3):298–312. <https://doi.org/10.1002/oca.939>
18. Iooss B, Lemaître P (2015) A review on global sensitivity analysis methods. In: Uncertainty management in simulation-optimization of complex systems. Springer, pp 101–122
19. Jacques JA, Greif P (1985) Numerical parameter identifiability and estimability: integrating identifiability, estimability, and optimal sampling design. *Math Biosci* 77(1):201–227
20. Jansen M (1999) Analysis of variance designs for model output. *Comput Phys Commun* 117(1):35–43
21. Kelley C (1999) Iterative methods for optimization, vol 18. SIAM, Philadelphia
22. Li R, Henson M, Kurtz M (2004) Selection of model parameters for off-line parameter estimation. *IEEE Trans Control Syst Technol* 12(3):12
23. Mahdi A, Meshkat N, Sullivan S (2014) Structural identifiability of viscoelastic mechanical systems. *PLoS ONE* 9(2):e86411
24. Mahdi A, Sturdy J, Ottesen J, Olufsen M (2013) Modeling the afferent dynamics of the baroreflex control system. *PLoS Comput Biol* 9(12):e1003384
25. Martins J, Kroo I, Alonso J (2000) An automated method for sensitivity analysis using complex variables. In: Proceeding of the 38th aerospace sciences meeting, p 0689
26. Miao H, Xia X, Perelson AS, Wu H (2011) On identifiability of nonlinear ode models and applications in viral dynamics. *SIAM Rev* 53:3–39
27. Morris M (1991) Factorial sampling plans for preliminary computational experiments. *Technometrics* 33(2):161–174
28. Olufsen MS, Ottesen JT (2013) A practical approach to parameter estimation applied to model predicting heart rate regulation. *J Math Biol* 67(1):39–68
29. Pope S, Ellwein L, Zapata C, Novak V, Kelley C, Olufsen M (2009) Estimation and identification of parameters in a lumped cerebrovascular model. *Math Biosci Eng* 6(1):93–115
30. Rall LB (1981) Automatic differentiation: techniques and applications. Lecture notes in computer science, vol 120. Springer, New York
31. Raue A, Kreutz C, Maiwald T, Bachmann J, Schilling M, Klingmüller U, Timmer J (2009) Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics* 25:1923–1929
32. Rodríguez-Fernández M, Banga JR, Doyle FJ (2012) Novel global sensitivity analysis methodology accounting for the crucial role of

- the distribution of input parameters: application to systems biology models. *Int J Robust Nonlinear Control* 22(10):1082–1102
33. Saltelli A, Annoni P, Azzini I, Campolongo F, Ratto M, Tarantola S (2010) Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Comput Phys Commun* 181(2):259–270
  34. Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M, Tarantola S (2008) *Global sensitivity analysis: the primer*. Wiley, Chichester
  35. Saltelli A, Ratto M, Tarantola S, Campolongo F, of Ispra European Commission JRC (2006) Sensitivity analysis practices: strategies for model-based inference. *Reliab Eng Syst Saf* 91(10):1109–1125
  36. Saltelli A, Tarantola S, Chan KPS (1999) A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics* 41(1):39–56. <https://doi.org/10.1080/00401706.1999.10485594>
  37. Silverman BW (2018) *Density estimation for statistics and data analysis*. Routledge, London
  38. Smith RC (2014) *Uncertainty quantification: theory, implementation, and applications*. SIAM, Philadelphia
  39. Sobol I (1967) On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Comput Math Math Phys* 7:784–802
  40. Sobol I (1993) Sensitivity estimates for nonlinear mathematical models. *Math Model Comput Exp* 1(4):407–414
  41. Sobol I (2001) Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Math Comput Simul* 55(1–3):271–280
  42. Transtrum M, Machta B, Sethna JP (2011) Geometry of nonlinear least squares with applications to sloppy models and optimization. *Phys Rev E* 83(3):36701
  43. Vrugt J, Braak CT, Diks C, Robinson B, Hyman J, Higdon D (2009) Accelerating markov chain monte carlo simulation by differential evolution with self-adaptive randomized subspace sampling. *Int J Nonlinear Sci Numer Simul* 10(3):273–290
  44. Wei P, Lu Z, Yuan X (2013) Monte carlo simulation for moment-independent sensitivity analysis. *Reliab Eng Syst Saf* 110:60–67. <https://doi.org/10.1016/j.res.2012.09.005>. <http://www.sciencedirect.com/science/article/pii/S0951832012001858>
  45. Wentworth MT, Smith RC, Banks HT (2015) Parameter selection and verification techniques based on global sensitivity analysis illustrated for an HIV model. *SIAM/ASA J Uncertain Quantif* 4(1):266–297
  46. Williams ND, Wind-Willassen O, Wright AA, Program R, Mehlsen J, Ottesen JT, Olufsen MS (2014) Patient-specific modelling of head-up tilt. *Math Med Biol* 31(4):365–392
  47. Wolfram: Gauss–Newton methods (2013). <http://reference.wolfram.com/mathematica/tutorial/UnconstrainedOptimizationGaussNewtonMethods.html>
  48. Yao K, Shaw B, Kou B, McAuley K, Bacon D (2003) Modeling ethylene/butene copolymerization with multi-site catalysts: parameter estimability and experimental design. *Polym React Eng* 11(3):563–588